

1- Probabilidad

1.1 - Espacios muestrales y eventos.

La Teoría de Probabilidades estudia los llamados *experimentos aleatorios*.

Ejemplos clásicos de experimentos aleatorios son los juegos de azar:

- a) tirar un dado y observar el número en la cara de arriba.
- b) tirar una moneda
- c) lanzar una moneda cuatro veces y contar el número total de caras obtenidas.
- d) lanzar una moneda cuatro veces y observar la sucesión de caras y cecas obtenidas.

Simbolizamos con ε a un experimento aleatorio.

Un experimento aleatorio tiene las siguientes características:

1-Se lo puede repetir bajo las mismas condiciones tantas veces como se desee.

2- No se puede predecir con exactitud el resultado de dicho experimento, pero se puede decir cuáles son los posibles resultados del mismo.

3- A medida que el experimento se repite, los resultados individuales parecen ocurrir en forma caprichosa. Pero si el experimento se repite un gran número de veces, y registramos la *proporción* de veces que ocurre un determinado resultado, veremos que esa proporción tiende a estabilizarse en un valor determinado a medida que aumenta el número de veces que se repite el experimento.

Por ejemplo, consideremos el experimento de lanzar un dado y observar el número de la cara superior. Supongamos que tiramos el dado N veces, y sea n el número de veces que sale el número 5 en los N tiros del dado. Entonces $\frac{n}{N}$ es la proporción de veces que sale el número 5 en los N tiros. Si el dado

es normal a medida que N aumenta, $\frac{n}{N}$ tiende a estabilizarse en un número que es $1/6$.

Observación: en los experimentos no aleatorios o *deterministas* se puede predecir con exactitud el resultado del experimento, es decir, las condiciones en las que se verifica un experimento determinan el resultado del mismo. Por ejemplo, si colocamos una batería en un circuito simple, el modelo matemático que posiblemente describiría el flujo observable de corriente sería $I = E/R$, que es la ley de Ohm. El modelo predice el valor de I al dar E y R . O sea, si se repite el experimento anterior cierto número de veces, empleando cada vez el mismo circuito, es decir manteniendo fijas E y R , esperaríamos observar el mismo valor de I .

A veces sucede que un experimento no es aleatorio estrictamente, pero resulta mucho más sencillo estudiarlo como si fuera aleatorio. Por ejemplo, si tiramos una moneda y observamos qué lado queda hacia arriba, el resultado sería predecible conociendo en forma precisa las velocidades iniciales de traslación y rotación, y las elasticidades de los materiales del piso y de la moneda. Pero la precisión con la que se necesitan conocer estos datos es casi imposible de obtener en la realidad, por lo que es más conveniente tratar al experimento como aleatorio.

El conjunto de todos los resultados posibles de un experimento aleatorio es el *espacio muestral*. Al espacio muestral lo anotamos con la letra S .

Por ejemplo,

- a) Si ε : tirar un dado y observar el número en la cara de arriba, entonces podemos tomar como espacio muestral a $S = \{1,2,3,4,5,6\}$

- b) Si ε : tirar una moneda, entonces $S = \{c, s\}$
- c) Si ε : lanzar una moneda tres veces y contar el número total de caras obtenidas entonces podemos considerar $S = \{0, 1, 2, 3\}$
- d) Si ε : lanzar una moneda tres veces y observar la sucesión de caras y cecas obtenidas, entonces $S = \{(c, c, c); (c, c, s); (c, s, c); (s, c, c); (c, s, s); (s, s, c); (s, c, s); (s, s, s)\}$
- e) Si ε : tirar un dado las veces necesarias hasta que sale un 6 por primera vez, y contar el número de tiros realizados, entonces $S = \{1, 2, 3, 4, \dots\} = N$, donde N es el conjunto de los números naturales.
- f) Si ε : medir el tiempo de vida de una lamparita eléctrica, entonces $S = \{t \in R, t \geq 0\}$ donde R es el conjunto de los números reales.

Observaciones:

- 1- La elección de S no es única, depende de lo que se quiera observar del experimento aleatorio.
- 2- El espacio muestral puede ser un conjunto finito, o infinito. A su vez si es infinito puede ser infinito numerable o no numerable. En e) el conjunto S es infinito numerable, en f) el conjunto S es infinito no numerable.

Se llama **evento** o **suceso** a todo subconjunto del espacio muestral. Por ejemplo,

- a) En el experimento dado en el ejemplo a), un evento de S sería $A = \{2, 4, 6\}$ pues $A \subset S$. Podemos expresar al evento A con palabras de la siguiente manera A : “sale un número par”
También $B = \{1, 2, 3\}$ es un evento al que podemos expresar verbalmente como B : “sale un número menor o igual que 3”
- b) En el experimento dado en el ejemplo d), un evento de S sería $C = \{(c, c, c); (c, c, s); (c, s, c); (s, c, c)\}$, el que en palabras se puede expresar como C : “salen por lo menos dos caras”
- c) En el experimento dado en el ejemplo f), un evento de S sería D : “la lamparita dura más de 1200 horas”, en notación de conjuntos $D = \{t \in R; t > 1200\}$

Observaciones:

- 1- En el ejemplo a) anterior, si al tirar el dado sale el número 2, entonces podemos decir que **A ocurrió** pues $2 \in A$. Pero también **B ocurrió** pues $2 \in B$. En cambio si al tirar el dado sale el número 4, entonces el evento A ocurrió pero **B no ocurrió**, pues $4 \notin B$.
- 2- El conjunto \emptyset es un evento (pues el conjunto \emptyset está incluido en todo conjunto, en particular $\emptyset \subset S$). Es el evento que **nunca ocurre**.
El espacio muestral S es un evento (pues todo conjunto está incluido en sí mismo), y S **siempre ocurre**.

Las operaciones habituales entre conjuntos se pueden aplicar a los eventos, dando como resultado nuevos eventos. Específicamente

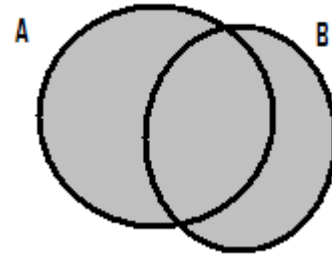
- 1- Si A y B son eventos, entonces $A \cup B$ es otro evento. $A \cup B$ ocurre si y solo si ocurre A o si ocurre B
- 2- Si A y B son eventos, entonces $A \cap B$ es otro evento. $A \cap B$ ocurre si y solo si ocurre A y ocurre B
- 3- Si A es un evento A^c es un evento. A^c ocurre si y solo si no ocurre A

Nota: recordar que las operaciones de unión, intersección, diferencia y complemento se definen de la siguiente manera:

1- $A \cup B$ es el conjunto formado por los elementos que están en A o en B (donde el o está en sentido inclusivo), en notación de conjuntos

$$A \cup B = \{x; x \in A \vee x \in B\}$$

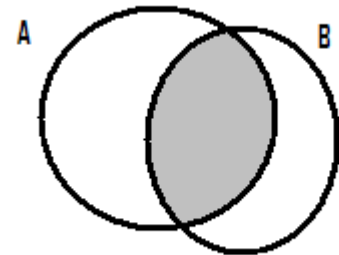
En la figura la zona en gris simboliza $A \cup B$



2- $A \cap B$ es el conjunto formado por los elementos que están en A y en B , en notación de conjuntos

$$A \cap B = \{x; x \in A \wedge x \in B\}$$

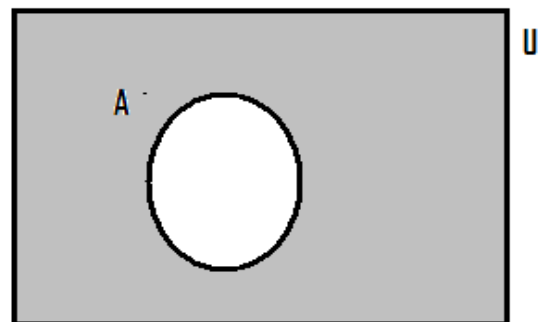
En la figura la zona en gris simboliza $A \cap B$



3- A^c es el conjunto formado por los elementos que están en el conjunto universal U y no están en A , en notación de conjuntos

$$A^c = \{x; x \in U \wedge x \notin A\}$$

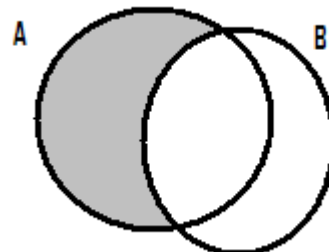
La zona en gris simboliza A^c



4- $A - B$ es el conjunto formado por los elementos que están en A y *no están* en B , en notación de conjuntos $A - B = \{x; x \in A \wedge x \notin B\}$,

La zona en gris simboliza $A - B$

Notar que $A - B = A \cap B^c$



Es útil recordar las siguientes propiedades sobre el álgebra de conjuntos:

1- Leyes de idempotencia

$$\text{a) } A \cup A = A \qquad \text{b) } A \cap A = A$$

2- Leyes asociativas

$$\text{a) } A \cup (B \cap C) = (A \cup B) \cap C \qquad \text{b) } A \cap (B \cup C) = (A \cap B) \cup C$$

3- Leyes conmutativas

$$\text{a) } A \cup B = B \cup A \qquad \text{b) } A \cap B = B \cap A$$

4- Leyes distributivas

$$\text{a) } A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \qquad \text{b) } A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

5- Leyes de identidad

$$\begin{aligned} \text{a) } A \cup \emptyset &= A & \text{b) } A \cap \emptyset &= \emptyset \\ \text{c) } A \cup U &= U & \text{d) } A \cap U &= A \end{aligned}$$

6- Leyes de complemento

$$\begin{aligned} \text{a) } A \cup A^c &= U & \text{b) } A \cap A^c &= \emptyset \\ \text{c) } (A^c)^c &= A & \text{d) } U^c &= \emptyset, \quad \emptyset^c = U \end{aligned}$$

7- Leyes de De Morgan

$$\text{a) } (A \cup B)^c = A^c \cap B^c \qquad \text{b) } (A \cap B)^c = A^c \cup B^c$$

La relación de inclusión entre un conjunto y otro y las operaciones anteriores con conjuntos lleva al siguiente

Teorema:

- a) $A \subset B$ es equivalente a $A \cap B = A$
- b) $A \subset B$ es equivalente a $A \cup B = B$
- c) $A \subset B$ es equivalente a $B^c \subset A^c$
- d) $A \subset B$ es equivalente a $A \cap B^c = \emptyset$,
- e) $A \subset B$ es equivalente a $B \cup A^c = U$

Sean A y B dos conjuntos. El **conjunto producto** de A y B , expresado $A \times B$, está formado por todos los pares ordenados (a, b) donde $a \in A$ y $b \in B$, es decir

$$A \times B = \{(a, b); a \in A \wedge b \in B\}$$

Se anota al producto $A \times A = A^2$

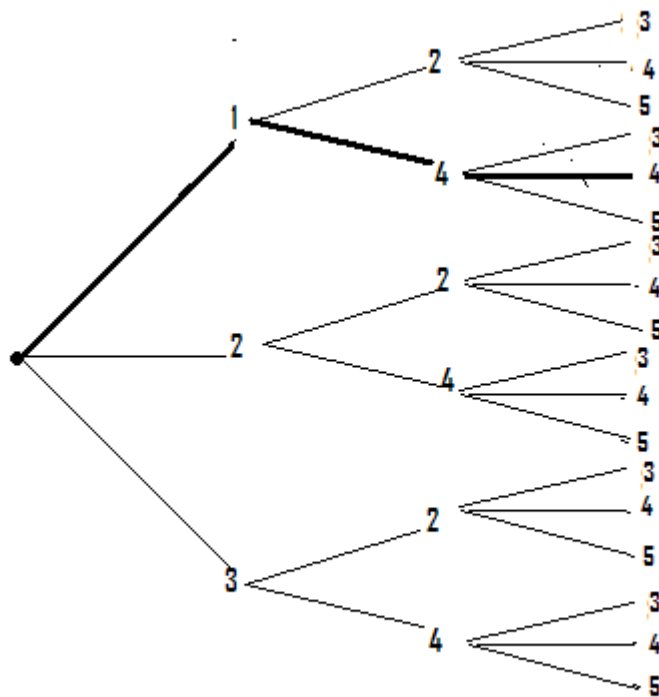
Por ejemplo, si $A = \{1, 2, 3\}$ y $B = \{7, 8\}$, entonces $A \times B = \{(1, 7); (1, 8); (2, 7); (2, 8); (3, 7); (3, 8)\}$

El concepto de conjunto producto se extiende a un número finito de conjuntos en forma natural. El conjunto producto de los conjuntos A_1, A_2, \dots, A_n se anota $A_1 \times A_2 \times \dots \times A_n$ y es igual a

$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n), a_i \in A_i, i = 1, 2, \dots, n\}$, es decir es el conjunto de todas las n -uplas (a_1, a_2, \dots, a_n) donde $a_i \in A_i$ para cada i .

Por ejemplo, si $A = \{1,2,3\}$, $B = \{2,4\}$, $C = \{3,4,5\}$, entonces un método conveniente para hallar el producto $A \times B \times C$ es por medio del denominado “diagrama de árbol” que se muestra a continuación, el cual se construye de izquierda a derecha.

$A \times B \times C$ consta de todas las ternas (o 3-uplas) formadas al seguir cada “camino” del árbol



Por ejemplo el camino con trazo más grueso indicaría la terna $(1, 4, 4)$

Las operaciones de unión e intersección también se pueden generalizar a más de dos conjuntos

Si A_1, A_2, \dots, A_n son n conjuntos, entonces

a) la unión de todos ellos se anota $A_1 \cup A_2 \cup \dots \cup A_n$ o también $\bigcup_{i=1}^n A_i$ y es igual a $A_1 \cup A_2 \cup \dots \cup A_n = \{x; \text{ existe } i \text{ tal que } x \in A_i\}$, es decir un elemento x pertenece a $A_1 \cup A_2 \cup \dots \cup A_n$ si x pertenece **a alguno de los conjuntos** A_i .

De forma análoga se define la unión de una secuencia infinita de conjuntos A_1, A_2, \dots , y se la anota con el símbolo $\bigcup_{i=1}^{\infty} A_i$

b) la intersección de todos ellos se anota $A_1 \cap A_2 \cap \dots \cap A_n$ o también $\bigcap_{i=1}^n A_i$ y es igual a $\bigcap_{i=1}^n A_i = \{x; x \in A_i \text{ para todo } i\}$, es decir un elemento x pertenece a $\bigcap_{i=1}^n A_i$ si x pertenece **a todos los conjuntos** A_i

De forma análoga se define la intersección de una secuencia infinita de conjuntos A_1, A_2, \dots , y se la anota con el símbolo $\bigcap_{i=1}^{\infty} A_i$

Si A y B son dos eventos tales que $A \cap B = \emptyset$, se dice que son disjuntos o **mutuamente excluyentes**.

Es decir si A y B son mutuamente excluyentes no pueden ocurrir a la vez. Por ejemplo, si se tira un dado y se observa el número de la cara superior, los eventos $A = \{1,3,5\}$ y $B = \{2,4,6\}$ son mutuamente excluyentes. Al tirar un dado sale un número par o un número impar, no pueden darse ambas cosas a la vez.

Los eventos con un solo elemento son eventos *elementales o simples*. Por ejemplo, volviendo al experimento ε : tirar un dado y observar el número en la cara de arriba, y $S = \{1,2,3,4,5,6\}$, entonces los conjuntos unitarios $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$ son eventos simples. Notar que dos eventos simples cualesquiera son mutuamente excluyentes.

Dado un evento A asociado a un experimento aleatorio ε . Supongamos que se repite n veces el experimento ε , y anotamos n_A al número de veces que ocurre A en la n repeticiones de ε . Se define

la **frecuencia relativa de A** , y se simboliza f_A , al cociente $\frac{n_A}{n}$. Es decir que f_A es la **proporción de veces que ocurre A en las n repeticiones de ε** .

La frecuencia relativa f_A tiene las siguientes propiedades:

- 1- $0 \leq f_A \leq 1$
- 2- $f_A = 1$ si y solo si A ocurre **cada vez** en las n repeticiones
- 3- $f_A = 0$ si y solo si A no ocurre **nunca** en las n repeticiones
- 4- Si A y B son dos eventos mutuamente excluyentes entonces $f_{A \cup B} = f_A + f_B$

Dado un evento A , se quiere asignar al mismo un número que indique qué tan **probable** es que A ocurra. A ese número lo definiríamos como la **probabilidad de A** . En ese sentido parecería que la frecuencia relativa f_A sería una buena elección. Pero nos encontramos con el siguiente problema, ¿cuántas veces deberíamos repetir el experimento aleatorio para definir f_A , es decir qué valor de n tomaríamos?

Por ejemplo en el experimento de tirar un dado consideremos el evento A : “sale el número 4”, si lo lanzamos 100 veces podríamos encontrar que $n_A = 14$, y si lo lanzamos nuevamente 100 veces podría ocurrir que n_A sea diferente del anterior por ejemplo podría A ocurrir 20 veces. Entonces, tendríamos dos valores diferentes para f_A , 0.14 y 0.2

Se dijo antes que en un experimento aleatorio a medida que n aumenta la frecuencia relativa de A tiende a estabilizarse en un número, pero no podemos en la práctica repetir el experimento infinitas veces.

Se quiere asignar a cada evento A un número que no dependa de la experimentación. Por este motivo procedemos como sigue:

Definición axiomática de probabilidad.

Sea ε un experimento aleatorio y S un espacio muestral asociado con ε . Con cada evento A asociamos un número real llamado probabilidad de A , que anotamos $P(A)$, el cual satisface las siguientes propiedades básicas o *axiomas*

La elección de estas propiedades está motivada por las características correspondientes de la frecuencia relativa.

$$1- 0 \leq P(A) \leq 1$$

$$2- P(S) = 1$$

$$3- \text{ Si } A \text{ y } B \text{ son eventos mutuamente excluyentes entonces } P(A \cup B) = P(A) + P(B)$$

$$4- \text{ Si } A_1, A_2, \dots, A_n, A_{n+1}, \dots \text{ es una secuencia de eventos tales que}$$

$$A_i \cap A_j = \emptyset \quad \text{si } i \neq j, \text{ entonces } P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Observación: supongamos el experimento de lanzar una moneda y el espacio muestral $S = \{c, s\}$. Notar que podemos escribir $S = \{c\} \cup \{s\}$, es decir como unión de eventos simples. Por los axiomas 2 y 3 de la definición de probabilidad tenemos que

$$1 = P(S) = P(\{c\}) + P(\{s\})$$

Es decir $1 = P(\{c\}) + P(\{s\})$, lo que implica que $P(\{c\}) = 1 - P(\{s\})$. No podemos deducir el valor de $P(\{c\})$ ni el valor de $P(\{s\})$ de las propiedades anteriores. Necesitamos información adicional, por ejemplo si el dado es normal. De ser así podemos plantear que $P(\{c\}) = P(\{s\})$, entonces

$$\left. \begin{array}{l} 1 = P(\{c\}) + P(\{s\}) \\ P(\{c\}) = P(\{s\}) \end{array} \right\} \Rightarrow P(\{c\}) = P(\{s\}) = 0.5$$

Podemos deducir de los axiomas otras propiedades útiles para el cálculo de probabilidades.

Propiedades de la probabilidad

$$1- P(\emptyset) = 0$$

Dem.) Siendo A un evento cualquiera podemos escribir $A \cup \emptyset = A$

Además A y \emptyset son mutuamente excluyentes, por lo tanto por axioma 3

$$P(A) = P(A \cup \emptyset) = P(A) + P(\emptyset)$$

O sea que

$$P(A) = P(A) + P(\emptyset) \Rightarrow P(\emptyset) = 0$$

$$2- \text{ Si } A^c \text{ es el evento complementario de } A, \text{ entonces } P(A^c) = 1 - P(A)$$

Dem.) Si A es un evento cualquiera, entonces podemos escribir $A \cup A^c = S$

Además, por definición de complemento de un conjunto, A y A^c son mutuamente excluyentes. Por lo tanto, por axioma 2 y por axioma 3

$$1 = P(S) = P(A) + P(A^c)$$

Despejando

$$P(A^c) = 1 - P(A)$$

3-

Si $A \subset B$ entonces $P(A) \leq P(B)$

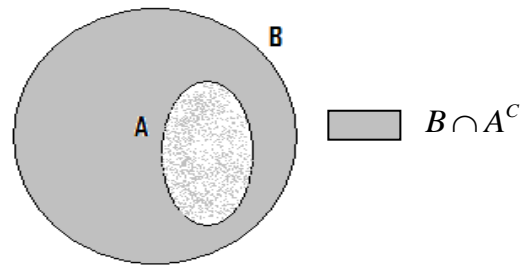
Dem.) Sean A y B dos eventos tales que $A \subset B$.

De la figura vemos que $B = A \cup (B \cap A^c)$

Y además A y $B \cap A^c$ son mutuamente excluyentes. Entonces

$$P(B) = P(A) + P(B \cap A^c)$$

Y como por axioma 1 tenemos que $P(B \cap A^c) \geq 0$, entonces $P(B) \geq P(A)$



4-

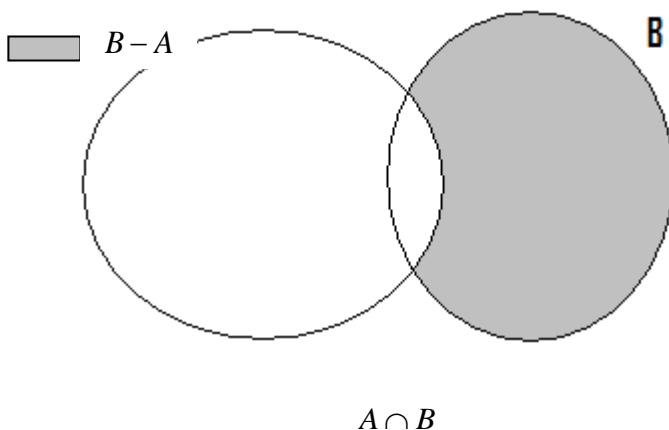
Si $A \subset B$, entonces $P(B - A) = P(B) - P(A)$

Dem.) Siguiendo los pasos de la demostración anterior llegamos a

$$P(B) = P(A) + P(B \cap A^c), \text{ lo que implica que } P(B \cap A^c) = P(B) - P(A)$$

Y como $B \cap A^c = B - A$, entonces $P(B - A) = P(B) - P(A)$.

Observación: *en general* vale la siguiente propiedad $P(B - A) = P(B) - P(A \cap B)$



5- Si A y B son dos eventos *cualesquiera*, entonces

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Dem.) Escribimos a $A \cup B$ como unión de partes disjuntas de la siguiente manera, (observar la figura)

$$A \cup B = B \cup (A \cap B^c)$$

$$\therefore P(A \cup B) = P(B) + P(A \cap B^c) \quad (1)$$

$$\text{hatched box} \quad A \cap B$$

$$\text{gray box} \quad A \cap B^c$$

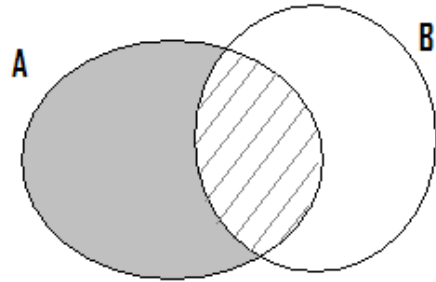
Y por la observación anterior

$$P(B \cap A^c) = P(B - A) = P(B) - P(A \cap B) \quad (2)$$

Por lo tanto, reemplazando (2) en (1):

$$P(A \cup B) = P(B) + P(A) - P(A \cap B)$$

Con lo que queda demostrada la propiedad.



Observaciones:

1- La propiedad anterior se puede generalizar para expresar la probabilidad de tres eventos

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Se llega a la igualdad anterior escribiendo $A \cup B = D$ y aplicando la propiedad 5, es decir:

$$\begin{aligned} P(A \cup B \cup C) &= P(D \cup C) = P(D) + P(C) - P(D \cap C) = \\ &= P(A \cup B) + P(C) - P((A \cup B) \cap C) \quad (3) \end{aligned}$$

Nuevamente aplicamos 5: $P(A \cup B) = P(B) + P(A) - P(A \cap B)$

Y además, aplicando las leyes distributivas $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

Entonces:

$$P((A \cup B) \cap C) = P((A \cap C) \cup (B \cap C)) = P(A \cap C) + P(B \cap C) - P((A \cap C) \cap (B \cap C))$$

Como $(A \cap C) \cap (B \cap C) = A \cap B \cap C$, reemplazando en (3) se llega a :

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

En general, se puede probar que si A_1, A_2, \dots, A_n son n eventos cualesquiera entonces

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_n)$$

2- Si A_1, A_2, \dots, A_n son n eventos tales que $A_i \cap A_j = \emptyset$ con $i \neq j$ entonces $P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$

Notar que la igualdad anterior es una generalización del axioma 3.

1.2 - El espacio muestral finito

Sea ε un experimento aleatorio y S un espacio muestral asociado a ε . Supongamos que S es un conjunto finito al que anotamos $S = \{a_1, a_2, \dots, a_n\}$, es decir consideramos que tiene n elementos.

Podemos escribir a S como unión de eventos elementales de la siguiente forma.

Si $A_i = \{a_i\}$ entonces $S = A_1 \cup A_2 \cup \dots \cup A_n$. Además sabemos que $P(S) = 1$, por lo tanto tenemos el resultado:

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1 \quad (4)$$

Es decir: **la suma de las probabilidades de los eventos elementales es igual a 1**

Si ahora tomamos un evento cualquiera B de S , lo anotamos $B = \{a_{i_1}, a_{i_2}, \dots, a_{i_k}\}$, entonces podemos escribir a B como unión de eventos simples: $B = A_{i_1} \cup A_{i_2} \cup \dots \cup A_{i_k}$. Por lo tanto

$$P(B) = P(A_{i_1} \cup A_{i_2} \cup \dots \cup A_{i_k}) = P(A_{i_1}) + P(A_{i_2}) + \dots + P(A_{i_k})$$

Es decir para calcular la probabilidad de B se suman las probabilidades de los eventos elementales que lo componen.

Si asumimos que $P(A_1) = P(A_2) = \dots = P(A_n) = p$, es decir todos los eventos elementales tienen la misma probabilidad de ocurrir, entonces reemplazando en (4):

$$\underbrace{p + p + \dots + p}_{n \text{ veces } p} = 1 \Rightarrow np = 1 \Rightarrow p = \frac{1}{n} = \frac{1}{\#S}$$

Donde el símbolo $\#S$ se lee: número de elementos de S , (en general $\#S$ indica el **cardinal** de S)

Cuando $P(A_1) = P(A_2) = \dots = P(A_n) = p$, se dice que S es **equiprobable**.

Si B es un evento cualquiera de S , lo anotamos $B = \{a_{i_1}, a_{i_2}, \dots, a_{i_k}\}$ (o sea que $\#B = k$), entonces escribimos a B como unión de eventos simples: $B = A_{i_1} \cup A_{i_2} \cup \dots \cup A_{i_k}$, y nuevamente

$$P(B) = P(A_{i_1} \cup A_{i_2} \cup \dots \cup A_{i_k}) = P(A_{i_1}) + P(A_{i_2}) + \dots + P(A_{i_k})$$

Como ya vimos que $P(A_i) = \frac{1}{n}$ para todo $i=1, 2, \dots, n$, se tiene que:

$$P(B) = P(A_{i_1}) + P(A_{i_2}) + \dots + P(A_{i_k}) = \underbrace{\frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n}}_{k \text{ veces}} = \frac{k}{n} = \frac{\#B}{\#S}$$

Es decir **si B es un evento de un espacio muestral equiprobable entonces**

$$P(B) = \frac{\#B}{\#S}$$

Ejemplo:

Supongamos que se tira una moneda normal tres veces, y se cuenta el número de caras obtenido luego de los tres tiros. Tomemos como espacio muestral a $S = \{0, 1, 2, 3\}$

¿Cuál es la probabilidad de que salgan al menos dos caras?

En este caso S **no es equiprobable**, pues

el evento $\{0\}$ ocurre de una sola forma: cuando en los tres tiros sale ceca

El evento $\{3\}$ ocurre de una sola forma: cuando en los tres tiros sale cara

Pero el evento $\{1\}$ ocurre cuando en los tres tiros sale una sola cara y eso puede darse de tres formas distintas (c, s, s) ; (s, s, c) o (s, c, s)

Análogamente para el evento $\{2\}$ puede darse de tres formas distintas:

(c, c, s) ; (c, s, c) o (s, c, c)

Por lo tanto $P(\{0\}) = P(\{3\})$ y $P(\{1\}) = P(\{2\})$. Si anotamos $P(\{0\}) = P(\{3\}) = p$ y $P(\{1\}) = P(\{2\}) = q$ entonces $q = 3p$ (5)

Además se cumple $P(\{0\}) + P(\{1\}) + P(\{2\}) + P(\{3\}) = 1$ (6)

Entonces de (5) y (6)

$$\left. \begin{array}{l} q = 3p \\ p + q + q + p = 1 \end{array} \right\} \Rightarrow \left. \begin{array}{l} q = 3p \\ 2p + 2q = 1 \end{array} \right\} \Rightarrow p = \frac{1}{8}, \quad q = \frac{3}{8}$$

De esta forma conocemos las probabilidades de todos los eventos elementales.

Si B : “salen al menos dos caras”, entonces $B = \{2, 3\}$

Para calcular $P(B)$ hacemos

$$P(B) = P(\{2\}) + P(\{3\}) = \frac{3}{8} + \frac{1}{8} = \frac{4}{8} = 0.5$$

Si tomamos como espacio muestral a

$$S = \{(c, c, c); (c, c, s); (c, s, c); (s, c, c); (c, s, s); (s, s, c); (s, c, s); (s, s, s)\}$$

Entonces S es **equiprobable** pues cada resultado (cada terna) se da una sola forma.

En este caso **podemos plantear directamente** $P(B) = \frac{\#B}{\#S} = \frac{4}{8} = 0.5$

1.3 - Métodos de conteo

Cuando se calculan probabilidades, algunas veces se necesita determinar el número de resultados en un espacio muestral o un evento. Se describirán diversos métodos con ese propósito.

La regla básica, que se conoce como principio fundamental de conteo, se presenta por medio del siguiente ejemplo:

Ejemplo: Cierta tipo de automóvil se encuentra disponible en tres colores: rojo, azul y verde, y puede tener un motor grande o pequeño. ¿De cuántas maneras puede un comprador elegir un automóvil?

Solución:

Hay tres opciones de color y dos opciones de motor. Una lista completa de las opciones se muestra en la siguiente tabla de 3×2 . El número total de opciones es $(3)(2) = 6$

	Rojo	Azul	Verde
Grande	Rojo y grande	Azul y grande	Verde y grande
Pequeño	Rojo y pequeño	Azul y pequeño	Verde y pequeño

Al generalizar el ejemplo, si hay n_1 elecciones de color y n_2 elecciones de motor, una lista completa de elecciones se puede escribir como una tabla de $n_1 \times n_2$ por lo que el número total de elecciones es $n_1 \cdot n_2$

Si una operación se puede realizar en n_1 maneras y si para cada una de esas maneras se puede realizar una segunda operación en n_2 maneras, entonces el número total de maneras en que se realizan las dos operaciones es $n_1 \cdot n_2$

Este razonamiento del principio fundamental del conteo se puede ampliar para cualquier número de operaciones.

El principio fundamental de conteo

Suponga que se pueden realizar k operaciones. Si hay n_1 maneras de realizar la primera operación y si para cada una de esas maneras hay n_2 maneras de realizar la segunda operación, y si para cada una de esas elecciones de esas de esas maneras de realizar las dos primeras operaciones hay n_3 maneras de realizar la tercera operación y así sucesivamente, entonces el número total de maneras de realizar la secuencia de las k operaciones es $n_1 \cdot n_2 \dots n_k$.

Ejemplo:

Cuando se hace un pedido de cierto tipo de computadora, hay tres elecciones de disco duro, cuatro de la cantidad de memoria, dos de la tarjeta de video y tres de monitor. ¿En cuántas maneras se puede solicitar una computadora?

Solución:

El número total es $(3)(4)(2)(3) = 72$

Podemos simbolizar cada operación con un guión y escribir debajo del mismo el número de formas en que puede realizarse

$$\begin{array}{cccc} \overset{1^\circ}{\rule{0.5cm}{0.4pt}} & \overset{2^\circ}{\rule{0.5cm}{0.4pt}} & \overset{3^\circ}{\rule{0.5cm}{0.4pt}} & \overset{4^\circ}{\rule{0.5cm}{0.4pt}} \\ 3 & 4 & 2 & 3 \end{array}$$

Permutaciones

Una **permutación** constituye un ordenamiento de un conjunto de elementos. Por ejemplo, hay 6 permutaciones de las letras A, B y C:

ABC ACB BAC BCA CAB CBA

Con solamente tres elementos, es fácil determinar el número de permutaciones, sólo con hacer una lista de todas ellas. Pero con un gran número de elementos esto último no sería factible. El principio fundamental de conteo se puede usar para determinar el número de permutaciones de cualquier conjunto de elementos. Por ejemplo, se puede determinar el número de permutaciones de un conjunto de tres elementos de la siguiente manera: hay tres

elecciones para colocar el primer elemento. Después que se hace la elección hay dos elecciones restantes para el elemento del segundo lugar. Entonces queda una elección para el elemento del último lugar. Por tanto, el número total de maneras de ordenar tres objetos es $(3)(2)(1) = 6$. Este razonamiento se puede generalizar. El número de permutaciones de un conjunto de n elementos es

$$n(n-1)(n-2)\dots(3)(2)(1)$$

Éste es el producto de los enteros del 1 al n . Este producto se puede escribir con el símbolo $n!$, que se lee “n factorial”.

Para cualquier entero positivo n , $n! = n(n-1)(n-2)\dots(3)(2)(1)$
También se define $0! = 1$

El número de permutaciones de n objetos es $n!$

Ejemplos:

1- Cinco personas están en hilera en un cine. ¿De cuántas formas diferentes se pueden ordenar?.

Solución:

El número de permutaciones de un conjunto de cinco personas es

$$5! = (5)(4)(3)(2)(1) = 120$$

2- a) ¿De cuántas formas se pueden ordenar en un estante 3 libros de Matemática y 4 de Física?

b) ¿De cuántas formas si los de Matemática deben estar juntos?

c) ¿De cuántas formas si los libros de Matemática deben estar juntos y los de Física también deben estar juntos?

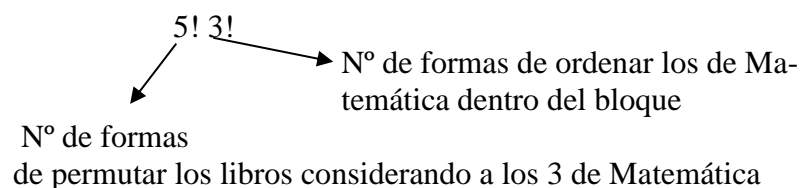
Solución:

a) $7!$

b) Si los de Matemática deben estar juntos, consideramos los 3 libros de Matemática como un solo libro, es decir pensamos que tenemos que ordenar 5 libros

$$\underbrace{M_1 \ M_2 \ M_3}_{M} \ F_1 \ F_2 \ F_3 \ F_4$$

Entonces considerando a los tres libros de Matemática en un bloque, la forma de ordenar los 7 libros es



como un solo libro.

- c) Razonando en forma similar al inciso anterior consideramos dos bloques: un bloque consiste en los 3 libros de Matemática, y el otro bloque consiste en los 4 libros de Física. El número de formas de permutar los bloques es $2!$. Dentro de cada bloque podemos permutar los libros: de $3!$ maneras a los de Matemática y de $4!$ formas a los de Física.
Por lo tanto la respuesta es $2! 3! 4!$.

A veces se está interesado en contar el número de permutaciones de los subconjuntos de cierto tamaño elegidos de un conjunto más grande. Esto se muestra en el siguiente ejemplo

Ejemplo:

Cinco salvavidas están disponibles para la guardia de un sábado por la tarde. Hay tres estaciones de salvavidas. ¿De cuántas maneras se pueden elegir y organizar los tres salvavidas entre las estaciones?

Solución:

Se usa el principio fundamental de conteo. Hay cinco maneras de elegir un salvavidas para que ocupe la primera estación, luego cuatro de elegir a un salvavidas para que ocupe la segunda estación y por último tres para elegir un salvavidas que ocupe la tercera estación. El número total de permutaciones de los tres salvavidas elegidos entre los cinco es $(5)(4)(3) = 60$.

El razonamiento usado para resolver el ejemplo anterior se puede generalizar. El número de permutaciones de k objetos elegidos de un grupo de n objetos es

$$n(n-1)\dots(n-k+1)$$

Esta expresión se puede simplificar utilizando la notación factorial:

El número de permutaciones de k objetos elegidos de un grupo de n elementos es

$$\frac{n!}{(n-k)!}$$

$$n(n-1)\dots(n-k+1) = \frac{n(n-1)\dots(n-k+1)(n-k)(n-k-1)\dots(3)(2)(1)}{(n-k)(n-k-1)\dots(3)(2)(1)} = \frac{n!}{(n-k)!}$$

Usamos el símbolo $P_{k,n} = \frac{n!}{(n-k)!}$

Los anteriores son casos de permutaciones donde los objetos se eligen a lo sumo una vez. Si un objeto puede elegirse más de una vez tenemos permutaciones con elementos repetidos

Ejemplo:

A la cumbre de una montaña conducen 5 caminos. ¿De cuántas formas puede realizar un turista el camino de ida y vuelta (subir y descender)?

Solución:

Hay 5 formas posibles de subir a la montaña y 5 formas de descender si consideramos que se puede elegir para el descenso el mismo camino que se tomó para el ascenso. Por lo tanto la respuesta es $(5)(5) = 25$

En general si se deben elegir k elementos de un conjunto de n objetos y cada objeto se puede elegir más de un vez entonces el número de formas posibles de realizar esta operación es

$$(n)(n).....(n) = n^k$$

El número de permutaciones de k objetos elegidos de un grupo de n elementos donde cada elemento se puede elegir más de una vez es

$$n^k$$

Simbolizamos

$$P_{k,n}^* = n^k$$

Combinaciones

En algunos casos, cuando se elige un conjunto de elementos de un conjunto más grande, no se tiene en cuenta el orden de los elementos elegidos, solo se consideran los elementos que se eligen. Por ejemplo, puede que no importe qué salvavidas ocupe cada estación, puede que solo sea importante la elección de tres salvavidas. A cada grupo distinto de elementos que se puede seleccionar, sin importar el orden, se le llama **combinación**.

Se mostrará cómo determinar el número de combinaciones de k elementos elegidos de un conjunto de n objetos. Se mostrará el razonamiento con el resultado del ejemplo anterior. En este ejemplo se mostró que hay 60 permutaciones de tres elementos elegidos entre cinco. Al denotar a los elementos por A, B, C, D, E en la figura siguiente se presenta una lista de las 60 permutaciones.

ABC	ABD	ABE	ACD	ACE	ADE	BCD	BCE	BDE	CDE
ACB	ADB	AEB	ADC	AEC	AED	BDC	BEC	BED	CED
BAC	BAD	BAE	CAD	CAE	DAE	CBD	CBE	DBE	DCE
BCA	BDA	BEA	CDA	CEA	DEA	CDB	CEB	DEB	DEC
CAB	DAB	EAB	DAC	EAC	EAD	DBC	EBC	EBD	ECD
CBA	DBA	EBA	DCA	ECA	EDA	DCB	ECB	EDB	EDC

Las 60 permutaciones están ordenadas en 10 columnas de 6 permutaciones cada una. Dentro de cada columna, los tres elementos son los mismos y la columna contiene las 6 permutaciones diferentes de esos tres elementos. Por lo tanto cada columna representa una combinación distinta de tres elementos elegidos entre cinco y hay diez combinaciones de este tipo. Por lo tanto la figura muestra que el número de combinaciones de tres elementos elegidos entre cinco se puede encontrar al

dividir el número de permutaciones de los tres elementos elegidos, o sea $5!/(5-3)!$, por el número de permutaciones de los tres elementos que es $3!$. En resumen, el número de combinaciones de los tres elementos elegidos es $\frac{5!}{3!(5-3)!}$.

A menudo el número de combinaciones de k elementos elegidos entre n se anota por el símbolo $\binom{n}{k}$ o también con el símbolo $C_{k,n}$.

El razonamiento utilizado para deducir el número de combinaciones de los tres elementos elegidos se puede generalizar para deducir una expresión para $\binom{n}{k}$.

El número de combinaciones de k elementos elegidos de un grupo de n elementos es

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Ejemplo: A cierto evento asisten 30 personas y se elegirá aleatoriamente a 5 para recibir premios. Estos últimos son iguales, así que el orden en que se elige a las personas no es importante. ¿Cuántos grupos diferentes de 5 personas se pueden elegir?

Solución:

En virtud de que el orden de las 5 personas elegidas no es importante, se tiene que calcular el número de combinaciones de 5 elegidas entre 30. Esto es

$$\binom{30}{5} = \frac{30!}{5!(30-5)!} = \frac{(30)(29)(28)(27)(26)}{(5)(4)(3)(2)(1)} = 142506$$

Elegir una combinación de k elementos de un conjunto de n elementos, divide a los n elementos en dos subconjuntos: k que fueron elegidos y $n-k$ que no fueron elegidos. A veces un conjunto se divide en más de dos subconjuntos. Por ejemplo, suponga que en una clase de 12 estudiantes se asignará un proyecto a los estudiantes para trabajar en equipos. Se formarán tres equipos, que constarán de 5, 4 y 3 estudiantes. Se puede calcular el número de maneras en las que se formarán los equipos del siguiente modo. Se considera el proceso para dividir la clase en 3 equipos como una secuencia de dos operaciones. La primera operación es seleccionar una combinación de 5 estudiantes para formar el equipo de 5. La segunda consiste en elegir una combinación de 4 estudiantes entre los 7 restantes, para formar el equipo de 4. El equipo de 3 automáticamente constará de los 3 estudiantes que quedan.

El número de maneras de realizar la primera operación es

$$\binom{12}{5} = \frac{12!}{5!7!}$$

Después de que se ha realizado la primera operación, el número de formas de realizar la segunda es

$$\binom{7}{4} = \frac{7!}{4!3!}$$

Por lo tanto el número total de maneras de realizar la secuencia de las dos operaciones es

$$\binom{12}{5} \binom{7}{4} = \frac{12!}{5!7!} \cdot \frac{7!}{4!3!} = \frac{12!}{5!4!3!} = 27720$$

Observe que el numerador en la respuesta final es el factorial del tamaño total del grupo, mientras que el denominador constituye el producto de los factoriales de los tamaños de los equipos elegidos de éste. Esto último es válido en general

El número de maneras de dividir un grupo de n elementos en grupos de k_1, \dots, k_r elementos, donde $k_1 + \dots + k_r = n$ es

$$\frac{n!}{k_1! \dots k_r!}$$

Ejemplo:

Se lanza un dado 20 veces. En virtud de que en 3 de las tiradas salió el número 1, en 5 de las tiradas salió el 2, en 4 salió el 3, en 2 el 4, en 3 salió el 5 y en 3 el 6, ¿cuántos arreglos diferentes de resultados hay?

Solución:

Hay 20 resultados, están divididos en 6 grupos: el grupo de 3 resultados en los que salió el 1, el de 5 en los que salió el 2 y así sucesivamente. El número de maneras de dividir los 20 resultados en 6 grupos de tamaños específicos es

$$\frac{20!}{3!5!4!2!3!3!}$$

Observación: Notar que en este ejemplo podemos razonar de la siguiente manera.

Se lanza un dado 20 veces, podemos interpretar entonces que tenemos 20 objetos donde hay 3 elementos iguales entre sí (los 3 número 1), 5 elementos iguales entre sí y diferentes de los anteriores (los 5 números 2), y así siguiendo.

Una posible forma de que esto ocurra sería que en los sucesivos tiros se hayan obtenido los números en el siguiente orden:

1 1 1 2 2 2 2 2 3 3 3 3 4 4 5 5 5 6 6 6

La pregunta es cuántas formas hay de **ordenar** estos 20 objetos. La respuesta no es $20!$ pues para esto deberían ser los 20 elementos **diferentes**.

Podemos pensar que hay un 1° tiro, un 2° tiro, etc., es decir tenemos 20 lugares diferentes y debemos elegir entre ellos, sin orden, 3 lugares, para ubicar los tiros en que salieron los número 1. Luego elegimos entre los 17 tiros que quedan 5 para ubicar los tiros en donde salieron los números 2, y así siguiendo.

Por lo tanto el número $\frac{20!}{3!5!4!2!3!3!}$ se podría interpretar como el número de formas de ordenar 20 elementos donde no todos son distintos.

En general $\frac{n!}{k_1! \dots k_r!}$ indicaría el número de formas de ordenar n elementos donde hay k_1 iguales entre sí, k_2 iguales entre sí y diferentes de los anteriores, ..., k_r iguales entre sí y diferentes de todos los anteriores donde $k_1 + \dots + k_r = n$

Ejemplo:

¿De cuántas formas pueden ordenarse las letras de la palabra INGENIERIA?

Solución:

La palabra INGENIERIA consta de 10 letras donde hay 3 letras I, 2 letras N, 2 letras E. Por lo tanto la respuesta es $\frac{10!}{3!2!2!}$

Principio de suma

Supongamos el siguiente ejemplo:

¿Cuántas cadenas de 8 bits comienzan con 101 o con 111?

Una cadena de 8 bits que comienza con 101 puede construirse con 5 operaciones: se elige el cuarto bit, luego el quinto bit y así siguiendo. Cada una de estas elecciones puede hacerse de 2 maneras, se elige un 0 ó se elige un 1. Por el principio fundamental del conteo, hay $2^5 = 32$ cadenas que comienzan con 101.

Análogamente, hay $2^5 = 32$ cadenas que comienzan con 111.

Si A es el conjunto de las cadenas de 8 bits que comienzan con 101 y B es el conjunto de las cadenas de 8 bits que comienzan con 111 entonces $A \cap B = \emptyset$. Por lo tanto el número de cadenas de 8 bits que comienzan con 101 o con 111 es el número de elementos de $A \cup B$, es decir $32 + 32 = 64$.

En general

Principio de suma

Si A_1, A_2, \dots, A_n son conjuntos tales que $A_i \cap A_j = \emptyset$ si $i \neq j$, entonces el número de elementos de $A_1 \cup A_2 \cup \dots \cup A_n$ es igual a la suma del número de elementos de cada conjunto, en símbolos

$$\#(A_1 \cup A_2 \cup \dots \cup A_n) = \#A_1 + \#A_2 + \dots + \#A_n$$

Ejemplos:

1- De un grupo de 4 hombres y 5 mujeres, ¿cuántos comités de 3 miembros son posibles

- a) sin restricciones?
- b) con 1 hombre y 2 mujeres?
- c) con más mujeres que hombres?

Solución:

a) En un comité no hay cargos, por lo tanto no importa el orden en el cual son elegidos los miembros, por lo tanto la respuesta es

$$\binom{9}{3} = \frac{9!}{3!(9-3)!} = \frac{9!}{3!6!} = \frac{(9)(8)(7)}{(3)(2)(1)} = 3 \times 4 \times 7 = 84$$

b) Si debe haber un hombre y 2 mujeres hay que contar el número de formas de elegir

un hombre entre 4 : $\binom{4}{1} = 4$

y el número de formas de elegir 2 mujeres entre 5:

$$\binom{5}{2} = \frac{5!}{2!3!} = \frac{(5)(4)(3)(2)(1)}{(2)(1)(3)(2)(1)} = 5 \times 2 = 10 .$$

Por el principio fundamental del conteo la respuesta es $4 \times 10 = 40$

c) Separamos en casos y aplicamos el principio de suma:

caso1: 2 mujeres y 1 hombre \longrightarrow 40

caso 2: 3 mujeres \longrightarrow $\binom{5}{3} = 10$

La respuesta es $40 + 10 = 50$

2- En el alfabeto Morse de puntos y rayas, ¿cuántos caracteres pueden formarse empleando a lo sumo 3 signos?

Solución:

Un carácter está formado por puntos y rayas.

Podríamos considerar 3 casos excluyentes:

caso1: carácter de 1 signo

En este caso el número de formas de construir un carácter es 2, se elige un punto o una raya

caso2: carácter de 2 signos

En este caso tenemos 2 posibles elecciones para cada signo, elegir un punto o una raya. Considerando que la elección para cada signo es una operación que se puede hacer de 2 formas distintas y que debemos realizar 2 operaciones (porque tenemos 2 signos) la respuesta es $(2)(2) = 2^2 = 4$

caso3: carácter de 3 signos

Nuevamente tenemos 2 posibles elecciones para cada signo, elegir un punto o una raya. Considerando que la elección para cada signo es una operación que se puede hacer de 2 formas distintas y que debemos realizar 3 operaciones (porque tenemos 3 signos) por lo tanto la respuesta es $(2)(2)(2) = 2^3 = 8$

Por el principio de suma el resultado final es $2 + 4 + 8 = 14$

3- Tres parejas de casados han comprado boletos para el teatro y se sientan en una fila formada por sólo 6 asientos.

a) ¿De cuántas maneras pueden sentarse de manera tal que Juan y Paula (marido y mujer) se sienten en los 2 asientos de la extrema izquierda?

b) ¿De cuántas maneras pueden sentarse de manera tal que Juan y Paula (marido y mujer) se sienten juntos?

- c) ¿De cuántas maneras pueden sentarse de manera tal que por lo menos una de las esposas termine sentada junto a su marido?

Solución:

- a) Simbolicemos la situación utilizando guiones para cada uno de los 6 asientos

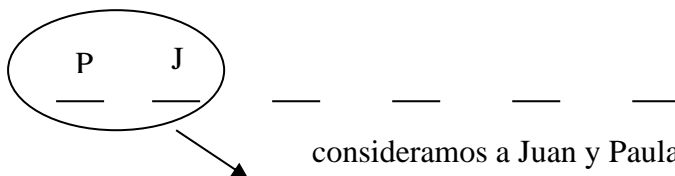
$\begin{array}{cccccc} P & J & & & & \\ \hline & & & & & \end{array}$

Permutamos a las 4 restantes personas pero además Juan y Paula pueden permutar entre sí, la respuesta es

$$4! 2!$$

donde $2!$ es el número de formas en que Juan y Paula pueden permutar entre sí.

- b) En este inciso a diferencia del anterior Juan y Paula pueden sentarse en cualquier asiento no necesariamente en el extremo izquierdo. Pensamos a Juan y Paula como una sola persona



consideramos a Juan y Paula como una sola persona, permutamos 5 personas

La respuesta sería $5! 2!$

- c) Podemos considerar 3 casos:

caso1: los integrantes de la pareja 1 están sentados juntos

caso2: los integrantes de la pareja 2 están sentados juntos

caso3: los integrantes de la pareja 3 están sentados juntos

Notar que **no son casos excluyentes**

Nos ayudamos con conjuntos escribiendo

A_i es el conjunto de todas las permutaciones de las 6 personas donde los miembros de la pareja i están sentados juntos con $i = 1, 2, 3$

Entonces observar que la respuesta es el número de elementos del conjunto

$A_1 \cup A_2 \cup A_3$ pero los conjuntos no son disjuntos de a dos.

Por lo tanto **no podemos aplicar el principio de suma**

Pero sí podemos escribir lo siguiente

$$\#(A_1 \cup A_2 \cup A_3) =$$

$$\#A_1 + \#A_2 + \#A_3 - \#(A_1 \cap A_2) - \#(A_1 \cap A_3) - \#(A_2 \cap A_3) + \#(A_1 \cap A_2 \cap A_3)$$

Además

$$\#A_1 + \#A_2 + \#A_3 = 5! 2! + 5! 2! + 5! 2!$$

$$\#(A_1 \cap A_2) = \#(A_1 \cap A_3) = \#(A_2 \cap A_3) = 4! \cdot 2! \cdot 2!$$

pues consideramos a los miembros de dos de las parejas en un bloque, es decir tenemos dos bloques y los dos integrantes de la pareja restante que pueden no estar sentados uno al lado del otro, en total 4 elementos para permutar. Además dentro de cada bloque los integrantes de cada pareja pueden permutar entre sí.

$$\#(A_1 \cap A_2 \cap A_3) = 3! \cdot 2! \cdot 2! \cdot 2! \quad \text{pues ahora tenemos 3 bloques}$$

El resultado final es $5! \cdot 2! + 5! \cdot 2! + 5! \cdot 2! - 3(4! \cdot 2! \cdot 2!) + 3! \cdot 2! \cdot 2! \cdot 2!$

1.4 – Cálculo de probabilidades en espacios muestrales finitos equiprobables

Ejemplos:

1- En el caso de espacios muestrales equiprobables es necesario recordar las **técnicas de conteo** para calcular el número de elementos de un conjunto sin enumerar sus elementos.

Por ejemplo:

Se tiene una urna con 15 bolillas distinguibles (pueden estar numeradas), de las cuales 10 son blancas y 5 son rojas. Se extraen **al azar** dos bolillas de la urna,

- ¿cuál es la probabilidad de extraer todas blancas?
- ¿cuál es la probabilidad de extraer exactamente una bolilla blanca?
- ¿cuál es la probabilidad de extraer al menos una bolilla blanca?

Supongamos que se extraen las dos bolillas **sin reemplazo** y **no importa el orden** en que son extraídas.

Al decir que se extraen al azar, se entiende que el espacio muestral del experimento es equiprobable.

Al decir que se extraen las bolillas sin reemplazo, significa que se extrae una bolilla y sin regresaarla a la urna se extrae la siguiente.

El espacio muestral de este experimento se puede tomar como $S = \{\{1,4\}, \{2,5\}, \{2,10\}, \dots\}$, es decir como el conjunto cuyos elementos son conjuntos, todos los conjuntos de dos elementos que se puedan formar a partir de un conjunto de 15 elementos. Por ejemplo, el resultado $\{1,4\}$ significa que se extrajeron las bolillas 1 y 4. ¿Cuántos elementos tiene S ?, recordando las técnicas de

conteo el número de elementos de S es igual al número combinatorio $\binom{15}{2} = \frac{15!}{2!(15-2)!}$

Para responder la pregunta **a)** anotemos A : “las dos bolillas son blancas”

Entonces $P(A) = \frac{\#A}{\#S}$. ¿Cuántos elementos tiene el conjunto A ?, tantos como formas haya de

extraer dos bolillas blancas. Se tiene que $\#A = \binom{10}{2} = \frac{10!}{2!(10-2)!}$.

$$\text{Por lo tanto } P(A) = \frac{\#A}{\#S} = \frac{\binom{10}{2}}{\binom{15}{2}} = \frac{\frac{10!}{2!(10-2)!}}{\frac{15!}{2!(15-2)!}} = \frac{3}{7}$$

Veamos ahora el inciso **b)**. Anotamos B : “se extrae exactamente una bolilla blanca”
Entonces

$$P(B) = \frac{\#B}{\#S} = \frac{\binom{10}{1}\binom{5}{1}}{\binom{15}{2}}$$

Y por último el inciso **c)**. Anotamos C : “se extrae al menos una bolilla blanca”
Podemos resolverlo de dos maneras:

La forma más fácil es calcular $P(C^c)$ donde C^c : “ninguna bolilla es blanca”

$$P(C^c) = \frac{\#C^c}{\#S} = \frac{\binom{5}{2}}{\binom{15}{2}} \Rightarrow P(C) = 1 - P(C^c) = 1 - \frac{\binom{5}{2}}{\binom{15}{2}} = 1 - \frac{2}{21} = \frac{19}{21}$$

Otra forma de resolverlo podría ser la siguiente, escribimos $C = C_1 \cup C_2$ donde

C_i : “se extraen i bolillas blancas exactamente” $i = 1, 2$

Notar que $C_1 \cap C_2 = \emptyset$. Por lo tanto $P(C) = P(C_1) + P(C_2)$

$$\text{Y } P(C_1) = \frac{\#C_1}{\#S} = \frac{\binom{10}{1}\binom{5}{1}}{\binom{15}{2}} \quad ; \quad P(C_2) = \frac{\#C_2}{\#S} = \frac{\binom{10}{2}\binom{5}{0}}{\binom{15}{2}}$$

$$\text{Por lo tanto } P(C) = \frac{\binom{10}{1}\binom{5}{1}}{\binom{15}{2}} + \frac{\binom{10}{2}\binom{5}{0}}{\binom{15}{2}} = \frac{10}{21} + \frac{9}{21} = \frac{19}{21}$$

2- Si en el ejemplo anterior importa el orden en que son extraídas las bolillas, entonces

$S = \{(1,4), (2,5), (2,10), (4,1), (5,2), \dots\}$, es decir que S es el conjunto de todos los **pares** (a,b) donde $a \neq b$, por ejemplo $(1,4)$ expresa el resultado: primero se extrajo la bolilla 1 y en segundo término se extrajo la bolilla 2. El número de elementos de S es $15 \times 14 = 210$.

En este caso

$$P(A) = \frac{10 \times 9}{15 \times 14} = \frac{3}{7};$$

Para calcular la probabilidad de B hay que distinguir si la bolilla blanca se extrae primero o se extrae en segundo término:

$$P(B) = \frac{10 \times 5 + 5 \times 10}{15 \times 14} = \frac{2(10 \times 5)}{15 \times 14}$$

Para calcular la probabilidad de C nuevamente podemos pensarlo de dos formas, pero si consideramos $C = C_1 \cup C_2$ en cada caso hay que tener en cuenta el orden en que se extraen las bolillas.

Es mas práctico hacer $P(C) = 1 - P(C^c) = 1 - \frac{5 \times 4}{15 \times 14}$

3- Si ahora **importa el orden** en que son extraídas las bolillas y además la extracción se hace **con reemplazo** (se extrae una bolilla y se la devuelve a la urna antes de extraer la siguiente), entonces S es el conjunto de pares (a, b) donde ahora puede ocurrir que $a = b$.

En este caso $\#S = 15 \times 15$, y calculamos

$$P(A) = \frac{10 \times 10}{15 \times 15} \quad ; \quad P(B) = \frac{10 \times 5 + 5 \times 10}{15 \times 15} = \frac{2(10 \times 5)}{15 \times 15} \quad ; \quad P(C) = 1 - P(C^c) = 1 - \frac{5 \times 5}{15 \times 15}$$

1.5 - Espacios muestrales infinitos

Sea S un espacio muestral **infinito numerable** es decir $S = \{a_1; a_2; a_3; \dots\}$. Como en el caso finito, a cada a_i asignamos un número $p_i = P(\{a_i\})$ tal que

$$a) \quad p_i \geq 0 \quad b) \quad \sum_{i=1}^{\infty} p_i = 1$$

La probabilidad $P(A)$ de un evento A es entonces la suma de las probabilidades de los eventos elementales que lo componen.

Los únicos espacios muestrales **infinitos no numerables** que consideraremos son aquellos de medida geométrica finita $m(S)$ tales como longitud, área o volumen, y en los cuales un punto se selecciona al azar. La probabilidad de un evento A , esto es aquella en la que el punto seleccionado pertenece a A , es entonces la relación entre $m(A)$ a $m(S)$, es decir

$$P(A) = \frac{\text{longitud de } A}{\text{longitud de } S} \quad \text{ó} \quad P(A) = \frac{\text{área de } A}{\text{área de } S} \quad \text{ó} \quad P(A) = \frac{\text{volumen de } A}{\text{volumen de } S}$$

Ejemplo:

En el interior de un círculo se selecciona un punto al azar. Hallar la probabilidad de que el punto quede más cercano al centro que a la circunferencia.

Tomamos como espacio muestral a $S = \{(x, y), x^2 + y^2 \leq r^2\}$, entonces

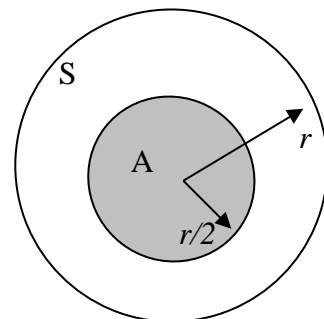
$$A = \left\{ (x, y), x^2 + y^2 \leq \left(\frac{r}{2}\right)^2 \right\}$$

$$\text{Por lo tanto } P(A) = \frac{\text{área de } A}{\text{área de } S} = \frac{\pi \left(\frac{r}{2}\right)^2}{\pi r^2} = \frac{1}{4}$$

Observación: si $A = \left\{ (x, y), x^2 + y^2 = \left(\frac{r}{2}\right)^2 \right\}$,

$$\text{entonces } P(A) = \frac{\text{área de } A}{\text{área de } S} = \frac{0}{\pi r^2} = 0, \text{ pero } A \neq \emptyset$$

Por lo tanto si un evento tiene probabilidad 0 eso no implica que sea el evento vacío.



Práctica
Espacios muestrales y eventos. Asignación de probabilidades

- 1) Un experimento implica lanzar un par de dados, uno verde y uno rojo, y registrar los números que salen. Si x es igual al resultado en el dado verde e y es el resultado en el dado rojo, describa el espacio muestral S
 - a) por extensión
 - b) por comprensión
- 2) Para el espacio muestral del ejercicio 1) liste los elementos del eventos:
 - a) A: “la suma de los números es mayor que 8”
 - b) B: “ocurre un dos en cualquiera de los dos dados”
 - c) C: “sale un número mayor que cuatro en el dado verde”
 - d) $A \cap C$
 - e) $A \cap B$
 - f) $B \cap C$
- 3) Dos estaciones de servicio, A y B, se encuentran en cierto cruce de caminos; en cada una hay seis surtidores de nafta. Consideramos el experimento en que el número de surtidores en uso en un día particular se determina para cada una de las estaciones.
 - a) Describir el espacio muestral del experimento. ¿Cuántos elementos tiene?
 - b) Expresar explícitamente los siguientes eventos:
 - i) C: el número de surtidores en uso es el mismo para ambas estaciones de servicio.
 - ii) D: el número total de surtidores en uso es cuatro.
 - iii) E: por lo menos un surtidor está en uso en cada estación de servicio.
 - iv) ¿Cuáles de los sucesos C, D o E son mutuamente excluyentes?
 - v) sean los eventos A_i : la estación de servicio A tiene en uso exactamente i surtidores, $i = 0, 1, 2, 3, 4, 5, 6$; y B_i : la estación de servicio B tiene en uso i surtidores exactamente, $i = 0, 1, 2, 3, 4, 5, 6$. Expresar los eventos anteriores en función de $A_i, B_i, i = 0, 1, 2, 3, 4, 5, 6$.
- 4) Suponga que los dos dados del ejercicio 1) son normales. Entonces cada resultado del espacio muestral S tienen la misma probabilidad de ocurrir (S es equiprobable). Encuentre las siguientes probabilidades:
 - a) $P(A)$; b) $P(B)$; $P(C)$; $P(A \cap C)$
- 5) Se construye un dado de manera que el 1 y el 2 ocurran con el doble de la frecuencia que se presenta el 5, el cual ocurre con una frecuencia 3 veces superior al 3, al 4, o al 6. Si se lanza una vez, encuentre la probabilidad de que
 - a) el número sea par.
 - b) el número sea mayor que 4.
- 6) Una clase consta de 10 hombres y 20 mujeres de los cuales la mitad de los hombres y la mitad de las mujeres tienen ojos castaños. Hallar la probabilidad de que una persona escogida al azar sea un hombre o tenga los ojos castaños.
- 7) A un individuo se le presentan tres vasos diferentes de bebida de cola, marcadas C, D, y P. Se le pide degustar las tres y luego ponerlas en lista de preferencia. Supongamos que la misma bebida de cola es la que en realidad se ha puesto en los tres vasos.

- a) ¿Cuáles son los eventos elementales en este experimento de clasificación, y probabilidad se asignaría a cada uno?
 - b) ¿Cuál es la probabilidad de que C sea clasificada primera?
 - c) ¿Cuál es la probabilidad de que C sea clasificada primera y D última?
- 8) Suponga que se lanzan dos monedas normales y que se observa el resultado de las caras que quedan hacia arriba. Si se tira primero una moneda y después la otra:
- a) Establezca los puntos muestrales de este experimento.
 - b) Asigne una probabilidad razonable a cada punto. ¿Son los puntos igualmente probables?
 - c) Sea A el evento de observar exactamente una vez cara y B el evento de observar al menos una cara. Obtenga los puntos muestrales de A y B.
 - d) A partir de la respuesta en c), calcule $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup B)$, y $P(A^c \cup B)$.
 - e) Repetir a), b), c), d) si se tiran las dos monedas simultáneamente, y el experimento consiste en observar el número de caras que se obtienen.

Practica (OPCIONAL)
Métodos de conteo

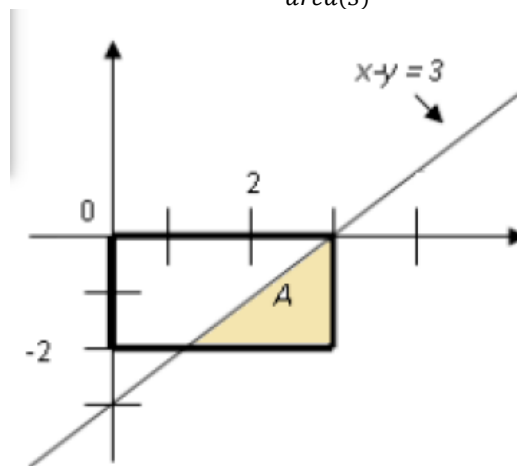
- 1) a) ¿Cuántas banderas diferentes de 3 bandas distintas se pueden formar con los colores verde, rojo, azul, blanco, amarillo y negro?
 - b) ¿Cuántas se pueden formar si se permite también la primera banda igual a la tercera?
- 2) a) ¿Cuántos números de cuatro cifras pueden formarse?
- b) ¿Cuántos con todas las cifras distintas?
 - c) ¿Cuántos con todas las cifras distintas y que terminen en cero?
- 3) ¿De cuántas maneras pueden disponerse las cinco letras de la palabra ARBOL?
- 4) Una torta necesita 8 ingredientes, ¿de cuántas maneras se pueden mezclar?
- 5) ¿De cuántas maneras se pueden colocar en fila 10 niños, si uno determinado de ellos no puede estar a la cabeza?
- 6) Se quieren disponer 4 libros de matemática distintos, 6 distintos de física y 2 distintos de química en un estante. ¿De cuántas maneras puede hacerse si:
- a) los libros de un mismo tema deben estar juntos?
 - b) solo es obligatorio que los libros de matemática estén juntos?
- 7) ¿En cuántas permutaciones de la palabra AYCXYDYTV las tres Y aparecen juntas?
- 8) ¿Cuántos números de 5 cifras pueden formarse con los dígitos 1, 2, 3, 4, 5, 6, 7, 8, 9 si:
- a) los números deben ser impares?
 - b) los números deben tener las dos primeras cifras pares?
- 9) ¿Cuántos números de 6 cifras se pueden formar a partir de las cifras del número 327332?
- 10) a) ¿De cuántas maneras pueden disponerse las letras de la palabra PAPANATA?
- b) ¿De cuántas, de manera que la T esté al principio?. ¿Y al final?
- 11) ¿De cuántas maneras puede elegirse entre 9 personas un comité de 5?

- 12) Dados 5 puntos del plano no alineados de a 3, ¿cuántos triángulos determinan?
- 13) Se elige un comité de 5 personas entre 30 hombres y 20 mujeres.
- ¿De cuántas maneras se puede hacer la elección?
 - ¿De cuántas si el número de mujeres debe sobrepasar al de hombres?
- 14) ¿Cuántas palabras de 4 consonantes distintas y 3 vocales distintas pueden formarse con 9 consonantes y 5 vocales?
- 15) Una planta de producción emplea 20 trabajadores en el turno de día, 15 trabajadores en el segundo turno y 10 en el turno noche. Un consultor de control de calidad selecciona 6 de estos trabajadores para hacerles una entrevista.
- ¿De cuántas maneras se pueden seleccionar 6 trabajadores que provengan del turno día?
 - ¿De cuántas maneras se pueden seleccionar 6 trabajadores que provengan del mismo turno?

Práctica
Asignación de probabilidades utilizando métodos de conteo
Espacios muestrales infinitos

- 1) Se escogen al azar 3 lámparas entre 15.
- ¿De cuántas formas se puede hacer la elección si:
 - no importa el orden en que son elegidas?
 - importa el orden en que son elegidas y la elección se hace sin sustitución?
 - importa el orden en que son elegidas y la elección se hace con sustitución?
 - De las 15 lámparas 5 son defectuosas. Para cada uno de los experimentos anteriores hallar la probabilidad de los siguientes eventos:
 - ninguna sea defectuosa
 - una exactamente sea defectuosa.
 - una por lo menos sea defectuosa.
- 2) Se seleccionan al azar dos cartas entre 10 cartas numeradas de 1 a 10. Hallar la probabilidad de que la suma sea impar si:
- las dos cartas se sacan juntas.
 - se sacan una tras otra sin sustitución.
 - las dos cartas se sacan una después de la otra con sustitución.
- 3) Una caja posee 6 tuercas y 10 tornillos. Si se escogen al azar 3 piezas, hallar la probabilidad de:
- seleccionar 3 tuercas
 - seleccionar exactamente 2 tuercas
 - seleccionar por lo menos una tuerca
 - seleccionar exactamente 2 tornillos
- 4) Tres moléculas de tipo A, tres de tipo B, tres de tipo C y tres de tipo D deben combinarse para formar una molécula en cadena. Una de estas moléculas en cadena es ABCDABCDABCD, y otra es BCDDAAABDBCC.
- ¿Cuántas de estas moléculas en cadena hay?.

- b) Supongamos que se selecciona al azar una molécula en cadena del tipo descrito. ¿Cuál es la probabilidad de que las tres moléculas de cada tipo terminen una junto a otra?
- 5) Tres parejas de casados han comprado boletos para el teatro y se sientan en una fila formada por solo seis asientos. Si toman sus asientos de un modo totalmente aleatorio
- a) ¿Cuál es la probabilidad de que Juan y Paula (marido y mujer), se sienten en los dos asientos de la extrema izquierda?
- b) ¿Cuál es la probabilidad de que Juan y Paula terminen sentados uno junto a otro?
- 6) Sobre la línea real \mathbb{R} , se seleccionan al azar los puntos a y b tales que $-2 \leq b \leq 0$ y $0 \leq a \leq 3$: Hallar la probabilidad de que la distancia entre a y b sea mayor que 3. (Ayuda: el espacio muestral consta de todos los pares ordenados $(a; b)$ de la región rectangular indicada en la figura de abajo. Por otra parte el evento $A = \{(a, b) \in S; |a - b| > 3\}$ consta de aquellos pares de S que caen debajo de la recta $x - y = 3$. Entonces $P(A) = \frac{\text{area}(A)}{\text{area}(S)}$)



- 7) Un ingeniero se dispone a visitar dos fábricas A y B. Los transportes que lo llevan a las fábricas parten del mismo lugar. A las horas en punto salen los que lo llevan a la fábrica A y a las horas y $\frac{1}{4}$ los que lo llevan a la fábrica B. Si sale de su casa sin preocuparse por la hora y toma el primer transporte que sale, ¿cuál es la probabilidad de que visite la fábrica A?. ¿Y la B?.

2 - Probabilidad condicional

2.1- Definición

Supongamos el experimento aleatorio de extraer al azar sin reemplazo dos bolillas de una urna que contiene 7 bolillas rojas y 3 blancas. Asumimos que las bolillas de un mismo color son distinguibles.

Consideramos los eventos A: “la primer bolilla extraída es blanca”, y B: “la segunda bolilla extraída es blanca”.

El espacio muestral S se puede pensar como el conjunto

$$S = \{(a, b); a = 1, 2, \dots, 10; b = 1, 2, \dots, 10; a \neq b\} \text{ y } \#S = 10 \times 9.$$

Es claro que $P(A) = \frac{3 \times 9}{10 \times 9} = \frac{3}{10}$. Pero si queremos calcular $P(B)$ no es tan directo. Podemos calcular

la **probabilidad de B sabiendo que A ocurrió**: es igual a $\frac{2}{9}$, ya que si A ocurrió, entonces en la urna quedaron 9 bolillas de las cuales 2 son blancas. La probabilidad anterior la anotamos $P(B/A)$ y se lee: probabilidad condicional de B dado A. Es decir $P(B/A) = \frac{2}{9}$.

Notar que podemos interpretar lo anterior de la siguiente forma: el espacio muestral original S se ha **reducido** al evento A, es decir **se toma a A como nuevo espacio muestral** para calcular la probabilidad de B.

También podemos interpretar que la probabilidad condicional de B dado A debería ser la proporción de veces que ocurre $A \cap B$ con respecto al número de veces que ocurre A. Pensando en términos de frecuencia relativa: $P(B/A) \approx \frac{n_{A \cap B}}{n_A}$.

Esta idea motiva la siguiente definición:

Sean A y B dos eventos de un espacio muestral S . La probabilidad condicional de B dado A se define como

$$P(B/A) = \frac{P(A \cap B)}{P(A)} \quad \text{si } P(A) \neq 0$$

Análogamente

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad \text{si } P(B) \neq 0$$

En algunos casos se puede calcular $P(B/A)$ directamente reduciendo el espacio muestral. En otros será necesario aplicar la definición anterior.

Observación: si A y B son eventos de un espacio muestral S **equiprobable**, entonces

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{\#(A \cap B)}{\#A}$$

Ejemplos:

- 1- En el experimento de extraer dos bolillas

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{3 \times 2}{10 \times 9}}{\frac{3}{10}} = \frac{2}{9}$$

- 2- Se tira un dado normal dos veces. Sean los eventos A: “la suma de los números obtenidos es 6” y B: “el primer número es igual a 4” Tenemos que $A = \{(1,5); (2,4); (3,3); (4,2); (5,1)\}$ y $B = \{(4,1); (4,2); (4,3); (4,4); (4,5); (4,6)\}$ Entonces para calcular $P(A/B)$ mediante la definición de probabilidad condicional

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{36}}{\frac{6}{36}} = \frac{1}{6}$$

También podemos calcularlo en forma directa, reduciendo el espacio muestral, de todos los pares del evento B, observamos cuáles cumplen con lo requerido por A, es decir de todos los pares de B, solo uno tiene la propiedad de que sus componentes suman 6, por lo tanto

$$P(A/B) = \frac{1}{6}$$

- 3- Se lanza una moneda normal tres veces.

Hallar la probabilidad de que salgan todas caras si sale alguna cara.

El espacio muestral reducido es el evento A: “sale alguna cara”

Tenemos que $A = \{(c, c, c); (c, c, s); (c, s, c); (s, c, c); (c, s, s); (s, s, c); (s, c, s)\}$

Y $B = \{(c, c, c)\}$

Por lo tanto $P(B/A) = \frac{1}{7}$

- 4- En cierta ciudad, 40% de la población tiene cabellos castaños, 25% tiene ojos castaños y 15% tiene cabellos y ojos castaños. Se escoge una persona al azar
- a) si tiene cabellos castaños, ¿cuál es la probabilidad de que también tenga ojos castaños?
 - b) Si tiene ojos castaños, ¿cuál es la probabilidad de que no tenga cabellos castaños?
 - c) ¿Cuál es la probabilidad de que no tenga ni cabellos ni ojos castaños?

Sean los eventos A: “la persona elegida al azar tiene ojos castaños”, B: “la persona elegida al azar tiene cabellos castaños”

Entonces $P(A) = 0.25$, $P(B) = 0.40$ y $P(A \cap B) = 0.15$

a) Se pide calcular $P(A/B)$:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0.15}{0.40}$$

$$b) P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{0.15}{0.25}$$

$$c) P(A^c \cap B^c) = P((A \cup B)^c) = 1 - P(A \cup B) = 1 - (0.25 + 0.40 - 0.15)$$

- 5- Sean los eventos A y B con $P(A) = 0.5$, $P(B) = \frac{1}{3}$ y $P(A \cap B) = \frac{1}{4}$.

Hallar: a) $P(A/B)$, b) $P(B/A)$, c) $P(A \cup B)$, d) $P(A^c/B^c)$, e) $P(B^c/A^c)$

$$\text{a) } P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{1}{3}} = \frac{3}{4} \quad \text{b) } P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

$$\text{c) } P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{4} = \frac{7}{12}$$

$$\text{d) } P(A^c/B^c) = \frac{P(A^c \cap B^c)}{P(B^c)}$$

$$\text{Calculamos: } P(B^c) = 1 - P(B) = 1 - \frac{1}{3} = \frac{2}{3} ;$$

$$P(A^c \cap B^c) = P((A \cup B)^c) = 1 - P(A \cup B) = 1 - \frac{7}{12} = \frac{5}{12}$$

Por la ley de De Morgan

$$\text{Entonces } P(A^c/B^c) = \frac{P(A^c \cap B^c)}{P(B^c)} = \frac{\frac{5}{12}}{\frac{2}{3}} = \frac{15}{24} = \frac{5}{8}$$

$$\text{e) } P(B^c/A^c) = \frac{P(A^c \cap B^c)}{P(A^c)} = \frac{P(A^c \cap B^c)}{1 - P(A)} = \frac{\frac{5}{12}}{1 - \frac{1}{2}} = \frac{5}{6}$$

Observaciones:

$$\text{a) Si } A \subset B \text{ entonces } P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

$$\text{b) Si } A \cap B = \emptyset \text{ entonces } P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{P(\emptyset)}{P(A)} = 0$$

c) Es fácil comprobar que $P(B/A)$ para A fijo, satisface los axiomas de la probabilidad, esto es:

$$1- 0 \leq P(B/A) \leq 1$$

$$2- P(S/A) = 1$$

3-Si B_1 y B_2 son eventos mutuamente excluyentes entonces

$$P((B_1 \cup B_2)/A) = P(B_1/A) + P(B_2/A)$$

4-Si $B_1, B_2, \dots, B_n, B_{n+1}, \dots$ es una secuencia de eventos tales que $B_i \cap B_j = \emptyset$ si $i \neq j$ entonces

$$P\left(\left(\bigcup_{i=1}^{\infty} B_i\right)/A\right) = \sum_{i=1}^{\infty} P(B_i/A)$$

Teorema de la multiplicación

Si A y B son dos eventos entonces $P(B/A) = \frac{P(A \cap B)}{P(A)}$ si $P(A) \neq 0$

Por lo tanto

$$P(A \cap B) = P(B/A) P(A) \quad (6)$$

Análogamente de $P(A/B) = \frac{P(A \cap B)}{P(B)}$ si $P(B) \neq 0$, se deduce

$$P(A \cap B) = P(A/B) P(B) \quad (7)$$

(6) y (7) se conocen como **teorema de la multiplicación**.

Consideremos otra vez el ejemplo de extraer dos bolillas al azar sin reemplazo de una urna que contiene 3 bolillas blancas y 7 rojas. Si A : “la primera bolilla extraída es blanca”, y B : “la segunda bolilla extraída es blanca”, entonces

$$P(A \cap B) = P(A)P(B/A) = \frac{3}{10} \times \frac{2}{9}$$

Si A_1, A_2, A_3 son tres eventos entonces

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2)$$

pues:

$$P(A_1 \cap A_2 \cap A_3) = P(B \cap A_3) = P(B)P(A_3/B) = P(A_1 \cap A_2)P(A_3/A_1 \cap A_2) =$$

\swarrow
 $B = A_1 \cap A_2$

\searrow
 Teorema de la multiplicación

$$= P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2)$$



Teorema de la multiplicación

El teorema de la multiplicación se puede generalizar a n eventos A_1, A_2, \dots, A_n :

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2) \dots P(A_n/A_1 \cap A_2 \cap \dots \cap A_{n-1}) \quad (8)$$

Ejemplos:

1- Una clase tiene 12 niños y 4 niñas. Si se escogen tres estudiantes de la clase al azar, ¿cuál es la probabilidad de que sean todos niños?

Solución:

Anotamos A_i : “el i -ésimo estudiante elegido es un niño” $i = 1, 2, 3$

Entonces la probabilidad pedida es

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2) = \frac{12}{16} \times \frac{11}{15} \times \frac{10}{14}$$

- 2- Los estudiantes de una clase se escogen al azar, uno tras otro, para presentar un examen.
- Si la clase consta de 4 niños y 3 niñas, ¿cuál es la probabilidad de que niños y niñas queden alternados?
 - Si la clase consta de 3 niños y 3 niñas, ¿cuál es la probabilidad de que niños y niñas queden alternados?

Solución:

- a) Nuevamente anotamos A_i : “el i-ésimo estudiante elegido es un niño”

Entonces la probabilidad pedida es, aplicando (8):

$$P(A_1 \cap A_2^c \cap A_3 \cap A_4^c \cap A_5 \cap A_6^c \cap A_7) = \frac{4}{7} \times \frac{3}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} \times \frac{1}{2} \times \frac{1}{1} = \frac{1}{35}$$

- b) Hay dos casos mutuamente excluyentes: el primer estudiante es un niño, y el primero es una niña.

Si el primero es un niño, entonces por (8) la probabilidad de que los estudiantes se alternen es

$$P(A_1 \cap A_2^c \cap A_3 \cap A_4^c \cap A_5 \cap A_6^c) = \frac{3}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} \times \frac{1}{2} \times \frac{1}{1} = \frac{1}{20}$$

Si el primero es una niña, entonces por (8) la probabilidad de que los estudiantes se alternen es

$$P(A_1^c \cap A_2 \cap A_3^c \cap A_4 \cap A_5^c \cap A_6) = \frac{3}{6} \times \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} \times \frac{1}{2} \times \frac{1}{1} = \frac{1}{20}$$

Entonces la probabilidad pedida es $\frac{1}{20} + \frac{1}{20} = \frac{1}{10}$

En el ejemplo inicial de extraer dos bolillas de una urna, todavía queda por resolver cómo calculamos la $P(B)$, siendo A : “la primer bolilla extraída es blanca”, y B : “la segunda bolilla extraída es blanca”

Podemos expresar al espacio muestral S como $S = A \cup A^c$

Además A y A^c son mutuamente excluyentes.

Por otro lado podemos escribir

$$B = B \cap S = B \cap (A \cup A^c) = (B \cap A) \cup (B \cap A^c) \quad \text{por la ley distributiva}$$

Entonces

$$P(B) = P(B \cap A) + P(B \cap A^c) = P(B/A)P(A) + P(B/A^c)P(A^c)$$

$$(B \cap A) \cap (B \cap A^c) = \emptyset \quad \text{teorema de la multiplicación}$$

Lo hecho en este ejemplo se generaliza en el siguiente teorema

2.2 - Teorema de la probabilidad total

Sean A_1, A_2, \dots, A_n eventos de un espacio muestral S que cumplen:

- $A_1 \cup A_2 \cup \dots \cup A_n = S$
- $A_i \cap A_j = \emptyset$ si $i \neq j$
- $P(A_i) > 0 \quad \forall i = 1, 2, \dots, n$ Se dice que A_1, A_2, \dots, A_n forman una **partición de S** Entonces para cualquier evento B de S

$$P(B) = P(B/A_1)P(A_1) + P(B/A_2)P(A_2) + \dots + P(B/A_n)P(A_n)$$

Dem.) Podemos escribir

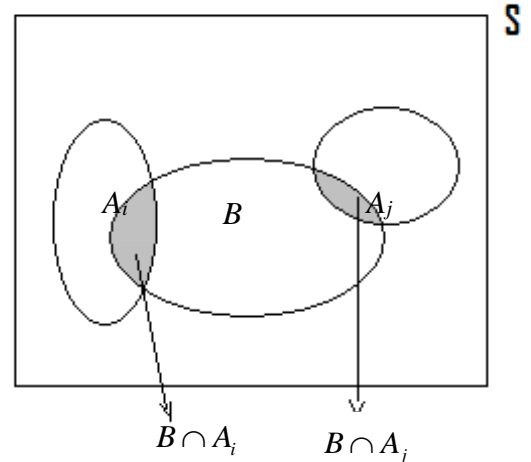
$$B = B \cap S = B \cap (A_1 \cup A_2 \cup \dots \cup A_n) = (B \cap A_1) \cup (B \cap A_2) \cup \dots \cup (B \cap A_n)$$

Ley distributiva de la \cap con respecto a la \cup

Además si $A_i \cap A_j = \emptyset$ con $i \neq j$, entonces
 $(B \cap A_i) \cap (B \cap A_j) = \emptyset$ si $i \neq j$ (ver figura)

Por lo tanto
 $P(B) = P(B \cap A_1) + P(B \cap A_2) + \dots + P(B \cap A_n) =$

Teorema de la multiplicación



$$= P(B / A_1)P(A_1) + P(B / A_2)P(A_2) + \dots + P(B / A_n)P(A_n)$$

Teorema de Bayes

Sean A_1, A_2, \dots, A_n eventos de un espacio muestral S que cumplen:

- a) $A_1 \cup A_2 \cup \dots \cup A_n = S$
- b) $A_i \cap A_j = \emptyset$ si $i \neq j$
- c) $P(A_i) > 0 \quad \forall i = 1, 2, \dots, n$

Entonces para cualquier evento B de S tal que $P(B) > 0$

$$P(A_k / B) = \frac{P(B / A_k)P(A_k)}{\sum_{i=1}^n P(B / A_i)P(A_i)} \quad k = 1, \dots, n$$

Dem.)

$$P(A_k / B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(B / A_k)P(A_k)}{P(B)} = \frac{P(B / A_k)P(A_k)}{\sum_{i=1}^n P(B / A_i)P(A_i)}$$

def. de prob. Condicional
teor. de la probabilidad total
teor. de multiplicación

Ejemplos:

- 1- Tres máquinas A, B, y C producen respectivamente 60%, 30% y 10% del número total de artículos de una fábrica. Los porcentajes de desperfectos de producción de estas máquinas son respectivamente 2%, 3% y 4%. Se selecciona un artículo al azar
 - a) ¿Cuál es la probabilidad de que sea defectuoso?
 - b) Si al seleccionar un artículo al azar resulta defectuoso, ¿cuál es la probabilidad de que el artículo hubiera sido producido por la máquina C?

Solución:

- a) Sean los eventos

A: "el artículo seleccionado fue producido por la máquina A"

B: "el artículo seleccionado fue producido por la máquina B"

C: "el artículo seleccionado fue producido por la máquina C"

D: "el artículo seleccionado es defectuoso"

Los datos que tenemos son los siguientes

$$P(A) = 0.6 \quad P(B) = 0.3 \quad P(C) = 0.1 \quad P(D/A) = 0.02 \quad P(D/B) = 0.03 \quad P(D/C) = 0.04$$

Se pide hallar la $P(D)$.

Se aplica el teorema de la probabilidad total tomando como partición de S a los eventos A , B y C .

Entonces

$$P(D) = P(D/A)P(A) + P(D/B)P(B) + P(D/C)P(C) = 0.02 \times 0.6 + 0.03 \times 0.3 + 0.04 \times 0.1$$

b) Se pide hallar $P(C/D)$. Aplicamos el teorema de Bayes

$$P(C/D) = \frac{P(C \cap D)}{P(D)} = \frac{P(D/C)P(C)}{P(D)} = \frac{0.04 \times 0.1}{0.02 \times 0.6 + 0.03 \times 0.3 + 0.04 \times 0.1} = \frac{4}{25}$$

2- Se nos dan tres urnas como sigue:

Una urna 1 contiene 3 bolas rojas y 5 blancas.

Una urna 2 contiene 2 bolas rojas y 1 blanca.

Una urna 3 contiene 2 bolas rojas y 3 blancas

Se selecciona una urna al azar y se saca una bola de la urna. Si la bola es roja, ¿cuál es la probabilidad de que proceda de la urna 1?

Solución:

Sean los eventos A_i : "se elige la urna i " $i = 1, 2, 3$

Entonces $P(A_i) = 1/3 \quad i = 1, 2, 3$

Además podemos tomar a A_1, A_2, A_3 como una partición del espacio muestral S .

Sea el evento B : "extraer bolilla roja"

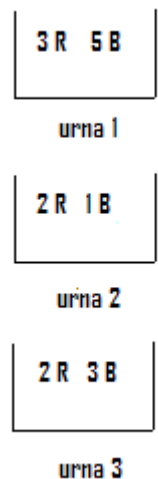
Se pide calcular $P(A_1/B)$

Entonces

$$P(A_1/B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{P(B/A_1)P(A_1)}{P(B/A_1)P(A_1) + P(B/A_2)P(A_2) + P(B/A_3)P(A_3)} =$$

\swarrow def. de prob. condicional \searrow Teorema de Bayes

$$= \frac{\frac{3}{8} \times \frac{1}{3}}{\frac{3}{8} \times \frac{1}{3} + \frac{2}{3} \times \frac{1}{3} + \frac{2}{5} \times \frac{1}{3}} = \frac{\frac{1}{8}}{\frac{173}{360}} = \frac{45}{173}$$



2.3 - Independencia

Dados dos eventos A y B , puede ocurrir que $P(B/A)$ y $P(B)$ sean diferentes, eso significa que saber que A ocurrió modifica la probabilidad de ocurrencia de B

En el ejemplo anterior $P(B/A_1) = \frac{3}{8} \neq P(B) = \frac{173}{360}$

Pero puede suceder que $P(B/A)$ y $P(B)$ sean iguales, en ese caso A y B son eventos independientes, saber que A ocurrió no afecta la probabilidad de ocurrencia de B .

Entonces, dos eventos A y B son **independientes** si $P(B/A) = P(B)$, y son **dependientes** de otro modo.

Notar que por el teorema de la multiplicación $P(A \cap B) = P(B/A)P(A)$ si $P(A) > 0$

Entonces

A y B son independientes $\Rightarrow P(A \cap B) = P(B/A)P(A) = P(B)P(A)$

Recíprocamente

Si $P(A \cap B) = P(B)P(A)$ entonces

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B)P(A)}{P(A)} = P(B) \quad \therefore \quad A \text{ y } B \text{ son independientes}$$

\downarrow
 si $P(A) > 0$

Por lo tanto:

$$A \text{ y } B \text{ son independientes si y solo si } P(A \cap B) = P(A)P(B) \quad (9)$$

Es decir podemos usar (9) como definición de independencia de eventos.

Ejemplos:

1-Se tira un dado normal dos veces, sean los eventos

A : “la suma de los números obtenidos es igual a 7”

B : “el primer número obtenido es 4”

¿Son A y B independientes?

Sabemos que el espacio muestral es el conjunto de 36 pares ordenados (a,b) donde tanto a como b pueden tomar los valores 1, 2, 3, 4, 5, 6.

Además $A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$ y $B = \{(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)\}$

Entonces

$$P(A \cap B) = \frac{1}{36} \quad P(A) = \frac{6}{36} = \frac{1}{6} \quad P(B) = \frac{6}{36} = \frac{1}{6}$$

Como $P(A \cap B) = \frac{1}{36} = P(A)P(B) = \frac{1}{6} \times \frac{1}{6}$ entonces A y B son independientes

Observación: si A fuera el evento A : “la suma de los números obtenidos es igual a 6”, entonces A y B son dependientes

2-Se tiene una urna con 10 bolillas blancas y 5 rojas. Se extraen al azar dos bolillas **con reemplazo** de la urna. Entonces los eventos

A : “la primer bolilla es blanca”

B : “la segunda bolilla es roja”

Son independientes

$$P(A \cap B) = \frac{10 \times 5}{15 \times 15} \quad P(A) = \frac{10 \times 15}{15 \times 15} = \frac{10}{15} \quad P(B) = \frac{5 \times 15}{15 \times 15} = \frac{5}{15}$$

Pero si la extracción se hace *sin reemplazo*, entonces A y B son *dependientes* pues

$$P(A \cap B) = P(B/A)P(A) = \frac{5}{14} \times \frac{10}{15} \quad P(A) = \frac{10}{15}$$

$$\text{y } P(B) = P(B/A)P(A) + P(B/A^c)P(A^c) = \frac{5}{14} \times \frac{10}{15} + \frac{4}{14} \times \frac{5}{15} = \frac{5}{15}$$

por lo tanto $P(A \cap B) \neq P(B)P(A)$

Notar que la diferencia está en que $P(B/A) = \frac{5}{14} \neq P(B) = \frac{5}{15}$

Observación: si en el ejemplo anterior la urna tiene 1000 bolillas blancas y 500 rojas, y se extraen *sin reemplazo* dos bolillas al azar entonces

$$P(B/A) = \frac{500}{1499} = 0.3335557... \quad \text{y} \quad P(B) = \frac{5}{15} = 0.3333333...$$

O sea que $P(B/A)$ y $P(B)$ son *casi* iguales.

Por lo tanto podemos asumir que A y B son independientes, aunque la extracción se haga sin reemplazo.

En la práctica, si N es el tamaño de la población y n el tamaño de la muestra extraída sin reemplazo, si $\frac{n}{N} < 0.05$ entonces podemos operar como si la extracción se hubiera hecho con reemplazo.

Si dos eventos A y B son independientes entonces A y B^c son independientes (10)

Dem.) se debe probar que $P(A \cap B^c) = P(A)P(B^c)$

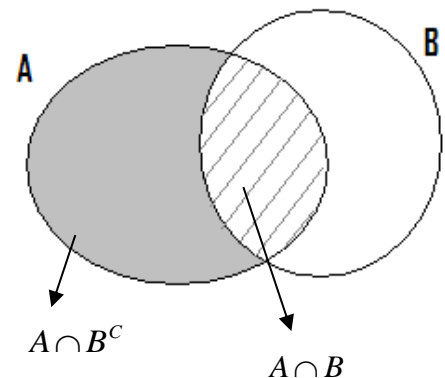
Para esto escribimos

$$A = (A \cap B) \cup (A \cap B^c)$$

Como $A \cap B$ y $A \cap B^c$ son mutuamente excluyentes entonces

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^c) \Rightarrow \\ \Rightarrow P(A \cap B^c) &= P(A) - P(A \cap B) \xrightarrow{\text{A y B independientes}} P(A) - P(A)P(B) = \\ &= P(A)(1 - P(B)) = P(A)P(B^c) \end{aligned}$$

Con lo que queda demostrada la propiedad.



Observación: si A y B son eventos independientes, entonces A^C y B^C son independientes. Se llega a este resultado aplicando (10) sobre A y B , y luego se aplica (10) nuevamente a A y B^C .

Independencia de más de dos eventos.

La noción de independencia de eventos se puede ampliar a n eventos de la siguiente manera:

Sean A_1, A_2, \dots, A_n eventos, se dice que son *independientes* si

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k}) \quad k = 2, \dots, n \quad (11)$$

Observaciones:

- 1- si $n = 2$ entonces (11) se reduce a la definición de dos eventos independientes.
- 2- si $n = 3$ entonces (11) significa que se deben cumplir las siguientes 4 condiciones:

- a) $P(A_1 \cap A_2) = P(A_1)P(A_2)$
- b) $P(A_1 \cap A_3) = P(A_1)P(A_3)$
- c) $P(A_2 \cap A_3) = P(A_2)P(A_3)$
- d) $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$

La condición d) significa que un evento es independiente de la intersección de los otros dos, por ejemplo $P(A_1 / A_2 \cap A_3) = P(A_1)$

Esto es porque en general por el teorema de la multiplicación vale que

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2 / A_1)P(A_3 / A_1 \cap A_2)$$

y por d)

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$$

entonces

$$P(A_1)P(A_2 / A_1)P(A_3 / A_1 \cap A_2) = P(A_1)P(A_2)P(A_3) \Rightarrow P(A_3 / A_1 \cap A_2) = P(A_3)$$

Ejemplos:

- 1- Las probabilidades de que tres hombres peguen en el blanco son, respectivamente, $\frac{1}{6}$, $\frac{1}{4}$, y $\frac{1}{3}$.

Cada uno dispara una vez al blanco.

- a) ¿Cuál es la probabilidad de que exactamente uno de ellos pegue en el blanco?
- b) Si solamente uno pega en el blanco, ¿cuál es la probabilidad de que sea el primer hombre?

Solución:

- a) consideremos los eventos A_i : “el hombre i -ésimo pega en el blanco” $i = 1, 2, 3$

$$P(A_1) = \frac{1}{6} \quad P(A_2) = \frac{1}{4} \quad P(A_3) = \frac{1}{3}$$

Sea el evento B : “exactamente un hombre pega en el blanco”

$$\text{Entonces } B = (A_1^C \cap A_2^C \cap A_3) \cup (A_1^C \cap A_2 \cap A_3^C) \cup (A_1 \cap A_2^C \cap A_3^C)$$

$$\text{Por lo tanto } P(B) = P(A_1^C \cap A_2^C \cap A_3) + P(A_1^C \cap A_2 \cap A_3^C) + P(A_1 \cap A_2^C \cap A_3^C)$$

$$\text{Y por independencia } P(B) = P(A_1^C)P(A_2^C)P(A_3) + P(A_1^C)P(A_2)P(A_3^C) + P(A_1)P(A_2^C)P(A_3^C) =$$

$$= \left(1 - \frac{1}{6}\right) \left(1 - \frac{1}{4}\right) \frac{1}{3} + \left(1 - \frac{1}{6}\right) \frac{1}{4} \left(1 - \frac{1}{3}\right) + \frac{1}{6} \left(1 - \frac{1}{4}\right) \left(1 - \frac{1}{3}\right) = \frac{1}{12} + \frac{5}{36} + \frac{5}{24} = \frac{31}{72}$$

b) Se pide calcular $P(A_1 / B)$

$$P(A_1 / B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{P(A_1 \cap A_2^c \cap A_3^c)}{P(B)} = \frac{\frac{1}{6} \left(1 - \frac{1}{4}\right) \left(1 - \frac{1}{3}\right)}{\frac{31}{72}} = \frac{6}{31}$$

2- Cierta tipo de proyectil da en el blanco con probabilidad 0.3, ¿ cuántos proyectiles deberán ser disparados para que haya al menos un 80% de probabilidad de pegar en el blanco?

Solución:

Escribimos A_i : “el proyectil i -ésimo da en el blanco” $i = 1, 2, \dots, n$

Se quiere que

$$P(A_1 \cup A_2 \cup \dots \cup A_n) > 0.8$$

Asumiendo independencia esto es equivalente a

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n) &= 1 - P((A_1 \cup A_2 \cup \dots \cup A_n)^c) = 1 - P(A_1^c \cap A_2^c \cap \dots \cap A_n^c) = \\ &= 1 - P(A_1^c)P(A_2^c) \dots P(A_n^c) = 1 - 0.7^n > 0.8 \end{aligned}$$

$$\text{Por lo tanto } 0.7^n < 0.2 \Rightarrow n \ln(0.7) < \ln(0.2) \Rightarrow n > \frac{\ln(0.2)}{\ln(0.7)} = 4.5123$$

Es decir se deben hacer por lo menos 5 disparos

Practica**Probabilidad condicional - Sucesos independientes****Teorema de la probabilidad total – Fórmula de Bayes.**

- 1) Se lanza un par de dados normales. Hallar la probabilidad de que la suma de sus números sea 10 o mayor si
 - a) aparece un 5 en el primer dado
 - b) aparece un 5 en uno de los dos dados por lo menos.
- 2) Se lanzan 3 monedas normales. Hallar la probabilidad de que sean todas caras si
 - a) la primera de las monedas es cara
 - b) una de las monedas es cara.
- 3) Se escogen dos dígitos al azar del 1 al 9. si la suma es par, hallar la probabilidad de que ambos números sean impares.
- 4) Sean los eventos A y B con $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ y $P(A \cap B) = \frac{1}{4}$. Hallar:
 - a) $P(A/B)$; b) $P(B/A)$; c) $P(A \cup B)$; d) $P(A^C / B^C)$; e) $P(B^C / A^C)$
- 5) Una clase tiene 12 niños y 4 niñas. Si se escogen 3 estudiantes de la clase al azar, ¿cuál es la probabilidad de que sean todos niños?
- 6) Una urna contiene 3 bolas rojas y 7 blancas. Se saca una bola de la urna y se reemplaza por una del otro color. Se saca de la urna una segunda bola.
 - a) Hallar la probabilidad de que la segunda bola sea roja
 - b) Si ambas bolas son del mismo color, ¿cuál es la probabilidad de que las dos sean blancas?
- 7) Una ciudad tiene dos carros de bomberos que operan en forma independiente. La probabilidad de que un carro específico esté disponible cuando se lo necesite es 0.96.
 - a) ¿Cuál es la probabilidad de que ninguno esté disponible cuando se les necesite?
 - b) ¿Cuál es la probabilidad de que un carro de bomberos esté disponible cuando se le necesite?
- 8) Una caja contiene 2 caramelos de coco y 3 de chocolate. Una segunda caja contiene 3 caramelos de coco, 2 caramelos de chocolate y 1 de dulce de leche. Si se saca un caramelo al azar de cada caja, encuentre la probabilidad de que
 - a) ambos caramelos sean de coco.
 - b) ningún caramelo sea de coco.
 - c) los dos caramelos sean diferentes.
- 9) En una prueba de opción múltiple, un estudiante contesta una pregunta que ofrece cuatro posibles respuestas, de las cuales sólo una es correcta. Suponga que la probabilidad de que el estudiante sepa la respuesta a la pregunta es 0.8 y que conteste al azar es 0.2
 - a) ¿Cuál es la probabilidad de que conteste correctamente la pregunta?
 - b) Si contesta correctamente la pregunta. ¿Cuál es la probabilidad de que realmente sepa la respuesta?

ta correcta?

10) Se lanza cinco veces un dado normal. Hallar la probabilidad de que:

- a) en ninguna tirada salga el 1
- b) salga el 1 una sola vez.
- c) salga el 1 al menos una vez.

11) a) Si $P(A / B) = 0.4$, $P(B) = 0.8$ y $P(A) = 0.6$, ¿puede decirse que los eventos A y B son independientes?

b) Si, $P(A / B) = 0.3$, $P(B) = 0.8$ y $P(A) = 0.3$, ¿puede decirse que los eventos A^C y B son independientes?

12) En una cierta estación de servicio, el 40% de los clientes utilizan nafta normal sin plomo, 35% utilizan nafta extra sin plomo, y el 25% utilizan nafta premium sin plomo. De los clientes que consumen nafta normal, solo 30% llenan sus tanques, de los que consumen nafta extra, 60% llenan sus tanques, en tanto que de los que usan premium, 50% llenan sus tanques.

- a) ¿Cuál es la probabilidad de que el siguiente cliente pida nafta extra sin plomo y llene su tanque?
- b) ¿Cuál es la probabilidad de que el siguiente cliente llene el tanque?
- c) Si el siguiente cliente llena el tanque, ¿cuál es la probabilidad de que pida nafta normal?
- d) ¿Qué propiedades utiliza para resolver los incisos a), b) y c)?

13) El 10% de los chips informáticos vendidos en el mercado son producidos por una empresa “pirata”.

Para un chip “pirata” la probabilidad de que sea defectuosos es del 50% mientras que si el chip no es “pirata” la probabilidad de que sea defectuoso desciende al 5%.

- a) Definir los sucesos convenientes, junto con sus probabilidades.
- b) Determinar el porcentaje total de chips defectuosos que salen al mercado.
- c) Se compra un chip y resulta ser defectuoso. Calcular la probabilidad de que proceda de la empresa “pirata”.

14) Se utilizan dos líneas de producción para empaquetar azúcar en bolsas de 5 kg. La línea 1 produce el doble de bolsas que la línea 2. Uno por ciento de las bolsas de la línea 1 están defectuosas ya que no cumplen con una especificación de calidad, mientras que 3% de las bolsas de la línea 2 están defectuosas. Se elige aleatoriamente una bolsa para inspeccionarla.

- a) ¿Cuál es la probabilidad de que provenga de la línea 1?
- b) ¿Cuál es la probabilidad de que esté defectuosa?
- c) Si la bolsa está defectuosa, ¿cuál es la probabilidad de que venga de la línea 1?
- d) Si la bolsa no está defectuosa, ¿cuál es la probabilidad de que venga de la línea 1?

3 - VARIABLES ALEATORIAS

3.1- Generalidades

En muchas situaciones experimentales se quiere asignar un número real a cada uno de los elementos del espacio muestral. Al describir el espacio muestral de un experimento un resultado individual no tiene que ser necesariamente un número, por ejemplo, al tirar una moneda y tomar como espacio muestral $S = \{c, s\}$, o al tirar un dado dos veces tomamos como espacio muestral a $S = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$, aquí S es un conjunto de pares ordenados.

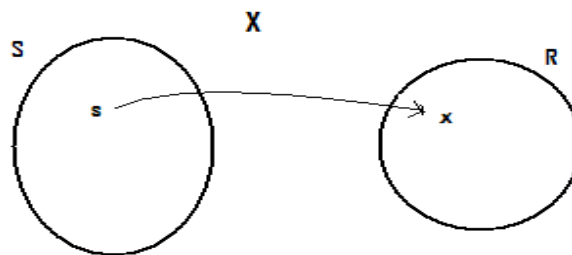
Definición: Sea ε un experimento aleatorio y S un espacio muestral asociado a él. Una **variable aleatoria** es una función que asigna a cada elemento de S un número real.

Notación: se anota a una variable aleatoria con letras mayúsculas X, Y, Z, W, \dots

Entonces, si X es una variable aleatoria de S en R

$$X : S \rightarrow R \quad \text{tal que} \quad X(s) = x$$

Con diagramas de Venn



Desde ahora en lugar de escribir variable aleatoria, escribiremos v.a.

Ejemplos:

1- Se tira una moneda tres veces

Sea X la v.a. X : “número de caras obtenidas luego de los tres tiros”

Si tomamos como espacio muestral

$$S = \{(c, c, c); (c, c, s); (c, s, c); (s, c, c); (c, s, s); (s, s, c); (s, c, s); (s, s, s)\}$$

entonces

$$X((c, c, c)) = 3$$

$$X((c, c, s)) = X((s, c, c)) = X((c, s, c)) = 2$$

$$X((c, s, s)) = X((s, c, s)) = X((s, s, c)) = 1$$

$$X((s, s, s)) = 0$$

La imagen de esta función es el conjunto $\{0,1,2,3\}$

Dada una v.a. X a su imagen se la anota R_X y se la denomina **rango o recorrido de X**

En el ejemplo anterior $R_X = \{0,1,2,3\}$

2- Se tira un dado tantas veces como sean necesarias hasta que sale el número 1 por primera vez.

Podemos simbolizar el espacio muestral de la siguiente manera $S = \{1, 01, 001, 0001, \dots\}$, por ejemplo 001 simboliza el resultado que en los dos primeros tiros no salió el número 1 y en el tercer tiro salió el 1.

Sea Y la v.a.:

Y : “número de tiros necesarios hasta que sale el 1 por primera vez”

Entonces $R_Y = \{1, 2, 3, 4, \dots\}$, es decir el rango de Y es el conjunto de los números naturales.

3- En el interior de un círculo de radio r y centro el origen de coordenadas, se elige un punto al azar.

Tomamos como espacio muestral a $S = \{(x, y), x^2 + y^2 \leq r^2\}$. Aquí S es infinito no numerable

Definimos la v.a. Z : “distancia del punto elegido al origen”

Entonces $R_Z = \{z; 0 \leq z \leq r\}$

Las variables aleatorias se clasifican según su rango.

Sea X es una v.a. con rango R_X . Si R_X es un conjunto *finito o infinito numerable* entonces se dice que X es una *v.a. discreta*. Si R_X es un conjunto *infinito no numerable* entonces X es una *v.a. continua*.

El rango R_X es considerado un nuevo *espacio muestral*, y sus subconjuntos son *eventos*.

Por ejemplo:

En el ejemplo 1, los eventos unitarios o elementales son $\{0\}; \{1\}; \{2\}; \{3\}$, *pero los anotamos* $\{X = 0\}; \{X = 1\}; \{X = 2\}; \{X = 3\}$

Otros eventos son, por ejemplo:

$\{X \leq 1\}$, es decir, salió a lo sumo una cara.

Notar que podemos escribir $\{X \leq 1\} = \{X = 0\} \cup \{X = 1\}$, o sea escribimos al evento como *unión de eventos elementales*

$\{X > 0\}$, es decir, salieron una o más caras. Tenemos que $\{X > 0\} = R_X - \{X = 0\}$

En el ejemplo 2, $\{Y \geq 4\}$ sería el evento “al menos 4 tiros son necesarios para que salga por primera vez el número 1”

$\{4 \leq Y \leq 6\}$ sería el evento “se necesitan entre 4 y 6 tiros para que salga el 1 por primera vez”

Notar que $\{4 \leq Y \leq 6\} = \{Y = 4\} \cup \{Y = 5\} \cup \{Y = 6\}$

En el ejemplo 3, $\left\{\frac{1}{3}r < Z < \frac{2}{3}r\right\}$ sería el evento “el punto elegido se encuentra a una distancia del

centro mayor que $\frac{1}{3}r$, pero menor que $\frac{2}{3}r$

Volviendo al ejemplo 1, notar que $B = \{X = 0\}$ ocurre en R_X si y solo si el evento $A = \{(s, s, s)\}$ ocurre en S . Se dice que A y B son *eventos equivalentes*.

De la misma forma los eventos

$A = \{(c, c, c)\}$ y $B = \{X = 3\}$ son equivalentes

$A = \{(c, c, s), (c, s, c), (s, c, c)\}$ y $B = \{X = 2\}$ son equivalentes

En general

siendo $A \subset S$ y $B \subset R_X$, A y B son *equivalentes* si $A = \{s \in S; X(s) \in B\}$

Si X es una v.a. de S en R , y R_X es el rango de X , para calcular la probabilidad de un evento B de R_X se busca el evento A en S equivalente a B y entonces $P(B) = P(A)$

Por ejemplo,

En el ejemplo 1, $P(B) = P(X = 0) = P(A) = P(\{(s, s, s)\}) = \frac{1}{8}$ si la moneda es normal.

$P(B) = P(X \leq 1) = P(\{(s, s, s); (c, s, s); (s, c, s); (s, s, c)\}) = \frac{4}{8} = 0.5$ si la moneda es normal

También podríamos haber planteado

$$P(B) = P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{1}{8} + \frac{3}{8} = 0.5$$

En el ejemplo 3, si $B = \left\{ \frac{1}{3}r \leq Z \leq \frac{2}{3}r \right\}$, entonces B es equivalente a

$A = \left\{ (x, y); \frac{1}{3}r \leq \sqrt{x^2 + y^2} \leq \frac{2}{3}r \right\}$, por lo tanto

$$P(B) = P(A) = \frac{\text{área de } A}{\text{área de } S} = \frac{\pi \left(\frac{2}{3}r \right)^2 - \pi \left(\frac{1}{3}r \right)^2}{\pi r^2} = \left(\frac{2}{3} \right)^2 - \left(\frac{1}{3} \right)^2 = \frac{3}{9} = \frac{1}{3}$$

Observación: en este ejemplo si $B = \left\{ Z = \frac{2}{3}r \right\}$, entonces

$$P(B) = P(A) = \frac{\text{área de } A}{\text{área de } S} = \frac{0}{\pi r^2} = 0$$

3.2 - Variables aleatorias discretas

Sea X una v.a. discreta. Anotamos su rango como $R_X = \{x_1, x_2, \dots, x_n\}$ si el rango es un conjunto finito de n elementos, y anotamos $R_X = \{x_1, x_2, \dots\}$ si el rango es un conjunto infinito numerable.

A cada x_i se le asigna un número $p(x_i) = P(X = x_i)$. Estos números deben satisfacer las condiciones siguientes

a) $p(x_i) \geq 0$ para todo i

b) $\sum_i p(x_i) = 1$

La función $p(x)$ que antes se definió, se llama **función de probabilidad o de frecuencia de la v.a. X** .

El conjunto de pares $(x_i, p(x_i))$ $i = 1, 2, \dots$ es la **distribución de probabilidad de X** .

Por ejemplo

1-Se tira una moneda normal tres veces, sea la v.a. X : “número de caras obtenidas”

Entonces $R_X = \{0, 1, 2, 3\}$

Para hallar la distribución de probabilidad de X supongamos que la probabilidad de salir cara es $\frac{1}{2}$ entonces

$$P(X = 0) = P(\{(s, s, s)\}) = \frac{1}{8}$$

$$P(X = 1) = P(\{(c, s, s); (s, c, s); (s, s, c)\}) = \frac{3}{8}$$

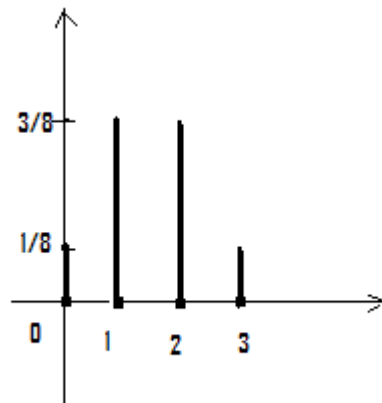
$$P(X = 2) = P(\{(c, c, s); (s, c, c); (c, s, c)\}) = \frac{3}{8}$$

$$P(X = 3) = P(\{(c, c, c)\}) = \frac{1}{8}$$

Se puede presentar la distribución de probabilidad de X en una tabla de la siguiente forma

x	0	1	2	3
$p(x)$	1/8	3/8	3/8	1/8

Un gráfico de la distribución de probabilidad de X sería



2-Se tira un dado normal. Sea X : “número que queda en la cara superior” Entonces $R_X = \{1, 2, 3, 4, 5, 6\}$
La función de distribución de X es

x	1	2	3	4	5	6
$p(x)$	1/6	1/6	1/6	1/6	1/6	1/6

Observación:

Sea X una v.a. discreta con rango finito $R_X = \{x_1, x_2, \dots, x_n\}$, donde cada x_i es un número entero y x_{i+1} es el consecutivo de x_i .

Si $P(X = x_i) = \frac{1}{n}$ para cada i entonces se dice que X tiene **distribución uniforme discreta**.

Por ejemplo podría ser $R_X = \{1, 2, 3, \dots, n-1, n\}$, en este caso X es uniforme discreta en el intervalo natural $[1, n]$.

La v.a. del ejemplo 2 anterior es uniforme discreta en el intervalo natural $[1, 6]$

Función de distribución acumulada

Sea X una v.a. con rango R_X . Se define la **función de distribución acumulada de X** (abreviamos F.d.a. de X) como

$$F(x) = P(X \leq x) \quad -\infty < x < \infty \quad (12)$$

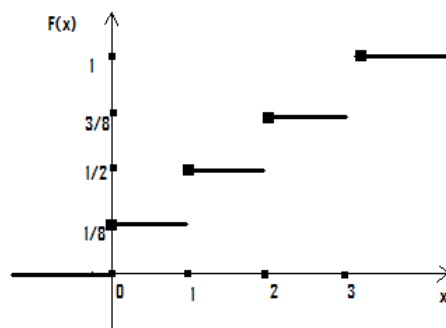
En el caso de ser X una v.a. discreta

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} p(x_i) \quad -\infty < x < \infty$$

Volviendo al ejemplo 1 anterior, la F.d.a. de X es

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{8} & \text{si } 0 \leq x < 1 \\ \frac{1}{8} + \frac{3}{8} & \text{si } 1 \leq x < 2 \\ \frac{1}{8} + \frac{3}{8} + \frac{3}{8} & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x \geq 3 \end{cases} \Rightarrow F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{8} & \text{si } 0 \leq x < 1 \\ \frac{1}{2} & \text{si } 1 \leq x < 2 \\ \frac{7}{8} & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x \geq 3 \end{cases}$$

La gráfica de la F.d.a. de X es



Observación: la F.d.a. de X es una función **escalonada**, los puntos de “salto” coinciden con los puntos del rango de X , y la magnitud del salto en x_i es igual a $P(X = x_i)$

En general si X es una v.a. discreta cualquiera, su F.d.a. será una función escalonada.

Además si x_1, x_2, \dots, x_n son los valores del rango de X ordenados de menor a mayor entonces

$$\begin{aligned} P(X = x_1) &= F(x_1) \\ P(X = x_i) &= F(x_i) - F(x_{i-1}) \quad i = 2, \dots, n \end{aligned}$$

Es decir, se puede obtener la función de distribución de X a partir de su F.d.a.

Para números cualesquiera a y b

1- Si $a \leq b$ entonces $P(a \leq X \leq b) = P(X \leq b) - P(X < a)$

2- Si $a < b$ entonces $P(a < X \leq b) = P(X \leq b) - P(X \leq a)$

3- Si $a < b$ entonces $F(a) \leq F(b)$ (es decir $F(x)$ es una función creciente)

Además se

cumple que

$$1 - \lim_{x \rightarrow \infty} F(x) = \lim_{x \rightarrow \infty} \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i} P(X = x_i) = 1$$

$$2 - \lim_{x \rightarrow -\infty} F(x) = \lim_{x \rightarrow -\infty} \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq -\infty} P(X = x_i) = 0$$

3.3 – Esperanza de una variable aleatoria discreta

Sea X una v.a. discreta con rango R_X . La *esperanza*, *valor medio* o *valor esperado de X* , lo anotamos $E(X)$, y se define como

$$E(X) = \sum_{x_i \in R_X} x_i P(X = x_i)$$

La sumatoria se hace sobre todos los posibles valores de X

Otra notación usual es μ_X o μ

Ejemplos:

1- Sea la v.a. X : “número que queda en la cara de arriba al tirar un dado normal”

$$R_X = \{1, 2, 3, 4, 5, 6\}$$

$$\text{Entonces } E(X) = \sum_{x=1}^6 xP(X = x) =$$

$$= 1P(X = 1) + 2P(X = 2) + 3P(X = 3) + 4P(X = 4) + 5P(X = 5) + 6P(X = 6) =$$

$$= 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{7}{2} = 3.5$$

2- Se tira una moneda normal tres veces, sea la v.a. X : “número de caras obtenidas”

$$\text{Entonces } R_X = \{0, 1, 2, 3\}$$

Calculamos la esperanza de X

$$E(X) = \sum_{x=0}^3 xP(X = x) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{3}{2} = 1.5$$

Observaciones:

1- La esperanza de una v.a. no tiene que coincidir necesariamente con algún valor del rango de la variable

2- En el ejemplo 1 donde el rango es finito y equiprobable, la esperanza de X coincide con el **promedio de los valores del rango de X**

- 3- Se puede interpretar a la esperanza de una v.a. como un **promedio “pesado” o “ponderado” de los valores del rango de la variable**, donde el “peso” de cada x_i es la probabilidad $P(X = x_i)$
- 4- Otra interpretación que se puede hacer de la esperanza es la siguiente: consideremos el ejemplo1, supongamos que tiramos el dado muchas veces, N veces, y entonces obtenemos una secuencia de N valores x_1, x_2, \dots, x_N donde cada x_i es un número natural del 1 al 6. Supongamos además que hacemos un promedio de esos N valores, y si llamamos n_i al número de veces que sale el número i tenemos que

$$\begin{aligned} \frac{x_1 + x_2 + \dots + x_N}{N} &= \frac{n_1 1 + n_2 2 + \dots + n_6 6}{N} = \\ &= \frac{n_1}{N} \times 1 + \frac{n_2}{N} \times 2 + \dots + \frac{n_6}{N} \times 6 \approx 1 \times P(X = 1) + 2 \times P(X = 2) + \dots + 6 \times P(X = 6) = E(X) \end{aligned}$$

Es decir si promediamos los N valores medidos de X , ese promedio tiende a $E(X)$ cuando $N \rightarrow \infty$,

pues $\frac{n_i}{N} \approx P(X = i)$ cuando N es grande.

Esperanza de una función

A veces importa hallar la esperanza de una **función de X** y no de X misma. Veamos un ejemplo. Un instructor de escritura técnica ha solicitado que cierto reporte sea entregado a la semana siguiente, agregando la restricción de que cualquier reporte que sobrepase las cuatro páginas será rechazado. Sea X : “número de páginas del reporte de cierto estudiante seleccionado al azar” Supongamos que X tenga la siguiente distribución de probabilidad

x	1	2	3	4
$p(x)$	0.01	0.19	0.35	0.45

Suponga que el instructor tarda \sqrt{X} minutos calificando un trabajo que consiste en X páginas. Claramente \sqrt{X} es **otra variable aleatoria**. ¿Cuál será su esperanza?, es decir ¿a qué es igual $E(\sqrt{X})$?

Para calcular la esperanza de una v.a. se necesita conocer su función de distribución de probabilidad, por lo tanto habría que hallar previamente la distribución de probabilidad de la v.a. $Y = \sqrt{X}$.

Está claro que si el rango de X es $R_X = \{1, 2, 3, 4\}$ entonces el rango de Y será $R_Y = \{\sqrt{1}, \sqrt{2}, \sqrt{3}, \sqrt{4}\}$.

Además

$$P(Y = \sqrt{1}) = P(X = 1) = 0.01$$

$$P(Y = \sqrt{2}) = P(X = 2) = 0.19$$

$$P(Y = \sqrt{3}) = P(X = 3) = 0.35$$

$$P(Y = \sqrt{4}) = P(X = 4) = 0.45$$

Por lo tanto

$$\begin{aligned} E(Y) &= \sqrt{1} \times P(Y = \sqrt{1}) + \sqrt{2} \times P(Y = \sqrt{2}) + \sqrt{3} \times P(Y = \sqrt{3}) + \sqrt{4} \times P(Y = \sqrt{4}) = \\ &= \sqrt{1} \times P(X = 1) + \sqrt{2} \times P(X = 2) + \sqrt{3} \times P(X = 3) + \sqrt{4} \times P(X = 4) = 1.78491 \end{aligned}$$

O sea

$$E(Y) = \sum_x \sqrt{x} P(X = x)$$

Lo visto en este ejemplo se puede generalizar en el siguiente

Teorema: Si X es una v.a. discreta con rango R_X y distribución de probabilidad $p(x)$, entonces la esperanza de cualquier función $h(X)$ es igual a

$$E(h(X)) = \sum_x h(x) p(x)$$

Ejemplo:

Un negocio de computadoras ha comprado tres computadoras de cierto tipo a \$500 cada una y las venderá a \$1000 cada una. El fabricante ha aceptado volver a comprar en \$200 cualquier computadora que no se haya vendido en un tiempo especificado.

Sea X : “número de computadoras vendidas”, y supongamos que la distribución de probabilidad de X es

x	0	1	2	3
$p(x)$	0.1	0.2	0.3	0.4

Si consideramos la v.a. Y : “utilidad obtenida”, entonces Y es una función de X , es decir $Y = h(X)$

Específicamente

$$Y = 1000X + 200(3 - X) - 1500 = 800X - 900$$

La utilidad esperada, es decir la $E(Y)$ será

$$\begin{aligned} E(Y) &= \sum_{x=0}^3 (800x - 900) P(X = x) = \\ &= (800 \times 0 - 900) P(X = 0) + (800 \times 1 - 900) P(X = 1) + (800 \times 2 - 900) P(X = 2) + (800 \times 3 - 900) P(X = 3) = \\ &= (-900) \times 0.1 + (-100) \times 0.2 + 700 \times 0.3 + 1500 \times 0.4 = \$700 \end{aligned}$$

Notar que aplicando propiedades de la notación Σ se puede plantear

$$E(Y) = \sum_{x=0}^3 (800x - 900) P(X = x) = 800 \sum_{x=0}^3 x P(X = x) - 900 \sum_{x=0}^3 P(X = x) = 800E(X) - 900$$

y calculando la esperanza de X , se llega al mismo resultado

Propiedades de la esperanza

En el ejemplo anterior tenemos que Y es una **función lineal de X** , es decir $Y = aX + b$ con a y b números reales.

En este caso vale entonces la siguiente propiedad

$$E(aX + b) = aE(X) + b$$

La demostración sigue los mismos pasos que en el ejemplo anterior

$$E(aX + b) = \sum_x (ax + b)P(X = x) = a \underbrace{\sum_x xP(X = x)}_{= E(X)} + b \underbrace{\sum_x P(X = x)}_{= 1} = aE(X) + b$$

Ejemplo:

En el ejemplo anterior donde $Y = 800X - 900$

Directamente calculamos

$$E(Y) = 800E(X) - 900$$

Y

$$E(X) = 0 \times 0.1 + 1 \times 0.2 + 2 \times 0.3 + 3 \times 0.4 = 2$$

En consecuencia

$$E(Y) = 800E(X) - 900 = 800 \times 2 - 900 = 700$$

Observaciones:

1- Para cualquier constante a , $E(aX) = aE(X)$

2- Para cualquier constante b , $E(X + b) = E(X) + b$

Varianza de una variable aleatoria

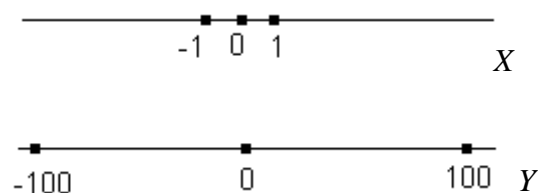
La esperanza de una v.a. mide dónde está centrada la distribución de probabilidad. Pero supongamos el siguiente ejemplo

Sean X e Y dos variables aleatorias con distribuciones dadas por

x	-1	1
$p(x)$	0.5	0.5

y	-100	100
$p(y)$	0.5	0.5

Es fácil verificar que $E(X) = E(Y) = 0$, pero los valores que toma la v.a. Y están más “alejados” de su esperanza que los valores de X .



Se busca una medida que refleje este hecho, se define entonces la varianza de una v.a.

Sea X una v.a. discreta con rango R_X , función de distribución de probabilidad $p(x)$ y esperanza $E(X) = \mu$,

Entonces la **varianza de X** , que anotamos $V(X)$, σ^2 o σ_X^2 es

$$V(X) = E[(X - \mu)^2] = \sum_{x \in R_X} (x - \mu)^2 p(x)$$

La **desviación estándar de X** es $\sigma_X = \sqrt{V(X)}$

Observaciones:

1- La varianza de una v.a. nunca es negativa

2- La cantidad $h(X) = (X - \mu)^2$ es el cuadrado de la desviación de X desde su media, y la varianza de X es la esperanza de la desviación al cuadrado. Si la mayor parte de la distribución de probabilidad está cerca de μ , entonces σ^2 será relativamente pequeña. Si hay valores de la variable alejada de μ que tengan alta probabilidad, entonces σ^2 será grande.

3- σ^2 está expresado en las unidades de medida de X al cuadrado, mientras que σ está expresada en las mismas unidades de medida que X .

Ejemplo:

En el caso de las variables aleatorias X e Y nombradas anteriormente,

$$V(X) = (-1-0)^2 \times 0.5 + (1-0)^2 \times 0.5 = 1 \quad \text{y} \quad \sigma_X = 1$$

$$V(Y) = (-100-0)^2 \times 0.5 + (100-0)^2 \times 0.5 = 100^2 \quad \text{y} \quad \sigma_Y = 100$$

Otra forma de escribir la varianza de una v.a., que facilita los cálculos es

$$\begin{aligned} V(X) &= \sum_{x \in R_X} (x - \mu)^2 p(x) = \sum_{x \in R_X} (x^2 - 2\mu x + \mu^2) p(x) = \sum_{x \in R_X} x^2 p(x) - 2\mu \sum_{x \in R_X} x p(x) + \mu^2 \sum_{x \in R_X} p(x) = \\ &= E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 \end{aligned}$$

Por lo tanto

$$V(X) = E(X^2) - \mu^2$$

Propiedades de la varianza

Las propiedades de la varianza de una v.a. son consecuencia de las propiedades de la esperanza de una v.a.

Si X es una v.a. discreta con rango R_X y distribución de probabilidad $p(x)$, entonces la varianza de cualquier función $h(X)$ es igual a

$$V(h(X)) = \sum_{x \in R_X} (h(x) - E(h(X)))^2 p(x)$$

Si $h(X)$ es una función lineal, entonces

$$V(aX + b) = a^2 V(X) \quad \text{y} \quad \sigma_{aX+b} = \sqrt{V(aX+b)} = |a| \sigma_X$$

Observaciones:

1- $V(aX) = a^2 V(X)$

2- $V(X + b) = V(X)$

Ejemplo:

En un ejemplo anterior donde X : “número de computadoras vendidas” y Y : “utilidad obtenida”, la

$$V(Y) \text{ sería } V(Y) = 800^2 V(X)$$

Necesitamos calcular $V(X)$

$$V(X) = E(X^2) - \mu^2$$

$$\text{Sabemos ya que } \mu = E(X) = 2$$

$$\text{Calculamos } E(X^2) = 0^2 \times 0.1 + 1^2 \times 0.2 + 2^2 \times 0.3 + 3^2 \times 0.4 = 5$$

En consecuencia

$$V(Y) = 800^2 V(X) = 800^2 (E(X^2) - \mu^2) = 800^2 (5 - 2^2) = 800$$

Práctica
Variables aleatorias y distribuciones de probabilidad discretas

- 1) Para cada variable aleatoria (v.a.) definida a continuación, describa el conjunto de valores posibles para la variable (recorrido o rango), e indique si la variable es discreta:
 - a) X: “ número de huevos que no están rotos en una caja de cartón estándar (de una docena) seleccionada al azar”
 - b) Y: “ el número de estudiantes de la lista de un grupo para un curso en particular, ausentes el primer día de clases”.
 - c) U: “ el número de veces que un inexperto tiene que intentar pegarle a una pelota de golf antes de lograrlo”.
 - d) X: “ peso de una persona elegida al azar”
 - e) Y: “ tensión (en libras por pulgada cuadrada) a la que una raqueta de tenis seleccionada al azar haya sido encordada”
- 2) Sea X: “ el número de neumáticos de un automóvil seleccionado al azar, que tengan baja presión”.
 - a) ¿Cuál de las siguientes tres funciones $p(x)$ es una función de distribución de probabilidad (fdp), y por qué no se permiten las otras dos?

x	0	1	2	3	4
p(x)	0.3	0.2	0.1	0.05	0.05
p(x)	0.4	0.1	0.1	0.1	0.3
p(x)	0.4	0.1	0.2	0.1	0.3

- b) Para la fdp legítima de la parte a) calcular: $P(2 \leq X \leq 4)$; $P(X \leq 2)$ y $P(X \neq 0)$.
 - c) Si $p(x) = c(5 - x)$, $x = 0, 1, 2, 3, 4$, ¿cuál es el valor de c ? (Sugerencia: $\sum_{x=0}^4 p(x) = 1$)
- 3) Un embarque de 10 automóviles extranjeros contiene 4 que tienen ligeras manchas de pintura. Si una agencia recibe 6 de estos automóviles al azar, sea
 X: “nº de automóviles que la agencia compra con manchas de pintura”.
 - a) Hallar la f.d.p. de X
 - b) Determine $P(X = 0)$; $P(X = 2)$; $P(X \leq 2)$; $P(X \geq 2)$
- 4) Un negocio de computadoras que atiende pedidos por correo tiene seis líneas telefónicas. Denotamos por X el número de líneas en uso en un momento especificado. Supongamos que la función de probabilidad de X está dada en la tabla siguiente:

x	0	1	2	3	4	5	6
p(x)	0.10	0.15	0.20	0.25	0.20	0.06	0.04

- a) Graficar la función de probabilidad de X.
- b) Hallar la función de distribución acumulada (Fda.) de X y graficarla.
- c) Calcular la probabilidad de los siguientes eventos:
 - i) A: “a lo sumo 3 líneas están en uso”
 - ii) B: “ menos de 3 líneas están en uso”
 - iii) C: “ por lo menos 3 líneas están en uso”
 - iv) D: “ entre 2 y 5 líneas inclusive están en uso”

- 5) El espesor de un entablado de madera (en pulgadas) que algún cliente ordena, es una variable aleatoria X que tiene la siguiente Fda.:

$$F(x) = \begin{cases} 0 & x < \frac{1}{8} \\ 0.2 & \frac{1}{8} \leq x < \frac{1}{4} \\ 0.9 & \frac{1}{4} \leq x < \frac{3}{8} \\ 1 & \frac{3}{8} \leq x \end{cases}$$

a) Hallar la función de distribución de X .
 b) Calcular: i) $P(X \leq 1/8)$
 ii) $P(X \leq 1/4)$
 iii) $P(X \leq 5/16)$
 iv) $P(X > 1/4)$
 v) $P(X \leq 1/2)$

- 6) La distribución de probabilidad de X : “nº de imperfecciones por 10 metros de una tela sintética en rollos continuos de ancho uniforme, está dada por:

x	0	1	2	3	4
$f(x)$	0.41	0.37	0.16	0.05	0.01

- a) Hallar la función de distribución acumulada de X
 b) Determine $F(2)$ y $F(3.1)$
- 7) Para las variables aleatorias de los ejercicios 3), 4), 5) y 6) hallar $E(X)$, $E(X^2)$, $V(X)$.
- 8) Un instructor de un grupo de escritura técnica ha solicitado que cierto reporte sea entregado a la semana siguiente, agregando la restricción de que cualquier reporte que supere cuatro páginas será rechazado. Sea Y = el número de páginas del reporte de cierto estudiante seleccionado al azar y supongamos que Y tiene la siguiente distribución

y	1	2	3	4
$p(y)$	0.01	0.19	0.35	0.45

- a) Calcular $E(Y)$, $V(Y)$ y σ_Y .
 b) Suponga que el instructor tarda \sqrt{Y} minutos calificando un trabajo que consiste en Y páginas. ¿Cuál es la cantidad esperada de tiempo ($E(\sqrt{Y})$) empleada en calificar un trabajo seleccionado al azar?
- 9) Una compañía proveedora de productos químicos tiene actualmente en existencia 100 libras de cierto producto, que vende a clientes en lotes de 5 libras. Sea X = número de lotes ordenados por un cliente seleccionado al azar, y suponga que X tiene la siguiente distribución

x	1	2	3	4
$p(x)$	0.2	0.4	0.3	0.1

- a) Calcular $E(X)$ y $V(X)$.
 b) Calcular el número esperado de libras sobrantes después de embarcar el pedido del siguiente cliente, y la varianza del número de libras restantes.

3.4- Variables aleatorias discretas importantes

Distribución binomial

Sea ε un experimento aleatorio. Sea A un evento asociado a ε y anotamos $P(A) = p$.

Supongamos un experimento aleatorio ε_0 que cumple los siguientes requisitos:

- 1- se realizan n repeticiones **independientes** de ε , donde n se fija de antemano.
- 2- las repeticiones son idénticas, y en cada repetición de ε observamos **si ocurre A o no ocurre A** (cuando A ocurre se dice que se obtuvo un “éxito”, caso contrario se obtuvo un “fracaso”)
- 3- la probabilidad de éxito es constante de una repetición a otra de ε , y es igual a p

Se dice entonces que ε_0 es un **experimento binomial**

Ejemplos:

- 1- Se tira una moneda 4 veces en forma sucesiva e independiente, y observamos en cada tiro si sale cara o no sale cara.

Entonces este es un experimento binomial pues:

ε sería el experimento “tirar una moneda”

A sería el evento “sale cara”

ε se repite en forma sucesiva e independiente $n = 4$ veces

$P(A) = p$ es la misma en cada tiro.

- 2- Se tiene una urna con 15 bolillas blancas y 5 verdes. Se extraen al azar **con reemplazo** tres bolillas y se observa si la bolilla extraída es blanca.

Entonces este es un experimento binomial pues:

ε sería el experimento “extraer al azar una bolilla de la urna”

A sería el evento “se extrae bolilla blanca”

ε se repite en forma sucesiva e independiente $n = 3$ veces

$P(A) = \frac{15}{20} = \frac{3}{4}$ es la misma en cada extracción.

- 3- Si en el ejemplo anterior se extraen las bolillas **sin reemplazo** entonces el experimento no es binomial, pues **falla la independencia**:

Si anotamos A_i : “se extrae bolilla blanca en la i – ésima extracción”, entonces

$$P(A_1) = \frac{15}{20} \quad ; \quad P(A_2) = P(A_2 / A_1)P(A_1) + P(A_2 / A_1^c)P(A_1^c) = \frac{14}{19} \times \frac{15}{20} + \frac{15}{19} \times \frac{5}{20} = \frac{15}{20}$$

Pero $P(A_2) = \frac{15}{20} \neq P(A_2 / A_1) = \frac{14}{19}$ por lo tanto las extracciones no son independientes

Observación: si en la urna hubiese 1500 bolillas blancas y 500 verdes y se extraen dos bolillas al azar sin reemplazo, entonces

$$P(A_2) = \frac{15}{20} = 0.75 \approx P(A_2 / A_1) = \frac{1499}{1999} = 0.74987$$

Por lo tanto en estas condiciones podemos asumir que el experimento es binomial

La variable aleatoria binomial y su distribución

En la mayoría de los experimentos binomiales, interesa el número total de éxitos, más que saber exactamente cuáles repeticiones produjeron los éxitos

Sea la v.a. X : “número de éxitos en las n repeticiones de ε ”

Entonces se dice que X es una v.a. binomial

Veamos cuál es la distribución de probabilidad de X , para esto primero tomamos un caso concreto: el ejemplo 1 anterior en el que se tira una moneda 4 veces. Supongamos que la probabilidad de cara es $\frac{3}{4}$

Aquí el rango de X sería $R_X = \{0,1,2,3,4\}$

Para facilitar la notación escribimos A_i : “sale cara en el i -ésimo tiro” $i = 1,2,3,4$

Por lo tanto

$$P(X=0) = P(A_1^c \cap A_2^c \cap A_3^c \cap A_4^c) \underset{\substack{\downarrow \\ \text{por independencia}}}{=} P(A_1^c)P(A_2^c)P(A_3^c)P(A_4^c) = \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \left(\frac{1}{4}\right)^4$$

Para calcular la $P(X=1)$ pensamos que hay cuatro casos posibles en los que se puede obtener exactamente una cara, que la cara salga en el 1º tiro, o en el 2º o en el 3º o en el 4º tiro. Notar que tenemos cuatro casos y eso **es igual a la cantidad de formas en que podemos elegir entre los 4 tiros uno de ellos en el cual sale cara, es decir tenemos** $\binom{4}{1} = \frac{4!}{1!3!} = 4$ **casos diferentes.**

$$\begin{aligned} P(X=1) &= \\ &= P(A_1 \cap A_2^c \cap A_3^c \cap A_4^c) + P(A_1^c \cap A_2 \cap A_3^c \cap A_4^c) + P(A_1^c \cap A_2^c \cap A_3 \cap A_4^c) + \\ &+ P(A_1^c \cap A_2^c \cap A_3^c \cap A_4) \end{aligned}$$

Cada término es igual a $p(1-p)^3$ por lo tanto $P(X=1) = 4p(1-p)^3$

Análogamente, para calcular $P(X=2)$ tenemos $\binom{4}{2} = \frac{4!}{2!2!} = 6$ casos en los que salen exactamente dos caras, por lo tanto

$$P(X=2) = P(A_1 \cap A_2 \cap A_3^c \cap A_4^c) + P(A_1 \cap A_2^c \cap A_3 \cap A_4^c) + \dots = \binom{4}{2} p^2 (1-p)^2$$

Pensando de la misma forma los otros casos se llega a

$$P(X=3) = \binom{4}{3} p^3 (1-p) \quad ; \quad P(X=4) = p^4$$

En general con un argumento análogo tenemos que $R_X = \{0,1,2,\dots,n\}$ y

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, 2, \dots, n$$

Notación: indicamos que X es una v.a. binomial **con parámetros n y p** con el símbolo $X \sim B(n, p)$

Dado que los números $P(X = k)$ corresponden a la distribución de una v.a., automáticamente cumplen que $\sum_{k=0}^n P(X = k) = 1$

De todas formas se podría hacer una verificación algebraica utilizando la fórmula del binomio de Newton

$$\sum_{k=0}^n P(X = k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1$$

Ejemplos:

- 1- En el ejemplo anterior en el que se tira una moneda 4 veces, calcular la probabilidad de obtener:
 - a) exactamente una cara
 - b) al menos una cara
 - c) a lo sumo una cara

Solución:

- a) tenemos que la v.a. X : “número de caras obtenido” es $B(4, 0.25)$

$$\text{se pide } P(X = 1) = \binom{4}{1} \left(\frac{1}{4}\right)^1 \left(1 - \frac{1}{4}\right)^3 = 4 \times \frac{1}{4} \times \left(\frac{3}{4}\right)^3 = 0.421875$$

- b) la probabilidad de obtener **al menos una cara** es

$$P(X \geq 1) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = \sum_{k=1}^4 \binom{4}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{4-k}$$

Pero más fácil es hacer

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{4}{0} \left(\frac{1}{4}\right)^0 \left(1 - \frac{1}{4}\right)^{4-0} = 1 - \left(\frac{3}{4}\right)^4 = 1 - 0.421875 = 0.578125$$

- c) la probabilidad de obtener **a lo sumo una cara** es

$$P(X \leq 1) = P(X = 0) + P(X = 1) = \binom{4}{0} \left(\frac{1}{4}\right)^0 \left(1 - \frac{1}{4}\right)^{4-0} + \binom{4}{1} \left(\frac{1}{4}\right)^1 \left(1 - \frac{1}{4}\right)^{4-1} = 0.84375$$

Observación: si $X \sim B(n, p)$ para calcular $P(X \leq k)$ en general se debe hacer

$$P(X \leq k) = \sum_{i=0}^k P(X = i) = P(X = 0) + P(X = 1) + \dots + P(X = k)$$

Notar que $P(X \leq k)$ es la F.d.a. de X evaluada en k , es decir $F(k) = P(X \leq k)$

Existen tablas de la función de distribución acumulada de la binomial para diferentes valores de n y p

Consultando estas tablas se puede obtener directamente el resultado del inciso c) buscando para $n = 4$ y $p = 0.25$

Además consultando las tablas podemos evaluar $P(X = k)$ haciendo

$$P(X = k) = F(k) - F(k - 1) \quad k = 1, 2, \dots, n$$

2- Supongamos que el 20% de todos los ejemplares de un texto en particular fallan en una prueba de resistencia a la encuadernación. Se seleccionan 15 ejemplares al azar.

Sea la v.a. X : “número de ejemplares que fallan en la prueba entre los 15 seleccionados”

a) ¿cuál es la probabilidad de que a lo sumo 8 fallen en la prueba?

b) ¿cuál es la probabilidad de que exactamente 8 fallen en la prueba?

c) ¿cuál es la probabilidad de que al menos 8 fallen en la prueba?

Solución:

a) Tenemos que $X \sim B(15, 0.2)$

$$P(X \leq 8) = \sum_{k=0}^8 P(X = k) = F(8) = 0.999$$

por tabla de la F.d.a.

$$b) P(X = 8) = F(8) - F(7) = 0.999 - 0.996 = 0.003$$

por tabla de la F.d.a.

$$c) P(X \geq 8) = 1 - P(X \leq 7) = 1 - F(7) = 1 - 0.996 = 0.004$$

por tabla de la F.d.a.

Observaciones:

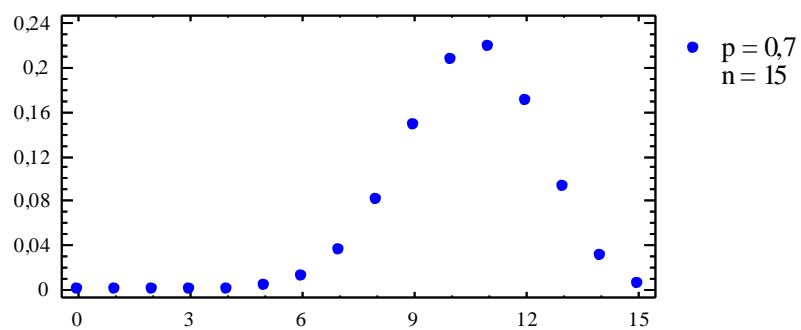
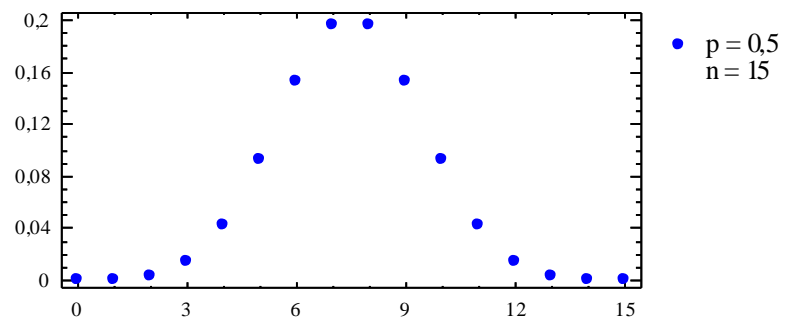
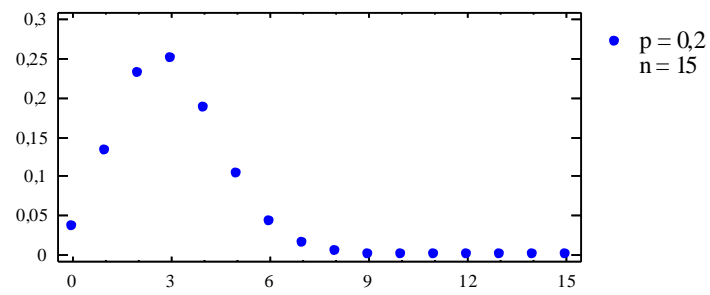
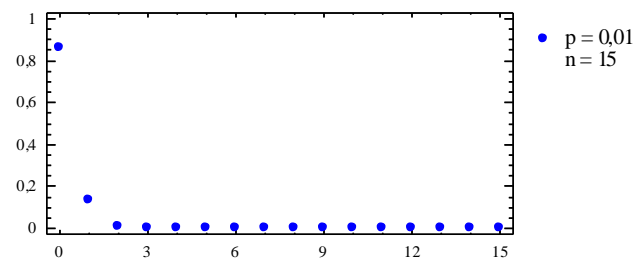
1- Si $X \sim B(1, p)$ entonces la v.a. X toma sólo dos valores 0 y 1 con probabilidades p y $1-p$ es decir podemos escribir

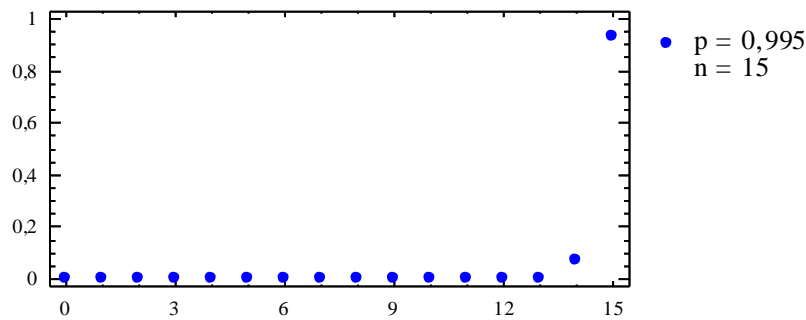
$$X = \begin{cases} 1 & \text{si al ejecutar } \varepsilon \text{ ocurre éxito} \\ 0 & \text{caso contrario} \end{cases} \quad \begin{aligned} P(X = 1) &= P(A) = p \\ P(X = 0) &= P(A^c) = 1 - p \end{aligned}$$

En este caso se dice que X tiene distribución de Bernoulli

En el caso de ser $X \sim B(n, p)$ se dice que se tienen “*n ensayos de Bernoulli*”

2- A continuación se muestra cómo varía la forma de la distribución a medida que p aumenta manteniendo n fijo en 15. Se grafica la distribución de frecuencia para $p = 0.01; 0.2, 0.5, 0.7$ y 0.995 . Observar que para $p = 0.5$ la distribución de frecuencia es simétrica.





Esperanza y varianza

Sea $X \sim B(n, p)$, entonces $E(X) = np$ y $V(X) = np(1-p)$

Dem.)

Aplicamos la definición:

$$E(X) = \sum_{k=0}^n k p \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

El primer término es cero, por lo tanto podemos comenzar la suma en $k=1$:

$$E(X) = \sum_{k=1}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^n \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}$$

Ponemos todo en términos de $k-1$:

$$E(X) = \sum_{(k-1)=0}^n \frac{n(n-1)!}{(k-1)![(n-1)-(k-1)]!} p p^{k-1} (1-p)^{[(n-1)-(k-1)]}$$

Sacamos fuera de la suma n y p que no dependen del índice k y hacemos el cambio de índice:
 $(k-1) \rightarrow s$:

$$E(X) = np \sum_{s=0}^{n-1} \frac{(n-1)!}{s![(n-1)-s]!} p^s (1-p)^{[(n-1)-s]} = np \sum_{s=0}^{n-1} \binom{n-1}{s} p^s (1-p)^{[(n-1)-s]}$$

Recordando el desarrollo del binomio de Newton

$$(a+b)^r = \sum_{s=0}^{r} \binom{r}{s} a^s b^{r-s}, \text{ tenemos (con } r=n-1\text{):}$$

$$E(X) = np [p + (1-p)]^{n-1} = np [1]^{n-1} = np$$

Veamos el cálculo de la varianza

$$V(X) = E(X^2) - [E(X)]^2.$$

Luego:

$$V(X) = E(X^2) - [np]^2. \text{ Nos queda calcular } E(X^2).$$

$$E(X^2) = \sum_{k=0}^n k^2 \cdot p(k) = \sum_{k=0}^n k^2 \cdot \binom{n}{k} \cdot p^k (1-p)^{n-k} = \sum_{k=0}^n k^2 \cdot \frac{n!}{k!(n-k)!} \cdot p^k (1-p)^{n-k}.$$

Para calcular esta suma tratamos de llevarla a la forma del desarrollo de un binomio de Newton. Como el primer término es cero comenzamos a sumar desde $k=1$. Además simplificamos k en el numerador con el factor k en $k!$ del denominador:

$$E(X^2) = \sum_{k=1}^n k \cdot \frac{n!}{(k-1)!(n-k)!} \cdot p^k (1-p)^{n-k} = \sum_{k=1}^n k \cdot \frac{n(n-1)!}{(k-1)!(n-k)!} \cdot p^k (1-p)^{n-k}.$$

Separamos en dos sumas:

$$\begin{aligned} E(X^2) &= \sum_{k=1}^n [1 + (k-1)] \cdot \frac{n(n-1)!}{(k-1)!(n-k)!} \cdot p^k (1-p)^{n-k} \\ &= np \sum_{(k-1)=0}^{n-1} \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} \cdot p^{(k-1)} (1-p)^{[(n-1)-(k-1)]} + \\ &\quad + n(n-1)p^2 \sum_{(k-2)=0}^{n-2} \frac{(n-2)!}{(k-1)![(n-2)-(k-2)]!} \cdot p^{(k-2)} (1-p)^{[(n-2)-(k-2)]} \end{aligned}$$

Esto es:

$$E(X^2) = np \sum_{s=0}^{n-1} k \cdot \binom{n-1}{s} \cdot p^s (1-p)^{[(n-1)-s]} + n(n-1)p^2 \sum_{r=0}^{n-2} k \cdot \binom{n-2}{r} \cdot p^r (1-p)^{[(n-2)-r]}$$

Las sumas corresponden al desarrollo de un binomio de Newton:

$$E(X^2) = np[p + (1-p)]^{n-1} + n(n-1)p^2[p + (1-p)]^{n-2} = np[1]^{n-1} + n(n-1)p^2[1]^{n-2}, \text{ es decir}$$

$$E(X^2) = np + n(n-1)p^2. \text{ Entonces:}$$

$$V(X) = E(X^2) - [np]^2 = np + n(n-1)p^2 - [np]^2 = np - np^2 \text{ o sea: } V(X) = np(1-p)$$

Distribución geométrica

Sea ε un experimento aleatorio. Sea A un evento asociado a ε y anotamos $P(A) = p$.

Supongamos un experimento aleatorio ε_0 que cumple los siguientes requisitos:

- 1- se realizan repeticiones independientes de ε , hasta que ocurre A por primera vez inclusive.
- 2- las repeticiones son idénticas, y en cada repetición de ε observamos *si ocurre A* o *no ocurre A* (cuando A ocurre se dice que se obtuvo un “éxito”, caso contrario se obtuvo un “fracaso”)
- 3- la probabilidad de éxito es constante de una repetición a otra de ε , y es igual a p

Sea la v.a. X : “número de repeticiones de ε hasta que ocurre A por primera vez inclusive”
Veamos cuál es la distribución de probabilidad de X

El rango es el conjunto $R_X = \{1, 2, 3, \dots\} = \mathbb{N}$

Además si anotamos A_i : "ocurre A en la i -ésima repetición de ε " , entonces

$$P(X = k) = P(A_1^C \cap A_2^C \cap \dots \cap A_{k-1}^C \cap A_k) = P(A_1^C)P(A_2^C) \dots P(A_{k-1}^C)P(A_k) = (1-p)^{k-1} p$$

↓
por independencia

En consecuencia

$$P(X = k) = (1-p)^{k-1} p \quad k = 1, 2, \dots$$

Notación: $X \sim G(p)$

Para verificar que $\sum_{k=1}^{\infty} P(X = k) = 1$ recordar que $\sum_{k=0}^{\infty} a^k = \frac{1}{1-a}$ si $|a| < 1$

$$\sum_{k=1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{k=1}^{\infty} (1-p)^{k-1} = p \frac{1}{1-(1-p)} = 1$$

Observaciones:

1- $P(X = k) = (1-p)^{k-1} p$ y $P(X = k+1) = (1-p)^k p$

Es decir $P(X = k+1) = (1-p)^{k-1} p(1-p) = P(X = k)(1-p) \quad \forall k = 1, 2, \dots$

Podemos interpretar que para cada k , los $a_k = P(X = k)$ son los términos de una sucesión geométrica con razón $1-p$ y primer término $a_1 = p$

2- La F.d.a. sería

$$F(x) = P(X \leq x) = \sum_{k=1}^{[x]} P(X = k)$$

donde $[x]$ indica parte entera de x

Si x es un entero positivo entonces recordando que la suma de los n primeros términos de una sucesión geométrica con razón r y término general $a_k = pr^{k-1}$ es $\sum_{k=1}^n a_k = \frac{a_1(1-r^n)}{1-r}$ tenemos que

$$F(x) = P(X \leq x) = \sum_{k=1}^{[x]} P(X = k) = \frac{p(1-r^{[x]})}{1-r} = \frac{p(1-(1-p)^{[x]})}{1-(1-p)} = 1 - (1-p)^{[x]}$$

↓
 $r = 1-p$

Por lo tanto

$$F(x) = \begin{cases} 1 - (1-p)^{[x]} & \text{si } x \geq 1 \\ 0 & \text{caso contrario} \end{cases}$$

- 3- Una variante en la definición de distribución geométrica es definir la v.a. Y : “**número de fracasos** hasta el primer éxito”, en este caso $R_Y = \{0, 1, 2, \dots\}$, es decir se incluye al cero.

La distribución de probabilidad o frecuencia sería en este caso

$$P(Y = k) = (1 - p)^k p \quad k = 0, 1, 2, \dots$$

Notar que la relación entre X e Y sería: $X = Y + 1$.

Es decir si adoptamos esta última definición **no incluimos** la repetición del experimento en el cual ocurre el primer éxito.

Ejemplos:

- 1- La probabilidad de que una computadora que corre cierto sistema operativo se descomponga en determinado día es de 0.1. Determinar la probabilidad de que la máquina se descomponga por primera vez en el duodécimo día, después de la instalación del sistema operativo

Solución:

Definimos la v.a. X : “número de días hasta que la computadora se descompone por primera vez”

Entonces $X \sim G(0.1)$

Se pide calcular la $P(X = 12)$

$$P(X = 12) = (1 - p)^{12-1} p = 0.9^{11} \times 0.1 = 0.031381$$

- 2- Una prueba de resistencia a la soldadura consiste en poner carga en uniones soldadas hasta que se dé una ruptura. Para cierto tipo de soldadura, 80% de las rupturas ocurre en la propia soldadura, mientras que otro 20% se da en las vigas. Se prueba cierto número de soldaduras.

Sea la v.a. X : “número de pruebas hasta que se produce la ruptura de la viga”

¿Qué distribución tiene X ? ¿Cuál es la probabilidad que en la tercera prueba se produzca la primera ruptura de la viga?

Solución:

Cada prueba es un “**ensayo de Bernoulli**”, con un éxito definido como la ruptura de una viga. Por lo tanto, la probabilidad de éxito es $p = 0.2$.

La v.a. X tiene una distribución geométrica con parámetro $p = 0.2$ es decir $X \sim G(0.2)$

Para calcular la probabilidad pedida hacemos $P(X = 3) = (1 - p)^{3-1} p = 0.8^2 \times 0.2 = 0.128$

Esperanza y varianza

Sea $X \sim G(p)$ entonces $E(X) = \frac{1}{p}$ y $V(X) = \frac{1-p}{p^2}$

Dem.) Llamamos $1 - p = q$

Planteamos

$$\sum_{k=1}^{\infty} k P(X = k) = \sum_{k=1}^{\infty} k p (1 - p)^{k-1} = p \sum_{k=1}^{\infty} k (1 - p)^{k-1} = p \sum_{k=1}^{\infty} k q^{k-1} =$$

Notar que $\frac{d}{dq}(q^k) = kq^{k-1}$, por lo tanto como $\sum_{k=0}^{\infty} a^k = \frac{1}{1-a}$ si $|a| < 1$

$$= p \sum_{k=1}^{\infty} \frac{d}{dq}(q^k) = p \frac{d}{dq} \left(\sum_{k=1}^{\infty} q^k \right) = p \frac{d}{dq} \left(\frac{q}{1-q} \right) = p \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p} \quad \therefore E(X) = \frac{1}{p}$$

Calculamos ahora la varianza

$$V(X) = E(X^2) - \mu^2 \quad \text{donde} \quad \mu = E(X) = \frac{1}{p}$$

$$E(X^2) = \sum_{k=1}^{\infty} k^2 P(X=k) = \sum_{k=1}^{\infty} k^2 p(1-p)^{k-1} = \sum_{k=1}^{\infty} k(k-1+1)p(1-p)^{k-1} = p \sum_{k=1}^{\infty} k(k-1)q^{k-1} + \sum_{k=1}^{\infty} kpq^{k-1}$$

Pero

$$p \sum_{k=1}^{\infty} k(k-1)q^{k-1} = pq \sum_{k=1}^{\infty} k(k-1)q^{k-2} = pq \sum_{k=1}^{\infty} \frac{d^2}{dq^2}(q^k) = pq \frac{d^2}{dq^2} \left(\sum_{k=1}^{\infty} q^k \right) = pq \frac{d^2}{dq^2} \left(\frac{q}{1-q} \right) =$$

$$= pq \frac{2}{(1-q)^3} = q \frac{2}{p^2}$$

$$\text{Y} \quad \sum_{k=1}^{\infty} kpq^{k-1} = \frac{1}{p}$$

Por lo tanto

$$V(X) = E(X^2) - \frac{1}{p^2} = q \frac{2}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q + p - 1}{p^2} = \frac{2(1-p) + p - 1}{p^2} = \frac{1-p}{p^2}$$

Distribución binomial negativa

La distribución binomial negativa constituye una extensión de la distribución geométrica.

Sea r un entero positivo.

Sea ε un experimento aleatorio. Sea A un evento asociado a ε y anotamos $P(A) = p$.

Supongamos un experimento aleatorio ε_0 que cumple los siguientes requisitos:

- 1- se realizan repeticiones independientes de ε , hasta que ocurre A por r -ésima vez inclusive.
- 2- las repeticiones son idénticas, y en cada repetición de ε observamos **si ocurre A o no ocurre A** (cuando A ocurre se dice que se obtuvo un “éxito”, caso contrario se obtuvo un “fracaso”)
- 3- la probabilidad de éxito es constante de una repetición a otra de ε , y es igual a p

Sea la v.a. X : “número de repeticiones de ε hasta que ocurre A por r -ésima vez, incluyendo la r -ésima vez que ocurre A ”

Veamos cuál es la distribución de probabilidad de X

El rango es el conjunto $R_X = \{r, r+1, r+2, r+3, \dots\}$

Para obtener una expresión genérica de la $P(X = k)$, notar que si en el k -ésimo ensayo ocurre éxito por r -ésima vez, entonces en los $k-1$ primeros ensayos ocurrieron $r-1$ éxitos. Si anotamos B : “en los primeros $k-1$ ensayos ocurran $r-1$ éxitos”, entonces

$$P(B) = \binom{k-1}{r-1} p^{r-1} (1-p)^{(k-1)-(r-1)} = \binom{k-1}{r-1} p^{r-1} (1-p)^{k-r}$$

Además si anotamos A : “ocurre A en la r -ésima repetición de ε ”, entonces $P(A) = p$ y A y B son independientes, por lo tanto

$$P(X = k) = P(B \cap A) = P(B)P(A) = \binom{k-1}{r-1} p^{r-1} (1-p)^{k-r} \times p = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

En resumen

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad k = r, r+1, r+2, \dots$$

Notación: $X \sim BN(r, p)$

Para verificar que $\sum_{k=0}^{\infty} P(X = k) = 1$ se deriva r veces la igualdad $\sum_{k=0}^{\infty} (1-p)^{k-1} = \frac{1}{p}$ la que se deduce

$$\text{de } p \sum_{k=0}^{\infty} (1-p)^{k-1} = 1$$

Ejemplo:

En una prueba de fuerza de soldadura, 80% de las pruebas da como resultado ruptura de soldadura, mientras que otro 20% da ruptura de la viga. Sea la v.a. X : “número de pruebas hasta la tercera ruptura de la viga inclusive”. ¿Cuál es la distribución de X ? Determinar la $P(X = 8)$

Solución:

Tenemos que $X \sim BN(3, 0.2)$

$$\text{Por lo tanto } P(X = 8) = \binom{8-1}{3-1} p^3 (1-p)^{8-3} = \binom{7}{2} 0.2^3 \times 0.8^5 = 0.05505$$

Observación: la distribución geométrica puede verse como un caso particular de la distribución binomial negativa con $r = 1$

Esperanza y varianza

$$\text{Si } X \sim BN(r, p) \text{ entonces } E(X) = \frac{r}{p} \quad \text{y} \quad V(X) = \frac{r(1-p)}{p^2}$$

Dem.) Se hará mas adelante

Distribución hipergeométrica (OPCIONAL)

Supongamos que tenemos una población o conjunto de N objetos o individuos (es decir tenemos una población finita).

Clasificamos a los objetos de la población en dos categorías. Hay M objetos de una categoría y $N-M$ de la otra categoría. Se suele decir que tenemos M “éxitos” y $N-M$ “fracasos”.

Se extraen al azar y sin reemplazo n objetos de dicha población. Es decir se extrae una muestra de n objetos de la población, de manera tal que es igualmente probable que se seleccione cada subconjunto de tamaño n .

Consideramos la v.a. X : “número de éxitos en la muestra extraída”

Se dice que X tiene una distribución hipergeométrica con parámetros n , M y N

Notación: $X \sim H(n, M, N)$

Veamos cuál es la distribución de X

Primero notar que una expresión para la $P(X = k)$, usando combinatoria es

$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} \quad \text{donde para que los números combinatorios estén bien definidos debe}$$

cumplirse $0 \leq k \leq M$ y $0 \leq n-k \leq N-M$. Pero estas condiciones son equivalentes a

$$\max(0, n - N + M) \leq k \leq \min(n, M)$$

Por lo tanto la distribución de probabilidad de X es

$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} \quad \max(0, n - N + M) \leq k \leq \min(n, M)$$

Se verifica que $\sum_{k=0}^n P(X = k) = 1$ pues los números $P(X = k)$ corresponden a la distribución de una v.a.

Ejemplo:

1- De 50 edificios en un parque industrial, 12 no cumplen el código eléctrico. Si se seleccionan aleatoriamente 10 edificios para inspeccionarlos, ¿cuál es la probabilidad de que exactamente tres de los diez no cumplan el código?

Solución:

Sea la v.a. X : “número de edificios seleccionados que violan el código”, entonces $X \sim H(10, 12, 50)$. Se pide calcular la $P(X = 3)$

$$P(X = 3) = \frac{\binom{12}{3} \binom{50-12}{10-3}}{\binom{50}{10}} = 0.2703$$

- 2- Un cargamento contiene 40 elementos. Se seleccionará de forma aleatoria y se probará 5 elementos. Si dos o más están defectuosos, se regresará el cargamento.
- si de hecho el cargamento contiene cinco elementos defectuosos, ¿cuál es la probabilidad de que sean aceptados?
 - si de hecho el cargamento contiene diez elementos defectuosos, ¿cuál es la probabilidad de que no sean aceptados?

Solución:

- a) Sea la v.a. X : “número de elementos defectuosos en la muestra”

En este caso $X \sim H(5, 5, 40)$. Hay que calcular la $P(X \leq 1)$

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$P(X = 0) = \frac{\binom{5}{0} \binom{40-5}{5-0}}{\binom{40}{5}} = 0.4933557 \quad P(X = 1) = \frac{\binom{5}{1} \binom{40-5}{5-1}}{\binom{40}{5}} = 0.3978675$$

$$\therefore P(X \leq 1) = 0.4933557 + 0.3978675 = 0.8912232$$

- b) Sea la v.a. X : “número de elementos defectuosos en la muestra”

En este caso $X \sim H(5, 10, 40)$. Hay que calcular la $P(X \geq 2)$

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1)$$

$$P(X = 0) = \frac{\binom{10}{0} \binom{40-10}{5-0}}{\binom{40}{5}} = 0.2165718 \quad P(X = 1) = \frac{\binom{10}{1} \binom{40-10}{5-1}}{\binom{40}{5}} = 0.416484$$

$$\therefore P(X \geq 2) = 0.3669442$$

Observación:

En el ejemplo anterior si el cargamento hubiese tenido 400 elementos se podría haber considerado en la parte a) a X con distribución binomial con parámetros $n = 5$ y $p = \frac{5}{400}$

En general, si el tamaño de la población N y el número de éxitos M crecen pero de manera tal que $\frac{M}{N} \rightarrow p$ y n es chico comparado con N , se puede verificar que

$$\lim_{N \rightarrow \infty} \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{donde } \frac{M}{N} = p$$

Por lo tanto, para una fracción fija de defectuosos $\frac{M}{N} = p$ la función de probabilidad hipergeométrica converge a la función de probabilidad binomial cuando N se hace grande.

Esperanza y varianza

$$\text{Si } X \sim H(n, M, N) \text{ entonces } E(X) = \frac{nM}{N} \text{ y } V(X) = n \frac{M}{N} \left(\frac{N-M}{N} \right) \left(\frac{N-n}{N-1} \right)$$

Dem.) La demostración se hará más adelante.

Distribución de Poisson

Una v.a. X con rango $R_X = \{0, 1, 2, \dots\}$ se dice tener **distribución de Poisson con parámetro λ** , si para algún $\lambda > 0$

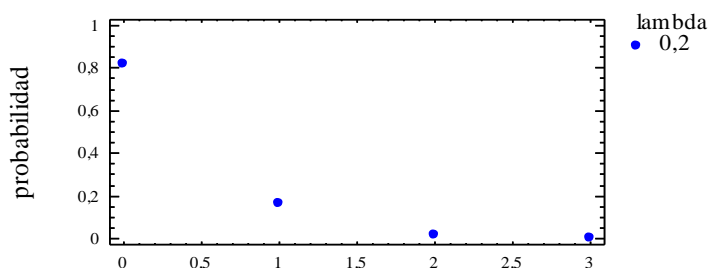
$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad k = 0, 1, 2, \dots$$

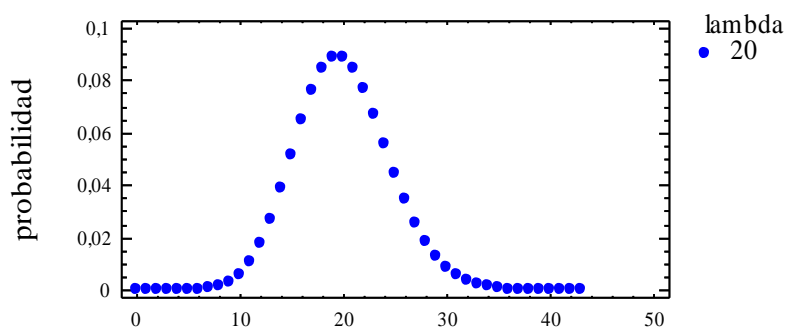
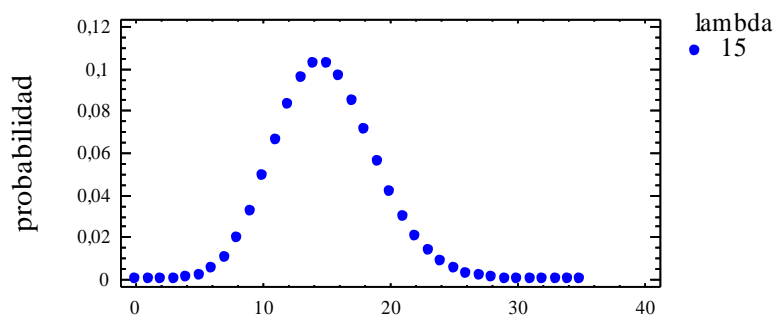
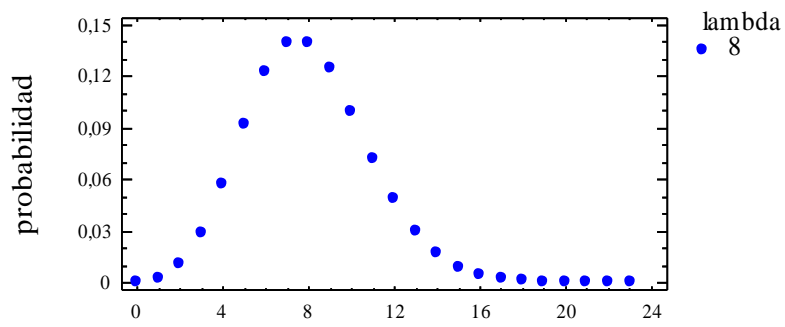
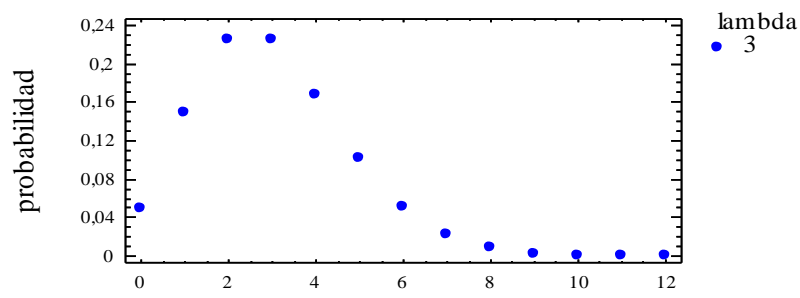
Es fácil verificar que $\sum_{k=0}^{\infty} P(X = k) = 1$ usando el hecho que $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$

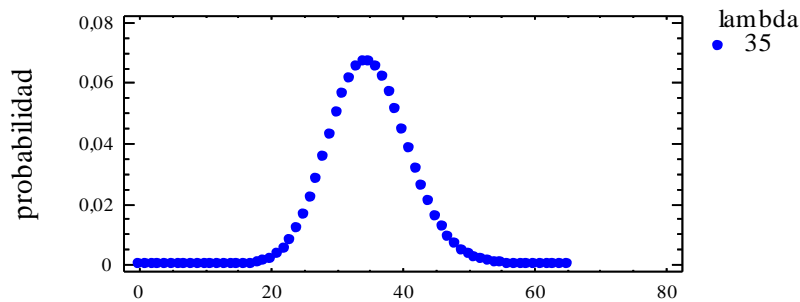
$$\sum_{k=0}^{\infty} P(X = k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

En los siguientes gráficos se ve como varía la forma de la distribución con los valores de λ

Notar que para valores de λ “pequeños” la distribución es asimétrica, a medida que λ aumenta, la distribución tiende a ser cada vez más simétrica





**Ejemplo:**

Considere escribir en un disco de computadora y luego enviar el escrito por un certificador que cuenta el número de pulsos faltantes. Suponga que este número X tiene una distribución de Poisson con parámetro igual a 0.2

- ¿Cuál es la probabilidad de que un disco tenga exactamente un pulso faltante?
- ¿Cuál es la probabilidad de que un disco tenga al menos dos pulsos faltantes?
- Si dos discos se seleccionan independientemente, ¿cuál es la probabilidad de que ninguno contenga algún pulso faltante?

Solución:

- a) Sea la v.a. X : “número de pulsos faltantes en un disco”

Entonces $X \sim P(0.2)$

$$\text{Se pide } P(X = 1) = e^{-0.2} \frac{0.2^1}{1!} = 0.163746$$

- b) Siendo X como en a) se pide calcular $P(X \geq 2)$

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1) = 1 - e^{-0.2} \frac{0.2^0}{0!} - e^{-0.2} \frac{0.2^1}{1!} = 0.01752$$

- c) Sea la v.a. Y : “número de discos sin pulsos faltantes”

Entonces $Y \sim B(2, p)$ donde $p = P(X = 0) = e^{-0.2}$

Por lo tanto se pide calcular $P(Y = 2)$

$$P(Y = 2) = \binom{2}{2} p^2 (1 - p)^0 = p^2 = (e^{-0.2})^2 = 0.67032$$

Esperanza y varianza

Si $X \sim P(\lambda)$ entonces $E(X) = \lambda$ y $V(X) = \lambda$

Dem.)

$$E(X) = \sum_{k=0}^{\infty} k P(X = k) = \sum_{k=0}^{\infty} k \frac{e^{-\lambda} \lambda^k}{k!} = \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} = \lambda \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^{k-1}}{(k-1)!} = \lambda \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

$$V(X) = E(X^2) - \mu^2 \quad \text{donde} \quad \mu = E(X) = \lambda$$

$$\begin{aligned}
E(X^2) &= \sum_{k=0}^{\infty} k^2 P(X=k) = \sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda} \lambda^k}{k!} = \sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda} \lambda^k}{k!} = \sum_{k=1}^{\infty} k \frac{e^{-\lambda} \lambda^k}{(k-1)!} = \sum_{k=1}^{\infty} (k-1+1) \frac{e^{-\lambda} \lambda^k}{(k-1)!} = \\
&= \sum_{k=1}^{\infty} (k-1) \frac{e^{-\lambda} \lambda^k}{(k-1)!} + \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} = \sum_{k=2}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k-2)!} + \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} = \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^{k+2}}{k!} + \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^{k+1}}{k!} = \\
&= \lambda^2 \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} + \lambda \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = \lambda^2 + \lambda
\end{aligned}$$

Por lo tanto

$$V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Aplicaciones de la distribución de Poisson (OPCIONAL)

La v.a. Poisson tiene un gran rango de aplicaciones, una de ellas es la aproximación para una v.a. binomial con parámetros n y p cuando n es grande y p es pequeño de manera tal que $np \rightarrow \lambda$, específicamente, sea $X \sim B(n, p)$ y sea $\lambda = np$, entonces

$$P(X=k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n(n-1)\dots(n-k+1)}{n^k} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n$$

Para n grande y p chico

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda} ; \quad \frac{n(n-1)\dots(n-k+1)}{n^k} \approx 1 ; \quad \left(1 - \frac{\lambda}{n}\right)^k \approx 1$$

Entonces, para n grande y p chico

$$P(X=k) \approx e^{-\lambda} \frac{\lambda^k}{k!}$$

Es decir cuando n es grande, p chico y np es “moderado” entonces la v.a. binomial con parámetros n y p tiene una distribución que se **aproxima** a la de una Poisson con parámetro $\lambda = np$

Ejemplo:

Supongamos que la probabilidad de que un artículo producido por cierta máquina sea defectuoso es 0.1. Hallar la probabilidad que una muestra de 10 artículos contenga a lo sumo un defectuoso.


Sea X : “número de artículos defectuosos en la muestra”

Podemos asumir que $X \sim B(10, 0.1)$

La probabilidad **exacta** pedida es

$$P(X \leq 1) = P(X=0) + P(X=1) = \binom{10}{0} (0.1)^0 (0.9)^{10-0} + \binom{10}{1} (0.1)^1 (0.9)^{10-1} = 0.7361$$

La **aproximación** de Poisson da

$$P(X \leq 1) = P(X=0) + P(X=1) \approx e^{-1} \frac{\lambda^0}{0!} + e^{-1} \frac{\lambda^1}{1!} = e^{-1} + e^{-1} \approx 0.7358$$


$$\lambda = np = 10 \times 0.1 = 1$$

Algunos autores sostienen que la aproximación de Poisson funciona bien cuando n es grande, p es chico y $np < 7$ (Mendenhall, Estadística matemática con aplicaciones), otros recomiendan usar la aproximación de Poisson a la binomial cuando $n \geq 100$ y $p \leq 0.01$ y $np \leq 20$ (Devore, Probabilidad para ingeniería y ciencias)

En la siguiente tabla se da un ejemplo de una aproximación de Poisson a la función de distribución de la binomial. Se tabula la $P(X = k)$ para algunos valores de k para las distribuciones binomial y Poisson con los parámetros que se indican

K	X=B(10 ⁵ , 2×10 ⁻⁵)	X=B(10 ³ , 4×10 ⁻⁵)	P(2)
0	0.135281	0.135281	0.135335
1	0.270671	0.270671	0.270671
2	0.270725	0.270725	0.270671
3	0.180483	0.180483	0.180447
4	0.0902235	0.0902235	0.0902235
5	0.036075	0.036075	0.0360894
6	0.0120178	0.0120178	0.0120298
7	0.0034309	0.0034309	0.00343709
8	0.000856867	0.000856867	0.000859272
9	0.000190186	0.000190186	0.000190949
10	0.000037984	0.000037984	0.0000381899
11	6.89513×10 ⁻⁶	6.89513×10 ⁻⁶	6.94361×10 ⁻⁶
12	1.14712×10 ⁻⁶	1.14712×10 ⁻⁶	1.15727×10 ⁻⁶
13	1.76127×10 ⁻⁷	1.76127×10 ⁻⁷	1.78041×10 ⁻⁷
14	2.51056×10 ⁻⁸	2.51056×10 ⁻⁸	2.54345×10 ⁻⁸
15	3.33937×10 ⁻⁹	3.33937×10 ⁻⁹	3.39126×10 ⁻⁹

Ejemplo:

En una prueba de tarjetas de circuitos, la probabilidad de que un diodo en particular falle es 0.01. Suponga que una tarjeta contiene 200 diodos.

- ¿Cuál es la probabilidad aproximada de que por lo menos 4 diodos fallen en una tarjeta seleccionada al azar?
- Si se embarcan cinco tarjetas a un cliente en particular, ¿cuál es la probabilidad de que por lo menos cuatro de ellas funcionen bien? (Una tarjeta funciona bien solo si todos sus diodos funcionan bien)

Solución:

- Sea la v.a. X: “número de diodos en una tarjeta que fallan”

Entonces $X \sim B(200, 0.01)$. Como n es grande y p chico aplicamos la aproximación Poisson con $np = 200 \times 0.01 = 2$

Se pide calcular la $P(X \geq 4)$

$$P(X \geq 4) = 1 - P(X \leq 3) \approx 1 - 0.857 = 0.143$$

↓
por tabla de la acumulada de la Poisson

- Sea la v.a. Y: “número de tarjetas entre 5 que funcionan bien”

Tenemos que $Y \sim B(5, p)$ donde $p = P(X = 0) \approx e^{-2} \frac{2^0}{0!} = e^{-2}$

Se pide calcular $P(Y \geq 4)$

$$P(Y \geq 4) = P(Y = 4) + P(Y = 5) = \binom{5}{4} (e^{-2})^4 (1 - e^{-2})^{5-4} + \binom{5}{5} (e^{-2})^5 (1 - e^{-2})^{5-5} = 5e^{-8} (1 - e^{-2}) + 5e^{-10}$$

Proceso de Poisson

Una aplicación importante de la distribución de Poisson se presenta en relación con el acontecimiento de eventos de un tipo particular en el tiempo. Por ejemplo, un evento podría ser un individuo entrando en un establecimiento en particular, o pulsos radiactivos registrados por un contador Geiger, o automóviles pasando por un cruce determinado.

Supongamos que tenemos eventos que ocurren en ciertos puntos aleatorios de tiempo, y asumimos que para alguna constante positiva λ las siguientes suposiciones se sostienen:

- 1- La probabilidad que exactamente 1 evento ocurra en un intervalo de longitud t es la misma para todos los intervalos de longitud t y es igual a $\lambda t + o(t)$ (donde $o(t)$ simboliza una función $f(t)$ tal que $\lim_{t \rightarrow 0} \frac{f(t)}{t} = 0$, por ejemplo $f(t) = t^2$ es $o(t)$, pero $f(t) = t$ no lo es)
- 2- La probabilidad que 2 o más eventos ocurran en un intervalo de longitud t es la misma para todos los intervalos de longitud t y es igual a $o(t)$.
- 3- Para cualesquiera enteros n, k_1, k_2, \dots, k_n y cualquier conjunto de n intervalos I_1, I_2, \dots, I_n que no se superpongan, si definimos los eventos E_i : "en el intervalo I_i ocurren exactamente k_i eventos" $i = 1, 2, \dots, n$, entonces los eventos E_1, E_2, \dots, E_n son independientes.

Las suposiciones 1 y 2 establecen que para pequeños valores de t , la probabilidad de que exactamente un evento ocurra en un intervalo de longitud t es igual a λt mas algo que es chico comparado con t , mientras que la probabilidad de que 2 o más eventos ocurran es pequeño comparado con t . la suposición 3 establece que lo que ocurra en un intervalo no tiene efecto (en la probabilidad) sobre lo que ocurrirá en otro intervalo que no se superponga.

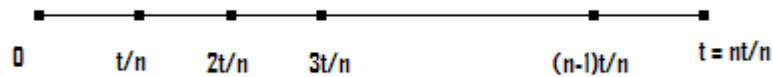
Bajo las suposiciones 1, 2 y 3 se puede probar que la v.a.

X : " número de eventos que ocurren en cualquier intervalo de longitud t " , tiene distribución Poisson con parámetro λt

Específicamente

$$P(X = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!} \quad k = 0, 1, 2, \dots$$

La idea de la demostración es la siguiente, partimos al intervalo $[0, t]$ en n subintervalos que no se superpongan cada uno de longitud $\frac{t}{n}$



Elegimos n suficientemente grande para que en cada subintervalo se tenga una ocurrencia exactamente o ninguna ocurrencia.

Sea la v.a. Y : “número de subintervalos en los que hay exactamente una ocurrencia”, entonces podemos asumir que $Y \sim B(n, p)$ donde p es la probabilidad que en un subintervalo hay exactamente una ocurrencia y si n es grande entonces la longitud del subintervalo $\frac{t}{n}$ es chica con lo cual por suposición 1 tenemos que $p \approx \lambda \frac{t}{n}$.

Entonces, utilizando la aproximación de Poisson a la binomial con parámetro $np = n\lambda \frac{t}{n} = \lambda t$ tenemos

$$P(X = k) = P(Y = k) \approx e^{-\lambda t} \frac{(\lambda t)^k}{k!} \quad k = 0, 1, 2, \dots$$

Observaciones:

- 1- Un **proceso temporal de Poisson** consiste en eventos que ocurren en el tiempo en forma aleatoria que cumplen con las suposiciones 1, 2 y 3.
- 2- El parámetro λ es la **tasa o rapidez del proceso**.
- 3- Si en lugar de observar eventos en el tiempo, consideramos observar eventos de algún tipo que ocurren en una región de dos o tres dimensiones, por ejemplo, podríamos seleccionar de un mapa una región R de un bosque, ir a esa región y contar el número de árboles. Cada árbol representaría un evento que ocurre en un punto particular del espacio. Bajo las suposiciones 1, 2 y 3, se puede demostrar que el número de eventos que ocurren en la región R tiene una distribución de Poisson con parámetro λa donde a es el área de R (λ se interpreta como la densidad del proceso). Se trata ahora de un **proceso espacial de Poisson**.

Ejemplos:

- 1- Suponga que aviones pequeños llegan a cierto aeropuerto según un proceso de Poisson, con tasa $\lambda = 8$ aviones por hora, de modo que el número de llegadas durante un período de t horas es una v.a. Poisson con parámetro $\lambda = 8t$.
 - a) ¿Cuál es la probabilidad de que exactamente 5 aviones pequeños lleguen durante un período de una hora? ¿Por lo menos 5?
 - b) ¿Cuál es la probabilidad de que por lo menos 20 aviones pequeños lleguen durante un período de 2 ½ hs? ¿De que a lo sumo 10 lleguen en ese período?

Solución:

- a) Sea la v.a. X : “número de aviones pequeños que llegan a cierto aeropuerto en una hora” Entonces $X \sim P(8)$. Por lo tanto

$$P(X = 5) = e^{-8} \frac{8^5}{5!} = 0.0916$$

O también usando la tabla de distribución acumulada para la Poisson

$$P(X = 5) = P(X \leq 5) - P(X \leq 4) = 0.191 - 0.099 = 0.092$$

Y la probabilidad de que lleguen al menos 5 aviones será

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.099 = 0.901$$

b) Sea la v.a. X : “número de aviones pequeños que llegan a cierto aeropuerto en $2 \frac{1}{2}$ horas”

Entonces $X \sim P(8 \times 2.5)$ es decir ahora $\lambda t = 8 \times 2.5 = 20$

$$P(X \geq 20) = 1 - P(X \leq 19) = 1 - 0.470 = 0.53$$

↓
por tabla de F.d.a.

Y por último calculamos por tabla $P(X \leq 10) = 0.010$

2- Se supone que el número de defectos en los rollos de tela de cierta industria textil es una v.a. Poisson con tasa 0.1 defectos por metro cuadrado..

a) ¿Cuál es la probabilidad de tener dos defectos en un metro cuadrado de tela?

b) ¿Cuál es la probabilidad de tener un defecto en 10 metros cuadrados de tela?

c) ¿Cuál es la probabilidad de que no halla defectos en 20 metros cuadrados de tela?

d) supongamos que el número de defectos está relacionado con la máquina que produce la tela, debido a desperfectos de la máquina el número de defectos varía en ciertos tramos del rollo. ¿Se puede asumir que el número de defectos sigue una distribución de Poisson?

Solución:

a) Sea X : “número de defectos en un metro cuadrado”. Entonces $X \sim P(0.1)$ pues
 $\lambda a = \lambda \times 1 = 0.1 \times 1 = 0.1$

$$P(X = 2) = e^{-0.1} \frac{0.1^2}{2!} = 0.004524$$

b) Si X : “número de defectos en 10 metros cuadrados”. Entonces $X \sim P(1)$ pues

$$\lambda a = \lambda \times 10 = 0.1 \times 10 = 1$$

$$P(X = 1) = e^{-1} \frac{1^1}{1!} = 0.3678794$$

c) X : “número de defectos en 20 metros cuadrados”. Entonces $X \sim P(2)$ pues

$$\lambda a = \lambda \times 20 = 0.1 \times 20 = 2$$

$$P(X = 0) = e^{-2} \frac{2^0}{0!} = 0.135$$

d) NO se puede asumir que el número de defectos sigue una distribución de Poisson, ya que las suposiciones que debe satisfacer un proceso de Poisson no se cumplirían.

Práctica

Distribución binomial, distribución geométrica, distribución binomial negativa o Pascal,

Distribución de Poisson

- 1) En cierto servicio telefónico, la probabilidad de que una llamada sea contestada en menos de 30 segundos es 0.75. Suponga que las llamadas son independientes.
 - a) Si una persona llama 10 veces, ¿cuál es la probabilidad de que exactamente 9 de las llamadas sean contestadas en un espacio de 30 segundos?
 - b) Si una persona llama 20 veces, ¿cuál es la probabilidad de que al menos 16 de las llamadas sean contestadas en un espacio de 30 segundos?
 - c) Si una persona llama 20 veces, ¿cuál es el número promedio de llamadas que serán contestadas en menos de 30 segundos?
- 2) Una persona pasa todas las mañanas a la misma hora por un cruce donde el semáforo está en verde el 20% de las veces. Suponga que cada mañana representa un ensayo independiente.
 - a) En cinco mañanas consecutivas, ¿cuál es la probabilidad de que el semáforo esté en verde exactamente un día?
 - b) En 20 mañanas, ¿cuál es la probabilidad de que el semáforo esté en verde exactamente 4 días?
 - c) En 20 mañanas, ¿cuál es la probabilidad de que el semáforo esté en verde más de 4 días?
 - d) En 20 mañanas, ¿cuál es el número esperado de veces que el semáforo estará en verde?
- 3) En referencia al ejercicio 2):
 - a) ¿Cuál es la probabilidad de que la primera mañana en que la luz del semáforo se encuentre en verde sea la cuarta mañana desde el inicio del experimento?
 - b) ¿Cuál es la probabilidad de que la luz del semáforo no se encuentre en verde durante 10 mañanas consecutivas?
- 4) Suponga que cada una de las llamadas que hace una persona a una estación de radio muy popular tiene una probabilidad de 0.02 de que la línea no esté ocupada. Suponga que las llamadas son independientes.
 - a) ¿Cuál es la probabilidad de que la primera llamada que entre sea la décima que realiza la persona?
 - b) ¿Cuál es la probabilidad de que sea necesario llamar más de 5 veces para hallar desocupada la línea?
 - c) ¿Cuál es el número promedio de llamadas que deben hacerse para hallar desocupada la línea?
- 5) La probabilidad de que la calibración de un transductor en un instrumento electrónico cumpla con las especificaciones del sistema de medición, es 0.6. Suponga que los intentos de calibración son independientes. ¿Cuál es la probabilidad de que se requieran como máximo 3 intentos para satisfacer las especificaciones del sistema de medición?
- 6) La escala electrónica de un proceso de llenado automático detiene la línea de producción después de haber detectado 3 paquetes con un peso menor que el especificado. Suponga que la probabilidad de llenar un paquete con un peso menor es 0.001 y que cada operación de llenado es independiente.

- a) ¿Cuál es el número promedio de operaciones de llenado antes de que se detenga la línea de producción?
 - b) ¿Cuál es la desviación estándar del número de operaciones de llenado antes de que se detenga la línea de producción?.
- 7) Suponga que el número de clientes que entran en un banco en una hora es una v.a. Poisson, y que $P(X = 0) = 0.05$. Hallar la media y la varianza de X .
- 8) A menudo, el número de llamadas telefónicas que llegan a un conmutador se modela como una v.a. Poisson. Suponga que, en promedio, se reciben 10 llamadas por hora.
- a) ¿Cuál es la probabilidad de que lleguen exactamente 5 llamadas en una hora?
 - b) ¿Cuál es la probabilidad de que se reciban 3 o menos llamadas en una hora?
 - c) ¿Cuál es la probabilidad de que se reciban exactamente 15 llamadas en dos horas?
 - d) ¿Cuál es la probabilidad de que lleguen exactamente 5 llamadas en 30 minutos?
- 9) Se supone que el número de defectos en los rollos de tela de cierta industria textil es una v.a. Poisson con una media de 0.1 defectos por metro cuadrado.
- a) ¿Cuál es la probabilidad de tener dos defectos en un metro cuadrado de tela?
 - b) ¿Cuál es la probabilidad de tener un defecto en 10 metros cuadrados de tela?
 - c) ¿Cuál es la probabilidad de que no halla defectos en 20 metros cuadrados de tela?
 - d) ¿Cuál es la probabilidad de que existan al menos 2 defectos en 10 metros cuadrados de tela?

3.5 – Variables aleatorias continuas

En la sección anterior se consideraron variables aleatorias discretas, o sea variables aleatorias cuyo rango es un conjunto finito o infinito numerable. Pero hay variables aleatorias cuyo rango son todos los números reales de un intervalo dado, (es decir es un conjunto infinito no numerable). Ejemplos de variables continuas podrían ser

X : “tiempo que tarda en llegar un colectivo a una parada”

Y : “tiempo de vida de un fusible”

Como ahora los valores de una v.a. continua no son contables no se puede hablar del i -ésimo valor de la v.a. X y por lo tanto $p(x_i) = P(X = x_i)$ pierde su significado. Lo que se hace es sustituir la función $p(x)$ definida sólo para x_1, x_2, \dots , por una función $f(x)$ definida para todos los valores x del rango de X . Por lo tanto se da la siguiente definición de v.a. continua

Sea X una v.a.. Decimos que es **continua** si existe una función no negativa f , definida sobre todos los reales $x \in (-\infty, \infty)$, tal que para cualquier conjunto B de números reales

$$P(X \in B) = \int_B f(x)dx$$

O sea que la probabilidad de que X tome valores en B se obtiene al integrar la función f sobre el conjunto B .

A la función f la llamamos función densidad de probabilidad (f.d.p.).

Observaciones:

1- Como X debe tomar algún valor real, entonces debe cumplirse que

$$1 = P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f(x)dx$$

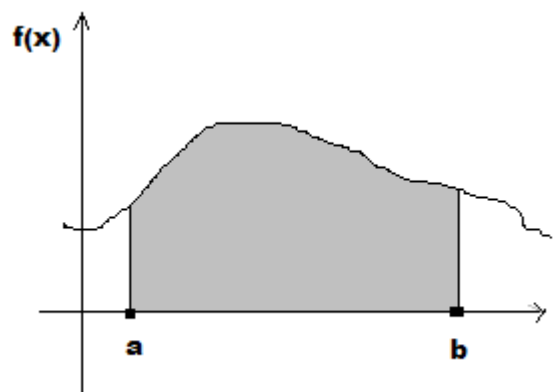
2- Si B es el intervalo real $[a, b] = \{x \in \mathbb{R}; a \leq x \leq b\}$ entonces

$$P(X \in B) = P(a \leq X \leq b) = \int_a^b f(x)dx$$

Notar que en este caso la probabilidad de que X tome valores en el intervalo $[a, b]$ es **el área bajo f entre a y b**

3- Si en la observación anterior $a = b$ entonces

$$P(X = a) = \int_a^a f(x)dx = 0$$



Es decir la probabilidad que una v.a. continua tome algún valor fijado es cero. Por lo tanto, para una v.a. continua

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = \int_a^b f(x)dx$$

Función de distribución acumulada

Sea X una v.a. continua. Se define la **función de distribución acumulada de X** (abreviamos F.d.a de X) como

$$F(x) = P(X \leq x) \quad -\infty < x < \infty$$

Si X tiene f.d.p. $f(x)$ entonces

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt \quad -\infty < x < \infty$$

Además

$$P(a \leq X \leq b) = \int_a^b f(x)dx = \int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx = F(b) - F(a)$$

Observaciones:

1- Si X es una v.a. con f.d.p. $f(x)$ y función de distribución acumulada $F(x)$ entonces

$$\frac{dF(x)}{dx} = \frac{d}{dx} \left(\int_{-\infty}^x f(t)dt \right) = f(x) \quad \text{donde } F(x) \text{ sea derivable}$$

Es decir, se puede obtener la función de densidad de X a partir de su F.d.a.

2- Como en el caso discreto vale

$$\text{Si } a < b \text{ entonces } F(a) \leq F(b) \text{ (es decir } F(x) \text{ es una función creciente)}$$

Y además se cumple que

$$\begin{aligned} \lim_{x \rightarrow \infty} F(x) &= \lim_{x \rightarrow \infty} \int_{-\infty}^x f(t)dt = \int_{-\infty}^{\infty} f(t)dt = 1 \\ \lim_{x \rightarrow -\infty} F(x) &= \lim_{x \rightarrow -\infty} \int_{-\infty}^x f(t)dt = \int_{-\infty}^{-\infty} f(t)dt = 0 \end{aligned}$$

Ejemplos:

1- Supongamos que X es una v.a. continua con f.d.p. dada por

$$f(x) = \begin{cases} C(4x - 2x^2) & \text{si } 0 \leq x \leq 2 \\ 0 & \text{caso contrario} \end{cases}$$

a) ¿Cuál es el valor de C ?

b) Hallar $P(X > 1)$

c) Hallar la F.d.a. de X

Solución:

a) Por lo dicho en la observación 1, se debe cumplir que $1 = \int_{-\infty}^{\infty} f(x)dx$, por lo tanto

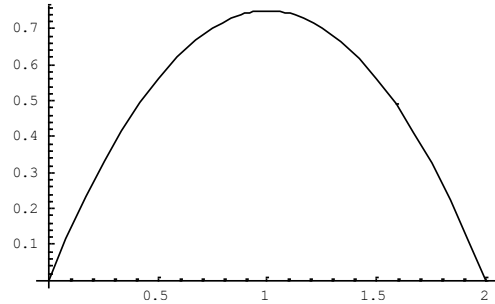
$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^0 0dx + \int_0^2 C(4x - 2x^2)dx + \int_2^{\infty} 0dx = \int_0^2 C(4x - 2x^2)dx$$

Entonces

$$C \int_0^2 (4x - 2x^2)dx = C \left(4 \frac{x^2}{2} - 2 \frac{x^3}{3} \right) \Big|_0^2 = C \frac{8}{3} = 1$$

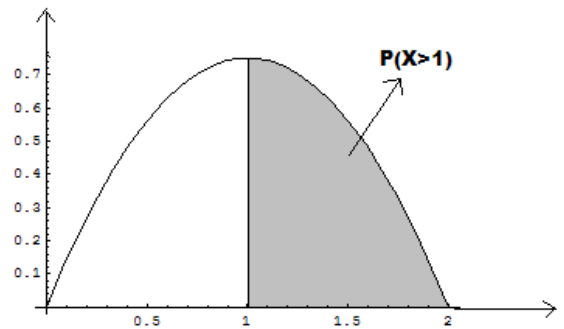
$$\therefore C = \frac{3}{8}$$

Es útil hacer un gráfico de la densidad



b) Para calcular la probabilidad que X sea mayor que 1, planteamos

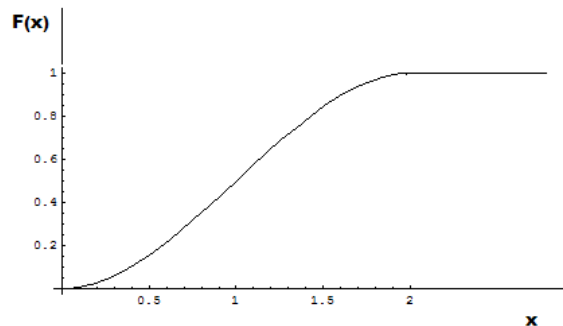
$$P(X > 1) = \int_1^2 \frac{3}{8}(4x - 2x^2)dx = \frac{1}{2}$$



c) Para calcular la F.d.a. notar que tenemos tres casos: $x < 0$, $0 \leq x \leq 2$ y $x > 2$, por lo tanto

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \int_0^x \frac{3}{8}(4t - 2t^2)dt & \text{si } 0 \leq x \leq 2 \\ 1 & \text{si } x > 2 \end{cases} \quad \text{es decir} \quad F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{3}{4}x^2\left(1 - \frac{x}{3}\right) & \text{si } 0 \leq x \leq 2 \\ 1 & \text{si } x > 2 \end{cases}$$

El gráfico de la F.d.a. es



Se podría haber calculado la $P(X > 1)$ a partir de la F.d.a. de la siguiente forma

$$P(X > 1) = 1 - P(X \leq 1) = 1 - F(1) = 1 - \frac{3}{4}1^2\left(1 - \frac{1}{3}\right) = 1 - \frac{3}{4} \times \frac{2}{3} = \frac{1}{2}$$

2- El tiempo de vida en horas que una computadora funciona antes de descomponerse es una v.a. continua con *f.d.p.* dada por

$$f(x) = \begin{cases} 0.01e^{-0.01x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

a) hallar la F.d.a. de X

b) ¿Cuál es la probabilidad que la computadora funcione entre 50 y 150 horas antes de descomponerse?

c) ¿Cuál es la probabilidad que una computadora se descomponga antes de registrar 100 horas de uso?

d) ¿Cuál es la probabilidad que exactamente 2 de 5 computadoras se descompongan antes de registrar 100 horas de uso?. Asumir que las computadoras trabajan en forma independiente.

Solución:

a) Hacemos un gráfico de la densidad y entonces observamos claramente que hay dos casos a considerar para calcular la F.d.a.: $x \geq 0$ y $x < 0$

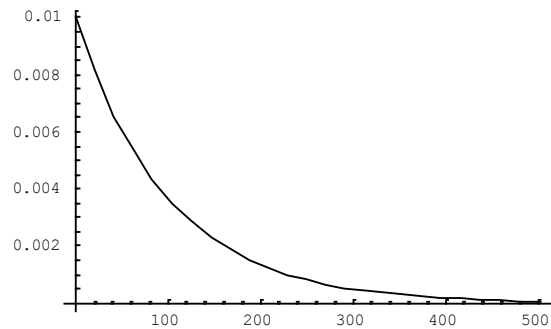
Si $x < 0$ entonces

$$F(x) = P(X \leq x) = \int_{-\infty}^0 0 dx = 0$$

Y si $x \geq 0$ tenemos

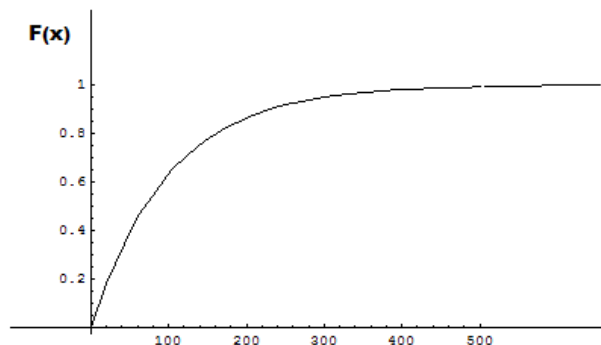
$$F(x) = P(X \leq x) = \int_{-\infty}^0 0 dt + \int_0^x 0.01e^{-0.01t} dt =$$

$$= 0 + 0.01 \left(\frac{e^{-0.01t}}{-0.01} \right) \Bigg|_0^x = 1 - e^{-0.01x}$$



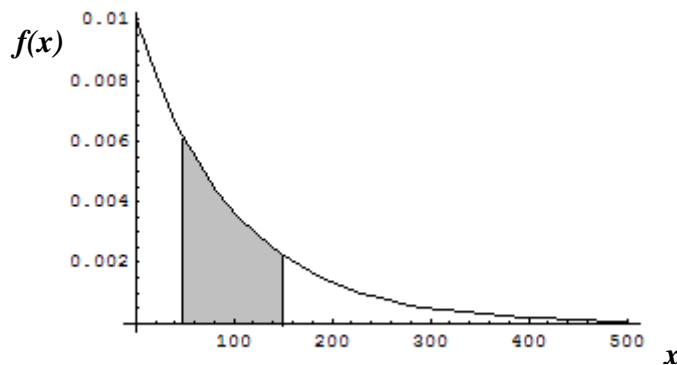
$$\therefore F(x) = \begin{cases} 1 - e^{-0.01x} & \text{si } x \geq 0 \\ 0 & \text{caso contrario} \end{cases}$$

La gráfica de la F.d.a. es



Se pide calcular $P(50 \leq X \leq 150)$, lo hacemos con la F.d.a.:

$$P(50 \leq X \leq 150) = F(150) - F(50) = e^{-0.01 \times 50} - e^{-0.01 \times 150} = 0.3834$$



c) Hay que calcular $P(X < 100)$

$$P(X < 100) = F(100) = 1 - e^{-0.01 \times 100} = 1 - e^{-1} \approx 0.63212$$

d) Podemos definir la v.a.

Y : “número de computadores entre 5 que se descomponen antes de las 100 horas de uso”

Entonces $Y \sim B(5, p)$ donde $p = P(X < 100)$

Por lo tanto hay que calcular

$$P(Y = 2) = \binom{5}{2} p^2 (1-p)^3 = \frac{5!}{2!3!} (1 - e^{-1})^2 (e^{-1})^3 \approx 0.198937$$

3.6 – Esperanza de una variable aleatoria continua

Para una v.a. discreta la $E(X)$ se definió como la suma de los $x_i p(x_i)$. Si X es una v.a. continua con f.d.p. $f(x)$, se define $E(X)$ sustituyendo la sumatoria por integración y $p(x_i)$ por $f(x)$.

La esperanza de una v.a. continua X con f.d.p. $f(x)$ se define como

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Ejemplo:

Para ciertas muestras de minerales la proporción de impurezas por muestra, es una v.a. X con f.d.p. dada por

$$f(x) = \begin{cases} \frac{3}{2}x^2 + x & \text{si } 0 < x < 1 \\ 0 & \text{caso contrario} \end{cases}$$

Hallar la esperanza de X

Solución:

Se plantea

$$E(X) = \int_0^1 x \left(\frac{3}{2} x^2 + x \right) dx = \left(\frac{3}{2} \frac{x^4}{4} + \frac{x^3}{3} \right) \Big|_0^1 = \left(\frac{3}{2} \times \frac{1}{4} + \frac{1}{3} \right) = \frac{17}{24}$$

A menudo se desea calcular la esperanza de una función de X , $Y = h(X)$, esto se puede hacer hallando previamente la densidad de Y y luego calcular $E(Y)$ aplicando la definición anterior.

Otra forma de calcular $E(Y)$ sin hallar la densidad de Y está dada por el siguiente

Teorema: Si X es una v.a. continua con $f.d.p.$ $f(x)$ y $h(X)$ es cualquier función de X , entonces

$$E(h(X)) = \int_{-\infty}^{\infty} h(x) f(x) dx$$

Dem.) sin demostración

Ejemplo: En el ejemplo anterior supongamos que el valor en dólares de cada muestra es $Y = h(X) = 5 - 0.5X$. Encontrar la esperanza de Y

Podemos hallar la esperanza de Y encontrando previamente su $f.d.p.$

Para esto se encuentra la $F.d.a.$ de Y , para luego hallar la densidad de Y derivando la $F.d.a.$

Anotamos $G(y)$ y $g(y)$ a la $F.d.a.$ de Y y a la densidad de Y respectivamente

$$G(y) = P(Y \leq y) = P(5 - 0.5X \leq y) = P\left(X \geq \frac{5-y}{0.5}\right) = 1 - P\left(X \leq \frac{5-y}{0.5}\right) = 1 - F\left(\frac{5-y}{0.5}\right)$$

Donde $F(x) = P(X \leq x)$

Entonces

$$g(y) = \frac{d}{dy} \left(1 - F\left(\frac{5-y}{0.5}\right) \right) = -f\left(\frac{5-y}{0.5}\right) \left(-\frac{1}{0.5} \right) = f\left(\frac{5-y}{0.5}\right) \left(\frac{1}{0.5} \right) =$$

$$= \begin{cases} \left(\frac{3}{2} \left(\frac{5-y}{0.5} \right)^2 + \frac{5-y}{0.5} \right) \left(\frac{1}{0.5} \right) & \text{si } 0 < \frac{5-y}{0.5} < 1 \\ 0 & \text{caso contrario} \end{cases}$$

O sea

$$g(y) = \begin{cases} \left(\frac{3}{2} \left(\frac{5-y}{0.5} \right)^2 + \frac{5-y}{0.5} \right) \left(\frac{1}{0.5} \right) & \text{si } \frac{9}{2} < y < 5 \\ 0 & \text{caso contrario} \end{cases}$$

Ahora calculamos la $E(Y)$

$$E(Y) = \int_{9/2}^5 y \left(\frac{3}{2} \left(\frac{5-y}{0.5} \right)^2 + \frac{5-y}{0.5} \right) \frac{1}{0.5} dy = \frac{223}{46}$$

Aplicando el teorema anterior los cálculos se reducen:

$$E(Y) = E(h(X)) = \int_0^1 \underbrace{(5 - 0.5x)}_{h(x)} \underbrace{\left(\frac{3}{2} x^2 + x \right)}_{f(x)} dx = \frac{223}{46}$$

Notar que de la misma forma que en el caso discreto, si $h(x) = ax + b$, es decir si h es una función lineal, aplicando las propiedades de linealidad de la integral tenemos

$$E(aX + b) = aE(X) + b$$

En el ejemplo anterior se podía encontrar la esperanza de Y haciendo

$$E(Y) = E(h(X)) = E(5 - 0.5X) = 5 - 0.5E(X) = 5 - 0.5 \times \frac{17}{24} = \frac{223}{46}$$

Varianza de una variable aleatoria continua

Sea X una v.a. continua con f.d.p. $f(x)$ y sea $E(X) = \mu$, entonces la varianza de X es

$$V(X) = \sigma_X^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

La interpretación de la varianza de una v.a. continua es la misma que para el caso discreto. Además sigue valiendo la igualdad

$$V(X) = E(X^2) - \mu^2$$

Pues en la demostración hecha para el caso discreto si sustituyen las sumatorias por integrales. Por la misma razón, también vale que

$$V(aX + b) = a^2 V(X) \quad \text{y} \quad \sigma_{aX+b} = |a| \sigma_X$$

Ejemplo:

Calculamos la varianza de $Y = 5 - 0.5X$

$$V(Y) = 0.5^2 V(X)$$

$$V(X) = E(X^2) - \mu^2 = E(X^2) - \left(\frac{17}{24} \right)^2$$

$$E(X^2) = \int_0^1 x^2 \left(\frac{3}{2} x^2 + x \right) dx = \left(\frac{3}{2} \frac{x^5}{5} + \frac{x^4}{4} \right) \Big|_0^1 = \left(\frac{3}{2} \times \frac{1}{5} + \frac{1}{4} \right) = \frac{11}{20}$$

$$\therefore V(X) = \frac{11}{20} - \left(\frac{17}{24}\right)^2 \Rightarrow V(Y) = 0.5^2 \left(\frac{11}{20} - \left(\frac{17}{24}\right)^2 \right) = \frac{139}{11520}$$

Práctica**Variables aleatorias continuas y distribuciones de probabilidad**

- 1) El tiempo total, medido en unidades de 100 horas, que un adolescente utiliza su estéreo en un período de un año es una v.a. continua X con f.d.p. dada por

$$f(x) = \begin{cases} x, & 0 < x < 1 \\ 2 - x, & 1 \leq x < 2 \\ 0, & \text{caso contrario} \end{cases}$$

Encuentre la probabilidad de que en un período de un año el adolescente utilice su estéreo

- menos de 120 horas
- entre 50 y 100 horas

- 2) Suponga que la distancia X entre un blanco puntual y un disparo dirigido al punto, en un juego de tiro al blanco accionado por monedas, es una v.a. continua con f.d.p. dada por:

$$f(x) = \begin{cases} 0.75(1 - x^2) & -1 \leq x \leq 1 \\ 0 & \text{caso contrario} \end{cases}$$

- Calcular $P(X > 0)$
- Calcular $P(-0.5 \leq X \leq 0.5)$
- Calcular $P(X < -0.25 \text{ ó } X > 0.25)$
- Hallar la F.d.a. de X .

- 3) Considere la función de densidad

$$f(x) = \begin{cases} k\sqrt{x} & 0 < x < 1 \\ 0 & \text{c.c.} \end{cases}$$

- Evalúe k
- Encuentre $F(x)$
- Evalúe $P(0.3 < X < 0.6)$ utilizando $F(x)$

- 4) Una barra de 12 pulgadas, que está sujeta por ambos extremos, debe someterse a un creciente cantidad de esfuerzo hasta que se rompa.

Sea X : “distancia desde el extremo izquierdo en el que ocurre la rotura” y suponga que la f.d.p. de X es

$$f(x) = \begin{cases} \frac{1}{24} x \left(1 - \frac{x}{12}\right) & 0 < x < 12 \\ 0 & \text{en caso contrario} \end{cases}$$

Calcular:

- F.d.a de X
- $P(X \leq 4)$; $P(X > 6)$; $P(4 < X < 6)$
- $E(X)$
- la probabilidad de que el punto de ruptura ocurra a más de 2 pulgadas del punto esperado de ruptura.

- 5) La Fda de la v.a. X está dada por

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x^2}{4} & 0 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

- Hallar la fdp de X . Calcular $E(X)$ y $V(X)$
- Calcular $P(X \leq 1)$ y $P(0.5 \leq X \leq 1)$

- 6) Para las variables aleatorias de los ejercicios anteriores hallar su esperanza y desviación estándar.
- 7) Si la temperatura a la que un cierto compuesto se funde es una variable aleatoria con valor medio de 120°C y desviación estándar de 2°C , ¿cuáles son la temperatura media y la desviación estándar medidas en $^{\circ}\text{F}$? (Sugerencia: $^{\circ}\text{F} = 1.8^{\circ}\text{C} + 32$).
- 8) Para la v.a. del ejercicio 1), sea la v.a. Y el número de kilowatts-hora que el adolescente gasta al año; se tiene que $Y = 60X^2 + 39X$. Calcule la esperanza de Y . Explique qué propiedad utiliza.

3.7 - Variables aleatorias continuas importantes

Distribución uniforme

La distribución continua más sencilla es análoga a su contraparte discreta.

Una v.a. continua X se dice que tiene **distribución uniforme en el intervalo** $[a, b]$, con $a < b$, si tiene función de densidad de probabilidad dada por

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{caso contrario} \end{cases}$$

La figura muestra la gráfica de la f.d.p.

Notación: $X \sim U[a, b]$

Es fácil verificar que $f(x)$ es una f.d.p. pues $f(x) \geq 0$ para todo x , y además

$$\int_{-\infty}^{\infty} f(x) dx = \int_a^b \frac{1}{b-a} dx = \frac{1}{b-a} x \Big|_a^b = \frac{1}{b-a} (b-a) = 1$$

La F.d.a. para $a \leq x \leq b$ sería

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt = \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}$$

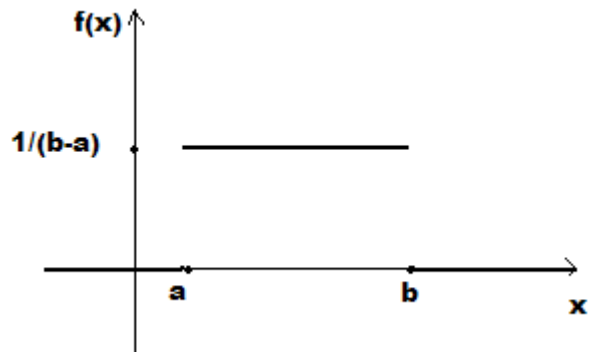
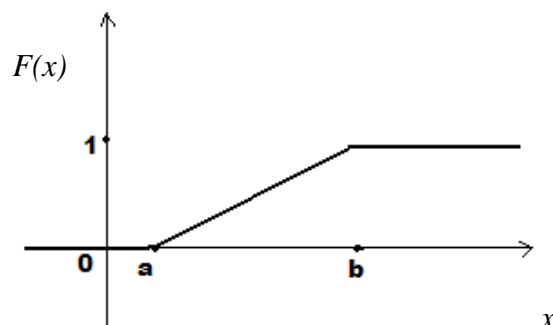
Para $x < a$, tenemos que $P(X \leq x) = \int_{-\infty}^x 0 dt = 0$

Y para $x > b$ tenemos que $P(X \leq x) = \int_{-\infty}^a 0 dt + \int_a^b \frac{1}{b-a} dt + \int_b^x 0 dt = 1$

Por lo tanto la F.d.a. es

$$F(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$$

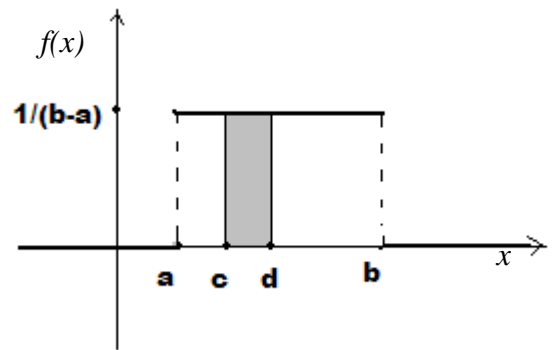
Y su gráfica es



Observación:

Si $X \sim U[a, b]$ y $a \leq c < d \leq b$, entonces

$$P(c \leq X \leq d) = F(d) - F(c) = \frac{d-a}{b-a} - \frac{c-a}{b-a} = \frac{d-c}{b-a}$$



Ejemplo:

Los colectivos de una determinada línea llegan a una parada en particular en intervalos de 15 minutos comenzando desde las 7 A.M. Esto es, ellos llegan a la parada a las 7, 7:15, 7:30, 7:45 y así siguiendo. Si un pasajero llega a la parada en un tiempo que se puede considerar una v.a. distribuida uniformemente entre 7 y 7:30, encontrar la probabilidad de que

- el pasajero espere menos de 5 minutos al colectivo
- el pasajero espere más de 10 minutos al colectivo

Solución:

Sea X : “tiempo en minutos desde las 7 hs en que el pasajero llega a la parada”

Entonces podemos considerar que $X \sim U[0, 30]$

- si el pasajero espera menos de 5 minutos al colectivo, entonces llega a la parada entre las 7:10 y 7:15 o entre las 7:25 y 7:30, entonces la probabilidad pedida es

$$P(10 \leq X \leq 15) + P(25 \leq X \leq 30) = \int_{10}^{15} \frac{1}{30} dx + \int_{25}^{30} \frac{1}{30} dx = \frac{1}{3}$$

O también se puede plantear directamente

$$P(10 \leq X \leq 15) + P(25 \leq X \leq 30) = \frac{15-10}{30} + \frac{30-25}{30} = \frac{2 \times 5}{30} = \frac{1}{3}$$

- Análogamente si debe esperar más de 10 minutos, deberá llegar entre las 7 y las 7:05 o entre las 7:15 y 7:20, por lo tanto la probabilidad pedida es

$$P(0 \leq X \leq 5) + P(15 \leq X \leq 20) = \frac{5}{30} + \frac{5}{30} = \frac{2 \times 5}{30} = \frac{1}{3}$$

Esperanza y varianza

Sea una X variable aleatoria continua distribuida uniformemente en el intervalo $[a, b]$, es decir

$$X \sim U[a, b]. \text{ Entonces, } E(X) = \frac{a+b}{2} \text{ y } V(X) = \frac{(b-a)^2}{12}.$$

Dem.)

Recordamos que la *f.d.p.* de una v.a. $X \sim U[a, b]$ es

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{caso contrario} \end{cases}$$

Entonces:

$$E(X) = \int_{-\infty}^{\infty} f(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{1}{b-a} \frac{x^2}{2} \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}$$

Notar que $(a+b)/2$ representa el punto medio del intervalo $[a, b]$ (como es de esperar por el significado de la distribución y de la esperanza):

Calculamos ahora la varianza de X

Deseamos calcular $V(X) = E(X^2) - [E(X)]^2$, es decir

$V(X) = E(X^2) - \left(\frac{a+b}{2}\right)^2$. Debemos obtener $E(X^2)$:

$$E(X^2) = \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{b-a} \cdot \frac{x^3}{3} \Big|_a^b = \frac{1}{3} \frac{b^3 - a^3}{b-a} = \frac{b^2 + b.a + a^2}{3}. \text{ Entonces:}$$

$$V(X) = \frac{b^2 + b.a + a^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

Distribución normal o gaussiana

Sea X una v.a. Decimos que tiene distribución normal con parámetros μ y σ si su *f.d.p.* es de la forma

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty \quad (13)$$

Donde $\mu \in R$ y $\sigma > 0$

Notación: $X \sim N(\mu, \sigma^2)$

Para darse una idea de la forma de la gráfica notar que:

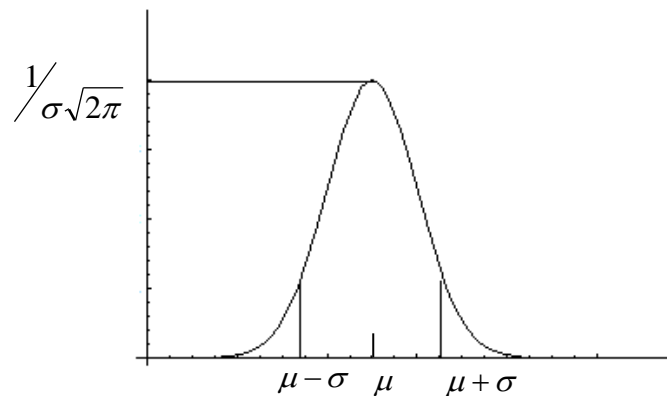
1- $f(x)$ es simétrica alrededor de μ , es decir $f(\mu+x) = f(\mu-x)$ para todo x

2- $\lim_{x \rightarrow \pm\infty} f(x) = 0$ (eje x asíntota horizontal)

3- Si planteamos $\frac{d}{dx} f(x) = 0 \Rightarrow x = \mu$. Se puede verificar que en $x = \mu$ la función tiene un máximo absoluto, $f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}$

4- Si planteamos $\frac{d^2}{dx^2} f(x) = 0 \Rightarrow x = \mu \pm \sigma$. Se puede verificar que en $x = \mu - \sigma$ y en $x = \mu + \sigma$ la función tiene dos puntos de inflexión, y además en el intervalo $(\mu - \sigma, \mu + \sigma)$ la función es cóncava hacia abajo y fuera de ese intervalo es cóncava hacia arriba

La gráfica de $f(x)$ tiene forma de campana

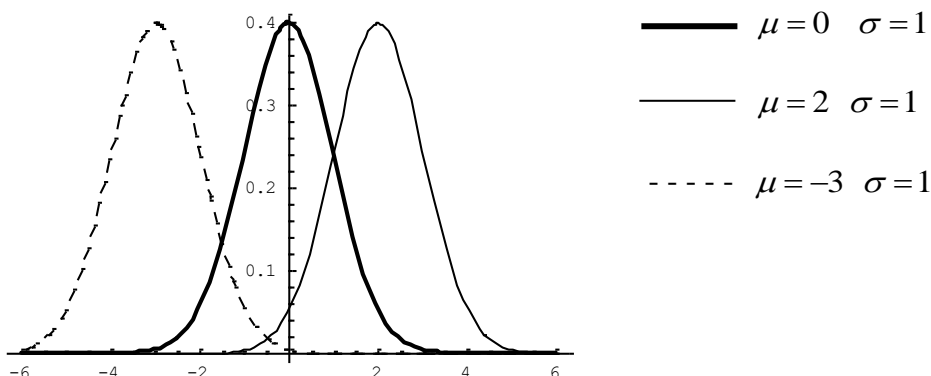


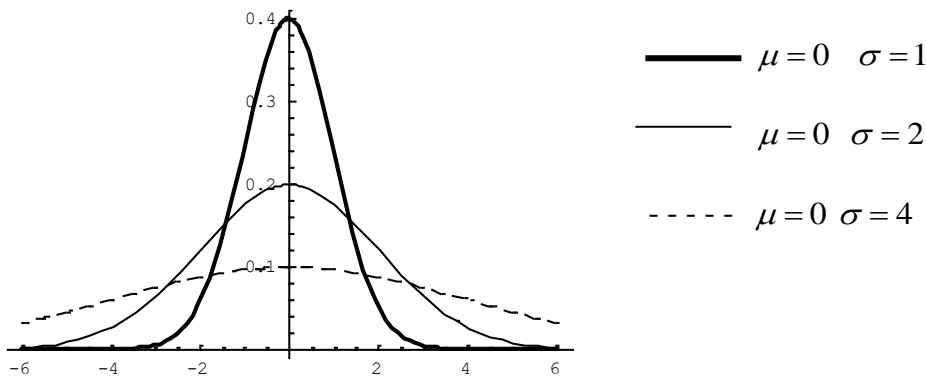
Observación:

Cuando μ varía la gráfica de la función se *traslada*, μ es un *parámetro de posición*.

Cuando σ aumenta, la gráfica se “achata”, cuando σ disminuye la gráfica se hace más “puntiaguda”, se dice que σ es un *parámetro de escala*.

En las siguientes figuras vemos cómo varía la gráfica de $f(x)$ con la variación de los parámetros





Se puede probar que $f(x)$ es una *f.d.p.* es decir que

a) $f(x) \geq 0$ para todo x

b) $\int_{-\infty}^{\infty} f(x)dx = 1$

Que a) es cierta es obvio; para probar b) es necesario recurrir al cálculo en dos variables (no lo demostramos).

Si $\mu = 0$ y $\sigma = 1$ entonces se dice que X tiene **distribución normal estándar**. Se anota $X \sim N(0,1)$

En este caso la *f.d.p.* se simboliza con $\varphi(x)$, es decir

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad -\infty < x < \infty$$

En este caso la gráfica de la densidad es simétrica con respecto al origen.

La *F.d.a.* de una v.a. normal estándar se anota $\Phi(x)$

$$\Phi(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

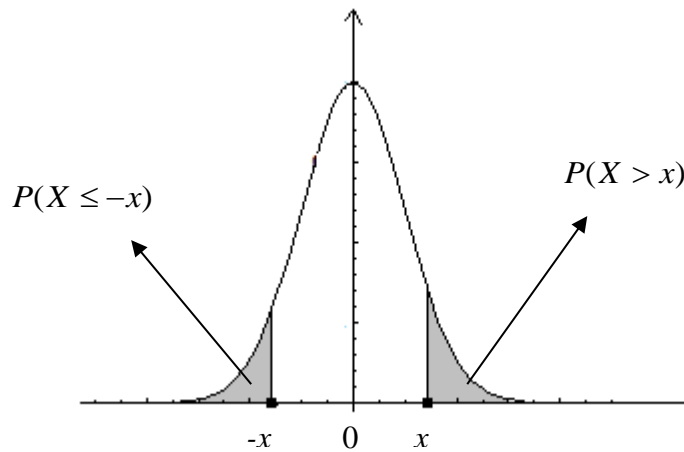
Esta integral no puede expresarse en términos de funciones elementales, por lo tanto se calcula $\Phi(x)$

Para valores específicos de x mediante una aproximación numérica.

Esto ya está hecho, existen tablas de la función de distribución acumulada de la normal estándar para valores de x que oscilan en general entre -4 y 4, pues para valores de x **menores que -4**, $\Phi(x) \approx 0$, y para valores de x **mayores que 4**, $\Phi(x) \approx 1$

Notar que como la $\varphi(x)$ es simétrica con respecto al origen entonces

$$\Phi(-x) = P(X \leq -x) = P(X > x) = 1 - P(X \leq x) = 1 - \Phi(x)$$



Por ejemplo, si $X \sim N(0,1)$ entonces utilizando la tabla de la F.d.a. de X

a) $P(X \leq 1.26) = \Phi(1.26) = 0.89616$

b) $P(X > 1.26) = 1 - P(X \leq 1.26) = 1 - \Phi(1.26) = 1 - 0.89616 = 0.10384$

c) $P(X > -1.37) = P(X \leq 1.37) = \Phi(1.37) = 0.91465$

d) $P(-1.25 < X < 0.37) = P(X < 0.37) - P(X < -1.25) = \Phi(0.37) - \Phi(-1.25) =$
 $= \Phi(0.37) - (1 - \Phi(1.25)) = 0.64431 - (1 - 0.89435) = 0.53866$

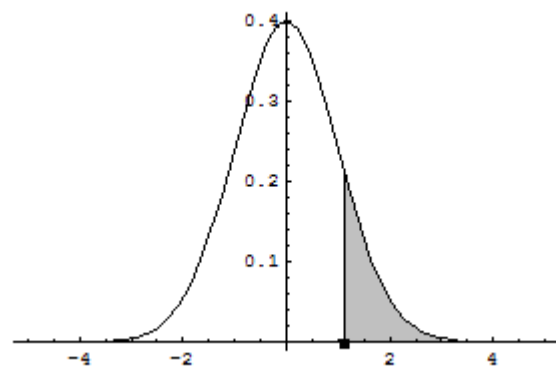
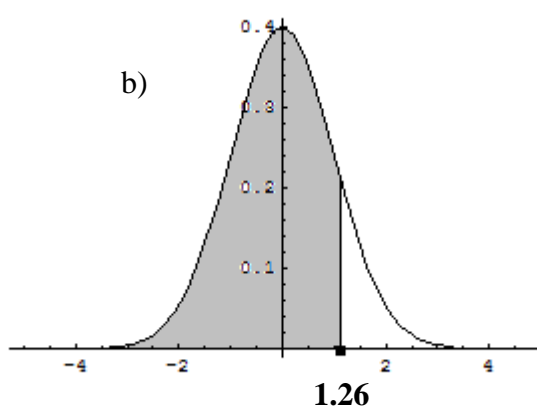
e) ¿Para qué valor x se cumple que $P(-x < X < x) = 0.95$?

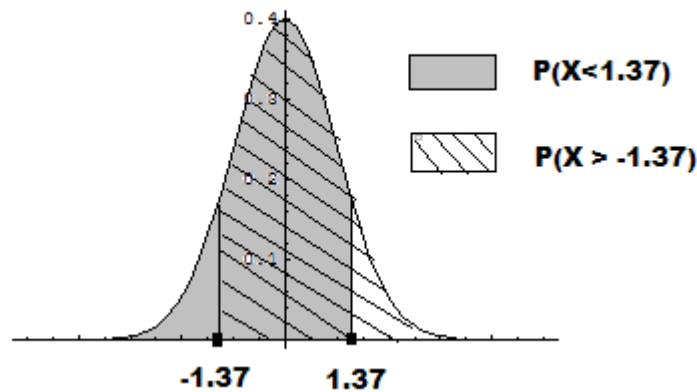
Tenemos que $P(-x < X < x) = \Phi(x) - \Phi(-x) = \Phi(x) - (1 - \Phi(x)) = 2\Phi(x) - 1$

Por lo tanto $2\Phi(x) - 1 = 0.95 \Rightarrow \Phi(x) = \frac{0.95 + 1}{2} = 0.975$

Observamos en la tabla de la F.d.a. que $x = 1.96$, pues $\Phi(1.96) = 0.975$

Para los incisos a), b) y c) se grafican las regiones correspondientes





Una propiedad importante de la distribución normal es que si $X \sim N(\mu, \sigma^2)$ entonces la v.a. $Y = aX + b$ con a y b números reales, $a \neq 0$, tiene también distribución normal pero con parámetros $a\mu + b$ y $a^2\sigma^2$, es decir

$$X \sim N(\mu, \sigma^2) \Rightarrow aX + b \sim N(a\mu + b, a^2\sigma^2) \quad (14)$$

Podemos demostrar lo anterior primero hallando la F.d.a. de Y

$$G(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F\left(\frac{y-b}{a}\right) \text{ donde } F \text{ es la F.d.a. de } X$$

\downarrow
 $a > 0$

Por lo tanto, la f.d.p. de Y la obtenemos derivando $G(y)$

$$g(y) = \frac{d}{dy} G(y) = \frac{d}{dy} F\left(\frac{y-b}{a}\right) = f\left(\frac{y-b}{a}\right) \frac{1}{a} = \frac{1}{a} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\frac{y-b}{a}-\mu}{\sigma}\right)^2} \quad \text{donde}$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Operando en el exponente se llega a

$$g(y) = \frac{1}{a} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-(a\mu+b)}{a\sigma}\right)^2}$$

Si $a < 0$ entonces

$$G(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \geq \frac{y-b}{a}\right) = 1 - F\left(\frac{y-b}{a}\right)$$

Y derivando con respecto a y obtenemos

$$g(y) = \frac{d}{dy} G(y) = \frac{d}{dy} \left(1 - F\left(\frac{y-b}{a}\right) \right) = -f\left(\frac{y-b}{a}\right) \frac{1}{a} = -\frac{1}{a} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-b-\mu}{\sigma}\right)^2}$$

O sea, para $a \neq 0$ la f.d.p. de Y es $g(y) = \frac{1}{|a|} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-(a\mu+b)}{a\sigma}\right)^2}$

Y comparando con (13) se deduce que $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$

Una consecuencia importante del resultado (14) es que

$$\text{si } X \sim N(\mu, \sigma^2) \text{ entonces } Y = \frac{X - \mu}{\sigma} \sim N(0,1) \quad (15)$$

Notar que Y se puede escribir como $Y = \frac{1}{\sigma} X + \left(-\frac{\mu}{\sigma}\right)$ es decir claramente Y es una función lineal de X

Por lo tanto aplicamos el resultado (14) con $a = \frac{1}{\sigma}$ y $b = -\frac{\mu}{\sigma}$ y llegamos a (15).

Si $X \sim N(\mu, \sigma^2)$ entonces la F.d.a. de X es

$$F(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

$F(x)$ no puede expresarse en términos de funciones elementales y sólo hay tablas de la F.d.a. de la normal estándar.

Para calcular $F(x)$ procedemos de la siguiente forma

$$F(x) = P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P\left(Y \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

\swarrow
 $Y \sim N(0,1)$

Ejemplos:

1- Si $X \sim N(3,9)$ entonces

a) $P(2 < X < 5) = P\left(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right) = \Phi\left(\frac{5-3}{3}\right) - \Phi\left(\frac{2-3}{3}\right) = \Phi\left(\frac{2}{3}\right) - \Phi\left(-\frac{1}{3}\right) = 0.3779$

b) $P(X > 0) = P\left(\frac{X-3}{3} > \frac{0-3}{3}\right) = 1 - \Phi(-1) = \Phi(1) = 0.8413$

c) $P(|X - 3| > 6) = 1 - P(|X - 3| \leq 6) = 1 - [P(-6 \leq X - 3 \leq 6)] = 1 - \left[P\left(\frac{-6}{3} \leq \frac{X-3}{3} \leq \frac{6}{3}\right)\right] =$
 $= 1 - [\Phi(2) - \Phi(-2)] = 2[1 - \Phi(2)] = 0.0456$

2- Hay dos máquinas para cortar corchos destinados para usarse en botellas de vino. La primera produce corchos con diámetros que están normalmente distribuidos con media de 3 cm y desviación estándar de 0.1 cm. La segunda máquina produce corchos con diámetros que tienen una distribución

normal con media de 3.04 cm y desviación estándar de 0.02 cm. Los corchos aceptables tienen diámetros entre 2.9 cm y 3.1 cm. ¿Cuál máquina tiene más probabilidad de producir un corcho aceptable?

Solución:

Sean las variables aleatorias

X : “diámetro de un corcho producido por la máquina 1”

Y : “diámetro de un corcho producido por la máquina 2”

Entonces $X \sim N(3, 0.1^2)$ y $Y \sim N(3.04, 0.02^2)$

Calculamos cuál es la probabilidad que la máquina 1 produzca un corcho aceptable

$$P(2.9 \leq X \leq 3.1) = P\left(\frac{2.9-3}{0.1} \leq \frac{X-3}{0.1} \leq \frac{3.1-3}{0.1}\right) = \Phi\left(\frac{3.1-3}{0.1}\right) - \Phi\left(\frac{2.9-3}{0.1}\right) = \\ = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = 0.6826$$

Análogamente para la máquina 2

$$P(2.9 \leq Y \leq 3.1) = P\left(\frac{2.9-3.04}{0.02} \leq \frac{Y-3.04}{0.02} \leq \frac{3.1-3.04}{0.02}\right) = \Phi\left(\frac{3.1-3.04}{0.02}\right) - \Phi\left(\frac{2.9-3.04}{0.02}\right) = \\ = \Phi(3) - \Phi(-7) = 0.9987 - 0 = 0.9987$$

Entonces es más probable que la máquina 2 produzca corchos aceptables.

3- El dispositivo automático de apertura de un paracaídas militar de carga se ha diseñado para abrir el paracaídas cuando éste se encuentre a 200 m de altura sobre el suelo. Supongamos que la altitud de apertura en realidad tiene una distribución normal con valor medio de 200 m y desviación estándar de 30 m. Habrá un daño al equipo si el paracaídas se abre a una altitud de menos de 100 m. ¿Cuál es la probabilidad de que haya daño a la carga en al menos uno de cinco paracaídas lanzados independientemente?

Solución:

Sea la v.a. X : “altitud de apertura en metros de un paracaídas”

Entonces $X \sim N(200, 30^2)$

$$\text{Calculamos } P(X < 100) = \Phi\left(\frac{100-200}{30}\right) = \Phi(-3.33) = 0.0004$$

Consideramos ahora la v.a. Y : “número de paracaídas entre 5 que se abren a menos de 100 metros”

Podemos considerar que $Y \sim B(5, 0.0004)$

Por lo tanto hay que calcular $P(Y \geq 1)$

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - \binom{5}{0} 0.0004^0 (1 - 0.0004)^{5-0} = 1 - (1 - 0.0004)^5 = 0.0019984$$

4- Supóngase que la resistencia a romperse (en Kgr) de fibras de yute está descrita por una v.a. continua X normalmente distribuida con $\mu = E(X) = 165$ Kgr y $\sigma^2 = V(X) = 9$ (Kgr)². suponiendo además que una muestra de esta fibra se considera defectuosa si $X < 162$. ¿Cuál es la probabilidad de que una fibra elegida al azar sea defectuosa?

Solución:

Deseamos conocer $P(X < 162)$

$$P(X < 162) = P\left(\frac{X - 165}{3} < \frac{162 - 165}{3}\right) = P\left(\frac{X - 165}{3} < -1\right) = \Phi(-1),$$

puesto que $Z = \frac{X - 165}{3} \sim N(0,1)$. Entonces

$$P(X < 162) = \Phi(-1) = 1 - \Phi(1).$$

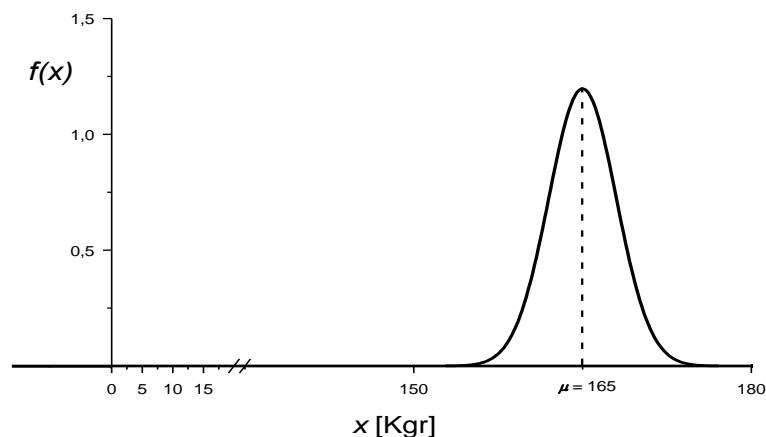
De la tabla tenemos $\Phi(1) = 0.8413 \rightarrow \Phi(-1) = 1 - \Phi(1) = 0.1587$.

Es decir $P(X < 162) = 0.1587$.

Observación:

Uno puede objetar el usar una distribución normal para describir a la v.a. X que representa la resistencia a romperse de la fibra ya que ésta es, obviamente, una cantidad no negativa, mientras que una v.a. normalmente distribuida puede tomar valores que varían entre $-\infty$ y $+\infty$. Sin embargo al modelar el problema con una normal (que aparentemente debería ser invalidada como modelo por lo señalado) vemos que les estamos asignando al suceso $\{X < 0\}$ una probabilidad prácticamente nula (ver también la figura siguiente):

$$P(X < 0) = P\left(\frac{X - 165}{3} < \frac{0 - 165}{3}\right) = \Phi(-55) = 1 - \Phi(55) \approx 1 - 1 = 0.$$



En casos como estos se justifica usar la distribución normal para modelar situaciones en que la variable aleatoria considerada puede tomar, por su significado, sólo valores positivos, aún cuando la normal permita tomar valores tanto positivos como negativos por cuanto la probabilidad de que la v.a. tome valores negativos es prácticamente nula.

Esperanza y varianza de una variable aleatoria con distribución normal

Sea $X \sim N(\mu, \sigma^2)$ entonces $E(X) = \mu$ y $V(X) = \sigma^2$

Dem.) Usaremos el resultado:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = 1$$

Calculamos la esperanza de X

$$E(X) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad \text{hacemos la sustitución}$$

$$t = \frac{x-\mu}{\sigma} \rightarrow x = (\sigma \cdot t + \mu), \quad dt = \frac{dx}{\sigma} \quad y \quad -\infty < t < +\infty.$$

Luego:

$$E(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma \cdot t + \mu) \cdot e^{-\frac{t^2}{2}} dt = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2}} dt + \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt$$

Considerado como función de t , el integrando de la primera integral $f(t) = t e^{-\frac{t^2}{2}}$ es una función impar de t , es decir, verifica $f(-t) = -f(t)$. En consecuencia la integral a lo largo de un intervalo simétrico con respecto al origen, como lo es $(-\infty, \infty)$ se anula.

El segundo sumando, por su parte, es justamente μ veces la integral $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = 1$. Entonces $E(X) = \mu$.

Veamos el cálculo de la varianza de X

$$V(X) = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2. \text{ Nos queda evaluar } E(X^2),$$

$$E(X^2) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x^2 \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx. \text{ Hacemos nuevamente la sustitución}$$

$$t = \frac{x-\mu}{\sigma} \rightarrow x = (\sigma \cdot t + \mu), \quad dt = \frac{dx}{\sigma} \quad y \quad -\infty < t < +\infty. \text{ Reemplazando:}$$

$$E(X^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma \cdot t + \mu)^2 \cdot e^{-\frac{t^2}{2}} dt = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 \cdot e^{-\frac{t^2}{2}} dt + \frac{2\sigma \cdot \mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t \cdot e^{-\frac{t^2}{2}} dt + \frac{\mu^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt$$

Ahora el integrando de la segunda integral es impar y por lo tanto se anula. Entonces

$$E(X^2) = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 \cdot e^{-\frac{t^2}{2}} dt + \mu^2.$$

Debemos calcular la integral $\int_{-\infty}^{+\infty} t^2 \cdot e^{-\frac{t^2}{2}} dt$. Integrando por partes

$$\text{con } \begin{cases} u = t & du = dt \\ dv = t e^{-\frac{t^2}{2}} dt & v = -e^{-\frac{t^2}{2}} \end{cases} \rightarrow \int_{-\infty}^{+\infty} t^2 \cdot e^{-\frac{t^2}{2}} dt = -t e^{-\frac{t^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt.$$

El corchete de Barrow se anula en ambos extremos y la integral es justamente $\sqrt{2\pi}I = \sqrt{2\pi}$. Entonces

: $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 \cdot e^{-\frac{t^2}{2}} dt = 1$. Por lo tanto $E(X^2) = \sigma^2 + \mu^2$ y, en consecuencia,

$$V(X) = E(X^2) - \mu^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$$

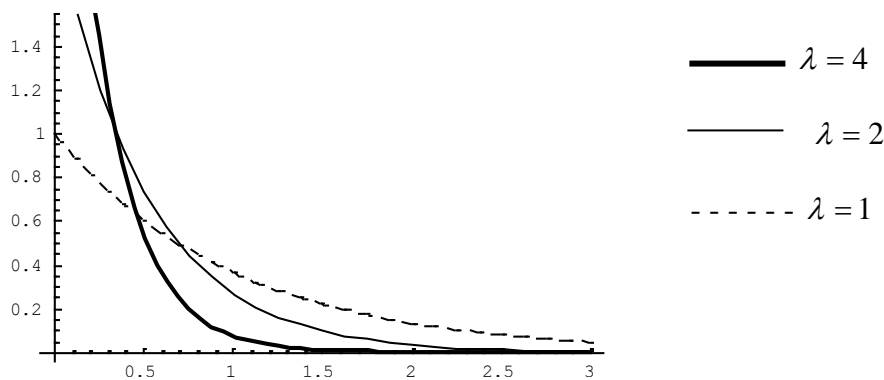
Distribución exponencial

Sea X una v.a. continua. Se dice que tiene distribución exponencial con parámetro λ si su f.d.p. es de la forma

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{caso contrario} \end{cases} \quad \text{donde } \lambda > 0$$

La distribución exponencial se utiliza algunas veces para modelar el tiempo que transcurre antes de que ocurra un evento. A menudo se lo llama **tiempo de espera**.

La siguiente figura muestra la gráfica de la densidad para diferentes valores del parámetro



Es fácil verificar que $f(x)$ es una densidad bien definida

a) Claramente $f(x) \geq 0$ para todo x

$$b) \int_0^{\infty} \lambda e^{-\lambda t} dt = \lambda \left(\frac{e^{-\lambda t}}{-\lambda} \right) \Big|_0^{\infty} = 1$$

La F.d.a. de una v.a. exponencial es sencilla:

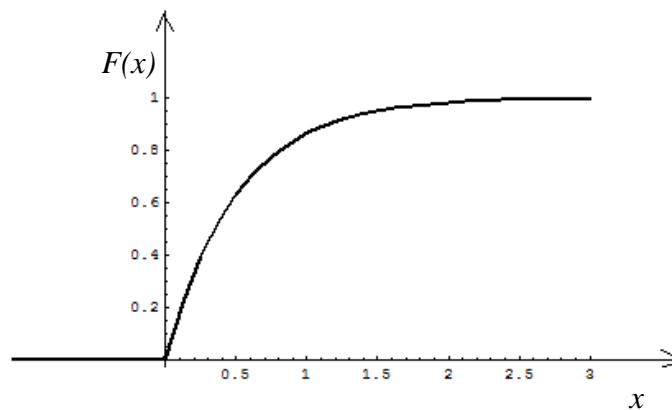
Si $x < 0$ entonces $F(x) = P(X \leq x) = 0$

$$\text{Si } x \geq 0 \text{ entonces } F(x) = P(X \leq x) = \int_0^x \lambda e^{-\lambda t} dt = \lambda \left(\frac{e^{-\lambda t}}{-\lambda} \right) \Big|_0^x = 1 - e^{-\lambda x}$$

Por lo tanto

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{caso contrario} \end{cases}$$

Su gráfica es



Notación: $X \sim \text{Exp}(\lambda)$

Ejemplo:

Supongamos que el tiempo, en segundos, de respuesta en cierta terminal de computadora en línea (es decir el tiempo transcurrido entre el fin de la consulta del usuario y el principio de la respuesta del sistema a esa consulta) se puede considerar como una variable aleatoria con distribución exponencial con parámetro $\lambda = 0.2$. Calcular

- la probabilidad de que el tiempo de respuesta sea a lo sumo 10 segundos.
- la probabilidad de que el tiempo de respuesta esté entre 5 y 10 segundos.

Solución:

Sea X la v.a., entonces $X \sim \text{Exp}(0.2)$

- Se pide calcular $P(X \leq 10)$

Por lo tanto

$$P(X \leq 10) = F(10) = 1 - e^{-0.2 \times 10} = 1 - 0.135 = 0.865$$

- $P(5 \leq X \leq 10) = F(10) - F(5) = (1 - e^{-0.2 \times 10}) - (1 - e^{-0.2 \times 5}) = 0.233$

Esperanza y varianza de la distribución exponencial

Sea $X \sim \text{Exp}(\lambda)$ entonces $E(X) = \frac{1}{\lambda}$ y $V(X) = \frac{1}{\lambda^2}$

Dem.)

Es útil calcular en general $E(X^k)$ el *momento de orden k con respecto al origen* siendo k un número natural

$$E(X^k) = \int_0^{+\infty} x^k f(x) dx = \lambda \int_0^{+\infty} x^k e^{-\lambda x} dx \quad (k = 0, 1, 2, \dots).$$

Integrando por partes:

$$\begin{cases} u = x^k \\ dv = \lambda e^{-\lambda x} dx \end{cases} \quad \begin{cases} du = kx^{k-1} dx \\ v = -e^{-\lambda x} \end{cases} \rightarrow \mu_k = -x^k e^{-\lambda x} \Big|_0^{+\infty} + k \int_0^{+\infty} x^{k-1} e^{-\lambda x} dx. \text{ El corchete de Barrow se}$$

anula y queda $E(X^k) = k\lambda \frac{1}{\lambda} \int_0^{+\infty} x^{k-1} e^{-\lambda x} dx$. Aplicando reiteradamente se llega finalmente a

$$E(X^k) = k(k-1)\dots 2 \cdot 1 \left(\frac{1}{\lambda}\right)^k \lambda \int_0^{+\infty} e^{-\lambda x} dx = k! \left(\frac{1}{\lambda}\right)^k$$

Por lo tanto tenemos para la esperanza y la varianza:

$$E(X) = \frac{1}{\lambda}$$

$$V(X) = E(X^2) - \mu^2 = 2\left(\frac{1}{\lambda}\right)^2 - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Propiedades de la distribución exponencial

1- Relación entre la distribución exponencial y el proceso de Poisson.

Sea T el tiempo de espera hasta el siguiente evento en un proceso de Poisson con parámetro λ . Veamos cuál es la F.d.a. de T .

Si $t < 0$ entonces claramente $F(t) = P(T \leq t) = 0$

Si $t \geq 0$ entonces para hallar $F(t) = P(T \leq t)$ consideramos el evento complementario de $\{T \leq t\}$.

Notar que $\{T > t\}$ si y solo si no ocurre ningún evento durante las siguientes t unidades de tiempo.

Si X : “número de eventos que ocurren en las siguientes t unidades de tiempo”, entonces

$\{T > t\}$ ocurre si y solo si $\{X = 0\}$ ocurre, por lo tanto $P(T > t) = P(X = 0)$

Como $X \sim P(\lambda t)$ entonces

$$P(T > t) = P(X = 0) = e^{-\lambda t} \frac{(\lambda t)^0}{0!} = e^{-\lambda t}$$

Por lo tanto

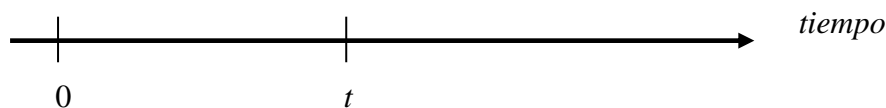
$$F(t) = P(T \leq t) = 1 - P(T > t) = 1 - e^{-\lambda t}$$

Como $F(t)$ es la F.d.a. de una v.a. exponencial, entonces $T \sim \text{Exp}(\lambda)$

Por lo tanto

Si los eventos siguen un proceso de Poisson con parámetro λ , y si T representa el tiempo de espera desde cualquier punto inicial hasta el próximo evento, entonces $T \sim \text{Exp}(\lambda)$

Consideremos un aplicación hidrológica de estas ideas



i) $X = \text{N}^\circ$ de inundaciones en un período $[0, t] \rightarrow X \sim P(\lambda t)$

- ii) $\lambda = \text{N}^\circ \text{ medio de inundaciones por unidad de tiempo} \rightarrow \lambda = \frac{E(X)}{t}$
- iii) $T = \text{Tiempo transcurrido entre inundaciones} \rightarrow T \sim \text{Exp}(\lambda)$
- iv) $E(T) = \text{Tiempo medio de retorno de las inundaciones} \rightarrow E(T) = \frac{1}{\lambda}$.

2- Propiedad falta de memoria

La distribución exponencial tiene una propiedad conocida como falta de memoria, que se muestra en el siguiente ejemplo:

El tiempo de vida, en años, de un circuito integrado particular tiene una distribución exponencial con parámetro 0.5. Encuentre la probabilidad de que el circuito dure más de tres años

Sea la v.a. X : “tiempo de vida, en años, de un circuito integrado particular”, entonces $X \sim \text{Exp}(0.5)$

$$Y \ P(X > 3) = 1 - F(3) = e^{-1.5} = 0.223$$

Supongamos ahora que actualmente un circuito tiene cuatro años y aún funciona. Se quiere hallar la probabilidad de que funcione tres años más.

Por lo tanto planteamos una probabilidad condicional

$$P(X > 7 / X > 4) = \frac{P(X > 7 \text{ y } X > 4)}{P(X > 4)} = \frac{P(X > 7)}{P(X > 4)} = \frac{e^{-0.5 \times 7}}{e^{-0.5 \times 4}} = e^{-0.5 \times (7-4)} = e^{-1.5} = 0.223$$

Observamos que $P(X > 7 / X > 4) = P(X > 7 - 4) = P(X > 3)$

En general, la probabilidad que se tenga que esperar t unidades adicionales, dado que ya se han esperado s unidades, es la misma que la probabilidad de que se tenga que esperar t unidades desde el inicio.

La distribución exponencial no “recuerda” cuánto tiempo se ha esperado.

En particular, si el tiempo de vida de un componente sigue una distribución exponencial, entonces la probabilidad de que un componente que tiene s unidades de tiempo dure t unidades de tiempo adicionales es la misma que la probabilidad de que un componente nuevo dure t unidades de tiempo. En otras palabras, un componente cuyo tiempo de vida siga una distribución exponencial no muestra ningún síntoma de los años o del uso.

Los cálculos hechos en el ejemplo anterior se pueden repetir para valores cualesquiera s y t y entonces se puede probar que

$$\text{si } X \sim \text{Exp}(\lambda) \text{ y } t \text{ y } s \text{ son números positivos, entonces } P(X > t + s / X > s) = P(X > t)$$

Práctica**Distribución uniforme. Distribución normal. Distribución exponencial**

- 1) Supongamos que tomo un colectivo para ir al trabajo, y que cada 5 minutos llega un colectivo a la parada de la esquina. Debido a las variaciones en la hora en que salgo de casa, no siempre llego al mismo tiempo a la parada de la esquina, así que mi tiempo de espera en minutos X al siguiente colectivo es una v.a. continua uniforme en el intervalo $(0,5)$.
 - a) ¿Cuál es la probabilidad de que yo espere entre 1 y 3 minutos?
 - b) ¿Cuál es la probabilidad de que yo espere por lo menos 4 minutos?

- 2) El espesor del borde de un componente de una aeronave está distribuido de manera uniforme entre 0.95 y 1.05 milímetros.
 - a) Obtenga la F.d.a. del espesor del borde.
 - b) Calcule la proporción de bordes cuyo espesor es mayor que 1.02 mm.
 - c) ¿Qué espesor está excedido por el 90% de los bordes?
 - d) Calcule la media y la varianza del espesor del borde.

- 3) Sea Z una v.a. normal estándar, (es decir $Z \sim N(0,1)$). Calcular las siguientes probabilidades:

a) $P(0 \leq Z \leq 2.17)$	e) $P(-1.50 \leq Z \leq 2.00)$
b) $P(-2.50 \leq Z \leq 0)$	f) $P(1.50 \leq Z)$
c) $P(-2.50 \leq Z \leq 2.50)$	g) $P(Z \leq 2.50)$
d) $P(Z \leq 1.37)$	h) $P(-1.75 \leq Z)$

- 4) Sea $Z \sim N(0,1)$, determinar el valor de la constante c para que:

a) $\Phi(c) = 0.9838$	d) $P(Z \geq c) = 0.016$
b) $P(Z \geq c) = 0.121$	e) $P(0 \leq Z \leq c) = 0.291$
c) $P(-c \leq Z \leq c) = 0.668$	

- 5) Si $X \sim N(80,10)$, calcular:

a) $P(X \leq 100)$	d) $P(X \leq 80)$
b) $P(65 \leq X \leq 100)$	e) $P(X \geq 70)$
c) $P(85 \leq X \leq 95)$	f) $P(X - 80 \leq 10)$

- 6) Un tipo particular de tanque de gasolina para un automóvil compacto está diseñado para contener 15 galones. Suponga que la capacidad real de un tanque escogido al azar de este tipo esté normalmente distribuido con media de 15 galones y desviación estándar de 0.2 galones.
 - a) ¿Cuál es la probabilidad de que un tanque seleccionado al azar contenga a lo sumo 14.8 galones?
 - b) ¿Cuál es la probabilidad de que un tanque seleccionado al azar contenga entre 14.7 y 15.1 galones?

- c) Si el automóvil en el que se instala un tanque seleccionado al azar recorre exactamente 25 millas por galón, ¿cuál es la probabilidad de que el automóvil pueda recorrer 370 millas sin reabastecerse?.
- 7) Hay dos máquinas para cortar corchos destinados para usarse en botellas de vino. La primera produce corchos con diámetros que están normalmente distribuidos con media de 3 cm y desviación estándar de 0.1 cm. La segunda máquina produce corchos con diámetros que tienen una distribución normal con media de 3.04 cm y desviación estándar de 0.02 cm. Los corchos aceptables tienen diámetros entre 2.9 cm y 3.1 cm. ¿Cuál máquina tiene más probabilidad de producir un corcho aceptable?
- 8) El dispositivo automático de apertura de un paracaídas militar de carga se ha diseñado para abrir el paracaídas cuando éste se encuentre a 200 m de altura sobre el suelo. Supongamos que la altitud de apertura en realidad tiene una distribución normal con valor medio de 200 m y desviación estándar de 30 m. Habrá un daño al equipo si el paracaídas se abre a una altitud de menos de 100 m. ¿Cuál es la probabilidad de que haya daño a la carga en al menos uno de cinco paracaídas lanzado independientemente?
- 9) La distribución del peso de paquetes enviados de cierto modo es normal con valor medio de 10 libras y desviación estándar de 2 libras. El servicio de paquetería desea establecer un valor de peso c , mas allá del cual habrá cargo extra. ¿Cuál valor de c es tal que 99% de todos los paquetes pesen por lo menos 1 libra abajo del peso con cargo extra?.
- 10) Suponga que X tiene una distribución exponencial con media 10. Calcular:
 a) $P(X > 10)$ b) $P(X > 20)$ c) x , tal que $P(X < x) = 0.95$
- 11) El número de arribos de taxis a un cruce muy concurrido tiene una distribución Poisson con tasa de 6 arribos por hora.
 a) ¿Cuál es la probabilidad de que una persona que esté en el cruce tenga que esperar más de una hora para tomar un taxi?
 b) Suponga que la persona ya esperó una hora; ¿cuál es la probabilidad de que llegue uno en los siguientes 10 minutos?
 c) Determine x , de modo tal que la probabilidad de que la persona espere más de x minutos para tomar un taxi sea 0.01.
- 12) El número de mensajes electrónicos que llegan a una computadora tiene una distribución Poisson con tasa de 0.5 mensajes por hora.
 a) ¿Cuál es la probabilidad de que la computadora no reciba mensajes en un periodo de 2 horas?
 b) Si la computadora no ha recibido ningún mensaje en las últimas 4 horas, ¿cuál es la probabilidad de recibir un mensaje en las 2 horas siguientes?
 c) ¿Cuál es el tiempo esperado entre el quinto y el sexto mensaje?

4- VARIABLES ALEATORIAS BIDIMENSIONALES

4.1 – Generalidades

Hasta ahora hemos considerado el caso de variables aleatorias unidimensionales. Esto es, el resultado del experimento de interés se registra como un único número real.

En muchos casos, sin embargo, nos puede interesar asociar a cada resultado de un experimento aleatorio, dos o más características numéricas. Por ejemplo, de los remaches que salen de una línea de producción nos puede interesar el diámetro X y la longitud Y . Teniendo en cuenta la inevitable variabilidad en las dimensiones de los remaches debido a las numerosas causas presentes en el proceso de fabricación, los podemos representar asociándoles dos variables aleatorias X e Y que pueden pensarse como una **variable aleatoria bidimensional**: (X, Y) .

Sea ε un experimento aleatorio y S un espacio muestral asociado a él. Sean $X: S \rightarrow R, Y: S \rightarrow R$, que a cada resultado $s \in S$ le asignan el par de números reales (x, y)

Llamaremos a (X, Y) **variable aleatoria bidimensional**.

Si en lugar de dos variables aleatorias, tenemos n variables aleatorias X_1, X_2, \dots, X_n , llamaremos a (X_1, X_2, \dots, X_n) **variable aleatoria n -dimensional**

En lo que sigue nos referiremos en particular a variables aleatorias n -dimensionales con $n=2$, es decir nos concentraremos en **variables aleatorias bidimensionales** por cuanto son las más simples de describir, fundamentalmente en relación a la notación. Pero debemos tener presente que las propiedades que estudiemos para ellas se pueden extender sin demasiada dificultad al caso general.

Al conjunto de valores que toma la variable aleatoria bidimensional (X, Y) lo llamaremos **recorrido** de la v.a. (X, Y) y lo indicaremos R_{XY} . En otras palabras $R_{XY} = \{(x, y): x = X(s) \text{ e } y = Y(s) \text{ con } s \in S\}$, es decir, es la imagen por (X, Y) del espacio muestral S .

Notar que el recorrido de (X, Y) es un subconjunto del espacio Euclidiano: $R_{XY} \subseteq R^2$. Como antes, puede considerarse al recorrido R_{XY} como un espacio muestral cuyos elementos son ahora pares de números reales.

Como con cualquier espacio muestral, según el número de elementos que lo constituyen, podemos clasificar a los recorridos R_{XY} en numerables (finitos o infinitos) y no-numerables.

Los recorridos numerables son, en general, de la forma

$$R_{XY} = \{(x_i, y_j) \text{ con } i = 1, 2, \dots, n \text{ y } j = 1, 2, \dots, m\} = \{(x_1, y_1), (x_1, y_2), \dots, (x_n, y_m)\} \quad (\text{finito})$$

$$R_{XY} = \{(x_i, y_j) \text{ con } i = 1, 2, \dots \text{ y } j = 1, 2, \dots\} = \{(x_1, y_1), (x_1, y_2), \dots\} \quad (\text{infinito numerable})$$

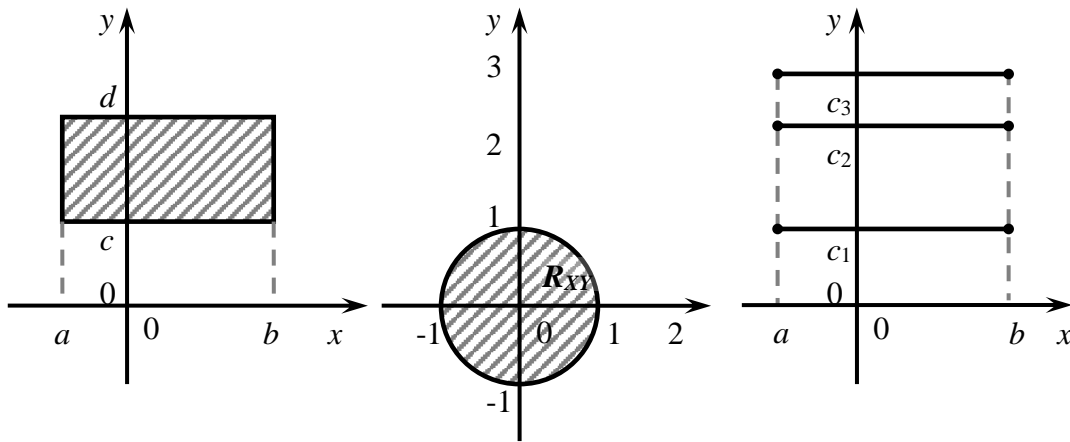
Los recorridos no numerables son regiones o subconjuntos no numerables del plano Euclidiano. Por ejemplo:

$$R_{XY} = \{(x, y): a \leq x \leq b; \quad c \leq y \leq d\} \quad (\text{no numerable})$$

$$R_{XY} = \{(x, y): x^2 + y^2 \leq 1\} \quad (\text{no numerable})$$

$$R_{XY} = \{(x, y_j): a \leq x \leq b, \quad y_j = c_1, c_2, c_3\} \quad (\text{no numerable "mixto"})$$

cuyas gráficas se pueden apreciar en la figura siguiente. Notar en el último recorrido, X es v.a. continua e Y discreta.



Clasificaremos a las variables aleatorias bidimensionales de la siguiente manera:

(X, Y) es v.a. **bidimensional discreta** si X e Y son discretas

(X, Y) es v.a. **bidimensional continua** si X e Y son continuas

El caso X continua, Y discreta (o viceversa) no lo consideramos.

4.2-Variables aleatorias bidimensionales discretas

Sea (X, Y) una variable aleatoria bidimensional discreta y sea R_{XY} su recorrido (numerable). Sea $p: R_{XY} \rightarrow R$ una función que a cada elemento (x_i, y_j) le asigna un número real $p(x_i, y_j)$ tal que $P(X = x_i, Y = y_j) = p(x_i, y_j) \quad \forall (x_i, y_j) \in R_{XY}$ y que verifica.

- a) $p(x_i, y_j) \geq 0 \quad \forall (x_i, y_j) \in R_{XY}$
- b) $\sum_{(x_i, y_j) \in R_{XY}} p(x_i, y_j) = \sum_i \sum_j p(x_i, y_j) = 1$

A esta función la llamaremos **función de probabilidad puntual conjunta** de la variable aleatoria bidimensional (X, Y) . En forma abreviada la designaremos **fdp conjunta**.

Ejemplos:

1-Dos líneas de producción, señaladas I y II, manufacturan cierto tipo de artículo a pequeña escala. Supóngase que la capacidad máxima de producción de la línea I es cinco artículos por día, mientras que para la línea II es 3 artículos/día. Debido a los innumerables factores presentes en todo proceso de producción, el número de artículos realmente producido por cada línea puede pensarse como una variable aleatoria. En conjunto podemos pensar en una variable aleatoria bidimensional (X, Y) discreta, donde la primera componente X corresponde a la producción de la línea I y la segunda componente Y a los artículos que salen de la línea II. La fdp conjunta correspondiente a variables aleatorias bidimensionales suele presentarse, por comodidad, como una tabla. Supongamos que la para la v.a. (X, Y) que nos interesa aquí la tabla correspondiente a $p(x_i, y_j)$ es

Y/X	0	1	2	3	4	5
0	0	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

¿Cuál es la probabilidad de que salgan más artículos de la línea I que de la línea II?

Antes de calcular la probabilidad que nos pide el problema, hagamos algunas consideraciones sobre la tabla que representa a $p(x_i, y_j)$.

Se trata de una tabla a doble entrada donde en la primera fila se indican los valores que puede tomar la v.a. X (en este caso $X=0,1,2,3,4,5$) y la primera columna indica los valores que puede tomar la variable Y ($0,1,2,3$). Para determinar el valor de la $p(x_i, y_j)$ cuando la v.a. (X, Y) toma el valor (x_i, y_j) consideramos el número que se encuentra en la columna correspondiente a $X = x_i$ y la fila correspondiente a $Y = y_j$. Por ejemplo: $p(4,2) = P(X = 4, Y = 2) = 0.05$.

Podemos verificar fácilmente que la *fdp* conjunta definida por esta bien definida. En efecto verifica las condiciones a) $p(x_i, y_j) \geq 0 \quad \forall (x_i, y_j) \in R_{XY}$ y b) $\sum_{(x_i, y_j) \in R_{XY}} p(x_i, y_j) = 1$.

Para contestar la pregunta del enunciado, consideremos el suceso $B \subset R_{XY}$ definido

B : “es el suceso que ocurre cuando la línea I produce más artículos que la línea II” o, $B = \{X > Y\}$. Luego:

$$P(B) = P(X > Y) = \sum_{y_j=0}^3 \sum_{x_i > y_j} p(x_i, y_j) = 0.01 + 0.03 + 0.05 + 0.07 + 0.09 + 0.04 + 0.05 + 0.06 + 0.08 + 0.05 + 0.05 + 0.06 + 0.06 + 0.05 = 0.75.$$

2- Hay tres cajas registradoras a la salida de un supermercado. Dos clientes llegan a las cajas en diferentes momentos cuando no hay otros clientes ante aquellas. Cada cliente escoge una caja al azar e independientemente del otro.

Sean las variables aleatorias X : “nº de clientes que escogen la caja 1” e Y : “nº de clientes que escogen la caja 2”. Hallar la *fdp* conjunta de (X, Y)

Podemos suponer que el espacio muestral original S es el conjunto de pares ordenados $S = \{(1,1); (1,2); (1,3); (2,1); (2,2); (2,3); (3,1); (3,2); (3,3)\}$ donde la primera componente del par indica la caja elegida por el cliente 1 y la segunda componente del par indica la caja elegida por el cliente 2.

Además notar que X como Y pueden tomar los valores 0, 1, 2

El punto muestral (3,3) es el único punto muestral que corresponde al evento $\{X = 0, Y = 0\}$

Entonces

$P(X = 0, Y = 0) = \frac{1}{9}$; pensando de forma análoga los otros casos:

$$P(X = 1, Y = 0) = \frac{2}{9}; \quad P(X = 2, Y = 0) = \frac{1}{9}; \quad P(X = 0, Y = 1) = \frac{2}{9}; \quad P(X = 1, Y = 1) = \frac{2}{9},$$

$$P(X = 0, Y = 2) = \frac{1}{9}; \quad P(X = 1, Y = 2) = P(X = 2, Y = 2) = 0$$

Disponemos estas probabilidades en una tabla de la siguiente forma

$Y \setminus X$	0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

4.2.1 - Funciones de distribución marginales de una v.a. (X, Y) discreta

En el ejemplo 1, supongamos que queremos saber cuál es la probabilidad de que el número de artículos producidos por la línea I sea 2, o sea $P(X = 2)$

Como el evento $\{X = 2\}$ es igual a $\{X = 2\} \cap (\{Y = 0\} \cup \{Y = 1\} \cup \{Y = 2\} \cup \{Y = 3\})$, y a su vez $\{X = 2\} \cap (\{Y = 0\} \cup \{Y = 1\} \cup \{Y = 2\} \cup \{Y = 3\}) = (\{X = 2\} \cap \{Y = 0\}) \cup (\{X = 2\} \cap \{Y = 1\}) \cup (\{X = 2\} \cap \{Y = 2\}) \cup (\{X = 2\} \cap \{Y = 3\})$
Entonces

$$\begin{aligned} P(X = 2) &= P(\{X = 2\} \cap \{Y = 0\}) + P(\{X = 2\} \cap \{Y = 1\}) + P(\{X = 2\} \cap \{Y = 2\}) \\ &\quad + P(\{X = 2\} \cap \{Y = 3\}) \\ &= P(X = 2, Y = 0) + P(X = 2, Y = 1) + P(X = 2, Y = 2) + P(X = 2, Y = 3) \\ &= \sum_{j=0}^3 P(X = 2, Y = j) \end{aligned}$$

Razonando de la misma forma podemos escribir

$$P(X = i) = \sum_{j=0}^3 P(X = i, Y = j) \quad i = 0, 1, \dots, 5$$

Es decir obtenemos la **función de distribución de probabilidad de X**

Análogamente obtenemos

$$P(Y = j) = \sum_{i=0}^5 P(X = i, Y = j) \quad j = 0, 1, 2, 3$$

Que es la **función de distribución de probabilidad de Y**

En general se las denomina **distribuciones marginales de X e Y**, y su definición sería la siguiente

Sea (X, Y) discreta y sea $p(x_i, y_j)$ ($i=1, 2, \dots, n, j=1, 2, \dots, m$) su función de probabilidad conjunta (Eventualmente n y/o m pueden ser ∞).

La función de probabilidad marginal de X es

$$p(x_i) = P(X = x_i) = \sum_{j=1}^m p(x_i, y_j) \quad (i=1, 2, \dots, n)$$

La función de probabilidad marginal de Y es

$$q(y_j) = P(Y = y_j) = \sum_{i=1}^n p(x_i, y_j) \quad (j=1, 2, \dots, m)$$

Observación: Remarcamos que la función de probabilidad marginal de X, es decir $p(x_i)$ calculada a partir de $p(x_i, y_j)$ en la forma indicada, coincide con la función de probabilidad de la variable aleatoria unidimensional X considerada en forma aislada. Análogamente la función de probabilidad marginal de Y, es decir $q(y_j)$ calculada a partir de $p(x_i, y_j)$ en la forma indicada, coincide con la función de probabilidad de variable aleatoria unidimensional Y considerada en forma aislada.

Ejemplo:

Siguiendo con el ejemplo 1,

$$\begin{aligned} p(5) &= P(X = 5) = p(5, 0) + p(5, 1) + p(5, 2) + p(5, 3) = 0.09 + 0.08 + 0.06 + 0.05 \\ &= 0.28 \end{aligned}$$

$$\begin{aligned}
 q(1) &= P(Y = 1) = p(0,1) + p(1,1) + p(2,1) + p(3,1) + p(4,1) + p(5,1) \\
 &= 0.01 + 0.02 + 0.04 + 0.05 + 0.06 \\
 &= 0.26
 \end{aligned}$$

Observemos que se verifica la condición de normalización para cada una de las marginales:

$$\begin{aligned}
 \sum_{x_i=0}^5 p(x_i) &= 0.03 + 0.08 + 0.16 + 0.21 + 0.24 + 0.28 = 1 \\
 \sum_{y_j=0}^3 q(y_j) &= 0.25 + 0.26 + 0.25 + 0.24 = 1
 \end{aligned}$$

4.2.2 - Funciones de probabilidades condicionales

Consideremos nuevamente el ejemplo de las dos líneas I y II que producen cierto artículo a pequeña escala. Definimos la v.a. (X, Y) cuya función de probabilidad conjunta $p(x_i, y_j)$ está dada por la tabla anterior que repetimos

Y/X	0	1	2	3	4	5	$q(y_j)$
0	0	0.01	0.03	0.05	0.07	0.09	0.25
1	0.01	0.02	0.04	0.05	0.06	0.08	0.26
2	0.01	0.03	0.05	0.05	0.05	0.06	0.25
3	0.01	0.02	0.04	0.06	0.06	0.05	0.24
$p(x_i)$	0.03	0.08	0.16	0.21	0.24	0.28	1

Supongamos que deseamos conocer la probabilidad de que la línea I produzca tres artículos sabiendo que la línea II ha fabricado dos. Tenemos que calcular una probabilidad condicional. Entonces

$$P(X = 3|Y = 2) = \frac{P(X = 3, Y = 2)}{P(Y = 2)} = \frac{p(3,2)}{q(2)} = \frac{0.05}{0.25} = 0.2$$

En general definimos la **función de probabilidad puntual de X condicional a Y** como sigue:

$p(x_i|y_j) = P(X = x_i|Y = y_j) = \frac{p(x_i, y_j)}{q(y_j)}$, es decir como el cociente de la función de probabilidad conjunta de (X, Y) y la función de probabilidad puntual marginal de Y .

Análogamente, definimos la **función de probabilidad puntual de Y condicional a X** :

$q(y_j|x_i) = P(Y = y_j|X = x_i) = \frac{p(x_i, y_j)}{p(x_i)}$, es decir como el cociente de la función de probabilidad puntual conjunta de (X, Y) y la función de probabilidad puntual marginal de X .

4.2.3 – Variables aleatorias discretas independientes

Ya se discutió el concepto de independencia entre dos eventos A y B . Esas mismas ideas podemos trasladarlas en relación a dos variables aleatorias X e Y que, eventualmente, podemos considerarlas como las componentes de una variable aleatoria bidimensional (X, Y) .

De acuerdo con esto, intuitivamente decimos que dos variables, X e Y , son independientes si el valor que toma una de ellas no influye de ninguna manera sobre el valor que toma la otra. Esto lo establecemos más formalmente:

Sea (X, Y) una variable aleatoria bidimensional discreta. Sea $p(x_i, y_j)$ su *fdp* conjunta y $p(x_i)$ y $q(y_j)$ las correspondientes *fdp* marginales de X e Y . Decimos que X e Y son **variables aleatorias independientes** si y sólo si

$$p(x_i, y_j) = p(x_i)q(y_j) \quad \forall (x_i, y_j) \in R_{XY}$$

Observación: Notar que para poder afirmar la independencia de X e Y debe cumplirse la factorización de la *fdp* conjunta como producto de las *fdp* marginales para **todos** los pares de valores de la v.a. (X, Y) . Por lo tanto, para verificar la independencia es necesario demostrar la validez de la factorización para todos los pares. En cambio, es suficiente encontrar un solo par que no la verifica, para afirmar, de acuerdo con la definición, que las variables X e Y son no independientes, es decir, que son dependientes. Esto es, para demostrar la dependencia es suficiente con encontrar un solo par que no verifique la factorización señalada.

Vimos que dos sucesos A y B son independientes si y sólo si $P(A|B) = P(A)$ y $P(B|A) = P(B)$ (donde por supuesto debía ser $P(A) \neq 0$ y $P(B) \neq 0$). En términos de variables aleatorias, esta forma de ver la independencia se manifiesta en la igualdad entre las *fdp* condicionales y las correspondientes *fdp* marginales, como demostramos en este

Teorema

a) Sea (X, Y) una variable aleatoria bidimensional discreta cuyas *fdp* conjunta, condicionales y marginales son, respectivamente, $p(x_i, y_j)$; $p(x_i|y_j)$, $q(y_j|x_i)$ y $p(x_i)$, $q(y_j)$.

Entonces, X e Y son variables aleatorias independientes si y sólo si

$$a_1) p(x_i|y_j) = p(x_i) \quad \forall (x_i, y_j) \in R_{XY}, \text{ o}$$

$$a_2) q(y_j|x_i) = q(y_j) \quad \forall (x_i, y_j) \in R_{XY}, \text{ que es equivalente a lo anterior}$$

Ejemplo:

Supongamos que una máquina se usa para un trabajo específico a la mañana y para uno diferente en la tarde. Representemos por X e Y el número de veces que la máquina falla en la mañana y en la tarde respectivamente. Supongamos que la tabla siguiente da la función de probabilidad conjunta $p(x_i, y_j)$ de la variable aleatoria bidimensional discreta (X, Y) .

Y/X	0	1	2	$q(y_j)$
0	0.1	0.2	0.2	0.5
1	0.04	0.08	0.08	0.2
2	0.06	0.12	0.12	0.3
$P(x_i)$	0.2	0.4	0.4	1

Deseamos saber si las variables aleatorias X e Y son independientes o dependientes.

Para demostrar que son independientes debemos probar que se verifica $\forall (x_i, y_j) \in R_{XY}$

$p(x_i, y_j) = p(x_i)q(y_j)$ Verificamos directamente que

$$\begin{aligned} p(0,0) &= 0.1 = p(0)q(0) = 0.2 \times 0.5 \\ p(0,1) &= 0.04 = p(0)q(1) = 0.2 \times 0.2 \\ p(0,2) &= 0.06 = p(0)q(2) = 0.2 \times 0.3 \\ p(1,0) &= 0.04 = p(1)q(0) = 0.4 \times 0.5 \\ p(1,1) &= 0.08 = p(1)q(1) = 0.4 \times 0.2 \end{aligned}$$

$$\begin{aligned}
 p(1,2) &= 0.12 = p(1)q(2) = 0.4 \times 0.3 \\
 p(2,0) &= 0.2 = p(2)q(0) = 0.4 \times 0.5 \\
 p(2,1) &= 0.08 = p(2)q(1) = 0.4 \times 0.2 \\
 p(2,2) &= 0.12 = p(2)q(2) = 0.4 \times 0.3
 \end{aligned}$$

Luego X e Y son independientes.

Podríamos haber usado las condiciones

a) $p(x_i|y_j) = p(x_i) \quad \forall (x_i, y_j) \in R_{XY}$, o su equivalente

b) $q(y_j|x_i) = q(y_j) \quad \forall (x_i, y_j) \in R_{XY}$. Veamos, como muestra para un solo valor, que se verifica

$p(2|1) = \frac{p(2,1)}{q(1)} = \frac{0.08}{0.2} = 0.4 = p(2)$. Para demostrar la independencia por este camino habría que demostrar que se cumple la condición para el resto de los pares de valores. Se deja este cálculo como ejercicio optativo.

Dem.)

Demostraremos solamente a). La equivalencia entre a) y b) la dejamos como ejercicio.

Para demostrar a) verificaremos la doble equivalencia entre ésta y la definición de v.a. independientes.

\Rightarrow)

Sean X e Y variables aleatorias independientes. Entonces $\forall (x_i, y_j) \in R_{XY}$

$$p(x_i|y_j) = \frac{p(x_i, y_j)}{q(y_j)} = \frac{p(x_i)q(y_j)}{q(y_j)} = p(x_i)$$

Aquí la primera igualdad es la definición de *fdp* condicional y la segunda sale de la definición de independencia al suponer que X e Y son independientes.

\Leftarrow)

Supongamos que se verifica a1). Entonces $\forall (x_i, y_j) \in R_{XY}$

$$p(x_i|y_j) = p(x_i) \rightarrow \frac{p(x_i, y_j)}{q(y_j)} = p(x_i) \rightarrow p(x_i, y_j) = p(x_i)q(y_j) \rightarrow X \text{ e } Y \text{ independientes}$$

Aquí, la primera implicación se debe a la definición de *fdp* condicional y la tercera a la definición de v.a. independientes.

4.3 - Variables aleatorias bidimensionales continuas

Sea (X, Y) una variable aleatoria bidimensional continua y sea R_{XY} su recorrido (no numerable). Sea $f: R_{XY} \rightarrow R$ una función que, a cada punto (x, y) de R_{XY} le asigna un número real $f(x, y)$ tal que $P(B) = \iint_B f(x, y) dx dy \quad \forall B \subseteq R$ y que verifica.

$$a) f(x, y) \geq 0 \quad \forall (x, y) \in R^2$$

$$b) \iint_{R^2} f(x, y) dx dy = 1.$$

A esta función la llamaremos **función de densidad de probabilidad conjunta** de la variable aleatoria bidimensional (X, Y) . En forma abreviada la designaremos también *fdp conjunta*.

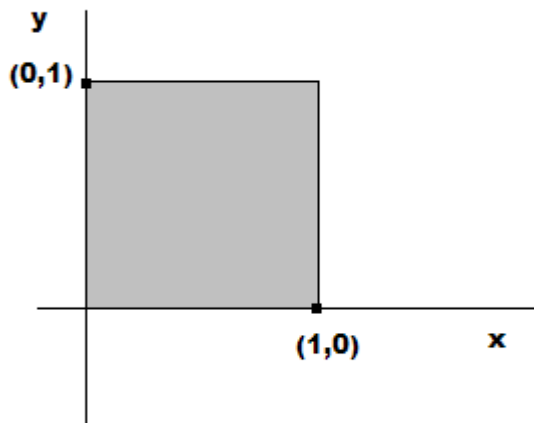
Ejemplo:

Supongamos que una partícula radiactiva se localiza aleatoriamente en un cuadrado con lados de longitud unitaria. Es decir, si se consideran dos regiones de la misma área, la partícula tendrá igual probabilidad de estar en cualquiera de ellas. Sean X e Y las coordenadas que localizan la partícula. Un

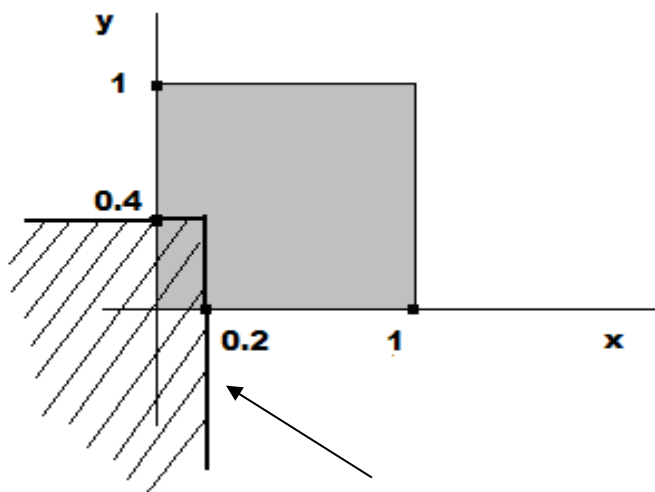
modelo adecuado para la distribución conjunta de X e Y sería considerar a (X, Y) continua con f_{dp} dada por

$$f(x, y) = \begin{cases} 1 & \text{si } 0 \leq x \leq 1; 0 \leq y \leq 1 \\ 0 & \text{caso contrario} \end{cases}$$

Es conveniente hacer un gráfico en el plano del dominio de la f_{dp}



Nos preguntamos cuál es la probabilidad de que la coordenada en x sea menor que 0.2 y la coordenada en y menor que 0.4, es decir, cuál es la $P(X \leq 0.2, Y \leq 0.4)$. Para calcularla, graficamos en el plano xy la **región que corresponde al evento intersección la región que es dominio de la f_{dp}**



$$\{X \leq 0.2, Y \leq 0.4\}$$

Por lo tanto

$$P(X \leq 0.2, Y \leq 0.4) = \int_0^{0.4} \int_0^{0.2} 1 dx dy = 0.2 \times 0.4 = 0.08$$

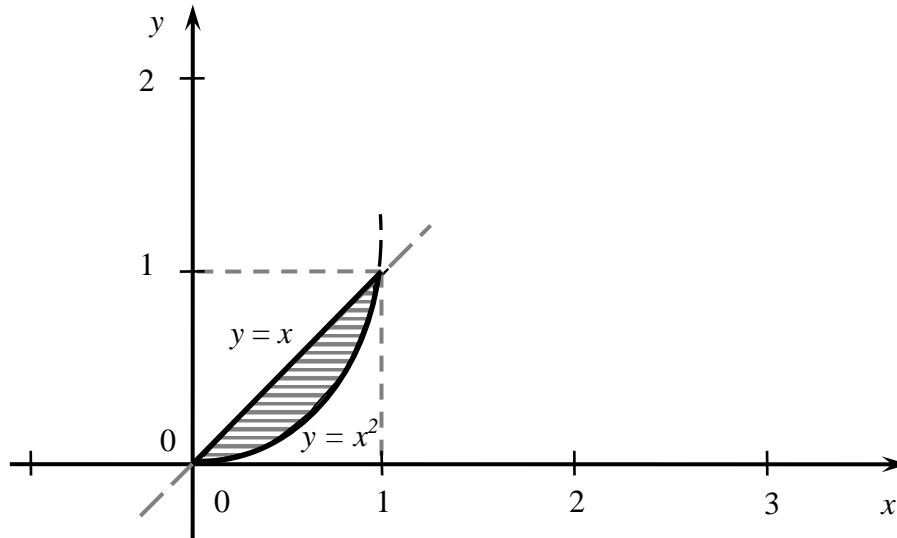
En este ejemplo la f_{dp} es un caso particular de v.a. bidimensional uniformemente distribuida. Diremos que una variable aleatoria bidimensional continua está **uniformemente distribuida** en la región R_{XY} del plano Euclidiano R si su función de densidad de probabilidad es

$$f(x, y) = \begin{cases} c & \text{para } (x, y) \in R_{XY} \\ 0 & \text{para } (x, y) \notin R_{XY} \end{cases}$$

Puesto que la condición de normalización exige $\iint_{R^2} f(x, y) dx dy = \iint_{R_{XY}} c dx dy = 1$ debe ser

$$c = \frac{1}{\text{área}(R_{XY})}$$

A una v.a. (X, Y) bidimensional continua uniformemente distribuida en su recorrido R_{XY} la indicaremos $X \sim U[R_{XY}]$. Por ejemplo, supongamos que la v.a. (X, Y) está distribuida uniformemente en el recorrido R_{XY} que se muestra en la figura. ¿Cuál es su *fdp* conjunta?



De la figura calculamos el área del recorrido:

$$\text{área}(R_{XY}) = \int_0^1 dx \int_{x^2}^x dy = \int_0^1 (x - x^2) dx = \frac{1}{6}. \text{ Por lo tanto}$$

$$f(x, y) = \begin{cases} 6 & \text{para } (x, y) \text{ tal que } 0 \leq x \leq 1, \quad x^2 \leq y \leq x \\ 0 & \text{caso contrario} \end{cases}$$

4.3.1 - Funciones de densidad marginales de una v.a. (X, Y) continua

En el ejemplo 3 supongamos que queremos hallar la probabilidad de que la coordenada x de la partícula sea menor o igual a 0.2, es decir $P(X \leq 0.2)$. Podemos escribir

$$P(X \leq 0.2) = P(X \leq 0.2, -\infty < Y < \infty) = \int_0^{0.2} \int_{-\infty}^{\infty} f(x, y) dy dx = \int_0^{0.2} \int_0^1 1 dy dx = 0.2$$

En general si queremos hallar $P(X \leq x)$ podemos plantear

$$P(X \leq x) = P(X \leq x, -\infty < Y < \infty) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x, y) dy dx = \int_{-\infty}^x g(x) dx$$

donde $\int_{-\infty}^{\infty} f(x, y) dy = g(x)$

Por definición de *fdp* debe ser $g(x)$ la *fdp* de la v.a. X

Análogamente

$$P(Y \leq y) = P(-\infty < X < \infty, Y \leq y) = \int_{-\infty}^y \int_{-\infty}^{\infty} f(x, y) dx dy = \int_{-\infty}^y h(x) dx$$

donde $\int_{-\infty}^{\infty} f(x, y) dy = h(x)$

Por definición de fdp debe ser $h(x)$ la fdp de la v.a. Y

En general:

Sea (X, Y) continua y sea $f(x, y)$ su función de densidad de probabilidad conjunta.

La **función de densidad de probabilidad marginal de X** es:

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

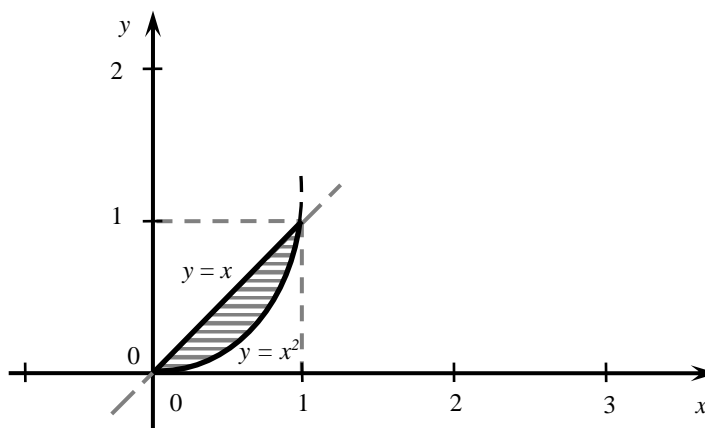
La **función de densidad de probabilidad marginal de Y** es:

$$h(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Observación: Remarcamos aquí también que la función de densidad de probabilidad marginal de X , es decir $g(x)$, calculada a partir de $f(x, y)$ en la forma indicada, coincide con la función de densidad de probabilidad de variable aleatoria unidimensional X considerada en forma aislada. Análogamente la función de densidad de probabilidad marginal de Y , es decir $h(y)$ calculada a partir de $f(x, y)$ en la forma indicada, coincide con la función de densidad de probabilidad de variable aleatoria unidimensional Y considerada en forma aislada. De manera que podemos calcular probabilidades como, por ejemplo

$$P(a \leq X \leq b) = P(a \leq X \leq b, -\infty < Y < \infty) = \int_a^b \int_{-\infty}^{\infty} f(x, y) dy dx = \int_a^b g(x) dx$$

Ejemplo: Consideremos nuevamente la v.a. continua (X, Y) uniformemente distribuida cuyo recorrido R_{XY} dibujamos otra vez



Ya vimos que la fdp está dada por

$$f(x, y) = \begin{cases} 6 & \text{para } (x, y) \text{ tal que } 0 \leq x \leq 1, \quad x^2 \leq y \leq x \\ 0 & \text{caso contrario} \end{cases}$$

Entonces las funciones de densidad de probabilidad de X e Y son

$$g(x) = \begin{cases} \int_{-\infty}^{\infty} f(x, y) dy = \int_{x^2}^x 6 dy = 6(x - x^2) & 0 \leq x \leq 1 \\ 0 & \text{caso contrario} \end{cases}$$

$$h(y) = \begin{cases} \int_{-\infty}^{\infty} f(x, y) dx = \int_y^{\sqrt{y}} 6 dx = 6(\sqrt{y} - y) & 0 \leq y \leq 1 \\ 0 & \text{caso contrario} \end{cases}$$

4.3.2 - Función de densidad condicional

Como ocurre con las variables aleatorias continuas en general, el definir la probabilidad condicional de ocurrencia de un valor dado de una de las variables aleatorias del par (X, Y) supuesto que ocurrió la otra, presenta las dificultades conocidas relacionadas con el hecho de que la probabilidad de un punto es cero. Entonces probabilidades tales como $P(X = x_i | Y = y_j)$ tienen el problema que $P(Y = y_j) = 0$.

Sea (X, Y) una variable aleatoria bidimensional continua cuya *fdp* conjunta es $f(x, y)$. Sean $g(x)$ y $h(y)$ las *fdp* marginales de X e Y , respectivamente. Definimos la **función de densidad de probabilidad de X condicional a que $Y=y$** , a la que denotaremos $g(x|y)$, de la siguiente manera:

$$g(x|y) = \frac{f(x, y)}{h(y)} \quad \text{con } h(y) > 0.$$

Análogamente, la **función de densidad de probabilidad de Y condicional a que $X=x$** , a la que denotaremos $h(y|x)$, se define:

$$h(y|x) = \frac{f(x, y)}{g(x)} \quad \text{con } g(x) > 0.$$

De acuerdo con estas definiciones, podemos calcular, por ejemplo,

$$P(a \leq X \leq b | Y = c) = \int_a^b g(x|c) dx = \frac{\int_a^b f(x, c) dx}{h(c)}$$

Observemos que si quisiéramos calcular esta probabilidad usando la *fdp* conjunta, es decir, refiriéndonos al recorrido completo, llegaríamos a una indeterminación:

$$P(a \leq X \leq b | Y = c) = \frac{P[(a \leq X \leq b) \cap (Y = c)]}{P(Y = c)} = \frac{\int_a^b dx \int_c^c dy f(x, y)}{\int_{-\infty}^{+\infty} dx \int_c^c dy f(x, y)} = \frac{0}{0}.$$

Notar la diferencia entre la función de densidad de probabilidad condicional $g(x|y)$ y la función de densidad de probabilidad marginal $g(x)$:

$$\int_a^b g(x) dx = P(a \leq X \leq b), \text{ mientras que } \int_a^b g(x|y) dx = P(a \leq X \leq b|Y = y).$$

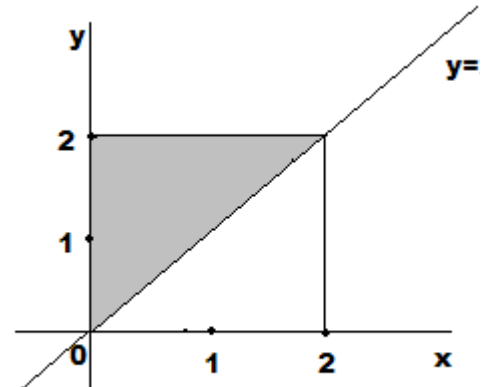
Ejemplo: Una máquina vendedora de refrescos se llena al principio de un día dado con una cantidad aleatoria Y , y se despacha durante el día una cantidad aleatoria X (medida en galones). No se le vuelve a surtir durante el día y entonces $X \leq Y$

Se ha observado que (X, Y) tienen la densidad conjunta

$$f(x, y) = \begin{cases} 1/2 & \text{si } 0 \leq x \leq y; 0 \leq y \leq 2 \\ 0 & \text{caso contrario} \end{cases}$$

¿Cuál es la probabilidad de que se venda menos de $1/2$ galón, dado que la máquina contiene 1 galón al inicio del día?

Solución: Primero es conveniente hacer un gráfico de la región del plano donde la densidad conjunta está definida



Hallamos la densidad condicional de X dado Y . Para esto primero encontramos la *fdp* marginal de Y

$$h(y) = \int_{-\infty}^{\infty} f(x, y) dx = \begin{cases} \int_0^y (1/2) dx & \text{si } 0 \leq y \leq 2 \\ 0 & \text{caso contrario} \end{cases} = \begin{cases} (1/2)y & \text{si } 0 \leq y \leq 2 \\ 0 & \text{caso contrario} \end{cases}$$

Entonces la densidad condicional es

$$h(x|y) = \frac{f(x, y)}{h(y)} = \begin{cases} \frac{1/2}{(1/2)y} & \text{si } 0 < x \leq y \leq 2 \\ 0 & \text{caso contrario} \end{cases} = \begin{cases} \frac{1}{y} & \text{si } 0 < x \leq y \leq 2 \\ 0 & \text{caso contrario} \end{cases}$$

La probabilidad que interesa es

$$P(X \leq 1/2 | Y = 1) = \int_{-\infty}^{1/2} f(x|y = 1) dx = \int_0^{1/2} (1) dx = \frac{1}{2}$$

Notar que si la máquina hubiera contenido 2 galones al principio del día, entonces

$$P(X \leq 1/2 | Y = 2) = \int_{-\infty}^{1/2} f(x|y = 2) dx = \int_0^{1/2} \frac{1}{2} dx = \frac{1}{4}$$

Así la probabilidad condicional que $X \leq 1/2$ depende de la elección de Y

4.3.3 – Variables aleatorias continuas independientes

Ya se discutió el concepto de independencia entre dos eventos A y B . Esas mismas ideas podemos trasladarlas en relación a dos variables aleatorias X e Y que, eventualmente, podemos considerarlas como las componentes de una variable aleatoria bidimensional (X, Y) .

De acuerdo con esto, intuitivamente decimos que dos variables, X e Y , son independientes si el valor que toma una de ellas no influye de ninguna manera sobre el valor que toma la otra. Esto lo establecemos más formalmente:

Sea (X, Y) una variable aleatoria bidimensional continua. Sea $f(x, y)$ su *fdp* conjunta y $g(x)$ y $h(y)$ las correspondientes *fdp* marginales de X e Y . Decimos que X e Y son **variables aleatorias independientes** si y sólo si

$$f(x, y) = g(x)h(y) \quad \forall (x, y) \in \mathbb{R}^2$$

Observación: Notar que para poder afirmar la independencia de X e Y debe cumplirse la factorización de la *fdp* conjunta como producto de las *fdp* marginales para **todos** los pares de valores de la v.a. (X, Y) . Por lo tanto, para verificar la independencia es necesario demostrar la validez de la factorización para todos los pares. En cambio, es suficiente encontrar un solo par que no la verifica, para afirmar, de acuerdo con la definición, que las variables X e Y son no independientes, es decir, que son dependientes. Esto es, para demostrar la dependencia es suficiente con encontrar un solo par que no verifique la factorización señalada.

Sea (X, Y) una variable aleatoria bidimensional continua cuyas *fdp* conjunta, condicionales y marginales son, respectivamente, $f(x, y)$; $g(x|y)$, $h(y|x)$ y $g(x)$, $h(y)$.

Entonces, X e Y son variables aleatorias independientes si y sólo si

$$a) g(x|y) = g(x) \quad \forall (x, y) \in \mathbb{R}^2, \text{ o}$$

$$b) h(y|x) = h(y) \quad \forall (x, y) \in \mathbb{R}^2, \text{ que es equivalente al anterior.}$$

Dem.)

Demostramos sólo a) y dejamos como ejercicio demostrar la equivalencia entre a) y b).

\Rightarrow)

Sean X e Y variables aleatorias independientes. Entonces $\forall (x, y) \in \mathbb{R}^2$

$$g(x|y) = \frac{f(x, y)}{h(y)} = \frac{g(x)h(y)}{h(y)} = g(x)$$

Aquí la primera igualdad es la definición de *fdp* condicional y la segunda sale de la definición de independencia al suponer que X e Y son independientes.

\Leftarrow)

Supongamos que se verifica b1). Entonces $\forall (x, y) \in \mathbb{R}^2$

$$g(x|y) = g(x) \rightarrow \frac{g(x, y)}{h(y)} = g(x) \rightarrow f(x, y) = g(x)h(y) \rightarrow X \text{ e } Y \text{ independientes.}$$

Aquí, la primera implicación se debe a la definición de *fdp* condicional y la tercera a la definición de v.a. independientes.

Ejemplos

1-Sean X e Y v.a. continuas que representan el tiempo de vida de dos dispositivos electrónicos. Supongamos que la *fdp* conjunta de la v.a. continua (X, Y) es:

$$f(x, y) = \begin{cases} e^{-(x+y)} & 0 \leq x < \infty, \quad 0 \leq y < \infty \\ 0 & \text{caso contrario} \end{cases}$$

Deseamos saber si X e Y son variables aleatorias independientes.
Calculamos las *fdp* marginales de X e Y :

$$g(x) = \int_{-\infty}^{+\infty} f(x, y) dy = \int_0^{+\infty} e^{-(x+y)} dy = e^{-x}$$

$$h(y) = \int_{-\infty}^{+\infty} f(x, y) dx = \int_0^{+\infty} e^{-(x+y)} dx = e^{-y}$$

Luego las marginales son:

$$g(x) = \begin{cases} e^{-x} & 0 \leq x < \infty \\ 0 & c. c \end{cases}$$

$$h(y) = \begin{cases} e^{-y} & 0 \leq y < \infty \\ 0 & c. c \end{cases}$$

Vemos que efectivamente

$$f(x, y) = g(x)h(y) = \begin{cases} e^{-(x+y)} = e^{-x}e^{-y} & 0 \leq x < \infty, \quad 0 \leq y < \infty \\ 0 & \text{caso contrario} \end{cases}$$

Es decir X e Y son v.a. independientes.

También podemos verificar la condición b_1):

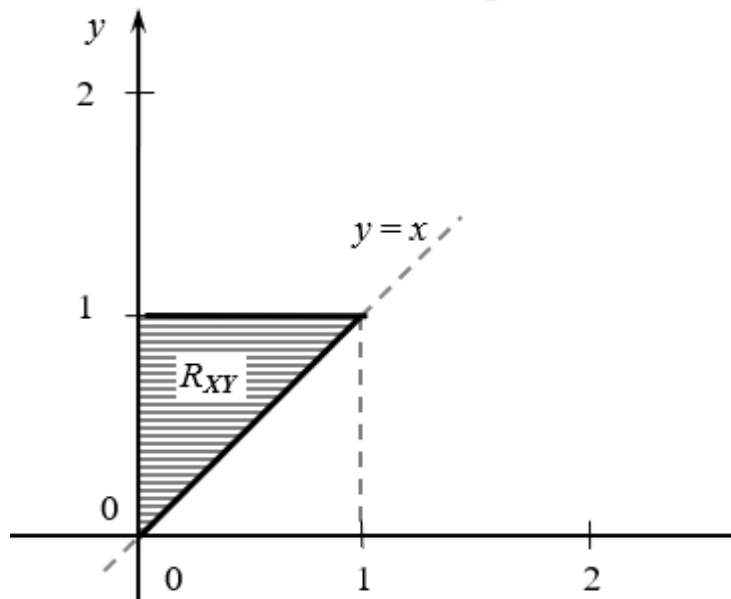
$$g(x|y) = \frac{f(x, y)}{h(y)} = \begin{cases} \frac{e^{-(x+y)}}{e^{-y}} = e^{-x} & 0 \leq x < \infty \\ 0 & \text{caso contrario} \end{cases} = g(x) \quad \forall (x, y) \in \mathbb{R}^2$$

2- Consideremos un ejemplo donde no se verifica la independencia.

Sea una v.a.b cuya *fdp* conjunta es

$$f(x, y) = \begin{cases} 8xy & 0 \leq x \leq y \leq 1 \\ 0 & \text{caso contrario} \end{cases}$$

En la figura siguiente mostramos el recorrido



Calculamos las *fdp* marginales de X e Y :

$$g(x) = \begin{cases} \int_x^1 8xy dy = 4x(1 - x^2) & 0 \leq x \leq 1 \\ 0 & \text{c.c} \end{cases}$$

$$h(y) = \begin{cases} \int_0^y 8xy dx = 4y^3 & 0 \leq y \leq 1 \\ 0 & \text{c.c} \end{cases}$$

y podemos apreciar que para $0 \leq x \leq y \leq 1$ en general es

$$f(x, y) = 8xy \neq 16x(1 - x^2)y^3 = g(x)h(y).$$

Luego X e Y son variables aleatorias dependientes.

Observaciones

1- De la definición de las *fdp* marginales, vemos que tanto en el caso discreto como en el continuo, la *fdp* conjunta determina unívocamente las *fdp* marginales. Es decir, si (X, Y) es discreta del conocimiento de la función de probabilidad conjunta $p(x_i, y_j)$ podemos determinar unívocamente las funciones de probabilidad $p(x_i)$ y $q(y_j)$. Análogamente, si (X, Y) es continua del conocimiento de la función de densidad de probabilidad conjunta $f(x, y)$ podemos determinar unívocamente las funciones de densidad $g(x)$ y $h(y)$. Sin embargo la inversa no se cumple en general. Es decir del conocimiento de $p(x_i)$ y $q(y_j)$ no se puede, en general, reconstruir $p(x_i, y_j)$ a menos que X e Y sean variables independientes en cuyo caso es $p(x_i, y_j) = p(x_i)q(y_j)$ y, análogamente, en el caso continuo, del conocimiento de $g(x)$ y $h(y)$ no se puede construir en general $f(x, y)$ excepto cuando X e Y son independientes en cuyo caso puedo escribir $f(x, y) = g(x)h(y)$

2- Podemos observar del último ejemplo, que puede ocurrir que aún cuando la función $f(x, y)$ ya esté escrita en forma factorizada como una función sólo de x por una función sólo de y , las variables X e Y no sean independientes puesto que el dominio es tal que hace que las variables sean dependientes. Así, en el ejemplo anterior, el recorrido $R_{XY} = \{(X, Y): 0 \leq x \leq y \leq 1\}$ condiciona los valores que toma x a los valores que toma y .

3- El concepto de independencia entre dos variables aleatorias se puede generalizar a n variables aleatorias X_1, X_2, \dots, X_n

4.4 - Función de una variable aleatoria bidimensional

Existen muchas situaciones en las que dado una variable aleatoria bidimensional nos interesa considerar otra variable aleatoria que es función de aquélla. Por ejemplo, supongamos que las variables aleatorias X e Y denotan la longitud y el ancho, respectivamente, de una pieza, entonces $Z = 2X + 2Y$ es una v.a. que representa el perímetro de la pieza, o la v.a. $W = X \cdot Y$ representa el área de la pieza. Tanto Z como W *son variables aleatorias*.

En general, sea S un espacio muestral asociado a un experimento probabilístico ε , sean $X: S \rightarrow R$ e $Y: S \rightarrow R$ dos variables aleatorias que definen una variable aleatoria bidimensional (X, Y) cuyo recorrido es R_{XY} , y sea una función de dos variables reales $H: R_{XY} \rightarrow R$ que a cada elemento (x, y) del recorrido R_{XY} le hace corresponder un número real $z = H(x, y)$, entonces la función compuesta $Z =$

$H(X, Y): S \rightarrow R$ es una variable aleatoria, puesto que a cada elemento $s \in S$ le hace corresponder un número real $z = H[X(s), Y(s)]$. Diremos que la variable aleatoria Z es **función de la variable aleatoria bidimensional** (X, Y) .

Algunas variables aleatorias que son función de variables aleatorias bidimensionales son $Z = X + Y$, $Z = X \cdot Y$, $Z = X/Y$, $Z = \min(X, Y)$, $Z = \max(X, Y)$, etc.

Lo anterior se puede generalizar si en lugar de dos variables aleatorias tenemos n variables aleatorias X_1, X_2, \dots, X_n , y $z = H(x_1, x_2, \dots, x_n)$ es una función de n variables a valores reales.

Ejemplos: (OPTATIVO)

1- Sea $Z \sim B(n, p)$

Podemos escribir a Z como suma de variables aleatorias de la siguiente forma.

Recordar que Z cuenta el número de éxitos en n repeticiones o ensayos del experimento ε

Si definimos

$$X_i = \begin{cases} 1 & \text{si en la } i\text{-ésima repetición de } \varepsilon \text{ ocurre éxito} \\ 0 & \text{caso contrario} \end{cases} \quad i = 1, 2, \dots, n$$

Notar que a cada X_i se le puede considerar $B(1, p)$, y además X_1, X_2, \dots, X_n son independientes

Podemos escribir $Z = X_1 + X_2 + \dots + X_n$

2- Sea Z v.a. binomial negativa con parámetros r y p , es decir $Z \sim BN(r, p)$

Si definimos

X_1 : “número de repeticiones del experimento requeridos hasta el 1º éxito”

X_2 : “número de repeticiones del experimento adicionales requeridos hasta el 2º éxito”

X_3 : “número de repeticiones del experimento adicionales requeridos hasta el 3º éxito”

Y en general

X_i : “número de repeticiones del experimento adicionales después del $(i-1)$ -ésimo éxito requeridos hasta el i -ésimo éxito”

Entonces cada variable tiene **distribución geométrica con parámetro p** y $Z = X_1 + X_2 + \dots + X_r$

Notar además que X_1, X_2, \dots, X_r son independientes

4.4.1 - Esperanza de una v.a. que es función de una v.a. bidimensional

Sea una variable aleatoria bidimensional (X, Y) cuya *fdp* conjunta es la función de probabilidad conjunta $p(x_i, y_j)$ si es discreta o la función de densidad de probabilidad conjunta $f(x, y)$ si es continua y sea una función real de dos variables $z = H(x, y)$ de manera que podemos definir una variable aleatoria Z que es función de la variable aleatoria bidimensional (X, Y) de la forma $Z = H(X, Y)$. Si la *fdp* de Z es $q(z_i)$, si Z es discreta, o $q(z)$ si es continua, entonces la esperanza matemática de Z es, de acuerdo con la definición general,

$$E(Z) = \sum_{x_i \in R_X} z_i \cdot q(z_i) \quad (\text{para } Z \text{ discreta})$$

$$E(Z) = \int_{-\infty}^{\infty} zq(z)dz \quad (\text{para } Z \text{ continua})$$

Nuevamente lo interesante es considerar la posibilidad de evaluar $E(Z)$ sin tener que calcular previamente la *fdp* de Z . El siguiente teorema nos muestra cómo hacerlo.

Teorema Sea (X, Y) una variable aleatoria bidimensional y sea $Z=H(X, Y)$ una variable aleatoria que es función de (X, Y) .

a) Si Z es variable aleatoria discreta que proviene de la variable aleatoria bidimensional discreta (X, Y) cuyo recorrido es R_{XY} y su *fdp* conjunta es $p(x_i, y_j)$, entonces:

$$E(Z) = E[H(X, Y)] = \sum_{(x_i, y_j) \in R_{XY}} H(x_i, y_j) p(x_i, y_j)$$

b) Si Z es variable aleatoria continua que proviene de la variable aleatoria continua bidimensional (X, Y) cuya *fdp* conjunta es $f(x, y)$, entonces:

$$E(Z) = E[H(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x, y) f(x, y) dx dy.$$

Dem.) sin demostración

Ejemplo:

Supongamos que debido a innumerables causas incontrolables la corriente i y la resistencia r de un circuito varían aleatoriamente de forma tal que pueden considerarse como variables aleatorias I y R independientes. Supongamos que las correspondientes *fdp* son:

$$g(i) = \begin{cases} 2i & 0 \leq i \leq 1 \\ 0 & \text{caso contrario} \end{cases} \quad h(r) = \begin{cases} \frac{r^2}{9} & 0 \leq r \leq 3 \\ 0 & \text{caso contrario} \end{cases}$$

Nos interesa considerar el voltaje $v = i \cdot r$ de manera que podemos definir la variable aleatoria $V = I \cdot R$. Específicamente deseamos conocer el valor esperado o esperanza matemática del voltaje: $E(V)$. Usando la propiedad establecida en el teorema anterior:

$E(V) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(i, r) \cdot f(i, r) di dr$. Ahora bien, puesto que consideramos que I y R son variables aleatorias independientes, la *fdp* conjunta de la v.a. bidimensional (I, R) es simplemente el producto de las marginales o sea de las *fdp* de las variables aleatorias I y R tomadas como variables aleatorias unidimensionales: $f(i, r) = g(i) \cdot h(r)$, es decir:

$$f(i, r) = \begin{cases} \frac{2}{9} i \cdot r^2 & \text{si } 0 \leq i \leq 1 \text{ y } 0 \leq r \leq 3 \\ 0 & \text{caso contrario} \end{cases}$$

Entonces

$$E(V) = \int_0^3 dr \int_0^1 di (i \cdot r) \cdot \frac{2}{9} i \cdot r^2 = \frac{2}{9} \int_0^3 r^3 dr \int_0^1 i^2 di = \frac{3}{2}$$

4.5 - Esperanza de una suma de variables aleatorias

Sean X e Y dos variables aleatorias **arbitrarias**.
Entonces $E(X + Y) = E(X) + E(Y)$.

Podemos generalizar la propiedad anterior a un número finito cualquiera de variables aleatorias:

Sean X_1, X_2, \dots, X_n n variables aleatorias **arbitrarias**. Entonces:

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) \quad \text{o, en notación más concentrada,}$$

$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i)$$

(leeremos: “la esperanza de la suma es la suma de las esperanzas”)

Observación: se deduce que la esperanza verifica la **propiedad lineal**:

$$E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i).$$

Ejemplos:

1- Vamos a aplicar algunas de las propiedades anteriores para calcular de una manera alternativa la esperanza matemática de una variable aleatoria X distribuida binomialmente.

Sea entonces una v.a. $X \sim B(n, p)$.

Ya vimos que podemos escribir $X = X_1 + X_2 + \dots + X_n$ donde cada X_i se la puede considerar $B(1, p)$, y además X_1, X_2, \dots, X_n son independientes

Entonces

$$E(X_i) = 1 \times P(X_i = 1) + 0 \times P(X_i = 0) = P(X_i = 1) = p \quad \text{para cualquier } i$$

Por lo tanto

$$E(X) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) = \underbrace{p + p + \dots + p}_{n \text{ veces}} = np$$

2- El espesor X de una cuña de madera (en milímetros) tiene una función de densidad de probabilidad

$$f(x) = \begin{cases} \frac{3}{4} - \frac{3(x-5)^2}{4} & 4 \leq x \leq 6 \\ 0 & \text{en otro lado} \end{cases}$$

a) Determine $E(X)$

b) Si Y denota el espesor de una cuña en pulgadas ($1\text{mm} = 0.0394$ pulgadas), determine $E(Y)$

c) Si se seleccionan tres cuñas de manera independiente y las apilamos una encima de otra, encuentre la media del espesor total.

a) Verifique el lector que

$$E(X) = \int_4^6 x \left(\frac{3}{4} - \frac{3}{4} \right)^2 dx = 5$$

b) $Y = 0.0394X$ entonces $E(Y) = E(0.0394X) = 0.0394E(X) = 0.197$

c) Notar que si X_i : “espesor de cuña i ”, $i = 1, 2, 3$ entonces $X = X_1 + X_2 + X_3$ es el espesor total

$$\text{Por lo tanto } E(X) = E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3) = 5 + 5 + 5 = 15$$

Observación: muchas veces es conveniente descomponer una variable aleatoria como suma de otras más simples para facilitar los cálculos

3- **Esperanza de una v.a. binomial negativa (OPTATIVO)**

Cuando se trató la v.a. binomial negativa se dijo cuál era su esperanza. Ahora damos una demostración

Sea X v.a. binomial negativa con parámetros r y p , es decir $X \sim BN(r, p)$

Si definimos

X_1 : “número de repeticiones del experimento requeridos hasta el 1° éxito”

X_2 : “número de repeticiones del experimento adicionales requeridos hasta el 2° éxito”

X_3 : “número de repeticiones del experimento adicionales requeridos hasta el 3° éxito”

Y en general

X_i : “número de repeticiones del experimento adicionales después del $(i-1)$ –ésimo éxito requeridos hasta

el i –ésimo éxito”

Entonces cada variable tiene **distribución geométrica con parámetro p** y $X = X_1 + X_2 + \dots + X_r$

Por lo tanto

$$E(X) = E(X_1 + X_2 + \dots + X_r) = E(X_1) + E(X_2) + \dots + E(X_r) = \underbrace{\frac{1}{p} + \frac{1}{p} + \dots + \frac{1}{p}}_{r \text{ veces}} = r \frac{1}{p} = \frac{r}{p}$$

En general **la esperanza de un producto de variables aleatorias no es igual al producto de las esperanzas**

Si (X, Y) es una variable aleatoria bidimensional tal que X e Y son variables aleatorias **independientes**, entonces: $E(X.Y) = E(X).E(Y)$

(leeremos:” la esperanza del producto es el producto de las esperanzas”).

Dem.) (caso continuo)

Sea $Z = H(X, Y) = X.Y$ y sea $f(x, y)$ la *fdp* conjunta de la v.a. (X, Y) . Entonces, usando el *teorema* que establece que $E(Z) = E[H(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x, y) f(x, y) dx dy$, tenemos:

$E(X.Y) = \int_{-\infty}^{\infty} (x.y) f(x, y) dx dy$. Pero siendo X e Y variables aleatorias independientes es $f(x, y) = g(x)h(y)$, donde $g(x)$ y $h(y)$ son las *fdp* marginales de X e Y que sabemos coinciden con las *fdp* de X e Y tomadas como variables aleatorias unidimensionales. Luego:

$$\begin{aligned} E(X.Y) &= \int_{-\infty}^{\infty} (x.y) g(x) h(y) dx dy = \int_{-\infty}^{\infty} x g(x) dx \int_{-\infty}^{\infty} y h(y) dy \\ &= E(X).E(Y) \quad , \end{aligned}$$

donde en la última igualdad tuvimos en cuenta la definición de esperanza de v.a. unidimensionales.

Ejemplo:

Nuevamente consideramos el siguiente ejemplo

Supongamos que debido a innumerables causas incontrolables la corriente i y la resistencia r de un circuito varían aleatoriamente de forma tal que pueden considerarse como variables aleatorias I y R independientes. Supongamos que las correspondientes *fdp* son:

$$g(i) = \begin{cases} 2i & 0 \leq i \leq 1 \\ 0 & \text{caso contrario} \end{cases} \quad h(r) = \begin{cases} \frac{r^2}{9} & 0 \leq r \leq 3 \\ 0 & \text{caso contrario} \end{cases}$$

Nos interesa considerar el voltaje $v = i.r$ de manera que podemos definir la variable aleatoria $V = I.R$. Hallar el valor esperado o esperanza matemática del voltaje: $E(V)$.

Como I y R son independientes, usando la propiedad anterior

$$E(V) = E(I)E(R)$$

$$E(I) = \int_0^1 i(2i) di = 2 \left. \frac{i^3}{3} \right|_0^1 = \frac{2}{3} \quad E(R) = \int_0^3 r \left(\frac{r^2}{9} \right) dr = \left. \frac{1}{9} \frac{r^4}{4} \right|_0^3 = \frac{1}{9} \times \frac{3^4}{4} = 1$$

$$\therefore E(V) = \frac{2}{3} \times 1 = \frac{2}{3}$$

4.6 - Varianza de una suma de variables aleatorias

$$V(X + Y) = V(X) + V(Y) + 2\sigma_{XY} \quad \text{con} \quad \sigma_{XY} = E(X.Y) - E(X).E(Y)$$

Dem.) Escribimos la varianza en su forma alternativa

$V(X + Y) = E([X + Y]^2) - [E(X + Y)]^2$. Desarrollamos los cuadrados y aplicamos la propiedad lineal de la esperanza:

$$\begin{aligned} V(X + Y) &= E(X^2 + 2.X.Y + Y^2) - [E(X) + E(Y)]^2 \\ &= E(X^2) + 2E(X.Y) + E(Y^2) - \{[E(X)]^2 + 2E(X)E(Y) + [E(Y)]^2\} \end{aligned}$$

Agrupando convenientemente:

$$\begin{aligned} V(X + Y) &= \{E(X^2) - [E(X)]^2\} + \{E(Y^2) - [E(Y)]^2\} + 2\{E(X.Y) - E(X)E(Y)\} \\ &= V(X) + V(Y) + 2\{E(X.Y) - E(X)E(Y)\} \quad , \text{ es decir} \end{aligned}$$

$$V(X + Y) = V(X) + V(Y) + 2\sigma_{XY}$$

$$\sigma_{XY} = E(X.Y) - E(X).E(Y) \text{ se la llama la } \textbf{covarianza de X e Y}.$$

Observaciones:

1- Teniendo presente la definición de la desviación estándar de una v.a. X : $\sigma_X = \sqrt{V(X)}$, vemos que a la propiedad anterior la podemos escribir:

$$V(X + Y) = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY}$$

2- Análogamente se prueba que $V(X - Y) = \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}$

3- X e Y son independientes, entonces $V(X + Y) = V(X - Y) = V(X) + V(Y)$

Esto es porque si las variables aleatorias X e Y son independientes, entonces $E(X.Y) = E(X).E(Y)$. Por lo tanto la covarianza vale cero : $\sigma_{XY} = E(X.Y) - E(X).E(Y) = 0$.

4- Podemos generalizar, usando el principio de inducción completa, al caso de n variables aleatorias independientes:

Si X_1, X_2, \dots, X_n son n variables aleatorias independientes entonces:

$$V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n) \text{ o, en forma más compacta, } V(\sum_{i=1}^n X_i) = \sum_{i=1}^n V(X_i).$$

5- Vemos que la esperanza de la suma de dos variables aleatorias X e Y es igual a la suma de las esperanzas $E(X + Y) = E(X) + E(Y)$ cualesquiera sean X e Y . En cambio la varianza de la suma de las variables aleatorias X e Y es, en general, igual a la suma de las varianzas, $V(X + Y) = V(X) + V(Y)$, sólo si X e Y son variables independientes.

Ejemplos:

1- Podemos ejemplificar la aplicación de las propiedades de la varianza, calculando nuevamente la varianza de una v.a. X distribuida binomialmente con parámetros n y p .

Sea entonces una v.a. $X \sim B(n, p)$. Vimos que se puede escribir:

$X = X_1 + X_2 + \dots + X_n$, donde las n variables aleatorias son independientes entre sí y tienen todas la misma distribución:

$$X_i \sim B(1, p) \quad \forall i = 1, 2, \dots, n$$

Entonces, tratándose de n variables aleatorias independientes

$V(X) = V(X_1) + V(X_2) + \dots + V(X_n)$ todas la varianzas son iguales y podemos escribir la suma como n veces una cualquiera de ellas:

$$V(X) = nV(X_i). \text{ Pero}$$

$$V(X_i) = E(X_i^2) - [E(X_i)]^2.$$

$$\text{Ya vimos que } E(X_i) = 1 \cdot p + 0(1 - p) = p$$

$$\text{Además es: } E(X_i^2) = 1^2 \cdot p + 0^2(1 - p) = p$$

$$\text{Entonces: } V(X_i) = E(X_i^2) - [E(X_i)]^2 = p - p^2 = p(1 - p).$$

Luego:

$$V(X) = nV(X_i) = np(1 - p)$$

que es el resultado que habíamos obtenido a partir de la definición y llevando las sumas involucradas a la forma del desarrollo de un binomio de Newton.

2- Varianza de una v.a. binomial negativa

Ya vimos que podemos escribir $X = X_1 + X_2 + \dots + X_r$, donde cada variable X_i tiene **distribución geométrica con parámetro p**

Por lo tanto

$$V(X) = V(X_1) + V(X_2) + \dots + V(X_r) = r \frac{1 - p}{p^2}$$

3- En el ejemplo de las cuñas de madera

Si X_i : “espesor de cuña i ”, $i = 1, 2, 3$ entonces $X = X_1 + X_2 + X_3$ es el espesor total

$$\text{Por lo tanto } V(X) = V(X_1) + V(X_2) + V(X_3) = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} = \frac{3}{5}$$

4.7 - Covarianza

Sean X e Y dos variables aleatorias. La **covarianza de X e Y** se define:

$$\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\} = E(X \cdot Y) - E(X) \cdot E(Y)$$

Notación: la notación usual para la covarianza de X e Y es σ_{XY} o $\text{Cov}(X, Y)$

La última igualdad surge de desarrollar el producto y aplicar las propiedades de la esperanza:

$$E\{[X - E(X)] \cdot [Y - E(Y)]\} = E\{X \cdot Y - X \cdot E(Y) - E(X) \cdot Y + E(X)E(Y)\}$$

Teniendo presente que $E(X)$ y $E(Y)$ son constantes:

$$E\{[X - E(X)] \cdot [Y - E(Y)]\} = E(X \cdot Y) - E(X) \cdot E(Y) - E(X)E(Y) + E(X)E(Y) = E(X \cdot Y) - E(X) \cdot E(Y).$$

La esperanza en la definición debe pensarse de la siguiente manera (suponiendo que la v.a. (X, Y) es continua):

$$Cov(X, Y) = E\{H(X, Y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x, y) f(x, y) dx dy \quad \text{con } H(x, y) = [x - E(X)] \cdot [y - E(Y)]$$

Ya vimos la siguiente propiedad

Si X e Y son variables aleatorias independientes, entonces $Cov(X, Y) = 0$.

Dem.)

Según vimos, si X e Y son variables aleatorias independientes, entonces $E(X \cdot Y) = E(X) \cdot E(Y)$, de donde se sigue la propiedad.

Propiedades de la covarianza

Las siguientes propiedades son útiles y su verificación se deja como ejercicio

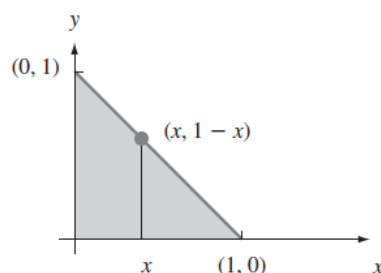
- 1- $Cov(a + bX, c + dY) = bdCov(X, Y)$
- 2- $Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)$
- 3- $Cov(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j) = \sum_{i=1}^n \sum_{j=1}^m Cov(X_i, Y_j)$
- 4- $Cov(X, X) = V(X)$

Ejemplo: (libro del autor Devore)

Una compañía de nueces comercializa latas de nueces combinadas de lujo que contienen almendras, nueces de acajú y cacahuates. Suponga que el peso neto de cada lata es exactamente de 1 lb, pero la contribución al peso de cada tipo de nuez es aleatoria. Como los tres pesos suman 1, un modelo de probabilidad conjunta de dos cualquiera da toda la información necesaria sobre el peso del tercer tipo.

Sea X :” el peso de las almendras en una lata seleccionada” y Y :” el peso de las nueces de acajú”. La función de densidad conjunta es

$$f(x, y) = \begin{cases} 24xy & 0 \leq x \leq 1, \quad 0 \leq y \leq 1, \quad 0 \leq x + y \leq 1 \\ 0 & \text{caso contrario} \end{cases}$$



Las densidades marginales son:

$$g(x) = \begin{cases} 12x(1-x)^2 & 0 \leq x \leq 1 \\ 0 & \text{caso contrario} \end{cases}$$

$$h(y) = \begin{cases} 12y(1-y)^2 & 0 \leq y \leq 1 \\ 0 & \text{caso contrario} \end{cases}$$

$$E(X) = E(Y) = \frac{2}{5}$$

$$E(X \cdot Y) = \int_0^1 \int_0^{1-x} xy \cdot 12xy dy dx = \frac{2}{15}$$

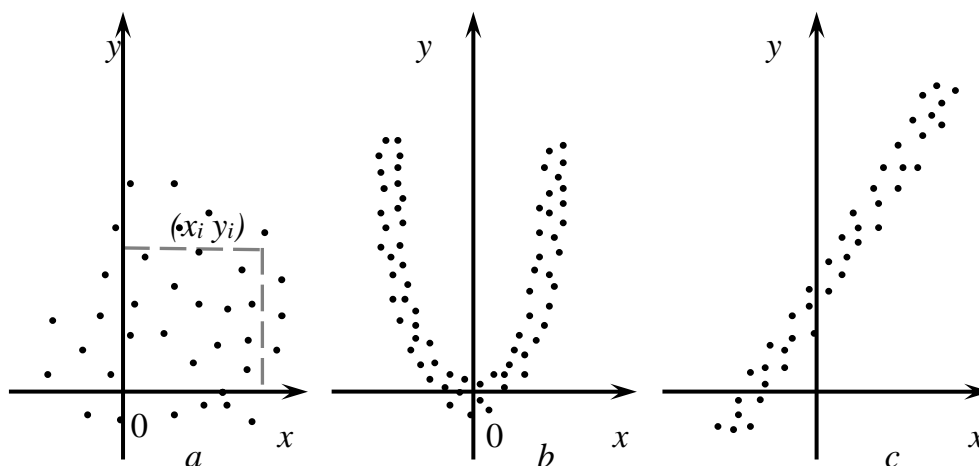
$$\text{Entonces } \text{Cov}(X, Y) = \frac{2}{15} - \frac{1}{5} \cdot \frac{1}{5} = -\frac{2}{75}$$

Una covarianza negativa se considera razonable en este caso porque más almendras contenidas en la lata implican menos nueces de acajú.

4.8 - Coefficiente de correlación lineal.

En realidad más que la covarianza aquí nos interesa considerar una cantidad relacionada con σ_{XY} y que según veremos nos dará información sobre el grado de asociación que existe entre X e Y . Más concretamente nos contará si existe algún grado de relación lineal entre X e Y . Esa cantidad es el coeficiente de correlación lineal.

En el mismo sentido en que podemos tener una idea aproximada sobre la probabilidad de un suceso A si repetimos el experimento y consideramos las ocurrencias de A en las n repeticiones, así podemos tener también una primera idea sobre la existencia de una relación funcional, específicamente una relación lineal, entre X e Y si consideramos un **diagrama de dispersión**. Consiste en dibujar pares de valores (x_i, y_j) medidos de la variable aleatoria (X, Y) en un sistema de coordenadas. En la figura mostramos diversas situaciones posibles.



De la figura a se deduciría que entre X e Y no hay ningún tipo de relación funcional. La figura b sugiere la posibilidad de que exista una relación funcional que corresponde a una parábola. La figura c , por su parte, sugiere una relación lineal entre X e Y . Este último es el comportamiento que nos interesa caracterizar. Con ese fin definimos el coeficiente de correlación lineal como sigue:

Sea (X, Y) una variable aleatoria bidimensional. Definimos el **coeficiente de correlación lineal entre X e Y** como $\rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$

En consecuencia:

$$\rho_{XY} = \frac{E\{[X-E(X)][Y-E(Y)]\}}{\sqrt{V(X)V(Y)}} = \frac{E(XY)-E(X)E(Y)}{\sqrt{V(X)V(Y)}}.$$

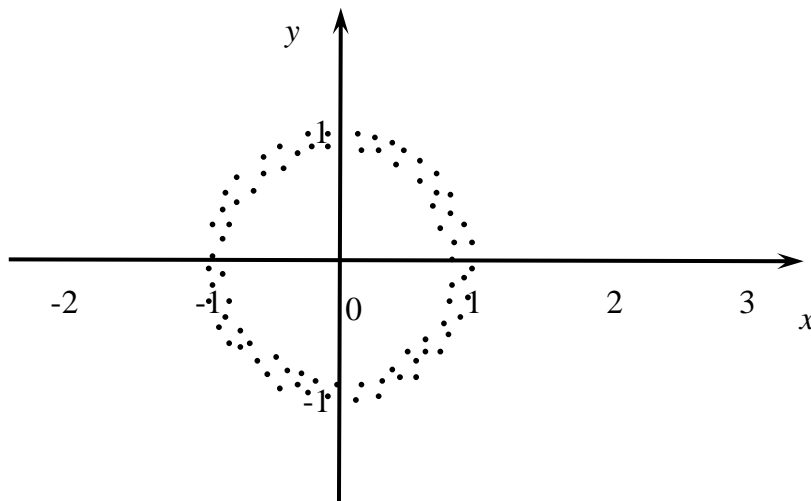
Daremos una serie de propiedades de ρ_{XY} que nos permitirán establecer más concretamente su significado.

Propiedad 1

Si X e Y son variables aleatorias independientes entonces $\rho_{XY} = 0$.

Dem.) inmediata a partir del hecho que si X e Y son independientes entonces $E(XY) = E(X)E(Y)$

Observación: La inversa no es necesariamente cierta. Puede ser que $\rho_{XY} = 0$ y sin embargo X e Y no sean variables aleatorias independientes. En efecto si tenemos una v.a. bidimensional (X, Y) que da lugar a un diagrama de dispersión como el que se muestra en la figura, veremos que correspondería a un coeficiente de correlación lineal $\rho_{XY} = 0$ y sin embargo la figura sugiere que entre X e Y existe la relación funcional $X^2 + Y^2 = 1$, es decir X e Y son v.a. dependientes. En realidad, como veremos, ρ_{XY} es una medida de la existencia de una relación lineal entre X e Y y una circunferencia se aleja mucho de una línea recta.



Propiedad 2 :

$$-1 \leq \rho_{XY} \leq 1$$

Dem.)

Si consideramos la v.a. $\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}$ entonces

$$0 \leq V\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = \frac{V(X)}{\sigma_X^2} + \frac{V(Y)}{\sigma_Y^2} + \frac{2Cov(X, Y)}{\sigma_X \sigma_Y} = 2(1 + \rho_{XY})$$

Implicando que $-1 \leq \rho_{XY}$

Por otro lado: $0 \leq V\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = \frac{V(X)}{\sigma_X^2} + \frac{V(Y)}{\sigma_Y^2} - \frac{2Cov(X,Y)}{\sigma_X\sigma_Y} = 2(1 - \rho_{XY})$

Implicando que $\rho_{XY} \leq 1$

$$\therefore -1 \leq \rho_{XY} \leq 1$$

Propiedad 3 :

Si $\rho_{XY}^2 = 1$, entonces con probabilidad 1 es $Y = A.X + B$ donde A y B son constantes.

Dem.) Si $\rho_{XY}^2 = 1$ entonces $\rho_{XY} = 1$ o $\rho_{XY} = -1$

Si $\rho_{XY} = -1$ entonces de la demostración anterior se deduce que

$V\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 2(1 + \rho_{XY}) = 0$, lo que implica que la v.a. $Z = \frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}$ tiene varianza cero. Según la interpretación de varianza podemos deducir (en forma intuitiva) que la v.a. **no tiene dispersión con respecto a su esperanza**, es decir la v.a. ***Z es una constante con probabilidad 1***

Por lo tanto esto implica que $Y = A.X + B$ con $A = -\frac{\sigma_Y}{\sigma_X} < 0$

Análogamente $\rho_{XY} = 1$ implica que $Y = A.X + B$ con $A = \frac{\sigma_Y}{\sigma_X} > 0$

Propiedad 4 :

Si X e Y son dos variables aleatorias tales que $Y = A.X + B$, donde A y B son constantes, entonces $\rho_{XY}^2 = 1$. Si $A > 0$ es $\rho_{XY} = 1$ y si $A < 0$ es $\rho_{XY} = -1$.

Dem.) se deja como ejercicio

Observación: Claramente las propiedades anteriores establecen que el coeficiente de correlación lineal es una medida del grado de linealidad entre X e Y .

Práctica**Variable aleatoria bidimensional****Funciones de variables aleatorias bidimensionales**

- 1) Cierta supermercado tiene una caja de salida común y una caja rápida. Denote por X_1 el número de clientes que están en espera en la caja común en un momento en particular del día, y por X_2 el número de clientes en espera en la caja rápida al mismo tiempo. Suponga que la fdp conjunta de X_1 y X_2 es como se indica en la tabla siguiente:

	X_1				
		0	1	2	3
X_2	0	0.08	0.07	0.04	0.00
	1	0.06	0.15	0.05	0.04
	2	0.05	0.04	0.10	0.06
	3	0.00	0.03	0.04	0.07
	4	0.00	0.01	0.05	0.06

- ¿Cuál es $P(X_1 = 1, X_2 = 1)$, esto es, la probabilidad de que haya exactamente un cliente en cada línea de espera?.
 - ¿Cuál es $P(X_1 = X_2)$, esto es, la probabilidad de que los números de clientes de las dos líneas de espera sean iguales?
 - Sea A el evento de que haya por lo menos dos clientes más en una línea de espera que en la otra. Expresa A en términos de X_1 y X_2 y calcule la probabilidad de este evento.
 - ¿Cuál es la probabilidad de que el número total de clientes de las dos líneas de espera sea exactamente cuatro?. ¿Y por lo menos cuatro?.
 - Determine la fdp. marginal de X_1 y la fdp marginal de X_2 .
 - Por inspección de las probabilidades $P(X_1 = 4)$, $P(X_2 = 0)$ y $P(X_1 = 4, X_2 = 0)$, ¿son X_1 y X_2 variables aleatorias independientes?
- 2) Se asignan aleatoriamente los contratos para dos construcciones a una o más de tres empresas A, B y C. Sea Y_1 el número de contratos asignados a la empresa A, e Y_2 el número de contratos asignados a la empresa B. Recuerdese que cada empresa puede recibir 0, 1 o 2 contratos.
- Encuentre la función de probabilidad conjunta para Y_1 e Y_2 .
 - Obtenga $F(1,0)$.
 - Determine la distribución de probabilidad marginal de Y_1 y de Y_2 .
 - ¿Son Y_1 e Y_2 independientes?.
- 3) Un ingeniero mide la cantidad, (por peso), de un contaminante particular en unas muestras de aire de cierto volumen recogidas sobre la chimenea de una central de energía eléctrica que funciona a carbón. Sea X la cantidad del contaminante por muestra recogida cuando no está funcionando cierto dispositivo de limpieza en la chimenea, y sea Y la cantidad de contaminante por muestra recogida bajo las mismas condiciones ambientales cuando el dispositivo está trabajando. Se observa que el comportamiento de la frecuencia relativa de X e Y puede tener el modelo :

$$f(x,y) = \begin{cases} K & \text{si } 0 \leq x \leq 2 ; 0 \leq y \leq 1 ; 2y \leq x \\ 0 & \text{en otro lado} \end{cases}$$

- Determine el valor de K.
 - Calcule $P(X \geq 3Y)$, es decir encuentre la probabilidad de que el dispositivo de limpieza reduzca la cantidad de contaminante en un tercio o más.
 - Si está funcionando el dispositivo limpiador, calcule la probabilidad de que la cantidad de contaminante en una muestra sea mayor que 0.5.
 - Dado que la cantidad de contaminante en una muestra obtenida cuando funciona el limpiador es **mayor de** 0.5, calcule la probabilidad de que la cantidad de contaminante, si no hubiera funcionado el dispositivo limpiador, hubiera sido mayor que 1.5
 - Dado que la cantidad de contaminante en una muestra obtenida cuando funciona el limpiador es **igual a** 0.5, calcule la probabilidad de que la cantidad de contaminante, si no hubiera funcionado el dispositivo limpiador, hubiera sido mayor que 1.5
 - ¿Son independientes las cantidades de contaminantes por muestra recogida con y sin el dispositivo de limpieza?.
- 4) Sean X e Y las proporciones de dos tipos diferentes de componentes de una muestra de una mezcla de productos químicos utilizada como insecticida.. Suponga que X e Y tienen la función de densidad conjunta dada por :

$$f(x,y) = \begin{cases} 2 & \text{si } 0 \leq x \leq 1 ; 0 \leq y \leq 1 ; 0 \leq x+y \leq 1 \\ 0 & \text{en cualquier otro punto} \end{cases}$$

- Encuentre $P(X \leq 3/4, Y \leq 3/4)$
 - Calcular $P(X \leq 1/2, Y \leq 1/2)$
 - Calcule $P(X \geq 1/2 | Y \leq 1/4)$
 - Calcule $P(X \geq 1/2 | Y = 1/4)$
 - ¿Son X e Y independientes ?
- 5) La duración Y para los fusibles de cierto tipo se representa por la fdp:

$$f(y) = \begin{cases} (1/3) e^{-y/3}, & y > 0 \\ 0 & \text{caso contrario} \end{cases}$$

(Las mediciones se indican en cientos de horas).

- Si dos de estos fusibles tienen duraciones independientes Y_1 e Y_2 , encuentre la fdp conjunta para Y_1 e Y_2 .
 - Un fusible de a) se encuentra en un sistema principal y el otro en un sistema de emergencia que entra en función solamente si falla el sistema principal. La duración efectiva total de los dos fusibles es entonces $Y_1 + Y_2$. Calcule $P(Y_1 + Y_2 \leq 1)$.
- 6) En un automóvil seleccionado al azar, se revisa el desgaste de cada neumático, y cada faro delantero se verifica para ver si está correctamente alineado. Denotemos por X el número de faros delanteros que necesitan ajuste y por Y el número de neumáticos defectuosos.
Si X e Y son independientes con las siguientes fdp

x	0	1	2
p(x)	0.5	0.3	0.2

y	0	1	2	3	4
p(y)	0.6	0.1	0.05	0.05	0.2

- a) Calcule $P(X \leq 1, Y \leq 1)$
b) ¿Cuál es la $P(X+Y=0)$, (la probabilidad de no defectos)?
c) Calcule $P(X+Y \leq 1)$
- 7) La diferencia entre el número de clientes en la línea de espera común y el número de ellos en la línea de espera de la caja rápida del ejercicio 1) es $X_1 - X_2$. Calcular la diferencia esperada.
- 8) a) Determine la covarianza de las variables Y_1 e Y_2 del ejercicio 2)
b) Determine la covarianza de las variables X e Y del ejercicio 3)

5- SUMA DE VARIABLES ALEATORIAS Y TEOREMA CENTRAL DEL LÍMITE

5.1 – Suma de variables aleatorias independientes

Cuando se estudiaron las variables aleatorias bidimensionales se habló de una *función de variable aleatoria bidimensional*. En particular se nombró la suma de n variables aleatorias, pero no se dijo nada sobre la *distribución* de esa v.a. suma.

Es a menudo importante saber cuál es la distribución de una suma de variables aleatorias independientes.

Consideramos algunos ejemplos en el caso discreto

1- Suma de variables aleatorias independientes con distribución Poisson

$$X \sim P(\lambda_1) ; Y \sim P(\lambda_2) ; X \text{ y } Y \text{ independientes} \Rightarrow X + Y \sim P(\lambda_1 + \lambda_2)$$

Dem.)

Consideramos el evento $\{X + Y = n\}$ como unión de eventos excluyentes

$\{X = k, Y = n - k\} \quad 0 \leq k \leq n$, entonces

$$P(X + Y = n) = \sum_{k=0}^n P(X = k, Y = n - k) = \sum_{k=0}^n P(X = k)P(Y = n - k) = \sum_{k=0}^n e^{-\lambda_1} \frac{\lambda_1^k}{k!} e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} =$$

X e Y independientes

$$= e^{-(\lambda_1 + \lambda_2)} \sum_{k=0}^n \frac{\lambda_1^k}{k!} \frac{\lambda_2^{n-k}}{(n-k)!} = \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} \lambda_1^k \lambda_2^{n-k} = \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} (\lambda_1 + \lambda_2)^n$$

Binomio de Newton

O sea $X + Y$ tiene distribución Poisson con parámetro $\lambda_1 + \lambda_2$

2- Suma de variables aleatorias binomiales independientes

$$X \sim B(n_1, p) ; Y \sim B(n_2, p) ; X \text{ y } Y \text{ independientes} \Rightarrow X + Y \sim B(n_1 + n_2, p)$$

Dem.)

Nuevamente consideramos el evento $\{X + Y = k\}$ como unión de eventos excluyentes

$\{X = i, Y = k - i\} \quad 0 \leq i \leq n_1$, entonces

$$P(X + Y = k) = \sum_{i=0}^{n_1} P(X = i, Y = k - i) = \sum_{i=0}^{n_1} P(X = i)P(Y = k - i) = \sum_{i=0}^{n_1} \binom{n_1}{i} p^i (1-p)^{n_1-i} \binom{n_2}{k-i} p^{k-i} (1-p)^{n_2-k+i} =$$

X e Y independientes

$$= p^k (1-p)^{n_1+n_2-k} \sum_{i=0}^{n_1} \binom{n_1}{i} \binom{n_2}{k-i}$$

En la expresión anterior si $j > r$ entonces $\binom{r}{j} = 0$

Por último usamos la siguiente identidad combinatoria
$$\sum_{i=0}^{n_1} \binom{n_1}{i} \binom{n_2}{k-i} = \binom{n_1+n_2}{k}$$

Y entonces

$$P(X+Y=k) = \binom{n_1+n_2}{k} p^k (1-p)^{n_1+n_2-k}$$

O sea $X+Y$ tiene distribución binomial con parámetros n_1+n_2 y p

Observación: en los dos casos anteriores se puede generalizar el resultado a n variables aleatorias independientes, usando el principio de inducción completa, es decir

1- Si X_1, X_2, \dots, X_n son n variables aleatorias independientes donde $X_i \sim P(\lambda_i)$ para todo

$$i=1, 2, \dots, n \text{ entonces } \sum_{i=1}^n X_i \sim P\left(\sum_{i=1}^n \lambda_i\right)$$

2- Si X_1, X_2, \dots, X_n son n variables aleatorias independientes donde $X_i \sim B(n_i, p)$ para todo

$$i=1, 2, \dots, n \text{ entonces } \sum_{i=1}^n X_i \sim B\left(\sum_{i=1}^n n_i, p\right)$$

Suma de variables aleatorias normales independientes

Si X e Y son dos variables aleatorias continuas independientes con densidades $g(x)$ y $h(y)$ respectivamente se puede probar (no lo demostraremos aquí) que la v.a. $Z = X + Y$ tiene densidad dada por

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} g(z-y)h(y)dy$$

Usando esto se puede demostrar el siguiente importante resultado:

Si X e Y son variables aleatorias independientes donde $X \sim N(\mu_1, \sigma_1^2)$ y $Y \sim N(\mu_2, \sigma_2^2)$ entonces $X+Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Por inducción completa se puede generalizar este resultado a n variables:

Si X_1, X_2, \dots, X_n son n variables aleatorias independientes donde $X_i \sim N(\mu_i, \sigma_i^2)$ para todo $i=1, 2, \dots, n$ entonces $\sum_{i=1}^n X_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$

De lo anterior y del hecho que $X \sim N(\mu, \sigma^2) \Rightarrow aX + b \sim N(a\mu + b, a^2\sigma^2)$ tenemos:

Si X_1, X_2, \dots, X_n son n variables aleatorias independientes donde $X_i \sim N(\mu_i, \sigma_i^2)$ para todo $i=1, 2, \dots, n$ entonces $\sum_{i=1}^n a_i X_i \sim N\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2\right)$ donde a_1, a_2, \dots, a_n son números reales

Se dice que $\sum_{i=0}^n a_i X_i$ es una **combinación lineal de variables aleatorias**.

Ejemplos:

1- La envoltura de plástico para un disco magnético está formada por dos hojas. El espesor de cada una tiene una distribución normal con media 1.5 milímetros y desviación estándar de 0.1 milímetros. Las hojas son independientes.

- Determine la media y la desviación estándar del espesor total de las dos hojas.
- ¿Cuál es la probabilidad de que el espesor total sea mayor que 3.3 milímetros?

Solución: Sean las variables aleatorias

X : “espesor de la hoja 1” e Y : “espesor de la hoja 2”

Entonces $X \sim N(1.5, 0.1^2)$; $Y \sim N(1.5, 0.1^2)$ y X e Y independientes

a) Si definimos la v.a. Z : “espesor total de las dos hojas”, entonces $Z = X + Y$

Por lo tanto $Z \sim N(1.5 + 1.5, 0.1^2 + 0.1^2)$ es decir $Z \sim N(3, 0.02)$

En consecuencia $E(Z) = 3$, $\sigma_Z = \sqrt{V(Z)} = \sqrt{0.02}$

b) Se pide calcular $P(Z > 3.3)$

$$P(Z > 3.3) = P\left(\frac{Z - 3}{\sqrt{0.02}} > \frac{3.3 - 3}{\sqrt{0.02}}\right) = 1 - \Phi\left(\frac{3.3 - 3}{\sqrt{0.02}}\right) = 1 - \Phi(2.12132) = 1 - 0.983 = 0.017$$

2-Tengo tres mensajes que atender en el edificio administrativo. Sea X_i : “el tiempo que toma el i -ésimo mensaje” ($i = 1, 2, 3$), y sea X_4 : “el tiempo total que utilizo para caminar hacia y desde el edificio y entre cada mensaje”. Suponga que las X_i son independientes, normalmente distribuidas, con las siguientes medias y desviaciones estándar:

$$\mu_1 = 15 \text{ min}, \sigma_1 = 4, \mu_2 = 5, \sigma_2 = 1, \mu_3 = 8, \sigma_3 = 2, \mu_4 = 12, \sigma_4 = 3$$

Pienso salir de mi oficina precisamente a las 10.00 a.m. y deseo pegar una nota en mi puerta que dice “regreso a las t a.m.” ¿A qué hora t debo escribir si deseo que la probabilidad de mi llegada después de t sea 0.01?

Solución: Definimos la v.a. Z : “tiempo transcurrido desde que salgo de mi oficina hasta que regreso”, entonces $T = X_1 + X_2 + X_3 + X_4$

Por lo tanto $T \sim N\left(\sum_{i=1}^4 \mu_i, \sum_{i=1}^4 \sigma_i^2\right)$, y se pide hallar t tal que $P(T > t) = 0.01$

$$\sum_{i=1}^4 \mu_i = 15 + 5 + 8 + 12 = 50 \text{ y } \sum_{i=1}^4 \sigma_i^2 = 4^2 + 1^2 + 2^2 + 3^2 = 30$$

$$\text{Entonces } P(T > t) = 1 - \Phi\left(\frac{t - 50}{\sqrt{30}}\right) = 0.01, \text{ es decir } \Phi\left(\frac{t - 50}{\sqrt{30}}\right) = 0.99$$

$$\text{Buscando en la tabla de la normal } \frac{t - 50}{\sqrt{30}} = 2.33 \Rightarrow t = 2.33 \times \sqrt{30} + 50 = 62.7619$$

3- El ancho del marco de una puerta tiene una distribución normal con media 24 pulgadas y desviación estándar de 1/8 de pulgada. El ancho de la puerta tiene una distribución normal con media 23.875 de pulgadas y desviación estándar de 1/16 de pulgadas. Suponer independencia.

- Determine la distribución, la media y la desviación estándar de la diferencia entre el ancho del marco y de la puerta.

- b) ¿Cuál es la probabilidad de que la diferencia entre el ancho del marco y de la puerta sea mayor que $\frac{1}{4}$ de pulgada?
 c) ¿Cuál es la probabilidad de que la puerta no quepa en el marco?.

Solución: Sean las variables aleatorias

X : “ancho del marco de la puerta en pulgadas”

Y : “ancho de la puerta en pulgadas”

Entonces $X \sim N(24, (1/8)^2)$, $Y \sim N(23.875, (1/16)^2)$, X e Y independientes

- a) Se pide la distribución de $X-Y$, $E(X-Y)$, $\sigma_{X-Y} = \sqrt{V(X-Y)}$

$$E(X-Y) = E(X) - E(Y) = 24 - 23.875 = 0.125$$

$$V(X-Y) = V(X) + V(Y) = \left(\frac{1}{8}\right)^2 + \left(\frac{1}{16}\right)^2 = \frac{5}{256} \quad \therefore \sigma_{X-Y} = \frac{\sqrt{5}}{16}$$

$$\text{Por lo tanto } X-Y \sim N\left(0.125, \left(\frac{\sqrt{5}}{16}\right)^2\right)$$

- b) Se pide la probabilidad $P(X-Y > 1/4)$

$$P(X-Y > 1/4) = 1 - \Phi\left(\frac{0.25 - 0.125}{\frac{\sqrt{5}}{16}}\right) = 1 - \Phi\left(\frac{2\sqrt{5}}{5}\right) = 1 - \Phi(0.8944) = 1 - 0.8133 = 0.1867$$

- c) Si la puerta no entra en el marco entonces se da el evento $\{X < Y\}$ o equivalentemente $\{X-Y < 0\}$, por lo tanto

$$P(X-Y < 0) = \Phi\left(\frac{0 - 0.125}{\frac{\sqrt{5}}{16}}\right) = \Phi\left(-\frac{2\sqrt{5}}{5}\right) = 1 - \Phi\left(\frac{2\sqrt{5}}{5}\right) = 0.1867$$

- 4- Supongamos que las variables aleatorias X e Y denotan la longitud y el ancho en cm, respectivamente, de una pieza.

Supongamos además que X e Y son independientes y que $X \sim N(2, 0.1^2)$, $Y \sim N(5, 0.2^2)$.

Entonces $Z = 2X + 2Y$ es una v.a. que representa el perímetro de la pieza.

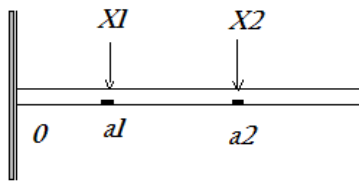
Calcular la probabilidad de que el perímetro sea mayor que 14.5 cm.

Solución: tenemos que $Z \sim N(2 \times 2 + 2 \times 5, 2^2 \times 0.1^2 + 2^2 \times 0.2^2)$, o sea $Z \sim N(14, 0.2)$

La probabilidad pedida es $P(Z > 14.5)$, entonces

$$P(Z > 14.5) = 1 - \Phi\left(\frac{14.5 - 14}{\sqrt{0.2}}\right) = 1 - \Phi\left(\frac{\sqrt{5}}{2}\right) = 1 - \Phi(1.1180) = 1 - 0.8810 = 0.119$$

- 5- Si se aplican dos cargas aleatorias X_1 y X_2 a una viga voladiza como se muestra en la figura siguiente, el momento de flexión en 0 debido a las cargas es $a_1 X_1 + a_2 X_2$.



- a) Suponga que X_1 y X_2 son v.a. independientes con medias 2 y 4 KLbs respectivamente, y desviaciones estándar 0.5 y 1.0 KLbs, respectivamente.

Si $a_1 = 5$ pies y $a_2 = 10$ pies, ¿cuál es el momento de flexión esperado y cuál es la desviación estándar del momento de flexión?

- b) Si X_1 y X_2 están normalmente distribuidas, ¿cuál es la probabilidad de que el momento de flexión supere 75 KLbs?

Solución: Sea la v.a. Z : “momento de flexión en 0”, entonces $Z = 5X_1 + 10X_2$

Por lo tanto

$$a) E(Z) = 5E(X_1) + 10E(X_2) = 5 \times 2 + 10 \times 4 = 50$$

$$V(Z) = 5^2 \times 0.5^2 + 10^2 \times 1^2 = 25 \times 0.25 + 10 \times 1 = \frac{65}{4} \quad \therefore \sigma_Z = \sqrt{\frac{65}{4}}$$

$$b) \text{ Si } X_1 \text{ y } X_2 \text{ están normalmente distribuidas, entonces } Z \sim N\left(50, \frac{65}{4}\right)$$

Por lo tanto

$$P(Z > 75) = 1 - \Phi\left(\frac{75 - 50}{\sqrt{\frac{65}{4}}}\right) = 1 - \Phi\left(\frac{10\sqrt{65}}{13}\right) = 1 - \Phi(6.20) \approx 1 - 1 = 0$$

Promedio de variables aleatorias normales independientes

Si X_1, X_2, \dots, X_n son n variables aleatorias independientes donde $X_i \sim N(\mu, \sigma^2)$ para todo

$i = 1, 2, \dots, n$ entonces la v.a. $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ tiene distribución normal con

media μ y varianza $\frac{\sigma^2}{n}$

Dem.) Notar que $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ es un caso particular de combinación lineal de variables aleatorias donde

$$a_i = \frac{1}{n} \text{ para todo } i = 1, 2, \dots, n$$

Además en este caso $\mu_i = \mu$ y $\sigma_i^2 = \sigma^2$ para todo $i = 1, 2, \dots, n$

Por lo tanto, \bar{X} tiene distribución normal con esperanza $\sum_{i=1}^n \frac{1}{n} \mu_i = \sum_{i=1}^n \frac{1}{n} \mu = \frac{1}{n} n \mu = \mu$ y varianza

$$\sum_{i=1}^n \left(\frac{1}{n}\right)^2 \sigma_i^2 = \sum_{i=1}^n \left(\frac{1}{n}\right)^2 \sigma^2 = \left(\frac{1}{n}\right)^2 n \sigma^2 = \frac{\sigma^2}{n}$$

Es decir, $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

Observación: a \bar{X} se lo llama **promedio muestral** o **media muestral**

Ejemplos:

1) El diámetro interno de un anillo de pistón seleccionado al azar es una v.a. con distribución normal con media 12 cm y desviación estándar de 0.04 cm.

a) Si \bar{X} es el diámetro promedio en una muestra de $n = 16$ anillos, calcule $P(11.99 \leq \bar{X} \leq 12.01)$

b) ¿Qué tan probable es que el diámetro promedio exceda de 12.01 cuando $n = 25$?

Solución:

a) Sean las variables aleatorias X_i : “diámetro del anillo i ” $i = 1, 2, \dots, 16$

Entonces $X_i \sim N(12, 0.04^2)$ para cada i .

Por lo tanto $\bar{X} \sim N\left(12, \frac{0.04^2}{16}\right)$. Entonces

$$\begin{aligned} P(11.99 \leq \bar{X} \leq 12.01) &= P\left(-\frac{11.99-12}{\sqrt{\frac{0.04^2}{16}}} \leq \frac{\bar{X}-12}{\sqrt{\frac{0.04^2}{16}}} \leq \frac{12.01-12}{\sqrt{\frac{0.04^2}{16}}}\right) = \\ &= \Phi\left(\frac{12.01-12}{\sqrt{\frac{0.04^2}{16}}}\right) - \Phi\left(\frac{11.99-12}{\sqrt{\frac{0.04^2}{16}}}\right) = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 = \\ &= 2 \times 0.8413 - 1 = 0.6826 \end{aligned}$$

b) En este caso $\bar{X} \sim N\left(12, \frac{0.04^2}{25}\right)$, entonces

$$P(\bar{X} > 12.01) = 1 - \Phi\left(\frac{12.01-12}{\sqrt{\frac{0.04^2}{25}}}\right) = 1 - \Phi(1.25) = 1 - 0.8944 = 0.1056$$

2) Una máquina embotelladora puede regularse de tal manera que llene un promedio de μ onzas por botella. Se ha observado que la cantidad de contenido que suministra la máquina presenta una distribución normal con $\sigma = 1$ onza. De la producción de la máquina un cierto día, se obtiene una muestra de 9 botellas llenas (todas fueron llenadas con las mismas posiciones del control operativo) y se miden las onzas del contenido de cada una.

- a) Determinar la probabilidad de que la media muestral se encuentre a lo más a 0.3 onzas de la media real μ para tales posiciones de control
- b) ¿Cuántas observaciones deben incluirse en la muestra si se desea que la media muestral esté a lo más a 0.3 onzas de μ con una probabilidad de 0.95?

Solución:

- a) Sean las variables aleatorias X_i : “contenido en onzas de la botella i ” $i = 1, 2, \dots, 9$

Entonces $X_i \sim N(\mu, 1)$ para cada i .

Por lo tanto $\bar{X} \sim N\left(\mu, \frac{1}{9}\right)$. Se desea calcular

$$\begin{aligned} P(|\bar{X} - \mu| \leq 0.3) &= P(-0.3 \leq \bar{X} - \mu \leq 0.3) = P\left(-\frac{0.3}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{0.3}{\sigma/\sqrt{n}}\right) = \\ &= P\left(-\frac{0.3}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{0.3}{\sigma/\sqrt{n}}\right) = P\left(-0.9 \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq 0.9\right) = \Phi(0.9) - \Phi(-0.9) = \\ &= 2\Phi(0.9) - 1 = 0.6318 \end{aligned}$$

- b) Ahora se pretende que

$$P(|\bar{X} - \mu| \leq 0.3) = P(-0.3 \leq \bar{X} - \mu \leq 0.3) = 0.95$$

Entonces

$$P(|\bar{X} - \mu| \leq 0.3) = P\left(-\frac{0.3}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{0.3}{\sigma/\sqrt{n}}\right) = P\left(-0.3\sqrt{n} \leq \frac{\bar{X} - \mu}{1/\sqrt{n}} \leq 0.3\sqrt{n}\right) = 0.95$$

Mediante la tabla de la acumulada de la normal estándar se tiene que

$$P\left(-0.3\sqrt{n} \leq \frac{\bar{X} - \mu}{1/\sqrt{n}} \leq 0.3\sqrt{n}\right) = 2\Phi(0.3\sqrt{n}) - 1 = 0.95 \Rightarrow \Phi(0.3\sqrt{n}) = 0.975 \Rightarrow (0.3\sqrt{n}) = 1.96$$

$$\text{O sea } n \approx \left(\frac{1.96}{0.3}\right)^2 = 42.68$$

Si tomamos $n = 43$, entonces $P(|\bar{X} - \mu| \leq 0.3)$ será un poco mayor que 0.95

5.2 - Teorema central del límite

Acabamos de ver que la suma de un número finito n de variables aleatorias independientes que están normalmente distribuidas es una variable aleatoria también normalmente distribuida. Esta propiedad reproductiva no es exclusiva de la distribución normal. En efecto, por ejemplo, ya vimos que existen variables aleatorias discretas que la cumplen, es el caso de la Poisson y la Binomial.

En realidad, la propiedad que le da a la distribución normal el lugar privilegiado que ocupa entre todas las distribuciones es el hecho de que la suma de un número muy grande, rigurosamente un número

infinito numerable, de variables aleatorias independientes con distribuciones **arbitrarias** (no necesariamente normales) es una variable aleatoria que tiene, aproximadamente, una distribución normal. Este es, esencialmente, el contenido del

Teorema central del límite (T.C.L.):

Sean X_1, X_2, \dots, X_n variables aleatorias independientes con $E(X_i) = \mu$ y $V(X_i) = \sigma^2$ para todo $i = 1, 2, \dots, n$, es decir **independientes idénticamente distribuidas**

Sea la v.a. $S_n = \sum_{i=1}^n X_i$ y sea $Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$.

Entonces $\lim_{n \rightarrow \infty} P(Z_n \leq z) = \Phi(z)$, esto es $\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq z\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-x^2/2} dx$

Dem.) sin demostración

Observaciones:

1- Notar que $E(S_n) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = n\mu$ y $V(S_n) = V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) = n\sigma^2$

Por lo tanto $Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$ es la v.a. S_n **estandarizada**

2- Notar que $P\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq z\right) = P\left(\frac{\frac{S_n - n\mu}{n}}{\frac{\sqrt{n\sigma^2}}{n}} \leq z\right) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z\right)$, por lo tanto también se puede enun-

ciar el Teorema central del límite de la siguiente forma

Sean X_1, X_2, \dots, X_n variables aleatorias independientes con $E(X_i) = \mu$ y $V(X_i) = \sigma^2$ para todo $i = 1, 2, \dots, n$, es decir **independientes idénticamente distribuidas**

Sea la v.a. promedio muestral $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ y sea $Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$.

Entonces $\lim_{n \rightarrow \infty} P(Z_n \leq z) = \Phi(z)$, esto es $\lim_{n \rightarrow \infty} P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-x^2/2} dx$

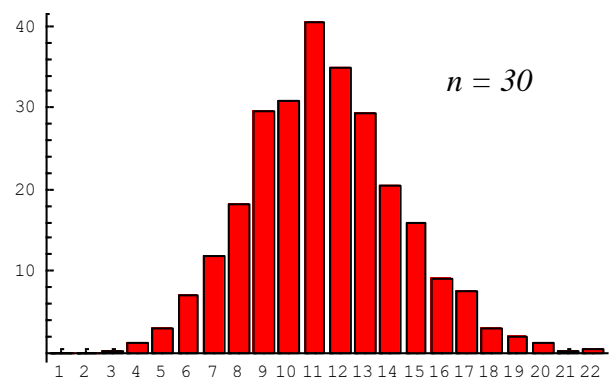
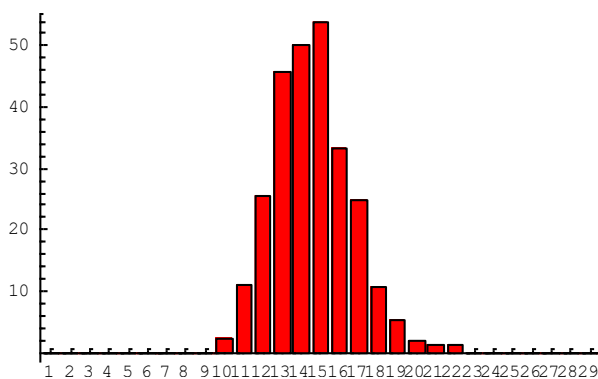
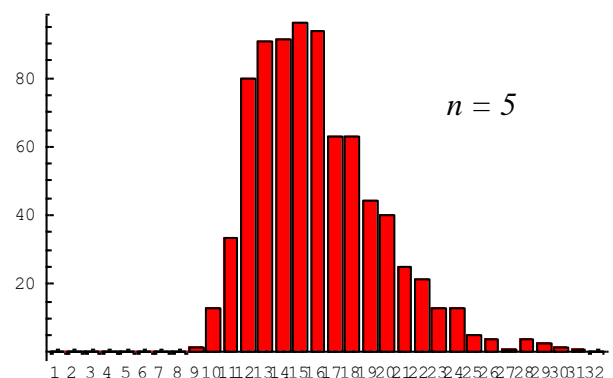
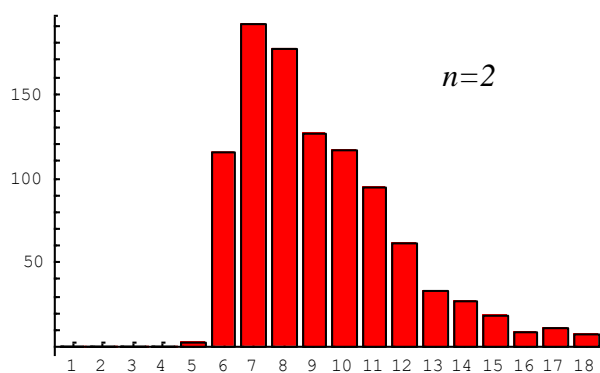
Donde $Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ es el **promedio muestral estandarizado**

3- Aunque en muchos casos el T.C.L. funciona bien para valores de n pequeños, en particular donde la población es continua y simétrica, en otras situaciones se requieren valores de n más grandes, dependiendo de la forma de la distribución de las X_i . En muchos casos de interés práctico, si $n \geq 30$, la aproximación normal será satisfactoria sin importar cómo sea la forma de la distribución de las X_i . Si $n < 30$, el T.C.L. funciona si la distribución de las X_i no está muy alejada de una distribución normal

4- Para interpretar el significado del T.C.L., se generan (por computadora) n valores de una v.a. exponencial con parámetro $\lambda = 0.5$, y se calcula el promedio de esos n valores. Esto se repite 1000 veces, por lo tanto tenemos 1000 valores de la v.a. \bar{X} .

Hacemos un **histograma de frecuencias** de \bar{X} , esto es, tomamos un intervalo (a,b) donde “caen” todos los valores de \bar{X} , y lo subdividimos en intervalos más chicos de igual longitud. La **frecuencia de cada subintervalo** es la cantidad de valores de \bar{X} que caen en dicho subintervalo. Se grafican estas frecuencias obteniéndose los gráficos siguientes que se pueden considerar una aproximación a la verdadera distribución de \bar{X} .

Se observa que a medida que aumenta el valor de n los gráficos se van haciendo más simétricos, pareciéndose a la gráfica de una distribución normal.



Ejemplos:

1- Supóngase que 30 instrumentos electrónicos D_1, D_2, \dots, D_{30} , se usan de la manera siguiente: tan pronto como D_1 falla empieza a actuar D_2 . Cuando D_2 falla empieza a actuar D_3 , etc. Supóngase que el tiempo de falla de D_i es una v.a. distribuida exponencialmente con parámetro $\lambda = 0.1$ por hora. Sea T el tiempo total de operación de los 30 instrumentos. ¿Cuál es la probabilidad de que T exceda 350 horas?

Solución:

Si X_i : “tiempo de falla del instrumento D_i ” $i = 1, 2, \dots, 30$

Entonces $X_i \sim \text{Exp}(0.1)$ para $i = 1, 2, \dots, 30$

El tiempo total de operación de los 30 instrumentos es $T = \sum_{i=1}^{30} X_i$, donde

$$E(T) = E\left(\sum_{i=1}^{30} X_i\right) = 30 \times E(X_i) = 30 \times \frac{1}{0.1} = 300$$

$$V(T) = V\left(\sum_{i=1}^{30} X_i\right) = 30 \times V(X_i) = 30 \times \frac{1}{0.1^2} = 3000$$

Entonces por T.C.L. $\frac{T - 300}{\sqrt{3000}} \sim N(0,1)$ *aproximadamente* pues $n = 30$

La probabilidad pedida es

$$P(T > 350) = P\left(\frac{T - 300}{\sqrt{3000}} > \frac{350 - 300}{\sqrt{3000}}\right) \underset{\text{T.C.L.}}{\approx} 1 - \Phi\left(\frac{350 - 300}{\sqrt{3000}}\right) = 1 - \Phi(0.9128) = 1 - 0.81859 = 0.18141$$

2- Suponga que el consumo de calorías por día de una determinada persona es una v.a. con media 3000 calorías y desviación estándar de 230 calorías. ¿Cuál es la probabilidad de que el promedio de consumo de calorías diario de dicha persona en el siguiente año (365 días) sea entre 2959 y 3050?

Solución:

Definimos las variables aleatorias

X_i : “cantidad de calorías que una persona consume en el día i ” $i = 1, 2, \dots, 365$

Se sabe que $E(X_i) = 3000$ y $V(X_i) = 230^2$

Si $\bar{X} = \frac{1}{365} \sum_{i=1}^{365} X_i$ entonces $E(\bar{X}) = 3000$ y $V(\bar{X}) = \frac{\sigma^2}{n} = \frac{230^2}{365}$

La probabilidad pedida es

$$P(2959 \leq \bar{X} \leq 3050) = P\left(\frac{2959 - 3000}{230/\sqrt{365}} \leq \frac{\bar{X} - 3000}{230/\sqrt{365}} \leq \frac{3050 - 3000}{230/\sqrt{365}}\right) \underset{\text{T.C.L.}}{\approx} \\ \approx \Phi\left(\frac{3050 - 3000}{230/\sqrt{365}}\right) - \Phi\left(\frac{2959 - 3000}{230/\sqrt{365}}\right) = \Phi(4.15) - \Phi(-3.40) \approx 1 - 0 = 1$$

Aplicaciones del Teorema central del límite

Aproximación normal a la distribución binomial

El Teorema central del límite se puede utilizar para aproximar las probabilidades de algunas variables aleatorias discretas cuando es difícil calcular las probabilidades exactas para valores grandes de los parámetros.

Supongamos que X tiene una distribución binomial con parámetros n y p . Para calcular $P(X \leq k)$

debemos hacer la suma $P(X \leq k) = \sum_{i=0}^k P(X = i)$ o recurrir a las tablas de la F.d.a., pero para valores

de n grandes no existen tablas, por lo tanto habría que hacer el cálculo en forma directa y muchas veces es laborioso.

Como una opción podemos considerar a X como suma de variables aleatorias más simples, específicamente, si definimos

$$X_i = \begin{cases} 1 & \text{si en la } i\text{-ésima repetición de } \varepsilon \text{ ocurre éxito} \\ 0 & \text{caso contrario} \end{cases} \quad i = 1, 2, \dots, n$$

entonces cada X_i se la puede considerar $B(1, p)$, y además X_1, X_2, \dots, X_n son independientes

Podemos escribir $X = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$ y **si n es grande** entonces X tendrá **aproximadamente** una distribución normal con parámetros np y $np(1-p)$, es decir

$$Z_n = \frac{X - n\mu}{\sqrt{n\sigma^2}} = \frac{X - n.p}{\sqrt{n.p(1-p)}} \approx N(0,1) \quad \text{si } n \text{ es lo suficientemente grande}$$

Observaciones:

1- La aproximación normal a la distribución binomial funciona bien aun cuando n no sea muy grande si p no está demasiado cerca de cero o de uno. En particular la aproximación normal a la binomial es buena si n es grande, $np > 5$ y $n(1-p) > 5$, **pero es más efectivo aplicar esta aproximación cuando $np > 10$ y $n(1-p) > 10$**

2- Corrección por continuidad.

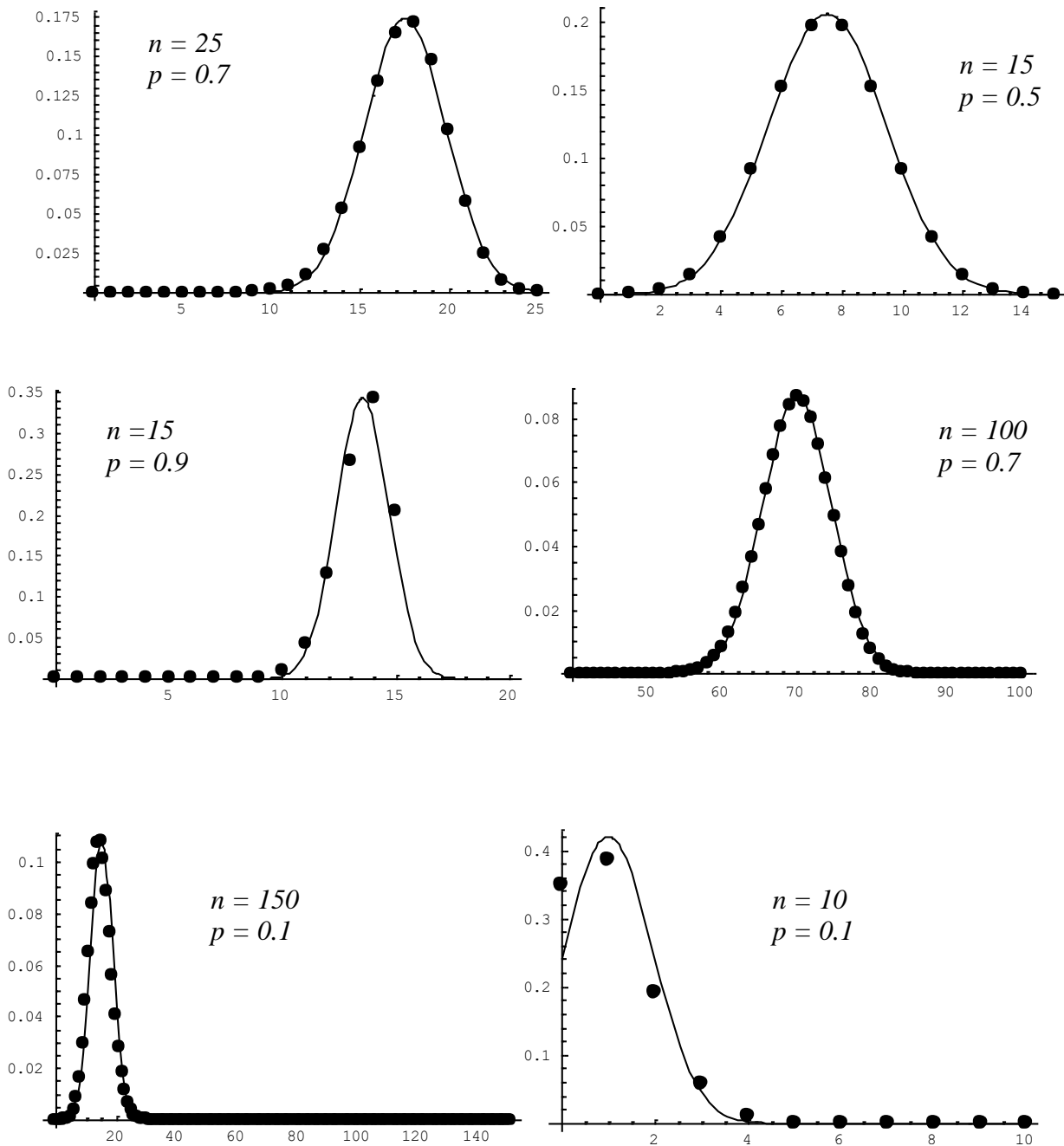
Acabamos de ver que si $X \sim B(n, p)$ entonces, para n suficientemente grande, podemos considerar que aproximadamente es $X \sim N[n.p, n.p(1-p)]$. El problema que surge de inmediato si deseo calcular, por ejemplo, la probabilidad de que $X = k$ (con k alguno de los valores posibles $0, 1, 2, \dots, n$) es que la binomial es una distribución discreta y tiene sentido calcular probabilidades como $P(X = k)$ mientras que la normal es una distribución continua y, en consecuencia, $P(X = k) = 0$ puesto que para una variable aleatoria continua la probabilidad de que ésta tome un valor aislado es cero. Esto se resuelve si se considera $P(X = k) \approx P\left(k - \frac{1}{2} \leq X \leq k + \frac{1}{2}\right)$

También se puede usar esta corrección para mejorar la aproximación en otros casos, específicamente en lugar de $P(X \leq k)$ calculamos

$$P(X \leq k) \approx P\left(X \leq k + \frac{1}{2}\right)$$

$$\text{Y en lugar de } P(X \geq k) \approx P\left(X \geq k - \frac{1}{2}\right)$$

En los gráficos siguientes se muestra para diferentes valores de n y p cómo aproxima la distribución $N(np, np(1-p))$ a la distribución $B(n, p)$



Ejemplos:

1- Sea $X \sim B(25, 0.4)$. Hallar las probabilidades exactas de que $X \leq 8$ y $X = 8$ y comparar estos resultados con los valores correspondientes encontrados por la aproximación normal.

Solución:

De la tabla de la *F.d.a.* de la binomial encontramos $P(X \leq 8) = 0.274$

Y $P(X = 8) = P(X \leq 8) - P(X \leq 7) = 0.274 - 0.154 = 0.120$

Ahora usamos la aproximación normal

$$P(X \leq 8) \approx P(X \leq 8.5) = P\left(\frac{X - np}{\sqrt{np(1-p)}} \leq \frac{8.5 - 10}{\sqrt{25 \times 0.4 \times 0.6}}\right) \approx \Phi(-0.61) = 0.2709$$

corrección por continuidad

Observar que el valor aproximado está muy cercano al valor exacto para $P(X \leq 8) = 0.274$

$$P(X = 8) \approx P(7.5 \leq X \leq 8.5) = P\left(\frac{7.5-10}{\sqrt{6}} \leq \frac{X-10}{\sqrt{6}} \leq \frac{8.5-10}{\sqrt{6}}\right) = P\left(-1.02 \leq \frac{X-10}{\sqrt{6}} \leq -0.61\right) = 0.2709 - 0.1593 = 0.1170$$

Nuevamente este valor aproximado está muy cerca del valor real de $P(X = 8) = 0.120$

2- Suponga que el 10% de todos los ejes de acero producidos por cierto proceso están fuera de especificaciones, pero se pueden volver a trabajar (en lugar de tener que enviarlos a la chatarra). Considere una muestra aleatoria de 200 ejes y denote por X el número entre ellos que estén fuera de especificaciones y se puedan volver a trabajar. ¿Cuál es la probabilidad (aproximada) de que X sea

- a) a lo sumo 30?
- b) menos de 30?
- c) entre 15 y 25 (inclusive)?

Solución:

Sea la v.a. X : “número de ejes fuera de especificaciones”

Entonces $X \sim B(200, 0.1)$, además $np = 200 \times 0.1 = 20 > 5$ y $n(1-p) = 200 \times (1-0.1) = 180 > 5$

Por lo tanto podemos aplicar la aproximación normal a la binomial

a) la probabilidad pedida es $P(X \leq 30)$

$$P(X \leq 30) \approx P(X \leq 30.5) = P\left(\frac{X - np}{\sqrt{np(1-p)}} \leq \frac{30.5 - 20}{\sqrt{18}}\right) \approx \Phi\left(\frac{30.5 - 20}{\sqrt{18}}\right) = \Phi(2.474) = 0.993244$$

b) La probabilidad pedida es $P(X < 30)$

Al ser X una v.a. **discreta** con distribución binomial $P(X < 30) = P(X \leq 29)$

$$P(X \leq 29) \approx P(X \leq 29.5) \approx \Phi\left(\frac{29.5 - 20}{\sqrt{18}}\right) = \Phi(2.2391) = 0.98745$$

c)

$$P(15 \leq X \leq 25) \approx P(14.5 \leq X \leq 25.5) \approx \Phi\left(\frac{25.5 - 20}{\sqrt{18}}\right) - \Phi\left(\frac{14.5 - 20}{\sqrt{18}}\right) = \Phi(1.2963) - \Phi(-1.2963) = 2\Phi(1.2963) - 1 = 2 \times 0.90147 - 1 = 0.80294$$

3- El gerente de un supermercado desea recabar información sobre la proporción de clientes a los que no les agrada una nueva política respecto de la aceptación de cheques. ¿Cuántos clientes tendría que incluir en una muestra si desea que la fracción de la muestra se desvíe a lo más en 0.15 de la verdadera fracción, con probabilidad de 0.98?

Solución:

Sea X : “número de clientes a los que no les agrada la nueva política de aceptación de cheques”

Entonces $X \sim B(n, p)$ donde p es desconocido y es la **verdadera proporción** de clientes a los que

no les agrada la nueva política de aceptación de cheques. El gerente tomará una muestra de n clientes

para “**estimar**” p con $\bar{X} = \frac{X}{n}$ ya que $\bar{X} = \frac{X}{n}$ es la proporción de clientes a los que no les agrada la

nueva política de aceptación de cheques **en la muestra de n clientes**. Si no se toman a **todos los clientes**, entonces $\bar{X} = \frac{X}{n}$ **no será igual a p** .

La pregunta es cuál debe ser n para que $\bar{X} = \frac{X}{n}$ se aleje del verdadero p en menos de 0.15 con probabilidad 0.98 por lo menos, o sea para que $P(|\bar{X} - p| \leq 0.15) \geq 0.98$

Entonces planteamos

$$P(|\bar{X} - p| \leq 0.15) = P(-0.15 \leq \bar{X} - p \leq 0.15) = P\left(\frac{-0.15n}{\sqrt{np(1-p)}} \leq \frac{X - np}{\sqrt{np(1-p)}} \leq \frac{0.15n}{\sqrt{np(1-p)}}\right) \approx \text{T.C.L.}$$

$$\approx \Phi\left(\frac{0.15n}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{-0.15n}{\sqrt{np(1-p)}}\right) = 2\Phi\left(\frac{0.15n}{\sqrt{np(1-p)}}\right) - 1 \geq 0.98$$

Por lo tanto $\Phi\left(\frac{0.15n}{\sqrt{np(1-p)}}\right) \geq \frac{0.98+1}{2} = 0.99$

Además $\frac{0.15n}{\sqrt{np(1-p)}} = \frac{0.15\sqrt{n}}{\sqrt{p(1-p)}} \geq \frac{0.15\sqrt{n}}{\sqrt{0.5(1-0.5)}} = 0.3\sqrt{n}$

Entonces debe cumplirse que $0.3\sqrt{n} \geq 2.33$ o sea $n \geq \left(\frac{2.33}{0.3}\right)^2 = 60.3211$

O sea **se debe tomar una muestra de al menos 61 clientes**

Aproximación normal a la distribución Poisson

Se puede probar aplicando Teorema central del límite que

Si $X \sim P(\lambda)$ entonces para λ suficientemente grande $\frac{X - \lambda}{\sqrt{\lambda}}$ tiene aproximadamente distribución $N(0,1)$

Es decir para λ suficientemente grande $\frac{X - \lambda}{\sqrt{\lambda}} \approx N(0,1)$

En la práctica si $\lambda \geq 30$ la aproximación es buena.

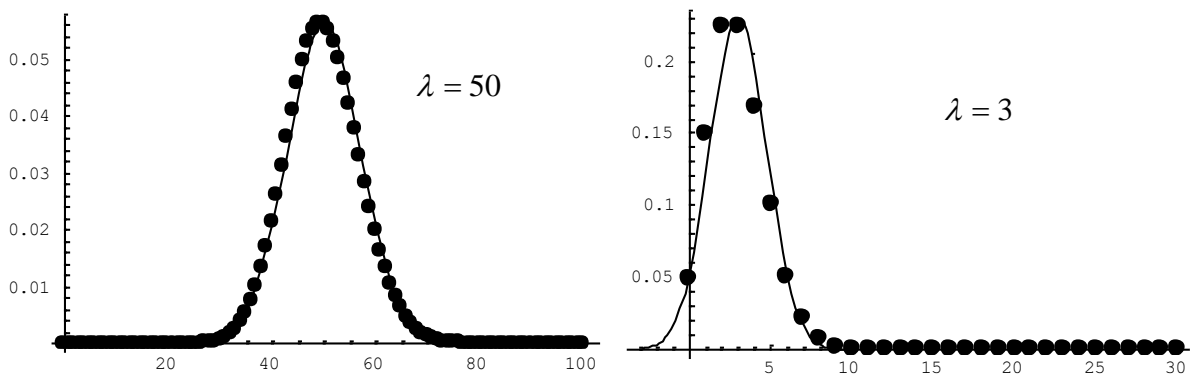
Observación: la demostración es sencilla si λ es igual a un número natural n pues, si consideramos las variables aleatorias $X_i \sim P(1)$ con $i = 1, 2, \dots, n$ independientes, entonces ya sabemos que

$$\sum_{i=1}^n X_i \sim P\left(\sum_{i=1}^n 1\right), \text{ es decir } \sum_{i=1}^n X_i \sim P(n)$$

Pero además por T.C.L. si n es grande $\sum_{i=1}^n X_i$ tiene aproximadamente distribución normal con parámetros $n\mu = n \times 1 = n$ y $n\sigma^2 = n \times 1 = n$

O sea la distribución de $\sum_{i=1}^n X_i$ que es exactamente Poisson con parámetro n , se puede aproximar con una $N(n, n)$, por lo tanto $\frac{X - n}{\sqrt{n}} \approx N(0,1)$ aproximadamente para valores de n suficientemente grandes

En los gráficos siguientes se muestra para diferentes valores de λ cómo aproxima la distribución $N(\lambda, \lambda)$ a la distribución $P(\lambda)$



Ejemplo:

El número de infracciones por estacionamiento en cierta ciudad en cualquier día hábil tiene una distribución de Poisson con parámetro $\lambda = 50$. ¿Cuál es la probabilidad aproximada de que:

- entre 35 y 70 infracciones se expidan en un día en particular?
- el número total de infracciones expedidas durante una semana de 5 días sea entre 225 y 275?

Solución:

Sea X : “número de infracciones por estacionamiento en cierta ciudad en cualquier día hábil”

Entonces $X \sim P(\lambda)$ donde $\lambda = 50$

Como $\lambda = 50$ entonces $\frac{X - 50}{\sqrt{50}} \approx N(0,1)$ (aproximadamente)

a) la probabilidad pedida es

$$P(35 \leq X \leq 70) \approx \Phi\left(\frac{70 - 50}{\sqrt{50}}\right) - \Phi\left(\frac{35 - 50}{\sqrt{50}}\right) = \Phi(2.8284) - \Phi(-2.12132) = 0.997599 - 0.017 = 0.9805$$

b) Sea Y : “número total de infracciones expedidas durante una semana de 5 días”

Entonces $Y \sim P(\lambda)$ donde $\lambda = 50 \times 5 = 250$

La probabilidad pedida es

$$P(225 \leq Y \leq 275) \approx \Phi\left(\frac{275 - 250}{\sqrt{250}}\right) - \Phi\left(\frac{225 - 250}{\sqrt{250}}\right) = \Phi(1.5811) - \Phi(-1.5811) = 2\Phi(1.5811) - 1 = 2 \times 0.94295 - 1 = 0.8859$$

Práctica

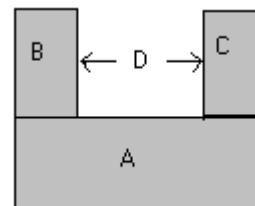
Suma de variables aleatorias - Teorema central del límite.

- 1) Un hueco cilíndrico es perforado con un bloque de acero y se fabrica un pistón cilíndrico que quepa en el hueco. El diámetro del hueco es una v.a. X con distribución normal con media 20 cm y desviación estándar 0.01 cm. El diámetro del pistón es una v.a. Y con distribución normal con media 19.90 cm y desviación estándar 0.02 cm. La holgura es otra v.a. Z y es la mitad de la diferencia entre los diámetros.

Si X e Y son independientes, ¿cuál es la probabilidad de que la holgura sea menor que 0.08 cm?

- 2) Un componente en forma de U esta formado por tres piezas A, B y C. La figura ilustra el componente.

La longitud de A tiene una distribución normal con media de 10 mm y desviación estándar de 0.1 mm. El espesor de las piezas B y C esta distribuido normalmente con media de 2 mm y desviación estándar de 0.05 mm. Suponga que todas las dimensiones son independientes.



a) Determine la media y la desviación estándar de la longitud del hueco D.

b) ¿Cuál es la probabilidad de que el hueco D sea menor que 5.9 mm?.

- 3) Una instalación de luz tiene dos focos A y B. La duración del foco A se puede considerar una v.a. X con distribución normal con media 800 hs. y desviación estándar de 100 hs. La duración del foco B se puede considerar una v.a. Y con distribución normal con media 900 hs. y desviación estándar de 150 hs. Suponga que las duraciones de los focos son independientes.

a) ¿Cuál es la probabilidad de que el foco B dure más que el foco A?.

(Sugerencia: piense cómo interpreta el evento $\{Y - X > 0\}$ y qué distribución tiene $Y - X$)

b) Otra instalación de luz tiene solo un foco. Se pone uno del tipo A y cuando se funde se instala otro de tipo B. ¿Cuál es la probabilidad de que la duración total de ambos sea mayor que 2000 hs?.

(Sugerencia: piense cómo interpreta el evento $\{Y+X > 2000\}$ y qué distribución tiene $Y + X$)

- 4) El peso de un caramelo pequeño tiene una distribución normal con media 2.835 gramos y desviación estándar de 0.2835 gramos. Suponga que se colocan 16 caramelos en un paquete y que los pesos de éstos son independientes.

a) ¿Cuáles son la media y la varianza del peso neto del paquete?

b) ¿Cuál es la probabilidad de que el peso neto del paquete sea menor que 45.5 gramos?

- 5) El tiempo para que un sistema automatizado localice una pieza en un almacén, tiene una distribución normal con media de 45 segundos y desviación estándar de 30 segundos. Suponga que se hacen pedidos independientes por 10 piezas.

a) ¿Cuál es la probabilidad de que el tiempo promedio necesario para localizar las 10 piezas sea mayor que 60 segundos? ¿Cuál es la probabilidad de que el tiempo total necesario para localizar las 10 piezas sea mayor que 600 segundos?

- 6) Suponga que la densidad del sedimento (g/cm) de un espécimen seleccionado al azar de cierta región está normalmente distribuida con media 2.65 y desviación estándar 0.85.

- a) Si se selecciona una muestra aleatoria de 25 especímenes, ¿cuál es la probabilidad de que la densidad promedio de sedimento muestral sea a lo sumo 3.00?. ¿Y entre 2.65 y 3.00?
 - b) ¿Qué tan grande se requeriría un tamaño muestral para asegurar que la primera probabilidad de la parte a) sea por lo menos 0.99?.
- 7) El centro de cálculo de una universidad dispone de un servidor para gestionar las páginas web personales de profesores y alumnos. Supongamos que la cantidad de memoria ocupada por una de estas páginas puede considerarse como una variable aleatoria con una media de 1.3 MB y una desviación estándar de 0.3. Si el servidor va a gestionar un total de 500 páginas, calcular aproximadamente la probabilidad de que la cantidad total de memoria necesaria supere los 660 MB. (Sugerencia: considerar las v.a. X_i : “cantidad de memoria ocupada por la página i ”, $i=1,2,\dots,500$)
- 8) El tiempo de vida (en horas) de un componente electrónico viene determinado por la siguiente función de densidad:
- $$f(x) = \begin{cases} 2ke^{-\frac{x}{5}} & x > 0 \\ 0 & \text{c. c.} \end{cases}$$
- a) Calcular k y la función de distribución acumulada asociada.
 - b) ¿Qué porcentaje de componentes de este tipo duran entre 2 y 10 horas?. ¿Y más de un día?. Determinar la esperanza y la varianza.
 - c) Si se consideran 40 componentes del tipo anterior, obtener aproximadamente la probabilidad de que la vida media de los 40 componentes esté comprendida entre 2 y 10 horas. (Sugerencia: considere las v.a. X_i : “duración en horas del componente electrónico i ”, $i = 1, 2, \dots, 40$)
- 9) La resistencia a la ruptura de un remache tiene un valor medio de 10000 lb/pulg² y una desviación estándar de 500 lb/pulg².
- a) ¿Cuál es la probabilidad de que la resistencia promedio a la ruptura de la muestra, para una muestra de 40 remaches, sea entre 9900 y 10200?
 - b) Si el tamaño de la muestra hubiera sido 15 en lugar de 40, ¿podría calcularse la probabilidad pedida en la parte a) a partir de la información dada?
- 10) Se procede a detener el funcionamiento de una máquina para repararla si en una muestra aleatoria de 100 artículos de la producción diaria de la máquina se encuentran por lo menos 15% de artículos defectuosos. (Suponga que la producción diaria consta de un gran número de artículos). Si realmente la máquina produce solo 10% de artículos defectuosos, encuentre la probabilidad de que se pare la máquina un día dado.
- 11) El gerente de un supermercado desea recabar información sobre la proporción de clientes a los que no les agrada una nueva política respecto de la aceptación de cheques. ¿Cuántos clientes tendría que incluir en una muestra si desea que la fracción de la muestra se desvíe a lo más en 0.15 de la verdadera fracción, con probabilidad de 0.98?.
- Responder utilizando Teorema central del límite.

1. Distribución Binomial

$$X \sim \text{Binom}(n, \theta)$$

$$p = P(X \leq x) = \sum_{k=0}^x \binom{n}{k} \theta^k (1-\theta)^{n-k} = 1 - \alpha$$

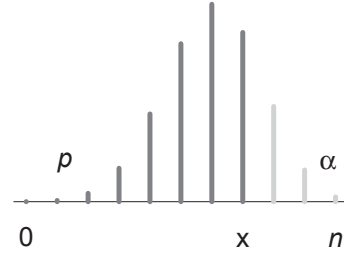


Tabla 1A. Probabilidades acumuladas p de la distribución binomial ($n = 5, 6, 7, 8, 9$).

x	θ														
	0.01	0.05	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9	0.95	0.99
$n = 5$	0	0.951	0.774	0.590	0.328	0.237	0.168	0.078	0.031	0.010	0.002	0.001	0.000	0.000	0.000
	1	0.999	0.977	0.919	0.737	0.633	0.528	0.337	0.188	0.087	0.031	0.016	0.007	0.000	0.000
	2	1.000	0.999	0.991	0.942	0.896	0.837	0.683	0.500	0.317	0.163	0.104	0.058	0.009	0.001
	3	1.000	1.000	1.000	0.993	0.984	0.969	0.913	0.813	0.663	0.472	0.367	0.263	0.081	0.023
	4	1.000	1.000	1.000	1.000	0.999	0.998	0.990	0.969	0.922	0.832	0.763	0.672	0.410	0.226
	5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 6$	0	0.941	0.735	0.531	0.262	0.178	0.118	0.047	0.016	0.004	0.001	0.000	0.000	0.000	0.000
	1	0.999	0.967	0.886	0.655	0.534	0.420	0.233	0.109	0.041	0.011	0.005	0.002	0.000	0.000
	2	1.000	0.998	0.984	0.901	0.831	0.744	0.544	0.344	0.179	0.070	0.038	0.017	0.001	0.000
	3	1.000	1.000	0.999	0.983	0.962	0.930	0.821	0.656	0.456	0.256	0.169	0.099	0.016	0.002
	4	1.000	1.000	1.000	0.998	0.995	0.989	0.959	0.891	0.767	0.580	0.466	0.345	0.114	0.033
	5	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.984	0.953	0.882	0.822	0.738	0.469	0.265
	6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 7$	0	0.932	0.698	0.478	0.210	0.133	0.082	0.028	0.008	0.002	0.000	0.000	0.000	0.000	0.000
	1	0.998	0.956	0.850	0.577	0.445	0.329	0.159	0.063	0.019	0.004	0.001	0.000	0.000	0.000
	2	1.000	0.996	0.974	0.852	0.756	0.647	0.420	0.227	0.096	0.029	0.013	0.005	0.000	0.000
	3	1.000	1.000	0.997	0.967	0.929	0.874	0.710	0.500	0.290	0.126	0.071	0.033	0.003	0.000
	4	1.000	1.000	1.000	0.995	0.987	0.971	0.904	0.773	0.580	0.353	0.244	0.148	0.026	0.004
	5	1.000	1.000	1.000	1.000	0.999	0.996	0.981	0.938	0.841	0.671	0.555	0.423	0.150	0.044
	6	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.992	0.972	0.918	0.867	0.790	0.522	0.302
	7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 8$	0	0.923	0.663	0.430	0.168	0.100	0.058	0.017	0.004	0.001	0.000	0.000	0.000	0.000	0.000
	1	0.997	0.943	0.813	0.503	0.367	0.255	0.106	0.035	0.009	0.001	0.000	0.000	0.000	0.000
	2	1.000	0.994	0.962	0.797	0.679	0.552	0.315	0.145	0.050	0.011	0.004	0.001	0.000	0.000
	3	1.000	1.000	0.995	0.944	0.886	0.806	0.594	0.363	0.174	0.058	0.027	0.010	0.000	0.000
	4	1.000	1.000	1.000	0.990	0.973	0.942	0.826	0.637	0.406	0.194	0.114	0.056	0.005	0.000
	5	1.000	1.000	1.000	0.999	0.996	0.989	0.950	0.855	0.685	0.448	0.321	0.203	0.038	0.006
	6	1.000	1.000	1.000	1.000	1.000	0.999	0.991	0.965	0.894	0.745	0.633	0.497	0.187	0.057
	7	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.983	0.942	0.900	0.832	0.570	0.337
	8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 9$	0	0.914	0.630	0.387	0.134	0.075	0.040	0.010	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.997	0.929	0.775	0.436	0.300	0.196	0.071	0.020	0.004	0.000	0.000	0.000	0.000	0.000
	2	1.000	0.992	0.947	0.738	0.601	0.463	0.232	0.090	0.025	0.004	0.001	0.000	0.000	0.000
	3	1.000	0.999	0.992	0.914	0.834	0.730	0.483	0.254	0.099	0.025	0.010	0.003	0.000	0.000
	4	1.000	1.000	0.999	0.980	0.951	0.901	0.733	0.500	0.267	0.099	0.049	0.020	0.001	0.000
	5	1.000	1.000	1.000	0.997	0.990	0.975	0.901	0.746	0.517	0.270	0.166	0.086	0.008	0.001
	6	1.000	1.000	1.000	1.000	0.999	0.996	0.975	0.910	0.768	0.537	0.399	0.262	0.053	0.008
	7	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.980	0.929	0.804	0.700	0.564	0.225	0.071
	8	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.990	0.960	0.925	0.866	0.613	0.370	0.086
	9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tabla 1B. Probabilidades acumuladas p de la distribución binomial ($n = 10, 11, 12, 13, 14$).

x	θ															
		0.01	0.05	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9	0.95	0.99
$n = 10$	0	0.904	0.599	0.349	0.107	0.056	0.028	0.006	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.996	0.914	0.736	0.376	0.244	0.149	0.046	0.011	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	2	1.000	0.988	0.930	0.678	0.526	0.383	0.167	0.055	0.012	0.002	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.999	0.987	0.879	0.776	0.650	0.382	0.172	0.055	0.011	0.004	0.001	0.000	0.000	0.000
	4	1.000	1.000	0.998	0.967	0.922	0.850	0.633	0.377	0.166	0.047	0.020	0.006	0.000	0.000	0.000
	5	1.000	1.000	1.000	0.994	0.980	0.953	0.834	0.623	0.367	0.150	0.078	0.033	0.002	0.000	0.000
	6	1.000	1.000	1.000	0.999	0.996	0.989	0.945	0.828	0.618	0.350	0.224	0.121	0.013	0.001	0.000
	7	1.000	1.000	1.000	1.000	1.000	0.998	0.988	0.945	0.833	0.617	0.474	0.322	0.070	0.012	0.000
	8	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.989	0.954	0.851	0.756	0.624	0.264	0.086	0.004
	9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.972	0.944	0.893	0.651	0.401	0.096
	10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 11$	0	0.895	0.569	0.314	0.086	0.042	0.020	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.995	0.898	0.697	0.322	0.197	0.113	0.030	0.006	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	2	1.000	0.985	0.910	0.617	0.455	0.313	0.119	0.033	0.006	0.001	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.998	0.981	0.839	0.713	0.570	0.296	0.113	0.029	0.004	0.001	0.000	0.000	0.000	0.000
	4	1.000	1.000	0.997	0.950	0.885	0.790	0.533	0.274	0.099	0.022	0.008	0.002	0.000	0.000	0.000
	5	1.000	1.000	1.000	0.988	0.966	0.922	0.753	0.500	0.247	0.078	0.034	0.012	0.000	0.000	0.000
	6	1.000	1.000	1.000	0.998	0.992	0.978	0.901	0.726	0.467	0.210	0.115	0.050	0.003	0.000	0.000
	7	1.000	1.000	1.000	1.000	0.999	0.996	0.971	0.887	0.704	0.430	0.287	0.161	0.019	0.002	0.000
	8	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.967	0.881	0.687	0.545	0.383	0.090	0.015	0.000
	9	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.970	0.887	0.803	0.678	0.303	0.102	0.005
	10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.980	0.958	0.914	0.686	0.431	0.105
	11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 12$	0	0.886	0.540	0.282	0.069	0.032	0.014	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.994	0.882	0.659	0.275	0.158	0.085	0.020	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	1.000	0.980	0.889	0.558	0.391	0.253	0.083	0.019	0.003	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.998	0.974	0.795	0.649	0.493	0.225	0.073	0.015	0.002	0.000	0.000	0.000	0.000	0.000
	4	1.000	1.000	0.996	0.927	0.842	0.724	0.438	0.194	0.057	0.009	0.003	0.001	0.000	0.000	0.000
	5	1.000	1.000	0.999	0.981	0.946	0.882	0.665	0.387	0.158	0.039	0.014	0.004	0.000	0.000	0.000
	6	1.000	1.000	1.000	0.996	0.986	0.961	0.842	0.613	0.335	0.118	0.054	0.019	0.001	0.000	0.000
	7	1.000	1.000	1.000	0.999	0.997	0.991	0.943	0.806	0.562	0.276	0.158	0.073	0.004	0.000	0.000
	8	1.000	1.000	1.000	1.000	1.000	0.998	0.985	0.927	0.775	0.507	0.351	0.205	0.026	0.002	0.000
	9	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.981	0.917	0.747	0.609	0.442	0.111	0.020	0.000
	10	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.980	0.915	0.842	0.725	0.341	0.118	0.006	0.000
	11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.986	0.968	0.931	0.718	0.460	0.114	0.000
	12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 13$	0	0.878	0.513	0.254	0.055	0.024	0.010	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.993	0.865	0.621	0.234	0.127	0.064	0.013	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	1.000	0.975	0.866	0.502	0.333	0.202	0.058	0.011	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.997	0.966	0.747	0.584	0.421	0.169	0.046	0.008	0.001	0.000	0.000	0.000	0.000	0.000
	4	1.000	1.000	0.994	0.901	0.794	0.654	0.353	0.133	0.032	0.004	0.001	0.000	0.000	0.000	0.000
	5	1.000	1.000	0.999	0.970	0.920	0.835	0.574	0.291	0.098	0.018	0.006	0.001	0.000	0.000	0.000
	6	1.000	1.000	1.000	0.993	0.976	0.938	0.771	0.500	0.229	0.062	0.024	0.007	0.000	0.000	0.000
	7	1.000	1.000	1.000	0.999	0.994	0.982	0.902	0.709	0.426	0.165	0.080	0.030	0.001	0.000	0.000
	8	1.000	1.000	1.000	1.000	0.999	0.996	0.968	0.867	0.647	0.346	0.206	0.099	0.006	0.000	0.000
	9	1.000	1.000	1.000	1.000	1.000	0.999	0.992	0.954	0.831	0.579	0.416	0.253	0.034	0.003	0.000
	10	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.989	0.942	0.798	0.667	0.498	0.134	0.025	0.000
	11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.987	0.936	0.873	0.766	0.379	0.135	0.007
	12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.990	0.976	0.945	0.746	0.487	0.122	0.000
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 14$	0	0.869	0.488	0.229	0.044	0.018	0.007	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.992	0.847	0.585	0.198	0.101	0.047	0.008	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	1.000	0.970	0.842	0.448	0.281	0.161	0.040	0.006	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.996	0.956	0.698	0.521	0.355	0.124	0.029	0.004	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	1.000	0.991	0.870	0.742	0.584	0.279	0.090	0.018	0.002	0.000	0.000	0.000	0.000	0.000
	5	1.000	1.000	0.999	0.956	0.888	0.781	0.486	0.212	0.058	0.008	0.002	0.000	0.000	0.000	0.000
	6	1.000	1.000	1.000	0.988	0.962	0.907	0.692	0.395	0.150	0.031	0.010	0.002	0.000	0.000	0.000
	7	1.000	1.000	1.000	0.998	0.990	0.969	0.850	0.605	0.308	0.093	0.038	0.012	0.000	0.000	0.000
	8	1.000	1.000	1.000	1.000	0.998	0.992	0.942	0.788	0.514	0.219	0.112	0.044	0.001	0.000	0.000
	9	1.000	1.000	1.000	1.000	1.000	0.998	0.982	0.910	0.721	0.416	0.258	0.130	0.009	0.000	0.000
	10	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.971	0.876	0.645	0.479	0.302	0.044	0.004	0.000
	11	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.960	0.839	0.719	0.552	0.158	0.030	0.000
	12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.992	0.953	0.899	0.802	0.415	0.153	0.008
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.993	0.982	0.956	0.771	0.512	0.131
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tabla 1C. Probabilidades acumuladas p de la distribución binomial ($n = 15, 16, 17, 18$).

		θ														
	x	0.01	0.05	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9	0.95	0.99
$n = 15$	0	0.860	0.463	0.206	0.035	0.013	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.990	0.829	0.549	0.167	0.080	0.035	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	1.000	0.964	0.816	0.398	0.236	0.127	0.027	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.995	0.944	0.648	0.461	0.297	0.091	0.018	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.999	0.987	0.836	0.686	0.515	0.217	0.059	0.009	0.001	0.000	0.000	0.000	0.000	0.000
	5	1.000	1.000	0.998	0.939	0.852	0.722	0.403	0.151	0.034	0.004	0.001	0.000	0.000	0.000	0.000
	6	1.000	1.000	1.000	0.982	0.943	0.869	0.610	0.304	0.095	0.015	0.004	0.001	0.000	0.000	0.000
	7	1.000	1.000	1.000	0.996	0.983	0.950	0.787	0.500	0.213	0.050	0.017	0.004	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.999	0.996	0.985	0.905	0.696	0.390	0.131	0.057	0.018	0.000	0.000	0.000
	9	1.000	1.000	1.000	1.000	0.999	0.996	0.966	0.849	0.597	0.278	0.148	0.061	0.002	0.000	0.000
	10	1.000	1.000	1.000	1.000	1.000	0.999	0.991	0.941	0.783	0.485	0.314	0.164	0.013	0.001	0.000
	11	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.982	0.909	0.703	0.539	0.352	0.056	0.005	0.000
	12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.973	0.873	0.764	0.602	0.184	0.036	0.000
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.965	0.920	0.833	0.451	0.171	0.010
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.987	0.965	0.794	0.537	0.140
	15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 16$	0	0.851	0.440	0.185	0.028	0.010	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.989	0.811	0.515	0.141	0.063	0.026	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.999	0.957	0.789	0.352	0.197	0.099	0.018	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.993	0.932	0.598	0.405	0.246	0.065	0.011	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.999	0.983	0.798	0.630	0.450	0.167	0.038	0.005	0.000	0.000	0.000	0.000	0.000	0.000
	5	1.000	1.000	0.997	0.918	0.810	0.660	0.329	0.105	0.019	0.002	0.000	0.000	0.000	0.000	0.000
	6	1.000	1.000	0.999	0.973	0.920	0.825	0.527	0.227	0.058	0.007	0.002	0.000	0.000	0.000	0.000
	7	1.000	1.000	1.000	0.993	0.973	0.926	0.716	0.402	0.142	0.026	0.007	0.001	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.999	0.993	0.974	0.858	0.598	0.284	0.074	0.027	0.007	0.000	0.000	0.000
	9	1.000	1.000	1.000	1.000	0.998	0.993	0.942	0.773	0.473	0.175	0.080	0.027	0.001	0.000	0.000
	10	1.000	1.000	1.000	1.000	1.000	0.998	0.981	0.895	0.671	0.340	0.190	0.082	0.003	0.000	0.000
	11	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.962	0.833	0.550	0.370	0.202	0.017	0.001	0.000
	12	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.989	0.935	0.754	0.595	0.402	0.068	0.007	0.000
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.982	0.901	0.803	0.648	0.211	0.043	0.001
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.974	0.937	0.859	0.485	0.189	0.011
	15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.990	0.972	0.815	0.560	0.149
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
$n = 17$	0	0.843	0.418	0.167	0.023	0.008	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.988	0.792	0.482	0.118	0.050	0.019	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.999	0.950	0.762	0.310	0.164	0.077	0.012	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.991	0.917	0.549	0.353	0.202	0.046	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.999	0.978	0.758	0.574	0.389	0.126	0.025	0.003	0.000	0.000	0.000	0.000	0.000	0.000
	5	1.000	1.000	0.995	0.894	0.765	0.597	0.264	0.072	0.011	0.001	0.000	0.000	0.000	0.000	0.000
	6	1.000	1.000	0.999	0.962	0.893	0.775	0.448	0.166	0.035	0.003	0.001	0.000	0.000	0.000	0.000
	7	1.000	1.000	1.000	0.989	0.960	0.895	0.641	0.315	0.092	0.013	0.003	0.000	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.997	0.988	0.960	0.801	0.500	0.199	0.040	0.012	0.003	0.000	0.000	0.000
	9	1.000	1.000	1.000	1.000	0.997	0.987	0.908	0.685	0.359	0.105	0.040	0.011	0.000	0.000	0.000
	10	1.000	1.000	1.000	1.000	0.999	0.997	0.965	0.834	0.552	0.225	0.107	0.038	0.001	0.000	0.000
	11	1.000	1.000	1.000	1.000	1.000	0.999	0.989	0.928	0.736	0.403	0.235	0.106	0.005	0.000	0.000
	12	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.975	0.874	0.611	0.426	0.242	0.022	0.001	0.000
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.954	0.798	0.647	0.451	0.083	0.009	0.000
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.988	0.923	0.836	0.690	0.238	0.050	0.001
	15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.981	0.950	0.882	0.518	0.208	0.012
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.992	0.977	0.833	0.582	0.157
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
$n = 18$	0	0.835	0.397	0.150	0.018	0.006	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.986	0.774	0.450	0.099	0.039	0.014	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.999	0.942	0.734	0.271	0.135	0.060	0.008	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.989	0.902	0.501	0.306	0.165	0.033	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.998	0.972	0.716	0.519	0.333	0.094	0.015	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	5	1.000	1.000	0.994	0.867	0.717	0.534	0.209	0.048	0.006	0.000	0.000	0.000	0.000	0.000	0.000
	6	1.000	1.000	0.999	0.949	0.861	0.722	0.374	0.119	0.020	0.001	0.000	0.000	0.000	0.000	0.000
	7	1.000	1.000	1.000	0.984	0.943	0.859	0.563	0.240	0.058	0.006	0.001	0.000	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.996	0.981	0.940	0.737	0.407	0.135	0.021	0.005	0.001	0.000	0.000	0.000
	9	1.000	1.000	1.000	0.999	0.995	0.979	0.865	0.593	0.263	0.060	0.019	0.004	0.000	0.000	0.000
	10	1.000	1.000	1.000	1.000	0.999	0.994	0.942	0.760	0.437	0.141	0.057	0.016	0.000	0.000	0.000
	11	1.000	1.000	1.000	1.000	1.000	0.999	0.980	0.881	0.626	0.278	0.139	0.051	0.001	0.000	0.000
	12	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.952	0.791	0.466	0.283	0.133	0.006	0.000	0.000
	13	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.985	0.906	0.667	0.481	0.284	0.028	0.002	0.000
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.967	0.835	0.694	0.499	0.098	0.011	0.000
	15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.992	0.940	0.865	0.729	0.266	0.058	0.001
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.986	0.961	0.901	0.550	0.226	0.014
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.994	0.982	0.850	0.603	0.165
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	

Tabla 1D. Probabilidades acumuladas p de la distribución binomial ($n = 19, 20, 21$).

x	θ														
	0.01	0.05	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9	0.95	0.99
$n = 19$	0	0.826	0.377	0.135	0.014	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.985	0.755	0.420	0.083	0.031	0.010	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.999	0.933	0.705	0.237	0.111	0.046	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.987	0.885	0.455	0.263	0.133	0.023	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.998	0.965	0.673	0.465	0.282	0.070	0.010	0.001	0.000	0.000	0.000	0.000	0.000
	5	1.000	1.000	0.991	0.837	0.668	0.474	0.163	0.032	0.003	0.000	0.000	0.000	0.000	0.000
	6	1.000	1.000	0.998	0.932	0.825	0.666	0.308	0.084	0.012	0.001	0.000	0.000	0.000	0.000
	7	1.000	1.000	1.000	0.977	0.923	0.818	0.488	0.180	0.035	0.003	0.000	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.993	0.971	0.916	0.667	0.324	0.088	0.011	0.002	0.000	0.000	0.000
	9	1.000	1.000	1.000	0.998	0.991	0.967	0.814	0.500	0.186	0.033	0.009	0.002	0.000	0.000
	10	1.000	1.000	1.000	1.000	0.998	0.989	0.912	0.676	0.333	0.084	0.029	0.007	0.000	0.000
	11	1.000	1.000	1.000	1.000	1.000	0.997	0.965	0.820	0.512	0.182	0.077	0.023	0.000	0.000
	12	1.000	1.000	1.000	1.000	1.000	0.999	0.988	0.916	0.692	0.334	0.175	0.068	0.002	0.000
	13	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.968	0.837	0.526	0.332	0.163	0.009	0.000
	14	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.990	0.930	0.718	0.535	0.327	0.035	0.002
	15	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.977	0.867	0.737	0.545	0.115	0.013	0.000
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.954	0.889	0.763	0.295	0.067	0.001
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.990	0.969	0.917	0.580	0.245	0.015
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.986	0.865	0.623	0.174
	19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 20$	0	0.818	0.358	0.122	0.012	0.003	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.983	0.736	0.392	0.069	0.024	0.008	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.999	0.925	0.677	0.206	0.091	0.035	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.984	0.867	0.411	0.225	0.107	0.016	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.997	0.957	0.630	0.415	0.238	0.051	0.006	0.000	0.000	0.000	0.000	0.000	0.000
	5	1.000	1.000	0.989	0.804	0.617	0.416	0.126	0.021	0.002	0.000	0.000	0.000	0.000	0.000
	6	1.000	1.000	0.998	0.913	0.786	0.608	0.250	0.058	0.006	0.000	0.000	0.000	0.000	0.000
	7	1.000	1.000	1.000	0.968	0.898	0.772	0.416	0.132	0.021	0.001	0.000	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.990	0.959	0.887	0.596	0.252	0.057	0.005	0.001	0.000	0.000	0.000
	9	1.000	1.000	1.000	0.997	0.986	0.952	0.755	0.412	0.128	0.017	0.004	0.001	0.000	0.000
	10	1.000	1.000	1.000	0.999	0.996	0.983	0.872	0.588	0.245	0.048	0.014	0.003	0.000	0.000
	11	1.000	1.000	1.000	1.000	0.999	0.995	0.943	0.748	0.404	0.113	0.041	0.010	0.000	0.000
	12	1.000	1.000	1.000	1.000	1.000	0.999	0.979	0.868	0.584	0.228	0.102	0.032	0.000	0.000
	13	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.942	0.750	0.392	0.214	0.087	0.002	0.000
	14	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.979	0.874	0.584	0.383	0.196	0.011	0.000
	15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.949	0.762	0.585	0.370	0.043	0.003
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.984	0.893	0.775	0.589	0.133	0.016
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.965	0.909	0.794	0.323	0.075	0.001
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.992	0.976	0.931	0.608	0.264	0.017
	19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.997	0.988	0.878	0.642	0.182
	20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 21$	0	0.810	0.341	0.109	0.009	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.981	0.717	0.365	0.058	0.019	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.999	0.915	0.648	0.179	0.075	0.027	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.981	0.848	0.370	0.192	0.086	0.011	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.997	0.948	0.586	0.367	0.198	0.037	0.004	0.000	0.000	0.000	0.000	0.000	0.000
	5	1.000	1.000	0.986	0.769	0.567	0.363	0.096	0.013	0.001	0.000	0.000	0.000	0.000	0.000
	6	1.000	1.000	0.997	0.891	0.744	0.551	0.200	0.039	0.004	0.000	0.000	0.000	0.000	0.000
	7	1.000	1.000	0.999	0.957	0.870	0.723	0.350	0.095	0.012	0.001	0.000	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.986	0.944	0.852	0.524	0.192	0.035	0.002	0.000	0.000	0.000	0.000
	9	1.000	1.000	1.000	0.996	0.979	0.932	0.691	0.332	0.085	0.009	0.002	0.000	0.000	0.000
	10	1.000	1.000	1.000	0.999	0.994	0.974	0.826	0.500	0.174	0.026	0.006	0.001	0.000	0.000
	11	1.000	1.000	1.000	1.000	0.998	0.991	0.915	0.668	0.309	0.068	0.021	0.004	0.000	0.000
	12	1.000	1.000	1.000	1.000	1.000	0.998	0.965	0.808	0.476	0.148	0.056	0.014	0.000	0.000
	13	1.000	1.000	1.000	1.000	1.000	0.999	0.988	0.905	0.650	0.277	0.130	0.043	0.001	0.000
	14	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.961	0.800	0.449	0.256	0.109	0.003	0.000
	15	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.987	0.904	0.637	0.433	0.231	0.014	0.000
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.963	0.802	0.633	0.414	0.052	0.003
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.989	0.914	0.808	0.630	0.152	0.019
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.973	0.925	0.821	0.352	0.085	0.001
	19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.981	0.942	0.635	0.283	0.019
	20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.991	0.891	0.659	0.190
	21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tabla 1E. Probabilidades acumuladas de la distribución binomial ($n = 22, 23$).

x	θ														
	0.01	0.05	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9	0.95	0.99
$n = 22$	0	0.802	0.324	0.098	0.007	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.980	0.698	0.339	0.048	0.015	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.999	0.905	0.620	0.154	0.061	0.021	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.978	0.828	0.332	0.162	0.068	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.996	0.938	0.543	0.323	0.165	0.027	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	5	1.000	0.999	0.982	0.733	0.517	0.313	0.072	0.008	0.000	0.000	0.000	0.000	0.000	0.000
	6	1.000	1.000	0.996	0.867	0.699	0.494	0.158	0.026	0.002	0.000	0.000	0.000	0.000	0.000
	7	1.000	1.000	0.999	0.944	0.838	0.671	0.290	0.067	0.007	0.000	0.000	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.980	0.925	0.814	0.454	0.143	0.021	0.001	0.000	0.000	0.000	0.000
	9	1.000	1.000	1.000	0.994	0.970	0.908	0.624	0.262	0.055	0.004	0.001	0.000	0.000	0.000
	10	1.000	1.000	1.000	0.998	0.990	0.961	0.772	0.416	0.121	0.014	0.003	0.000	0.000	0.000
	11	1.000	1.000	1.000	1.000	0.997	0.986	0.879	0.584	0.228	0.039	0.010	0.002	0.000	0.000
	12	1.000	1.000	1.000	1.000	0.999	0.996	0.945	0.738	0.376	0.092	0.030	0.006	0.000	0.000
	13	1.000	1.000	1.000	1.000	1.000	0.999	0.979	0.857	0.546	0.186	0.075	0.020	0.000	0.000
	14	1.000	1.000	1.000	1.000	1.000	1.000	0.993	0.933	0.710	0.329	0.162	0.056	0.001	0.000
	15	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.974	0.842	0.506	0.301	0.133	0.004	0.000
	16	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.992	0.928	0.687	0.483	0.267	0.018	0.001
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.973	0.835	0.677	0.457	0.062	0.004
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.992	0.932	0.838	0.668	0.172	0.022	0.000
	19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.979	0.939	0.846	0.380	0.095	0.001
	20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.985	0.952	0.661	0.302	0.020
	21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.993	0.902	0.676	0.198
	22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 23$	0	0.794	0.307	0.089	0.006	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.978	0.679	0.315	0.040	0.012	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.998	0.895	0.592	0.133	0.049	0.016	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.974	0.807	0.297	0.137	0.054	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.995	0.927	0.501	0.283	0.136	0.019	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	5	1.000	0.999	0.977	0.695	0.468	0.269	0.054	0.005	0.000	0.000	0.000	0.000	0.000	0.000
	6	1.000	1.000	0.994	0.840	0.654	0.440	0.124	0.017	0.001	0.000	0.000	0.000	0.000	0.000
	7	1.000	1.000	0.999	0.928	0.804	0.618	0.237	0.047	0.004	0.000	0.000	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.973	0.904	0.771	0.388	0.105	0.013	0.001	0.000	0.000	0.000	0.000
	9	1.000	1.000	1.000	0.991	0.959	0.880	0.556	0.202	0.035	0.002	0.000	0.000	0.000	0.000
	10	1.000	1.000	1.000	0.997	0.985	0.945	0.713	0.339	0.081	0.007	0.001	0.000	0.000	0.000
	11	1.000	1.000	1.000	0.999	0.995	0.979	0.836	0.500	0.164	0.021	0.005	0.001	0.000	0.000
	12	1.000	1.000	1.000	1.000	0.999	0.993	0.919	0.661	0.287	0.055	0.015	0.003	0.000	0.000
	13	1.000	1.000	1.000	1.000	1.000	0.998	0.965	0.798	0.444	0.120	0.041	0.009	0.000	0.000
	14	1.000	1.000	1.000	1.000	1.000	0.999	0.987	0.895	0.612	0.229	0.096	0.027	0.000	0.000
	15	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.953	0.763	0.382	0.196	0.072	0.001	0.000
	16	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.983	0.876	0.560	0.346	0.160	0.006	0.000
	17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.946	0.731	0.532	0.305	0.023	0.001
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.981	0.864	0.717	0.499	0.073	0.005
	19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.946	0.863	0.703	0.193	0.026	0.000
	20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.984	0.951	0.867	0.408	0.105	0.002
	21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.988	0.960	0.685	0.321	0.022
	22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.994	0.911	0.693	0.206
	23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tabla 1F. Probabilidades acumuladas de la distribución binomial ($n = 24, 25$).

x	θ														
	0.01	0.05	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9	0.95	0.99
$n = 24$	0	0.786	0.292	0.080	0.005	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.976	0.661	0.292	0.033	0.009	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.998	0.884	0.564	0.115	0.040	0.012	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.970	0.786	0.264	0.115	0.042	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.994	0.915	0.460	0.247	0.111	0.013	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	5	1.000	0.999	0.972	0.656	0.422	0.229	0.040	0.003	0.000	0.000	0.000	0.000	0.000	0.000
	6	1.000	1.000	0.993	0.811	0.607	0.389	0.096	0.011	0.001	0.000	0.000	0.000	0.000	0.000
	7	1.000	1.000	0.998	0.911	0.766	0.565	0.192	0.032	0.002	0.000	0.000	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.964	0.879	0.725	0.328	0.076	0.008	0.000	0.000	0.000	0.000	0.000
	9	1.000	1.000	1.000	0.987	0.945	0.847	0.489	0.154	0.022	0.001	0.000	0.000	0.000	0.000
	10	1.000	1.000	1.000	0.996	0.979	0.926	0.650	0.271	0.053	0.004	0.001	0.000	0.000	0.000
	11	1.000	1.000	1.000	0.999	0.993	0.969	0.787	0.419	0.114	0.012	0.002	0.000	0.000	0.000
	12	1.000	1.000	1.000	1.000	0.998	0.988	0.886	0.581	0.213	0.031	0.007	0.001	0.000	0.000
	13	1.000	1.000	1.000	1.000	0.999	0.996	0.947	0.729	0.350	0.074	0.021	0.004	0.000	0.000
	14	1.000	1.000	1.000	1.000	1.000	0.999	0.978	0.846	0.511	0.153	0.055	0.013	0.000	0.000
	15	1.000	1.000	1.000	1.000	1.000	1.000	0.992	0.924	0.672	0.275	0.121	0.036	0.000	0.000
	16	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.968	0.808	0.435	0.234	0.089	0.002	0.000
	17	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.989	0.904	0.611	0.393	0.189	0.007	0.000
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.960	0.771	0.578	0.344	0.028	0.001
	19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.987	0.889	0.753	0.540	0.085	0.006
	20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.958	0.885	0.736	0.214	0.030	0.000
	21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.988	0.960	0.885	0.436	0.116	0.002
	22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.991	0.967	0.708	0.339	0.024
	23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.995	0.920	0.708	0.214
	24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$n = 25$	0	0.778	0.277	0.072	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.974	0.642	0.271	0.027	0.007	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	2	0.998	0.873	0.537	0.098	0.032	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	1.000	0.966	0.764	0.234	0.096	0.033	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	4	1.000	0.993	0.902	0.421	0.214	0.090	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	5	1.000	0.999	0.967	0.617	0.378	0.193	0.029	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	6	1.000	1.000	0.991	0.780	0.561	0.341	0.074	0.007	0.000	0.000	0.000	0.000	0.000	0.000
	7	1.000	1.000	0.998	0.891	0.727	0.512	0.154	0.022	0.001	0.000	0.000	0.000	0.000	0.000
	8	1.000	1.000	1.000	0.953	0.851	0.677	0.274	0.054	0.004	0.000	0.000	0.000	0.000	0.000
	9	1.000	1.000	1.000	0.983	0.929	0.811	0.425	0.115	0.013	0.000	0.000	0.000	0.000	0.000
	10	1.000	1.000	1.000	0.994	0.970	0.902	0.586	0.212	0.034	0.002	0.000	0.000	0.000	0.000
	11	1.000	1.000	1.000	0.998	0.989	0.956	0.732	0.345	0.078	0.006	0.001	0.000	0.000	0.000
	12	1.000	1.000	1.000	1.000	0.997	0.983	0.846	0.500	0.154	0.017	0.003	0.000	0.000	0.000
	13	1.000	1.000	1.000	1.000	0.999	0.994	0.922	0.655	0.268	0.044	0.011	0.002	0.000	0.000
	14	1.000	1.000	1.000	1.000	1.000	0.998	0.966	0.788	0.414	0.098	0.030	0.006	0.000	0.000
	15	1.000	1.000	1.000	1.000	1.000	1.000	0.987	0.885	0.575	0.189	0.071	0.017	0.000	0.000
	16	1.000	1.000	1.000	1.000	1.000	1.000	0.996	0.946	0.726	0.323	0.149	0.047	0.000	0.000
	17	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.978	0.846	0.488	0.273	0.109	0.002	0.000
	18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.993	0.926	0.659	0.439	0.220	0.009	0.000
	19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.971	0.807	0.622	0.383	0.033	0.001
	20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.991	0.910	0.786	0.579	0.098	0.007	0.000
	21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.967	0.904	0.766	0.236	0.034	0.000
	22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.991	0.968	0.902	0.463	0.127	0.002
	23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.993	0.973	0.729	0.358	0.026
	24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.928	0.723	0.222
	24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

2. Distribución Poisson

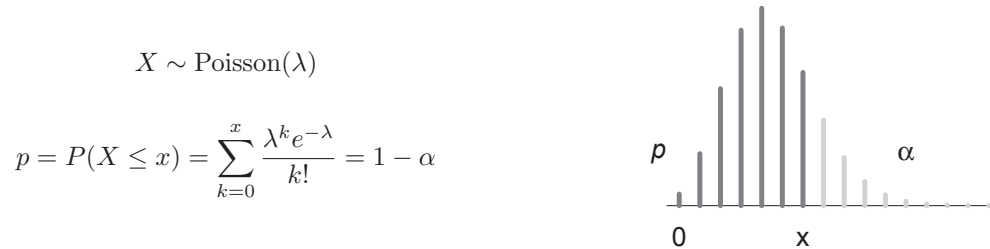


Tabla 2A. Probabilidades acumuladas p de la distribución Poisson.

x	λ									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.905	0.819	0.741	0.670	0.607	0.549	0.497	0.449	0.407	0.368
1	0.995	0.982	0.963	0.938	0.910	0.878	0.844	0.809	0.772	0.736
2	1.000	0.999	0.996	0.992	0.986	0.977	0.966	0.953	0.937	0.920
3	1.000	1.000	1.000	0.999	0.998	0.997	0.994	0.991	0.987	0.981
4	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.998	0.996
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Tabla 2B. Probabilidades acumuladas p de la distribución Poisson.

x	λ										
	2	3	4	5	6	7	8	9	10	15	20
0	0.135	0.050	0.018	0.007	0.002	0.001	0.000	0.000	0.000	0.000	0.000
1	0.406	0.199	0.092	0.040	0.017	0.007	0.003	0.001	0.000	0.000	0.000
2	0.677	0.423	0.238	0.125	0.062	0.030	0.014	0.006	0.003	0.000	0.000
3	0.857	0.647	0.433	0.265	0.151	0.082	0.042	0.021	0.010	0.000	0.000
4	0.947	0.815	0.629	0.440	0.285	0.173	0.100	0.055	0.029	0.001	0.000
5	0.983	0.916	0.785	0.616	0.446	0.301	0.191	0.116	0.067	0.003	0.000
6	0.995	0.966	0.889	0.762	0.606	0.450	0.313	0.207	0.130	0.008	0.000
7	0.999	0.988	0.949	0.867	0.744	0.599	0.453	0.324	0.220	0.018	0.001
8	1.000	0.996	0.979	0.932	0.847	0.729	0.593	0.456	0.333	0.037	0.002
9	1.000	0.999	0.992	0.968	0.916	0.830	0.717	0.587	0.458	0.070	0.005
10	1.000	1.000	0.997	0.986	0.957	0.901	0.816	0.706	0.583	0.118	0.011
11	1.000	1.000	0.999	0.995	0.980	0.947	0.888	0.803	0.697	0.185	0.021
12	1.000	1.000	1.000	0.998	0.991	0.973	0.936	0.876	0.792	0.268	0.039
13	1.000	1.000	1.000	0.999	0.996	0.987	0.966	0.926	0.864	0.363	0.066
14	1.000	1.000	1.000	1.000	0.999	0.994	0.983	0.959	0.917	0.466	0.105
15	1.000	1.000	1.000	1.000	0.999	0.998	0.992	0.978	0.951	0.568	0.157
16	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.989	0.973	0.664	0.221
17	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.995	0.986	0.749	0.297
18	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.993	0.819	0.381
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.997	0.875	0.470
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.917	0.559
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.947	0.644
22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.967	0.721
23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.981	0.787
24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.989	0.843
25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.994	0.888
26	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.922
27	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.948
28	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.966
29	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.978
30	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.987
31	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.992
32	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.995
33	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.997
34	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
35	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999
36	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

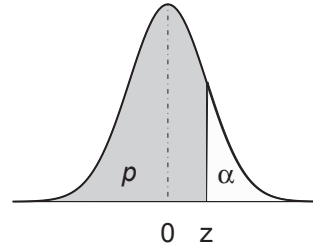
3. Distribución Normal Estándar

$$Z \sim N(0, 1)$$

$$p = P(Z \leq z) = \int_{-\infty}^z \phi_Z(u) du = 1 - \alpha$$

donde

$$\phi_Z(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$$



Nota: Si $X \sim N(\mu, \sigma^2)$, entonces $Z = (X - \mu)/\sigma \sim N(0, 1)$. Luego,

$$P(X \leq x) = P\left(Z \leq \frac{x - \mu}{\sigma}\right)$$

Tabla 3A. Probabilidades acumuladas p de la distribución normal estándar.

z	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00
-3.4	0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
-3.3	0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005
-3.2	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007
-3.1	0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010
-3.0	0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013
-2.9	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019
-2.8	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026
-2.7	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035
-2.6	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047
-2.5	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062
-2.4	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082
-2.3	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107
-2.2	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139
-2.1	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179
-2.0	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228
-1.9	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287
-1.8	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359
-1.7	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446
-1.6	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548
-1.5	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668
-1.4	0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808
-1.3	0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968
-1.2	0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151
-1.1	0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357
-1.0	0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587
-0.9	0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841
-0.8	0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119
-0.7	0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420
-0.6	0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743
-0.5	0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085
-0.4	0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446
-0.3	0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821
-0.2	0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207
-0.1	0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602
-0.0	0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000

Tabla 3B. Probabilidades acumuladas p de la distribución normal estándar.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

ESTADISTICA

1-Estadística Descriptiva

1. 1 – Introducción

El campo de la estadística tiene que ver con la recopilación, organización, análisis y uso de datos para tomar decisiones razonables basadas en tal análisis.

Al recoger datos relativos a las características de un grupo de individuos u objetos, sean alturas y pesos de estudiantes de una universidad o tuercas defectuosas producidas por una fábrica, suele ser imposible o poco práctico observar todo el grupo, en especial si es muy grande. En vez de examinar el grupo entero, llamado **población** o universo, se examina una pequeña parte del grupo, llamada **muestra**.

En muchos problemas estadísticos es necesario utilizar una muestra de observaciones tomadas de la población de interés con objeto de obtener conclusiones sobre ella. A continuación se presenta la definición de algunos términos

Una **población** está formada por la totalidad de las observaciones en las cuales se tiene cierto interés.

Una **muestra** es un subconjunto de observaciones seleccionada de una población

Si una muestra es representativa de una población, es posible inferir importantes conclusiones sobre la población a partir del análisis de la muestra. La parte de la estadística que trata sobre las condiciones bajo las cuales tal inferencia es válida se llama **estadística inductiva** o **inferencia estadística**. Ya que dicha inferencia no es del todo exacta, el lenguaje de las probabilidades aparecerá al establecer nuestras conclusiones.

La parte de la estadística que estudia la muestra sin inferir alguna conclusión sobre la población es la **estadística descriptiva**.

En particular la estadística descriptiva trata sobre los métodos para recolectar, organizar y resumir datos.

La estadística descriptiva puede a su vez dividirse en dos grandes áreas: métodos gráficos y métodos numéricos.

En lo referente a la notación, n representa el número de observaciones en un conjunto de datos, las observaciones están representadas por una variable con subíndice (por ejemplo x_1, x_2, \dots, x_n). Así la representación de los cinco valores, $n = 5$, de la velocidad de un chip de computadora en MHz medida por un ingeniero, será: $x_1 = 481.5$, $x_2 = 493.7$, $x_3 = 471.8$, $x_4 = 486.4$, $x_5 = 496.2$,

1.2 – Distribución de frecuencias e histogramas

Supongamos que los siguientes datos representan la vida de 40 baterías para automóvil similares, registradas al décimo de año más cercano. Las baterías se garantizan por tres años.

2.2 4.1 3.5 4.5 3.2 3.7 3.0 2.6 3.4 1.6 3.1 3.3 3.8 3.1 4.7 3.7
 2.5 4.3 3.4 3.6 2.9 3.3 3.9 3.1 3.3 3.1 3.7 4.4 3.2 4.1 1.9 3.4
 4.7 3.8 3.2 2.6 3.9 3.0 4.2 3.5

Para organizar los datos buscamos el mínimo y el máximo de la muestra, en este caso el mínimo es 1.6 y el máximo es 4.7

Elegimos un intervalo (a, b) que contenga todos los datos, por ejemplo $a = 1.5$ y $b = 5.0$.

Dividimos el intervalo (a, b) en subintervalos que pueden ser de igual longitud, pero no necesariamente, y contamos cuántas observaciones caen en cada subintervalo, esa será la **frecuencia** del intervalo. Para esto debemos decidir cuántos subintervalos utilizaremos. En general se puede usar la regla de tomar aproximadamente \sqrt{n} subintervalos. Los subintervalos se llaman **intervalos de clase** o simplemente **clases**. Resulta satisfactorio utilizar no menos de 5 clases ni más de 20.

En el ejemplo $\sqrt{40} \approx 6$, entonces 6 o 7 clases será una elección satisfactoria.

Como $b - a = 5.0 - 1.5 = 3.5$, si tomamos $r = 7$ clases entonces la longitud de cada una sería $(b - a)/r = 0.5$.

Construimos una **tabla de frecuencias** de manera tal que, por ejemplo, en el intervalo (1.5 , 2.0) están las observaciones **mayores** a 1.5 y **menores o iguales que** 2.0.

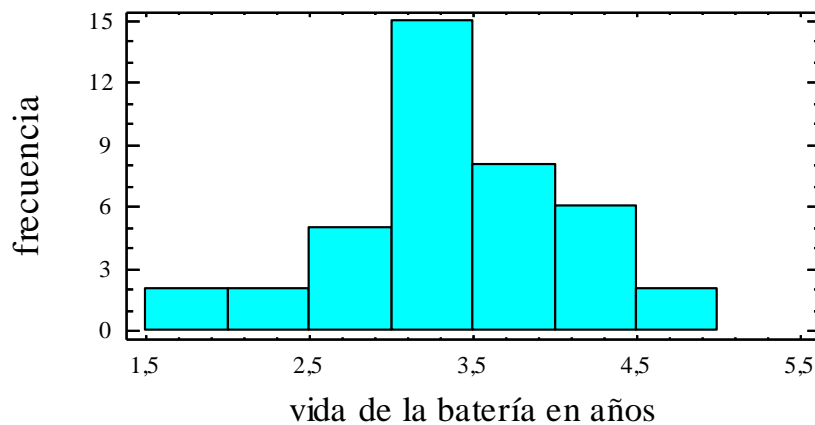
Los extremos de los intervalos de clase son los **límites de clase inferior y superior**.

Intervalo de clase	Marca de clase	Frecuencia f	Frecuencia relativa	Frecuencia acumulada	Frecuencia Acumulada relativa
1.5 – 2.0	1.75	3	0.075	2	0.075
2.0 – 2.5	2.25	1	0.025	4	0.1
2.5 – 3.0	2.75	5	0.125	9	0.225
3.0 – 3.5	3.25	15	0.375	24	0.6
3.5 – 4.0	3.75	8	0.2	32	0.8
4.0 – 4.5	4.25	6	0.15	38	0.95
4.5 – 5.0	4.75	2	0.05	40	1.000

El punto medio de cada clase es la **marca de clase**. La longitud de cada intervalo de clase es el **ancho de clase**.

El **gráfico de la tabla de frecuencias** es el **histograma**. Se construye en un sistema de ejes cartesianos. Sobre el eje de abscisas se marcan los límites de clase, y en cada clase se construye un rectángulo cuya base es el intervalo de clase y el **área del mismo** debe ser proporcional a la **frecuencia de la clase**. Si los intervalos de clase tienen el **mismo ancho** se puede construir cada rectángulo de manera que su **altura sea igual a la frecuencia** de la clase correspondiente. Estos histogramas son más fáciles de interpretar.

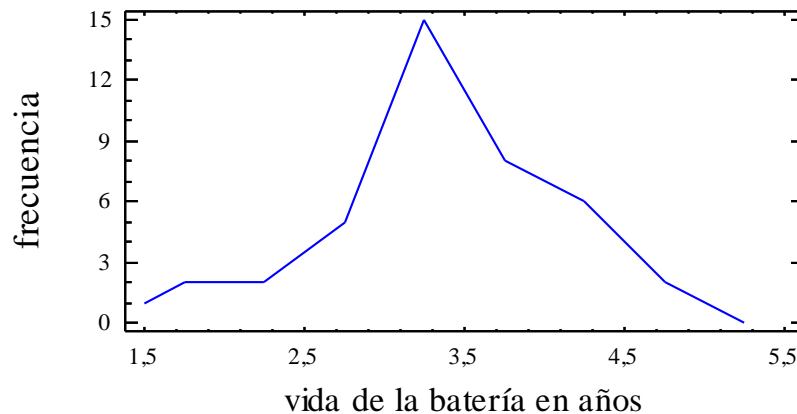
En la figura siguiente se muestra el histograma referido a la tabla de frecuencias anterior generado por el paquete Statgraphics.



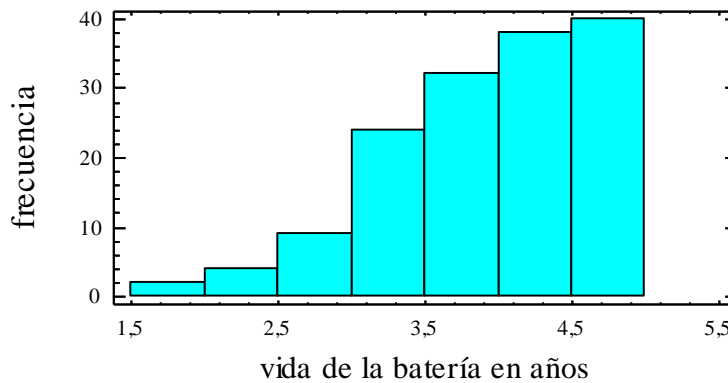
Si se tienen pocos datos los histogramas pueden cambiar de apariencia al variar el número de clases y el ancho de las mismas.

Si se hubiera graficado el **histograma de frecuencias relativas** el aspecto sería el mismo con la diferencia de la notación en el eje de ordenadas.

También se puede graficar un **polígono de frecuencias** al unir los puntos medios del lado superior de cada rectángulo con segmentos y agregar en los extremos dos clases adicionales de frecuencia cero como indica la siguiente figura



La figura siguiente muestra un **histograma de frecuencias acumuladas** disponible en el paquete Statgraphics. En esta gráfica la altura de cada rectángulo representa el número total de observaciones que son menores o iguales al límite superior de la clase respectiva.



Los histogramas son útiles al proporcionar una impresión visual del aspecto que tiene la distribución de las mediciones, así como información sobre la dispersión de los datos.

Al construir una tabla de frecuencias se pierde información, sin embargo esa pérdida de información es a menudo pequeña si se le compara con la concisión y la facilidad de interpretación ganada al utilizar la distribución de frecuencias y el histograma.

Las distribuciones acumuladas también son útiles en la interpretación de datos; por ejemplo en la figura anterior puede leerse de inmediato que existen aproximadamente 25 baterías con duración menor o igual a 3.5 años.

1.3 – Diagrama de tallo y hoja

El diagrama de tallo y hoja es una buena manera de obtener una presentación visual informativa del conjunto de datos x_1, x_2, \dots, x_n , donde cada número x_i está formado al menos por dos dígitos. Para construir un diagrama de este tipo los números x_i se dividen en dos partes: un ***tallo***, formada por uno o más dígitos principales, y una ***hoja***, la cual contiene el resto de los dígitos. Para ilustrar lo anterior consideramos los datos que especifican la vida de 40 baterías para automóvil dados anteriormente. Dividimos cada observación en dos partes de manera que el tallo representa el dígito entero que antecede al decimal, y la hoja corresponde a la parte decimal del número. Por ejemplo, para el número 3.7 el dígito 3 designa el tallo, y el 7 la hoja. Para nuestros datos los cuatro tallos 1, 2, 3 y 4 se listan verticalmente del lado izquierdo de la tabla, en tanto que las hojas se registran en el lado derecho correspondiente del valor del tallo adecuado. Entonces la hoja 6 del número 1.6 se registra enfrente del tallo 1, la hoja 5 del número 2.5 enfrente del tallo 2, y así sucesivamente.

<i>Tallo</i>	<i>Hoja</i>
1	6 9
2	2 5 6 6 9
3	0 0 1 1 1 1 2 2 2 3 3 3 4 4 4 5 5 6 7 7 7 8 8 9 9
4	1 1 2 3 4 5 7 7

Podríamos aumentar el número de tallos para obtener una forma mas adecuada de la distribución de los datos, para esto escribimos dos veces cada valor del tallo y después registramos las hojas 0, 1, 2, 3 y 4 enfrente del valor del tallo adecuado donde aparezca por primera vez, y las hojas 5, 6, 7, 8 y 9 enfrente de este mismo valor del tallo donde aparece por segunda vez.

En la tabla siguiente se ilustra el nuevo diagrama de tallo y hoja donde a los tallos que corresponden a las hojas 0 a 4 se les anotó un símbolo *, y al tallo correspondiente a las hojas 5 a 9 se les anotó el símbolo ·.

<i>Tallo</i>	<i>Hoja</i>
1	6 9
2*	2
2·	5 6 6 9
3*	0 0 1 1 1 1 2 2 2 3 3 3 4 4 4
3·	5 5 6 7 7 7 8 8 9 9
4*	1 1 2 3 4
4·	5 7 7

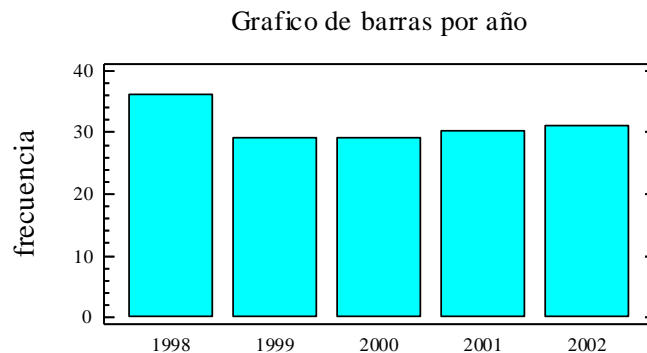
En cualquier problema específico, se debe decidir cuáles son los valores del tallo adecuados. Se trata de una decisión que se toma algo arbitrariamente, aunque nos guiamos por el tamaño de nuestra muestra. Por lo general se eligen entre 5 y 20 tallos. Cuanto menor es el número de datos disponibles, menor será la elección del número de tallos.

Observación:

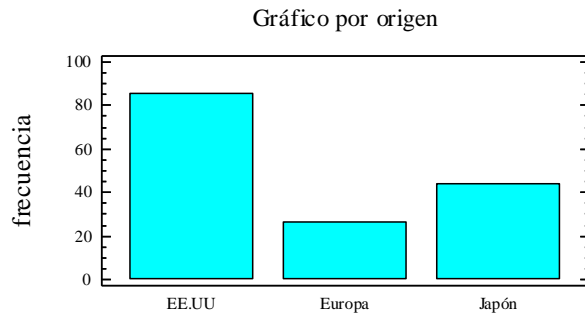
Las tablas de frecuencia y los histogramas también pueden emplearse en **datos cualitativos o categóricos**, es decir la muestra no consiste de valores numéricos (**datos cuantitativos**) sino que los datos se ordenan en **categorías** y se registra cuántas observaciones caen en cada categoría (las categorías pueden ser *masculino*, *femenino* o *fumador*, *no fumador* o clasificar según nivel educativo: *primario*, *secundario*, *terciario*, *universitario*, *ninguno*). Cuando los datos son categóricos las clases se dibujan con el mismo ancho.

Por ejemplo

El gráfico siguiente corresponde a una muestra de 155 autos clasificados según el año de fabricación: 1998, 1999, 2000, 2001 y 2002. Notar que el modelo 1998 es el de mayor frecuencia.



En el siguiente gráfico se clasifican los autos según el origen: americano, europeo o japonés. Notar que hay mayoría de autos americanos.



1.4 – Medidas Descriptivas

1.4.1 – Medidas de Localización

Del mismo modo que las gráficas pueden mejorar la presentación de los datos, las descripciones numéricas también tienen gran valor. Se presentan varias medidas numéricas importantes para describir las características de los datos.

Una característica importante de un conjunto de números es su **localización** o **tendencia central**.

Media

La medida más común de localización o centro de un grupo de datos es el promedio aritmético ordinario o media. Ya que casi siempre se considera a los datos como una muestra, la media aritmética se conoce como **media muestral**.

Si las observaciones de una muestra de tamaño n son x_1, x_2, \dots, x_n entonces la **media muestral** es

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

Ejemplo:

La media muestral de la vida útil en años de una batería de las 40 observaciones dadas en ejemplos anteriores es

$$\bar{x} = \frac{\sum_{i=1}^{40} x_i}{40} = \frac{2.2 + 4.1 + 3.5 + 4.5 + 3.2 + \dots + 3.5}{40} = 3.4125$$

La media tiene como ventaja su fácil cálculo e interpretación, pero tiene como desventaja el hecho de distorsionarse con facilidad ante la presencia de **valores atípicos** en los datos. Si en el ejemplo anterior tenemos $x_{40} = 350$ en lugar de 3.5 entonces $\bar{x} = 12.075$ dando así la idea errónea que los datos en su mayor parte se concentran alrededor de 12.075

Mediana

Otra medida de tendencia central es la **mediana**, o punto donde la muestra se divide en dos partes iguales.

Sean $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ una muestra escrita en orden creciente de magnitud; esto es $x_{(1)}$ denota la observación más pequeña, $x_{(2)}$ la segunda observación más pequeña, ..., y $x_{(n)}$ la observación más grande. Entonces la **mediana** \tilde{x} se define como la observación del lugar $\frac{n+1}{2}$ si n es impar, o el promedio de las observaciones de los lugares $\frac{n}{2}$ y $\frac{n}{2} + 1$ si n es par. En términos matemáticos

$$\tilde{x} = \begin{cases} x_{((n+1)/2)} & n \text{ impar} \\ \frac{x_{(n/2)} + x_{((n/2)+1)}}{2} & n \text{ par} \end{cases}$$

La ventaja de la mediana es que los valores extremos no tienen mucha influencia sobre ella.

Ejemplo:

Supóngase que los valores de una muestra son

1, 3, 4, 2, 7, 6, 8

Ordenamos los valores de menor a mayor: 1, 2, 3, 4, 6, 7, 8

Como son $n = 7$ valores y 7 es impar entonces la mediana es el valor del lugar $\frac{7+1}{2} = 4$, es decir $x_{(4)} = 4$

La media muestral es $\bar{x} = 4.4$. Ambas cantidades proporcionan una medida razonable de la tendencia central de los datos.

Si ahora tenemos la muestra 1, 2, 3, 4, 2450, 7, 8, entonces la media muestral es $\bar{x} = 353.6$

En este caso la media no dice mucho con respecto a la tendencia central de la mayor parte de los datos. Sin embargo la mediana sigue siendo $x_{(4)} = 4$, y ésta es una medida de tendencia central más significativa para la mayor parte de las observaciones.

Moda

La **moda** es la observación que se presenta con mayor frecuencia en la muestra

Por ejemplo la moda de los siguientes datos 3, 6, 9, 3, 5, 8, 3, 10, 4, 6, 3, 1
Es 3, porque este valor ocurre 4 veces y ningún otro lo hace con mayor frecuencia.
Puede existir más de una moda. Por ejemplo considérense las observaciones

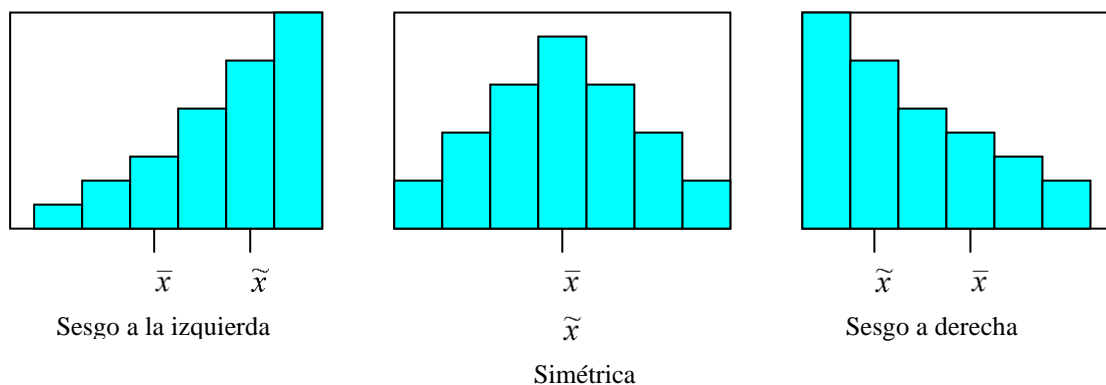
3, 6, 9, 3, 5, 8, 3, 10, 4, 6, 3, 1, 6, 2, 5, 6

La moda son 3 y 6, ya que ambos valores se presentan cuatro veces., y ningún otro lo hace con mayor frecuencia. Se dice en este caso que los datos son *bimodales*.

Si los datos son simétricos entonces la media y la mediana coinciden. Si además los datos tienen una sola moda entonces la media, la mediana y la moda coinciden.

Si los datos están *sesgados* (esto es, son asimétricos, con una larga cola en uno de los extremos), entonces la media, la mediana y la moda no coinciden. Generalmente se encuentra que $\text{moda} < \text{mediana} < \text{media}$ si la distribución está sesgada a la derecha, mientras que $\text{moda} > \text{mediana} > \text{media}$ si la distribución está sesgada hacia la izquierda.

Eso se representa en la siguiente figura



Percentiles y cuartiles

La mediana divide los datos de una muestra en dos partes iguales. También es posible dividir los datos en más de dos partes. Cuando se divide un conjunto ordenado de datos en cuatro partes iguales, los puntos de división se conocen como **cuartiles**. El *primer cuartil* o *cuartil inferior*, q_1 , es un valor que tiene aproximadamente la cuarta parte (25%) de las observaciones por debajo de él, y el 75% restante, por encima de él. El *segundo cuartil*, q_2 , tiene aproximadamente la mitad (50%) de las observaciones por debajo de él. El segundo cuartil coincide con la mediana. El *tercer cuartil* o *cuartil superior*, q_3 , tiene aproximadamente las tres cuartas partes (75%) de las observaciones por debajo de él. Como en el caso de la mediana, es posible que los cuartiles no sean únicos. Por simplicidad en este caso, si más de una observación cumple con la definición se utiliza el promedio de ellas como cuartil.

Ejemplo:

En 20 automóviles elegidos aleatoriamente se tomaron las emisiones de hidrocarburos en velocidad al vacío, en partes por millón (ppm)

141 359 247 940 882 494 306 210 105 880 200 223 188 940 241 190 300
435 241 380

Primero ordenamos los datos de menor a mayor:

105, 141, 188, 190, 200, 210, 223, 241, 241, 247, 300, 306, 359, 380, 435, 494, 880, 882, 940, 940

Buscamos la mediana o segundo cuartil, como $n = 20$ y es número par entonces la mediana es el promedio de la observaciones que se encuentran en los lugares $\frac{n}{2} = 10$ y

$$\frac{n}{2} + 1 = 11, \text{ es decir } \tilde{x} = q_2 = \frac{247 + 300}{2} = 273.5$$

Ahora buscamos el primer cuartil, para esto tomamos las primeras 10 observaciones

105, 141, 188, 190, 200, 210, 223, 241, 241, 247

y de éstas calculamos la mediana, por lo tanto $q_1 = \frac{200 + 210}{2} = 205$

Análogamente, para calcular el tercer cuartil, tomamos las últimas 10 observaciones y calculamos la mediana de éstas

300, 306, 359, 380, 435, 494, 880, 882, 940, 940

$$q_3 = \frac{435 + 494}{2} = 464.5$$

Cuando un conjunto ordenado de datos se divide en cien partes iguales, los puntos de división reciben el nombre de **percentiles**. En términos matemáticos el $100k$ – ésimo percentil p_k se define:

El $100k$ – ésimo percentil p_k es un valor tal que al menos el $100k\%$ de las observaciones son menores o iguales a él, y $100(1-k)\%$ son mayores o iguales a él.

Notar que $p_{0.25} = q_1$, $p_{0.5} = q_2$, $p_{0.75} = q_3$

Una regla práctica para calcular los percentiles de un conjunto de n datos es la siguiente: para calcular p_k hacemos el producto nk

si nk es un número entero i entonces $p_k = \frac{x_{(i)} + x_{(i+1)}}{2}$

si nk no es un número entero entonces tomamos la parte entera de nk : $\lfloor nk \rfloor = i$ y entonces $p_k = x_{(i+1)}$

Con los datos anteriores calculamos $p_{0.88}$

$n = 20$ $k = 0.88$ $nk = 20 \times 0.88 = 17.6$ la parte entera es $\lfloor 17.6 \rfloor = 17$ entonces $p_{0.88} = x_{18} = 882$

Calculamos $p_{0.10}$:

$n = 20$ $k = 0.10$ $nk = 20 \times 0.10 = 2$ entonces

$$p_{0.10} = \frac{x_{(2)} + x_{(3)}}{2} = \frac{141 + 188}{2} = 164.5$$

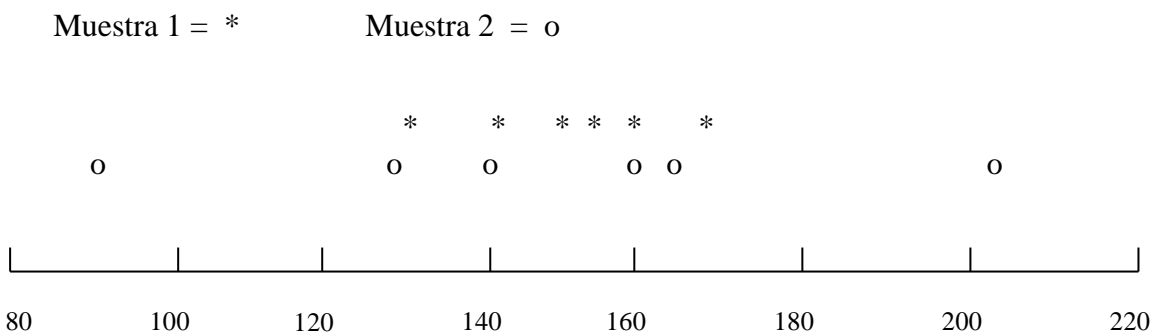
1.4.2 - Medidas de variabilidad

La localización o tendencia central no necesariamente proporciona información suficiente para describir datos de manera adecuada. Por ejemplo, supongamos que tenemos dos muestras de resistencia a la tensión (en psi) de aleación de aluminio-litio:

Muestra 1: 130, 150, 145, 158, 165, 140

Muestra 2: 90, 128, 205, 140, 165, 160

La media en ambas muestras es 148 psi. Sin embargo la dispersión o variabilidad de la muestra 2 es mucho mayor que la de la muestra 1. Para ilustrar esto hacemos para cada muestra un **diagrama de puntos**:



Rango de la muestra y rango intercuartílico

Una medida muy sencilla de variabilidad es el **rango de la muestra**, definido como la diferencia entre las observaciones más grande y más pequeña. Es decir

$$\text{rango} = \max(x_i) - \min(x_i)$$

Para las muestras anteriores

Muestra 1 \longrightarrow $\text{rango} = 165 - 130 = 35$

Muestra 2 \longrightarrow $\text{rango} = 205 - 90 = 115$

Está claro que a mayor rango, mayor variabilidad en los datos.

El rango ignora toda la información que hay en la muestra entre las observaciones más chica y más grande. Por ejemplo las muestras 1, 4, 6, 7, 9 y 1, 5, 5, 5, 9 tienen el mismo rango (rango = 8). Sin embargo en la segunda muestra sólo existe variabilidad en los valores de los extremos, mientras que en la primera los tres valores intermedios cambian de manera considerable.

Al igual que las observaciones máxima y mínima de una muestra llevan información sobre la variabilidad, el **rango intercuartílico** definido como $q_3 - q_1$ puede emplearse como medida de variabilidad.

$$RIC = q_3 - q_1$$

Para los datos de las emisiones de hidrocarburos en velocidad al vacío, el rango intercuartílico es $q_3 - q_1 = 464.5 - 205 = 259.5$.

El rango intercuartílico es menos sensible a los valores extremos de la muestra que el rango muestral.

Varianza muestral y desviación estándar muestral

Si x_1, x_2, \dots, x_n es una muestra de n observaciones, entonces la **varianza muestral** es

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

La **desviación estándar muestral**, s , es la raíz cuadrada positiva de la varianza muestral

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

Por ejemplo, para los datos de las emisiones de hidrocarburos en velocidad al vacío

141 359 247 940 882 494 306 210 105 880 200 223 188 940 241 190 300
435 241 380

la media muestral es $\bar{x} = 395.1$ ppm y la varianza muestral es

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{(141 - 395.1)^2 + (359 - 395.1)^2 + (247 - 395.1)^2 + \dots + (380 - 395.1)^2}{19} = 78998.5 \text{ ppm}^2$$

$$\text{y la desviación estándar es } s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = \sqrt{78998.5} = 281.067 \text{ ppm}$$

Observaciones:

- 1) La varianza muestral está en las unidades de medida de la variable al cuadrado. Por ejemplo si los datos están medidos en metros, las unidades de la varianza son metros al cuadrado. La desviación estándar tiene la propiedad de estar expresada en las mismas unidades de medida de las observaciones.

2) Se puede expresar la varianza muestral como

$$s^2 = \frac{\sum_{i=1}^n x_i^2 - n \bar{x}^2}{n-1}$$

pues:

$$\begin{aligned} s^2 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2)}{n-1} = \frac{\sum_{i=1}^n x_i^2 - \sum_{i=1}^n 2x_i\bar{x} + \sum_{i=1}^n \bar{x}^2}{n-1} = \\ &= \frac{\sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + \sum_{i=1}^n \bar{x}^2}{n-1} = \frac{\sum_{i=1}^n x_i^2 - 2\bar{x}n \frac{\sum_{i=1}^n x_i}{n} + n\bar{x}^2}{n-1} = \frac{\sum_{i=1}^n x_i^2 - 2\bar{x}n\bar{x} + n\bar{x}^2}{n-1} = \\ &= \frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n-1} \end{aligned}$$

En el ejemplo anterior volvemos a calcular la varianza muestral pero con esta última expresión

$$s^2 = \frac{\sum_{i=1}^n x_i^2 - n \bar{x}^2}{n-1} = \frac{4623052 - (20)(395.1)^2}{19} = 78998.5 \text{ ppm}^2$$

Coeficiente de variación

El *coeficiente de variación muestral* es

$$cv = \frac{s}{\bar{x}}$$

El coeficiente de variación muestral es útil cuando se compara la variabilidad de dos o más conjuntos de datos que difieren de manera considerable en la magnitud de las observaciones. Ejemplo: con un micrómetro se realizan mediciones del diámetro de un balero, que tienen una media de 4.03 mm y una desviación estándar de 0.012 mm; con otro micrómetro se toman mediciones de la longitud de un tornillo, que tienen una media de 1.76 pulgadas y una desviación estándar de 0.0075 pulgadas.

Los coeficientes de variación son

$$\text{balero} \longrightarrow cv = \frac{0.012}{4.03} = 0.003$$

tornillo \longrightarrow $cv = \frac{0.0075}{1.76} = 0.004$

en consecuencia, las mediciones hechas con el primer micrómetro tienen una variabilidad relativamente menor que las efectuadas con el otro micrómetro.

1.5 – Diagramas de caja

El **diagrama de caja** es una presentación visual que describe al mismo tiempo varias características importantes de un conjunto de datos, tales como el centro, la dispersión, la desviación de la simetría y la presencia de valores atípicos.

El diagrama de caja presenta los tres cuartiles, y los valores mínimo y máximo de los datos sobre un rectángulo en posición horizontal o vertical. El rectángulo delimita el rango intercuartílico con la arista izquierda ubicada en el primer cuartil, y la arista derecha ubicada en el tercer cuartil. Se dibuja una línea a través del rectángulo en la posición que corresponde al segundo cuartil. De cualquiera de las aristas del rectángulo se extiende una línea, o **bigote**, que va hacia los valores extremos. Éstas son observaciones que se encuentran entre cero y 1.5 veces el rango intercuartílico a partir de las aristas del rectángulo. Las observaciones que están entre 1.5 y 3 veces el rango intercuartílico a partir de las aristas del rectángulo reciben el nombre de **valores atípicos**. Las observaciones que están más allá de tres veces el rango intercuartílico a partir de las aristas del rectángulo se conocen como **valores atípicos extremos**. En ocasiones se emplean diferentes símbolos para identificar los dos tipos de valores atípicos. A veces los diagramas de caja reciben el nombre de **diagramas de caja y bigotes**.

Ejemplo: volvemos a los datos que representan la vida de 40 baterías para automóvil similares, registradas al décimo de año más cercano.

2.2 4.1 3.5 4.5 3.2 3.7 3.0 2.6 3.4 1.6 3.1 3.3 3.8 3.1 4.7 3.7
 2.5 4.3 3.4 3.6 2.9 3.3 3.9 3.1 3.3 3.1 3.7 4.4 3.2 4.1 1.9 3.4
 4.7 3.8 3.2 2.6 3.9 3.0 4.2 3.5

La figura siguiente presenta un diagrama de caja obtenido con el paquete Statgraphics

tamaño de muestra = 40

Media = 3.4125

Mediana = 3.4

Mínimo = 1.6

Máximo = 4.7

Rango = 3.1

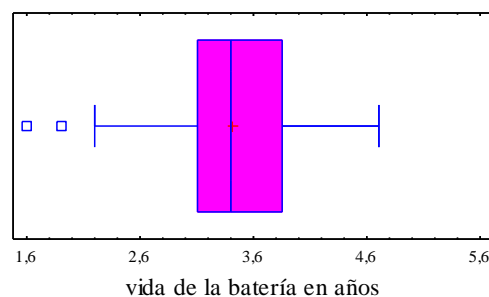
1° cuartil = 3.1

3° cuartil = 3.85

rango intercuartílico (RIC) = 0.75

1.5 RIC = 1.125

3 RIC = 2.25



El diagrama indica que la distribución de la vida en años de una batería de automóvil es bastante simétrica con respecto al valor central ya que los bigotes izquierdo y derecho, así como la longitud de los rectángulos izquierdo y derecho alrededor de la mediana, son casi los mismos. También se observa la existencia de dos valores atípicos de rango medio en el extremo izquierdo de los datos.

Los diagramas de caja son muy útiles para hacer comparaciones gráficas entre conjuntos de datos, ya que tienen un gran impacto visual y son fáciles de comprender.

Ejemplo: En 20 automóviles elegidos aleatoriamente, se tomaron las emisiones de hidrocarburos en velocidad al vacío, en partes por millón, para modelos de 1985 y 1995.

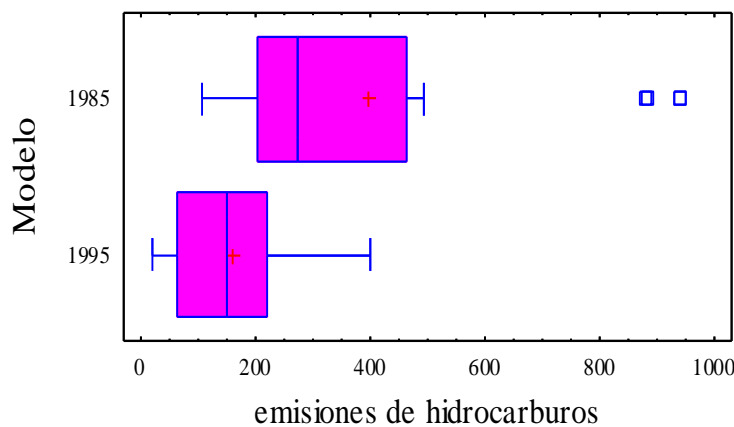
Modelos 1985:

141 359 247 940 882 494 306 210 105 880 200 223 188 940 241 190 300
435 241 380

Modelos 1995:

140 160 20 20 223 60 20 95 360 70 220 400 217 58 235 380
200 175 85 65

La figura siguiente presenta los diagramas de caja comparativos generados por el paquete Statgraphics



Modelo 1985

Media = 395.1

Mediana = 273.5

Mínimo = 105

Máximo = 940

Rango = 835

1° cuartil = 205

3° cuartil = 464.5

rango intercuartílico (*RIC*) = 259.5

Modelo 1995

Media = 160.15

Mediana = 150

Mínimo = 20

Máximo = 400

Rango = 380

1° cuartil = 62.5

3° cuartil = 221.5

rango intercuartílico (*RIC*) = 159

Se observa que han disminuido las emisiones de hidrocarburos del modelo 1985 al 1995. Además los datos correspondientes al modelo 1985 tienen mayor variabilidad que los del modelo 1995. Para el modelo 1985 hay dos autos con emisiones muy altas.

Ejemplo:

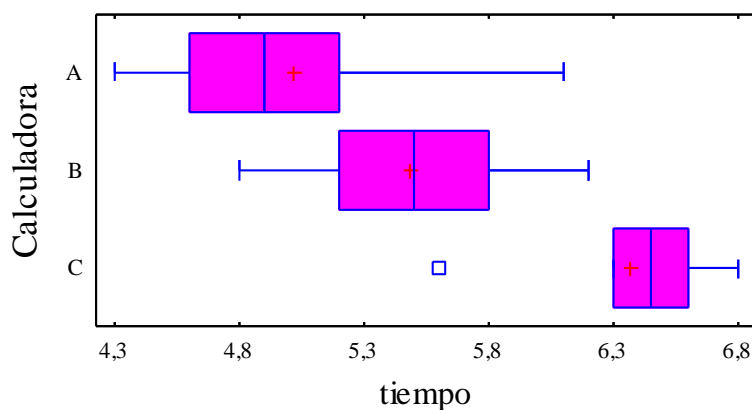
Los siguientes datos representan los tiempos de operación, en horas, para tres tipos de calculadoras científicas de bolsillo, antes de que requieran recarga

Calculadora A: 4.9 6.1 4.3 4.6 5.2

Calculadora B: 5.5 5.4 6.2 5.8 5.5 5.2 4.8

Calculadora C: 6.4 6.8 5.6 6.5 6.3 6.6

La figura siguiente presenta los diagramas de caja comparativos generados por el paquete Statgraphics. Queda a cargo del lector comentar las diferencias entre las calculadoras A, B y C.



Calculadora	tamaño de muestra	Media	Mediana	Mínimo	Máximo	q1	q3	RIC
A	5	5,02	4,9	4,3	6,1	4,6	5,2	0,6
B	7	5,48571	5,5	4,8	6,2	5,2	5,8	0,6
C	6	6,36667	6,45	5,6	6,8	6,3	6,6	0,3

Ejemplo:

El contenido de nicotina, en miligramos, en 40 cigarrillos de cierta marca se registraron como sigue:

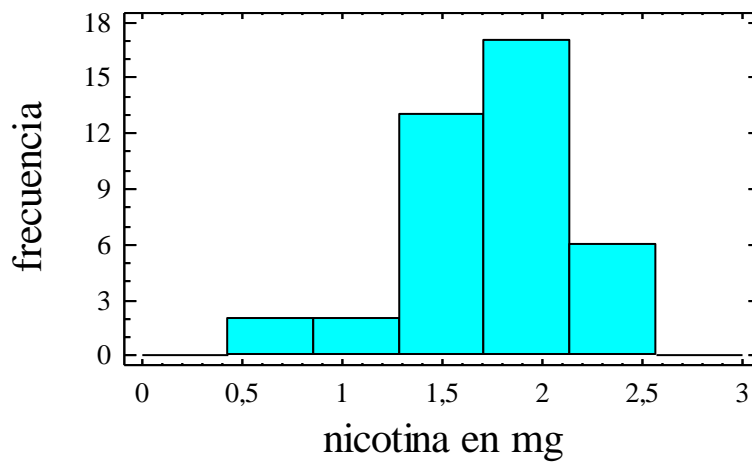
1.09 1.92 2.31 1.79 2.28 1.74 1.47 1.97 0.85 1.24 1.58 2.03
 1.70 2.17 2.55 2.11 1.86 1.90 1.68 1.51 1.64 0.72 1.69 1.85
 1.82 1.79 2.46 1.88 2.08 1.67 1.37 1.93 1.40 1.64 2.09 1.75
 1.63 2.37 1.75 1.69

Se muestra la tabla de frecuencias y el histograma de los datos según Statgraphics.

Se tomaron 7 clases, donde el límite inferior de la primera clase es 0 y el límite superior de la última clase es 3.

Clase	limite inferior	limite superior	Marca de clase	frecuencia	frecuencia relativa	frec. acum.	frec. ac. relativa
1	0,0	0,428571	0,214286	0	0,0000	0	0,0000
2	0,428571	0,857143	0,642857	2	0,0500	2	0,0500
3	0,857143	1,28571	1,07143	2	0,0500	4	0,1000

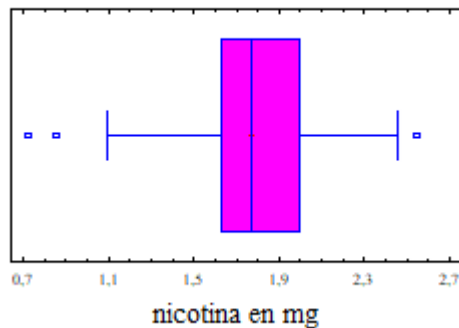
4	1,28571	1,71429	1,5	13	0,3250	17	0,4250
5	1,71429	2,14286	1,92857	17	0,4250	34	0,8500
6	2,14286	2,57143	2,35714	6	0,1500	40	1,0000
7	2,57143	3,0	2,78571	0	0,0000	40	1,0000



A continuación se dan las medidas descriptivas y el gráfico de caja y bigote:

Tamaño de muestra = 40
 Media = 1,77425
 Mediana = 1,77
 Moda =
 Varianza = 0,152456
 Desviación estándar = 0,390456
 Mínimo = 0,72
 Máximo = 2,55
 Rango = 1,83
 1° cuartil = 1,635
 3° cuartil = 2
 Rango intercuartílico = 0,365
 Coef. de variación = 22,0068%

gráfico de caja y bigote



Practica**Estadística Descriptiva**

- 1) Un artículo publicado en *Technometrics* (Vol. 19,1977, pag.425) presenta los datos siguientes sobre el octanaje de varias mezclas de gasolina:

88.5	87.7	83.4	86.7	87.5	91.5	88.6	100.3	96.5	93.3
94.7	91.1	91.0	94.2	87.8	89.9	88.3	87.6	84.3	86.7
84.3	86.7	88.2	90.8	88.3	98.8	94.2	92.7	93.2	91.0
90.1	93.4	88.5	90.1	89.2	88.3	85.3	87.9	88.6	90.9
89.0	96.1	93.3	91.8	92.3	90.4	90.1	93.0	88.7	89.9
89.8	89.6	87.4	88.4	88.9	91.2	89.3	94.4	92.7	91.8
91.6	90.4	91.1	92.6	89.8	90.6	91.1	90.4	89.3	89.7
90.3	91.6	90.5	93.7	92.7	92.2	92.2	91.2	91.0	92.2
90.0	90.7								

- a) Construya una tabla de frecuencias (absoluta y relativa) y un histograma utilizando ocho clases.
- b) Construya una tabla de frecuencias (absoluta y relativa) y un histograma con 16 clases. Compare la forma del histograma con la que tiene el histograma de la parte a). ¿Los dos histogramas presentan información similar?.
- 2) Los datos siguientes representan el número de ciclos transcurridos hasta que se presenta una falla en una prueba de piezas de aluminio sujetas a un esfuerzo alternante repetido de 21000 psi, a 18 ciclos por segundo:

1115	1567	1223	1782	1055	798	1016	2100	910	1501
1310	1883	375	1522	1764	1020	1102	1594	1730	1238
1540	1203	2265	1792	1330	865	1605	2023	1102	990
1502	1270	1910	1000	1608	2130	706	1315	1578	1468
1258	1015	1018	1820	1535	1421	2215	1269	758	1512
1315	845	1452	1940	1781	1109	785	1260	1416	1750
1085	1674	1890	1120	1750	1481	885	1888	1560	1642

- a) Construya una tabla de frecuencias (absoluta y relativa) y un histograma.
- b) Construya un polígono de frecuencias.
- c) ¿Existe evidencia de que una pieza “sobrevivirá” más allá de los 2000 ciclos?. Justifique su respuesta.
- 3) Las siguientes mediciones corresponden a las temperaturas de un horno registradas en lotes sucesivos de un proceso de fabricación de semiconductores (las unidades son °F): 953, 950, 948, 955, 951, 949, 957, 954, 955.
- Calcule:
- a) La media muestral de estos datos.
- b) La mediana muestral de estos datos
- c) ¿En cuánto puede incrementarse la mayor medición de temperatura sin que cambie la mediana muestral?.

- 4) Se toman ocho mediciones del diámetro interno de los anillos para los pistones del motor de un automóvil. Los datos (en mm) son: 74.001, 74.003, 74.015, 74.000, 74.005, 74.002, 74.005, 74.004.
- Encuentre la media y la mediana de estos datos.
 - Suponga que se elimina la observación más grande (74.015 mm). Calcule la media y la mediana muestrales para los datos restantes. Compare sus resultados con los obtenidos en la parte a).
- 5) Hallar la media y mediana para los datos de los ejercicios 1) y 2). Considerar los datos agrupados.
- 6) Los siguientes datos son las temperaturas de unión de los *O-rings* (en grados F), en cada prueba de lanzamiento o de un lanzamiento real, del motor del cohete del transbordador espacial (tomados de *Presidential Commission on the Space Shuttle Challenger Accident*, Vol.1, pags. 129-131): 84, 49, 61, 40, 83, 67, 45, 66, 70, 69, 80, 58, 68, 60, 67, 72, 73, 70, 57, 63, 70, 78, 52, 67, 53, 67, 75, 61, 70, 81, 76, 79, 75, 76, 58, 31.
- Calcule la media y la mediana muestrales.
 - Encuentre los cuartiles inferior y superior de la temperatura.
 - Encuentre los percentiles quinto y noveno de la temperatura.
 - Elimine la observación más pequeña (31°F) y vuelva a calcular lo que se pide en los incisos a), b) y c). ¿Qué efecto tiene la eliminación de este punto?
- 7) La contaminación de una pastilla de silicio puede afectar de manera importante la calidad de la producción de circuitos integrados. De una muestra de 10 pastillas se obtienen las siguientes concentraciones de oxígeno: 3.15, 2.68, 4.31, 2.09, 3.82, 2.94, 3.47, 3.39, 2.81, 3.61. Calcule:
- La varianza muestral.
 - La desviación estándar muestral.
 - El rango de la muestra.
- 8) Considere los datos para anillos de pistón, del ejercicio 4). Calcule:
- La varianza muestral.
 - La desviación estándar muestral.
 - El rango de la muestra.
 - Suponga que se elimina la observación más grande (74.015). Calcule la varianza muestral, la desviación estándar muestral, el rango de la muestra. Compare los resultados con los obtenidos en los incisos anteriores. Para esta medición en particular, ¿cuán sensibles son la varianza muestral, la desviación muestral y el rango de la muestra?
- 9) Considere los datos de los ejercicios 1) y 2). En cada caso calcule (para datos agrupados):
- La varianza muestral.
 - La desviación estándar muestral.
 - El rango de la muestra.
- 9) Considere los datos del ejercicio 5).
- La varianza muestral.
 - La desviación estándar muestral.
 - El rango de la muestra y rango intercuartílico.
 - Construya un diagrama de caja y discuta sobre la forma de la distribución y la posible presencia de valores atípicos.

- 10) Se tienen las millas por galón de 154 autos, los que se clasifican según su origen. Se obtienen así las siguientes tres muestras:

Muestra 1, consiste en las millas por galón de 85 autos de origen americano:

36.1, 19.9, 19.4, 20.2, 19.2, 20.5, 20.2, 25.1, 20.5, 19.4, 20.6,
20.8, 18.6, 18.1, 19.2, 17.7, 18.1, 17.5, 30, 30.9, 23.2, 23.8, 21.5,
19.8, 22.3, 20.2, 20.6, 17, 17.6, 16.5, 18.2, 16.9, 15.5, 19.2, 18.5,
35.7, 27.4, 23, 23.9, 34.2, 34.5, 28.4, 28.8, 26.8, 33.5, 32.1, 28,
26.4, 24.3, 19.1, 27.9, 23.6, 27.2, 26.6, 25.8, 23.5, 30, 39, 34.7,
34.4, 29.9, 22.4, 26.6, 20.2, 17.6, 28, 27, 34, 31, 29, 27, 24, 23,
38, 36, 25, 38, 26, 22, 36, 27, 27, 32, 28, 31.

Muestra 2, consiste en las millas por galón de 25 autos de origen europeo:

43.1, 20.3, 17, 21.6, 16.2, 31.5, 31.9, 25.4, 27.2, 37.3, 41.5, 34.3,
44.3, 43.4, 36.4, 30.4, 40.9, 29.8, 35, 33, 34.5, 28.1, 30.7, 36, 44.

Muestra 3, consiste en las millas por galón de 44 autos de origen japonés:

32.8, 39.4, 36.1, 27.5, 27.2, 21.1, 23.9, 29.5, 34.1, 31.8, 38.1,
37.2, 29.8, 31.3, 37, 32.2, 46.6, 40.8, 44.6, 33.8, 32.7, 23.7, 32.4,
39.1, 35.1, 32.3, 37, 37.7, 34.1, 33.7, 32.4, 32.9, 31.6, 25.4, 24.2,
37, 31, 36, 36, 34, 38, 32, 38, 32

- Hallar media, mediana y moda de cada muestra.
- Hallar rango, rango intercuartílico, desviación estándar de cada muestra.
- Hallar el coeficiente de variación para cada muestra.
- Haga un gráfico de cajas simultáneas.
- ¿Qué puede concluir sobre el consumo de los autos europeos y japoneses con respecto al de los autos americanos?

- 11) Los precios de 155 autos se clasifican según el origen de los mismos, obteniéndose las siguientes tres muestras:

Muestra 1, consiste en el precio de 85 autos de origen americano:

1900, 3300, 3125, 2850, 2800, 3275, 2375, 2275, 2700, 2300,
3300, 2425, 2700, 2425, 3900, 4400, 2525, 3000, 2100, 2250,
3200, 2400, 3925, 3200, 2975, 3150, 3325, 4650, 4850, 5725,
4025, 5225, 4825, 4100, 4725, 2700, 2725, 9900, 4050, 2625,
2775, 3525, 3625, 3525, 3325, 2900, 3625, 3525, 3625, 3770,
3200, 4250, 5100, 5175, 4950, 4550, 4900, 3600, 3450, 3850,
4100, 6675, 6875, 5450, 5625, 5275, 5500, 5175, 5650, 6250,
5650, 5225, 5150, 4350, 4550, 7300, 7450, 6000, 6100, 5475,
8400, 5975, 5650, 4925, 6050.

Muestra 2, consiste en el precio de 26 autos de origen europeo:

4475, 5875, 4200, 5450, 3675, 3100, 4675, 2275, 7000, 5900,
3900, 3825, 7975, 14275, 2575, 7000, 5000, 5650, 4600, 8500,
8225, 8550, 5200, 5075, 2400, 15475.

Muestra 3, consiste en el precio de 44 autos de origen japonés:

2200, 2725, 2250, 2975, 2775, 3700, 2975, 3425, 2750, 2750,
3850, 3525, 5500, 4675, 4050, 3975, 3350, 3300, 3925, 3625,

8150, 7250, 4700, 4400, 4000, 3950, 3775, 4475, 3975, 5550, 4650, 5825, 5650, 9475, 8375, 4900, 5100, 6350, 6500, 5425, 4625, 4875, 5075, 7700.

- a) Hallar media, mediana y moda de cada muestra.
- b) Hallar rango, rango intercuartílico, desviación estándar de cada muestra.
- c) Hallar el coeficiente de variación para cada muestra.
- d) Haga un gráfico de cajas simultáneas.
- e) ¿Hay diferencias de precio entre los diferentes orígenes?.

12) En relación al ejercicio 11), ahora se clasifica el precio de los 155 autos anteriores según el año del modelo, obteniéndose las siguientes cinco muestras:

Muestra 1, precios de 36 autos modelo año 78:

2400, 1900, 2200, 2725, 2250, 3300, 3125, 2850, 2800, 3275, 2375, 2275, 2700, 2300, 3300, 2425, 2700, 2425, 3900, 4400, 2525, 3000, 2100, 2975, 2775, 2250, 3700, 3200, 2400, 2975, 4475, 5875, 4200, 5450, 3675, 3425.

Muestra 2, precios de 29 autos modelo año 79:

3925, 3200, 2975, 3150, 3325, 4650, 4850, 5725, 4025, 5225, 4825, 4100, 4725, 3100, 2750, 2700, 2725, 15475, 9900, 4675, 4050, 2625, 2775, 2750, 2275, 3525, 3625, 3525, 3325.

Muestra 3, precios de 29 autos modelo año 80:

7000, 3850, 2900, 3525, 3625, 3525, 3625, 3700, 5900, 5500, 4675, 4050, 3975, 3350, 3200, 3300, 3900, 3825, 7975, 14275, 3925, 2575, 3625, 7000, 8150, 7250, 5000, 4250, 4700.

Muestra 4, precios de 30 autos modelo año 81:

5100, 5175, 4950, 4550, 4900, 4400, 3600, 4000, 3950, 3775, 4475, 3975, 3450, 3850, 4100, 5650, 4600, 5550, 4650, 5825, 5650, 8500, 8225, 8550, 9475, 8375, 6675, 6875, 5450, 5625.

Muestra 5, precios de 31 autos modelo año 82:

5275, 5500, 5175, 5650, 6250, 5650, 5225, 5150, 5200, 4900, 5100, 4350, 4550, 6350, 6500, 5425, 4625, 4875, 5075, 7300, 7450, 6000, 6100, 7700, 5475, 8400, 5975, 5075, 5650, 4925, 6050.

- a) Hallar media, mediana y moda de cada muestra.
- b) Hallar rango, rango intercuartílico, desviación estándar de cada muestra.
- c) Hallar el coeficiente de variación para cada muestra.
- d) Haga un gráfico de cajas simultáneas.
- e) ¿Ha variado el precio de los autos durante los años considerados?.

2-Estimación puntual

2.1 – Introducción

Supongamos la siguiente situación: en una fábrica se producen artículos, el interés está en la producción de un día, específicamente, de todos los artículos producidos en un día nos interesa una característica determinada, si el artículo es o no defectuoso. Sea p la proporción de artículos defectuosos en la **población**, es decir en la producción de un día.

Tomamos una **muestra** de 25 artículos, podemos definir la v.a. X : “número de artículos defectuosos en la muestra”, y podemos asumir que $X \sim B(25, p)$.

En **Probabilidades** se **conocían todos los datos sobre la v.a. X** , es decir conocíamos p . De esa forma podíamos responder preguntas como: ¿cuál es la probabilidad que entre los 25 artículos halla 5 defectuosos?. Si, por ejemplo, $p=0.1$ entonces calculábamos $P(X=5)$ donde $X \sim B(25, 0.1)$.

En **Estadística** **desconocemos las características de X** total o parcialmente, y a partir de la muestra de 25 artículos tratamos de **inferir** información sobre la distribución de X , o dicho de otra forma tratamos de inferir información sobre la **población**.

Por ejemplo, en estadística sabemos que X tiene distribución binomial pero **desconocemos p** , y a partir de la muestra de 25 artículos trataremos de hallar información sobre p .

En Estadística nos haremos preguntas tales como: si en la muestra de 25 artículos se encontraron 5 defectuosos, ¿ese hecho me permite inferir que el verdadero p es 0.1?.

El campo de la **inferencia estadística** está formado por los métodos utilizados para tomar decisiones o para obtener conclusiones sobre el o los parámetros de una población. Estos métodos utilizan la información contenida en una **muestra** de la población para obtener conclusiones.

La inferencia estadística puede dividirse en dos grandes áreas: **estimación de parámetros** y **pruebas de hipótesis**.

2.2 – Muestreo aleatorio

En muchos problemas estadísticos es necesario utilizar una muestra de observaciones tomadas de la población de interés con objeto de obtener conclusiones sobre ella. A continuación se presenta la definición de algunos términos

Una **población** está formada por la totalidad de las observaciones en las cuales se tiene cierto

En muchos problemas de inferencia estadística es poco práctico o imposible, observar toda la población, en ese caso se toma una parte o subconjunto de la población

Una **muestra** es un subconjunto de observaciones seleccionada de una población

Para que las inferencias sean válidas, la muestra debe ser representativa de la población. Se selecciona una muestra aleatoria como el resultado de un mecanismo aleatorio. En consecuencia, la selección de una muestra es un experimento aleatorio, y cada observación de la muestra es el valor observado de una variable aleatoria. Las observaciones en la población determinan la distribución de probabilidad de la variable aleatoria.

Para definir muestra aleatoria, sea X la v.a. que representa el resultado de tomar una observación de la población. Sea $f(x)$ la f.d.p. de la v.a. X . supongamos que cada observación en la muestra se obtiene de manera independiente, bajo las mismas condiciones. Es decir, las observaciones de la muestra se obtienen al observar X de manera independiente bajo condiciones que no cambian, digamos n veces.

Sea X_i la variable aleatoria que representa la i -ésima observación. Entonces X_1, X_2, \dots, X_n constituyen una muestra aleatoria, donde los valores numéricos obtenidos son x_1, x_2, \dots, x_n . Las variables aleatorias en una muestra aleatoria son independientes, con la misma distribución de probabilidad $f(x)$ debido a que cada observación se obtiene bajo las mismas condiciones. Es decir las funciones de densidad marginales de X_1, X_2, \dots, X_n son todas iguales a $f(x)$ y por independencia, la distribución de probabilidad conjunta de la muestra aleatoria es el producto de las marginales $f(x_1)f(x_2)\dots f(x_n)$

Las variables aleatorias (X_1, X_2, \dots, X_n) constituyen una **muestra aleatoria** de tamaño n de una v.a. X si X_1, X_2, \dots, X_n son independientes idénticamente distribuidas

El propósito de tomar una muestra aleatoria es obtener información sobre los parámetros desconocidos de la población. Por ejemplo, se desea alcanzar una conclusión acerca de la proporción de artículos defectuosos en la producción diaria de una fábrica. Sea p la proporción de artículos defectuosos en la población, para hacer una inferencia con respecto a p , se selecciona una muestra aleatoria (de un tamaño apropiado) y se utiliza la proporción observada de artículos defectuosos en la muestra para estimar p .

La proporción de la muestra \hat{p} se calcula dividiendo el número de artículos defectuosos en la muestra por el número total de artículos de la muestra. Entonces \hat{p} es una función de los valores observados en la muestra aleatoria. Como es posible obtener muchas muestras aleatorias de una población, el valor de \hat{p} cambiará de una a otra. Es decir **\hat{p} es una variable aleatoria**. Esta variable aleatoria se conoce como **estadístico**.

Un **estadístico** es cualquier función de la muestra aleatoria

Estadísticos usuales

Sea X_1, X_2, \dots, X_n una muestra aleatoria de una v.a. X donde $E(X) = \mu$ y $V(X) = \sigma^2$

Si desconocemos μ un estadístico que se utiliza para estimar ese parámetro es la **media o promedio muestral**

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Análogamente si se desconoce σ^2 un estadístico usado para tener alguna información sobre ese parámetro es la **varianza muestral** que se define como $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

Otro estadístico es la **desviación estándar muestral** $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$

Como un estadístico es una variable aleatoria, éste tiene una distribución de probabilidad, esperanza y varianza.

Una aplicación de los estadísticos es obtener **estimaciones puntuales** de los parámetros desconocidos de una distribución. Por ejemplo como se dijo antes se suelen estimar la media y la varianza de una población.

Cuando un estadístico se utiliza para estimar un parámetro desconocido se lo llama **estimador puntual**. Es habitual simbolizar en forma genérica a un parámetro con la letra θ y al estadístico que se utiliza como estimador puntual de θ , simbolizarlo con $\hat{\theta}$.

Por lo tanto $\hat{\theta}$ es una función de la muestra aleatoria: $\hat{\theta} = h(X_1, X_2, \dots, X_n)$

Al medir la muestra aleatoria se obtienen x_1, x_2, \dots, x_n , y entonces **el valor que toma** $\hat{\theta}$ es $\hat{\theta} = h(x_1, x_2, \dots, x_n)$ y se denomina **estimación puntual de** θ .

El objetivo de la estimación puntual es seleccionar un número, a partir de los valores de la muestra, que sea el valor más probable de θ .

Por ejemplo, supongamos que X_1, X_2, X_3, X_4 es una muestra aleatoria de una v.a. X . Sabemos que X tiene distribución normal pero desconocemos μ .

Tomamos como **estimador** de μ al promedio muestral \bar{X} , es decir $\hat{\mu} = \bar{X}$

Tomamos la muestra (medimos X_1, X_2, X_3, X_4) y obtenemos $x_1 = 24, x_2 = 30, x_3 = 27, x_4 = 32$

Entonces la **estimación puntual** de μ es $\bar{x} = \frac{24 + 30 + 27 + 32}{4} = 28.25$

Si la varianza σ^2 de X también es desconocida, un estimador puntual usual de σ^2 es la varianza muestral, es decir $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, para la muestra dada la estimación de σ^2 es 12.25.

Otro parámetro que a menudo es necesario estimar es la proporción p de objetos de una población que cumplen una determinada característica.

En este caso el estimador puntual de p sería $\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i$ donde

$$X_i = \begin{cases} 1 & \text{si la } i\text{-ésima observación tiene la característica de interés} \\ 0 & \text{caso contrario} \end{cases} \quad i = 1, 2, \dots, n$$

Por lo tanto $\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i$ es la **proporción de objetos en la muestra** cumplen la característica de interés

Puede ocurrir que se tenga más de un estimador para un parámetro, por ejemplo para estimar la media muestral se pueden considerar el promedio muestral, o también la semisuma entre X_1 y X_n , es decir $\hat{\mu} = \frac{X_1 + X_n}{2}$. En estos casos necesitamos de algún **criterio para decidir cuál es mejor estimador** de μ .

2.3 – Criterios para evaluar estimadores puntuales

Lo que se desea de un estimador puntual es **que tome valores “próximos” al verdadero parámetro**.

Se dice que el estimador puntual $\hat{\theta}$ es un **estimador insesgado** del parámetro θ si $E(\hat{\theta}) = \theta$ cualquiera sea el valor verdadero de θ

Podemos exigir que el estimador $\hat{\theta}$ tenga una distribución cuya media sea θ .

La diferencia $E(\hat{\theta}) - \theta$ se conoce como **sesgo de estimador** $\hat{\theta}$. Anotamos $b(\hat{\theta}) = E(\hat{\theta}) - \theta$

Notar que si un estimador es insesgado entonces su sesgo es cero

Ejemplos:

1- Sea X_1, X_2, \dots, X_n una muestra aleatoria de una v.a. X donde $E(X) = \mu$ y $V(X) = \sigma^2$

Si desconocemos μ un estadístico que se utiliza usualmente para estimar este parámetro es la

media o promedio muestral $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Veamos si es un estimador insesgado de μ . Debemos ver si $E(\bar{X}) = \mu$.

Usamos las propiedades de la esperanza, particularmente la propiedad de linealidad.

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} E\left(\sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i).$$

Pero, tratándose de las componentes de una muestra aleatoria es:

$$E(X_i) = E(X) = \mu \quad \forall i = 1, 2, \dots, n. \quad \text{Luego:}$$

$$E(\bar{X}) = \frac{1}{n} n\mu = \mu.$$

2- Sea X una variable aleatoria asociada con alguna característica de los individuos de una población y sean $E(X) = \mu$ y $V(X) = \sigma^2$. Sea $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ la varianza muestral (con

$\bar{X} = \left(\sum_{i=1}^n X_i\right)/n$ la esperanza muestral) para una muestra aleatoria de tamaño n ,

(X_1, X_2, \dots, X_n) . Entonces $E(S^2) = \sigma^2$ es decir $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ es un estimador insesgado de $V(X) = \sigma^2$ pues:

$$E(S^2) = E\left(\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2\right) = \frac{1}{n-1} E\left(\sum_{i=1}^n (X_i - \bar{X})^2\right).$$

Reescribiremos la suma de una forma más conveniente. Sumamos y restamos μ y desarrollamos el cuadrado:

$$\begin{aligned}
\sum_{i=1}^n (X_i - \bar{X})^2 &= \sum_{i=1}^n (X_i - \mu + \mu - \bar{X})^2 = \sum_{i=1}^n ([X_i - \mu] + [\mu - \bar{X}])^2 = \\
&= \sum_{i=1}^n \left\{ [X_i - \mu]^2 + 2[X_i - \mu][\mu - \bar{X}] + [\mu - \bar{X}]^2 \right\} \\
&= \sum_{i=1}^n [X_i - \mu]^2 + 2[\mu - \bar{X}] \sum_{i=1}^n [X_i - \mu] + n[\mu - \bar{X}]^2 = \\
&= \sum_{i=1}^n [X_i - \mu]^2 + 2[\mu - \bar{X}]n[\bar{X} - \mu] + n[\mu - \bar{X}]^2 = \sum_{i=1}^n [X_i - \mu]^2 - 2n[\mu - \bar{X}]^2 + n[\mu - \bar{X}]^2.
\end{aligned}$$

Esto es:

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n [X_i - \mu]^2 - n[\mu - \bar{X}]^2$$

Entonces:

$$\begin{aligned}
E(S^2) &= \frac{1}{n-1} E\left(\sum_{i=1}^n (X_i - \bar{X})^2\right) = \frac{1}{n-1} E\left(\sum_{i=1}^n [X_i - \mu]^2 - n[\mu - \bar{X}]^2\right) = \\
&= \frac{1}{n-1} \left(\sum_{i=1}^n E[X_i - \mu]^2 - nE[\bar{X} - \mu]^2\right) = \\
&= \frac{1}{n-1} \left(\sum_{i=1}^n V(X_i) - nE[\bar{X} - E(\bar{X})]^2\right) = \frac{1}{n-1} \left(\sum_{i=1}^n V(X_i) - nV(\bar{X})\right) = \frac{1}{n-1} \left(n\sigma^2 - n\frac{\sigma^2}{n}\right),
\end{aligned}$$

donde en la última igualdad tuvimos en cuenta que $V(X_i) = V(X) = \sigma^2 \quad \forall i = 1, 2, \dots, n$ y que

$V(\bar{X}) = \frac{\sigma^2}{n}$. Luego llegamos a lo que se deseaba demostrar: $E(S^2) = \sigma^2$.

3- Supongamos que tomamos como estimador de σ^2 a $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$

Entonces notar que podemos escribir $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n-1}{n} \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{n-1}{n} S^2$

Por lo tanto $E(\hat{\sigma}^2) = E\left(\frac{n-1}{n} S^2\right) = \frac{n-1}{n} E(S^2) = \frac{n-1}{n} \sigma^2 \neq \sigma^2$

Es decir $\hat{\sigma}^2$ no es un estimador insesgado de σ^2 , **es sesgado**, y su sesgo es

$$b(\hat{\sigma}^2) = E(\hat{\sigma}^2) - \sigma^2 = \left(\frac{n-1}{n}\right) \sigma^2 - \sigma^2 = -\frac{1}{n} \sigma^2$$

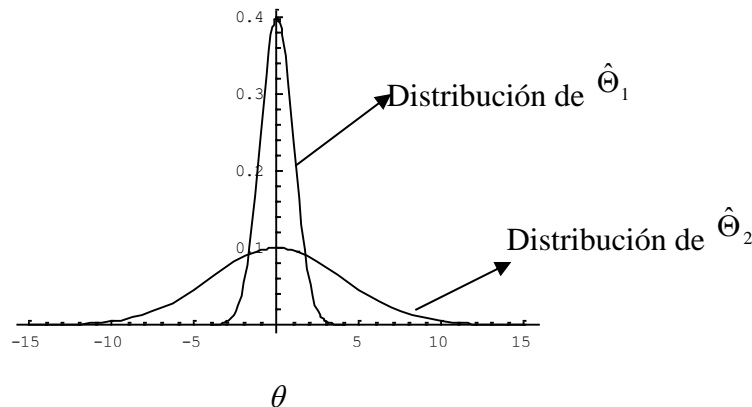
Como el sesgo es negativo el estimador tiende a subestimar el valor de verdadero parámetro

En ocasiones hay más de un estimador insesgado de un parámetro θ

Por lo tanto necesitamos un método para seleccionar un estimador entre varios estimadores insesgados.

Varianza y error cuadrático medio de un estimador puntual

Supongamos que $\hat{\theta}_1$ y $\hat{\theta}_2$ son dos estimadores insesgados de un parámetro θ . Esto indica que la distribución de cada estimador está centrada en el verdadero parámetro θ . Sin embargo las varianzas de estas distribuciones pueden ser diferentes. La figura siguiente ilustra este hecho.



Como $\hat{\theta}_1$ tiene menor varianza que $\hat{\theta}_2$, entonces es más probable que el estimador $\hat{\theta}_1$ produzca una estimación más cercana al verdadero valor de θ . Por lo tanto si tenemos dos estimadores insesgados se seleccionará aquel que tenga menor varianza.

Ejemplo: Sea X_1, X_2, \dots, X_n una muestra aleatoria de una v.a. X donde $E(X) = \mu$ y $V(X) = \sigma^2$. Suponemos μ desconocido.

Estimamos al parámetro μ con la **media o promedio muestral** $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Sabemos que es

un estimador insesgado de μ . Anotamos $\hat{\mu}_1 = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

Supongamos que tomamos otro estimador para μ , lo anotamos $\hat{\mu}_2 = \frac{X_1 + X_n}{2}$

Entonces como

$$E(\hat{\mu}_2) = E\left(\frac{X_1 + X_n}{2}\right) = \frac{1}{2}(E(X_1) + E(X_n)) = \frac{1}{2}(\mu + \mu) = \frac{1}{2}2\mu = \mu,$$

$\hat{\mu}_2 = \frac{X_1 + X_n}{2}$ es también un estimador insesgado de μ

¿Cuál de los dos estimadores es mejor?

Calculamos la varianza de cada uno utilizando las propiedades de la varianza.

Ya sabemos cuál es la varianza de $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ (se la halló para T.C.L.):

$$V(\bar{X}) = V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} V\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i),$$

donde en la última igualdad hemos tenido en cuenta que, por tratarse de una muestra aleatoria, las X_i con $i=1,2,\dots,n$ son variables aleatorias independientes y, en consecuencia, la varianza de la suma de ellas es la suma de las varianzas. Si tenemos en cuenta que además todas tienen la misma distribución que X y por lo tanto la misma varianza:

$$V(X_i) = V(X) = \sigma^2 \quad \forall i = 1, 2, \dots, n, \text{ tenemos}$$

$$V(\bar{X}) = \frac{1}{n^2} n \sigma^2 = \frac{\sigma^2}{n}.$$

Análogamente calculamos la varianza de $\hat{\mu}_2 = \frac{X_1 + X_n}{2}$:

$$V(\hat{\mu}_2) = V\left(\frac{X_1 + X_n}{2}\right) = \frac{1}{4}(V(X_1) + V(X_n)) = \frac{1}{4}(\sigma^2 + \sigma^2) = \frac{\sigma^2}{2}$$

Vemos que si $n > 2$ entonces $V(\hat{\mu}_1) < V(\hat{\mu}_2)$. Por lo tanto si $n > 2$ es mejor estimador $\hat{\mu}_1$

Supongamos ahora que $\hat{\Theta}_1$ y $\hat{\Theta}_2$ son dos estimadores de un parámetro θ y **alguno de ellos no es insesgado**.

A veces es necesario utilizar un estimador sesgado. En esos casos puede ser importante el **error cuadrático medio** del estimador.

El **error cuadrático medio** de un estimador $\hat{\Theta}$ de un parámetro θ está definido como

$$ECM(\hat{\Theta}) = E\left[(\hat{\Theta} - \theta)^2\right]$$

El error cuadrático medio puede escribirse de la siguiente forma:

$$ECM(\hat{\Theta}) = V(\hat{\Theta}) + (b(\hat{\Theta}))^2$$

Dem.) Por definición $ECM(\hat{\Theta}) = E\left[(\hat{\Theta} - \theta)^2\right]$. Sumamos y restamos el número $E(\hat{\Theta})$:

$$ECM(\hat{\Theta}) = E\left[(\hat{\Theta} - E(\hat{\Theta}) + E(\hat{\Theta}) - \theta)^2\right], \text{ y desarrollamos el cuadrado:}$$

$$ECM(\hat{\Theta}) = E\left[(\hat{\Theta} - E(\hat{\Theta}) + E(\hat{\Theta}) - \theta)^2\right] = E\left[(\hat{\Theta} - E(\hat{\Theta}))^2 + (E(\hat{\Theta}) - \theta)^2 + 2(\hat{\Theta} - E(\hat{\Theta}))(E(\hat{\Theta}) - \theta)\right] =$$

Aplicamos propiedades de la esperanza:

$$= E\left[(\hat{\Theta} - E(\hat{\Theta}))^2\right] + \underbrace{(E(\hat{\Theta}) - \theta)^2}_{b(\hat{\Theta})^2} + 2(E(\hat{\Theta}) - \theta) \underbrace{E(\hat{\Theta} - E(\hat{\Theta}))}_0 = V(\hat{\Theta}) + (b(\hat{\Theta}))^2$$

El error cuadrático medio es un criterio importante para comparar estimadores.

Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son dos estimadores de un parámetro θ .

La eficiencia relativa de $\hat{\theta}_2$ con respecto a $\hat{\theta}_1$ se define como $\frac{ECM(\hat{\theta}_1)}{ECM(\hat{\theta}_2)}$

Si la eficiencia relativa es menor que 1 entonces $\hat{\theta}_1$ tiene menor error cuadrático medio que $\hat{\theta}_2$

Por lo tanto $\hat{\theta}_1$ es más eficiente que $\hat{\theta}_2$

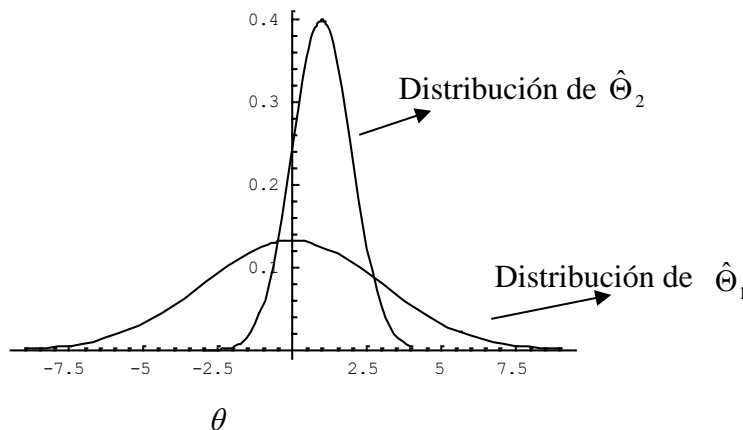
Observaciones:

1- Si $\hat{\theta}$ es un estimador insesgado de θ , entonces $ECM(\hat{\theta}) = V(\hat{\theta})$

2- A veces es preferible utilizar estimadores sesgados que estimadores insesgados, si es que tienen un error cuadrático medio menor.

En el error cuadrático medio se consideran tanto la varianza como el sesgo del estimador.

Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son dos estimadores de un parámetro θ , tales que $E(\hat{\theta}_1) = \theta$; $E(\hat{\theta}_2) \neq \theta$ y $V(\hat{\theta}_2) < V(\hat{\theta}_1)$, habría que calcular el error cuadrático medio de cada uno, y tomar el que tenga menor error cuadrático medio. Pues puede ocurrir que $\hat{\theta}_2$, aunque sea sesgado, al tener menor varianza tome valores más cercanos al verdadero parámetro que $\hat{\theta}_1$



Ejemplo:

Supóngase que $\hat{\theta}_1$, $\hat{\theta}_2$ y $\hat{\theta}_3$ son dos estimadores de un parámetro θ , y que $E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$; $E(\hat{\theta}_3) \neq \theta$, $V(\hat{\theta}_1) = 10$, $V(\hat{\theta}_2) = 6$ y $E[(\hat{\theta}_3 - \theta)^2] = 4$. Haga una comparación de estos estimadores. ¿Cuál prefiere y por qué?

Solución: Calculamos el error cuadrático medio de cada estimador

$ECM(\hat{\theta}_1) = V(\hat{\theta}_1) = 10$ pues $\hat{\theta}_1$ es insesgado

$ECM(\hat{\theta}_2) = V(\hat{\theta}_2) = 6$ pues $\hat{\theta}_2$ es insesgado

$$ECM(\hat{\Theta}_3) = E[(\hat{\Theta}_3 - \theta)^2] = 4 \text{ es dato}$$

En consecuencia $\hat{\Theta}_3$ es el mejor estimador de los tres dados porque tiene menor error cuadrático medio.

Consistencia de estimadores puntuales

Sea $\hat{\Theta}_n$ un estimador del parámetro θ , basado en una muestra aleatoria (X_1, X_2, \dots, X_n) de tamaño n . Se dice que $\hat{\Theta}_n$ es un estimador consistente de θ si

$$\lim_{n \rightarrow \infty} P(|\hat{\Theta}_n - \theta| \geq \varepsilon) = 0 \quad \text{para todo } \varepsilon > 0$$

Observación:

Este tipo de convergencia, que involucra a una sucesión de variables aleatorias, se llama **convergencia en probabilidad** y es la misma que consideramos en relación a la **ley de los grandes números**. Suele escribirse también $\hat{\Theta}_n \xrightarrow{P} \theta$.

Este tipo de convergencia debe distinguirse de la considerada en relación al teorema central del límite. En este último caso teníamos una sucesión de distribuciones: $F_{Z_n}(z) = P(Z_n \leq z)$ y se considera el límite $\lim_{n \rightarrow \infty} F_{Z_n}(z) = \lim_{n \rightarrow \infty} P(Z_n \leq z) = \Phi(z)$.

Se habla, entonces, de **convergencia en distribución** y suele indicarse $Z_n \xrightarrow{d} Z \sim N(0,1)$.

Teorema. Sea $\hat{\Theta}_n$ un estimador del parámetro θ basado en una muestra aleatoria (X_1, X_2, \dots, X_n) . Si $\lim_{n \rightarrow \infty} E(\hat{\Theta}_n) = \theta$ y $\lim_{n \rightarrow \infty} V(\hat{\Theta}_n) = 0$, entonces $\hat{\Theta}_n$ es un estimador consistente de θ .

Dem.)

Utilizamos la desigualdad de Chebyshev $\forall \varepsilon > 0$:

$$P(|\hat{\Theta}_n - \theta| \geq \varepsilon) \leq \frac{E(\hat{\Theta}_n - \theta)^2}{\varepsilon^2} = \frac{1}{\varepsilon^2} ECM(\hat{\Theta}_n) = \frac{1}{\varepsilon^2} [V(\hat{\Theta}_n) + b(\hat{\Theta}_n)^2]$$

Entonces, al tomar el límite $\lim_{n \rightarrow \infty}$ y teniendo presente que $\lim_{n \rightarrow \infty} E(\hat{\Theta}_n) = \theta$ y $\lim_{n \rightarrow \infty} V(\hat{\Theta}_n) = 0$, vemos que

$$\lim_{n \rightarrow \infty} P(|\hat{\Theta}_n - \theta| \geq \varepsilon) = 0 \quad \forall \varepsilon > 0, \text{ es decir } \hat{\Theta}_n \text{ es un estimador convergente de } \theta.$$

Ejemplo:

Sea X una variable aleatoria que describe alguna característica numérica de los individuos de una población y sean $\mu = E(X)$ y $\sigma^2 = V(X)$ la esperanza poblacional y la varianza

poblacional, respectivamente. Sea $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ la esperanza muestral basada en una muestra aleatoria (X_1, X_2, \dots, X_n) . Entonces \bar{X} es un estimador consistente de la esperanza poblacional $\mu = E(X)$.

Sabemos que

$$a) E(\bar{X}) = \mu = E(X) \quad \forall n$$

$$b) V(\bar{X}) = \frac{\sigma^2}{n} = \frac{V(X)}{n} \quad \forall n$$

La propiedad a) ya me dice que \bar{X} es un estimador insesgado de $\mu = E(X)$.

Por otra parte si a) vale para todo n , también vale en particular en el límite $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} E(\bar{X}) = \mu = E(X).$$

Además, de b) deducimos inmediatamente que

$$\lim_{n \rightarrow \infty} V(\bar{X}) = 0.$$

Por lo tanto vemos que \bar{X} es un estimador consistente de $\mu = E(X)$.

2.4 – Métodos de estimación puntual

Los criterios anteriores establecen propiedades que es deseable que sean verificadas por los estimadores. Entre dos estimadores posibles para un dado parámetro poblacional es razonable elegir aquél que cumple la mayor cantidad de criterios o alguno en particular que se considera importante para el problema que se esté analizando. Sin embargo estos criterios no nos enseñan por sí mismos a construir los estimadores. Existen una serie de métodos para construir estimadores los cuales en general se basan en principios básicos de razonabilidad. Entre éstos podemos mencionar:

- *Método de los momentos*

- *Método de máxima verosimilitud*

Método de los momentos

Se puede probar el siguiente resultado:

Ley débil de los grandes números:

Sean (X_1, X_2, \dots, X_n) n variables aleatorias independientes todas las cuales tienen la misma esperanza $\mu = E(X)$ y varianza $\sigma^2 = V(X)$. Sea $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Entonces

$$\lim_{n \rightarrow \infty} P(|\bar{X} - \mu| \geq \varepsilon) = 0$$

Decimos que \bar{X} converge a μ en probabilidad y lo indicamos: $\bar{X} \xrightarrow{p} \mu$.

Definimos los **momentos de orden k de una variable aleatoria** como:

$$\mu_k = E(X^k) = \sum_{x_i \in R_X} x_i^k p(x_i) \quad (k = 0, 1, 2, \dots) \quad \text{Si } X \text{ es discreta}$$

$$\mu_k = E(X^k) = \int_{-\infty}^{+\infty} x^k f(x) dx \quad (k = 0, 1, 2, \dots) \quad \text{Si } X \text{ es continua,}$$

y definimos los correspondientes momentos **muestrales de orden k** como:

$$M_k = \frac{1}{n} \sum_{i=1}^n X_i^k \quad (k = 0, 1, 2, \dots),$$

Entonces la ley débil de los grandes números se puede generalizar:

$$\lim_{n \rightarrow \infty} P(|M_k - \mu_k| \geq \varepsilon) = 0 \quad (k = 0, 1, 2, \dots).$$

De acuerdo con esto parece razonable estimar los momentos poblacionales de orden k mediante los momentos muestrales de orden k : $\mu_k \sim M_k \quad (k = 0, 1, 2, \dots)$.

Supongamos, entonces, una variable aleatoria X y supongamos que la distribución de X depende de r parámetros $\theta_1, \theta_2, \dots, \theta_r$, esto es la *fdp* poblacional es $p(x_i, \theta_1, \theta_2, \dots, \theta_r)$ si X es discreta o $f(x, \theta_1, \theta_2, \dots, \theta_r)$ si es continua. Sean $\mu_1, \mu_2, \dots, \mu_r$ los primeros r momentos poblacionales:

$$\mu_k = E(X^k) = \sum_{x_i \in R_X} x_i^k p(x_i, \theta_1, \theta_2, \dots, \theta_r) \quad (k = 1, 2, \dots, r) \quad \text{Si } X \text{ es discreta}$$

$$\mu_k = E(X^k) = \int_{-\infty}^{+\infty} x^k f(x, \theta_1, \theta_2, \dots, \theta_r) dx \quad (k = 1, 2, \dots, r) \quad \text{Si } X \text{ es continua,}$$

y sean

$M_k = \frac{1}{n} \sum_{i=1}^n X_i^k \quad (k = 1, 2, \dots, r)$ los r primeros momentos muestrales para una muestra de tamaño n (X_1, X_2, \dots, X_n). Entonces **el método de los momentos consiste en plantear el sistema de ecuaciones:**

$$\begin{cases} \mu_1 = M_1 \\ \mu_2 = M_2 \\ \vdots \\ \mu_r = M_r \end{cases}$$

Es decir

$$\left\{ \begin{array}{l} \sum_{x_i \in R_X} x_i p(x_i, \theta_1, \theta_2, \dots, \theta_r) = \frac{1}{n} \sum_{i=1}^n X_i^1 \\ \sum_{x_i \in R_X} x_i^2 p(x_i, \theta_1, \theta_2, \dots, \theta_r) = \frac{1}{n} \sum_{i=1}^n X_i^2 \\ \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ \sum_{x_i \in R_X} x_i^r p(x_i, \theta_1, \theta_2, \dots, \theta_r) = \frac{1}{n} \sum_{i=1}^n X_i^r \end{array} \right. \quad \text{Si } X \text{ es discreta,}$$

o

$$\left\{ \begin{array}{l} \int_{-\infty}^{+\infty} x f(x, \theta_1, \theta_2, \dots, \theta_r) dx = \frac{1}{n} \sum_{i=1}^n X_i^1 \\ \int_{-\infty}^{+\infty} x^2 f(x, \theta_1, \theta_2, \dots, \theta_r) dx = \frac{1}{n} \sum_{i=1}^n X_i^2 \\ \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ \int_{-\infty}^{+\infty} x^r f(x, \theta_1, \theta_2, \dots, \theta_r) dx = \frac{1}{n} \sum_{i=1}^n X_i^r \end{array} \right. \quad \text{Si } X \text{ es continua.}$$

Resolviendo este sistema de ecuaciones para los parámetros desconocidos $\theta_1, \theta_2, \dots, \theta_r$ en función de la muestra aleatoria (X_1, X_2, \dots, X_n) obtenemos los estimadores:

$$\left\{ \begin{array}{l} \hat{\Theta}_1 = H_1(X_1, X_2, \dots, X_n) \\ \hat{\Theta}_2 = H_2(X_1, X_2, \dots, X_n) \\ \vdots \\ \hat{\Theta}_r = H_r(X_1, X_2, \dots, X_n) \end{array} \right.$$

Observación:

En la forma que presentamos aquí el método necesitamos conocer la forma de la *fdp* poblacional, por lo tanto estamos frente a un caso de estimación puntual **paramétrica**.

Ejemplos:

1- Sea X una variable aleatoria. Supongamos que X tiene **distribución gama con parámetros σ y λ** : $X \sim \Gamma(\sigma, \lambda)$, es decir su *fdp* está dada por:

$$f(x) = \begin{cases} \frac{1}{\sigma \Gamma(\lambda)} \left(\frac{x}{\sigma} \right)^{\lambda-1} e^{-\frac{x}{\sigma}} & x > 0 \\ 0 & \text{demás valores} \end{cases}$$

con $\sigma > 0$; $\lambda > 0$ y $\Gamma(\lambda) = \int_0^{\infty} x^{\lambda-1} e^{-x} dx$.

Sea (X_1, X_2, \dots, X_n) una muestra aleatoria de tamaño n . Deseamos calcular los estimadores de σ y λ dados por el método de los momentos.

Solución:

Como tenemos dos parámetros desconocidos a estimar, planteamos el sistema de ecuaciones:

$$\begin{cases} \mu_1 = M_1 \\ \mu_2 = M_2 \end{cases}$$

Se puede probar que

$$\begin{aligned} \mu_1 &= \lambda \cdot \sigma \\ \mu_2 &= \lambda^2 \cdot \sigma^2 + \lambda \cdot \sigma^2 \end{aligned}$$

Tenemos, entonces, el sistema de ecuaciones

$$\begin{cases} \lambda \cdot \sigma = \frac{1}{n} \sum_{i=1}^n X_i \\ \lambda^2 \cdot \sigma^2 + \lambda \cdot \sigma^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 \end{cases} \Rightarrow \begin{cases} \lambda \cdot \sigma = \bar{X} \\ \lambda^2 \cdot \sigma^2 + \lambda \cdot \sigma^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 \end{cases}$$

Reemplazando en la segunda ecuación: $\bar{X}^2 + \sigma \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i^2 \Rightarrow \sigma = \frac{\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2}{\bar{X}}$

Y despejando λ de la primera ecuación y reemplazando la expresión hallada para σ

$$\begin{cases} \hat{\lambda} = \frac{n \bar{X}^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \\ \hat{\sigma} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n \bar{X}} \end{cases}$$

2- Sea (X_1, X_2, \dots, X_n) una muestra aleatoria de tamaño n de una v.a. X donde $X \sim U[0, \theta]$, θ desconocido. Hallar el estimador de θ por el método de los momentos.

Solución:

Planteamos la ecuación: $\mu_1 = M_1$

Sabemos que $\mu_1 = E(X) = \frac{0+\theta}{2} = \frac{\theta}{2}$. Entonces $\frac{\theta}{2} = \bar{X} \Rightarrow \hat{\Theta} = 2\bar{X}$

Observación: notar que el estimador $\hat{\Theta} = 2\bar{X}$ es un estimador consistente de θ , pues

$$E(\hat{\Theta}) = E(2\bar{X}) = 2E(\bar{X}) = 2\frac{\theta}{2} = \theta \quad \text{y} \quad V(\hat{\Theta}) = V(2\bar{X}) = 4V(\bar{X}) = 4\frac{(\theta-0)^2}{12n} = \frac{\theta^2}{3n} \xrightarrow{n \rightarrow \infty} 0$$

3- Sea (X_1, X_2, \dots, X_n) una muestra aleatoria de una v.a. $X \sim N(\mu, \sigma^2)$.

Encuentra los estimadores de μ y σ por el método de momentos.

Solución:

Planteamos las ecuaciones

$$\begin{cases} \mu_1 = M_1 \\ \mu_2 = M_2 \end{cases} \Rightarrow \begin{cases} \mu = \bar{X} \\ E(X^2) = \frac{1}{n} \sum_{i=1}^n X_i^2 \end{cases}$$

pero en general es válido que $V(X) = E(X^2) - \mu^2 \Rightarrow E(X^2) = V(X) + \mu^2$

Entonces las ecuaciones quedan

$$\begin{cases} \mu = \bar{X} \\ \sigma^2 + \mu^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 \end{cases} \Rightarrow \begin{cases} \hat{\mu} = \bar{X} \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 \end{cases}$$

4- Sea (X_1, X_2, \dots, X_n) una muestra aleatoria de una v.a. $X \sim N(0, \sigma^2)$.

Hallar un estimador por el método de los momentos de σ^2

Solución: en este caso no es conveniente plantear $\mu_1 = M_1$ pues quedaría

la ecuación $0 = \bar{X}$ que no conduce a nada.

Entonces podemos plantear $\mu_2 = M_2$ es decir

$$E(X^2) = \frac{1}{n} \sum_{i=1}^n X_i^2 \Rightarrow \sigma^2 + 0 = \frac{1}{n} \sum_{i=1}^n X_i^2 \Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2$$

Observación: si $\hat{\Theta}$ es un estimador por el método de los momentos de un parámetro θ , el estimador de los momentos de $g(\theta)$ es $g(\hat{\Theta})$, si $g(x)$ es una función inyectiva.

Por ejemplo, en el ejemplo anterior un estimador de σ por el método de los momentos sería

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2}. \text{ Notar que } g(x) = \sqrt{x} \text{ es inyectiva para los reales positivos.}$$

Método de máxima verosimilitud

Uno de los mejores métodos para obtener un estimador puntual de un parámetro es el método de máxima verosimilitud.

Supongamos que X es una v.a. discreta con función de distribución de probabilidad $p(x, \theta)$, donde θ es un parámetro desconocido. Sean x_1, x_2, \dots, x_n los valores observados de una muestra aleatoria de tamaño n .

Se define la **función de verosimilitud** como la función de distribución conjunta de las observaciones:

$$L(x_1, x_2, \dots, x_n, \theta) = P(X_1 = x_1)P(X_2 = x_2) \dots P(X_n = x_n) = p(x_1, \theta) \cdot p(x_2, \theta) \dots p(x_n, \theta)$$

Notar que la función de verosimilitud es una función de θ .

El estimador de máxima verosimilitud de θ es aquel valor de θ que maximiza la función de verosimilitud

La interpretación del método sería: el estimador de máxima verosimilitud es aquel valor del parámetro que maximiza la probabilidad de ocurrencia de los valores muestrales

La adaptación para el caso en que X es una v.a. continua sería la siguiente

Supongamos que X es una v.a. continua con función de densidad de probabilidad $f(x, \theta)$, donde θ es un parámetro desconocido. Sean x_1, x_2, \dots, x_n los valores observados de una muestra aleatoria de tamaño n .

Se define la **función de verosimilitud** como la función de distribución conjunta de las observaciones:

$$L(x_1, x_2, \dots, x_n, \theta) = f(x_1, \theta) \cdot f(x_2, \theta) \dots f(x_n, \theta)$$

La función de verosimilitud es una función de θ .

El estimador de máxima verosimilitud de θ es aquel valor de θ que maximiza la función de verosimilitud

Notación: abreviamos estimador de máxima verosimilitud con EMV

Ejemplos:

1- Sea (X_1, X_2, \dots, X_n) una muestra aleatoria de una v.a. $X \sim B(1, p)$

Por ejemplo, se eligen al azar n objetos de una línea de producción, y cada uno se clasifica como defectuoso (en cuyo caso $x_i = 1$) o no defectuoso (en cuyo caso $x_i = 0$).

Entonces $p = P(X_i = 1)$, es decir es la verdadera proporción de objetos defectuosos en la producción total.

Queremos hallar el EMV de p

Solución:

Si $X \sim B(1, p)$ entonces $P(X = k) = \binom{1}{k} p^k (1-p)^{1-k} \quad k = 0, 1$

Planteamos la función de verosimilitud

$$L(x_1, x_2, \dots, x_n; p) = p(x_1; p) p(x_2; p) \dots p(x_n; p) = [p^{x_1} (1-p)^{1-x_1}] [p^{x_2} (1-p)^{1-x_2}] \dots [p^{x_n} (1-p)^{1-x_n}]$$

Esto puede escribirse:

$$L(x_1, x_2, \dots, x_n; p) = p^{\sum_{i=1}^n x_i} (1-p)^{n - \sum_{i=1}^n x_i}$$

Para maximizar la función de verosimilitud y **facilitar los cálculos** tomamos el logaritmo natural de L . Pues maximizar L es equivalente a maximizar $\ln(L)$ y al tomar logaritmos transformamos productos en sumas.

Entonces

$$\ln(L(x_1, x_2, \dots, x_n; p)) = \left(\sum_{i=1}^n x_i \right) \ln p + \left(n - \sum_{i=1}^n x_i \right) \ln(1-p)$$

Y ahora podemos maximizar la función derivando e igualando a cero

$$\frac{\partial \ln L(x_1, x_2, \dots, x_n; p)}{\partial p} = \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p} = 0$$

de donde despejando p

$$p = \frac{\sum_{i=1}^n x_i}{n} = \bar{x} \quad \text{la proporción de defectuosos en la muestra}$$

Por lo tanto se toma como estimador a $\hat{p} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

2- El tiempo de fallar T de una componente tiene una distribución exponencial con parámetro λ :

$T \sim \text{Exp}(\lambda)$, es decir la fdp es

$$f(t; \lambda) = \begin{cases} \lambda e^{-\lambda t} & 0 \leq t < \infty \\ 0 & \text{demás valores} \end{cases}$$

Recordemos que la esperanza y varianza son:

$$E(T) = 1/\lambda \quad \text{y} \quad V(T) = 1/\lambda^2, \text{ respectivamente.}$$

Se desea calcular el estimador de máxima verosimilitud del parámetro λ para una muestra de tamaño n .

Solución:

La función de probabilidad es:

$$L(t_1, t_2, \dots, t_n; \lambda) = f(t_1; \lambda) f(t_2; \lambda) \dots f(t_n; \lambda) = [\lambda e^{-\lambda t_1}] \times [\lambda e^{-\lambda t_2}] \times \dots \times [\lambda e^{-\lambda t_n}],$$

que puede escribirse:

$$L(t_1, t_2, \dots, t_n; \lambda) = (\lambda)^n e^{-\lambda \sum_{i=1}^n t_i}$$

Nuevamente tomamos logaritmo natural

$$\ln L(t_1, t_2, \dots, t_n; \lambda) = n \ln \lambda - \lambda \sum_{i=1}^n t_i$$

$$\frac{\partial \ln L(t_1, t_2, \dots, t_n; \lambda)}{\partial \lambda} = n \frac{1}{\lambda} - \sum_{i=1}^n T_i = 0$$

de donde podemos despejar λ :

$$\lambda = \frac{n}{\sum_{i=1}^n t_i} = \bar{t}, \quad \text{entonces el estimador de } \lambda \text{ es } \hat{\lambda} = \frac{n}{\sum_{i=1}^n T_i}$$

El método de máxima verosimilitud presenta, algunas veces, dificultades para maximizar la función de verosimilitud debido a que la ecuación obtenida a partir de $\frac{d}{d\theta} L(\theta) = 0$ no resulta fácil de resolver. O también puede ocurrir que los métodos de cálculo para maximizar $L(\theta)$ no son aplicables.

Por ejemplo:

Sea (X_1, X_2, \dots, X_n) una muestra aleatoria de tamaño n de una v.a. X donde $X \sim U[0, \theta]$, θ desconocido. Hallar el estimador de θ por el método máxima verosimilitud.

Solución:

La f.d.p. de X es

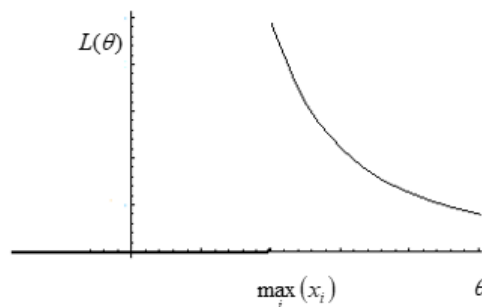
$$f(x) = \begin{cases} \frac{1}{\theta} & \text{si } 0 < x < \theta \\ 0 & \text{caso contrario} \end{cases}$$

Planteamos la función de verosimilitud

$$L(x_1, x_2, \dots, x_n, \theta) = \begin{cases} \frac{1}{\theta^n} & \text{si } 0 < x_i < \theta \quad \forall i \\ 0 & \text{caso contrario} \end{cases} = \begin{cases} \frac{1}{\theta^n} & \text{si } \max_i(x_i) < \theta \\ 0 & \text{caso contrario} \end{cases}$$

Si derivamos con respecto a θ obtenemos $\frac{d}{d\theta} \theta^{-n} = -\frac{n}{\theta^{n+1}}$ que es siempre menor que cero. Por lo tanto la función de verosimilitud es una función decreciente para todos los $\theta > \max_i(x_i)$

Si hacemos un gráfico de la función de verosimilitud



Vemos que donde la función tiene el máximo hay una discontinuidad no evitable.

Por lo tanto $\hat{\Theta} = \max_i(x_i)$

El método de máxima verosimilitud puede emplearse en el caso donde hay más de un parámetro desconocido para estimar. En ese caso la función de verosimilitud es una función de varias variables.

Específicamente si tenemos para estimar k parámetros $\theta_1, \theta_2, \dots, \theta_k$, entonces la función de verosimilitud es una función de k variables $L(x_1, x_2, \dots, x_n, \theta_1, \theta_2, \dots, \theta_k)$ y los estimadores de máxima verosimilitud $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$ se obtienen al plantear (si existen las derivadas parciales) y resolver el sistema de k ecuaciones con k incógnitas $\theta_1, \theta_2, \dots, \theta_k$

$$\frac{d}{d\theta_i} L(x_1, x_2, \dots, x_n, \theta_1, \theta_2, \dots, \theta_k) = 0 \quad i = 1, 2, \dots, k$$

Ejemplo:

La variable aleatoria X tiene distribución $N(\mu, \sigma^2)$ con μ y σ^2 ambos parámetros desconocidos para los cuales se desea encontrar los estimadores máxima verosimilitud. La *fdp* es

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty,$$

La función de verosimilitud para una muestra aleatoria de tamaño n es

$$\begin{aligned} L(x_1, x_2, \dots, x_n; \mu, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x_2-\mu}{\sigma}\right)^2} \dots \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x_n-\mu}{\sigma}\right)^2} = \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^n \left(\frac{x_i-\mu}{\sigma}\right)^2} \end{aligned}$$

Luego

$$\ln L(x_1, x_2, \dots, x_n; \mu, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma}\right)^2$$

y el sistema de ecuaciones de verosimilitud queda:

$$\begin{cases} \frac{\partial \ln L(x_1, x_2, \dots, x_n; \mu, \sigma^2)}{\partial \mu} = \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma}\right) = 0 \\ \frac{\partial \ln L(x_1, x_2, \dots, x_n; \mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2} \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^4} = 0 \end{cases}$$

Resolvemos con respecto a μ y σ^2 :

$$\begin{cases} \mu = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x} \\ \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \end{cases}$$

Entonces los estimadores máxima verosimilitud de μ y σ^2 son

$$\begin{cases} \hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X} \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \end{cases}$$

Propiedades de los estimadores máxima verosimilitud

- 1- Los EMV pueden ser *sesgados*, pero en general si $\hat{\Theta}$ es el EMV de un parámetro θ basado en una muestra de tamaño n , entonces $\lim_{n \rightarrow \infty} E(\hat{\Theta}) = \theta$, es decir son *asintóticamente insesgados*
- 2- Bajo condiciones bastantes generales se puede probar que los EMV son *consistentes*
- 3- Bajo condiciones bastantes generales se puede probar que los EMV *asintóticamente tienen varianza mínima*
- 4- Los EMV cumplen la *propiedad de invarianza* es decir:
si $\hat{\Theta}$ es un EMV de un parámetro θ , el EMV de $g(\theta)$ es $g(\hat{\Theta})$, si $g(x)$ es una función inyectiva.

Ejemplos:

1- Si consideramos nuevamente la situación considerada en el Ejemplo 2, donde teníamos una v.a. T cuya distribución es una exponencial: $T \sim \text{Exp}(\lambda)$, entonces, si queremos el EMV de la varianza poblacional, podemos calcularlo recordando que $V(T) = 1/\lambda^2$, es decir,

$$V(T) = g(\lambda) = 1/\lambda^2. \text{ Vimos que } \hat{\lambda} = \frac{n}{\sum_{i=1}^n T_i} = \frac{1}{\bar{T}}. \text{ Por lo tanto el EMV de la varianza es } \hat{\sigma}^2 = \frac{1}{\hat{\lambda}^2}$$

.

2- Sea X_1, X_2, \dots, X_n una muestra aleatoria de una v.a. $B(1, p)$. Un EMV de p es

$$\hat{p} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Se selecciona una muestra aleatoria de n cascos para ciclistas fabricados por cierta compañía. Sea X : “el número entre los n que tienen defectos”, y $p = P(\text{el casco tiene defecto})$. Supongamos que solo se observa X (el número de cascos con defectos).

Si $n = 20$ y $x = 3$, es la estimación de p es $\hat{p} = \frac{3}{20}$

El E.M.V. de la probabilidad $(1-p)^5$, de que ninguno de los siguientes cinco cascos que se examinen tenga defectos será $(1-\hat{p})^5$ y su estimación en este caso $\left(1 - \frac{3}{20}\right)^5$

Práctica**Estimación puntual****Método de máxima verosimilitud - Método de momentos**

- 1) Suponga que se tiene una muestra aleatoria de tamaño $2n$ tomada de una población X , que $E(X)=\mu$ y $V(X)=\sigma^2$. Sean

$$\bar{X}_1 = \frac{1}{2n} \sum_{i=1}^{2n} X_i \quad \text{y} \quad \bar{X}_2 = \frac{1}{n} \sum_{i=1}^n X_i$$

Dos estimadores de μ . ¿Cuál es el mejor estimador de μ ? Explique su elección.

- 2) Sea X_1, X_2, \dots, X_7 una muestra aleatoria de una población que tiene media μ y varianza σ^2 . Considere los siguientes estimadores de μ :

$$\hat{\Theta}_1 = \frac{X_1 + X_2 + \dots + X_7}{7} \quad \text{y} \quad \hat{\Theta}_2 = \frac{2X_1 - X_6 + X_4}{2}$$

- ¿Alguno de estos estimadores es insesgado?
 - ¿Cuál estimador es el “mejor”? ¿En qué sentido es mejor?
- 3) Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n .
- Demuestre que \bar{X}^2 es un estimador sesgado de μ^2 .
 - Determine la magnitud del sesgo de este estimador.
 - ¿Qué sucede con el sesgo a medida que aumenta el tamaño n de la muestra?
- 4) El número diario de desconexiones accidentales de un servidor sigue una distribución de Poisson. En cinco días se observan: 2, 5, 3, 3, 7 desconexiones accidentales.
- Obtenga el estimador de máxima verosimilitud de λ . Obtenga la estimación de λ a partir de la muestra dada.
 - Encuentre el estimador de máxima verosimilitud de la probabilidad de que ocurran 3 o más desconexiones accidentales y encuentre la estimación de dicha probabilidad a partir de los datos.
- 5) a) Sea X_1, X_2, \dots, X_n una muestra aleatoria de una v.a. $B(1, p)$. Hallar un estimador de máxima verosimilitud (E.M.V.) de p .
- b) Se selecciona una muestra aleatoria de n cascos para ciclistas fabricados por cierta compañía. Sea $\sum_{i=1}^n X_i$ el número entre los n que tienen defectos y $p = P(\text{el casco tiene defecto})$. Supongamos que solo se observa $\sum_{i=1}^n X_i$ (el número de cascos con defectos).
- Si $n = 20$ y $\sum_{i=1}^n X_i = 3$, ¿cuál es la estimación de p ?
 - Si $n = 20$ y $\sum_{i=1}^n X_i = 3$, ¿cuál es el E.M.V. de la probabilidad $(1-p)^5$, de que ninguno de los siguientes cinco cascos que se examinen tenga defectos?

- 6) Denotemos por X la proporción de tiempo asignado que un estudiante seleccionado al azar emplea trabajando en cierta prueba de actitud, y supongamos que la f.d.p. de X

es:

$$f(x) = \begin{cases} (\theta + 1)x^\theta, & 0 \leq x \leq 1 \\ 0, & \text{c.c} \end{cases} \quad \text{donde } \theta > -1$$

Una muestra aleatoria de diez estudiantes produce la siguiente información:

0.92, 0.79, 0.90, 0.65, 0.86, 0.47, 0.73, 0.97, 0.94, 0.77.

- a) Utilice el método de los momentos para obtener un estimador de θ y luego calcule la estimación para esta información.
- b) Obtenga el E.M.V. de θ y luego calcule la estimación para la información dada.

7) Sea X_1, X_2, \dots, X_n una muestra aleatoria de una v.a. $N(\mu, \sigma^2)$.

- a) Hallar los estimadores de μ y σ por el método de momentos.
- b) Hallar los estimadores de μ y σ por el método de máxima verosimilitud.
- c) Se determina la resistencia al corte de cada una de diez soldaduras eléctricas por puntos de prueba, dando los siguientes datos (lb/plg²):
392, 376, 401, 367, 389, 362, 409, 415, 358, 375.
Si se supone que la resistencia al corte esta normalmente distribuida, estime el verdadero promedio de resistencia al corte y desviación estándar de resistencia al corte usando el método de máxima verosimilitud y el método de momentos.

8) En una prueba 294 de 300 aisladores cerámicos soportaron cierto choque térmico.

- a) Obtenga el estimador y la estimación de máxima verosimilitud de la probabilidad de que un aislante cerámico sobrevivirá a un choque térmico.
- b) Suponga que un dispositivo contiene tres aislantes cerámicos y todos deben sobrevivir al choque, con la finalidad de que el dispositivo funcione. Encuentre el estimador y la estimación de máxima verosimilitud de la probabilidad de que los tres sobrevivirán a un choque térmico.

3-Intervalos de confianza

3.1 – Introducción

Se ha visto como construir a partir de una muestra aleatoria un estimador puntual de un parámetro desconocido. En esos casos necesitábamos dar algunas características del estimador, como por ejemplo si era insesgado o su varianza.

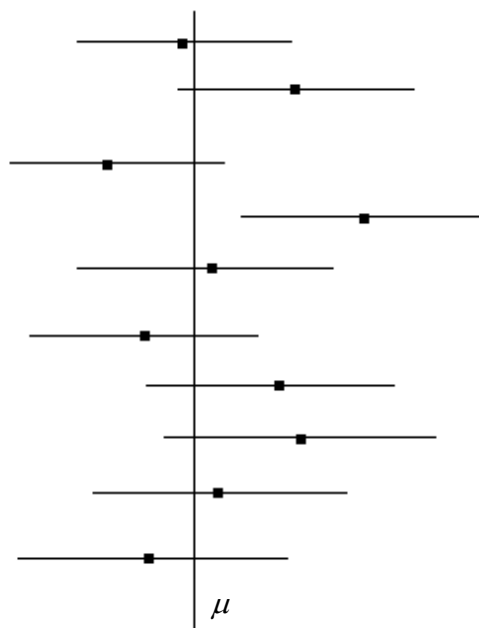
A veces resulta más conveniente dar un **intervalo de valores posibles** del parámetro desconocido, de manera tal que dicho intervalo contenga al verdadero parámetro con determinada probabilidad.

Específicamente, a partir de una muestra aleatoria se construye un intervalo $(\hat{\theta}_1, \hat{\theta}_2)$ donde los extremos $\hat{\theta}_1$ y $\hat{\theta}_2$ son dos estadísticos, tal que $P(\theta \in (\hat{\theta}_1, \hat{\theta}_2)) = 1 - \alpha$ donde θ es el parámetro desconocido a estimar y α es un valor real entre cero y uno dado de antemano. Por ejemplo si $\alpha = 0.05$, se quiere construir un intervalo $(\hat{\theta}_1, \hat{\theta}_2)$ tal que $P(\theta \in (\hat{\theta}_1, \hat{\theta}_2)) = 0.95$, o escrito de otra forma $P(\hat{\theta}_1 \leq \theta \leq \hat{\theta}_2) = 0.95$

Esta probabilidad tiene el siguiente significado: como $\hat{\theta}_1$ y $\hat{\theta}_2$ son estadísticos, los valores que ellos toman varían con los valores de la muestra, es decir si x_1, x_2, \dots, x_n son los valores medidos de la muestra entonces el estadístico $\hat{\theta}_1$ tomará el valor θ_1 y el estadístico $\hat{\theta}_2$ tomará el valor θ_2 . Si medimos nuevamente la muestra obtendremos ahora valores x_1', x_2', \dots, x_n' y por lo tanto $\hat{\theta}_1$ tomará el valor θ_1' y el estadístico $\hat{\theta}_2$ tomará el valor θ_2' , diferentes en general de los anteriores. Esto significa que si medimos la muestra 100 veces obtendremos 100 valores diferentes para $\hat{\theta}_1$ y $\hat{\theta}_2$ y por lo tanto obtendremos 100 intervalos distintos, de los cuales aproximadamente 5 de ellos no contendrán al verdadero parámetro.

Al valor $1 - \alpha$ se lo llama **nivel de confianza** del intervalo. También se suele definir como nivel de confianza al $(1 - \alpha)100\%$

La construcción repetida de un intervalo de confianza para μ se ilustra en la siguiente figura



3.2 – Intervalo de confianza para la media de una distribución normal, varianza conocida.

El método general para construir intervalos de confianza es el siguiente llamado *método del pivote*:

Supongamos el siguiente caso particular, sea (X_1, X_2, \dots, X_n) una muestra aleatoria de tamaño n de una v.a. X donde $X \sim N(\mu, \sigma^2)$, σ^2 conocido, se quiere construir un intervalo de confianza para μ de nivel $1 - \alpha$. Supongamos $\alpha = 0.05$.

1- tomamos un estimador puntual de μ , sabemos que $\hat{\mu} = \bar{X}$ es un estimador con buenas propiedades.

2- a partir de $\hat{\mu} = \bar{X}$ construimos el estadístico $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$. Notar que Z (*pivote*) contiene al

verdadero parámetro μ y que bajo las condiciones dadas $Z \sim N(0,1)$

3- como conocemos la distribución de Z , podemos plantear: hallar un número z tal que $P(-z \leq Z \leq z) = 0.95$

Por la simetría de la distribución normal estándar podemos escribir

$$P(-z \leq Z \leq z) = \Phi(z) - \Phi(-z) = 2\Phi(z) - 1 = 0.95 \Rightarrow \Phi(z) = 0.975 \Rightarrow z = 1.96$$

$$\text{Por lo tanto } P(-1.96 \leq Z \leq 1.96) = P\left(-1.96 \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq 1.96\right) = 0.95$$

Despejamos μ :

$$\begin{aligned} P\left(-1.96 \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq 1.96\right) &= P\left(-1.96 \frac{\sigma}{\sqrt{n}} \leq \bar{X} - \mu \leq 1.96 \frac{\sigma}{\sqrt{n}}\right) = \\ &= P\left(-1.96 \frac{\sigma}{\sqrt{n}} - \bar{X} \leq -\mu \leq 1.96 \frac{\sigma}{\sqrt{n}} - \bar{X}\right) = P\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}\right) = 0.95 \end{aligned}$$

Entonces

$$P\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}\right) = P\left(\mu \in \left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}}; \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}\right)\right) = 0.95$$

Es decir el intervalo de confianza para μ es $\left(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}}; \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}\right)$ y tiene nivel de confianza 0.95 o 95%.

$$\text{Aquí } \hat{\Theta}_1 = \bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} \text{ y } \hat{\Theta}_2 = \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}$$

Repetimos el procedimiento anterior y construimos un intervalo de confianza para μ con nivel de confianza $1 - \alpha$

1-Partimos de la esperanza muestral $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ para una muestra aleatoria (X_1, X_2, \dots, X_n) de tamaño n . Sabemos que es un estimador insesgado y consistente de μ .

2-Construimos el estadístico

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

La variable aleatoria Z cumple las condiciones necesarias de un pivote

Para construir un intervalo de confianza al nivel de confianza $1-\alpha$ partiendo del pivote Z , comenzamos por plantear la ecuación

$$P(-z \leq Z \leq z) = 1-\alpha,$$

donde la incógnita es el número real z .

Si reemplazamos la v.a. Z por su expresión tenemos:

$$P\left(-z \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq z\right) = P\left(-z \frac{\sigma}{\sqrt{n}} \leq \bar{X} - \mu \leq z \frac{\sigma}{\sqrt{n}}\right) = P\left(-\bar{X} - z \frac{\sigma}{\sqrt{n}} \leq -\mu \leq -\bar{X} + z \frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

Multiplicando todos los miembros de la desigualdad por -1 (el orden de los miembros se invierte) llegamos a:

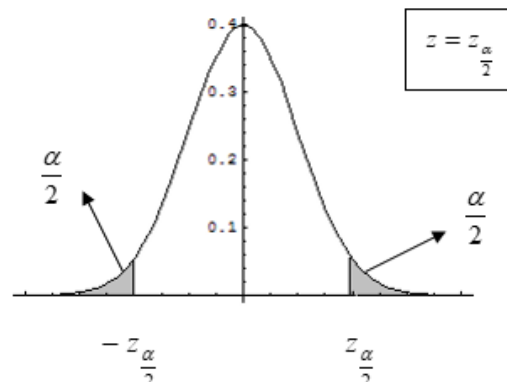
$$P\left(\bar{X} - z \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z \frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

Evidentemente, si definimos

$$\begin{cases} \hat{\Theta}_1 = \bar{X} - z \frac{\sigma}{\sqrt{n}} \\ \hat{\Theta}_2 = \bar{X} + z \frac{\sigma}{\sqrt{n}} \end{cases}, \text{ hemos construido dos estadísticos } \hat{\Theta}_1 \text{ y } \hat{\Theta}_2 \text{ tales que } P(\hat{\Theta}_1 \leq \mu \leq \hat{\Theta}_2) =$$

$1-\alpha$,

es decir hemos construido el intervalo de confianza bilateral deseado $[\hat{\Theta}_1, \hat{\Theta}_2]$. Todos los elementos que forman los estadísticos $\hat{\Theta}_1$ y $\hat{\Theta}_2$ son conocidos ya que el número z verifica la ecuación anterior, es decir (ver figura):



$$P(-z \leq Z \leq z) = \Phi(z) - \Phi(-z) = 1 - \alpha \quad \text{donde } \Phi(z) \text{ es la Fda para la v.a. } Z \sim N(0,1)$$

Recordando que $\Phi(-z) = 1 - \Phi(z)$, esta ecuación queda:

$$\Phi(z) - \Phi(-z) = 2\Phi(z) - 1 = 1 - \alpha, \quad \text{o bien (ver figura anterior),}$$

$$\Phi(z) = 1 - \frac{\alpha}{2} \quad \text{o de otra forma } P(Z > z) = \frac{\alpha}{2}.$$

Al valor de z que verifica esta ecuación se lo suele indicar $z_{\frac{\alpha}{2}}$. En consecuencia, el intervalo de confianza bilateral al nivel de significación $1 - \alpha$ queda:

$$[\hat{\Theta}_1, \hat{\Theta}_2] = \left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$$

En consecuencia:

Si (X_1, X_2, \dots, X_n) una muestra aleatoria de tamaño n de una v.a. X donde $X \sim N(\mu, \sigma^2)$, σ^2 conocido, un intervalo de confianza para μ de nivel $1 - \alpha$ es

$$\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] \quad (8.1)$$

Ejemplo:

Un ingeniero civil analiza la resistencia a la compresión del concreto. La resistencia está distribuida aproximadamente de manera normal, con varianza 1000 (psi)^2 . Al tomar una muestra aleatoria de 12 especímenes, se tiene que $\bar{x} = 3250 \text{ psi}$.

- Construya un intervalo de confianza del 95% para la resistencia a la compresión promedio.
- Construya un intervalo de confianza del 99% para la resistencia a la compresión promedio. Compare el ancho de este intervalo de confianza con el ancho encontrado en el inciso a).

Solución:

La v. a. de interés es X_i : “resistencia a la compresión del concreto en un espécimen i ”
Tenemos una muestra de $n = 12$ especímenes.

Asumimos que $X_i \sim N(\mu, \sigma^2)$ para $i = 1, 2, 3, \dots, 12$ con $\sigma^2 = 1000$

a) Queremos un intervalo de confianza para μ de nivel 95%. Por lo tanto $\alpha = 0.05$

El intervalo a utilizar es $\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$.

Buscamos en la tabla de la normal estándar el valor de $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

Reemplazando:

$$\left[3250 - 1.96 \times \frac{\sqrt{1000}}{\sqrt{12}}, 3250 + 1.96 \times \frac{\sqrt{1000}}{\sqrt{12}} \right] = \left[3232.10773, 3267.89227 \right]$$

b) repetimos lo anterior pero ahora $\alpha = 0.01$

El intervalo a utilizar es $\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$.

Buscamos en la tabla de la normal estándar el valor de $z_{\frac{\alpha}{2}} = z_{0.005} = 2.58$

Reemplazando:

$$\left[3250 - 2.58 \times \frac{\sqrt{1000}}{\sqrt{12}}, 3250 + 2.58 \times \frac{\sqrt{1000}}{\sqrt{12}} \right] = \left[3226.44793, 3273.55207 \right]$$

La longitud del intervalo encontrado en a) es: 35.78454

La longitud del intervalo encontrado en b) es: 47.10414

Notar que la seguridad de que el verdadero parámetro se encuentre en el intervalo hallado es mayor en el intervalo b) que en el a), pero la longitud del intervalo b) es mayor que la del intervalo a).

Al aumentar el nivel de confianza se perdió **precisión en la estimación**, ya que a menor longitud hay mayor precisión en la estimación.

En general la longitud del intervalo es $L = 2z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

Notar que:

a) si n y σ están fijos, a medida que α disminuye tenemos que $z_{\frac{\alpha}{2}}$ aumenta, por lo tanto

L aumenta.

b) si α y σ están fijos, entonces a medida que n aumenta tenemos que L disminuye.

Podemos plantearnos la siguiente pregunta relacionada con el ejemplo anterior: ¿qué tamaño n de muestra se necesita para que el intervalo tenga nivel de confianza 99% y longitud la mitad de la longitud del intervalo hallado en a)?

Solución: el intervalo hallado en a) tiene longitud 35.78454, y queremos que el nuevo intervalo tenga longitud 17.89227 aproximadamente. Planteamos:

$$L = 2z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq 17.89227 \quad \Rightarrow \quad 2 \times 2.58 \times \frac{\sqrt{1000}}{\sqrt{n}} \leq 17.89227$$

Despejando n :

$$\left(2 \times 2.58 \times \frac{\sqrt{1000}}{17.89227} \right)^2 \leq n \quad \Rightarrow \quad n \geq 83.170$$

O sea, hay que tomar por lo menos 84 especímenes para que el intervalo tenga la longitud pedida.

En general, si queremos hallar n tal que $L = 2z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq l$, donde l es un valor dado, entonces despejando n

$$n \geq \left(\frac{2z_{\frac{\alpha}{2}} \sigma}{l} \right)^2$$

Si estimamos puntualmente al parámetro μ con \bar{X} estamos cometiendo un error en la estimación menor o igual a $\frac{L}{2} = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$, que se conoce como **precisión del estimador**

Ejemplo: Se estima que el tiempo de reacción a un estímulo de cierto dispositivo electrónico está distribuido normalmente con desviación estándar de 0.05 segundos. ¿Cuál es el número de mediciones temporales que deberá hacerse para que la confianza de que el error de la estimación de la esperanza no exceda de 0.01 sea del 95%?

Nos piden calcular n tal que $\frac{L}{2} = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < 0.01$ con $\alpha = 0.05$.

Por lo tanto $n \geq \left(z_{0.025} \frac{0.05}{0.01} \right)^2$.

Además $z_{0.025} = 1.96$. Entonces $n \geq \left(z_{0.975} \frac{0.05}{0.01} \right)^2 = (1.96 \times 5)^2 = 96.04$.

O sea hay que tomar por lo menos 97 mediciones temporales.

Para muestras tomadas de una población normal, o para muestras de tamaño $n \geq 30$, de una población cualquiera, el intervalo de confianza dado anteriormente en (8.1), proporciona buenos resultados.

En el caso de que la población de la que se extrae la muestra no sea normal pero $n \geq 30$, el nivel de confianza del intervalo (8.1) es **aproximadamente** $1 - \alpha$.

Pero para muestras pequeñas tomadas de poblaciones que no son normales no se puede garantizar que el nivel de confianza sea $1 - \alpha$ si se utiliza (8.1).

Ejemplo:

Supongamos que X representa la duración de una pieza de equipo y que se probaron 100 de esas piezas dando una duración promedio de 501.2 horas. Se sabe que la desviación estándar poblacional es $\sigma = 4$ horas. Se desea tener un intervalo del 95% de confianza para la esperanza poblacional $E(X) = \mu$.

Solución:

En este caso, si bien no conocemos cuál es la distribución de X tenemos que el tamaño de la muestra es $n = 100 > 30$ (muestra grande) por lo tanto el intervalo buscado es

$$\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$$

Puesto que $1 - \alpha = 0.95 \rightarrow \alpha = 1 - 0.95 = 0.05 \rightarrow \frac{\alpha}{2} = 0.025$

De la tabla de la normal estandarizada obtenemos $z_{0.025} = 1.96$. Entonces reemplazando:

$$\left[\bar{X} - 1.96 \frac{4}{\sqrt{100}}, \bar{X} + 1.96 \frac{4}{\sqrt{100}} \right]$$

Para el valor particular $\bar{x} = 501.2$ tenemos el intervalo

$$\left[\bar{x} - 1.96 \frac{4}{\sqrt{100}}, \bar{x} + 1.96 \frac{4}{\sqrt{n}} \right] = \left[501.2 - 1.96 \frac{4}{10}, 501.2 + 1.96 \frac{4}{10} \right] = \left[500.4, 502.0 \right].$$

Al establecer que $\left[500.4, 502.0 \right]$ es un intervalo al 95% de confianza de μ estamos diciendo que la probabilidad de que el intervalo $\left[500.4, 502.0 \right]$ contenga a μ es 0.95. O, en otras palabras, la probabilidad de que la muestra aleatoria (X_1, X_2, \dots, X_n) tome valores tales que el intervalo aleatorio $\left[\bar{X} - 1.96 \frac{4}{\sqrt{100}}, \bar{X} + 1.96 \frac{4}{\sqrt{100}} \right]$ defina un intervalo numérico que contenga al parámetro fijo desconocido μ es 0.95.

3.3 - Intervalo de confianza para la media de una distribución normal, varianza desconocida

Nuevamente como se trata de encontrar un intervalo de confianza para μ nos basamos en la esperanza muestral $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ que sabemos es un buen estimador de μ . Pero ahora no podemos usar como pivote a

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

porque desconocemos σ y una condición para ser pivote es que, excepto por el parámetro a estimar (en este caso μ), todos los parámetros que aparecen en él deben ser conocidos. Entonces proponemos como pivote una variable aleatoria definida en forma parecida a Z pero reemplazando σ por un estimador adecuado.

Ya vimos que la varianza muestral definida

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2,$$

donde \bar{X} es la esperanza muestral, es un estimador insesgado de la varianza poblacional $V(X)$, es decir, $E(S^2) = V(X) = \sigma^2 \quad \forall n$. Entonces estimamos σ con S y proponemos como pivote a la variable aleatoria

$$T = \frac{\bar{X} - \mu}{S / \sqrt{n}}.$$

Pero para poder usar a T como pivote debemos conocer su distribución.

Se puede probar que la distribución de T es una distribución llamada **Student con parámetro $n-1$** .

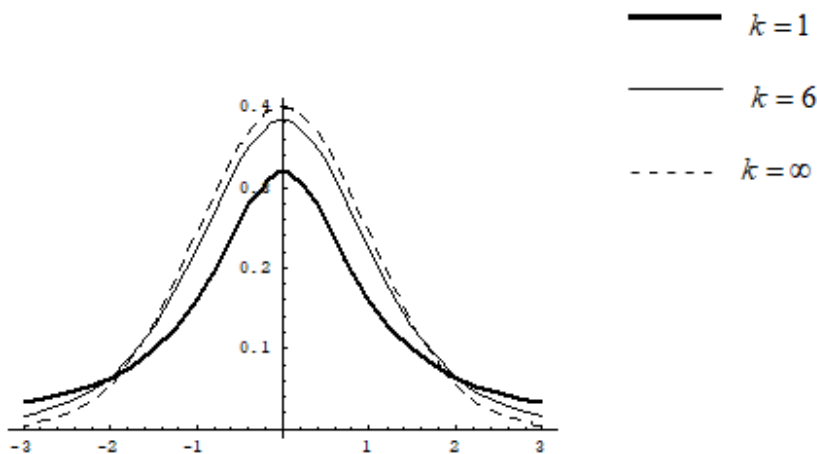
Nota: Una v.a. continua tiene distribución **Student con k grados de libertad**, si su *f.d.p.* es de la forma

$$f(x) = \frac{\Gamma\left[\frac{(k+1)}{2}\right]}{\sqrt{\pi k} \Gamma\left(\frac{k}{2}\right)} \frac{1}{\left[\left(\frac{x^2}{k}\right) + 1\right]^{\frac{k+1}{2}}} \quad -\infty < x < \infty$$

Notación: $T \sim t_k$

La gráfica de la *f.d.p.* de la distribución Student tiene forma de campana como la normal, pero tiende a cero más lentamente. Se puede probar que cuando $k \rightarrow \infty$ la *f.d.p.* de la Student tiende a la *f.d.p.* de la $N(0, 1)$.

En la figura siguiente se grafica $f(x)$ para diferentes valores de k



Anotaremos $t_{\alpha,k}$ al cuantil de la Student con k grados de libertad que deja bajo la *f.d.p.* a derecha un área de α , y a su izquierda un área de $1 - \alpha$.

Luego, para construir el intervalo de confianza buscado a partir del pivote T procedemos como en los casos anteriores:

Comenzamos por plantear la ecuación

$$P(-t \leq T \leq t) = 1 - \alpha,$$

donde la incógnita es el número real t .

Si reemplazamos la v.a. T por su expresión, tenemos sucesivamente (multiplicando por S/\sqrt{n} y restando \bar{X}):

$$P\left(-t \leq \frac{\bar{X} - \mu}{S/\sqrt{n}} \leq t\right) = P\left(-t \frac{S}{\sqrt{n}} \leq \bar{X} - \mu \leq t \frac{S}{\sqrt{n}}\right) = P\left(-\bar{X} - t \frac{S}{\sqrt{n}} \leq -\mu \leq -\bar{X} + t \frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

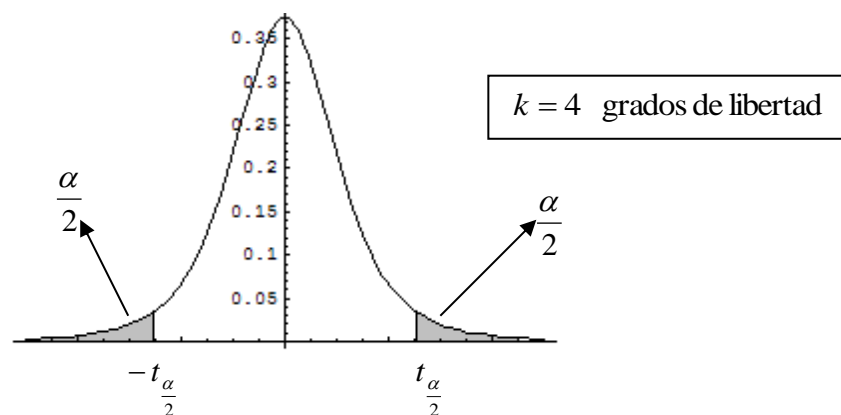
Multiplicando todos los miembros de la desigualdad por -1 (el orden de los miembros se invierte) llegamos a:

$$P\left(\bar{X} - t \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t \frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

Evidentemente, si definimos

$$\begin{cases} \hat{\Theta}_1 = \bar{X} - t \frac{S}{\sqrt{n}} \\ \hat{\Theta}_2 = \bar{X} + t \frac{S}{\sqrt{n}} \end{cases}, \text{ hemos construido dos estadísticos } \hat{\Theta}_1 \text{ y } \hat{\Theta}_2 \text{ tales que } P(\hat{\Theta}_1 \leq \mu \leq \hat{\Theta}_2) = 1 - \alpha,$$

veamos quien es el número t que verifica la ecuación, es decir (ver figura):



$$P(-t \leq T \leq t) = F(t) - F(-t) = 1 - \alpha \quad \text{donde } F(t) \text{ es la Fda para la v.a. } T \sim t_{n-1}.$$

Por la simetría de la distribución t de Student se deduce fácilmente de la figura anterior que $F(-t) = 1 - F(t)$, entonces:

$$F(t) - F(-t) = 2F(t) - 1 = 1 - \alpha, \quad \text{o bien (ver figura anterior),}$$

$$F(t) = 1 - \frac{\alpha}{2}.$$

Al valor de t que verifica esta ecuación se lo suele indicar $t_{\frac{\alpha}{2}, n-1}$. En consecuencia, el intervalo de confianza bilateral al nivel de significación $1-\alpha$ queda:

$$\left[\bar{X} - t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}} \right] \quad \text{con} \quad F\left(t_{\frac{\alpha}{2}, n-1}\right) = 1 - \frac{\alpha}{2}.$$

En consecuencia:

Si (X_1, X_2, \dots, X_n) una muestra aleatoria de tamaño n de una v.a. X donde $X \sim N(\mu, \sigma^2)$, σ^2 desconocido, un intervalo de confianza para μ de nivel $1-\alpha$ es

$$\left[\bar{X} - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right] \quad (8.2)$$

Ejemplo:

Se hicieron 10 mediciones sobre la resistencia de cierto tipo de alambre que dieron valores x_1, x_2, \dots, x_{10} tales que $\bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i = 10.48$ ohms y $S = \sqrt{\frac{1}{9} \sum_{i=1}^{10} (x_i - \bar{x})^2} = 1.36$ ohms. Supóngase que $X \sim N(\mu, \sigma^2)$.

Se desea obtener un intervalo de confianza para la esperanza poblacional μ al 90 %.

Tenemos que $1-\alpha = 0.90 \rightarrow \alpha = 0.1 \rightarrow \alpha/2 = 0.05$

De la Tabla de la t de Student tenemos que $t_{0.05, 9} = 1.8331$. Entonces el intervalo de confianza buscado es:

$$\left[\bar{X} - t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2}, n-1} \frac{S}{\sqrt{n}} \right] = \left[10.48 - 1.8331 \frac{1.36}{\sqrt{10}}, 10.48 + 1.8331 \frac{1.36}{\sqrt{10}} \right]$$

Esto es: $[9.69, 11.27]$.

Si la muestra aleatoria se toma de una distribución normal, σ^2 es desconocido y el tamaño de la muestra grande, entonces se puede probar que al reemplazar σ por S , el estadístico

$$Z = \frac{\bar{X} - \mu}{S / \sqrt{n}} \sim N(0,1) \quad \text{aproximadamente}$$

y puedo construir el intervalo para μ como antes:

$$\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right], \text{ pero su nivel es aproximadamente } 1-\alpha$$

3.4 – Intervalo de confianza para la varianza de una distribución normal

Supongamos que se quiere hallar un intervalo de confianza para la varianza σ^2 de una distribución normal.

Sea (X_1, X_2, \dots, X_n) una muestra aleatoria de una v.a. X , donde $X \sim N(\mu, \sigma^2)$.

Tomamos como estimador puntual de σ^2 a $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

Luego a partir de este estimador puntual construimos el estadístico $X = \frac{(n-1)S^2}{\sigma^2}$

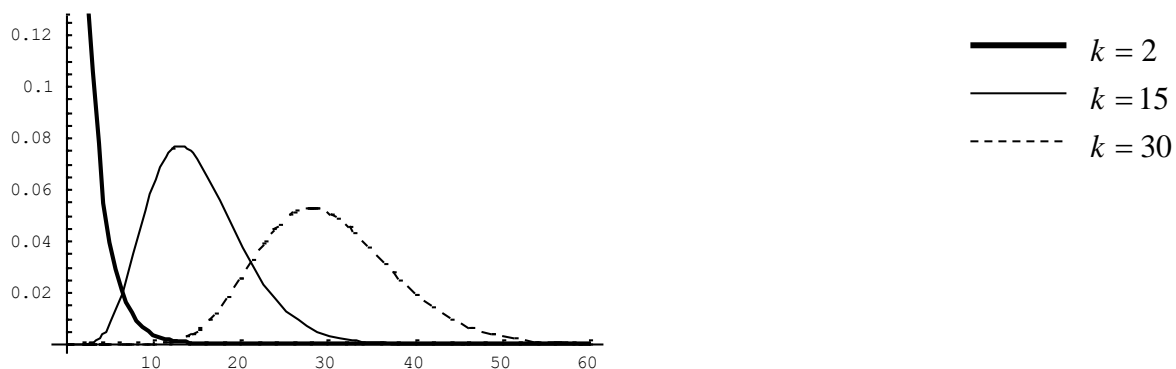
Este estadístico contiene al parámetro desconocido a estimar σ^2 y tiene una distribución conocida, se puede probar que X tiene una distribución llamada **ji-cuadrado con $n-1$ grados de libertad**

Observación: Si X es una v.a. continua se dice que tiene distribución **ji-cuadrado con k grados de libertad** si su f.d.p. es

$$f(x) = \frac{1}{2^{k/2} \Gamma\left(\frac{k}{2}\right)} x^{(k/2)-1} e^{-x/2} \quad x > 0$$

Notación: $X \sim \chi_k^2$

La distribución ji-cuadrado es asimétrica. En la figura siguiente se grafica la densidad para diferentes valores de k



Anotaremos $\chi_{\alpha, k}^2$ al cuantil de la ji-cuadrado con k grados de libertad que deja bajo la fdp a derecha un área de α , y a su izquierda un área de $1 - \alpha$.

Propiedades:

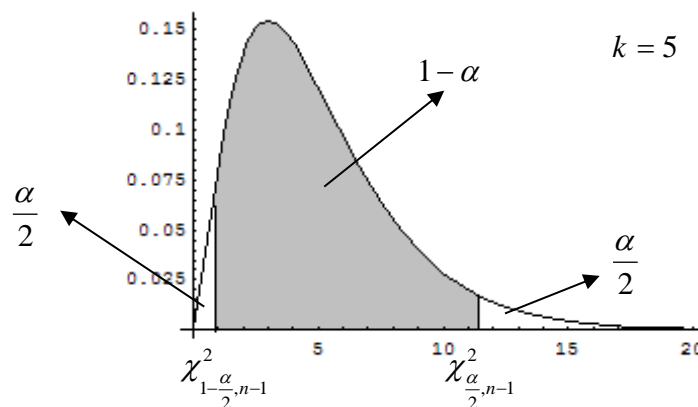
- 1- Se puede probar que si X_1, X_2, \dots, X_n son variables aleatorias independientes con distribución $N(0,1)$ entonces $Z = X_1^2 + X_2^2 + \dots + X_n^2$ tiene distribución ji-cuadrado con n grados de libertad.
- 2- Si X_1, X_2, \dots, X_n son variables aleatorias independientes tal que X_i tiene distribución ji-cuadrado con k_i grados de libertad, entonces $Z = X_1 + X_2 + \dots + X_n$ tiene distribución ji-cuadrado con k grados de libertad donde $k = k_1 + k_2 + \dots + k_n$

3- Si $X \sim \chi_k^2$ entonces **para k grande** $\sqrt{2X} \sim N(\sqrt{2k-1}, 1)$ aproximadamente.

Para desarrollar el intervalo de confianza planteamos hallar dos números a y b tales que

$$P(a \leq X \leq b) = 1 - \alpha \quad \text{es decir} \quad P\left(a \leq \frac{(n-1)S^2}{\sigma^2} \leq b\right) = 1 - \alpha$$

Se puede probar que la mejor elección de a y b es: $a = \chi_{1-\frac{\alpha}{2}, n-1}^2$ y $b = \chi_{\frac{\alpha}{2}, n-1}^2$



Por lo tanto

$$P\left(\chi_{1-\frac{\alpha}{2}, n-1}^2 \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi_{\frac{\alpha}{2}, n-1}^2\right) = 1 - \alpha$$

y despejando σ^2 se llega a

$$P\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}, n-1}^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2}\right) = 1 - \alpha$$

Entonces

Si (X_1, X_2, \dots, X_n) es una muestra aleatoria de una v.a. X , donde $X \sim N(\mu, \sigma^2)$, un intervalo de confianza para σ^2 de nivel $1 - \alpha$ es

$$\left(\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}, n-1}^2}; \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2}\right) \quad (8.8)$$

Observación: un intervalo de confianza para σ de nivel $1 - \alpha$, es $\left(\sqrt{\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}, n-1}^2}}; \sqrt{\frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2}}\right)$

Ejemplo:

Un fabricante de detergente líquido está interesado en la uniformidad de la máquina utilizada para llenar las botellas. De manera específica, es deseable que la desviación estándar σ del proceso de llenado sea menor que 0.15 onzas de líquido; de otro modo, existe un porcentaje mayor del deseable de botellas con un contenido menor de detergente. Supongamos que la distribución del volumen de llenado es aproximadamente normal. Al tomar una muestra aleatoria de 20 botellas, se obtiene una varianza muestral $S^2 = 0.0153$. Hallar un intervalo de confianza de nivel 0.95 para la verdadera varianza del volumen de llenado.

Solución:

La v.a. de interés es X : “volumen de llenado de una botella”

Se asume que $X \sim N(\mu, \sigma^2)$ con σ desconocido.

Estamos en las condiciones para aplicar (8.8)

Tenemos que $1 - \alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow \chi^2_{\frac{\alpha}{2}, n-1} = \chi^2_{0.025, 19} = 32.85$ y

$$\chi^2_{\frac{\alpha}{2}, n-1} = \chi^2_{0.025, 19} = 32.85$$

Además $S^2 = 0.0153$

Por lo tanto el intervalo es

$$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}, n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}} \right) = \left(\frac{(20-1) \times 0.0153}{32.85}, \frac{(20-1) \times 0.0153}{8.91} \right) = (0.00884; 0.0326)$$

Y un intervalo para σ es $(\sqrt{0.00884}; \sqrt{0.0326}) = (0.09; 0.1805)$

Por lo tanto con un nivel de 0.95 los datos **no apoyan la afirmación que $\sigma < 0.15$**

3.5 – Intervalo de confianza para una proporción

Sea una población de tamaño N (eventualmente puede ser infinito) de cuyos individuos nos interesa cierta propiedad A . Supongamos que la probabilidad de que un individuo de la población verifique A es $p = P(A)$. El significado del parámetro p es, en consecuencia, el de proporción de individuos de la población que verifican la propiedad A . Podemos definir una variable aleatoria X_i que mide a los individuos de la población la ocurrencia o no de la propiedad A .

La variable aleatoria tendrá la distribución:

$$p(x) = \begin{cases} p(1) = P(X_i = 1) = p \\ p(0) = P(X_i = 0) = 1 - p, \end{cases}$$

es decir, X_i es una v.a. que toma sólo dos valores: 1 (si el individuo verifica A) con probabilidad p y 0 (cuando no verifica A) con probabilidad $1-p$. Esto es equivalente a decir que X_i tiene una distribución binomial con parámetros 1 y p : $X_i \sim B(1, p)$.

Supongamos que consideramos una muestra aleatoria (X_1, X_2, \dots, X_n) de tamaño n . Si formamos el estadístico $X = X_1 + X_2 + \dots + X_n$, es evidente que esta v.a. mide el número de individuos de la muestra de tamaño n que verifican la propiedad A. Por lo tanto por su significado X es una v.a. cuya distribución es binomial con parámetros n y p : $X \sim B(n, p)$. De acuerdo con esto, la variable aleatoria \hat{P} definida: $\hat{P} = \frac{X}{n}$ representa la proporción de individuos de la muestra que verifican la propiedad A.

Observemos que siendo $X_i \sim B(1, p)$ es $E(X_i) = p$. Y, dado que $X \sim B(n, p)$, también es $E(\hat{P}) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = \frac{1}{n} np = p$, es decir \hat{P} es un estimador insesgado de p . Esto es de esperar pues $\hat{P} = \frac{X}{n} = \frac{1}{n} \sum_{i=1}^n X_i$.

Pero además, es fácil ver que \hat{P} es estimador consistente de p . En efecto, tenemos que $E(\hat{P}) = p$, pero también es

$$V(\hat{P}) = V\left(\frac{X}{n}\right) = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}.$$

Deseamos construir un intervalo de confianza de p . Es razonable basarnos en el estimador insesgado \hat{P} . Consideramos como pivote a la variable aleatoria

$$Z = \frac{\hat{P} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

cuya distribución es, para n suficientemente grande, aproximadamente $N(0,1)$.

En efecto:

$$\text{Siendo } \hat{P} = \frac{X_1}{n} + \frac{X_2}{n} + \dots + \frac{X_n}{n}, \quad \text{es } E(\hat{P}) = \sum_{i=1}^n E\left(\frac{X_i}{n}\right) = p \quad \text{y} \quad V(\hat{P}) = \sum_{i=1}^n V\left(\frac{X_i}{n}\right) = \frac{p(1-p)}{n}$$

Por lo tanto:

$$Z = \frac{\hat{P} - p}{\sqrt{\frac{p(1-p)}{n}}} \underset{n \text{ grande}}{\sim} N(0,1),$$

El pivote puede ponerse en una forma más conveniente si tenemos en cuenta que, según vimos recién, \hat{P} es estimador consistente de p y en consecuencia, en el denominador reemplazamos el parámetro desconocido p por su estimador \hat{P} , y se puede probar que :

$$Z = \frac{\hat{P} - p}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}} \approx N(0,1). \quad \text{aproximadamente si } n \text{ es grande}$$

Partiendo de este pivote podemos seguir los mismos pasos de los casos anteriores para llegar al siguiente intervalo de confianza al nivel $1 - \alpha$ de p :

$$\left[\hat{P} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}(1-\hat{P})}{n}}, \hat{P} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \right] \quad \text{con} \quad \Phi\left(z_{\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2}.$$

Entonces

Si \hat{P} es la proporción de observaciones de una muestra aleatoria de tamaño n que verifican una propiedad de interés, entonces un intervalo de confianza para la proporción p de la población que cumple dicha propiedad de nivel aproximadamente $1 - \alpha$ es

$$\left[\hat{P} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}(1-\hat{P})}{n}}, \hat{P} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \right] \quad (8.10)$$

Observaciones:

1- Este procedimiento depende de la aproximación normal a la distribución binomial. Por lo tanto el intervalo (8.10) se puede utilizar si $n\hat{P} > 10$ y $n(1-\hat{P}) > 10$, es decir, **la muestra debe contener un mínimo de diez éxitos y diez fracasos.**

2- La longitud del intervalo es $L = 2z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$, pero esta expresión está en función de \hat{P}

Si nos interesa hallar un valor de n de manera tal que la longitud L sea menor que un valor determinado, podemos hacer dos cosas:

a) tomar una muestra preliminar, con ella estimar p con \hat{P} y de la expresión anterior despejar n , lo que lleva a

$$L = 2z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \leq l \quad \Rightarrow \quad n \geq \left(\frac{2z_{\frac{\alpha}{2}}}{l} \right)^2 \hat{P}(1-\hat{P})$$

b) si no tomamos una muestra preliminar, entonces acotamos $\hat{P}(1-\hat{P}) \leq 0.5 \times (1-0.5)$, entonces

$$L = 2z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \leq 2z_{\frac{\alpha}{2}} \sqrt{\frac{0.5(1-0.5)}{n}} \leq l \quad \Rightarrow \quad n \geq \left(\frac{z_{\frac{\alpha}{2}}}{l} \right)^2$$

Ejemplo:

Un fabricante de componentes compra un lote de dispositivos de segunda mano y desea saber la proporción de la población que están fallados. Con ese fin experimenta con 140 dispositivos elegidos al azar y encuentra que 35 de ellos están fallados.

- a) Calcular un intervalo de confianza del 99% para la proporción poblacional p .
 b) ¿De qué tamaño deberá extraerse la muestra a fin de que la proporción muestral no difiera de la proporción poblacional en más de 0.03 con un 95% de confianza?

Solución:

a) El tamaño de la muestra es $n = 140$ (muestra grande)

La proporción muestral es $\hat{p} = \frac{35}{140} = 0.25$

El nivel de confianza es $1 - \alpha = 0.99 \rightarrow \alpha = 0.01 \rightarrow \frac{\alpha}{2} = 0.005$.

De la tabla de la normal estandarizada vemos que $z_{0.005} = 2.58$. Entonces el intervalo buscado es:

$$\left[0.25 - 2.58 \sqrt{\frac{0.25(1-0.25)}{140}}, 0.25 + 2.58 \sqrt{\frac{0.25(1-0.25)}{140}} \right] = [0.15558, 0.34441]$$

b) Buscamos el tamaño n de la muestra tal que con un 95% de confianza la proporción muestral \hat{p} esté a una distancia 0.03 de la proporción poblacional p , es decir buscamos n tal que

$\frac{L}{2} \leq 0.03$, por lo tanto como $\alpha = 0.05 \rightarrow \frac{\alpha}{2} = 0.025$ **si tomamos la muestra anterior como preliminar :**

$$n \geq \left(\frac{2z_{\frac{\alpha}{2}}}{l} \right)^2 \hat{p}(1 - \hat{p}) = \left(\frac{2 \times 1.96}{2 \times 0.03} \right)^2 0.25(1 - 0.25) = 800.3333$$

Por lo tanto hay que tomar una muestra de tamaño por lo menos 801. como ya se tomó una muestra de tamaño 140, hay que tomar otra adicional de tamaño $801 - 140 = 661$

Supongamos que no tomamos una muestra inicial, entonces directamente planteamos

$$n \geq \left(\frac{z_{\frac{\alpha}{2}}}{l} \right)^2 = \left(\frac{1.96}{2 \times 0.03} \right)^2 = 1067.1111$$

Entonces hay que tomar una muestra de tamaño 1068 por lo menos.

Practica**Intervalos de confianza**

- 1) Se sabe que la duración en horas, de un foco de 75 watts tiene una distribución aproximadamente normal, con una desviación estándar de 25 horas. Se toma una muestra aleatoria de 20 focos, la cual resulta tener una duración promedio de $\bar{x} = 1014$ horas.
 - a) Construya un intervalo de confianza del 95% para la duración media.
 - b) Supóngase que se desea una confianza del 95% en que el error en la estimación de la duración media sea menor que 5 horas. ¿Qué tamaño de muestra debe utilizarse?
 - c) Supóngase que se desea que el ancho total del intervalo de confianza sea de 6 horas, con una confianza del 95%. ¿Qué tamaño de muestra debe utilizarse para este fin?

- 2) Un ingeniero civil analiza la resistencia a la compresión del concreto. La resistencia está distribuida aproximadamente de manera normal, con varianza 1000 (psi)^2 . Al tomar una muestra aleatoria de 12 especímenes, se tiene que $\bar{x} = 3250$ psi.
 - a) Construya un intervalo de confianza del 95% para la resistencia a la compresión media.
 - b) Construya un intervalo de confianza del 99% para la resistencia a la compresión media. Compare el ancho de este intervalo de confianza con el ancho encontrado en el inciso a).

- 3) a) Con base en pruebas de comportamiento de una gran muestra de uniones soldadas, se calculó un intervalo de confianza de 90% para la media de la dureza Rockwell B de cierto tipo de soldadura de (83.2 , 84.1). Determine un intervalo de confianza de 95% para la media de la dureza Rockwell B de este tipo de soldadura.
- b) Un intervalo de confianza de 95% para una media poblacional se calcula de una muestra de tamaño 50. Se calculará otro intervalo de confianza de 95% para una muestra de tamaño 200 extraída de la misma población. Justificando su elección, elija la mejor respuesta que complete el espacio en blanco:
El intervalo de una muestra de tamaño 50 será aproximadamentedel intervalo de la muestra de tamaño 200.
 - b1) un octavo de ancho
 - b2) un cuarto de ancho
 - b3) la mitad de ancho
 - b4) del mismo ancho
 - b5) dos veces el ancho
 - b6) cuatro veces el ancho
 - b7) ocho veces el ancho

- 4) Las siguientes mediciones se registraron para el tiempo de secado, en horas, de cierta marca de pintura látex:
 3.4 2.5 4.8 2.9 3.6 2.8 3.3 5.6 3.7 2.8 4.4 4.0 5.2 3.0 4.8
 - a) Suponiendo que las mediciones representan una muestra aleatoria de una población normal, encuentre un intervalo de confianza de nivel 99% para la media de los tiempos de secado. ¿Está de acuerdo con la afirmación de que el tiempo medio de secado podría ser 3.50 horas ?. ¿Por qué?.
 - b) Encuentre un intervalo de confianza del 99% para la varianza del tiempo de secado

- 5) Un proceso novedoso para elaborar gasolina ecológica toma biomasa en la forma de sacarosa y la convierte en gasolina usando reacciones catalíticas. En un paso en un proceso de

la planta piloto, un ingeniero químico mide la salida de cadenas de carbono de longitud tres. Nueve corridas con el mismo catalizador dieron los rendimientos (en galones)

0.63 2.64 1.85 1.68 1.09 1.67 0.73 1.04 0.68

Suponga que el rendimiento tiene una distribución normal.

- a) Obtenga un intervalo de confianza del 95% para el verdadero rendimiento medio del proceso de la planta piloto.
 - b) Encuentre un intervalo de confianza del 90% para la varianza del rendimiento del proceso de la planta piloto.
- 6) Un fabricante de calculadoras electrónicas está interesado en estimar la fracción de unidades defectuosas producidas. Se toma una muestra aleatoria de 800 calculadoras, de las cuales 10 resultan defectuosas. Calcule un intervalo de confianza del 99% para la fracción de calculadoras defectuosas.
- 7) En una muestra aleatoria de 85 soportes para el cigüeñal de un motor de automóvil, 10 tienen un terminado que es más rugoso de lo que las especificaciones permiten.
- a) Calcular un intervalo de confianza del 95% para la verdadera proporción de soportes en la población que exceden las especificaciones.
 - b) ¿Cuán grande debe ser la muestra si se desea tener una confianza del 95% de que el error de estimación sea menor que 0.05?.
- 8) En una muestra aleatoria de 100 baterías producidas por cierto método, el promedio del tiempo de vida fue de 150 horas y la desviación estándar de 25 horas.
- a) Determine un intervalo de confianza de 95% para la media del tiempo de vida de las baterías producidas por este método.
 - b) Un ingeniero afirma que la media del tiempo de vida está entre 147 y 153 horas. ¿Con qué nivel de confianza se puede hacer esta afirmación?.

4- Test o prueba de hipótesis

4.1 – Introducción

Hasta ahora hemos estudiado el problema de estimar un parámetro desconocido a partir de una muestra aleatoria.

En muchos problemas se requiere tomar una decisión entre aceptar o rechazar una proposición sobre algún parámetro. Esta proposición recibe el nombre de **hipótesis estadística**, y el procedimiento de toma de decisión sobre la hipótesis se conoce como **prueba o test de hipótesis**.

Como se emplean distribuciones de probabilidad para representar poblaciones, también podemos decir que una hipótesis estadística es una proposición sobre la distribución de probabilidad de una variable aleatoria, donde la hipótesis involucra a uno más parámetros de esta distribución.

Por ejemplo, supongamos que cierto tipo de motor de automóvil emite una media de 100 mg de óxidos de nitrógeno (NO_x) por segundo con 100 caballos de fuerza. Se ha propuesto una modificación al diseño del motor para reducir las emisiones de NO_x . El nuevo diseño se producirá si se demuestra que la media de su tasa de emisiones es menor de 100 mg/s. Se construye y se prueba una muestra de 50 motores modificados. La media muestral de emisiones de NO_x es de 92 mg/s, y supongamos que se puede asumir que σ es 21 mg/s.

La variable aleatoria de interés en este caso es X : “tasa de emisión de un motor modificado tomado al azar”.

La preocupación de los fabricantes consiste en que los motores modificados no puedan reducir todas las emisiones; es decir que la media poblacional pudiera ser 100 o mayor que 100.

Entonces, la pregunta es: ¿es factible que esta muestra pueda provenir de una v.a. con media 100 o mayor?

Éste es el tipo de preguntas que las pruebas de hipótesis están diseñadas para responder. Veremos cómo construir una prueba de hipótesis, pero podemos decir que en general se basa en construir a partir de la muestra aleatoria un **estadístico**, y según el valor que tome este **estadístico de prueba** se aceptará o se rechazará la hipótesis.

Se ha observado una muestra con media $\bar{X} = 92$.

Hay dos interpretaciones posibles de esta observación:

- 1- La media poblacional es realmente mayor o igual que 100, y la media muestral es menor que 100 debido a la **variabilidad** propia de la variable aleatoria \bar{X}
 - 2- La media poblacional es en realidad menor que 100, y la media muestral refleja este hecho.
- Estas dos explicaciones tienen nombres: la primera se llama **hipótesis nula**; la segunda es la **hipótesis alternativa**.

En la mayoría de las situaciones la hipótesis nula dice que el efecto que indica la muestra es atribuible solamente a la variación aleatoria del estadístico de prueba.

La hipótesis alternativa establece que el efecto que indica la muestra es verdadero.

Para hacer las cosas más precisas, todo se expresa mediante símbolos. La hipótesis nula se denota por H_0 , la hipótesis alternativa se denota con H_1 . Como es usual la media poblacional se anota μ . Por lo tanto se tiene

$$H_0 : \mu \geq 100 \quad \text{contra} \quad H_1 : \mu < 100 \quad (\text{hipótesis alternativa unilateral})$$

Esencialmente, para realizar una prueba de hipótesis se pone la hipótesis nula en juicio. Se asume que H_0 es verdadera, de la misma manera como se empieza en un juicio bajo el supuesto de que un acusado es inocente. La muestra aleatoria proporciona la evidencia.

Las hipótesis son siempre proposiciones **sobre los parámetros de la población o distribución bajo estudio**, no proposiciones sobre la muestra.

Otros tipos de hipótesis que podrían formularse son

$$H_0 : \mu \leq 100 \quad \text{contra} \quad H_1 : \mu > 100 \quad (\text{hipótesis alternativa unilateral})$$

o

$$H_0 : \mu = 100 \quad \text{contra} \quad H_1 : \mu \neq 100 \quad (\text{hipótesis alternativa bilateral})$$

En el ejemplo tenemos X_1, X_2, \dots, X_{50} muestra aleatoria de la v.a. X definida anteriormente.

Como estamos haciendo una hipótesis sobre la media poblacional es razonable tomar como estadístico de prueba a \bar{X} . El valor observado de la media muestral es $\bar{X} = 92$.

Si el valor de \bar{X} es muy “menor” que 100 entonces se considera que hay evidencia en contra H_0 y se la rechaza, aceptando la hipótesis alternativa.

Si el valor de \bar{X} no es “muy menor” que 100 entonces se considera que no hay evidencia en contra H_0 y se rechaza la hipótesis alternativa.

Ya veremos como construir una **regla de decisión**, supongamos ahora que tenemos la siguiente regla:

$$\begin{cases} \text{se rechaza } H_0 & \text{si } \bar{X} < 95 \\ \text{se acepta } H_0 & \text{si } \bar{X} \geq 95 \end{cases}$$

El intervalo $(95, \infty)$ es la **zona de aceptación**.

La región $(-\infty, 95)$ es la **zona de rechazo o región crítica**.

Mientras que 95 es el **punto crítico**.

Como estamos tomando una decisión basados en el valor de un estadístico podemos cometer dos tipos de errores: rechazar H_0 cuando ésta es verdadera, es decir el estadístico toma valores en la zona de rechazo cuando H_0 es verdadera; o aceptar H_0 cuando ésta es falsa, es decir que el estadístico tome valores en la zona de aceptación cuando H_0 es falsa.

El primero se conoce como **error de tipo I**, y el segundo como **error de tipo II**.

Debido a que la decisión se basa en variables aleatorias es posible asociar probabilidades a los errores de tipo I y II, específicamente anotamos

$$\alpha = P(\text{error de tipo I})$$

$$\beta = P(\text{error de tipo II})$$

A $\alpha = P(\text{error de tipo I})$ se lo conoce como **nivel de significancia del test**.

Para calcular estas probabilidades debemos conocer la distribución del estadístico de prueba en el caso de ser H_0 verdadera, es decir debemos conocer la distribución del estadístico de prueba “bajo H_0 ”.

En el ejemplo anterior la muestra es grande, ya sabemos que por T.C.L. el estadístico

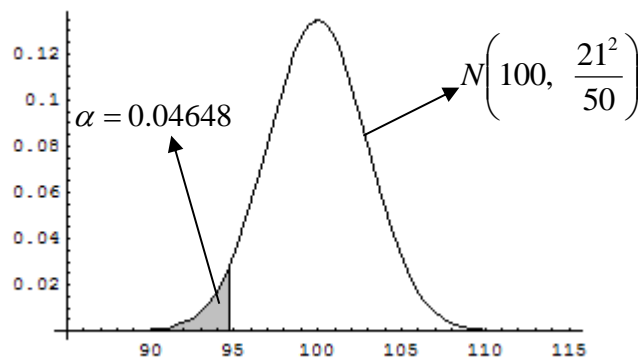
$$Z = \frac{\bar{X} - 100}{\sigma / \sqrt{n}} \approx N(0,1) \quad \text{si } H_0 \text{ es verdadera, o sea } Z = \frac{\bar{X} - 100}{21 / \sqrt{50}} \approx N(0,1)$$

Entonces para calcular α planteamos:

$$\begin{aligned} \alpha &= P(\text{error de tipo I}) = P\left(\text{rechazar } H_0 / H_0 \text{ es V}\right) = P(\bar{X} < 95 / \mu = 100) = \\ &= P\left(\frac{\bar{X} - 100}{21 / \sqrt{50}} < \frac{95 - 100}{21 / \sqrt{50}}\right) \approx \Phi\left(\frac{95 - 100}{21 / \sqrt{50}}\right) = \Phi(-1.6835) = 1 - 0.95352 = 0.04648 \end{aligned}$$

Esto significa que el 4.64% de las muestras aleatorias conducirán al rechazo de la hipótesis $H_0 : \mu \geq 100$ cuando el verdadero μ sea mayor o igual que 100.

En este caso el gráfico de la zona de rechazo es



Del gráfico anterior vemos que podemos reducir α al aumentar la zona de aceptación. Por ejemplo supongamos que ahora la regla de decisión es

$$\begin{cases} \text{se rechaza } H_0 & \text{si } \bar{X} < 93 \\ \text{se acepta } H_0 & \text{si } \bar{X} \geq 93 \end{cases}$$

$$\text{Entonces } \alpha = P(\text{error de tipo I}) = P\left(\text{rechazar } H_0 / H_0 \text{ es V}\right) = P(\bar{X} < 93 / \mu = 100) =$$

$$= P\left(\frac{\bar{X}-100}{21/\sqrt{50}} < \frac{93-100}{21/\sqrt{50}}\right) \approx \Phi\left(\frac{93-100}{21/\sqrt{50}}\right) = \Phi(-2.357) = 1 - 0.99061 = 0.00939$$

También se puede reducir α aumentando el tamaño de la muestra. Supongamos que $n = 85$, entonces

$$\begin{aligned}\alpha &= P(\text{error de tipo I}) = P(\text{rechazar } H_0 / H_0 \text{ es V}) = P(\bar{X} < 95 / \mu = 100) = \\ &= P\left(\frac{\bar{X}-100}{21/\sqrt{85}} < \frac{95-100}{21/\sqrt{85}}\right) \approx \Phi\left(\frac{95-100}{21/\sqrt{85}}\right) = \Phi(-2.195) = 1 - 0.98574 = 0.01426\end{aligned}$$

También es importante examinar la probabilidad de cometer error de tipo II, esto es

$$\beta = P(\text{error de tipo II}) = P(\text{aceptar } H_0 / H_0 \text{ es falsa})$$

Pero en este caso para llegar a un valor numérico necesitamos tener una alternativa específica pues en nuestro ejemplo:

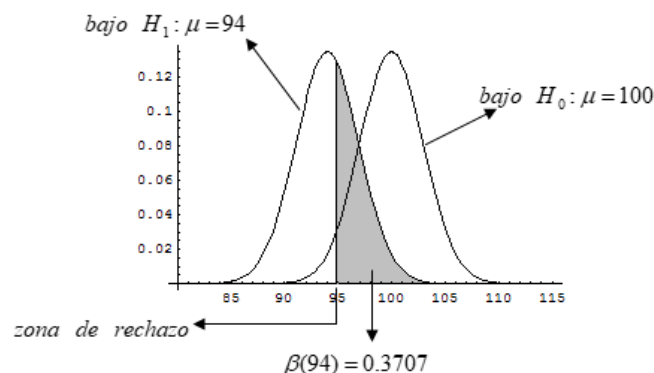
$$\begin{aligned}\beta &= P(\text{error de tipo II}) = P(\text{aceptar } H_0 / H_0 \text{ es falsa}) = P(\bar{X} \geq 95 / \mu \neq 100) = \\ &= P\left(\frac{\bar{X}-\mu}{21/\sqrt{50}} \geq \frac{95-\mu}{21/\sqrt{50}}\right) = 1 - \Phi\left(\frac{95-\mu}{21/\sqrt{50}}\right) = \beta(\mu)\end{aligned}$$

Donde anotamos con μ a la verdadera media poblacional **desconocida**.

Podemos entonces calcular β para un valor particular de μ , por ejemplo nos puede interesar como se comporta el test cuando la verdadera media es $\mu = 94$, entonces

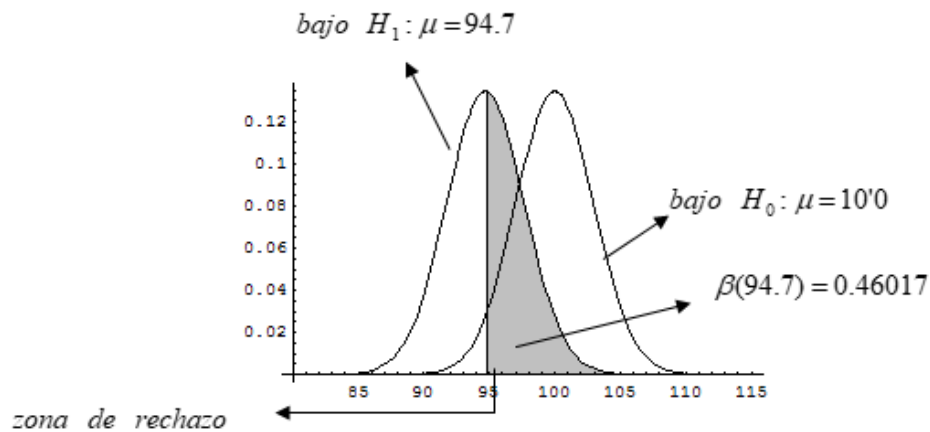
$$\beta(94) = P\left(\frac{\bar{X}-94}{21/\sqrt{50}} \geq \frac{95-94}{21/\sqrt{50}}\right) = 1 - \Phi\left(\frac{95-94}{21/\sqrt{50}}\right) = 1 - \Phi(0.3367) = 1 - 0.62930 = 0.3707$$

Gráficamente:



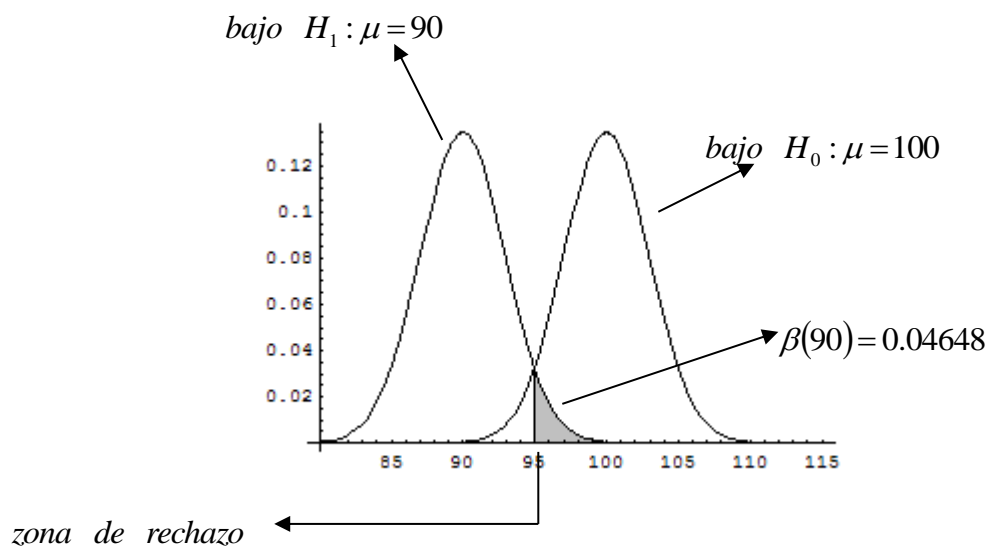
La probabilidad β de cometer error de tipo II **crece** a medida que el valor verdadero de μ se acerca al valor hipotético. Por ejemplo si el verdadero valor de μ fuera 94.7 entonces

$$\beta(94.7) = P\left(\frac{\bar{X} - 94.7}{21/\sqrt{50}} \geq \frac{95 - 94.7}{21/\sqrt{50}}\right) = 1 - \Phi\left(\frac{95 - 94.7}{21/\sqrt{50}}\right) = 1 - \Phi(0.101015) = 1 - 0.53983 = 0.46017$$



Además, la probabilidad β de cometer error de tipo II **disminuye** a medida que el valor verdadero de μ se aleja del valor hipotético. Por ejemplo si el verdadero valor de μ fuera 90 entonces

$$\beta(90) = P\left(\frac{\bar{X} - 90}{21/\sqrt{50}} \geq \frac{95 - 90}{21/\sqrt{50}}\right) = 1 - \Phi\left(\frac{95 - 90}{21/\sqrt{50}}\right) = 1 - \Phi(1.6835) = 1 - 0.95352 = 0.04648$$



También se puede reducir la probabilidad de cometer error de tipo II con el tamaño de la muestra. Por ejemplo si $n = 85$ entonces y $\mu = 94$

$$\beta(94) = P\left(\frac{\bar{X} - 94}{21/\sqrt{85}} \geq \frac{95 - 94}{21/\sqrt{85}}\right) = 1 - \Phi\left(\frac{95 - 94}{21/\sqrt{85}}\right) = 1 - \Phi(0.4390) = 1 - 0.67003 = 0.32997$$

Lo que se ha visto en los ejemplos anteriores se puede generalizar. Podemos recalcar los siguientes puntos importantes:

- 1- El tamaño de la región crítica, y en consecuencia la probabilidad α de cometer error de tipo I, siempre pueden reducirse mediante una selección apropiada de los valores críticos.
- 2- Los errores tipo I y II están relacionados. Una disminución en la probabilidad en un tipo de error siempre da como resultado un aumento en la probabilidad del otro, siempre que el tamaño de la muestra no cambie.
- 3- En general, un aumento en el tamaño de la muestra reduce tanto a α como a β , siempre que los valores críticos se mantengan constantes.
- 4- Cuando la hipótesis nula es falsa, β aumenta a medida que el valor verdadero del parámetro tiende al valor hipotético propuesto por la hipótesis nula. El valor de β disminuye a medida que aumenta la deferencia entre el verdadero valor medio y el propuesto.

En general el investigador controla la probabilidad α del error de tipo I cuando selecciona los valores críticos. Por lo tanto el rechazo de la hipótesis nula de manera errónea se puede fijar de antemano. Eso hace que rechazar la hipótesis nula sea una **conclusión fuerte**.

La probabilidad β de error de tipo II no es constante, sino que depende del valor verdadero del parámetro. También depende β del tamaño de la muestra que se haya seleccionado. Como β está en función del tamaño de la muestra y del valor verdadero del parámetro, la decisión de aceptar la hipótesis nula se la considera una **conclusión débil**, a menos que se sepa que β es aceptablemente pequeño. Por lo tanto **cuando se acepta H_0 en realidad se es incapaz de rechazar H_0 . No se puede rechazar H_0 pues no hay evidencia en contra H_0 .**

Un concepto importante es el siguiente:

La **potencia** de un test es la probabilidad de rechazar la hipótesis nula. La simbolizamos $\pi(\mu)$. Para los **valores de μ tal que la alternativa es verdadera** se tiene

$$\pi(\mu) = P\left(\text{rechazar } H_0 / H_0 \text{ es falsa}\right) = 1 - \beta(\mu)$$

Las pruebas estadísticas se comparan mediante la comparación de sus propiedades de potencia. La potencia es una medida de la **sensibilidad** del test, donde por sensibilidad se entiende la capacidad de una prueba para detectar diferencias.

En el ejemplo anterior, la sensibilidad de la prueba para detectar la diferencia entre una tasa de emisión media de 100 y otra de 94 es $\pi(94) = 1 - \beta(94) = 1 - 0.3707 = 0.6293$. Es decir si el valor verdadero de la tasa de emisión media es 94, la prueba rechazará de manera correcta H_0 y detectará esta diferencia el 62.93% de las veces. Si el investigador piensa que este valor es bajo entonces el investigador puede aumentar α o el tamaño de la muestra.

4.2 – Prueba de hipótesis sobre la media, varianza conocida

Veamos ahora cómo construir una regla de decisión sobre la media de una población.

Supongamos que la variable aleatoria de interés X tiene una media μ y una varianza σ^2 conocida. Asumimos que X tiene distribución normal, es decir $X \sim N(\mu, \sigma^2)$.

Nuevamente, como en el ejemplo introductorio, es razonable tomar como estadístico de prueba al promedio muestral \bar{X} . Bajo las suposiciones hechas tenemos que $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

Supongamos que tenemos las hipótesis

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu \neq \mu_0$$

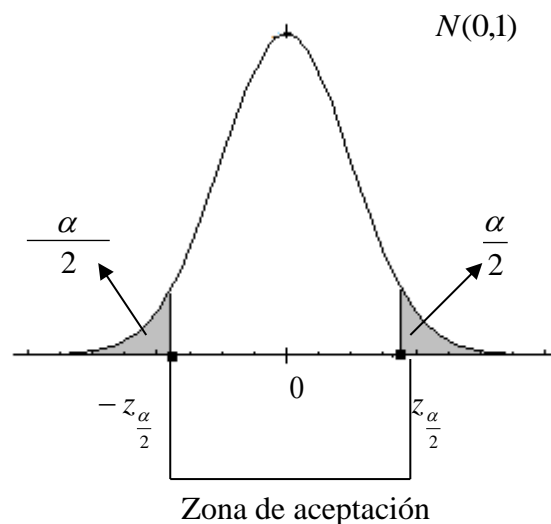
Donde μ_0 es una constante específica. Se toma una muestra aleatoria X_1, X_2, \dots, X_n de la población.

Si $H_0 : \mu = \mu_0$ es verdadera, entonces $\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$, por lo tanto el estadístico

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad \text{tiene distribución } N(0,1) \text{ si } H_0 : \mu = \mu_0 \text{ es verdadera}$$

Tomamos a Z como **estadístico de prueba**

$$\text{Si } H_0 : \mu = \mu_0 \text{ es verdadera entonces } P\left(-z_{\frac{\alpha}{2}} \leq Z \leq z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$



Es evidente que una muestra que produce un valor del estadístico de prueba que cae en las colas de la distribución de Z será inusual si $H_0 : \mu = \mu_0$ es verdadera, por lo tanto esto es un indicador que H_0 es falsa.

Entonces la regla de decisión es:

$$\begin{cases} \text{rechazar } H_0 & \text{si } |Z| > z_{\frac{\alpha}{2}} \\ \text{aceptar } H_0 & \text{si } |Z| \leq z_{\frac{\alpha}{2}} \end{cases}$$

Notar que la probabilidad que la estadística de prueba tome un valor que caiga en la zona de rechazo si H_0 es verdadera es igual a α , es decir la probabilidad de cometer error de tipo I es α pues

$$\begin{aligned} P(\text{error de tipo I}) &= P(\text{rechazar } H_0 / H_0 \text{ es } V) = P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| > z_{\frac{\alpha}{2}} \middle/ \mu = \mu_0\right) = \\ &= P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > z_{\frac{\alpha}{2}}\right) + P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -z_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha \end{aligned}$$

Ejemplo:

El porcentaje deseado de SiO_2 en cierto tipo de cemento aluminoso es 5.5. Para probar si el verdadero promedio de porcentaje es 5.5 para una planta de producción en particular, se analizaron 16 muestras obtenidas de manera independiente. Supongamos que el porcentaje de SiO_2 en una muestra está normalmente distribuido con $\sigma = 3$, y que $\bar{x} = 5.25$.

¿Indica esto de manera concluyente que el verdadero promedio de porcentaje difiere de 5.5? Utilice $\alpha = 0.01$

Solución:

La v.a. de interés es X : “porcentaje de SiO_2 en cierto tipo de cemento aluminoso”

Asumimos que $X \sim N(\mu, 3^2)$

Podemos plantear las hipótesis

$$H_0 : \mu = 5.5 \quad \text{contra} \quad H_1 : \mu \neq 5.5$$

Tenemos una muestra de tamaño $n = 16$ que dio un promedio muestral $\bar{x} = 5.25$

Como $\alpha = 0.01$ entonces $z_{\frac{\alpha}{2}} = z_{0.005} = 2.575$

$$\text{Por lo tanto la regla de decisión es } \begin{cases} \text{rechazar } H_0 & \text{si } \left|\frac{\bar{X} - 5.5}{3/\sqrt{16}}\right| > 2.575 \\ \text{aceptar } H_0 & \text{si } \left|\frac{\bar{X} - 5.5}{3/\sqrt{16}}\right| \leq 2.575 \end{cases}$$

$$\text{El estadístico } \left|\frac{\bar{X} - 5.5}{3/\sqrt{16}}\right| \text{ toma el valor } z_0 = \left|\frac{5.25 - 5.5}{3/\sqrt{16}}\right| = 0.333333$$

Como $z_0 = 0.333333 < 2.575 = z_{\frac{0.01}{2}}$ se acepta H_0

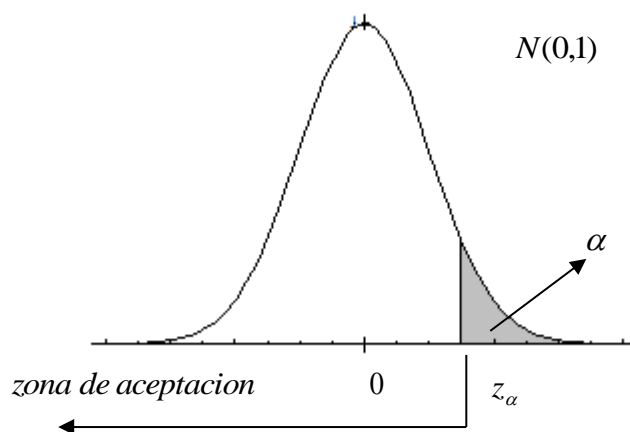
También podemos desarrollar tests o pruebas de hipótesis para el caso de que la hipótesis alternativa es unilateral.

Supongamos las hipótesis

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu > \mu_0$$

En este caso la región crítica debe colocarse en la cola superior de la distribución normal estándar y el rechazo de H_0 se hará cuando el valor calculado de z_0 sea muy grande, esto es la regla de decisión será

$$\begin{cases} \text{rechazar } H_0 & \text{si } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha \\ \text{aceptar } H_0 & \text{si } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_\alpha \end{cases}$$

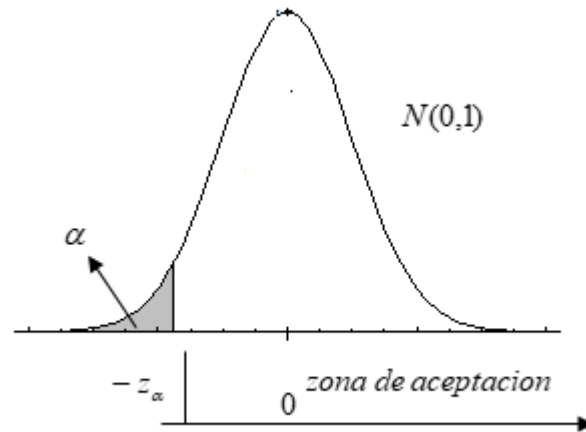


De manera similar para las hipótesis

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu < \mu_0$$

se calcula el valor del estadístico de prueba z_0 y se rechaza H_0 si el valor de z_0 es muy pequeño, es decir la regla de decisión será

$$\begin{cases} \text{rechazar } H_0 & \text{si } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -z_\alpha \\ \text{aceptar } H_0 & \text{si } \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq -z_\alpha \end{cases}$$

Ejemplo:

Se sabe que la duración, en horas, de un foco de 75 watts tiene una distribución aproximadamente normal, con una desviación estándar de $\sigma = 25$ horas. Se toma una muestra aleatoria de 20 focos, la cual resulta tener una duración promedio de $\bar{x} = 1040$ horas. ¿Existe evidencia que apoye la afirmación de que la duración promedio del foco es mayor que 1000 horas?. Utilice $\alpha = 0.05$.

Solución:

La v.a. de interés es X: “duración en horas de un foco tomado al azar”

Asumimos $X \sim N(\mu, 25^2)$

Podemos plantear las hipótesis

$$H_0 : \mu = 1000 \quad \text{contra} \quad H_1 : \mu > 1000$$

Tenemos una muestra de tamaño $n = 20$ que dio un promedio muestral $\bar{x} = 1040$

Como $\alpha = 0.05$ entonces $z_\alpha = z_{0.05} = 1.645$

Por lo tanto la regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } \frac{\bar{X} - 1000}{25/\sqrt{20}} > 1.645 \\ \text{aceptar } H_0 & \text{si } \frac{\bar{X} - 1000}{25/\sqrt{20}} \leq 1.645 \end{cases}$$

El estadístico toma el valor $Z = \frac{\bar{X} - 1000}{25/\sqrt{20}}$ toma el valor $z_0 = \frac{1040 - 1000}{25/\sqrt{20}} = 7.1554$

Como $z_0 = 7.1554 > 1.645 = z_{0.05}$ se rechaza H_0

4.3 P- valor

Hasta ahora se dieron los resultados de una prueba de hipótesis estableciendo si la hipótesis nula fue o no rechazada con un valor especificado de α o nivel de significancia.

A menudo este planteamiento resulta inadecuado, ya que no proporciona ninguna idea sobre si el valor calculado del estadístico está apenas en la región de rechazo o bien ubicado dentro de ella. Además, esta forma de establecer los resultados impone a otros usuarios el nivel de significancia predeterminado.

Para evitar estas dificultades, se adopta el enfoque del **p-valor**. El **valor p** o **p-valor** es la probabilidad de que el estadístico de prueba tome un valor que sea al menos tan extremo como el valor observado del estadístico de prueba cuando la hipótesis nula es verdadera. Es así como el p-valor da mucha información sobre el peso de la evidencia contra H_0 , de modo que el investigador pueda llegar a una conclusión para cualquier nivel de significancia especificado.

La definición formal del p-valor es la siguiente:

El **valor p** es el nivel de significancia más pequeño que conduce al rechazo de la hipótesis nula H_0

Para las pruebas de distribuciones normales presentadas hasta el momento, es sencillo calcular el p-valor.

Si z_0 es el valor calculado del estadístico de prueba Z , entonces el p-valor es

a) si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$

$$p\text{-valor} = P(|Z| > |z_0|) = 1 - P(|Z| < |z_0|) = 1 - [\Phi(|z_0|) - \Phi(-|z_0|)] = 1 - [2\Phi(|z_0|) - 1] = 2[1 - \Phi(|z_0|)]$$

b) si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu > \mu_0$

$$p\text{-valor} = P(Z > z_0) = 1 - P(Z \leq z_0) = 1 - \Phi(z_0)$$

c) si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu < \mu_0$

$$p\text{-valor} = P(Z < z_0) = \Phi(z_0)$$

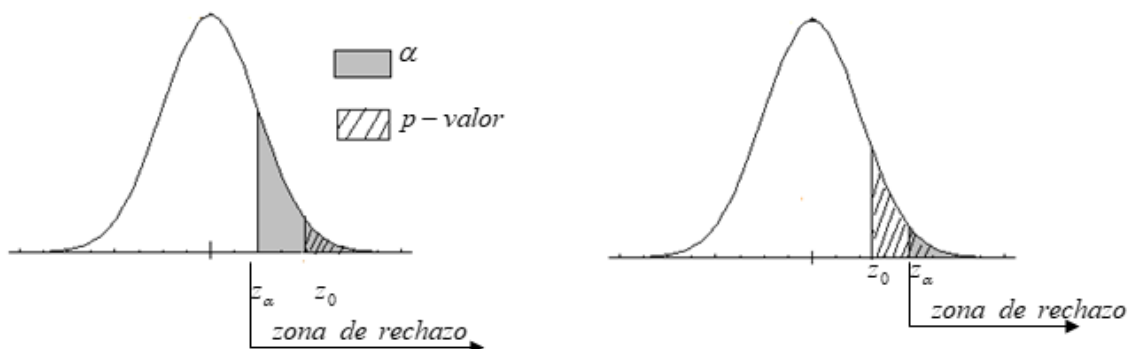
Un p-valor muy chico significa mucha evidencia en contra de H_0 ; un p-valor alto significa que no hay evidencia en contra H_0

Notar que:

Si $\alpha < p\text{-valor}$ entonces se acepta H_0 con nivel de significancia α

Si $\alpha > p\text{-valor}$ entonces se rechaza H_0 con nivel de significancia α

Esto se ilustra en las siguientes figuras:



Ejemplos:

1- En el ejemplo anteúltimo referido al porcentaje deseado de SiO_2 en cierto tipo de cemento aluminoso las hipótesis eran: $H_0 : \mu = 5.5$ contra $H_1 : \mu \neq 5.5$; y el estadístico de prueba tomó el valor

$$z_0 = 0.333333 < 2.575 = z_{\frac{0.01}{2}}; \text{ por lo tanto se aceptaba } H_0.$$

En esta caso $p\text{-valor} = P(|Z| > |z_0|) = 2[1 - \Phi(|z_0|)] = 2[1 - \Phi(0.33333)] = 2[1 - 0.62930] = 0.7414$

Como el p-valor es muy alto no hay evidencia en contra H_0 . Se necesitaría tomar un valor de α mayor a 0.7414 para rechazar H_0 .

2- En el último ejemplo, sobre la duración, en horas, de un foco de 75 watts, las hipótesis eran

$H_0 : \mu = 1000$ contra $H_1 : \mu > 1000$; y el estadístico Z tomó el valor

$$z_0 = 7.1554 > 1.645 = z_{0.05};$$

por lo tanto se rechazaba H_0 .

En este caso

$$p\text{-valor} = P(Z > z_0) = 1 - \Phi(z_0) = 1 - \Phi(7.1554) \approx 0$$

Como el p-valor es casi cero hay mucha evidencia en contra de H_0 . Prácticamente para ningún valor de α se acepta H_0

4.4 - Error de tipo II y selección del tamaño de la muestra

En la prueba de hipótesis el investigador selecciona directamente la probabilidad del error de tipo I. Sin embargo, la probabilidad β de cometer error de tipo II depende del tamaño de la muestra y del valor verdadero del parámetro desconocido.

Supongamos las hipótesis

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu \neq \mu_0$$

Entonces si anotamos con μ al valor verdadero del parámetro

$$\beta = P(\text{aceptar } H_0 / H_0 \text{ es falsa}) = P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \leq z_{\frac{\alpha}{2}} \mid \mu \neq \mu_0\right)$$

Como la hipótesis nula es falsa, entonces $\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ **no tiene distribución** $N(0,1)$

Por lo tanto hacemos lo siguiente:

$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu + \mu - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}}; \text{ y ahora como } \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1) \text{ pues se estandarizó a}$$

\bar{X} con el verdadero μ , entonces

$$\begin{aligned}\beta &= P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \leq z_{\frac{\alpha}{2}} \mid \mu \neq \mu_0\right) = P\left(-z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}} \mid \mu \neq \mu_0\right) = \\ &= P\left(-z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}} \mid \mu \neq \mu_0\right) = P\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) = \\ &= \Phi\left(z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) = \Phi\left(z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right)\end{aligned}$$

En consecuencia

Si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$, entonces

$$\beta(\mu) = \Phi\left(z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right)$$

Para un valor específico de μ y un valor de α dado, podemos preguntarnos qué tamaño de muestra se necesita para que β sea menor que un valor dado en particular β_0 .

Por ejemplo si $\mu - \mu_0 > 0$ entonces podemos aproximar $\Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) \approx 0$, y planteamos que

$\beta(\mu) = \Phi\left(z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) < \beta_0$. Buscamos en la tabla de la $N(0,1)$ para qué z se cumple que $\Phi(z) = \beta_0$, lo anotamos $-z_{\beta_0}$, y entonces podemos escribir

$$z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n} < -z_{\beta_0} \Rightarrow z_{\frac{\alpha}{2}} + z_{\beta_0} < \frac{(\mu - \mu_0)}{\sigma}\sqrt{n} \Rightarrow n > \frac{\left(z_{\frac{\alpha}{2}} + z_{\beta_0}\right)^2 \sigma^2}{(\mu - \mu_0)^2}$$

En el caso de ser $\mu - \mu_0 < 0$ entonces podemos aproximar $\Phi\left(z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) \approx 1$, y planteamos que

$$\beta(\mu) = 1 - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right) < \beta_0. \text{ Es decir } 1 - \beta_0 < \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma}\sqrt{n}\right)$$

Buscamos en la tabla de la $N(0,1)$ para qué z se cumple que $\Phi(z) = 1 - \beta_0$, lo anotamos z_{β_0} , y entonces podemos escribir

$$-z_{\frac{\alpha}{2}} - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n} > z_{\beta_0} \Rightarrow z_{\frac{\alpha}{2}} + z_{\beta_0} < -\frac{(\mu - \mu_0)}{\sigma} \sqrt{n} \stackrel{\mu - \mu_0 < 0}{\Rightarrow} n > \frac{\left(z_{\frac{\alpha}{2}} + z_{\beta_0}\right)^2 \sigma^2}{(\mu - \mu_0)^2}$$

En consecuencia queda la misma fórmula que la anterior
Por lo tanto

Si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$, entonces

$$n > \frac{\left(z_{\alpha/2} + z_{\beta_0}\right)^2 \sigma^2}{(\mu - \mu_0)^2}$$

En forma análoga se puede probar que si las hipótesis son

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu > \mu_0$$

Entonces

$$\begin{aligned} \beta &= P(\text{aceptar } H_0 / H_0 \text{ es falsa}) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\alpha} \mid \mu \neq \mu_0\right) = \\ &= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}} \leq z_{\alpha} \mid \mu \neq \mu_0\right) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) = \Phi\left(z_{\alpha} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) = \Phi\left(z_{\alpha} - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n}\right) \end{aligned}$$

Entonces

Si las hipótesis son : $H_0 : \mu = \mu_0$ contra $H_1 : \mu > \mu_0$ entonces

$$\beta(\mu) = \Phi\left(z_{\alpha} - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n}\right)$$

Y si tenemos las hipótesis $H_0 : \mu = \mu_0$ contra $H_1 : \mu < \mu_0$

$$\begin{aligned} \beta &= P(\text{aceptar } H_0 / H_0 \text{ es falsa}) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq -z_{\alpha} \mid \mu \neq \mu_0\right) = \\ &= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}} \geq -z_{\alpha} \mid \mu \neq \mu_0\right) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq -z_{\alpha} - \frac{(\mu - \mu_0)}{\sigma/\sqrt{n}}\right) = 1 - \Phi\left(-z_{\alpha} - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n}\right) \end{aligned}$$

Entonces

Si las hipótesis son : $H_0 : \mu = \mu_0$ contra $H_1 : \mu < \mu_0$ entonces

$$\beta(\mu) = 1 - \Phi\left(-z_\alpha - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n}\right)$$

Y además con una deducción análoga al caso de alternativa bilateral:

Si las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu > \mu_0$, (o $H_1 : \mu < \mu_0$) entonces

$$n > \frac{(z_\alpha + z_{\beta_0})^2 \sigma^2}{(\mu - \mu_0)^2}$$

Ejemplos:

1- En el ejemplo referido al porcentaje deseado de SiO_2 en cierto tipo de cemento aluminoso las hipótesis eran: $H_0 : \mu = 5.5$ contra $H_1 : \mu \neq 5.5$; y el estadístico de prueba tomó el valor $z_0 = 0.333333 < 2.575 = z_{\frac{0.01}{2}}$; por lo tanto se aceptaba H_0 . Teníamos $n = 16$ y $\sigma = 3$

Si el verdadero promedio de porcentaje es $\mu = 5.6$ y se realiza una prueba de nivel $\alpha = 0.01$ con base en $n = 16$, ¿cuál es la probabilidad de detectar esta desviación?

¿Qué valor de n se requiere para satisfacer $\alpha = 0.01$ y $\beta(5.6) = 0.01$?

Solución:

La probabilidad de detectar la desviación es la potencia del test cuando $\mu = 5.6$, es decir

$$\pi(5.6) = P\left(\text{rechazar } H_0 / H_0 \text{ es falsa}\right) = 1 - \beta(5.6)$$

Como estamos con hipótesis alternativa bilateral, calculamos

$$\begin{aligned} \beta(5.6) &= \Phi\left(z_{\frac{\alpha}{2}} - \frac{(5.6 - \mu_0)}{\sigma} \sqrt{n}\right) - \Phi\left(-z_{\frac{\alpha}{2}} - \frac{(5.6 - \mu_0)}{\sigma} \sqrt{n}\right) = \\ &= \Phi\left(2.575 - \frac{(5.6 - 5.5)}{3} \sqrt{16}\right) - \Phi\left(-2.575 - \frac{(5.6 - 5.5)}{3} \sqrt{16}\right) = \Phi(2.441) - \Phi(-2.708) = \\ &= 0.99266 - (1 - 0.99664) = 0.9893 \quad \Rightarrow \quad \pi(5.6) = 0.0107 \end{aligned}$$

Ahora se quiere hallar n tal que $\beta(5.6) = 0.01$, como el test es bilateral podemos usar directamente la fórmula con $z_{\beta_0} = z_{0.01} = 2.33$

$$n > \frac{\left(z_{\frac{\alpha}{2}} + z_{\beta_0}\right)^2 \sigma^2}{(\mu - \mu_0)^2} = \frac{(2.575 + 2.33)^2 3^2}{(5.6 - 5.5)^2} = 21653.1225 \quad \Rightarrow \quad n \geq 21654$$

2- En el último ejemplo, sobre la duración, en horas, de un foco de 75 watts, las hipótesis eran

$H_0 : \mu = 1000$ contra $H_1 : \mu > 1000$; y el estadístico Z tomó el valor $z_0 = 7.1554 > 1.645 = z_{0.05}$; por lo tanto se rechazaba H_0 .

En este caso $\sigma = 25$ y $n = 20$

Si la verdadera duración promedio del foco es 1050 horas, ¿cuál es la probabilidad de error de tipo II para la prueba?

¿Qué tamaño de muestra es necesario para asegurar que el error de tipo II no es mayor que 0.10 si la duración promedio verdadera del foco es 1025 hs. ?

Solución:

Como las hipótesis son $H_0 : \mu = 1000$ contra $H_1 : \mu > 1000$ entonces

$$\beta(\mu) = \Phi\left(z_\alpha - \frac{(\mu - \mu_0)}{\sigma} \sqrt{n}\right) = \Phi\left(1.645 - \frac{(1050 - 1000)}{25} \sqrt{20}\right) = \Phi(-7.29927) \neq 0$$

Para hallar n tal que $\beta(1025) \leq 0.1$ aplicamos la fórmula con $z_{\beta_0} = z_{0.1} = 1.285$

$$n > \frac{(z_\alpha + z_{\beta_0})^2 \sigma^2}{(\mu - \mu_0)^2} = \frac{(1.645 + 1.285)^2 25^2}{(1025 - 1000)^2} = 8.584 \Rightarrow n \geq 9$$

4.5- Relación entre test de hipótesis e intervalos de confianza

Existe una estrecha relación entre la prueba de hipótesis bilateral sobre un parámetro μ y el intervalo de confianza de nivel $1 - \alpha$ para μ .

Específicamente supongamos que tenemos las hipótesis

$$H_0 : \mu = \mu_0 \quad \text{contra} \quad H_1 : \mu \neq \mu_0$$

La regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } \left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| > z_{\frac{\alpha}{2}} \\ \text{aceptar } H_0 & \text{si } \left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \leq z_{\frac{\alpha}{2}} \end{cases}$$

$$\text{Aceptar } H_0 \quad \text{si} \quad \left| \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \leq z_{\frac{\alpha}{2}} \quad \text{es equivalente a:} \quad \text{aceptar } H_0 \quad \text{si} \quad -z_{\frac{\alpha}{2}} \leq \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}}; \text{ y}$$

esto es a su vez equivalente, despejando μ_0 , a:

$$\text{aceptar } H_0 \quad \text{si} \quad \bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \text{ es decir si}$$

$$\mu_0 \in \left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$$

aceptar H_0 si $\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$; es decir si

$$\mu_0 \in \left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$$

Pero resulta que $\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$ es el intervalo de confianza que se construiría para el verdadero parámetro μ de nivel $1 - \alpha$.

Por lo tanto la regla de decisión queda:

$$\begin{cases} \text{rechazar } H_0 & \text{si } \mu_0 \notin \left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] \\ \text{aceptar } H_0 & \text{si } \mu_0 \in \left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] \end{cases}$$

Ejemplo:

En el ejemplo referido al porcentaje deseado de SiO_2 en cierto tipo de cemento aluminoso las hipótesis eran: $H_0: \mu = 5.5$ contra $H_1: \mu \neq 5.5$;

y teníamos $n = 16$; $\sigma = 3$; un promedio muestral $\bar{x} = 5.25$

Como $\alpha = 0.01$ entonces $z_{\frac{\alpha}{2}} = z_{0.005} = 2.575$

Construimos un intervalo de confianza de nivel $1 - \alpha = 1 - 0.01 = 0.99$

$$\left[\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right] = \left[5.25 - 2.575 \frac{3}{\sqrt{16}}; 5.25 + 2.575 \frac{3}{\sqrt{16}} \right] = [3.31875; 7.18125]$$

Entonces la regla de decisión es:

$$\begin{cases} \text{rechazar } H_0 & \text{si } 5.5 \notin [3.31875; 7.18125] \\ \text{aceptar } H_0 & \text{si } 5.5 \in [3.31875; 7.18125] \end{cases}$$

Como $5.5 \in [3.31875; 7.18125]$, entonces se acepta H_0 .

4.6 – Prueba de hipótesis sobre la media, varianza desconocida para muestras grandes

Hasta ahora se ha desarrollado el procedimiento de test de hipótesis para la hipótesis nula $H_0: \mu = \mu_0$ suponiendo que σ^2 es conocida, pero en la mayoría de las situaciones prácticas σ^2

es desconocida. En general si $n \geq 30$, entonces la varianza muestral S^2 está próxima a σ^2 en la mayor parte de las muestras, de modo que es posible sustituir S^2 por σ^2 . Es decir el estadístico

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \approx N(0,1) \quad \text{aproximadamente, si } n \geq 30 \quad \text{si } H_0: \mu = \mu_0$$

Además, si no podemos decir que la muestra aleatoria proviene de una población normal, sea σ^2 conocida o no, por T.C.L. los estadísticos

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \approx N(0,1) \quad \text{aproximadamente, si } n \geq 30 \quad \text{si } H_0: \mu = \mu_0$$

Y

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \approx N(0,1) \quad \text{aproximadamente, si } n \geq 30 \quad \text{si } H_0: \mu = \mu_0$$

Las pruebas de hipótesis tendrán entonces un nivel de significancia **aproximadamente de α**

Ejemplo:

Un inspector midió el volumen de llenado de una muestra aleatoria de 100 latas de jugo cuya etiqueta afirmaba que contenían 12 oz. La muestra tenía una media de volumen de 11.98 oz y desviación estándar de 0.19 oz. Sea μ la verdadera media del volumen de llenado para todas las latas de jugo recientemente llenadas con esta máquina. El inspector probará $H_0: \mu = 12$ contra $H_1: \mu \neq 12$

- Determinar el p-valor
- ¿Piensa que es factible que la media del volumen de llenado es de 12 oz?

Solución:

La v.a. de interés sería X : “volumen de llenado de una lata tomada al azar”

No se especifica ninguna distribución para X . Anotamos $E(X) = \mu$ y $V(X) = \sigma^2$, ambas desconocidas.

Se toma una muestra de $n = 100$ latas y se obtiene $\bar{x} = 11.98$ y $s = 0.19$

Las hipótesis son $H_0: \mu = 12$ contra $H_1: \mu \neq 12$

El estadístico de prueba es

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{\bar{X} - 12}{S/\sqrt{100}} \quad \text{y si } H_0: \mu = 12 \text{ es verdadera entonces } Z \approx N(0,1)$$

El estadístico Z toma el valor $z_0 = \frac{11.98 - 12}{0.19/\sqrt{100}} = -1.0526$

Como la hipótesis alternativa es bilateral entonces

$$p\text{-valor} = P(|Z| > |z_0|) \approx 2[1 - \Phi(1.0526)] = 2[1 - 0.85314] = 0.29372$$

Como el p-valor es mayor que 0.05 se considera que **no hay evidencia en contra de $H_0: \mu = 12$**

Por lo tanto **es factible que la media del volumen de llenado sea de 12 oz**

4.7 – Prueba de hipótesis sobre la media de una distribución normal, varianza desconocida

Cuando se prueban hipótesis sobre la media μ de una población cuando σ^2 es desconocida es posible utilizar los procedimientos de prueba dados anteriormente siempre y cuando el tamaño de la muestra sea grande ($n \geq 30$). Estos procedimientos son aproximadamente válidos sin importar si la población de interés es normal o no. Pero si la muestra es pequeña y σ^2 es desconocida debe suponerse que la distribución de la variable de interés es normal.

Específicamente, supongamos que la v.a. de interés tiene distribución $N(\mu, \sigma^2)$ donde μ y σ^2 son desconocidas.

Supongamos las hipótesis $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$

Sea $X_1; X_2, \dots, X_n$ una muestra aleatoria de tamaño n de la v.a. X y sean \bar{X} y S^2 la media y la varianza muestrales respectivamente.

El procedimiento se basa en el estadístico

$$T = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$$

El cual, *si la hipótesis nula es verdadera*, tiene distribución *Student con n-1 grados de libertad*. Entonces, para un nivel α prefijado, la regla de decisión es

$$\left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } |T| > t_{\frac{\alpha}{2}, n-1} \\ \text{aceptar } H_0 \text{ si } |T| \leq t_{\frac{\alpha}{2}, n-1} \end{array} \right. \quad \text{es decir} \quad \left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } \left| \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \right| > t_{\frac{\alpha}{2}, n-1} \\ \text{aceptar } H_0 \text{ si } \left| \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \right| \leq t_{\frac{\alpha}{2}, n-1} \end{array} \right.$$

La lógica sigue siendo la misma, si el estadístico de prueba toma un valor inusual, entonces se considera que hay evidencia en contra H_0 y se rechaza la hipótesis nula. Como ahora la distribución del estadístico es Student, nos fijamos si T toma un valor t_0 en las colas de la distribución Student con $n-1$ grados de libertad.

Si la alternativa es $H_1 : \mu > \mu_0$ entonces la regla de decisión es $\left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } T > t_{\alpha, n-1} \\ \text{aceptar } H_0 \text{ si } T \leq t_{\alpha, n-1} \end{array} \right.$

Si la alternativa es $H_1 : \mu < \mu_0$ entonces la regla de decisión es $\left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } T < -t_{\alpha, n-1} \\ \text{aceptar } H_0 \text{ si } T \geq -t_{\alpha, n-1} \end{array} \right.$

Ejemplo:

Antes de que una sustancia se pueda considerar segura para enterrarse como residuo se deben caracterizar sus propiedades químicas. Se toman 6 muestras de lodo de una planta de tratamiento de agua residual en una región y se les mide el pH obteniéndose una media muestral de 6.68 y una desviación estándar muestral de 0.20. ¿Se puede concluir que la media del pH es menor que 7.0? Utilizar $\alpha = 0.05$ y suponer que la muestra fue tomada de una población normal.

Solución:

La v.a. de interés es X : “pH de una muestra de lodo tomada al azar”

Asumimos que X tiene distribución $N(\mu, \sigma^2)$

Las hipótesis serían $H_0 : \mu = 7.0$ contra $H_1 : \mu < 7.0$

El estadístico de prueba es $T = \frac{\bar{X} - 7.0}{S / \sqrt{6}}$ y toma el valor $t_0 = \frac{6.68 - 7.0}{0.20 / \sqrt{6}} = -3.919$

Buscamos en la tabla de la distribución Student $t_{\alpha, n-1} = t_{0.05, 5} = 2.015$

Entonces como $t_0 = -3.919 < -t_{\alpha, n-1} = -t_{0.05, 5} = -2.015$ se rechaza H_0 , por lo tanto hay evidencia que $\mu < 7.0$

P-valor de un test t

En este caso el cálculo del p-valor se realiza considerando:

Si t_0 es el valor calculado del estadístico de prueba T , entonces el p-valor es

a) las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu \neq \mu_0$

$$p\text{-valor} = P(|T| > |t_0|) = 1 - P(|T| \leq |t_0|) = 2(1 - P(T \leq t_0))$$

b) las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu > \mu_0$

$$p\text{-valor} = P(T > t_0) = 1 - P(T \leq t_0)$$

c) las hipótesis son $H_0 : \mu = \mu_0$ contra $H_1 : \mu < \mu_0$

$$p\text{-valor} = P(T \leq t_0)$$

Para calcular el p-valor en una prueba t nos encontramos con la dificultad que las tablas de la Student no son completas, por lo tanto en algunas ocasiones se deberá **acotar** el p-valor

En el ejemplo anterior para calcular el p-valor de la prueba como es un test con alternativa unilateral

$$p\text{-valor} = P(T \leq t_0) = P(T \leq -3.919)$$

Buscamos en la tabla de la distribución Student la fila donde figuran $\nu = 5$ grados de libertad y vemos que el valor 3.919 no está tabulado.

Pero $3.365 < 3.919 < 4.032$, y $P(T_5 > 3.365) = 0.01$ y $P(T_5 > 4.032) = 0.005$

Por lo tanto $0.005 < P(T_5 > 3.919) < 0.01$, es decir

$$0.005 < p\text{-valor} = P(T_5 < -3.919) < 0.01$$

Podemos deducir que existe evidencia de que la media del pH es menor que 7.0

4.8 – Tests de hipótesis sobre la varianza

Supongamos que se desea probar la hipótesis de que la varianza de una población normal es igual a un valor específico, por ejemplo σ_0^2 .

Sea (X_1, X_2, \dots, X_n) una muestra aleatoria de tamaño n de una v.a. X , donde $X \sim N(\mu, \sigma^2)$.

Tomamos como estimador puntual de σ^2 a $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

Luego a partir de este estimador puntual construimos el estadístico $X = \frac{(n-1)S^2}{\sigma^2}$

Este estadístico contiene al parámetro desconocido a estimar σ^2 y ya sabemos que tiene una distribución llamada **ji-cuadrado con n-1 grados de libertad**

Supongamos las hipótesis

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{contra} \quad H_1 : \sigma^2 \neq \sigma_0^2$$

Tomamos como estadístico de prueba a

$$X = \frac{(n-1)S^2}{\sigma_0^2} \quad \text{y} \quad \text{si} \quad H_0 : \sigma^2 = \sigma_0^2 \quad \text{es verdadera, entonces} \quad X = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi_{n-1}^2$$

Nuevamente, el razonamiento es: si el estadístico X que bajo $H_0 : \sigma^2 = \sigma_0^2$ tiene distribución χ_{n-1}^2 toma un valor “inusual”, se considera que hay evidencia en contra H_0

Recordar que la distribución χ_{n-1}^2 es asimétrica. Entonces la regla de decisión es

$$\left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } X > \chi_{\frac{\alpha}{2}, n-1}^2 \text{ ó } X < \chi_{1-\frac{\alpha}{2}, n-1}^2 \\ \text{aceptar } H_0 \text{ si } \chi_{1-\frac{\alpha}{2}, n-1}^2 \leq X \leq \chi_{\frac{\alpha}{2}, n-1}^2 \end{array} \right. \quad \text{donde} \quad X = \frac{(n-1)S^2}{\sigma_0^2}$$

$$\text{Si } H_1 : \sigma^2 > \sigma_0^2 \text{ entonces la regla de decisión es } \left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } X > \chi_{\alpha, n-1}^2 \\ \text{aceptar } H_0 \text{ si } X \leq \chi_{\alpha, n-1}^2 \end{array} \right.$$

$$\text{Si } H_1 : \sigma^2 < \sigma_0^2 \text{ entonces la regla de decisión es } \left\{ \begin{array}{l} \text{rechazar } H_0 \text{ si } X < \chi_{1-\alpha, n-1}^2 \\ \text{aceptar } H_0 \text{ si } X \geq \chi_{1-\alpha, n-1}^2 \end{array} \right.$$

Para calcular el p-valor, si el estadístico X tomó el valor x_0 , y teniendo en cuenta que no hay simetría en la distribución ji-cuadrado, hacemos:

$$\text{Si } H_1 : \sigma^2 > \sigma_0^2 \text{ entonces } p\text{-valor} = P(X > x_0)$$

$$\text{Si } H_1 : \sigma^2 < \sigma_0^2 \text{ entonces } p\text{-valor} = P(X < x_0)$$

$$\text{Si } H_1 : \sigma^2 \neq \sigma_0^2 \text{ entonces } p\text{-valor} = 2 \min \left(P(X < x_0), P(X > x_0) \right)$$

Ejemplo:

Consideremos nuevamente el ejemplo visto en la sección de intervalos de confianza para la varianza sobre la máquina de llenado de botellas. Al tomar una muestra aleatoria de 20 botellas se obtiene una varianza muestral para el volumen de llenado de $s^2 = 0.0153 \text{ oz}^2$.

Si la varianza del volumen de llenado es mayor que 0.01 oz^2 , entonces existe una proporción inaceptable de botellas que serán llenadas con una cantidad menor de líquido. ¿Existe evidencia en los datos muestrales que sugiera que el fabricante tiene un problema con el llenado de las botellas? Utilice $\alpha = 0.05$

Solución:

La variable de interés es X : “volumen de llenado de una botella tomada al azar”

Asumimos $X \sim N(\mu, \sigma^2)$

Los datos son $s^2 = 0.0153$ de una muestra de tamaño $n = 20$

Las hipótesis son $H_0 : \sigma^2 = 0.01$ contra $H_1 : \sigma^2 > 0.01$

$$\alpha = 0.05 \rightarrow \chi_{\alpha, n-1}^2 = \chi_{0.05, 19}^2 = 30.14$$

El estadístico de prueba es $X = \frac{(n-1)S^2}{\sigma_0^2} = \frac{19 \times S^2}{0.01}$ y toma el valor

$$x_0 = \frac{19 \times S^2}{0.01} = \frac{19 \times 0.0153}{0.01} = 29.07$$

Como $x_0 = 29.07 < \chi_{0.05, 19}^2 = 30.14$ entonces no hay evidencia fuerte de que la varianza del volumen de llenado sea menor que 0.01

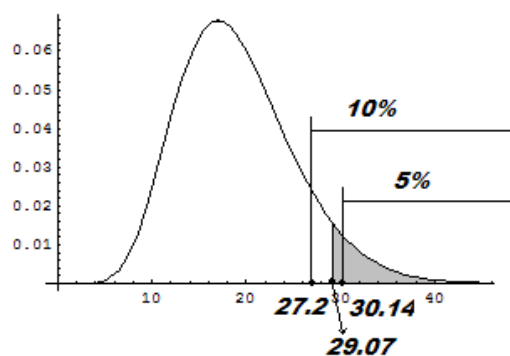
Para calcular el p-valor

$$p\text{-valor} = P(X > x_0) = P(X > 29.07)$$

Buscamos en la tabla de la distribución ji-cuadrado y vemos que en la fila con $\nu = 19$ no figura 29.07, pero $27.20 < 29.07 < 30.14$, y además

$$\begin{cases} P(X > 27.20) = 0.10 \\ P(X > 30.14) = 0.05 \end{cases} \Rightarrow 0.05 < p\text{-valor} < 0.10$$

En la figura siguiente se ilustra la situación



4.9 – Tests de hipótesis sobre una proporción

En muchos problemas se tiene interés en una variable aleatoria que sigue una distribución binomial. Por ejemplo, un proceso de producción que fabrica artículos que son clasificados como aceptables o defectuosos. Lo más usual es modelar la ocurrencia de artículos defectuosos con la distribución binomial, donde el parámetro binomial p representa la proporción de artículos defectuosos producidos.

En consecuencia, muchos problemas de decisión incluyen una prueba de hipótesis con respecto a p .

Consideremos las hipótesis

$$H_0 : p = p_0 \text{ contra } H_1 : p \neq p_0$$

Supongamos que consideramos una muestra aleatoria (X_1, X_2, \dots, X_n) de tamaño n , donde X_i tiene una distribución binomial con parámetros 1 y p : $X_i \sim B(1, p)$.

Ya sabemos que $X = X_1 + X_2 + \dots + X_n$, es una v.a. cuya distribución es binomial con parámetros n y p : $X \sim B(n, p)$. De acuerdo con esto, la variable aleatoria \hat{P} definida: $\hat{P} = \frac{X}{n}$ representa la proporción de individuos de la muestra que verifican la propiedad de interés. Además

$$E(\hat{P}) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(X) = \frac{1}{n} np = p, \text{ y } V(\hat{P}) = V\left(\frac{X}{n}\right) = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}$$

Consideramos el estadístico de prueba

$$Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Si $H_0 : p = p_0$ es verdadera entonces $Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \approx N(0,1)$ aproximadamente por

T.C.L.

Por lo tanto la regla de decisión es

$$\begin{cases} \text{rechazar } H_0 & \text{si } |Z| > z_{\frac{\alpha}{2}} \\ \text{aceptar } H_0 & \text{si } |Z| \leq z_{\frac{\alpha}{2}} \end{cases} \quad \text{donde } Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Si $H_1 : p > p_0$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } Z > z_\alpha \\ \text{aceptar } H_0 & \text{si } Z \leq z_\alpha \end{cases}$

Si $H_1 : p < p_0$ entonces la regla de decisión es $\begin{cases} \text{rechazar } H_0 & \text{si } Z < -z_\alpha \\ \text{aceptar } H_0 & \text{si } Z \geq -z_\alpha \end{cases}$

Observaciones:

1- La prueba descrita anteriormente requiere que la proporción muestral esté normalmente distribuida. Esta suposición estará justificada siempre que $np_0 > 10$ y $n(1-p_0) > 10$, donde p_0 es la proporción poblacional que se especificó en la hipótesis nula.

2- También se podía haber tomado como estadístico de prueba a $Z = \frac{X - np_0}{\sqrt{np_0(1-p_0)}}$ donde

$$X \sim B(n, p)$$

Ejemplo:

Un fabricante de semiconductores produce controladores que se emplean en aplicaciones de motores automovilísticos. El cliente requiere que la fracción de controladores defectuosos en uno de los pasos de manufactura críticos no sea mayor que 0.05, y que el fabricante demuestre esta característica del proceso de fabricación con este nivel de calidad, utilizando $\alpha = 0.05$. El fabricante de semiconductores toma una muestra aleatoria de 200 dispositivos y encuentra que 4 de ellos son defectuosos. ¿El fabricante puede demostrar al cliente la calidad del proceso?

Solución:

Sea la v.a. X : “número de controladores defectuosos en la muestra”

Entonces $X \sim B(200, p)$ donde p es la proporción de controladores defectuosos en el proceso

Las hipótesis son $H_0 : p = 0.05$ contra $H_1 : p < 0.05$

Como $\alpha = 0.05$ entonces $-z_\alpha = -z_{0.05} = -1.645$

El estadístico de prueba es $Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\hat{P} - 0.05}{\sqrt{\frac{0.05(1-0.05)}{200}}}$ y toma el valor $z_0 = -1.95$

Como $z_0 = -1.95 < -z_\alpha = -z_{0.05} = -1.645$ entonces se rechaza H_0 , y se concluye que la fracción de controladores defectuosos es menor que 0.05.

Calculamos el p-valor

$$p\text{-valor} = P(Z < z_0) = P(Z < -1.95) \approx \Phi(-1.95) = 0.0256$$

4.10 - Valor de β y selección del tamaño de la muestra

Podemos obtener expresiones aproximadas para la probabilidad de cometer error de tipo II de manera análoga a las obtenidas para los test para la media

Si $H_1 : p \neq p_0$ entonces

$$\begin{aligned} \beta(p) &= P\left(\text{aceptar } H_0 / H_0 \text{ es falsa}\right) \approx \\ &\approx \Phi\left(\frac{p_0 - p + z_{\frac{\alpha}{2}} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) - \Phi\left(\frac{p_0 - p - z_{\frac{\alpha}{2}} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) \end{aligned}$$

Si $H_1 : p < p_0$ entonces

$$\beta(p) = P\left(\text{aceptar } H_0 / H_0 \text{ es falsa}\right) \approx 1 - \Phi\left(\frac{p_0 - p - z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

Si $H_1: p > p_0$ entonces

$$\beta(p) = P\left(\text{aceptar } H_0 / H_0 \text{ es falsa}\right) \approx \Phi\left(\frac{p_0 - p + z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

Estas ecuaciones pueden resolverse para encontrar el tamaño aproximado de la muestra n para que con un nivel de significancia de α la probabilidad de cometer error de tipo II sea menor o igual que un valor específico β_0 . Las ecuaciones se deducen como en casos anteriores y son

$$\text{Si } H_1: p \neq p_0 \text{ entonces } n \geq \left(\frac{\frac{z_\alpha \sqrt{p_0(1-p_0)} + z_{\beta_0} \sqrt{p(1-p)}}{2}}{p - p_0}\right)^2$$

$$\text{Si } H_1: p < p_0 \text{ ó } H_1: p > p_0 \text{ entonces } n \geq \left(\frac{z_\alpha \sqrt{p_0(1-p_0)} + z_{\beta_0} \sqrt{p(1-p)}}{p - p_0}\right)^2$$

Ejemplo:

Volviendo al ejemplo anterior, supongamos que la verdadera proporción de componentes defectuosos en el proceso es $p = 0.03$, ¿cuál es el valor de β si $n = 200$ y $\alpha = 0.05$?

Solución:

Ya que la alternativa es $H_1: p < p_0$ aplicamos la fórmula

$$\begin{aligned} \beta(p) &= P\left(\text{aceptar } H_0 / H_0 \text{ es falsa}\right) \approx 1 - \Phi\left(\frac{p_0 - p - z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right) = \\ &= 1 - \Phi\left(\frac{0.05 - 0.03 - 1.645 \sqrt{\frac{0.05(1-0.05)}{200}}}{\sqrt{\frac{0.03(1-0.03)}{200}}}\right) = 1 - \Phi(-0.44) = 0.67 \end{aligned}$$

Como la probabilidad de aceptar que el proceso tiene la calidad deseada cuando en realidad $p = 0.03$ es bastante alta, podemos preguntar qué tamaño de muestra se necesita para que en el test anterior sea $\beta < 0.1$ si la verdadera proporción de defectuosos es $p = 0.03$. En este caso aplicamos la fórmula donde $z_{\beta_0} = z_{0.1} = 1.28$

$$n \geq \left(\frac{z_\alpha \sqrt{p_0(1-p_0)} + z_{\beta_0} \sqrt{p(1-p)}}{p - p_0}\right)^2 = \left(\frac{1.645 \sqrt{0.05(1-0.05)} + 1.28 \sqrt{0.03(1-0.03)}}{0.03 - 0.05}\right)^2 \approx 832$$

La muestra requerida es muy grande, pero la diferencia a detectar $p - p_0 = 0.03 - 0.05$ es bastante pequeña.

Practica**Test de Hipótesis**

- 1) Se hace una prueba de la hipótesis $H_0: \mu \leq 10$ contra $H_1: \mu > 10$. Para cada una de las situaciones siguientes, determine si la decisión fue correcta u ocurrieron errores de tipo I o II:
 - a) siendo el verdadero $\mu = 8$, H_0 es rechazada ;
 - b) siendo el verdadero $\mu = 10$, H_0 no es rechazada
 - c) siendo el verdadero $\mu = 14$, H_0 no es rechazada ;
 - d) siendo el verdadero $\mu = 12$, H_0 es rechazada

- 2) Un proceso de fabricación produce cojinetes de bola con diámetros que tienen una distribución normal y una desviación estándar de $\sigma = 0.04$ cm. Los cojinetes de bola que tienen diámetros que son muy pequeños o muy grandes son indeseables. Para poner a prueba la hipótesis nula de que $\mu = 0.5$ cm se selecciona al azar una muestra de 25 y se encuentra que la media muestral es 0.51
 - a) Establezca las hipótesis nula y alternativa tales que el rechazo de la hipótesis nula implicará que los cojinetes de bola son indeseables.
 - b) Con $\alpha = 0.02$, ¿cuál es el valor crítico para el estadístico de prueba?. Realice el test.

- 3) Cuando está operando adecuadamente, una planta química tiene una media de producción diaria de por lo menos 740 toneladas. La producción se mide en una muestra aleatoria simple de 60 días. La muestra tenía una media de 715 toneladas por día y desviación estándar de 24 toneladas por día. Sea μ la media de la producción diaria de la planta. Un ingeniero prueba que

$$H_0: \mu \geq 740 \text{ contra } H_1: \mu < 740.$$
 - a) Determine el p-valor
 - b) ¿Piensa que es factible que la planta esté operando adecuadamente o está convencido de que la planta no funciona en forma adecuada?. Explique su razonamiento.

- 4) a) Se realiza una prueba de hipótesis y el P-valor es 0.03. Justificando su respuesta, diga si las siguientes afirmaciones son verdaderas o falsas:
 - a1) H_0 se rechaza a un nivel de 5%
 - a2) H_0 se rechaza a un nivel de 2%
 - a3) H_0 no se rechaza a un nivel de 10%
 b) Se diseña un programa de tratamiento de aguas residuales para producir agua tratada con pH de 7. Sea μ la media del pH del agua tratada mediante dicho proceso. Se medirá el pH de 20 muestras de agua y se realizará una prueba de hipótesis

$$H_0: \mu = 7 \text{ contra } H_1: \mu \neq 7.$$
 Suponga que se sabe por experimentos previos que el pH del agua es normal con desviación estándar 0.5.
 - b1) Si la prueba se hace a un nivel de 5%, ¿cuál es la región de rechazo?
 - b2) Si la media muestral del pH es 6.77, ¿se rechaza H_0 a un nivel de 10%?
 - b3) Si la media muestral del pH es 6.77, ¿se rechaza H_0 a un nivel de 1%?
 - b4) Si el valor 1.79 representa un punto crítico, ¿cuál es el nivel de la prueba?

- 5) Un ingeniero está probando la resistencia a la compresión del concreto. Prueba 12 muestras de concreto y obtiene un promedio de 2.259 psi y un desvío estándar de 0.035 psi. Suponga que la resistencia a la compresión sigue una distribución normal.
- a) Construya un intervalo de confianza al 95% para la resistencia media.
 - b) Utilizando el intervalo del inciso a) pruebe la hipótesis
 $H_0: \mu = 2.250$ contra $H_1: \mu \neq 2.250$
con nivel de significancia $\alpha = 0.05$.
- 6) Se publica un informe sobre las cifras del número anual de kilowatts-hora que gastan varios aparatos electrodomésticos. Se afirma que una aspiradora gasta un promedio de 46 kilowatts-hora por año. Si una muestra aleatoria de 12 hogares que se incluye en un estudio planeado indica que las aspiradoras gastan un promedio de 42 kilowatts-hora con una desviación estándar de 11.9 kilowatts-hora por año, en un nivel de significancia de 0.05, ¿esto sugiere que las aspiradoras gastan, en promedio, menos de 46 kilowatts-hora anualmente?. Suponga que la población de kilowatts-hora es normal.
- 7) Suponga que ha comprado una máquina de llenado para bolsas de dulces que contendrá 16 onzas de éstos. Suponga que los pesos de las bolsas llenas están distribuidos en forma normal. Una muestra aleatoria de diez bolsas produce los siguientes datos (en onzas): 15.87, 16.02, 15.78, 15.83, 15.69, 15.81, 16.04, 15.81, 15.92, 16.10. Con base en estos datos, ¿puede concluir que la media del peso de llenado es, en realidad, menor que 16 onzas con un nivel de significancia de 0.05?
- 8) Un fabricante de baterías para automóvil afirma que sus baterías durarán, en promedio, 3 años con una varianza de 1 año. Si 5 de estas baterías tienen duraciones de 1.9, 2.4, 3.0, 3.5 y 4.2 años, construya un intervalo de confianza de 95% para la varianza. Suponga que la población de duraciones de las baterías se distribuye de forma aproximadamente normal. ¿Se puede concluir que la afirmación del fabricante de que la varianza es 1 es válida? Justifique su respuesta.
- 9) Una muestra aleatoria de 6 vigas de acero tiene una resistencia a la compresión promedio de 58392 psi (libras por pulgada cuadrada) con una desviación estándar de 648 psi. Suponiendo normalidad
- a) Use esta información y el nivel de significancia $\alpha = 0.05$ para probar si la verdadera resistencia a la compresión media del acero del que provino la muestra es de 58000 psi.
 - b) Pruebe la hipótesis nula $\sigma = 600$ psi contra la alternativa $\sigma > 600$ psi. Utilice $\alpha = 0.05$.
- 10) Un fabricante de estaciones de trabajo de computadora está probando un nuevo proceso de ensamble automatizado. El proceso actual tiene una tasa de defectos de 5%. En una muestra de 400 estaciones de trabajo ensambladas con el nuevo proceso, 15 tenían defecto. ¿Se puede concluir que el nuevo proceso tiene una tasa menor de defectos?. Calcule el p-valor.

5 – REGRESIÓN LINEAL SIMPLE

5.1 – Introducción

En muchos problemas existe una relación entre dos o más variables, y resulta de interés estudiar la naturaleza de esa relación. El **análisis de regresión** es la técnica estadística para el modelado y la investigación de la relación entre dos o más variables. Veamos un ejemplo.

Los resortes se usan en aplicaciones por su capacidad para alargarse (contraerse) bajo carga. La rigidez de un resorte se mide con la *constante del resorte*, que es la longitud del resorte que se alargará por unidad de la fuerza o de la carga. Para asegurarse de que un resorte dado funciona adecuadamente es necesario calcular la constante de resorte con exactitud y precisión.

En este experimento hipotético un resorte se cuelga verticalmente con un extremo fijo, y los pesos se cuelgan uno tras otro del otro extremo. Después de colgar cada peso se mide la longitud del resorte. Sean x_1, x_2, \dots, x_n los pesos, y sea l_i la longitud del resorte bajo la carga x_i .

La ley de Hooke establece que

$$l_i = \beta_0 + \beta_1 x_i$$

donde β_0 representa la longitud del resorte cuando no tiene carga y β_1 es la constante del resorte.

Sea y_i la longitud **medida** del resorte bajo la carga x_i . Debido al error de medición y_i será diferente de la longitud verdadera l_i . Se escribe como

$$y_i = l_i + \varepsilon_i$$

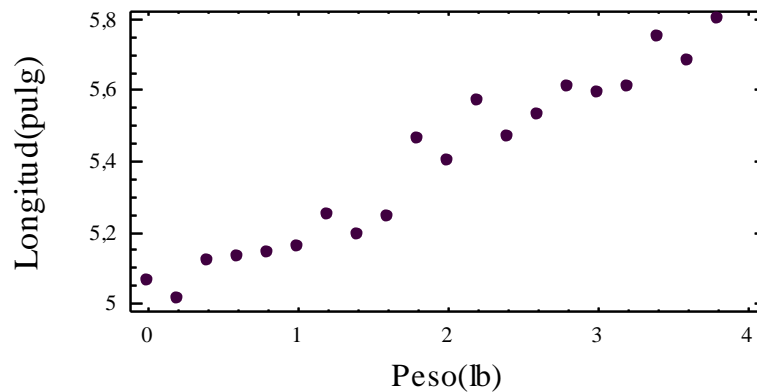
donde ε_i es el error en la i -ésima medición. Al combinar ambas ecuaciones se obtiene

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (1)$$

En la ecuación (1), y_i es la **variable dependiente**, x_i es la **variable independiente**, β_0 y β_1 son los **coeficientes de regresión**, y ε_i se denomina **error**. A la ecuación (1) se la llama **modelo de regresión lineal simple**.

La tabla siguiente presenta los resultados del experimento y la figura el **diagrama de dispersión** de y contra x .

<i>Peso (lb)</i>	<i>Longitud medida (pulg)</i>	<i>Peso (lb)</i>	<i>Longitud medida (pulg)</i>
x	y	x	y
0,0	5,06	2,0	5,40
0,2	5,01	2,2	5,57
0,4	5,12	2,4	5,47
0,6	5,13	2,6	5,53
0,8	5,14	2,8	5,61
1,0	5,16	3,0	5,59
1,2	5,25	3,2	5,61
1,4	5,19	3,4	5,75
1,6	5,24	3,6	5,68
1,8	5,46	3,8	5,80



La idea es utilizar estos datos para *estimar* los coeficientes de regresión. Si no hubiese error en la medición, los puntos se encontrarían en una línea recta con pendiente β_1 y ordenada al origen β_0 , y estas cantidades serían fáciles de determinar. La idea es entonces que los puntos están dispersos de manera aleatoria alrededor de una recta que es la recta de regresión lineal $l = \beta_0 + \beta_1 x$.

En general podemos decir que al fijar el valor de x observamos el valor de la variable Y . Si bien x es fijo, el valor de Y está afectado por el *error aleatorio* ε . Por lo tanto ε *determina las propiedades de Y* . Escribimos en general

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

donde x es, por ahora, una variable no aleatoria, ε *es la v.a. del error* y *asumimos que*

$$E(\varepsilon) = 0 \quad \text{y} \quad V(\varepsilon) = \sigma^2$$

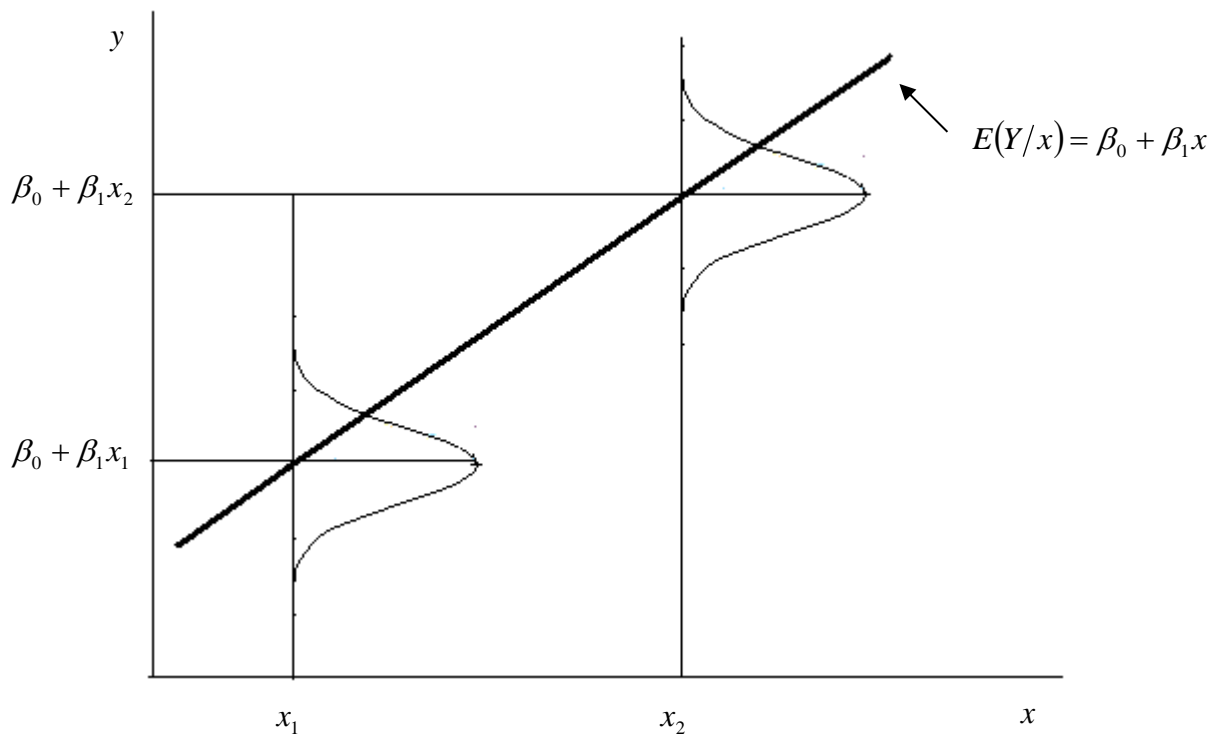
Entonces Y es una variable aleatoria tal que

$$E(Y/x) = E(\beta_0 + \beta_1 x + \varepsilon) = \beta_0 + \beta_1 x + E(\varepsilon) = \beta_0 + \beta_1 x$$

$$V(Y/x) = V(\beta_0 + \beta_1 x + \varepsilon) = V(\varepsilon) = \sigma^2$$

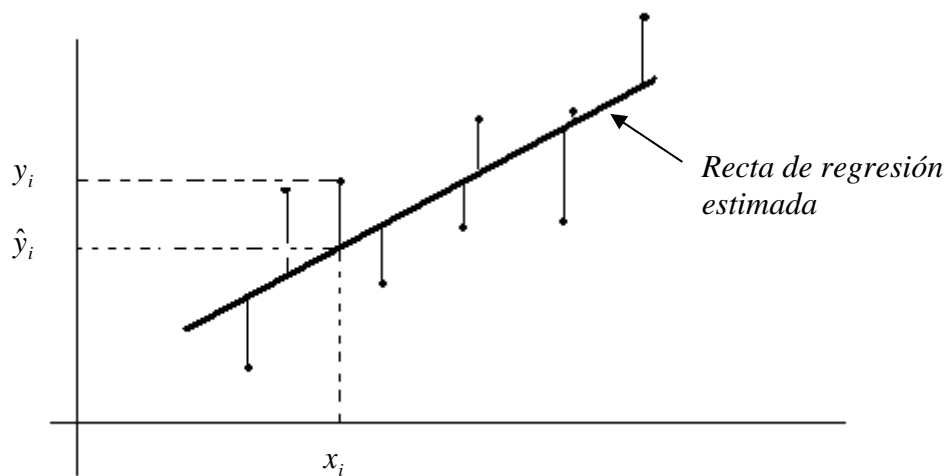
En consecuencia, el modelo de regresión verdadero $E(Y/x) = \beta_0 + \beta_1 x$ es una recta de valores promedio.

Notar que lo anterior implica que existe una distribución de valores de Y para cada x , y que la varianza de esta distribución es la misma para cada x . La siguiente figura ilustra esta situación. Notar que se utilizó una distribución normal para describir la variación aleatoria en ε . Por lo tanto la distribución de Y también será normal. La varianza σ^2 determina la variabilidad en las observaciones Y . por lo tanto, cuando σ^2 es pequeño, los valores observados de Y caen cerca de la línea, y cuando σ^2 es grande, los valores observados de Y pueden desviarse considerablemente de la línea. Dado que σ^2 es constante, la variabilidad en Y para cualquier valor de x es la misma.



5.2 – Regresión lineal simple- Estimación de parámetros

Para estimar los coeficientes de regresión se utiliza el **método de mínimos cuadrados**. Supongamos que se tienen n pares de observaciones $(x_1, y_1); (x_2, y_2); \dots; (x_n, y_n)$. Realizamos una gráfica representativa de los datos y una recta como posible recta de regresión. Anotamos a la **recta de regresión estimada** con $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$



Las estimaciones de β_0 y β_1 deben dar como resultado una línea que en algún sentido se “ajuste mejor” a los datos. El método de mínimos cuadrados consiste en estimar β_0 y β_1 de manera tal que se minimice la suma de los cuadrados de las desviaciones verticales mostradas en la figura anterior.

La suma de los cuadrados de las desviaciones de las observaciones con respecto a la recta de regresión es

$$L = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Los estimadores de mínimos cuadrados de β_0 y β_1 , que anotamos $\hat{\beta}_0$ y $\hat{\beta}_1$, deben satisfacer las siguientes ecuaciones

$$\begin{cases} \frac{\partial L}{\partial \beta_0} = -2 \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \\ \frac{\partial L}{\partial \beta_1} = -2 \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = 0 \end{cases} \quad (2)$$

Después de simplificar las expresiones anteriores, se llega a

$$\begin{cases} n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \\ \hat{\beta}_0 \sum_{i=1}^n x_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \end{cases} \quad (3)$$

Las ecuaciones (3) reciben el nombre de **ecuaciones normales de mínimos cuadrados**.

La solución de estas ecuaciones dan como resultado las **estimaciones de mínimos cuadrados** $\hat{\beta}_0$ y $\hat{\beta}_1$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (4)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - \frac{\left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n}}{\sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n}} \quad (5)$$

donde $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$ y $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

Las diferencias $e_i = y_i - \hat{y}_i$ con $i = 1, \dots, n$ se llaman **residuos**. El residuo e_i describe el error en el ajuste del modelo en la i -ésima observación y_i .

Para agilizar la notación son útiles los siguientes símbolos

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{\left(\sum_{i=1}^n x_i\right)^2}{n} \quad (6)$$

$$S_{xy} = \sum_{i=1}^n y_i (x_i - \bar{x}) = \sum_{i=1}^n x_i y_i - \frac{\left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n} \quad (7)$$

Entonces con esta notación podemos escribir $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$

Ejemplo:

Ajustamos un modelo de regresión lineal a los datos del ejemplo anterior. La estimación de la constante del resorte es $\hat{\beta}_1$ y $\hat{\beta}_0$ la estimación de la longitud sin carga.

De la tabla obtenemos

$$\bar{x} = 1.9 \quad \bar{y} = 5.3885$$

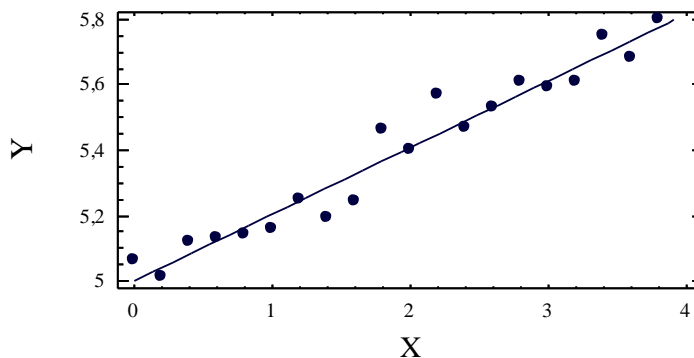
$$S_{xx} = 26.6 \quad S_{xy} = 5.4430$$

$$\text{Entonces } \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{5.4430}{26.6} = 0.2046 \quad \text{y} \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 5.3885 - 0.2046 \times 1.9 = 4.9997$$

La ecuación de la recta estimada es

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \quad \Rightarrow \quad \hat{y} = 4.9997 + 0.2046x$$

La figura siguiente muestra el gráfico de dispersión con la recta de regresión estimada



Po-

demos utilizar la recta de regresión estimada para predecir la longitud del resorte bajo una carga

determinada, por ejemplo con una carga de 1.3 lb:

$$\hat{y} = 4.9997 + 0.2046(1.3) = 5.27 \text{ pulg.}$$

Podemos también estimar la longitud del resorte bajo una carga de 1.4 lb:

$$\hat{y} = 4.9997 + 0.2046(1.4) = 5.29 \text{ pulg.}$$

Notar que la longitud medida para una carga de 1.4 lb es 5.19 pulg., pero la estimación de mínimos cuadrados de 5.29 pulg. Está basada en todos los datos y es más precisa (tiene menor incertidumbre). Más adelante calcularemos la varianza de estos estimadores.

Observaciones:

1- Las estimaciones de mínimos cuadrados $\hat{\beta}_1$ y $\hat{\beta}_0$ son valores de variables aleatorias y dicho valor varía con las muestras. Los coeficientes de regresión β_0 y β_1 son constantes desconocidas que estimamos con $\hat{\beta}_1$ y $\hat{\beta}_0$.

2- Los residuos e_i no son lo mismo que los errores ε_i . Cada residuo es la diferencia $e_i = y_i - \hat{y}_i$ entre el valor observado y el valor ajustado, y se pueden calcular a partir de los datos. Los errores ε_i representan la diferencia entre los valores medidos y_i y los valores $\beta_0 + \beta_1 x_i$. Como los valores verdaderos de β_0 y β_1 no se conocen entonces, los errores no se pueden calcular.

3- ¿Qué sucede si se quisiera estimar la longitud del resorte bajo una carga de 100 lb? La estimación de mínimos cuadrados es $\hat{y} = 4.9997 + 0.2046(100) = 25.46 \text{ pulg.}$ Pero esta estimación no es confiable, pues ninguno de los pesos en el conjunto de datos es tan grande. Es probable que el resorte se deformara, por lo que la ley de Hooke no valdría. Para muchas variables las relaciones lineales valen dentro de cierto rango, pero no fuera de él. Si se quiere saber cómo respondería el resorte a una carga de 100 lb se deben incluir pesos de 100 lb o mayores en el conjunto de datos.

Por lo tanto **no hay que extrapolar una recta ajustada fuera del rango de los datos**. La relación lineal puede no ser válida ahí.

5.3 – Propiedades de los estimadores de mínimos cuadrados y estimación de σ^2

Los *estimadores* de β_1 y β_0 los anotamos

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x} \qquad \hat{\beta}_1 = \frac{S_{xY}}{S_{xx}} = \frac{\sum_{i=1}^n Y_i (x_i - \bar{x})}{S_{xx}} \quad (8)$$

Como $\hat{\beta}_1$ y $\hat{\beta}_0$ son estimadores de β_1 y β_0 respectivamente, son variables aleatorias, por lo tanto podemos calcular su esperanza y varianza. Como estamos asumiendo que x no es v.a. entonces $\hat{\beta}_1$ y $\hat{\beta}_0$ son funciones de la v.a. Y .

Recordemos que el modelo es $Y = \beta_0 + \beta_1 x + \varepsilon$, si medimos n veces la variable Y tenemos

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

donde **asumimos** $E(\varepsilon_i) = 0$; $V(\varepsilon_i) = \sigma^2$ $i = 1, 2, \dots, n$ y $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ **independientes**

Por lo tanto

$$E(Y/x_i) = E(\beta_0 + \beta_1 x_i + \varepsilon) = \beta_0 + \beta_1 x_i + E(\varepsilon) = \beta_0 + \beta_1 x_i$$

$$V(Y/x_i) = V(\beta_0 + \beta_1 x_i + \varepsilon_i) = V(\varepsilon_i) = \sigma^2$$

Consideramos $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^n Y_i(x_i - \bar{x})}{S_{xx}}$. Podemos ver a $\hat{\beta}_1$ como una combinación lineal de las variables Y_i , entonces

$$\begin{aligned} E(\hat{\beta}_1) &= E\left(\frac{S_{xy}}{S_{xx}}\right) = E\left(\frac{\sum_{i=1}^n Y_i(x_i - \bar{x})}{S_{xx}}\right) = \frac{1}{S_{xx}} E\left(\sum_{i=1}^n Y_i(x_i - \bar{x})\right) = \frac{1}{S_{xx}} \sum_{i=1}^n E(Y_i)(x_i - \bar{x}) = \\ &= \frac{1}{S_{xx}} \sum_{i=1}^n (\beta_0 + \beta_1 x_i)(x_i - \bar{x}) = \frac{1}{S_{xx}} \left\{ \beta_0 \sum_{i=1}^n (x_i - \bar{x}) + \beta_1 \sum_{i=1}^n x_i(x_i - \bar{x}) \right\} = \frac{1}{S_{xx}} \beta_1 S_{xx} = \beta_1 \end{aligned}$$

Notar que $\sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - n\bar{x} = \sum_{i=1}^n x_i - n \left(\frac{\sum_{i=1}^n x_i}{n} \right) = 0$

y $\sum_{i=1}^n x_i(x_i - \bar{x}) = \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x}) = S_{xx}$

Por lo tanto $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^n Y_i(x_i - \bar{x})}{S_{xx}}$ es un estimador insesgado de β_1

Veamos ahora la varianza de $\hat{\beta}_1$

$$\begin{aligned} V(\hat{\beta}_1) &= V\left(\frac{S_{xy}}{S_{xx}}\right) = V\left(\frac{\sum_{i=1}^n Y_i(x_i - \bar{x})}{S_{xx}}\right) = \frac{1}{S_{xx}^2} V\left(\sum_{i=1}^n Y_i(x_i - \bar{x})\right) = \frac{1}{S_{xx}^2} \sum_{i=1}^n V(Y_i)(x_i - \bar{x})^2 = \\ &= \frac{1}{S_{xx}^2} \sum_{i=1}^n \sigma^2 (x_i - \bar{x})^2 = \frac{1}{S_{xx}^2} \sigma^2 S_{xx} = \frac{\sigma^2}{S_{xx}} \end{aligned}$$

Por lo tanto

$$\boxed{E(\hat{\beta}_1) = \beta_1 \quad \text{y} \quad V(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}} \quad (9)$$

Con un enfoque similar calculamos la esperanza y la varianza de $\hat{\beta}_0$

$$\begin{aligned} E(\hat{\beta}_0) &= E(\bar{Y} - \hat{\beta}_1 \bar{x}) = E(\bar{y}) - E(\hat{\beta}_1) \bar{x} = E\left(\frac{\sum_{i=1}^n Y_i}{n}\right) - \beta_1 \bar{x} = \frac{1}{n} \sum_{i=1}^n E(Y_i) - \beta_1 \bar{x} = \\ &= \frac{1}{n} \sum_{i=1}^n (\beta_0 + \beta_1 x_i) - \beta_1 \bar{x} = \beta_0 + \beta_1 \bar{x} - \beta_1 \bar{x} = \beta_0 \end{aligned}$$

Calculamos la varianza de $\hat{\beta}_0$, para esto planteamos:

$$V(\hat{\beta}_0) = V(\bar{Y} - \hat{\beta}_1 \bar{x}) = V(\bar{Y}) + V(\hat{\beta}_1)(\bar{x})^2 - 2Cov(\bar{Y}, \hat{\beta}_1 \bar{x})$$

Tenemos que

$$V(\bar{Y}) = V\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(Y_i) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n}$$

Y

$$\begin{aligned} Cov(\bar{Y}, \hat{\beta}_1 \bar{x}) &= \bar{x} Cov(\bar{Y}, \hat{\beta}_1) = \bar{x} Cov\left(\frac{1}{n} \sum_{i=1}^n Y_i, \sum_{i=1}^n Y_i \frac{(x_i - \bar{x})}{S_{xx}}\right) = \bar{x} \frac{1}{n S_{xx}} \sum_{i=1}^n Cov(Y_i, Y_i (x_i - \bar{x})) = \\ &= \bar{x} \frac{1}{n S_{xx}} \sum_{i=1}^n (x_i - \bar{x}) Cov(Y_i, Y_i) = \bar{x} \frac{1}{n S_{xx}} \sum_{i=1}^n (x_i - \bar{x}) \sigma^2 = \bar{x} \frac{\sigma^2}{n S_{xx}} \sum_{i=1}^n (x_i - \bar{x}) = 0 \end{aligned}$$

$Cov(Y_i, Y_j) = 0$ por indep.

Por lo tanto

$$V(\hat{\beta}_0) = V(\bar{Y} - \hat{\beta}_1 \bar{x}) = V(\bar{Y}) + V(\hat{\beta}_1)(\bar{x})^2 - 2Cov(\bar{Y}, \hat{\beta}_1 \bar{x}) = \frac{\sigma^2}{n} + \bar{x}^2 \frac{\sigma^2}{S_{xx}} - 0 = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)$$

Entonces

$$\boxed{E(\hat{\beta}_0) = \beta_0 \quad \text{y} \quad V(\hat{\beta}_0) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)} \quad (10)$$

Necesitamos estimar la varianza desconocida σ^2 que aparece en las expresiones de $V(\hat{\beta}_0)$ y $V(\hat{\beta}_1)$.

Los residuos $e_i = y_i - \hat{y}_i$ se emplean para estimar σ^2 . La suma de los cuadrados de los residuos es

$$SS_R = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (11)$$

Puede demostrarse que $E\left(\frac{SS_R}{\sigma^2}\right) = n - 2$, en consecuencia $E\left(\frac{SS_R}{n - 2}\right) = \sigma^2$.

Entonces se toma como estimador de σ^2 a

$$\hat{\sigma}^2 = \frac{SS_R}{n - 2} \quad (12)$$

Puede obtenerse una fórmula más conveniente para el cálculo de SS_R , para esto primero notar que las ecuaciones normales (2) se pueden escribir como

$$\begin{cases} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \\ \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = 0 \end{cases} \Rightarrow \begin{cases} \sum_{i=1}^n e_i = 0 \\ \sum_{i=1}^n e_i x_i = 0 \end{cases}$$

Entonces

$$\begin{aligned} SS_R &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \hat{y}_i)(y_i - \hat{y}_i) = \sum_{i=1}^n e_i (y_i - \hat{y}_i) = \sum_{i=1}^n e_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = \\ &= \sum_{i=1}^n e_i (y_i - \hat{\beta}_0) - \sum_{i=1}^n \hat{\beta}_1 e_i x_i = \sum_{i=1}^n e_i (y_i - \hat{\beta}_0) = \sum_{i=1}^n e_i (y_i - \bar{y} - \hat{\beta}_1 \bar{x}) = \sum_{i=1}^n e_i (y_i - \bar{y}) \end{aligned}$$

Por lo tanto

$$\begin{aligned} SS_R &= \sum_{i=1}^n e_i (y_i - \bar{y}) = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)(y_i - \bar{y}) = \sum_{i=1}^n (y_i - \bar{y} + \hat{\beta}_1 \bar{x} - \hat{\beta}_1 x_i)(y_i - \bar{y}) = \\ &= \sum_{i=1}^n (y_i - \bar{y})(y_i - \bar{y}) - \sum_{i=1}^n \hat{\beta}_1 (x_i - \bar{x})(y_i - \bar{y}) = S_{yy} - \hat{\beta}_1 S_{xy} \end{aligned}$$

También se puede escribir

$$SS_R = S_{yy} - \hat{\beta}_1 S_{xy} = S_{yy} - \frac{S_{xy}}{S_{xx}} S_{xy} = S_{yy} - \frac{S_{xy}^2}{S_{xx}}$$

En resumen $SS_R = S_{yy} - \hat{\beta}_1 S_{xy} \quad \text{ó} \quad SS_R = S_{yy} - \frac{S_{xy}^2}{S_{xx}} \quad (13)$

Por lo tanto $\hat{\sigma}^2 = \frac{S_{yy} - \frac{S_{xy}^2}{S_{xx}}}{n - 2}$

Y si anotamos a la **desviación estándar estimada de $\hat{\beta}_0$ y $\hat{\beta}_1$** con $s_{\hat{\beta}_0}$ y $s_{\hat{\beta}_1}$ respectivamente entonces

$$s_{\hat{\beta}_1} = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \quad \text{y} \quad s_{\hat{\beta}_0} = \sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)} \quad (14)$$

Ejemplo:

En el ejemplo anterior se calculó, $\bar{x} = 1.9$, $\bar{y} = 5.3885$, $S_{xx} = 26.6$, $S_{xy} = 5.4430$.

Calculamos ahora $S_{yy} = \sum_{i=1}^{20} (y_i - \bar{y})^2 = 1.1733$ y entonces

$$\hat{\sigma}^2 = \frac{S_{yy} - \frac{S_{xy}^2}{S_{xx}}}{n-2} = \frac{1.1733 - \frac{5.4430^2}{26.6}}{18} = 0.003307$$

$$s_{\hat{\beta}_1} = \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} = \sqrt{\frac{0.003307}{26.6}} = \sqrt{0.000124} = 0.0111$$

$$s_{\hat{\beta}_0} = \sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)} = \sqrt{0.003307 \left(\frac{1}{20} + \frac{1.9^2}{26.6} \right)} = 0.02478219$$

Observación:

La varianza de $\hat{\beta}_0$ y $\hat{\beta}_1$ **se puede disminuir** tomando valores x_i muy dispersos con respecto a \bar{x} pues de esta forma aumenta S_{xx}

Para construir intervalos de confianza para los coeficientes de regresión o para construir pruebas de hipótesis con respecto a β_0 o β_1 necesitamos asumir que **los errores ε_i tienen distribución normal**. Entonces $\varepsilon_i \sim N(0, \sigma^2)$

Observación:

Si $\varepsilon_i \sim N(0, \sigma^2)$ entonces, como $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, resulta que $Y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$. **Se pueden calcular entonces los EMV de los parámetros y llegaríamos a que son los mismos que los encontrados usando mínimos cuadrados. De modo que la función que cumple la suposición de normalidad de los ε_i no es otra que la de justificar el uso del método de mínimos cuadrados, que es el más sencillo de calcular.**

Ya vimos que $\hat{\beta}_0$ y $\hat{\beta}_1$ pueden considerarse combinaciones lineales de las Y_i , por lo tanto $\hat{\beta}_0$ y $\hat{\beta}_1$ son combinación lineal de variables aleatorias independientes con distribución normal y eso implica que

$$\hat{\beta}_0 \sim N\left(\beta_0, \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)\right) \quad \text{y} \quad \hat{\beta}_1 \sim N\left(\beta_1, \frac{\sigma^2}{S_{xx}}\right) \quad (15)$$

Y entonces

$$\frac{\hat{\beta}_0 - \beta_0}{\sqrt{\sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)}} \sim N(0,1) \quad \text{y} \quad \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\sigma^2}{S_{xx}}}} \sim N(0,1) \quad (16)$$

Bajo la suposición que los errores tienen distribución normal, se puede probar que

$$\frac{SS_R}{\sigma^2} \sim \chi_{n-2}^2 \quad (17)$$

Y también se puede probar que

$$\frac{\hat{\beta}_0 - \beta_0}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)}} \sim t_{n-2} \quad \text{y} \quad \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}} \sim t_{n-2} \quad (18)$$

5.4 – Inferencias estadísticas sobre los parámetros de regresión

Suponemos que los errores tienen distribución normal, con media cero, varianza σ^2 y son independientes.

Inferencias sobre β_1

Tests de hipótesis sobre β_1

Se desea probar la hipótesis de que la pendiente β_1 es igual a una constante, por ejemplo β_{10} . Supongamos las hipótesis

$$H_0 : \beta_1 = \beta_{10} \quad \text{contra} \quad H_0 : \beta_1 \neq \beta_{10}$$

El estadístico de prueba es $T = \frac{\hat{\beta}_1 - \beta_{10}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$ que bajo H_0 tiene distribución Student con $n-2$

grados de libertad.

Por lo tanto la regla de decisión es
$$\begin{cases} \text{rechazar } H_0 & \text{si } |T| > t_{\frac{\alpha}{2}, n-2} \\ \text{aceptar } H_0 & \text{si } |T| \leq t_{\frac{\alpha}{2}, n-2} \end{cases}$$

Si $H_1 : \beta_1 > \beta_{10}$ se rechaza $H_0 : \beta_1 = \beta_{10}$ si $T > t_{\alpha, n-2}$

Si $H_1 : \beta_1 < \beta_{10}$ se rechaza $H_0 : \beta_1 = \beta_{10}$ si $T < -t_{\alpha, n-2}$

Un caso especial importante es cuando $H_0 : \beta_1 = 0$ contra $H_0 : \beta_1 \neq 0$

Estas hipótesis están relacionadas con la **significancia de la regresión**.

Aceptar $H_0 : \beta_1 = 0$ es equivalente a concluir que no hay ninguna relación lineal entre x e Y .

Si $H_0 : \beta_1 = 0$ se rechaza implica que x tiene importancia al explicar la variabilidad en Y .

También puede significar que el modelo lineal es adecuado, o que aunque existe efecto lineal pueden obtenerse mejores resultados agregando términos polinomiales de mayor grado en x .

Ejemplos:

1- El fabricante del resorte de los datos de la ley de Hooke afirma que la constante del resorte β_1 es al menos 0.23 pulg/lb. Se ha calculado que la constante del resorte es $\hat{\beta}_1 = 0.2046$ pulg/lb. ¿Se puede concluir que la afirmación del fabricante es falsa?

Solución:

Se requiere una prueba de hipótesis para contestar la pregunta. Las hipótesis serían

$$H_0 : \beta_1 = 0.23 \quad \text{contra} \quad H_0 : \beta_1 < 0.23$$

El estadístico de prueba es
$$T = \frac{\hat{\beta}_1 - \beta_{10}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}} = \frac{\hat{\beta}_1 - 0.23}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$$

Se calculó anteriormente $\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} = 0.0111$, entonces el valor t_0 que toma el estadístico es

$$t_0 = \frac{0.2046 - 0.23}{0.0111} = -2.28$$

Calculamos el p-valor recordando que bajo $H_0 : \beta_1 = 0.23$, $T \sim t_{n-2}$:

$$p\text{-valor} = P(T < -2.28)$$

Vemos en la tabla de la distribución Student que en la fila $\nu = 18$ grados de libertad

$$\begin{cases} P(T > 2.101) = 0.025 \\ P(T > 2.552) = 0.01 \end{cases} \Rightarrow 0.01 < p\text{-valor} < 0.025$$

Por lo tanto se rechaza $H_0 : \beta_1 = 0.23$

2- La capacidad de una unión soldada de elongarse bajo tensión está afectada por el compuesto químico del metal de soldadura. En un experimento para determinar el efecto del contenido de carbono (x) sobre la elongación (y) se alongaron 39 soldaduras hasta la fractura, y se midió tanto el contenido de carbono (en partes por mil) como la elongación (en %). Se calcularon los siguientes resúmenes estadísticos:

$$S_{xx} = 0.6561 \quad ; \quad S_{xy} = -3.9097 \quad ; \quad \hat{\sigma} = 4.3319$$

Suponiendo que x e y siguen un modelo lineal, calcular el cambio estimado en la elongación debido a un aumento de una parte por mil en el contenido de carbono. ¿Se debe utilizar el modelo lineal para pronosticar la elongación del contenido de carbono?

Solución:

El modelo lineal es $y = \beta_0 + \beta_1 x + \varepsilon$, y el cambio de elongación debido a un aumento de una parte por mil en el contenido de carbono es β_1 .

Las hipótesis serían $H_0 : \beta_1 = 0$ contra $H_0 : \beta_1 \neq 0$

La hipótesis nula establece que incrementar el contenido de carbono no afecta la elongación, mientras que la hipótesis alternativa establece que sí afecta la elongación.

El estadístico de prueba $|T| = \frac{\hat{\beta}_1 - \beta_{10}}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}} = \frac{\hat{\beta}_1}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}}$ si $H_0 : \beta_1 = 0$ es verdadera tiene distribu-

ción Student con $n - 2$ grados de libertad.

$$\text{Calculamos } \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^n y_i (x_i - \bar{x})}{S_{xx}} = \frac{-3.9097}{0.6561} = -5.959$$

$$\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} = \frac{\hat{\sigma}}{\sqrt{S_{xx}}} = \frac{4.3319}{\sqrt{0.6561}} = 5.348$$

$$\text{El valor que toma el estadístico de prueba es } t_0 = \frac{-5.959}{5.348} = 1.114$$

$$\text{Y } p\text{-valor} = P(|T| > 1.114) > 2 \times 0.10 = 0.20$$

Por lo tanto no hay evidencia en contra de la hipótesis nula. No se puede concluir que el modelo lineal sea útil para pronosticar la elongación a partir del contenido de carbono.

Intervalos de confianza para β_1

Podemos construir intervalos de confianza para β_1 de nivel $1 - \alpha$ utilizando el hecho que el

estadístico $\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}} \sim t_{n-2}$. El intervalo sería

$$\left[\hat{\beta}_1 - t_{\frac{\alpha}{2}, n-2} \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}; \hat{\beta}_1 + t_{\frac{\alpha}{2}, n-2} \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \right] \quad (19)$$

Ejemplo:

Determinar un intervalo de confianza de nivel 0.95 para la constante del resorte de los datos de la ley de Hooke.

Solución:

Se calculó antes $\hat{\beta}_1 = 0.2046$ y $\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} = 0.0111$

El número de grados de libertad es $20 - 2 = 18$, y $\alpha = 0.05$ por lo tanto

$$t_{\frac{\alpha}{2}, n-2} = t_{0.025, 18} = 2.101$$

Por lo tanto el intervalo es

$$[0.2046 - 2.101(0.0111), 0.2046 + 2.101(0.0111)] = [0.181; 0.228]$$

Inferencias sobre β_0

De manera similar a lo visto sobre β_1 , se pueden deducir intervalos de confianza y tests de hipótesis para β_0

Específicamente, si tenemos las hipótesis

$$H_0 : \beta_0 = \beta_{00} \quad \text{contra} \quad H_0 : \beta_0 \neq \beta_{00}$$

El estadístico de prueba es $T = \frac{\hat{\beta}_0 - \beta_{00}}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)}}$ y bajo $H_0 : \beta_0 = \beta_{00}$ tenemos que $T \sim t_{n-2}$

Por lo tanto la regla de decisión es
$$\begin{cases} \text{rechazar } H_0 & \text{si } |T| > t_{\frac{\alpha}{2}, n-2} \\ \text{aceptar } H_0 & \text{si } |T| \leq t_{\frac{\alpha}{2}, n-2} \end{cases}$$

Si $H_1 : \beta_0 > \beta_{00}$ se rechaza $H_0 : \beta_0 = \beta_{00}$ si $T > t_{\alpha, n-2}$

Si $H_1 : \beta_0 < \beta_{00}$ se rechaza $H_0 : \beta_0 = \beta_{00}$ si $T < -t_{\alpha, n-2}$

Intervalos de confianza de nivel $1 - \alpha$ se deducen de manera análoga a lo visto anteriormente,

donde usamos el hecho que el estadístico $T = \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)}} \sim t_{n-2}$

$$\text{El intervalo es } \left[\hat{\beta}_0 - t_{\frac{\alpha}{2}, n-2} \sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)}; \hat{\beta}_0 + t_{\frac{\alpha}{2}, n-2} \sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)} \right] \quad (20)$$

Ejemplo:

En los datos de la ley de Hooke determine un intervalo de confianza de nivel 0.99 para la longitud del resorte no cargado.

Solución:

La longitud del resorte no cargado es β_0 . Se ha calculado anteriormente $\hat{\beta}_0 = 4.9997$ y

$$s_{\hat{\beta}_0} = \sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)} = 0.02478219$$

El número de grados de libertad es $20 - 2 = 18$ y como $\alpha = 0.01$ entonces

$$t_{\frac{\alpha}{2}, n-2} = t_{0.005, 18} = 2.878$$

Por lo tanto el intervalo es

$$[4.9997 - 2.878(0.02478219), 4.9997 + 2.878(0.02478219)] = [4.9283; 5.071023]$$

5.5 – Intervalo de confianza para la respuesta media

A menudo es de interés *estimar* mediante un intervalo de confianza $\beta_0 + \beta_1 x_0$, es decir estimar la media $E(Y/x_0)$ para un valor específico x_0 .

Un estimador puntual razonable para $\beta_0 + \beta_1 x_0$ es $\hat{\beta}_0 + \hat{\beta}_1 x_0$.

Sabemos que $E(\hat{\beta}_0 + \hat{\beta}_1 x_0) = \beta_0 + \beta_1 x_0$.

Como de costumbre necesitamos construir un estadístico a partir de $\hat{\beta}_0 + \hat{\beta}_1 x_0$ que contenga al parámetro de interés, (en este caso $\beta_0 + \beta_1 x_0$) y del cual conozcamos la distribución de probabilidad.

Pensamos en el estadístico $\frac{\hat{\beta}_0 + \hat{\beta}_1 x_0 - E(\hat{\beta}_0 + \hat{\beta}_1 x_0)}{\sqrt{V(\hat{\beta}_0 + \hat{\beta}_1 x_0)}}$

Nos falta calcular $V(\hat{\beta}_0 + \hat{\beta}_1 x_0)$. Para esto nuevamente observamos que $\hat{\beta}_0 + \hat{\beta}_1 x_0$ es una combinación lineal de las variables Y_i

$$\begin{aligned} \hat{\beta}_0 + \hat{\beta}_1 x_0 &= \bar{Y} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 x_0 = \frac{1}{n} \sum_{i=1}^n Y_i + \hat{\beta}_1 (x_0 - \bar{x}) = \frac{1}{n} \sum_{i=1}^n Y_i + \frac{S_{xy}}{S_{xx}} (x_0 - \bar{x}) = \\ &= \frac{1}{n} \sum_{i=1}^n Y_i + \frac{\sum_{i=1}^n Y_i (x_i - \bar{x})}{S_{xx}} (x_0 - \bar{x}) = \sum_{i=1}^n Y_i \left[\frac{1}{n} + \frac{(x_i - \bar{x})}{S_{xx}} (x_0 - \bar{x}) \right] \end{aligned}$$

Por lo tanto:

$$\begin{aligned} V(\hat{\beta}_0 + \hat{\beta}_1 x_0) &= V\left(\sum_{i=1}^n Y_i \left[\frac{1}{n} + \frac{(x_i - \bar{x})}{S_{xx}} (x_0 - \bar{x}) \right] \right) = \sum_{i=1}^n V(Y_i) \left[\frac{1}{n} + \frac{(x_i - \bar{x})}{S_{xx}} (x_0 - \bar{x}) \right]^2 = \\ &= \sum_{i=1}^n \sigma^2 \left[\frac{1}{n} + \frac{(x_i - \bar{x})}{S_{xx}} (x_0 - \bar{x}) \right]^2 = \sum_{i=1}^n \sigma^2 \left[\frac{1}{n^2} + \frac{(x_i - \bar{x})^2}{S_{xx}^2} (x_0 - \bar{x})^2 + 2 \frac{(x_i - \bar{x})}{n S_{xx}} (x_0 - \bar{x}) \right] = \end{aligned}$$

$$= \sigma^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \sum_{i=1}^n (x_i - \bar{x})^2 + 2 \frac{(x_0 - \bar{x})}{n S_{xx}} \sum_{i=1}^n (x_i - \bar{x}) \right] =$$

Notar que $\sum_{i=1}^n (x_i - \bar{x}) = 0$ y $\sum_{i=1}^n (x_i - \bar{x})^2 = S_{xx}$ entonces

$$= \sigma^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]$$

Por lo tanto

$$\hat{\beta}_0 + \hat{\beta}_1 x_0 \sim N \left(\beta_0 + \beta_1 x_0; \sigma^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right] \right) \quad (21)$$

Como σ^2 es desconocido lo reemplazamos por $\hat{\sigma}^2 = \frac{SS_R}{n-2}$, y puede probarse que

$\frac{\hat{\beta}_0 + \hat{\beta}_1 x_0 - (\beta_0 + \beta_1 x_0)}{\sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]}}$ tiene distribución Student con $n-2$ grados de libertad

Razonando como en casos anteriores, el intervalo de confianza para $\beta_0 + \beta_1 x_0$ de nivel $1-\alpha$ es

$$\left[\hat{\beta}_0 + \hat{\beta}_1 x_0 - t_{\frac{\alpha}{2}, n-2} \sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]}; \hat{\beta}_0 + \hat{\beta}_1 x_0 + t_{\frac{\alpha}{2}, n-2} \sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]} \right] \quad (22)$$

Ejemplo:

Mediante los datos de la ley de Hooke calcular un intervalo de confianza de nivel 0.95 para la longitud media de un resorte bajo una carga de 1.4 lb

Solución:

Para aplicar (10.22) necesitamos calcular $\hat{\beta}_0 + \hat{\beta}_1 x_0$; $\hat{\sigma}^2$; \bar{x} ; S_{xx} .

En este caso $x_0 = 1.4$ y $\alpha = 0.05$, por lo tanto $t_{\frac{\alpha}{2}, n-2} = t_{0.025, 18} = 2.101$

Ya tenemos calculado de ejemplos anteriores:

$$\hat{\sigma} = 0.0575$$

$$\bar{x} = 1.9$$

$$S_{xx} = 26.6$$

$$\hat{\beta}_0 = 4.9997 \text{ y } \hat{\beta}_1 = 0.2046$$

$$\text{De aquí ya calculamos } \hat{\beta}_0 + \hat{\beta}_1 x_0 = 4.9997 + 0.2046 \times 1.4 = 5.286$$

Entonces el intervalo es:

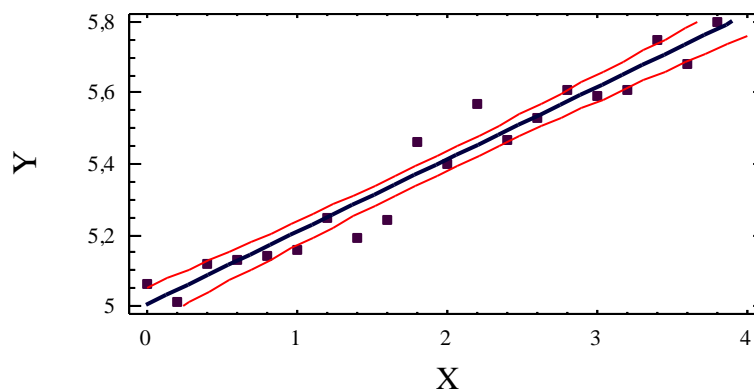
$$\left[5.286 - 2.101 \sqrt{0.0575^2 \left[\frac{1}{20} + \frac{(1.4 - 1.9)^2}{26.6} \right]}; 5.286 + 2.101 \sqrt{0.0575^2 \left[\frac{1}{20} + \frac{(1.4 - 1.9)^2}{26.6} \right]} \right] = [5.26; 5.32]$$

Observaciones:

1- Notar que el ancho del intervalo de confianza para $E(Y/x_0)$ depende del valor de x_0 . El ancho del intervalo es mínimo cuando $x_0 = \bar{x}$ y crece a medida que $|x_0 - \bar{x}|$ aumenta.

2- Al repetir los cálculos anteriores para varios valores diferentes de x_0 pueden obtenerse intervalos de confianza para cada valor correspondiente de $E(Y/x_0)$.

En la figura siguiente se presenta el diagrama de dispersión con la recta estimada y los correspondientes intervalos de confianza de nivel 0.95 graficados con las líneas inferior y superior referidos al ejemplo anterior. Se origina entonces una **banda de confianza** que envuelve a la recta estimada.



5.6 – Intervalos de predicción para futuras observaciones

Una aplicación importante de un modelo de regresión es la predicción de observaciones nuevas o futuras de Y , correspondientes a un nivel especificado de la variable x .

Si x_0 es el valor de x de interés, entonces una **estimación puntual** de la observación

$$Y_0 = \beta_0 + \beta_1 x_0 + \varepsilon_0 \text{ es } \hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0.$$

Para hallar un intervalo de predicción para $Y_0 = \beta_0 + \beta_1 x_0$ de nivel $1 - \alpha$ debemos construir un estadístico a partir de $\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$.

Primero notamos que si Y_0 es una nueva observación, entonces Y_0 es **independiente de las observaciones utilizadas para desarrollar el modelo de regresión**.

Consideramos $Y_0 - \hat{Y}_0$. Calculamos su esperanza y varianza:

$$E(Y_0 - \hat{Y}_0) = E(\beta_0 + \beta_1 x_0 + \varepsilon_0 - (\hat{\beta}_0 + \hat{\beta}_1 x_0)) = \beta_0 + \beta_1 x_0 + E(\varepsilon_0) - (\beta_0 + \beta_1 x_0) = 0$$

$$\begin{aligned} V(Y_0 - \hat{Y}_0) &= V(Y_0) + V(\hat{Y}_0) = V(\beta_0 + \beta_1 x_0 + \varepsilon_0) + V(\hat{\beta}_0 + \hat{\beta}_1 x_0) = \sigma^2 + \sigma^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right] = \\ &= \sigma^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right] \end{aligned}$$

Por lo tanto

$$Y_0 - \hat{Y}_0 \sim N\left(0; \sigma^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right] \right) \quad (23)$$

En consecuencia

$$\frac{Y_0 - \hat{Y}_0}{\sqrt{\sigma^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]}} \sim N(0; 1) \quad (24)$$

Si reemplazamos σ^2 por su estimación $\hat{\sigma}^2$ se puede probar que

$$\frac{Y_0 - \hat{Y}_0}{\sqrt{\hat{\sigma}^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]}} \sim t_{n-2} \quad (25)$$

Por el argumento usual llegamos al siguiente **intervalo de predicción de nivel $1 - \alpha$ para Y_0** :

$$\left[\hat{Y}_0 - t_{\frac{\alpha}{2}, n-2} \hat{\sigma}^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]; \hat{Y}_0 + t_{\frac{\alpha}{2}, n-2} \hat{\sigma}^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right] \right] \quad (26)$$

Ejemplo:

Calcular el intervalo de predicción con nivel 0.95 para la elongación de un resorte bajo una carga de 1.4 lb.

Solución:

El intervalo es

$$\left[5.286 - 2.101 \sqrt{0.0575^2 \left[1 + \frac{1}{20} + \frac{(1.4 - 1.9)^2}{26.6} \right]}; 5.286 + 2.101 \sqrt{0.0575^2 \left[1 + \frac{1}{20} + \frac{(1.4 - 1.9)^2}{26.6} \right]} \right] = [5.16165; 5.41034]$$

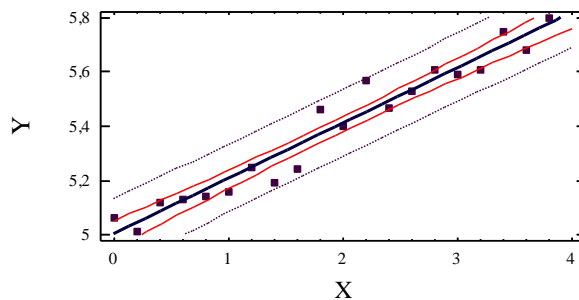
Observaciones:

1- Un **intervalo de confianza** es un intervalo que contiene, con un nivel de confianza fijado, un parámetro determinado de interés. Un **intervalo de predicción** es un intervalo que contiene, con un nivel de confianza fijado, una variable aleatoria de interés.

2- El ancho del intervalo de predicción es mínimo cuando $x_0 = \bar{x}$ y crece a medida que $|x_0 - \bar{x}|$ aumenta.

Al comparar (10.26) con (10.22) se observa que el intervalo de predicción en el punto x_0 siempre es más grande que el intervalo de confianza en x_0 . Esto se debe a que el intervalo de predicción depende tanto del error del modelo ajustado como del error asociado con las observaciones futuras.

3- Al repetir los cálculos anteriores para varios valores diferentes de x_0 pueden obtenerse los intervalos de predicción. En la figura siguiente se presenta el diagrama de dispersión con la recta estimada y los correspondientes intervalos de confianza y de predicción de nivel 0.95 graficados con las líneas inferior y superior referidos al ejemplo anterior. Se originan entonces una **banda de confianza** (línea continua) y otra **banda de predicción** (línea entrecortada) que envuelven a la recta estimada. Esto ilustra que los intervalos de confianza son menos amplios que los intervalos de predicción.



5.7

– Índice de ajuste

Si consideramos el ajuste por mínimos cuadrados de los pares de datos (x_i, Y_i) al modelo

$$Y = \beta_0 + \varepsilon$$

Entonces es fácil verificar que el estimador de mínimos cuadrados de β_0 es \bar{Y} , y la suma de

residuos al cuadrado es $S_{YY} = \sum_{i=1}^n (Y_i - \bar{Y})^2$. Por otro lado si consideramos el modelo lineal

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

Entonces tenemos un valor de $SS_R = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ que será menor o igual a $S_{YY} = \sum_{i=1}^n (Y_i - \bar{Y})^2$.

La cantidad R^2 se define como

$$R^2 = 1 - \frac{SS_R}{S_{YY}} \quad (27)$$

y es llamado **coeficiente de determinación**. Vemos que R^2 será cero si $\beta_1 = 0$ y será uno si

$SS_R = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = 0$, lo que significa ajuste lineal perfecto.

En general $0 \leq R^2 \leq 1$. El valor de R^2 se interpreta como la proporción de variación de la respuesta Y que es explicada por el modelo. La cantidad $\sqrt{R^2}$ es llamada **índice de ajuste**, y es a menudo usada como un indicador de qué tan bien el modelo de regresión ajusta los datos. Pero un valor alto de R no significa necesariamente que el modelo de regresión sea correcto.

Ejemplo:

En el ejemplo de la ley de Hooke, tenemos $S_{yy} = \sum_{i=1}^{20} (y_i - \bar{y})^2 = 1.1733$, y

$$SS_R = 1.1733 - \frac{5.4430^2}{26.6} = 0.059526$$

$$\text{Por lo tanto } R^2 = 1 - \frac{SS_R}{S_{YY}} = 1 - \frac{0.059526}{1.1733} = 0.949266$$

El índice de ajuste R es a menudo llamado **coeficiente de correlación muestral**. Si la variable fijada x es una **variable aleatoria**, entonces tendríamos una v.a. bidimensional (X, Y) con una distribución de probabilidad conjunta, y tenemos una muestra de pares (X_i, Y_i) $i = 1, \dots, n$. Supongamos que estamos interesados en estimar ρ el coeficiente de correlación entre X e Y . Es decir

$$\rho = \frac{E[(X - E(X))(Y - E(Y))]}{\sqrt{V(X)V(Y)}}$$

Es razonable estimar

$$E[(X - E(X))(Y - E(Y))] \quad \text{con} \quad \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

$$V(X) \quad \text{con} \quad \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \text{y} \quad V(Y) \quad \text{con} \quad \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$$

Por lo tanto un estimador natural de ρ es

$$\hat{\rho} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} = \frac{S_{XY}}{\sqrt{S_{XX} S_{YY}}} = R \quad (28)$$

Es decir *el índice de ajuste estima la correlación entre X e Y*

Si X es una variable aleatoria, entonces se observan pares independientes (X_i, Y_i) con $i = 1, \dots, n$ que cumplen el modelo

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Si asumimos que X_i y ε_i son independientes y que las ε_i tienen todas la misma distribución con $E(\varepsilon_i) = 0$, entonces $E(Y_i/X_i) = \beta_0 + \beta_1 X_i$

Si además suponemos que $\varepsilon_i \sim N(0, \sigma^2)$ entonces se puede probar que los estimadores de máxima verosimilitud para los parámetros β_0 y β_1 son

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

y

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n Y_i (X_i - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \frac{S_{XY}}{S_{XX}}$$

Es decir son los mismos estimadores a los dados por el método de mínimos cuadrados en el caso de suponer que X es una variable matemática.

También se puede probar que bajo las suposiciones hechas (17) y (18) siguen siendo válidas. Las distribuciones de los estimadores dependen ahora de las distribuciones de las X_i . Puede probarse que siguen siendo insesgados, y que su distribución condicional en las X_i es normal, pero en general su distribución no será normal.

5.8 – Análisis de residuos

El ajuste de un modelo de regresión requiere varias suposiciones. La estimación de los parámetros del modelo requiere la suposición de que los errores son variables aleatorias independientes con media cero y varianza constante. Las pruebas de hipótesis y la estimación de intervalos requieren que los errores estén distribuidos de manera normal. Además se supone que el grado del modelo es correcto, es decir, si se ajusta un modelo de regresión lineal simple, entonces se supone que el fenómeno en realidad se comporta de una manera lineal.

Se debe considerar la validez de estas suposiciones como dudosas y examinar cuán adecuado es el modelo que se propone. A continuación se estudian métodos que son útiles para este propósito.

Los residuos de un modelo de regresión son $e_i = y_i - \hat{y}_i$ $i = 1, 2, \dots, n$. A menudo el análisis de los residuos es útil para verificar la hipótesis de que los errores tienen una distribución que es aproximadamente normal con varianza constante, y también para determinar la utilidad que tiene la adición de más términos al modelo.

Es posible **estandarizar** los residuos mediante el cálculo de $\frac{e_i}{\sqrt{\hat{\sigma}^2}}$ $i = 1, 2, \dots, n$.

También se puede probar que la varianza del i -ésimo residuo e_i es igual a

$$V(e_i) = \sigma^2 \left[1 - \left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right) \right]$$

Y entonces podemos considerar al i -ésimo **residuo estudentizado** que se define como

$$r_i = \frac{e_i}{\sqrt{\hat{\sigma}^2 \left[1 - \left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right) \right]}}$$

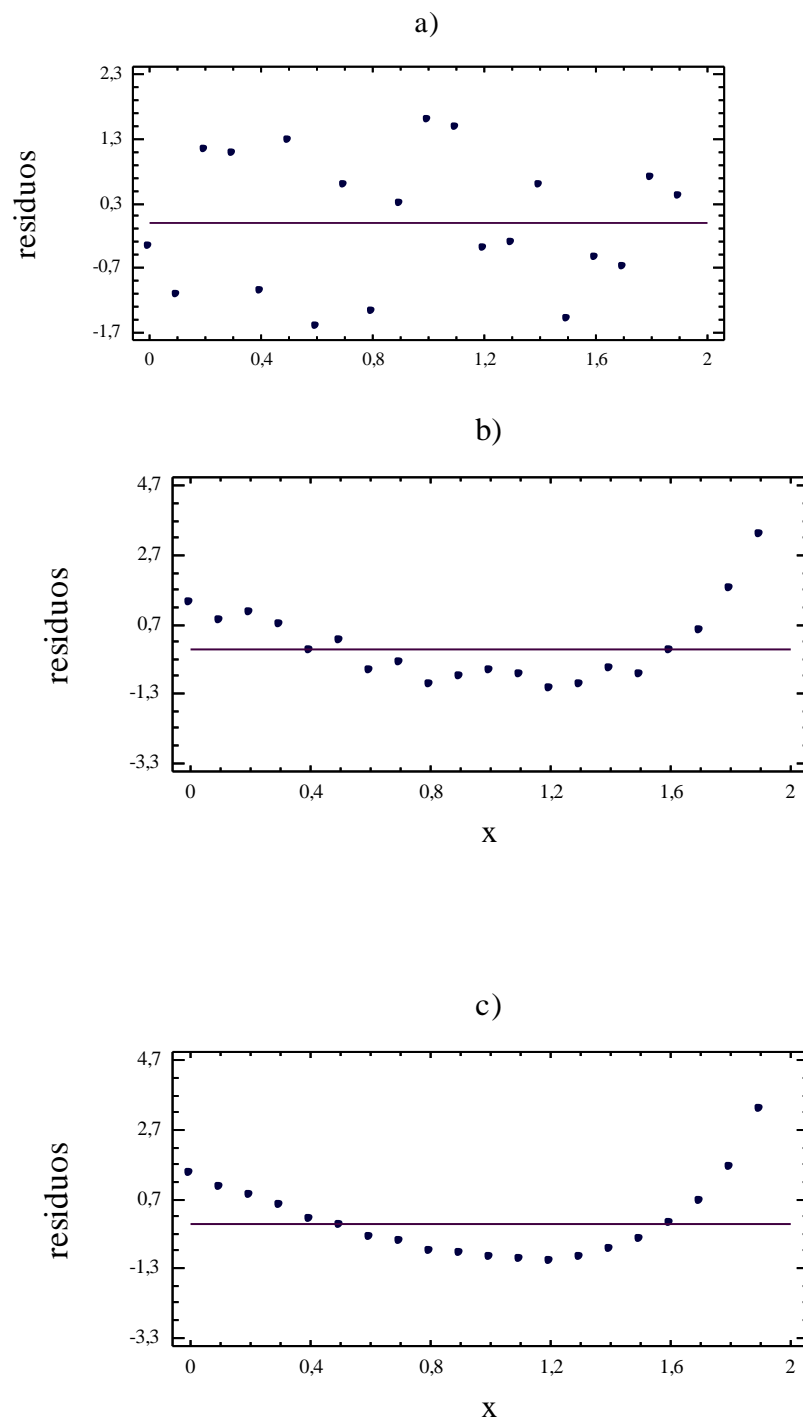
y tiene desviación estándar unitaria.

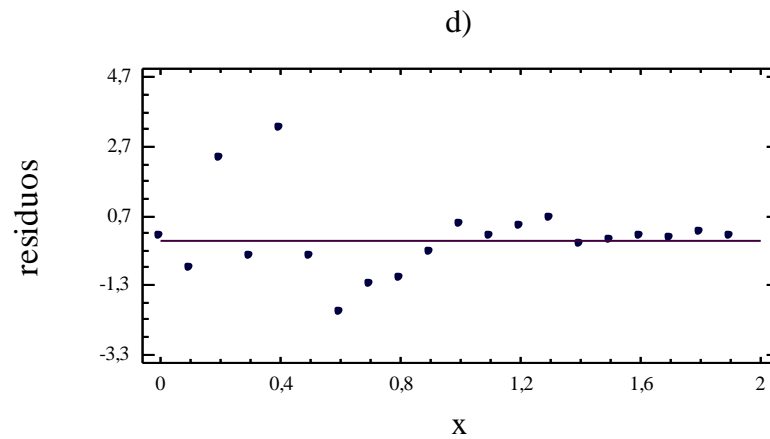
Si los errores tienen una distribución normal, entonces aproximadamente el 95% de los residuos estandarizados deben caer en el intervalo $(-2; 2)$. Los residuos que se alejan mucho de este intervalo pueden indicar la presencia de un **valor atípico**, es decir, una observación que no es común con respecto a los demás datos.

A menudo es útil hacer una gráfica de residuos contra la variable independiente x . En este caso la gráfica tendría que ser una nube de puntos sin ningún patrón en el intervalo $(-2; 2)$; pues $e_i = y_i - \hat{y}_i$ sería lo que queda de y_i al quitarle la influencia de x_i . Si en la gráfica aparece algún patrón quiere decir que no estamos quitando de las y toda la influencia de las x .

Patrones usuales para las gráficas de residuos suelen ser los de las siguientes figuras: en la figura a) se representa la situación ideal, una nube de puntos sin ningún patrón en el intervalo $(-2; 2)$. Las figuras b), c) y d) representan anomalías. Si los residuos aparecen como en b) o c) indican que el modelo es inadecuado. La figura d) muestra residuos que indican que la varianza de las observaciones varía con la magnitud de x . Comúnmente se utiliza una transformación de datos sobre la respuesta y para eliminar este problema. Las transformaciones más utilizadas para estabilizar la varianza son \sqrt{y} , $\ln(y)$ o $1/y$.

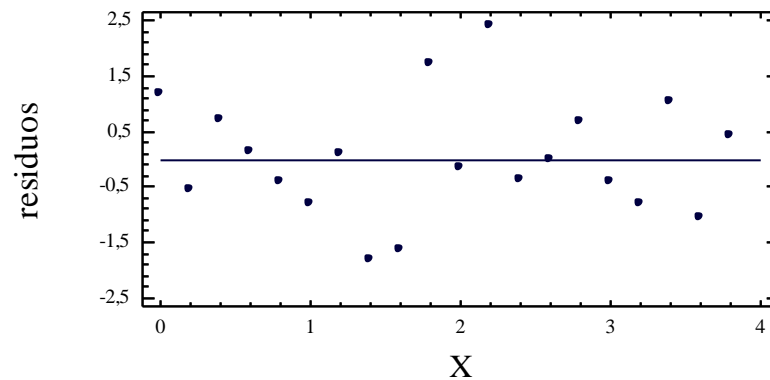
En la figura d) la varianza de las observaciones disminuye con el aumento de x





Ejemplo:

Para los datos sobre la ley de Hooke la gráfica de residuos es



Para el caso en que (X, Y) es una v.a. bidimensional, no siempre se está interesado en la relación lineal que defina $E(Y/X)$. Si no, únicamente saber si X e Y son variables aleatorias independientes. Si asumimos que la distribución conjunta de (X, Y) es una distribución llamada **normal bivariada**, entonces **probar que $\rho = 0$ es equivalente a probar que X e Y son independientes**.

Se puede probar que si la distribución conjunta de (X, Y) es **normal bivariada**, entonces R es el estimador de máxima verosimilitud de ρ . Pero es difícil obtener la distribución de probabilidad para R . Se puede superar esta dificultad en muestras bastante grandes al utilizar el hecho que el estadístico $\left(\frac{1}{2}\right)\ln\left(\frac{1+R}{1-R}\right)$ tiene aproximadamente una distribución normal con media

$$\mu = \left(\frac{1}{2}\right)\ln\left(\frac{1+\rho}{1-\rho}\right) \text{ y varianza } \sigma^2 = \frac{1}{n-3}.$$

Por lo tanto para probar la hipótesis $H_0 : \rho = \rho_0$ podemos utilizar el estadístico de prueba

$$Z = \frac{\left(\frac{1}{2}\right)\ln\left(\frac{1+R}{1-R}\right) - \left(\frac{1}{2}\right)\ln\left(\frac{1+\rho_0}{1-\rho_0}\right)}{\frac{1}{\sqrt{n-3}}} \quad (29)$$

Para construir intervalos de confianza de nivel $1 - \alpha$ para ρ , se despeja en $\mu = \left(\frac{1}{2}\right)\ln\left(\frac{1+\rho}{1-\rho}\right)$ el coeficiente ρ y se llega a

$$\rho = \frac{e^{2\mu} - 1}{e^{2\mu} + 1} \quad (30)$$

Ejemplo:

En un estudio de los tiempos de reacción, el tiempo de respuesta a un estímulo visual (x) y el tiempo de respuesta a un estímulo auditivo (y) se registraron para cada una de 10 personas. Los tiempos se midieron en minutos. Se presentan en la siguiente tabla.

x	161	203	235	176	201	188	228	211	191	178
y	159	206	241	163	197	193	209	189	169	201

- Determinar un intervalo de confianza de nivel 0.95 para la correlación entre los tiempos de reacción.
- Determinar el p-valor para $H_0 : \rho = 0.3$ contra $H_1 : \rho > 0.3$

Solución:

$$a) \text{ Se calcula } R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} = \frac{S_{XY}}{\sqrt{S_{XX} S_{YY}}} = 0.8159$$

$$\text{Luego calcula } \left(\frac{1}{2}\right)\ln\left(\frac{1+R}{1-R}\right) = \left(\frac{1}{2}\right)\ln\left(\frac{1+0.8159}{1-0.8159}\right) = 1.1444$$

Como $\left(\frac{1}{2}\right)\ln\left(\frac{1+R}{1-R}\right)$ está distribuido normalmente con varianza $\sigma^2 = \frac{1}{n-3}$, el intervalo para

$$\mu = \left(\frac{1}{2}\right)\ln\left(\frac{1+\rho}{1-\rho}\right) \text{ es}$$

$$\left[1.1444 - 1.96\left(\frac{1}{\sqrt{10-3}}\right); 1.1444 + 1.96\left(\frac{1}{\sqrt{10-3}}\right) \right] = [0.4036; 1.8852]$$

Para hallar el intervalo para ρ aplicamos la transformación (10.30) y se obtiene

$$\frac{e^{2(0.4036)} - 1}{e^{2(0.4036)} + 1} < \rho < \frac{e^{2(1.8852)} - 1}{e^{2(1.8852)} + 1} \Rightarrow 0.383 < \rho < 0.955$$

b) Si $H_0 : \rho = 0.3$ es verdadera entonces el estadístico

$$Z = \frac{\left(\frac{1}{2}\right)\ln\left(\frac{1+R}{1-R}\right) - \left(\frac{1}{2}\right)\ln\left(\frac{1+0.3}{1-0.3}\right)}{\frac{1}{\sqrt{10-3}}} \quad \text{tiene aproximadamente distribución } N(0,1)$$

El valor observado de $\left(\frac{1}{2}\right)\ln\left(\frac{1+R}{1-R}\right)$ es 1.1444, por lo tanto el estadístico toma el valor $z_0 = 2.2088$

Entonces $p\text{-valor} = P(Z > 2.2088) \approx 0.0136$. Entonces se rechaza $H_0 : \rho = 0.3$ y se concluye que $\rho > 0.3$

Práctica**Regresión lineal simple**

- 1) Se utiliza regresión lineal para analizar los datos de un estudio donde se investigó la relación que existe entre la temperatura de la superficie de una carretera (x) y la deformación del pavimento (y). El resumen de cantidades es el siguiente:

$$n = 20, \sum y_i = 12.75, \sum y_i^2 = 8.86, \sum x_i = 1478, \sum x_i^2 = 143215.8, \sum x_i y_i = 1083.67$$

- a) Calcular las estimaciones de mínimos cuadrados de la pendiente y la ordenada al origen. Hacer un gráfico de la recta de regresión, y estimar σ^2 .
- b) Utilice la ecuación de la recta ajustada para predecir la deformación del pavimento observada cuando la temperatura de la superficie sea 85°F.

- 2) Se tiene la siguiente información sobre la relación entre una medida de la corrosión del hierro (Y) y la concentración de NaPO_4 (X , en ppm)

x	2.50	5.03	7.60	11.60	13.00	19.60	26.20	33.00	40.00	50.00	55.00
y	7.68	6.95	6.30	5.75	5.01	1.43	0.93	0.72	0.68	0.65	0.56

- a) Construya un gráfico de dispersión de los datos. ¿Parece ser razonable el modelo de regresión lineal?
- b) Calcule la ecuación de la recta de regresión estimada, utilícela para pronosticar el valor de la rapidez de corrosión que se observaría para una concentración de 33 ppm, y calcule el residuo correspondiente.
- c) Estime la desviación estándar de observaciones alrededor de la recta de regresión verdadera.
- d) ¿Se puede concluir que el modelo de regresión lineal simple especifica una relación útil entre las dos variables?. Establezca y pruebe las hipótesis adecuadas al nivel de significación de 0.05.

- 3) En pruebas diseñadas para medir el efecto de cierto aditivo en el tiempo de secado de pintura se obtuvieron los siguientes datos

Concentración de aditivo (%)	4.0	4.2	4.4	4.6	4.8	5.0	5.2	5.4	5.6	5.8
Tiempo de secado (horas)	8.7	8.8	8.3	8.7	8.1	8.0	8.1	7.7	7.5	7.2

Se tiene el siguiente resumen estadístico:

$$\bar{x} = 4.9 \quad \bar{y} = 8.11 \quad S_{xx} = 3.29967 \quad S_{xy} = -2.75 \quad S_{yy} = 2.58894$$

- a) Realizar un gráfico de dispersión y estimar la recta de regresión lineal
- b) Estimar la varianza
- c) Pronostique el tiempo de secado para una concentración de 4.4%
- d) ¿Puede utilizarse la recta de mínimos cuadrados para pronosticar el tiempo de secado respecto a una concentración de 7%?
- e) ¿Para qué concentración pronosticaría un tiempo de secado de 8.2 horas?

- 4) Un químico está calibrando un espectrómetro que se utilizará para medir la concentración de monóxido de carbono en muestras atmosféricas. Para comprobar la calibración, se miden muestras de concentración conocida.

Las concentraciones verdaderas (x) y las medidas (y) están dadas en la tabla siguiente:

x (ppm)	0	10	20	30	40	50	60	70	80	90	100
Y (ppm)	1	11	21	28	37	48	56	68	75	86	96

Para comprobar la calibración se ajusta un modelo lineal $y = \beta_0 + \beta_1 x + \varepsilon$. Idealmente, el valor de β_0 debe ser 0 y el valor de β_1 debe ser 1.

- Calcule los estimadores de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$
 - ¿Se puede rechazar la hipótesis nula $H_0: \beta_0 = 0$? Utilice $\alpha = 0.05$
 - ¿Se puede rechazar la hipótesis nula $H_0: \beta_1 = 1$? Utilice $\alpha = 0.05$
 - ¿Los datos proporcionan suficiente evidencia para concluir que la máquina está fuera de calibración?
- 5) En un experimento para investigar la relación entre el diámetro de un clavo (x) y su fuerza retirada final (y), se colocaron clavos de forma anular enhebrados en madera de abeto de Douglas, y después se midieron sus fuerzas de retirada en N/mm. Se obtuvieron los resultados siguientes para 10 diámetros diferentes (en mm):

x	2.52	2.87	3.05	3.43	3.68	3.76	3.76	4.50	4.50	5.26
y	54.74	59.01	72.92	50.85	54.99	60.56	69.08	77.03	69.97	90.70

- Calcule la recta de mínimos cuadrados para predecir la fuerza a partir del diámetro
 - Determine el intervalo de confianza de nivel 0.95 para la media de la fuerza de retirada de clavos de 4 mm de diámetro
 - Determine el intervalo de predicción de nivel 0.95 para la media de la fuerza de retirada de clavos de 4 mm de diámetro
 - ¿Puede concluir que la media de la fuerza de retirada de clavos de 4 mm de diámetro es 60 N/mm con un nivel de significancia de 0.05?
- 6) Un comerciante realizó un estudio para determinar la relación que hay entre los gastos de la publicidad semanal y las ventas. Registró los datos siguientes:
- Costos de publicidad (en \$):* 40, 20, 25, 20, 30, 50, 40, 20, 50, 40, 25, 50.
Ventas (en \$): 385, 400, 395, 365, 475, 440, 490, 420, 560, 525, 480, 510.
- Haga un gráfico de dispersión
 - Encuentre la recta de regresión estimada para pronosticar las ventas semanales, a partir de los gastos de publicidad.
 - Estime las ventas semanales cuando los costos de la publicidad sean 35\$. ¿Es válido estimar las ventas semanales cuando los costos de la publicidad sean 75\$?
 - Pruebe la hipótesis de que $\beta_1 = 6$ contra la alternativa de que $\beta_1 < 6$, utilice $\alpha = 0.025$
 - Construya un intervalo de confianza de 95% para la media de las ventas semanales cuando se gastan 45\$ en publicidad.
 - Construya un intervalo de predicción de 95% para la media de las ventas semanales cuando se gastan 45\$ en publicidad.
 - ¿Qué proporción de la variabilidad total en las ventas está explicada por el costo en publicidad?
- 7) Considere los siguientes datos sobre el número de millas para ciertos automóviles, en millas por galón (mpg) y su peso en libras (wt)

Modelo	GMC	Geo	Honda	Hyundai	Infiniti	Isuzu	Jeep	Land	Lexus	Linclon
Wt (x)	4520	2065	2440	2290	3195	3480	4090	4535	3390	3930
Mpg (y)	15	29	31	28	23	21	15	13	22	18

- Estime la recta de regresión lineal.
 - Estime las millas para un vehículo que pesa 4000 libras.
 - Suponga que los ingenieros de Honda afirman que, en promedio, el Civic (o cualquier otro modelo de vehículo que pese 2440 libras) recorre mas de 30 mpg. Con base en los resultados del análisis de regresión, ¿es esta afirmación creíble?, ¿por qué?
Sugerencia: calcule el intervalo de confianza para la media mpg cuando el peso es de 2440 libras, con $\alpha = 0.05$
 - Los ingenieros de diseño para el Lexus ES300 tienen por objetivo lograr 18 mpg como ideal para dicho modelo (o cualquier otro que pese 3390 libras), aunque se espera que haya cierta variación. ¿Es probable que sea realista ese objetivo?
Sugerencia: calcule el intervalo de predicción para las mpg cuando el peso es de 3390 libras, con $\alpha = 0.05$
 - ¿Qué proporción de la variabilidad total en el millaje está explicada por el peso del motor?
- 8) Los valores siguientes son 26 lecturas sobre la congestión del transito y la concentración de monóxido de carbono efectuadas en un sitio de muestreo para determinar la calidad del aire de cierta ciudad. Enunciando las suposiciones necesarias , resolver los siguientes incisos :
- Prepare un diagrama de dispersión.
 - Calcule el coeficiente de correlación de la muestra.
 - Pruebe que $H_0 : \rho=0$ al nivel de significación de 0.05 y saque sus conclusiones.
 - Construya el intervalo de confianza del 95% para ρ .
 - ¿Hay suficiente evidencia de una correlación mayor que 0.95 ?. Utilice $\alpha=0.05$.

congestion de transito (automoviles por hora), X	CO (ppm), Y	congestion de transito (automoviles por hora), X	CO (ppm), Y
100	8.8	375	13.2
110	9.0	400	14.5
125	9.5	425	14.7
150	10.0	450	14.9
175	10.5	460	15.1
190	10.5	475	15.5
200	10.5	500	16.0
225	10.6	525	16.3
250	11.0	550	16.8
275	12.1	575	17.3
300	12.1	595	18.0
325	12.5	600	18.4

$$P[Z > Z_\alpha] = \int_{Z_\alpha}^{\infty} \frac{1}{\sqrt{2\pi}} \exp^{-z^2/2} dz = \alpha$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2910	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1597	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0995
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0394	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.00798	0.00776	0.00755	0.00734	0.00714	0.00695	0.00676	0.00657	0.00639
2.5	0.00621	0.00604	0.00587	0.00570	0.00554	0.00539	0.00523	0.00508	0.00494	0.00480
2.6	0.00466	0.00453	0.00440	0.00427	0.00415	0.00402	0.00391	0.00379	0.00368	0.00357
2.7	0.00347	0.00336	0.00326	0.00317	0.00307	0.00298	0.00289	0.00280	0.00272	0.00264
2.8	0.00256	0.00248	0.00240	0.00233	0.00226	0.00219	0.00212	0.00205	0.00199	0.00193
2.9	0.00187	0.00181	0.00175	0.00169	0.00164	0.00159	0.00154	0.00149	0.00144	0.00139
3.0	0.00135	0.00131	0.00126	0.00122	0.00118	0.00114	0.00111	0.00107	0.00103	0.00100
3.1	0.00097	0.00094	0.00091	0.00087	0.00084	0.00082	0.00079	0.00076	0.00074	0.00071
3.2	0.00069	0.00066	0.00064	0.00062	0.0006	0.00058	0.00056	0.00054	0.00052	0.00050
3.3	0.00048	0.00047	0.00045	0.00043	0.00042	0.00041	0.00032	0.00038	0.00036	0.00035
3.4	0.00034	0.00032	0.00031	0.0003	0.00029	0.00028	0.00027	0.00026	0.00025	0.00024
3.5	0.00023	0.00022	0.00022	0.00021	0.00020	0.00019	0.00019	0.00018	0.00017	0.00017
3.6	0.00016	0.00015	0.00015	0.00014	0.00014	0.00013	0.00013	0.00012	0.00012	0.00011
3.7	0.00011	0.00010	0.00010	0.00010	0.00009	0.00009	0.00009	0.00008	0.00008	0.00008
3.8	0.00007	0.00007	0.00007	0.00006	0.00006	0.00006	0.00006	0.00005	0.00005	0.00005
3.9	0.00005	0.00005	0.00004	0.00004	0.00004	0.00004	0.00004	0.00004	0.00003	0.00003

Tabla IV. Valores críticos de la distribución t de Student:
Abcisas $t_{\alpha;\nu}$ que dejan a su derecha un área α en una t con ν grados de libertad.

ν	0.4	0.25	0.1	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	22.326	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

Tabla V. Valores críticos de la distribución χ^2 de Pearson:
Abcisas $\chi^2_{\alpha;\nu}$ que dejan a su derecha un área α bajo la χ^2 con ν grados de libertad.

ν	0.995	0.99	0.975	0.95	0.9	0.75	0.5	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	-	-	-	-	0.016	0.102	0.455	1.32	2.71	3.84	5.02	6.63	7.88	10.8
2	0.010	0.020	0.051	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6	13.8
3	0.072	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8	16.3
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9	18.5
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7	20.5
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5	22.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3	24.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0	26.1
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6	27.9
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2	29.6
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8	31.3
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3	32.9
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8	34.5
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3	36.1
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8	37.7
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3	39.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7	40.8
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2	42.3
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6	43.8
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0	45.3
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4	46.8
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8	48.3
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2	49.7
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6	51.2
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9	52.6
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3	54.1
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6	55.5
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0	56.9
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3	58.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7	59.7