Niveladora de estadística - 2022

Probabilidad

1 Técnicas de Conteo

1.1 Principio de Multiplicación

Ejemplo 1 Una heladeria ofrece 3 gustos de helados, vainilla, chocolate y frutilla, y dos posible recipientes, vasito o cucurucho. Asumiendo que solo se puede elegir un sabor, ¿De cuántas formas posibles puedo elegir el helado?

Solución Hay tres gustos y por cada gusto hay dos posibles recipientes, entonces hay 3 * 2 combinaciones posibles. Notar que en este problema NO importa el orden. Por ejemplo, la combinacion Cucurucho seguida por Chocolate es la misma que la combinacion Chocolate seguida por Cucurucho. El ejemplo 1 se generaliza mediante el siguiente principio:

Principio de Multiplicación Si hay m maneras de hacer la tarea 1 y n maneras de hacer la tarea 2 entonces la realización conjunta de la tarea 1 seguida de la tarea 2 se puede hacer de $m.n$ formas.

Ejemplo 2 Supongamos que un menú permite seleccionar una entrada entre 4, una comida caliente entre 3 y un postre entre 5.

1. De cuántas formas puede elegir un menú un cliente?
2. De cuántas formas podrá hacerlo si quiere que el salpicón de ave y la suprema de pollo no aparezcan en el mismo menú?

1.2 Número Factorial

Ejemplo 3 Norma compró 5 libros para leer durante las vacaciones y quiere establecer un orden de lectura. De cuántas maneras puede hacerlo?

Solución Podemos argumentar mediante el principio de multiplicación. Para seleccionar el primer libro que leerá, Norma tiene 5 opciones; cualquiera haya sido ésta, tiene 4 posibles opciones para el segundo libro y así sucesivamente. Por lo tanto, mediante el principio de multiplicación la cantidad de formas en que puede organizar su lectura es igual a 5*4*3*2*1.

Ejemplo 4 Un grupo musical grabó 11 canciones con las que editará un nuevo disco. ¿De cuántas formas pueden elegir la secuencia de temas?

Generalizando estos ejercicios vamos a definir el siguiente objeto:

Número Factorial Dados n objetos distintos, siendo n un número natural cualquiera ¿De cuántas formas pueden ordenarse?

En primer lugar notamos que los dos problemas anteriores admiten el mismo planteo para $n = 5$ y $n = 11$ respectivamente. Esto nos induce a tratar de resolver esta situación general mediante un razonamiento análogo al utilizado en los casos particulares y así lo haremos.

Para seleccionar el primer elemento tenemos n posibilidades, y cualquiera sea nuestra eleccion tenemos $n-1$ formas de elegir el segundo. Luego hay $n*(n-1)$ maneras de elegir el primero y el segundo objeto. Por cada elección de los dos primeros tendremos $n-2$ posibilidades para el próximo y así sucesivamente. Entonces por aplicación reiterada del principio de multiplicación, concluimos que la cantidad de maneras de ordenar n objetos es

$$n! = n*(n-1)\dots 2*1$$

1.3 Número de Variación de n objetos tomados de a m.

Ejemplo 5 Tres personas se suben a un colectivo en el cual hay seis asientos libres. De cuántas maneras pueden ocuparlos?

Ejemplo 6 Para intervenir en un torneo de tenis de dobles mixtos, es necesario formar un equipo con tres parejas, debiendose elegir los jugadores entre los integrantes de un grupo constituido por 6 hombres y 3 mujeres. De cuántas maneras pude seleccionarse el equipo?

Como podemos observar, en los dos problemas el resultado fue el mismo. Esto no es casual. En ambos casos debimos efectuar una selección ordenada de 3 elementos (los asientos a ocupar, los hombres que jugarán) entre 6 posibles. Por que hablamos de una selección ordenada? La respuesta es que en ambos casos, no solo interesaba cuales eran los elementos elegidos sino también en que orden lo habían sido. Por ejemplo, en el primer problema, el orden es el que determina que asiento ocupará cada persona, y en el segundo, cual será el compañero de cada una de las damas.
Luego, ambos problemas son meras variantes de la siguiente situación general.De cuántas maneras pueden elegirse ordenadamente tres objetos tomados entre 6? Vamos a generalizar la pregunta ¿De cuántas maneras pueden elegirse ordenadamente m objetos tomados entre n? Para elegir el primer objeto tenemos n posibilidades, una vez elegido este hay $n-1$ posibilidades para el segundo, luego $n-2$ para el tercero y así sucesivamente. Al disponernos a elegir el último, notemos que ya fueron elegidos $m-1$ objetos, por lo tanto el número de elecciones posible para el mismo es $n-(m-1) = n-m+1$. Volviendo a usar el principio de multiplicación tenemos que el número total de elecciones ordenadas es

$$n*(n-1)*(n-2).....(n-m+1)$$

a este valor lo llamamos número de variación de n objetos tomados de a m. Notemos que

$$n*(n-1)*(n-2).....(n-m+1) = \frac{n!}{(n-m)!}$$

1.4 Número Combinatorio

Ejemplo 7 Para intervenir en un torneo de tenis masculino es necesario formar un equipo con tres personas debiendose elegir los jugadores entre los integrantes de un grupo constituido por 6 hombres. ¿De cuántas maneras pude seleccionarse el equipo?

Solución Este problema es parecido al problema 6 pero tenemos que seleccionar 3 caballeros entre 6 sin considerar el orden de los mismos. La cantidad de elecciones ordenadas de tres caballeros es 6*5*4. Sin embargo al considerar las elecciones ordenadas estamos contando cada equipo tantas veces como maneras hay de permutar los 3 caballeros elegidos. Esto es 3!=6 veces. Por ejemplo si P Q y R son los tres hombres seleccionados para formar el equipo, las elecciones PQR , PRQ, QPR, QRP, RPQ, RQP originan el mismo equipo. Luego la cantidad de equipos que puden formarse es

$$\frac{6*5*4}{3!} = \frac{6!}{3!3!}$$

En general dado un conjunto de n objetos y dado un número m menor o igual a n ¿De cuántas maneras pueden elegirse m objetos entre los n de un conjunto dado? la respuesta es el número combinatorio dado por

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

Ejemplo 8 Para integrar una comisión, se deberán elegir 4 personas entre un grupo formado por 8 hombres y 5 mujeres. ¿De cuántas maneras puede hacerse la eleccion? ¿Y si imponemos la condición de que dos de sus miembros por lo menos sean mujeres?

1.5 Problemas de conteo

Muchos problemas de conteo se pueden ver como problemas equivalentes a uno de los dos tipos de experimentos siguientes.

Muestreo con reposición:

En un bolillero hay n bolillas idénticas excepto por estar numeradas $\{1, 2, \ldots, n\}$. El experimento consiste en sortear, de a una por vez, k bolillas del bolillero, reponiendo la bolilla sorteada al bolillero cada vez. El resultado del experimento es la $k-$tupla ordenada $(a_1, \ldots, a_k)$, siendo a_1 el primer numero sorteado, a_2 el segundo, etc.

$$\# \text{ resultados posibles} = n^k$$

Muestreo sin reposición :

En un bolillero hay n bolillas idénticas excepto por estar numeradas $\{1, 2, \ldots, n\}$. El experimento consiste en sortear, de a una por vez, k bolillas del bolillero, sin reponer la bolilla sorteada al bolillero cada vez. El resultado del experimento es la k-tupla ordenada $(b_1, ..., b_k)$, siendo b_1 el primer núumero sorteado, b_2 el segundo, etc.

$$\# \text{ resultados posibles} = n(n-1)(n-2)....(n-k+1)$$

2 Teoría de Conjuntos

Operaciones entre conjuntos Dado un conjunto Ω y $A, B \subset \Omega$.

Unión de conjuntos $A \cup B = \{\omega \in \Omega : \omega \in A \text{ o } \omega \in B\}$

Intersección de conjuntos $A \cap B = \{\omega \in \Omega : \omega \in A \text{ y } \omega \in B\}$

Complemento de un conjunto $A^c = \{\omega \in \Omega : \omega \notin A\}$

Diremos que A y B son disjuntos si $A \cap B = \emptyset$

Ejemplo 9 Dado el siguiente conjunto $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ y los subconjuntos $A = \{\text{n pares}\}$ y $B = \{\text{n mayores a 4}\}$ calcular $A \cup B$, $A \cap B$, A^c, B^c, $(A \cup B)^c$, $(A \cap B)^c$

Leyes de Morgan dados A y $B \in \Omega$ vale que

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

3 Probabilidad

La probabilidad de un conjunto es un valor numérico que cuantifica la posibilidad de ocurrencia del evento.

Ejemplos de experimentos

1. Tirar un dado
2. Tirar dos dados
3. Tirar una moneda
4. Medir una persona elegida al azar
5. El número de alumnos que aprobará esta materia
6. El número de tiros al arco hasta que emboque un tiro
7. El precio del dolar manana

Definición El espacio muestral asociado a un experimento es el conjunto formado por todos los posibles resultados del mismo. Denotaremos al espacio muestral por la letra griega Ω

Ejemplos de espacios muestrales

1. Tirar un dado. $\Omega = \{1, 2, 3, 4, 5, 6\}$
2. Tirar dos dados. $\Omega = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6) \dots\}$
3. Tirar una moneda $\Omega = \{C, S\}$
4. Medir una persona elegida al azar $\Omega = \mathbb{R}_{\geq 0}$

5. El número de alumnos que aprobará esta materia $\Omega = \{1, 2, 3, 4, 5, 6, 7 \ldots, 100\}$

6. El número de tiros al arco hasta que emboque un tiro $\Omega = \mathbb{N}$

7. El precio del dolar manana $\Omega = \mathbb{R}_{\geq 80}$

Definición Un evento es un subconjunto del espacio muestral.

Ejemplos de eventos

1. Eventos: $A =$ sale un número par $= \{2, 4, 6\}$, $B =$ sale un $1 = \{1\}$

2. Evento: $A =$ la suma de los dos dados es $7 =$
   $\{(1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3)\}$

3. Evento: $A =$ Sale cara $= \{C\}$

4. Evento; $A =$ la persona mide más de $1.65 = (1.65; +\infty)$

Definición Un evento simple es un subconjunto del espacio muestral que tiene un único elemento.

Ejemplos de eventos

1. Eventos simples: $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$

2. Eventos simples $\{(1, 1)\}, \{(1, 2)\}, \{(1, 3)\}, \{(1, 4)\}, \ldots$

3. Eventos simples $\{C\}\{S\}$

4. Eventos simples $\{1.53\}$....

Propiedad Hay tantos eventos simples como elementos en Ω.

Definición de frecuencia relativa de un evento Sea Ω un espacio muestral asociado a un experimento y sea A un evento de Ω. Supongamos que repetimos n veces el experimento. Definimos la frecuencia relativa del evento A como

$$f_A = \text{Frecuencia Relativa de } A = \frac{\text{cantidad de veces que ocurre } A}{n}$$

Propiedades de la frecuencia relativa

$f_\Omega = 1$

Para todo evento A de Ω vale que $f_A \geq 0$.

Dados A y B eventos disjuntos de Ω, es decir $A \cap B = \emptyset$, entonces

$$f_{A \cup B} = f_A + f_B$$

Teorema de frecuencia relativa: si n es grande ($n > 30$)

$$f_A \approx \text{Probabilidad de que ocurra } A.$$

La probabilidad de eventos asociados a un experimento es algo que se parece mucho a la frecuencia relativa del evento (si hay muchas muchas repeticiones)

Axiomas de probabilidad Dado un experimento y su espacio muestral Ω, decimos que la función $P : \{\text{subconjuntos de } \Omega\} \to \mathbb{R}$ es una probabilidad si verifica

$A1$ Para todo $A \subset \Omega$ se tiene $P(A) \geq 0$

$A2$ $P(\Omega) = 1$

A_3 Si A y B son eventos disjuntos entonces

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

Más general, si $A_1, A_2 \ldots$ son eventos disjuntos dos a dos ($A_i \cap A_j = \emptyset$) entonces

$$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

Propiedades de la probabilidad (se deducen de los tres aximas $A_1 A_2$ y A_3)

1. $P(\emptyset) = 0$
2. Para todo evento A vale que $0 \leq P(A) \leq 1$
3. Dados A y B eventos $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
4. $P(A^c) = 1 - P(A)$
5. Si Ω es finito y $\{a_1\}$ .....$\{a_k\}$ son todos los eventos simples de Ω entonces

$$P(\{a_1\}) + P(\{a_2\}) + ...P(\{a_k\}) = 1$$

Asignación de probabilidades

¿Cómo calculamos probabilidades? Supongamos que el espacio muestral Ω asociado a cierto experimento es finito. En este caso, una manera simple de trabajar es asignar probabilidades a los eventos simples, ya que cualquier evento A será unión eventos simples y éstos son obviamente mutuamente excluyentes. Es decir si Ω es finito y $\{a_1\}$ .....$\{a_k\}$ son todos los eventos simples de Ω entonces resulta que

$$P(A) = \sum_{a_i \in A} P(\{a_i\})$$

Veamos una aplicación de este hecho

Ejemplo 9 Se arroja un dado equilibrado. ¿Cuál es la probabilida de cada evento simple? En este caso, $\Omega = \{1, 2, 3, 4, 5, 6\}$ y los eventos simples son $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$. Por suponer que el dado está equilibrado todos los eventos simples tienen igual probabilidad luego si llamamos $p = P(\{1\})$ resulta que por la propiedad 7

$$P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) + P(\{5\}) + P(\{6\}) = 6p = 1$$

de donde $p = \frac{1}{6}$. Si deseamos calcular la probabilidad del evento $A =$ el resultado es par, usando que $A = \{2\} \cup \{4\} \cup \{6\}$ se obtiene $P(A) = P(\{2\}) + P(\{2\}) + P(\{2\}) = 3/6 = 1/2$

Ejemplo 10 Supongamos ahora que se arroja un dado en el cual la probabilidad de las caras pares es el doble que la probabilidad de las caras impares. ¿Cuál es la probabilidad de cada evento simple? ¿Cuál es la probabilidad de sacar una cara par? ¿De sacar un número mayor a 3?

Espacios equiprobables Si en un experimento todos los eventos simples tienen igual probabilidad decimos que el espacio muestral es de equiprobabilidad y se puede ver que para todo evento A vale que

$$P(A) = \frac{\text{Cantidad de elementos de } A}{\text{Cantidad de elementos de } \Omega} = \frac{\#A}{\#\Omega}$$

Veremos algunos ejemplos

Ejemplo 11 De una urna que contiene 2 bolillas blancas y 3 rojas se extraen 2 bolillas con reposición.
a) ¿Cuál es la probabilidad de que se extraiga al menos una bolilla roja?
b) ¿Cuál es la probabilidad de que la primera bolilla extraída sea roja y la segunda blanca?

Ejemplo 11 Consideremos el ejemplo anterior pero suponiendo ahora que las extracciones se realizan sin reposición.

Ejemplo 12 Supongamos que tiramos dos dados (de 6 caras, equilibrados).

1. Describir un espacio muestral de este experimento que sea equiprobable.
2. Calcular la probabilidad de que en las dos tiradas hayan salido un 4 y un 1 en algun orden.
3. Calcular la probabilidad de que en las tiradas hayan salido dos 5.
4. Calcular la probabilidad de que el número obtenido en la segunda tirada sea estrictamente mayor al obtenido en la primera.

Ejemplo 13 El 65% del personal de una universidad está formado por mujeres y un 12% del personal son mujeres y menores de 24 años. Además el 20% del personal es menor de 24 años. Se selecciona un empleado de la universidad al azar. Hallar la probabilidad de que el empleado seleccionado

1. tenga 24 años o más .
2. tenga 24 años o más y sea un hormbre,

3. sea hombre.

Ejemplo 14 Se selecciona al azar un alumno de cierta universidad y señalamos como A al evento en que dicho individuo tiene una tarjeta de crédito Visa y como B el evento análago para una MasterCard. Supongamos que $P(A) = 0.5, P(B) = 0.4$, y $P(A \cap B) = 0.25$.

1. Calcule la probabilidad de que el individuo seleccionado tenga al menos una de las dos tarjetas.

2. ¿Cuál es la probabilidad de que el alumno no tenga ninguna de las dos tarjetas?

3. Describa en términos de A y B el evento ¨el alumno seleccionado tiene una tarjeta Visa pero no una MasterCard¨ y a continuación calcule la probabilidad de ese evento.

Ejemplo 15 Un comité está formado por cinco mexicanos, dos brasileros, tres argentinos y dos bolivianos. Se toma al azar un subcomité de cuatro personas.

1. Describir el espacio muestral asociado a este experimento.

2. ¿Cuál es la probabilidad de que todos los integrantes del subcomité sean de nacionalidades diferentes?

3. ¿Cuál es la probabilidad de ningún integrante del subcomité sea argentino?

4. ¿Cuál es la probabilidad de que todos los integrantes del subcomité sean mexicanos?

Ejemplo 16 Una familia, formada por una mamá un papá y tres hermanos, se forma al azar en hilera y se toma una fotografía.

1. Describir el espacio muestral asociado a este experimento.

2. ¿Cuál es la probabilidad de que el padre esté en el centro?

3. ¿ Cuál es la probabilidad de que la madre y el padre estén juntos?

Ejemplo 17 Se arroja un dado seis veces. ¿Cuál de las siguientes opciones ofrece una mejor chance de ganar?

1. Ganar \$100 si sale al menos un uno.

2. Ganar \$100 si sale un uno todas las veces.

3. Ganar \$100 si sale la secuencia 1,2,3,4,5,6.

4. Ganar \$100 si los dos primeros números que salen son iguales.

5. Ganar \$100 si no sale ni un uno ni un dos en las seis tiradas.

3.1 Probabilidad Condicional

Ejemplos y definición

Ejemplo 1: Consideremos una urna que contiene 4 bolillas rojas y 5 negras. De las 4 bolillas rojas, 2 son lisas y 2 rayadas y de las 5 bolillas blancas, 4 son lisas y una sola es rayada. Supongamos que se extrae una bolilla y, sin que la hayamos mirado, alguien nos dice que la bolilla es roja, ¿cuál es la probabilidad de que la bolilla sea rayada?

Solución: Sean los eventos A: la bolilla es rayada y B: la bolilla es roja. Obviamente, sin ninguna información previa, $P(A) = 3/9 = 1/3$ y $P(B) = 4/9$. Sin embargo, como sabemos que la bolilla es roja, la probabilidad de que sea rayada es $\frac{1}{2}$, ya que, de las rojas la mitad es lisa y la mitad rayada. Observemos, que al saber que la bolilla es roja el espacio muestral se reduce.

En general, dado un experimento y su espacio muestral asociado, queremos determinar cómo afecta a la probabilidad de A el hecho de saber que ha ocurrido otro evento B

Denotaremos con el símbolo $P(A|B)$ a la probabilidad de que ocurra A sabiendo que ocurrió el evento B. A dicha probabilidad la llamaremos probabilidad condicional de A dado B

Por ejemplo considere el evento A en que una persona seleccionada al azar padece de alopecia (calvicie). La probabilidad de A es 0.29. Sin embargo, suponga que la persona seleccionada es una mujer (el evento B) entonces la probabilidad de que ocurra A será mucho menor que 0.29 pues, tipicamente, las mujeres no padecen esa enfermedad. En símbolos: $P(A) = 0.29$ mientras que $P(A|B) << 0.29$

Definición: Sean A y B eventos tales que $P(B) > 0$, la probabilidad del evento A condicional a la ocurrencia del evento B es

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

Ejemplo 2: Consideremos una población en la que cada individuo es clasificado según dos criterios: es o no portador de cierta enfermedad y pertenece o no a cierto grupo de riesgo que denominaremos R. Se tiene la siguiente tabla:

	Portador	Sano
Pertenece al grupo de riesgo	3	17
No pertenece al grupo de riesgo	3	977

Se toma una persona al azar de esta población

1. ¿Cúal es la probabilidad de pertenezca al grupo de riesgo? ¿y de qué sea portador?
2. ¿Cual es la probabilidad de que sea portador y pertenezca al grupo de riesgo?
3. Sabiendo que la persona pertenece al grupo de riesgo ¿cuál es la probabilidad de que sea portador?

4. Sabiendo que la persona no pertenece al grupo de riesgo ¿cuál es la probabilidad de que sea portador?

Ejemplo 3 Una persona arroja dos dados equilibrados. Calcular la probabilidad de que la suma sea 7 dado que

1. la suma es impar.
2. la suma es mayor que 6.
3. el número del segundo dado es par.
4. el número de alguno de los dados es impar.
5. los números de los dados son iguales.

3.2 Teorema de multiplicación

Teorema de multiplicación: Dados dos eventos A y B, tales que $P(B) > 0$, vale que

$$P(A \cap B) = P(A|B)P(B)$$

Si además, $P(A) > 0$, entonces

$$P(A \cap B) = P(B|A)P(A).$$

Ejemplo 4: Supongamos que en el ejemplo 1 se extraen dos bolas sin reposición ¿Cuál es la probabilidad de extraer una bolilla roja y una negra, en ese orden?

Solución: Sean R_1: la primera bolilla es roja y N_2: la segunda bolilla es negra. Debemos calcular $P(R_1 \cap N_2)$. Aplicando la regla del producto

$$P(R_1 \cap N_2) = P(N_2|R_1)P(R_1) = \frac{5}{8}\frac{4}{9} = \frac{5}{18}.$$

Nota: el teorema de multiplicación es muy útil si una trabaja con un experimento que consta de varias etapas ya que el mismo se puede generalizar. Veamos el siguiente ejemplo.

Ejemplo 5: Supongamos que en el ejemplo 1 se extraen tres bolas sin reposición ¿Cuál es la probabilidad de extraer una bolilla roja, una blanca y una roja en ese orden?

Ejemplo 6: Cierta empresa se dedica a la venta de computadoras. De sus ventas, el 50% son de la marca 1, 30% de la marca 2 y 20% de la marca 3. Cada fabricante ofrece un año de garantía. Se sabe que el 25% de las computadoras de la marca 1 requieren reparación durante el período de garantía, mientras que los porcentajes correspondientes para las marcas 2 y 3 son 20% y 10%.

1. ¿Cuál es la probabilidad de que un comprador de computadoras de esta empresa seleccionado al azar adquiera un equipo de la marca 1 que necesitará reparación durante el período de garantía?
2. ¿Cuál es la probabilidad de que un cliente de esta empresa seleccionado al azar adquiera un equipo que necesitará de reparación durante el período de garantía?

3.3 Teorema de probabilidad total

Teorema de Probabilidad Total: supongamos que $A_1, A_2 \ldots A_k$ son eventos de cierto espacio muestral Ω tales que $A_1 \cup A_2 \ldots \cup A_k = \Omega$ y además para todo $i, j \in \{1, 2...k\}$ vale que $A_i \cap A_j = \emptyset$. En tal caso resulta que para cualquier evento B de Ω

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) \ldots + P(B|A_k)P(A_k) = \sum_{i=1}^{k} P(B|A_i)P(A_i)$$

Ejemplo 7: Un reconocido hospital tiene tres sedes llamadas Almagro, Belgrano y San Martín. En la sede Almagro el 13% de los pacientes diabéticos están en estado crítico. En la sede Belgrano el 15 % de los pacientes diabéticos están en estado crítico mientras que en la sede San Martín el 23% de los pacientes diabéticos está en estado crítico. Además, de la poblacion total de pacientes diabéticos del hospital un 27% se atiende en la sede Almagro, un 33% en la sede Belgrano y un 40% en la sede San Martín. Si se toma al azar un paciente diabético del hospital, cualcular la probabilidad de que esté en estado crítico.

3.4 Teorema de Bayes

Ejemplo 8 Si en el ejemplo 7 se toma un paciente diabético del hospital al azar y resulta que está en estado crítico ¿Cuál es la probabilidad de que se atienda en la sede San Martín?

Teorema de Bayes: Sea $A_1, A_2 \ldots A_k$ una partición del espacio muestral Ω, es decir, $A_1 \cup A_2 \ldots \cup A_k = \Omega$ y además para todo $i, j \in \{1, 2...k\}$ vale que $A_i \cap A_j = \emptyset$. Sea B un evento cualquiera tal que $P(B) > 0$, entonces

$$P(A_j|B) = \frac{P(B \cap A_j)}{P(B)} = \frac{P(B|A_j)P(A_j)}{P(B)}$$

3.5 Independencia

Ejemplo 9 De una urna que contiene 4 bolillas negras y 6 blancas se extren dos bolas con reposición ¿Cuál es la probabilidad de que la segunda bolilla sea blanca sabiendo que la primera fue negra?

La noción intuitiva de independecia de dos eventos es que la ocurrencia de uno no dice nada acerca de la ocurrencia del otro. A partir de ahi, uno puede pensar que A y B son independientes si la probabilidad de que ocurra A dado que ocurrió B es la misma que la probabilidad de que A ocurra. De ahi sale la definición de independencia.

Definición 1: Sea B un evento tal que $P(B) > 0$. Diremos que A y B son independientes si

$$P(A|B) = P(A).$$

Si bien esta cuenta respeta la noción de independencia entre eventos, queda por determinar que pasa cuando $P(B) = 0$....en tal caso tenemos la siguiente definición

Definición 2: Los eventos A y B son independientes si

$$P(A \cap B) = P(A)P(B).$$

Ejemplo 10 De un mazo de 40 cartas españolas se extrae una al azar. Consideremos los siguientes eventos: A: la carta es una copa o espada, B=la carta no es copa y C= la carta es copa u oro. ¿Son A y B independientes? ¿Son A y C independientes?

Propiedades:

1. Si los eventoss A y B son excluyentes, es decir $P(A \cap B) = 0$, entonces A y B no son independientes.

2. Si $P(B) = 0$ entonces B es independiente de cualquier evento A tal que $P(A) > 0$.

3. Si A y B son idependientes, entonces

   (a) A y B^c son independientes

   (b) A^c y B^c son independientes

   (c) A^c y B son independientes

Ejemplo 11 Para ingresar a una universidad se deben aprobar tres exámenes: uno de matemática, uno de historia y uno de deportes. Juan sabe que, siendo egresado del colegio X, la probabilidad de aprobar el examen de matemática es 0.6, el de historia es 0.7 y el de deportes es 0.3. (La aprobación de cada uno de los exámenes es independiente de la aprobación de los demás).

1. Hallar la probabilidad de que Juan no ingrese a la universidad.

2. Primero se rinden los exámenes de matemática e historia. Los alumnos que aprueban estos exámenes rinden el examen de deportes. Calcular la probabilidad de que Juan rinda el examen de deportes.

3. Pedro, que viene del colegio Y, sabe que la probalidad de aprobar los exámenes de matemática y de historia es 0.3 y que la probabilidad de aprobar el examen matemática y no el de historia es 0.4.

   (a) ¿Cuál es la probabilidad de que apruebe el examen de matemática?

   (b) Si además sabe que la probabilidad de aprobar el examen de historia y no el de matemática es 0.1. ¿Cuál es la probabilidad de que Pedro o Juan rindan el examen de deportes?

4 Variables aleatorias discretas

Definición- Ejemplos

Cuando hacemos un experimento, tipicamente queremos mirar ciertos resultados destacados del mismo. Por ejemplo, si tomamos con reposición 100 tormillos al azar de la producción total, podemos estar interesados en la cantidad de tornillos, entre los 100 elegidos, que son defectuosos. Otro ejempo sería tirar dos veces un dado, en dicho contexto podríamos estar interesados en la suma de los puntos obtenidos. A esta cantidad de interés se la denomina variable aleatoria. Variable porque toma distintos valores y aleatoria porque el valor observado no puede ser predicho antes de la realización del experimento, aunque sí se sabe cuáles son sus posibles valores. Vamos a formalizar un poco estas ideas...

Definición Dado un experimento y Ω su espacio muestral asociado, X es una variable aleatoria si $X : \Omega \to \mathbb{R}$ es una función. De esta manera si $\omega \in \Omega$, entonces $X(w)$ es un número.

Ejemplos: Se arroja dos veces un dado equilibrado. Posibles v.a. asociadas con este experimento son:

- $X :=$ número de caras pares
- Y : máximo puntaje
- Z : suma de puntos de ambos dados

Calculemos el valor de X, Y y Z para distintos $\omega \in \Omega = \{(x_1, x_2)/x_i \in \{1, 2, 3 \ldots 6\}, i = 1, 2\}$

- $X((2,5)) = 1, \quad Y((2,5)) = 5, \quad Z((2,5)) = 7$
- $X((1,3)) = 0, \quad Y((1,3)) = 3, \quad Z((1,3)) = 4$
- $X((2,2)) = 2, \quad Y((2,2)) = 2, \quad Z((2,2)) = 4$

Otros ejemplos de variables aleatorias son $T =$: tiempo (en años) que tarda un alumno en terminar la carrera, $A =$: altura de una persona tomada al azar, $V =$: cantidad de llamadas telefónicas recibidas.

Definición: Llamamos rango de una variable aleatoria X a la imagen de la misma, es decir al conjunto de valores que puede tomar la variable. Denotamos al rango mediante R_X.

Ejemplos

- $X :=$ número de caras pares, $\quad R_X = \{0, 1, 2\}$.
- Y : máximo puntaje, $\quad R_Y = \{1, 2 \ldots 6\}$.
- Z : suma de puntos de ambos dados, $\quad R_Z = \{2 \ldots, 12\}$.

Definición: Una v.a. es discreta si toma un número finito o infinito numerable de valores. En otras palabras, una variable es discreta si su rango es finito o finito numerable.

Por ejemplo las variables X, Y y Z definidas en el primer ejemplo resultan ser discretas, también es discreta la variable $V =$: cantidad de llamadas telefónicas diarias recibidas, mietras que $A =$: altura de una persona tomada al azar no es discreta.

4.1 Función de probabilidad puntual.

Ejemplo 1: En el caso del ejemplo 1), ¿cómo calcularíamos la probabilidad de que la v.a. Z tome el valor 5?

Definición: Sea $X : \Omega \to \mathbb{R}$ una variable aleatoria. Llamamos función de probabilidad puntual de X a la función $p_X : \mathbb{R} \to [0, 1]$ definida por

$$p_X(x) = P\left(\{\omega \in \Omega : X(\omega) = x\}\right).$$

De ahora en adelante usaremos la siguiente notación

$$\{\omega \in \Omega : X(\omega) = x\} = \{X = x\}$$

Ejemplo 2: Si $X :=$ número de caras pares al tirar dos veces un dado equilibrado. Hallar p_X.

Propiedades: p_X verifica:

- $0 \leq p_X(x) \leq 1$.
- $\sum_{x \in R_X} p_X(x) = 1$.

Ejemplo 2 Una empresa proovedora de televisión por cable ofrece a sus clientes 5 promociones. Definimos $C =$ número de promociones que contrata un cliente de la empresa. La función de probabilidad puntual de C viene dada por

k	1	2	3	4	5
$p_C(k)$	0.375	0.275	0.175	0.1	0.075

Se toma un cliente de esta empresa al azar, calcular la probabilidad de que

a) contrate dos promociones.

b) contrate por lo menos dos promociones.

Ejemplo 3 Un juego, llamado Suma Siete, consiste en tirar dos dados. Para jugar hay que pagar una ficha. Si la suma de ambos resultados es 7 el casino te devuelve la ficha apostada y te paga 4 fichas más. Si la suma no es siete, el casino se queda con la ficha apostada. Supongamos que cada ficha sale \$100. Sea G la variable aleatoria que cuenta el dinero ganado en cada apuesta. Hallar la función de probabilidad puntual de la variable G.

4.2 Función de distribución

La función de distribución acumulada de una v.a discreta X con función de probabilidad puntual p_X se define como

$$F_X(x) = P(X \leq x) = \sum_{y \leq x, y \in R_X} p_X(y),$$

es decir, $F_X(x)$ es la probabilidad de que la variable aleatoria tome valores menores o iguales a x.

Ejemplo 4 Volviendo al ejemplo 2, hallemos función de distribución acumulada de la variable aleatoria $X :=$ número de caras pares al tirar dos veces un dado equilibrado.
¿Qué características tiene F_X? Observemos que se trata de una función escalera, no decreciente que toma valores entre 0 y 1.

Propiedades de la función de distribución acumulada

1. $\forall x \in \mathbb{R}$, $F_X(x) \in [0,1]$
2. F_X es monótona no decreciente, es decir, si $x_1 < x_2$ entonces $F_X(x_1) \leq F_X(x_2)$.
3. F_X es continua a derecha, es decir, para todo x vale que $\lim\limits_{t \to x^+} F_X(t) = F_X(x)$
4. $\lim\limits_{x \to +\infty} F_X(x) = 1 \qquad \lim\limits_{x \to -\infty} F_X(x) = 0$
5. En cada x el valor del salto es la probabilidad puntual, es decir
$$p_X(x) = F_X(x) - F_X(x^-),$$
donde
$$F_X(x^-) = \lim_{t \to x^-} F_X(t).$$
En particular, notar que si $R_X = \{x_1, x_2 \dots x_k\}$, donde $x_1 < x_2 \dots < x_k$, resulta que
$$p_X(x_i) = F_X(x_i) - F_X(x_i^-), \text{ para } i = 2, ..k \text{ y } p_X(x_i) = F_X(x_1).$$

Propiedad Sean $a, b \in \mathbb{R}$ con $a \leq b$, entonces

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

$$P(a \leq X \leq b) = F_X(b) - F_X(a^-)$$

$$P(a < X < b) = F_X(b^-) - F_X(a)$$

$$P(a \leq X < b) = F_X(b^-) - F_X(a^-)$$

Ejemplo 5 Sea X una variable aleatoria con función de distribución

$$F_X(x) = \begin{cases} 0 & \text{si } x < 1 \\ 0.3 & \text{si } 1 \leq x < 3 \\ 0.4 & \text{si } 3 \leq x < 6 \\ 0.6 & \text{si } 6 \leq x < 12 \\ 1 & \text{si } 12 \leq x \end{cases}$$

a) Hallar la función de probabilidad puntual de X.

b) Calcular de dos maneras: utilizando la función de distribución y utilizando la función de probabilidad puntual, las siguientes probabilidades:

$$P(3 < X \leq 6) \qquad P(3 \leq X \leq 6) \qquad P(X \geq 4) \qquad P(X \geq 6)$$

4.3 Variables aleatorias discretas famosas

4.3.1 Variable Aleatoria Binomial

Definición Supongamos que un experimento puede dar dos resultados llamados éxito y fracaso. Considerar $\Omega = \{E, F\}$ su espacio muestral asociado y supongamos que $P(\{E\}) = p$. Llamamos variable aleatoria Bernoulli de parametro p a la variable $X : \Omega \to \{0, 1\}$ definida por

$$X = \begin{cases} 1 & \text{si ocurrió un éxito} \\ 0 & \text{si ocurrió un fracaso.} \end{cases}$$

Es decir $X(\{E\}) = 1$ y $X(\{F\}) = 0$. Para denotar que X es una variable Bernoulli de parámetro p pondremos: $X \sim Ber(p)$

Ejemplo 1 El 37% de los alumnos de UBA fuma. Tomamos 3 alumnos de UBA con reposición y definimos la variable X = cantidad de alumnos entre los tres elegidos que fuma. Hallar R_X y p_X.

Definición Supongamos que ocurre lo siguiente:

1) Un experimento consiste de n pruebas, siendo n fijo.

2) Las pruebas son idénticas y en cada prueba hay sólo dos resultados posibles, que denominaremos Éxito (E) y Fracaso (F). Una prueba de este tipo se denomina ensayo de Bernoulli.

3) Las pruebas son independientes, es decir que el resultado de una prueba no influye sobre el de las otras.

4) La probabilidad de Éxito $P(\{E\}) = p$ se mantiene constante en todas las pruebas.

Definimos la variable aleatoria X = cantidad de exitos en las n pruebas. X se llama variable aleatoria binomial de parámetros n y p (en símbolos: $X \sim Bi(n, p)$) .
Se verifica

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

Observación: Una variable Bernoulli de parámetro p es una binomial de parámetros $n = 1$ y p.

4.3.2 Variable aleatoria geométrica

Ejemplo 3 Una rueda de ruleta está dividida en 38 secciones, de las cuales 18 son rojas, 18 son negras y las 2 restantes son verdes. Sea X = número necesario de juegos hasta obtener una sección verde en jugadas independientes. ¿Cuál es la probabilidad de que sean necesarias 3 jugadas? Hallar la función de probabilidad puntual de X.

Definición Supongamos que se repite en forma independiente un experimento que puede dar éxito o fracaso con probabilidad de éxito igual a p en todas las repeticiones. Se define la v.a. X: número de repeticiones hasta obtener el primer éxito. X se llama variable aleatoria geométrica de parámetro p y se denota con $X \sim G(p)$. Se verifica

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

Ejemplo 4 En el ejemplo de la ruleta ¿Cuál es la probabilidad de que sean necesarios al menos 3 tiros hasta obtener la primer sección verde? ¿Cuál es la probabilidad de que sean necesarias un número impar de tiradas para obtener la primer sección verde?

4.3.3 Variable aleatoria hipergeométrica

Ejemplo 5 De una urna que contiene 2 bolillas blancas y 7 negras se extraen 3 bolillas sin reposición y se define X: número de bolillas blancas extraídas. Calcular la probabilidad de que X sea 2.

Definición: Se toman n individuos de una población que tiene N elementos clasificados en dos categorías: sanos y defectuosos, habiendo D defectuosos. Definimos la variable $X =$ cantidad de indiviudos defectuosos entre los n seleccionados. X se llama variable aleatoria hipergeométrica de parámetros n, N, D y se denota mediante $X \sim H(n, N, D)$. Tenemos que

$$p_X(k) = P(X = k) = \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}} \text{ donde } \quad max\{n-(N-D), 0\} \leq k \leq min\{D, n\}$$

4.3.4 Variable aleatoria Poisson

Definición X es una variable aleatoria Poisson de parámetro λ(un número positivo) si

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} \quad k \in \mathbb{N}.$$

Ejemplo 6 Se sabe que el número de microorganismos por gramo de una cierta muestra de suelo diluida en agua destilada, sigue una distribución de Poisson de parámetro $\lambda = 8$. Si una preparación con un gramo de esta dilución se vuelve turbia, este gramo contiene al menos un microorganismo. Sabiendo que una preparación se volvió turbia, halle la probabilidad de que tenga:

a) un sólo microorganismo

b) menos de tres microorganismos.

c) más de dos microorganismos

4.4 Ejercicios

1. Una persona va todos los días en colectivo a su trabajo. Cierto mes trabaja 20 días. Si el colectivo pasa en 10 minutos o menos, llega al trabajo en horario y si demora más de 10 minutos en pasar, llega tarde. Sabemos que la probabilidad de que pase en 10 minutos o menos es 0,9. Se le arma lío si no llega en horario en más de 3 ocasiones en el mes. Se asume independencia entre los distintos días. ¿Cuál es la probabilidad de que se le arme lío?

2. Un geólogo ha recolectado 10 especímenes de roca basáltica y 12 de granito. Un asistente de laboratorio selecciona 3 de los especímenes para analizarlos.

   a) ¿Cuál es la función de probabilidad puntual del número de especímenes de basalto seleccionados para ser analizados?

   b) ¿Cuál es la probabilidad de que todos los especímenes seleccionados para el análisis sean del mismo tipo de roca (es decir, todos graníticos o todos basálticos)?

5 Variables aleatorias continuas

Diferencia con las discretas

Supongamos que medimos el sueldo en bruto de 55 docentes universitarios. Los primeros 10 valores obtenidos fueron los siguientes

$$42.356,76 \;\; 39.234,354 \;\; 28.143,98 \;\; 65.765,98 \;\; 43.765,98 \;\; 36.053,87 \;\; 78.709,76 \;\; 20.102,89 \;\; 24.760,90$$

¿Se pude hacer una tabla de frecuencia? ¿Qué es lo que falla??

Recuerdo: Una variable es continua si entre dos valores del rango de la misma hay infinitos valores. Luego una variable continua puede tomar tantos valores que al seleccionar una muestra de dicha variable ningún valor se repite!!!! Para presentar la versión continua de la función de probabilidad puntual necesitamos recordar el concepto de frecuencia relativa.

Sea Ω un espacio muestral asociado a un experimento y sea A un evento de Ω. Supongamos que repetimos n veces el experimento.

Definición:

$$f_A = \text{Frecuencia Relativa de } A = \frac{\text{cantidad de veces que ocurre } A}{n}$$

Teorema de frecuencia relativa: si n es grande

$$f_A \approx P(A).$$

Ejemplo

- El experimento es tirar un dado.
- $A = \{\text{ sale un 1 }\}$.
- Repetimos 1587 veces el experimento.
- 259 veces salió un 1.

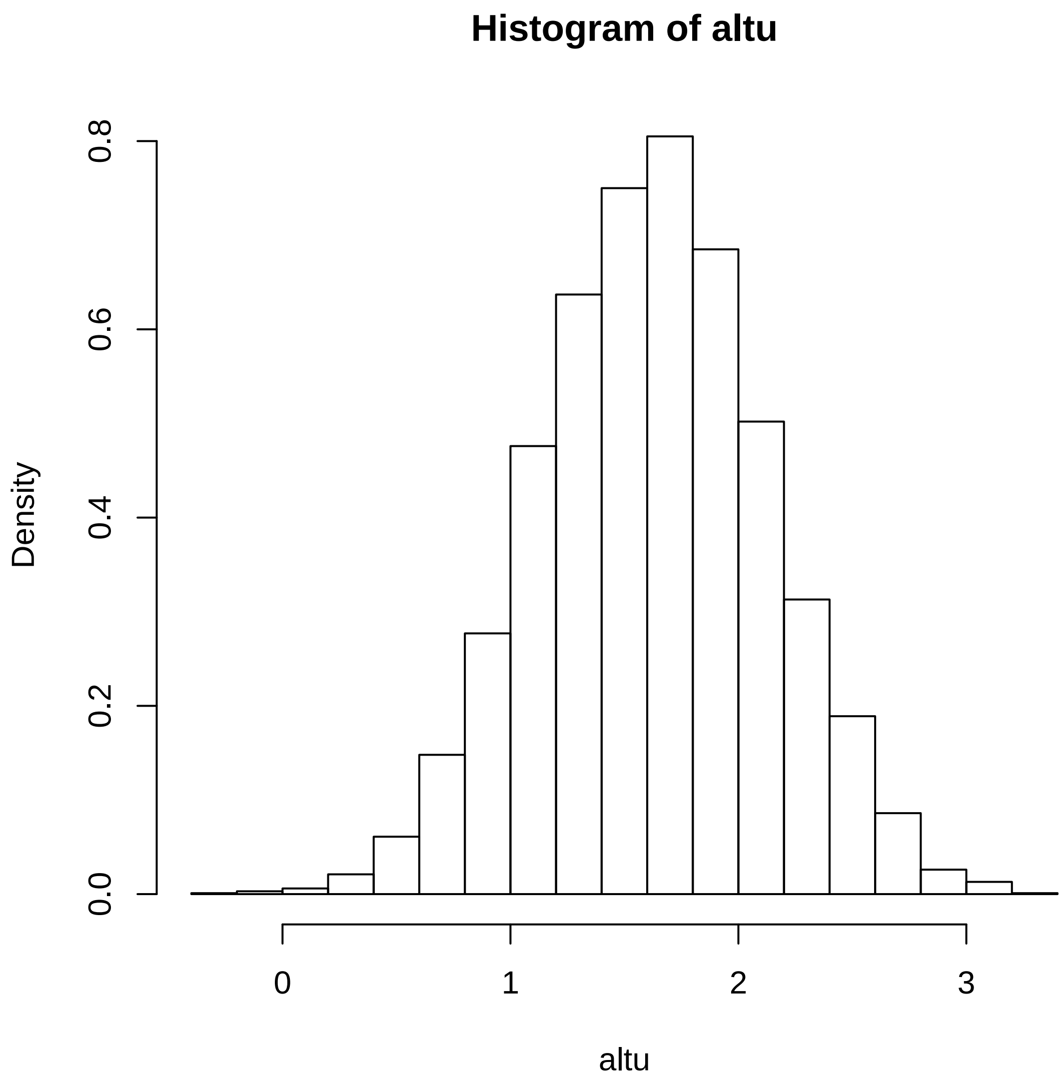
$$f_{R,A} = \frac{\text{cantidad de veces que ocurre } A}{n} = \frac{259}{1587} = 0.163201.$$

$$P(A) = \frac{1}{6} = 0.1666667$$

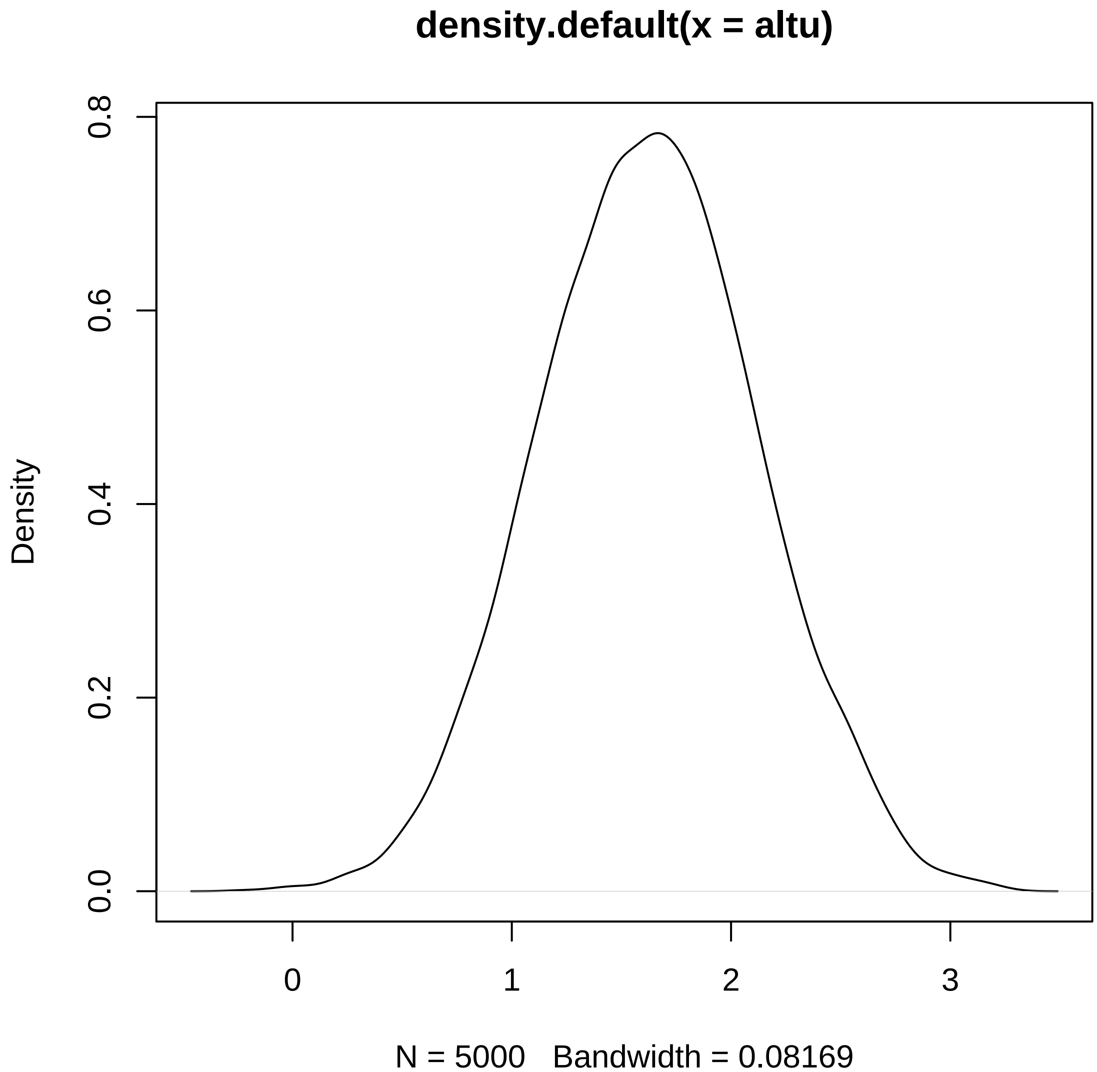
5.1 Función de densidad

Mediante el siguiente ejemplo queremos introducir la nocion de densidad de una variable aleatoria.

- $X = \{$ altura de una argentina tomada al azar $\}$.
- X es una variables aleatoria.
- Tomamos la altura de 5489 mujeres argentinas en todo el país: $X_1, X_2 .... X_{5489}$ (mediciones)
- Con las mediciones construimos el siguiente histograma:



Una función de densidad de la variable X es cualquier función que se aproxime al histograma de la variable X



Notar que

$$\frac{\#\{X_i \text{en}\,(a,b)\,\}}{5489} \approx P(X \in\ (a,b)\,) \approx \text{área de un rectángulo}\,(a,b) \approx \int_a^b f(x)dx$$

siendo f una densidad para X.

En general dada una variable aleatoria continua el histograma de la variable se construye haciendo un histograma de frecuencia relativa para mediciones independientes de dicha variable. Una función de densidad para la variable es cualquier función que aproxime al histograma de la misma.

Aclaración: No vamos a trabajar con integrales. Solo las escribiremos en la siguiente definición, ya que se debe entender el objeto pero no realizar cuentas con integrales.

Definición Dada X una variable aleatoria continua y f una densidad para X entonces

1. $f(x) \geq 0$
2. $\int_{\mathbb{R}} f(x)dx = 1$

3. $P(X \in \ (a,b)\ ) = \int_a^b f(x)dx$

Ejercicios

1. La fracción de alcohol X en cierto compuesto puede considerarse una v.a. donde X tiene función de densidad

$$f_X(x) = c(1-x) \quad I_{[0,1]}(x)$$

   (a) Determinar c.(sin integrar, usando area de triangulos)
   (b) Se elige un compuesto al azar. Hallar la probabilidad de que la fracción de alcohol en dicho compuesto esté entre $\frac{1}{3}$ y $\frac{2}{3}$.
   (c) Supóngase que el precio de venta del compuesto depende del contenido del alcohol: si $x < \frac{1}{3}$ el precio es 10\$, $\frac{1}{3} < x < \frac{2}{3}$ el precio es 20\$ y si $x > \frac{2}{3}$ el precio es 30\$. Hallar la función de probabilidad puntual del precio de venta del producto.

2. El diámetro D (expresado en dm) del tronco de cierta especie de árboles es una variable aleatoria con función de densidad

$$f_D(x) = k\, x\, I_{(0,10)}(x).$$

   a) Hallar el valor de la constante k.
   b) ¿Cuál es la probabilidad de que el diámetro de un árbol de esa especie elegido al azar mida entre 4 y 6 dm?
   c) Idem b) sabiendo que el diámetro mide más de 5 dm.
   d) En un área del bosque hay 3 árboles de esa especie. Calcular la probabilidad de que exactamente 2 de ellos tengan el diámetro entre 4 y 6 dm.

5.2 Variables aleatorias continuas famosas

5.2.1 Distribución Uniforme:

X tiene distribución uniforme en el intervalo $[a,b]$, si su función de densidad es

$$f_X(y) = \begin{cases} \frac{1}{b-a} & a \leq x < b \\ 0 & \text{en otro caso.} \end{cases}$$

Notación: $X \sim U(a,b)$.

Ejercicio 1 El tiempo, en minutos, que una persona permanece en la cola de un supermercado se puede modelar mediante una variable aleatoria $U(1,10)$.

1. Calcular la probabilidad de que una persona esté en la cola menos de 6 minutos.

2. Calcular la probabilidad de que una persona esté entre 3 y seis minutos

3. ¿Cuál es el tiempo medio esperado?

5.2.2 Distribución normal:

Definición Una variable aleatoria X tiene una **distribución normal** o **Gaussiana** con de parámetros μ y σ, $X \sim N(\mu, \sigma^2)$ si X puede tomar cualquier valor en la recta real y para cualquier $a < b$ vale que

$$P(a \leq X \leq b) = \int_a^b f_X(x)\, dx$$

siendo

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

- La función $f_X(x)$ es la **función de densidad** de la distribución normal.
- Cuando $\mu = 0$ y $\sigma = 1$ decimos que X tiene una distribución *normal estándard.*
- Si $X \sim N(\mu, \sigma^2)$, entonces X es una variable aleatoria continua pues puede tomar valores en cualquier punto de $(-\infty, +\infty)$.

Para muchas variables X que medimos en economía o en negocios, como

- la demanda mensual X de un cierto artículo en un comercio, por ejemplo el número de botas de mujer solicitadas en una zapatería en un mes,
- el logaritmo del ingreso mensual X de un hogar elegido al azar en la ciudad de Buenos Aires
- el retorno mensual X de una cierta cartera de inversión

es razonable suponer que tienen una distribución **aproximadamente normal**, esto es

$$P(a \leq X \leq b) \approx \int_a^b f_X(x)\, dx$$

para

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

y para cierto μ y $\sigma > 0$.

Para una variable aleatoria con distribución $N(\mu, \sigma^2)$, la integral

$$\int_a^b f_X(x)\, dx = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\, dx$$

no se puede resolver por lo métodos usuales de integración.

Sin embargo, para este caso existen varias formas prácticas de calcular los valores de la **distribución acumulada**

$$F_X(u) = P(X \leq u) = \int_{-\infty}^u f_X(x)\, dx$$

Una vez que sabemos calcular $P(X \leq u)$ para cualquier u, podemos calcular $P(a \leq X \leq b)$ con la fórmula

$$P(a \leq X \leq b) = F_X(b) - F_X(a)$$

Las distribución acumulada $F_X(u)$ de una variable aleatoria con distribución normal estándar se puede calcular por medio de:

- tablas que aparecen en los apéndices de casi todos los libros introductorios de estadística
- Excel (Add-In "Analysis ToolPack")
- distintos paquetes con software para realizar analisis estadí stico como Stata, R, SAS, SPSS
- applets disponibles en la web

Teorema de estandarización

$$X \sim N(\mu, \sigma^2) \Leftrightarrow Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

Hay dos tipos de preguntas

Problema 1 Dado x calcular $P(X \leq x) =?$.

Problema 2 Dado p calcular $P(X \leq ?) = p$.

Ejemplo Supongamos que la demanda mensual de pares de botas de un cierto estilo durante el mes de Julio a nuestra fábrica es una variable aleatoria X que sigue (aproximadamente) una distribución $N(1500, 100^2)$.

1. Si durante el mes de Julio producimos 1600 pares de botas, ¿cuál es la probabilidad de satisfacer la demanda?
2. Si durante el mes de Julio queremos satisfacer la demanda con una probabilidad del 95%, ¿cuántos pares de botas deberíamos producir?

Solución:

1.

$$\begin{aligned} P(X \leq 1600) &= P\left(\frac{X - 1500}{100} \leq \frac{1600 - 1500}{100}\right) \\ &= P\left(Z \leq \frac{1600 - 1500}{100}\right) = P(Z \leq 1) = 0.8413 \end{aligned}$$

Si fabricamos 1600 pares de botas tendremos una probabilidad del 84% de cubrir la demanda.

2.

$$0.95 = P(X \leq x) = P\left(\frac{X - 1500}{100} \leq \frac{x - 1500}{100}\right) = P\left(Z \leq \frac{x - 1500}{100}\right)$$

Obtenemos $1.65 = \frac{x-1500}{100}$ y entonces $x = 1500 + 1.65 \cdot 100 = 1665$.

Deberíamos producir 1665 pares para satisfacer la demanda con un 95% de probabilidad.

Teorema

$$X \sim N(\mu, \sigma^2) \Leftrightarrow Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$$

Ejemplo: cambio de unidades Supongamos que el peso medido en kg. de una persona elegida al azar de una población es una variable aleatoria $X \sim N(\mu, \sigma^2)$. Llamemos Y al peso de la misma persona, pero expresado en gramos. Entonces

$$Y = 1000X$$

$$Y \sim N(1000\mu, 1000^2\sigma^2)$$

5.3 Esperanza y varianza de sumas de variables aleatorias independientes

Recordemos que si $X_1, X_2, ..., X_n$ son variables aleatorias

$$E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$$

si además son independientes

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$$

Proposición: Sean $X_1, X_2, ..., X_n$ son v.a. independientes con la misma esperanza y varianza. $E(X_i) = \mu$ y $Var(X_i) = \sigma^2$. entonces

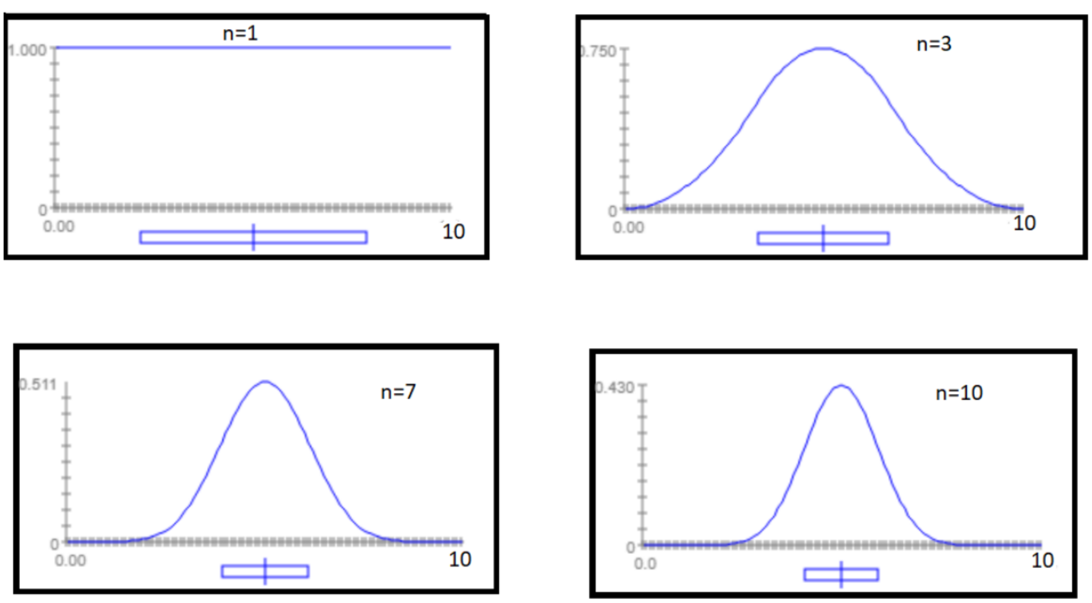
$$E(\sum_{i=1}^{n} X_i) = n\mu \quad , \quad Var(\sum_{i=1}^{n} X_i) = n\sigma^2$$

$$E\big(\frac{\sum_{i=1}^{n} X_i}{n}\big) = \mu \quad , \quad Var\big(\frac{\sum_{i=1}^{n} X_i}{n}\big) = \frac{\sigma^2}{n}.$$

6 Teorema central del límite

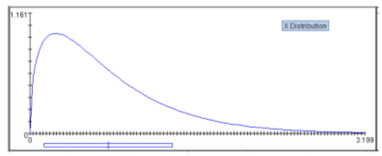
Ejemplos Sean $X_1, ..X_n$ una muestra aleatoria i.i.d.

1. X_i con distribución uniforme en el intervalo $(0, 10)$ En los gráficos se observa la distribución de la media muestral para distintos tamanos de muestras.

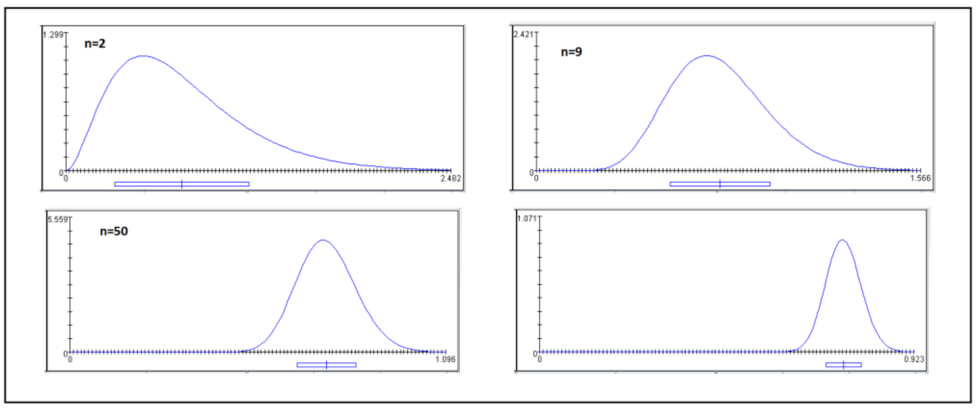


2. X_i con distribución chi cuadrada.

Distribución de X_i

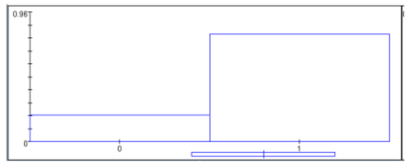


Distribución de $\overline{X}$

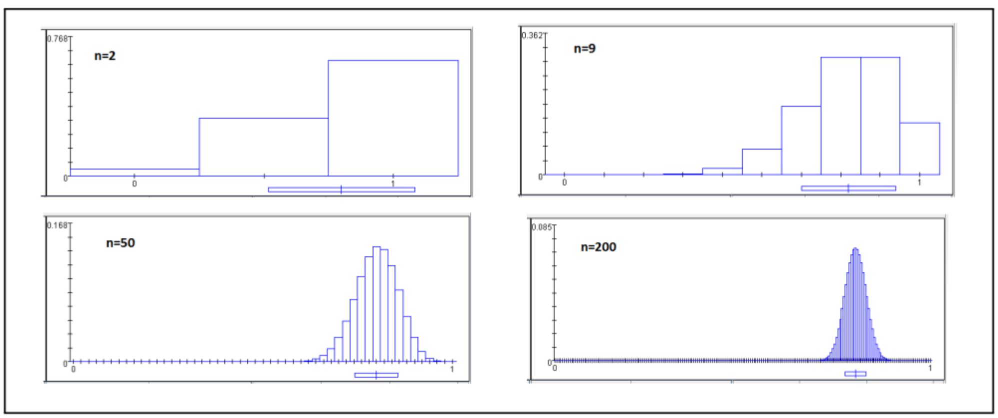


3. X_i con distribución Bernoulli, con probabilidad de éxito $0,7$

Distribución de $X_i \sim Ber\,(0.7)$



Distribución de $\overline{X}$



Teorema central del límite:

Sean $X_1, X_2, ..., X_n, ...$ variables aleatorias independientes e identicamente distribuidas (iid). Llamamos $E(X_i) = \mu$ y $Var(X_i) = \sigma^2$. Entonces para n grande ($n \geq 30$) las variables aleatorias $\sum_{i=1}^{n} X_i$ y $\dfrac{\sum_{i=1}^{n} X_i}{n}$ tienen distribuciones aproximadamente normales:

$$\sum_{i=1}^{n} X_i \approx N(n\mu, n\sigma^2) \quad y \quad \frac{\sum_{i=1}^{n} X_i}{n} \approx N(\mu, \frac{\sigma^2}{n})$$

Ejemplo Una compañía aérea sabe que el peso de la valija de un cliente es una va con media $\mu = 20\,kg$ y varianza $\sigma^2 = 100\,kg$.
La compañía decide no pesar el equipaje de sus clientes y les permite despachar 2 valijas por persona. Si en un avión viajan 200 personas y la bodega soporta $10000\,kg$ ¿cuál es la probabilidad de sobrecarga?
Solución
Sean $X_1, ..., X_{400}$ los pesos de las valijas, $E(X_i) = 20$ y $Var(X_i) = 100$ Luego por el TCL vale que

$$\sum_{i=1}^{400} X_i \approx N(400 * 20, 400 * 100) = N(8000, 40000)$$

luego

$$P(\text{sobrepeso}) = P\left(\sum_{i=1}^{400} X_i > 10000\right)$$

$$= P\left(\frac{\sum_{i=1}^{400} X_i - 8000}{\sqrt{40000}} > \frac{10000 - 8000}{200}\right) \approx 1 - \Phi(10) \cong 0$$

Ejemplo Una compañía de seguros aseguró 50 pozos petrolíferos contra posibles daños ambientales. La ganancia X en miles de dólares de cada una de estas pólizas se puede modelar como una v.a. continua de media $\mu = 17.5$ y varianza $\sigma^2 = 1.69$. Si las ganancias $X_1, X_2, ..., X_{50}$ obtenidas son independientes. ¿Cuál es la probabilidad de que la ganancia promedio supere los 18000 dólares?.
Solución: Sean $X_1, ..., X_{50}$ vs.as. iid con $E(X_i) = 17.5$ y $Var(X_i) = 1.69$, como 50 es suficientemente grande por TCL

$$P\left(\bar{X}_{50} > 18\right) = 1 - P\left(\bar{X}_{50} \leq 18\right) =$$

$$1 - P\left(\frac{\bar{X}_{50} - 17.5}{\sqrt{1.69/50}} \leq \frac{18 - 17.5}{\sqrt{1.69/50}}\right) = 1 - \Phi(2.72) = 0.0033$$

Aproximación de la distribución Binomial por la Normal

Sea $X \sim Bi(n, p)$ se puede ver que entonces $X = X_1 + X_2 + ... + X_n$, con $X_i \sim Ber(p)$ v.a.iid.
Si n es grande, por TCL, $X \approx N(np, np(1-p))$.

Ejemplo Sea $X \sim Bi(100, 0.2)$, $E(X) = 20$ y $Var(X) = 16$, entonces $X \approx N(20, 16)$. Si queremos calcular $P(X \leq 25)$ podemos hacerlo exactamente (sumando las 26 probabilidades puntuales),

$$P(X \leq 25) = P(\frac{X - 20}{4} \leq \frac{25 - 20}{4}) = P(Z \leq 1.25) \approx \Phi(1.25) = 0.8944$$

Corrección por continuidad

La aproximación anterior no es muy buena. El valor real es $P(X \leq 25) = 0.912$. El problema es que estamos aproximando una v.a. discreta por una continua y tenemos que

las probabilidades $P(X \leq 25)$, $P(X \leq 25.5)$, $P(X \leq 25.98)$ son todas iguales pero sus aproximaciones son distintas.
Utilizamos entonces la corrección por continuidad que consiste en elegir el valor del medio de todos los posibles.

$$P(X \leq 25) = P(X \leq 25.5) = P(\frac{X-20}{4} \leq \frac{25.5-20}{4}) = P(Z \leq 1.375) \approx \Phi(1.375) = 0.9147$$

Esto nos permite además calcular aproximadamente una probabilidad puntual de una v.a. discreta, de la siguiente manera:

$$\begin{aligned} P(X = 25) =& P(24.5 \leq X \leq 25.5) = P(\frac{24.5-20}{4} \leq \frac{X-20}{4} \leq \frac{25.5-20}{4}) = \\ &= P(1.125 \leq Z \leq 1.375) \approx \Phi(1.375) - \Phi(1.125) = 0.0461 \end{aligned}$$