ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

BLOQUE ESTADÍSTICA

Jose M. Framiñán Torres

Paz Pérez Gonzalez

Gabriel Villa Caro

Victor Fernández-Viagas Escudero

Curso 2017/2018

Escuela Técnica Superior de Ingeniería

Índice general

Capítulo 1

Introducción a la Teoría de la Probabilidad

1.1. Conceptos Básicos

1.1.1. Experimento aleatorio, espacio muestral y suceso

Definición (Experimento aleatorio): Un **experimento aleatorio** es un experimento en el que:

1. Todos los resultados posibles del experimento se conocen de antemano.
2. La realización del experimento produce un resultado no predecible de antemano.
3. El experimento puede ser repetido bajo condiciones idénticas.

Al resultado de un experimento aleatorio lo denotaremos por ω. Al conjunto de todos los resultados posibles del experimento lo denotaremos por Ω.

Definición (Espacio muestral): Dado un experimento aleatorio, definimos **espacio muestral** E como la colección de todos los subconjuntos que se pueden formar con elementos de Ω. A cada elemento $S \in E$ se le denomina **suceso**.
Si $S = \{\omega\}$, entonces se dice que S es un **suceso elemental**.

Nótese que si Ω es finito, entonces E es finito, y viceversa.

1.1.2. Operaciones con sucesos

1. Intersección de sucesos. $S = S_1 \cap S_2$ es un suceso, y es la ocurrencia de S_1 y S_2. Ver Figura 1.1.

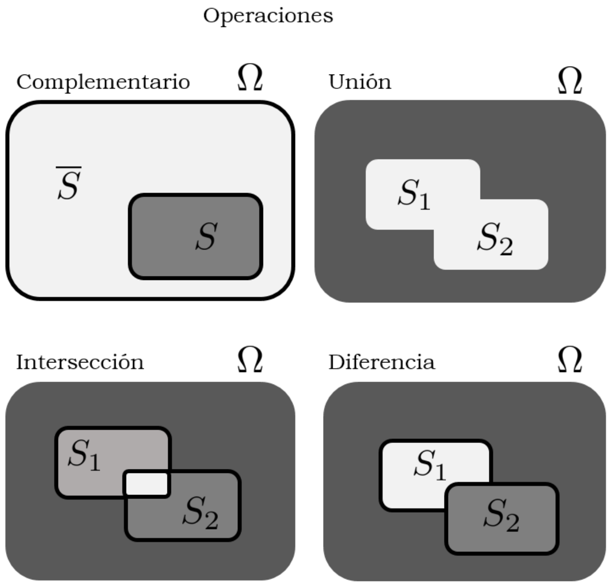

Figura 1.1: Operaciones entre sucesos

Nota: $\Omega \cap \varnothing = \varnothing$

2. Unión de sucesos. $S = S_1 \cup S_2$ es un suceso, y es la ocurrencia de S_1 ó S_2. Ver Figura 1.1.

   Nota: $\Omega \cup \varnothing = \Omega$

3. Suceso complementario. $\overline{S}$ es un suceso, y es la no ocurrencia de S. $S \cap \overline{S} = \varnothing$ y $S \cup \overline{S} = \Omega$. Ver Figura 1.1.

   Nota: $\overline{\Omega} = \varnothing$ y $\overline{\varnothing} = \Omega$

4. Diferencia entre sucesos. $S = S_1 - S_2 = S_1 \cap \overline{S_2}$. Es la ocurrencia de S_1 y no de S_2. Ver Figura 1.1.

1.1.3. Relaciones entre sucesos

Definición (Sucesos disjuntos): Dada una colección numerable de sucesos $\{S_k\}_{k \in K}$, con $K \subseteq \mathbb{N}$ diremos que son **mutuamente excluyentes** o **disjuntos** si $S_k \cap S_l = \varnothing$, $\forall k \neq l$, $k, l \in K$.

En particular, dos sucesos S_1 y S_2 son mutuamente excluyentes si se verifica que $S_1 \cap S_2 = \varnothing$. Ver Figura 1.2.

Definición (Implicación de sucesos): Un suceso S_1 **implica** otro suceso S_2 cuando $S_1 \subseteq S_2$

Como consecuencia de lo anterior, si S_1 implica S_2, entonces $S_1 \cap S_2 = S_1$. Ver Figura 1.2.

Definición (Sucesos exhaustivos): Dada una colección numerable de sucesos $\{S_k\}_{k \in K}$, con $K \subseteq \mathbb{N}$ diremos que son **exhaustivos** si $\bigcup_{k \in K} S_k = \Omega$

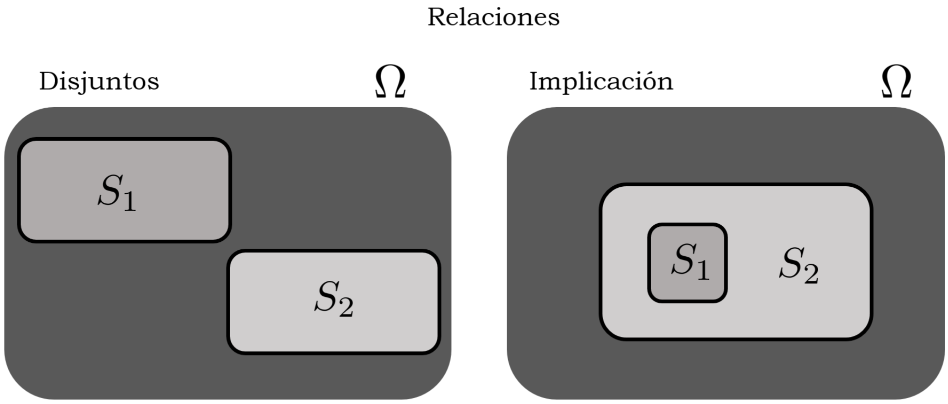

Figura 1.2: Relaciones entre sucesos

Sean $S_1, S_2 \in E$ dos sucesos cualesquiera. Se verifica que

$$S_1 = (S_1 \cap S_2) \cup (S_1 \cap \overline{S_2})$$

(ver figura 1.3)

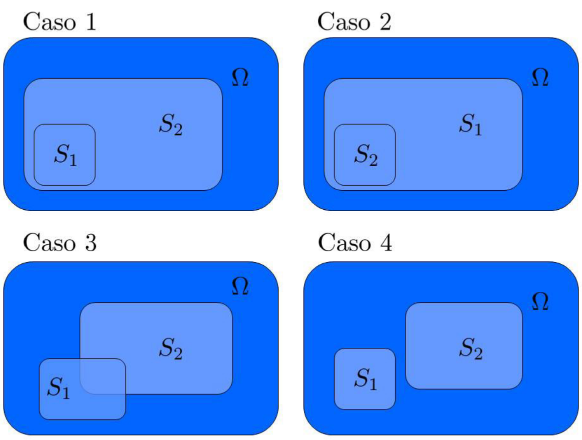

Figura 1.3: $S_1 = (S_1 \cap \overline{S_2}) \cup (S_1 \cap S_2)$

Además, $S_1 \cap S_2$ y $S_1 \cap \overline{S_2}$ son sucesos mutuamente excluyentes.

1.2. Probabilidad

Definición (Probabilidad): Dado el espacio muestral E, si la función $P : E \to \mathbb{R}$ verifica:

1. $P(S) \geq 0, \forall S \in E$
2. $P(\Omega) = 1$
3. Para cualquier colección numerable de sucesos de E, $\{S_k\}$ mutuamente excluyentes, entonces $P(\bigcup S_k) = \sum P(S_k)$ (es decir, P es numerablemente aditiva).

entonces diremos que P es una **medida de probabilidad**, o simplemente **probabilidad**.

1.2.1. Propiedades de la probabilidad

1. $P(\overline{S}) = 1 - P(S), \forall S \in E$
2. $P(S) \leq 1, \forall S \in E$
3. $P(\varnothing) = 0$
4. Si $S_1, S_2 \in E$ con $S_1 \subseteq S_2$, entonces $P(S_1) \leq P(S_2)$
5. $P(S_1 - S_2) = P(S_1) - P(S_1 \cap S_2), \forall S_1, S_2 \in E$
6. $P(S_1 \cup S_2) = P(S_1) + P(S_2) - P(S_1 \cap S_2), \forall S_1, S_2 \in E$. En general, dados $S_1, \ldots, S_n \in E$, se verifica que $P(\bigcup S_i) = \sum P(S_i) - \sum P(S_i \cap S_j) + \sum P(S_i \cap S_j \cap S_k) - \cdots + (-1)^{n+1} P(S_1 \cap S_2 \cap \cdots \cap S_n)$.

1.2.2. Probabilidad condicionada

Definición (Probabilidad condicionada): Dados A, B dos sucesos con $P(B) > 0$. Se define la **probabilidad** de A **condicionada** a B como

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

La probabilidad de A condicionada a B es la probabilidad de que ocurra A sabiendo que B ha sucedido.

Con la definición de probabilidad condicionada podemos hallar la probabilidad de la intersección, ya que $P(A \cap B) = P(A \mid B)P(B)$.

Algunas consecuencias inmediatas de la definición son:

1. En general, $P(A \mid B) \neq P(B \mid A)$
2. Si $P(B) = 1 \Rightarrow P(A \mid B) = P(A \cap B)$

3. $P(A \cap B) \leq P(A \mid B)$

4. $P(A \mid B) + P(\bar{A} \mid B) = 1$

5. Si $B \subseteq A \Rightarrow P(A \mid B) = 1$

La definición de la probabilidad condicionada se puede aplicar de forma recursiva dando lugar a la denominada **regla de la multiplicación**.

Regla de la multiplicación: Dados los sucesos $A_1, \ldots, A_n$ tales que $P(A_1 \cap A_2 \cap \cdots \cap A_n) > 0$, entonces

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots P(A_n \mid A_1 \cap \cdots \cap A_{n-1})$$

1.2.3. Probabilidad total. Fórmula de Bayes

Teorema de la probabilidad total:

Dada una colección numerable de sucesos $\{A_k\}_{k \in K}$, con $K \subseteq \mathbb{N}$ exhaustivos y mutuamente excluyentes, es decir,

1. $\bigcup_{k \in K} A_k = \Omega$

2. $A_k \cap A_l = \varnothing$, $\forall k \neq l$, $k, l \in K$

y con $P(A_k) > 0$, $\forall k \in K$, se verifica para todo suceso A que

$$P(A) = \sum_{k \in K} P(A \mid A_k)P(A_k)$$

Fórmula de Bayes: Dada una colección numerable de sucesos $\{A_k\}_{k \in K}$, con $K \subseteq \mathbb{N}$ exhaustivos y mutuamente excluyentes, es decir,

1. $\bigcup_{k \in K} A_k = \Omega$

2. $A_k \cap A_l = \varnothing$, $\forall k \neq l$, $k, l \in K$

y con $P(A_k) > 0$, $\forall k \in K$, y dado A con $P(A) > 0$, se verifica

$$P(A_k \mid A) = \frac{P(A_k \cap A)}{P(A)} = \frac{P(A \mid A_k)P(A_k)}{\sum_{k \in K} P(A \mid A_k)P(A_k)} \quad \forall k \in K$$

1.2.4. Independencia

Definición (Sucesos independientes): Diremos que dos **sucesos** A, B son **independientes** cuando

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

En este caso, si $P(B) > 0$, entonces $P(A \mid B) = P(A)$, es decir, la probabilidad de A no cambia si sucede B o no.

1.3. Sucesos equiprobables

Regla de Laplace: Sea Ω el conjunto finito de resultados de un experimento aleatorio. Si P probabilidad verifica que $P(\{\omega\}) = c$, $\forall\{\omega\} \in \Omega$, con c constante, entonces

$$P(S) = \frac{|S|}{|\Omega|} \quad \forall S \in \Omega$$

con $|S|$ el número de sucesos elementales en S.

1.4. Ejercicios Tema 1

Ejercicio 1.1. *Dados S_1 y S_2 dos sucesos de E, comprobar gráficamente que se verifican las denominadas **Leyes de Morgan**:*

1. $\overline{S_1 \cup S_2} = \overline{S_1} \cap \overline{S_2}$.

2. $\overline{S_1 \cap S_2} = \overline{S_1} \cup \overline{S_2}$.

Ejercicio 1.2. 1. *Dados dos sucesos A y B, con $P(B) < 1$, demostrar que*

$$P(A \mid \overline{B}) = \frac{P(A) - P(A \cap B)}{1 - P(B)}$$

2. *Si $A \subset B$, hallar $P(A|B) + P(A|\bar{B})$.*

Ejercicio 1.3. *Dada una colección numerable de sucesos $\{A_k\}_{k \in K}$, con $K \subseteq \mathbb{N}$ exhaustivos y mutuamente excluyentes, con $P(A_k) > 0$, $\forall k \in K$, y dado A con $P(A) > 0$ (condiciones de la fórmula de Bayes), demostrar que:*

$$\sum_{k \in K} P(A_k \mid A) = 1$$

Ejercicio 1.4. *Sean A y B dos sucesos independientes.*

1. *Utilizando una de las leyes de Morgan (ver cuestión 1.1), calcular $P(\overline{A} \cap \overline{B})$ en función de $P(A)$ y $P(B)$.*

2. *Calcular $P(\overline{A}) \cdot P(\overline{B})$ en función de $P(A)$ y $P(B)$.*

3. *¿Son $\overline{A}$ y $\overline{B}$ independientes? Justificar la respuesta.*

Ejercicio 1.5. *Cada vez que se realiza determinado experimento aleatorio, la probabilidad de que arroje un resultado concreto al que se denomina* ***éxito*** *es p. Por tanto, $q = 1 - p$ es la probabilidad de que no arroje dicho resultado (****fracaso****). Demostrar:*

1. *Si $F_i := \{$El número de fracasos obtenidos antes de obtener un éxito es $i\}$ $(i = 0, 1, \ldots)$, entonces $P(F_i) = q^i p$*

2. *Si $N_i := \{$El número de experimentos realizados antes de obtener un éxito es $i\}$ $(i = 1, 2, \ldots)$, entonces $P(N_i) = q^{i-1} p$*

3. *Si $E5_i := \{$El número de éxitos obtenidos al realizar el experimento 5 veces es $i\}$ $(i = 0, 1, \ldots, 5)$, entonces $P(E5_i) = \binom{5}{i} p^i q^{5-i}$*

4. *Si $En_i := \{$El número de éxitos obtenidos al realizar el experimento n veces es $i\}$ $(i = 0, 1, \ldots, n)$, entonces $P(En_i) = \binom{n}{i} p^i q^{n-i}$*

5. *Si $On_i := \{$El número de fracasos obtenidos al realizar el experimento n veces es $i\}$ $(i = 0, 1, \ldots, n)$, entonces $P(On_i) = \binom{n}{i} p^{n-i} q^i$*

1.5. Problemas Tema 1

Problema 1.1. *** Un estudiante se ha estudiado 10 temas de un libro de 15 temas. Cada tema tiene el mismo número de páginas. Si abre el libro al azar por cualquier página cuatro veces, se pide:*

1. *Probabilidad de que el estudiante se haya estudiado todos los temas de la página por la que ha abierto el libro las cuatro veces.*

2. *Probabilidad de que el estudiante no se haya estudiado los temas de la página por la que ha abierto el libro las cuatro veces.*

3. *Probabilidad de que el estudiante se haya estudiado dos de los temas de la página por la que ha abierto el libro las cuatro veces.*

Resultado

1. 0.1975
2. 0.0123
3. 0.2963

Problema 1.2. *** Un estudiante se ha estudiado 10 temas de un temario de 15 temas. Si el profesor elige 4 temas al azar para que entren en el examen, se pide:*

1. *Probabilidad de que el estudiante se haya estudiado los cuatro temas.*
2. *Probabilidad de que el estudiante no se haya estudiado ninguno de los cuatro temas.*
3. *Probabilidad de que el estudiante se haya estudiado dos de los cuatro temas.*

Resultado

1. 0.1538
2. 0.0037
3. 0.3297

Problema 1.3. *** En un estudio realizado entre 90 productos, se tiene que 30 cumplen la característica A, 30 otra característica B, y 10 verifican los dos tipos de características.*

Determinar que un producto al azar presente:

1. *Exactamente k características de las dos estudiadas, con $k = 0, 1, 2$*
2. *Al menos k características de las dos estudiadas, con $k = 0, 1, 2$.*
3. *No más de k características de las dos estudiadas, con $k = 0, 1, 2$.*

Resultado

1. 0.4444, 0.4444, 0.1111.
2. 1, 0.5555, 0.1111.
3. 0.4444, 0.8888, 1.

Problema 1.4. **** En una bolsa hay 30 tornillos y 20 puntas. Extraemos tres unidades una tras otra.*

1. *Hallar la probabilidad de que el tercer elemento sea un tornillo.*
2. *Hallar la probabilidad de que el tercer elemento sea un tornillo si en cada extracción devolvemos el elemento si es un tornillo, apartándolo en caso contrario.*

3. *En el mismo supuesto anterior, hallar la probabilidad de que los tres elementos sean iguales.*

Resultado

1. 0.5998
2. 0.6099
3. 0.2741

Problema 1.5. **** Una caja I contiene dos piezas buenas y una pieza defectuosa. Una segunda caja II contiene tres piezas buenas y dos defectuosas.*

1. *Se traslada una pieza de I a II y a continuación se extrae una pieza de II que resulta ser buena. Calcular la probabilidad de que la pieza trasladada fuera buena.*
2. *Suponiendo que en lugar de una, se trasladan simultáneamente dos piezas de I a II, y que la pieza extraída de II es defectuosa. ¿Cuál es la probabilidad de que de las dos piezas trasladadas hubiese exactamente una defectuosa?*

Resultado

1. 0.7272
2. 0.75

Problema 1.6. ***** La mayoría de fenómenos físicos se miden empleando pruebas cuya fiabilidad no es total. Así, se emplea informalmente el concepto de 'falso positivo' si la prueba arroja un resultado positivo (detecta el fenómeno) cuando en realidad dicho fenómeno no tiene lugar. De forma análoga, se habla de 'falso negativo' si la prueba arroja un resultado negativo (no detecta el fenómeno) cuando en realidad dicho fenómeno sí tiene lugar.*
Suponga un determinado fenómeno que se da con una probabilidad p y que se mide con una prueba cuyas probabilidades de 'falso positivo' y 'falso negativo' son α y β, respectivamente. Se pide:

1. *Probabilidad de que la prueba detecte el fenómeno si éste ocurre en realidad.*
2. *Probabilidad de que el fenómeno no ocurra y la prueba lo detecte.*
3. *Probabilidad de que el fenómeno no ocurra y la prueba no lo detecte.*

4. *Probabilidad de que la prueba detecte el fenómeno.*

5. *Probabilidad de que no ocurra el fenómeno en realidad aunque la prueba lo ha detectado.*

Resultado

1. $1-\beta$
2. $\alpha(1-p)$
3. $(1-\alpha)\cdot(1-p)$
4. $(1-\beta)p+\alpha(1-p)$
5. $\frac{\alpha(1-p)}{(1-\beta)p+\alpha(1-p)}$

Problema 1.7. ***** Un modelo de avión que sale de la línea de ensamblado debe realizar dos pruebas funcionales A y B. El avión tiene una probabilidad p de superar la prueba funcional A y, si la supera, la probabilidad de superar la prueba B es de p_1. Si no supera la prueba A, la probabilidad de superar la prueba B es de p_2. Se pide:*

1. *Probabilidad de que el modelo supere las pruebas A y B.*

2. *Probabilidad de que supere la prueba B.*

3. *Probabilidad de que haya superado la prueba A sabiendo que superó la prueba B.*

4. *Probabilidad de que no supere ni la prueba A ni la prueba B.*

Resultado

1. $p\cdot p_1$
2. $p\cdot p_1+(1-p)p_2$
3. $\frac{p\cdot p_1}{p\cdot p_1+(1-p)p_2}$
4. $(1-p)(1-p_2)$

Problema 1.1. *En el aeropuerto de Sevilla, se sabe que la probabilidad de que haya niebla si la temperatura bajó de los $0\,^{\circ}\mathrm{C}$ la noche anterior es de 0.8. Si la temperatura no baja de $0\,^{\circ}\mathrm{C}$, dicha probabilidad se reduce a 0.01.*

La probabilidad de que la temperatura baje de cero grados por la noche en el aeropuerto de Sevilla es de 0.02.

Finalmente, se sabe que la probabilidad de que un vuelo despegue cuando hay niebla y la temperatura bajó de $0^o C$ es 0.3.

1. *¿Cuál es la probabilidad de que esta mañana haya niebla y anoche la temperatura haya bajado de cero grados?*
2. *¿Cuál es la probabilidad de que no haya niebla si la temperatura bajó de cero grados? ¿y de que no haya niebla si la temperatura no bajó de cero grados?*
3. *Hallar la probabilidad de que un vuelo despegue, haya niebla y la temperatura baje de los $0^o C$.*
4. *Calcular la probabilidad de que haya niebla.*
5. *Probabilidad de que haya hecho menos de cero grados si al día siguiente hubo niebla.*
6. *Probabilidad de que haya hecho más de cero grados si al día siguiente no hubo niebla.*

Resultado

1. 0.016
2. 0.99
3. 0.0048
4. 0.0258
5. 0.62
6. 0.9959

Capítulo 2

Variable aleatoria unidimensional

2.1. Variables aleatorias reales

Definición (Variable aleatoria real): Sea un experimento aleatorio con un conjunto de resultados Ω y un espacio muestral E. Se define **variable aleatoria real** como la función $X : \Omega \to \mathbb{R}$ que verifica

$$X^{-1}(I) = \{\omega \in \Omega / X(\omega) \in I\} \in E$$

Con ω un resultado de dicho experimento, e $I \subseteq \mathbb{R}$ un intervalo de $\mathbb{R}$.

Es preciso hacer notar que la definición anterior es un caso restringido de la definición formal de variable aleatoria en el que se han asumido una serie de hipótesis. No obstante, estas hipótesis se verifican en todas las aplicaciones prácticas, por lo que se omite una definición más rigurosa.

Simplificando aún más a efectos prácticos, una variable aleatoria es una función que asocia un valor real a cada resultado de un experimento aleatorio y en la que cualquier intervalo en $\mathbb{R}$ corresponde a algún suceso.

Teorema: Dada una variable aleatoria X y una función $g : \mathbb{R} \to \mathbb{R}$, se tiene que $g(X) : \Omega \to \mathbb{R}$ es también una variable aleatoria.

Si tenemos en cuenta que una variable aleatoria es una función que asigna valores a los resultados de un experimento aleatorio, parece intuitivo que una función $g(x)$ que transforma esos valores en otros también lo sea. No obstante, la función g debe cumplir determinadas condiciones, por lo que el enunciado anterior no es completamente riguroso. Sin embargo, a efectos de las aplicaciones que nos interesan dentro del ámbito de la asignatura, asumiremos que cualquier función de $\mathbb{R}$ en $\mathbb{R}$ las cumplirá.

Definición (Probabilidad Inducida): Dada $X : \Omega \rightarrow \mathbb{R}$ una variable aleatoria, se define la **probabilidad inducida** por X a la función $P_X : \mathbb{R} \rightarrow \mathbb{R}$ tal que

$$P_X(I) = P(X^{-1}(I)),\ \forall I \subseteq \mathbb{R}$$

Para simplificar la expresión de probabilidad inducida introducimos la siguiente notación:

$$P[a < X \leq b] = P_X((a,b]) = P(X^{-1}((a,b]))$$

que indica la probabilidad del suceso correspondiente a que la variable aleatoria tome valores en el intervalo $(a,b]$.

Análogamente, es posible extender la notación anterior a otros intervalos o a un punto en $\mathbb{R}$:

- $P[X \leq x] = P_X((-\infty,x]) = P(X^{-1}((-\infty,x]))$, para denotar la probabilidad del suceso correspondiente a que la variable aleatoria tome valores en el intervalo $(-\infty,x]$. Nótese que x es un parámetro, en este caso la cota superior de valores que tomará la variable aleatoria X.
- $P[X < x] = P_X((-\infty,x)) = P(X^{-1}((-\infty,x)))$, para denotar la probabilidad del suceso correspondiente a que la variable aleatoria tome valores en el intervalo $(-\infty,x)$.
- $P[X = x] = P_X(x) = P(X^{-1}(x))$, para denotar la probabilidad del suceso correspondiente a que la variable aleatoria tome el valor x.

2.2. Función de distribución

Definición (Función de distribución): Se define **función de distribución** de una variable aleatoria X a la función $F_X : \mathbb{R} \rightarrow \mathbb{R}$ siguiente:

$$F_X(x) = F(x) = P_X((-\infty,x]) = P[X \leq x],\ \forall x \in \mathbb{R}$$

El valor que toma la función de distribución en x es igual a la probabilidad de que la variable aleatoria X tome valores menores o iguales a x.

2.2.1. Propiedades

La función de distribución tiene las siguientes propiedades:

1. $0 \leq F \leq 1$

$$F(+\infty) = \lim_{x \to +\infty} F(x) = 1$$

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0$$

2. F es monótona creciente.

3. F es continua por la derecha.

4. Existe límite por la izquierda en todo punto, que denotaremos por $F(x^-)$, es decir

$$F(x^-) = \lim_{y \to x^-} F(y)$$

Esta propiedad implica que, si F no es continua, sus puntos de discontinuidad serán de salto.

Definición (Variables aleatorias idénticamente distribuidas): Sean dos **variables aleatorias** X e Y con función de distribución F_X y F_Y respectivamente. Diremos que dichas variables aleatorias están **idénticamente distribuidas** si $F_X(x) = F_Y(x)\ \forall x \in \mathbb{R}$.

2.2.2. Cálculo de probabilidades a partir de la función de distribución

Nótese que las propiedades anteriores permiten calcular cualquier probabilidad asociada a la variable aleatoria, siempre que se conozca su función de distribución. Así, $\forall x, y \in \mathbb{R}$ se tiene:

1. $P[X \leq x] = F(x)$

2. $P[X > x] = 1 - F(x)$

3. $P[X < x] = F(x^-)$

4. $P[X = x] = P[X \leq x] - P[X < x] = F(x) - F(x^-)$

5. $P[x < X \leq y] = P[X \leq y] - P[X \leq x] = F(y) - F(x)$

2.3. Clasificación de las variables aleatorias

2.3.1. Variable Aleatoria Discreta

Si una función de distribución F no es continua, se puede demostrar que el conjunto de discontinuidades (saltos) es numerable. Sea $\{x_k\}_{k \in K}$ el conjunto ordenado de dichos puntos de discontinuidad, es decir $x_k \leq x_{k+1}$, $\forall k \in K$. Diremos que la variable aleatoria X asociada a esa función de distribución es una **variable aleatoria discreta (VAD)**. En ese caso, sin pérdida de generalidad podemos representar la imagen de X como $X(\Omega) = \{x_k\}_{k \in K}$. La función de distribución de las variables aleatorias discretas se expresa de la siguiente forma:

$$F(x) = P[X \leq x] = \sum_{x_k \leq x} P[X = x_k] = \sum_{x_k \leq x} p_k,\ \forall x \in \mathbb{R}$$

Las variables aleatorias discretas cumplen las siguientes propiedades:

1. La magnitud del salto de F en cada punto x_k equivale a la probabilidad de que la variable aleatoria X tome el valor x_k. A este valor se denomina **función de probabilidad** p_k, es decir:

$$p_k = P[X = x_k] = F(x_k) - F(x_k^-) > 0, \ \forall x_k \in X(\Omega)$$

2. $\sum_{\forall x_k} p_k = 1$

3. Para $a, b \in \mathbb{R}$, $P[a \leq X \leq b] = \sum_{a \leq x_k \leq b} p_k$

2.3.2. Variable Aleatoria Continua

Si una función de distribución F es absolutamente continua, diremos que X es una **variable aleatoria continua (VAC)**.

Como F es absolutamente continua, existe una función $f : \mathbb{R} \to \mathbb{R}$ finita, integrable, no negativa tal que

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} f(u)du \quad x \in \mathbb{R}$$

Definimos f como la **función de densidad** de X.

Las variables aleatorias continuas verifican las siguientes propiedades:

1. Como F es continua $\forall x \in \mathbb{R}$, se tiene que $P[X = x] = 0, \ \forall x \in \mathbb{R}$.

2. f verifica que $\int_{-\infty}^{+\infty} f(x)dx = 1$

3. Para $a, b \in \mathbb{R}$, $P[a \leq X \leq b] = \int_a^b f(x)dx$

2.4. Características

2.4.1. Esperanza matemática

La esperanza matemática o media poblacional de una variable aleatoria X es un número real que representa el valor esperado de dicha variable aleatoria. Se denota por $E[X]$ o μ. Su expresión dependerá del tipo de variable aleatoria.

Definición (Esperanza de una VAD): Dada X una variable aleatoria discreta que toma los valores $\{x_k\}_{k \in K}$, con K conjunto numerable, y $p_k = P[X = x_k]$, se define la **esperanza matemática** de X como

$$E[X] = \sum_{\forall x_k} x_k p_k$$

Técnicamente, para que exista la esperanza de una VAD se requiere que la serie sea absolutamente convergente (es decir, $\sum_{\forall x_k} |x_k| p_k < +\infty$). No obstante, a efectos de las aplicaciones que nos interesan dentro del ámbito de la asignatura, asumiremos que siempre se cumplirá.

Definición (Esperanza de una VAC): Dada X una VAC con función de densidad f, la esperanza matemática de X se define como

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

Técnicamente se requiere que la integral sea absolutamente convergente (es decir, $\int_{-\infty}^{\infty} |x| f(x) dx < +\infty$). No obstante, a efectos de las aplicaciones que nos interesan dentro del ámbito de la asignatura, asumiremos que todas las variables aleatorias lo cumplirán.

Teorema: Dada $g : \mathbb{R} \to \mathbb{R}$ una función bajo las condiciones expuestas en 2.1 en la definición de función de una variable aleatoria:

- Si X es una variable aleatoria discreta bajo las condiciones de su definición de esperanza, se tiene que $E[g(X)] = \sum_{\forall x_k} g(x_k) p_k$.
- X es una variable aleatoria continua bajo las condiciones de su definición de esperanza, se tiene que $E[g(X)] = \int_{-\infty}^{+\infty} g(x) f(x) dx$.

Nótese que el teorema anterior permite expresar $E[g(X)]$ en función de la función de probabilidad, o de la de densidad dependiendo del tipo de variable aleatoria, de la variable aleatoria X, y no de la de $g(X)$. Esto permitiría una definición generalizada de la esperanza, siendo $E[X]$ un caso particular para $g(X) = X$.

- $E[ag(X) + bh(X)] = aE[g(X)] + bE[h(X)]$, con $a, b \in \mathbb{R}$
- Para $g(X) = X$ y $h(X) = 1$ tenemos que $E[aX + b] = aE[X] + b$, con $a, b \in \mathbb{R}$
- Para $a = 0$ (ó $g(X) = 0$) y $h(X) = 1$, $E[b] = b$, con $b \in \mathbb{R}$

2.4.2. Varianza matemática

Definición (Varianza): La **varianza matemática** de una variable aleatoria X se define por:

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

Así, la varianza matemática de una variable aleatoria discreta es

$$V[X] = \sum_{\forall x_k} (x_k - E[X])^2 p_k$$

La varianza matemática de una variable aleatoria continua es

$$V[X] = \int_{-\infty}^{\infty} (x - E[X])^2 f(x) dx$$

La varianza mide la distancia cuadrática esperada de la variable aleatoria a su esperanza. La raíz cuadrada positiva de ésta se denomina **desviación típica**, representada por $\sigma[X] = +\sqrt{V[X]}$. La desviación típica expresa esa distancia en las mismas unidades de la variable aleatoria.

La varianza matemática verifica las siguientes propiedades:

1. $V[X] = E[X^2] - (E[X])^2$. Esta propiedad es muy importante a efectos de cálculo de la varianza.
2. $V[aX + b] = a^2 V[X]$

En ocasiones resulta de interés expresar la desviación típica de una VA X en relación a su esperanza. Dicho cociente se denomina **coeficiente de variación** $cv(X)$:

$$cv(X) = \frac{\sigma[X]}{E[X]}$$

2.4.3. Momentos y función generatriz de momentos

Definición (Momento de orden n**)**: Se define el **momento de orden** n, con $n > 0$ de una variable aleatoria X como:

$$m_n = E[X^n]$$

Definición (Función generatriz de momentos): Dada una variable aleatoria X, se define la **función generatriz de momentos** de X como:

$$\varphi(t) = E\left[e^{tX}\right]$$

La función generatriz de momentos (asumiendo que existe), genera los momentos de la variable aleatoria. Es decir, se puede calcular la esperanza de X elevada a cualquier valor natural n a partir de ésta de la siguiente forma:

Teorema: Si φ es n veces derivable en $t = 0$, se verifica que:

$$\varphi^{n)}(0) = E[X^n]$$

Como consecuencia, se tiene que:

1. $E[X] = \varphi'(0)$
2. $V[X] = \varphi''(0) - (\varphi'(0))^2$

Teorema: Dos variables aleatorias están idénticamente distribuidas (siguen la misma distribución) si y solo si tienen la misma función generatriz de momentos.

2.5. Ejercicios Tema 2

Ejercicio 2.1. *La **distribución de Weibull** es una de las distribuciones de tiempo más usadas en la ingeniería, en particular en fiabilidad. Modela la variable aleatoria $X = \{\text{Tiempo entre fallos en un sistema}\}$, cuando la tasa de fallos es proporcional a una potencia del tiempo. Su función de densidad viene dada por:*

$$f(x) = \begin{cases} \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

con $\alpha > 0$, $\beta > 0$.

1. *$\beta < 1$ indica que la tasa de fallos decrece con el tiempo.*
2. *$\beta = 1$ indica que la tasa de fallos es constante en el tiempo.*
3. *$\beta > 1$ indica que la tasa de fallos crece con el tiempo.*

Demostrar que la función de distribución de $X \sim W(\alpha, \beta)$ es

$$F(x) = \begin{cases} 1 - e^{-(x/\alpha)^{\beta}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Ejercicio 2.2. *Una variable aleatoria X se dice que sigue una **distribución de Pareto** si tiene la siguiente función de densidad:*

$$f(x) = \begin{cases} \frac{\alpha x_m^{\alpha}}{x^{\alpha+1}} & x > x_m \\ 0 & x < x_m \end{cases}$$

siendo $\alpha > 0$ y $x_m > 0$

Se pide:

1. *Demostrar que en efecto se trata de una función de densidad.*
2. *Demostrar que $E[X] = \frac{\alpha x_m}{\alpha-1}$ para $\alpha > 1$.*
3. *Demostrar que $V[X] = \frac{\alpha x_m^2}{(\alpha-1)^2(\alpha-2)}$ para $\alpha > 2$.*

Ejercicio 2.3. *Sea X una variable aleatoria, y $\varphi_X(t)$ su función generatriz de momentos. Demostrar:*

1. *$\varphi_X(0) = 1$.*
2. *Si $Y = kX$, entonces $\varphi_Y(t) = \varphi_X(kt)$.*
3. *Si $Y = X - a$, entonces $\varphi_Y(t) = \varphi_X(t)e^{-at}$.*

4. *Si* $Y = X - a$, *entonces* $E[Y^2] = E[X^2] + a^2 - 2aE[X]$.

5. *Si* $Y = aX + b$, *entonces* $\varphi_Y(t) = e^{bt}\varphi_X(at)$.

Ejercicio 2.4. *Sea* X *una variable aleatoria con* $E[X] = \mu$, *y* $V[X] = \sigma^2$. *Si* $Y = \frac{X-\mu}{\sigma}$, *demostrar:*

1. $E[Y] = 0$.

2. $V[Y] = 1$.

Ejercicio 2.5. *Sea* X *una variable aleatoria discreta que toma valores* $\{0, 1, 2, \dots\}$, *con la siguiente función de probabilidad:*

$$p_k = P[X = k] = \frac{1}{3}\left(\frac{2}{3}\right)^k$$

Nota: Usar las siguientes fórmulas:

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}, \quad |r| < 1$$

$$\sum_{k=0}^{n} r^k = \frac{r^{n+1} - r^0}{r - 1}$$

$$\sum_{k=1}^{\infty} k r^{k-1} = \frac{1}{(1-r)^2}, \quad |r| < 1$$

1. *Demostrar que* $\sum_{k=0}^{\infty} p_k = 1$.

2. *Demostrar que la función de distribución de* X *es*

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - \left(\frac{2}{3}\right)^{Ent[x]+1} & x \geq 0 \end{cases}$$

   con $Ent[x]$ *la parte entera de* $x \in \mathbb{R}$.

3. *Usando la definición de esperanza, demostrar que* $E[X] = 2$.

4. *Demostrar que* $\varphi_X(t) = \frac{1}{3-2e^t}$.

5. *Comprobar que* $E[X] = 2$ *a partir del apartado anterior, y demostrar que* $V[X] = 6$.

Ejercicio 2.6. *** Se dice que una variable aleatoria es simétrica respecto al origen si* $P[X \geq x] = P[X \leq -x]$, $\forall x \in \mathbb{R}$. *Demostrar, para* X *variable aleatoria simétrica:*

1. *Si* X *es una variable aleatoria discreta,* $P[X = x_k] = P[X = -x_k]$.

2. *Si* X *es una variable aleatoria discreta,* $E[X] = 0$.

3. *Si* X *es una variable aleatoria continua,* $f(x) = f(-x)$.

4. *Si* X *es una variable aleatoria continua,* $E[X] = 0$.

Ejercicio 2.7. *** Sea X una variable aleatoria continua con función de densidad:*

$$f(x) = \begin{cases} x & si\ 0 < x \leq 1 \\ 2-x & si\ 1 \leq x < 2 \\ 0 & en\ otro\ caso \end{cases}$$

1. *Hallar la función de distribución de X.*
2. *Hallar $E[X]$.*
3. *Hallar $V[X]$.*
4. *Hallar $P[0.3 < X \leq 1.5]$ y $P[0.3 < X < 1.5]$*

Ejercicio 2.8. *** Sea X variable aleatoria con función de distribución:*

$$F(x) = \begin{cases} 0 & si\ x < 0 \\ kx^2 & si\ 0 \leq x < k \\ 1 & si\ x \geq k \end{cases}$$

1. *Determinar el valor de k para que X sea continua.*
2. *Hallar $E[X]$.*
3. *Hallar $V[X]$.*
4. *Hallar $P[\frac{1}{3} < X \leq \frac{1}{2}]$ y $P[\frac{1}{3} < X < \frac{1}{2}]$.*

Ejercicio 2.9. *Sea una variable aletoria discreta que toma los valores $\{0, 1, \ldots\}$. Demostrar que $E[X] = \sum_{k=0}^{\infty} P[X > k]$[1].*

Ejercicio 2.10. *Sea X una VA con la siguiente función de probabilidad: $P[X = i] = \frac{k}{2^i}$, con $i = 1, 2, \ldots$. Se pide:*

1. *Encontrar el valor de k que hace que la expresión anterior sea una función de probabilidad.*
2. *Determinar la función generatriz de momentos de X.*
3. *Determinar $E[X]$ y $V[X]$.*

Ejercicio 2.11. *** La función de densidad de una variable aleatoria X es*

$$f(x) = \begin{cases} \frac{1}{2}e^{-\frac{x}{2}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

[1] A esta propiedad se la denomina *suma de la cola*

1. *¿Cuál es la función de distribución de X?*
2. *¿Cuál es la probabilidad de que X esté entre 1 y 2?*
3. *¿Cuál es la esperanza de X?*
4. *¿Cuál es la varianza de X?*

2.6. Problemas Tema 2

Problema 2.1. *Sea X una variable aleatoria discreta que toma los valores $\{0,1,2,3\}$ y de la que se conoce*

$$\begin{array}{cc} P[X<1]=\frac{1}{4} & P[1\leq X\leq 2]=\frac{1}{2} \\ P[\frac{1}{2}<X\leq 3]=\frac{3}{4} & E[X]=\frac{11}{8} \end{array}$$

1. *Determinar la función de probabilidad de X*
2. *Determinar la función de distribución de X.*
3. *Determinar la función de probabilidad de $Y=|X-1|+X$.*
4. *Hallar $E[Y]$*

Resultado

1. $P[X=0]=\frac{1}{4}$ $\quad P[X=1]=\frac{3}{8}$ $\quad P[X=2]=\frac{1}{8}$ $\quad P[X=3]=\frac{1}{4}$
2. La función de distribución es:

$$F(x)=\begin{cases} 0 & si\ x<0 \\ \frac{1}{4} & si\ 0\leq x<1 \\ \frac{5}{8} & si\ 1\leq x<2 \\ \frac{3}{4} & si\ 2\leq x<3 \\ 1 & si\ x\geq 3 \end{cases}$$

3. $P[Y=1]=\frac{5}{8}$ $\quad P[Y=3]=\frac{1}{8}$ $\quad P[Y=5]=\frac{1}{4}$
4. $E[Y]=\frac{9}{4}$

Problema 2.2. *En un control de calidad, la fracción defectuosa del proceso es p, con $p\in(0,1)$.*

1. *Modelar la variable aleatoria que represente el estado de una pieza, y su función de probabilidad.*

2. *Hallar la función de distribución de la variable modelada.*

3. *Hallar el valor esperado del estado de una pieza.*

4. *Hallar su varianza.*

5. *Hallar su función generatriz de momentos.*

Resultado

1. $X = \begin{cases} 0 & \text{pieza buena} \\ 1 & \text{pieza defectuosa} \end{cases}$

   Sea $q = 1 - p$, $p_k = P[X = k] = kp + (1 - k)q$, $k \in \{0, 1\}$.

2. $F(x) = \begin{cases} 0 & x < 0 \\ q & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$

3. $E[X] = p$

4. $V[X] = pq$

5. $\varphi(t) = q + pe^t$

Problema 2.3. *En un control de calidad con fracción defectuosa del proceso* $p \in (0, 1)$*, se cuentan el número de piezas defectuosas* X *para cada lote de* n *piezas. La probabilidad de que haya exactamente* k *piezas defectuosas de las* n *controladas es:*

$$p_k = P[X = k] = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}, \; k = 0, \ldots, n$$

Nota: Para todo x, y *se verifica que*

$$\sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k y^{n-k} = (x+y)^n$$

1. *Usando la expresión de la nota, demostrar que* p_k *es función de probabilidad.*

2. *Usando la expresión de la nota, demostrar que la función generatriz de momentos es* $\varphi(t) = (q + pe^t)^n$*, con* $q = 1 - p$*.*

3. *Usando el apartado anterior, hallar el valor esperado del número de piezas defectuosas en un lote.*

4. *Hallar su varianza.*

Supongamos que la fracción defectuosa del proceso es $p = \frac{1}{100}$. *Si el tamaño de un lote es* $n = 10$, *hallar:*

5. *La probabilidad de que haya tres piezas defectuosas.*
6. *La probabilidad de que haya menos de dos piezas defectuosas.*
7. *La probabilidad de que al menos haya una pieza defectuosa.*
8. *La probabilidad de que haya más de una pieza defectuosa.*
9. *El número esperado de piezas defectuosas.*
10. *¿Cuál es la varianza de esta variable aleatoria?*

Resultado

1. Ver teoría y aplicar la expresión dada.
2. Ver teoría y aplicar la expresión dada.
3. $E[X] = np$
4. $V[X] = npq$
5. 0.0001118
6. 0.9957
7. 0.09562
8. 0.0043
9. 0.1
10. 0.099

Problema 2.4. *En una competición deportiva, la probabilidad de que un equipo empate un partido es* p_1 *y de que gane* p_2. *Si pierde no le dan puntos, si empata le dan un punto, y si gana son dos puntos.*

1. *Modelar la variable aleatoria que represente los puntos que le dan al equipo tras un partido, y su función de probabilidad.*
2. *Hallar la función de distribución de la variable modelada.*
3. *Hallar el valor esperado de los puntos del equipo tras un partido.*
4. *Hallar su varianza.*
5. *Hallar su función generatriz de momentos.*

Resultado

1. $X = \begin{cases} 0 & \text{pierden el partido} \\ 1 & \text{empatan el partido} \\ 2 & \text{ganan el partido} \end{cases}$

   $$p_k = P[X = x_k] = \begin{cases} p_0 = 1 - (p_1 + p_2) & k = 0 \\ p_1 & k = 1 \\ p_2 & k = 2 \end{cases}$$

2. $F(x) = \begin{cases} 0 & x < 0 \\ p_0 & 0 \leq x < 1 \\ p_0 + p_1 = 1 - p_2 & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$

3. $E[X] = p_1 + 2p_2$

4. $V[X] = p_1 + 4p_2 - (p_1 + 2p_2)^2$

5. $\varphi(t) = p_0 + p_1 e^t + p_2 e^{2t}$

Problema 2.5. **** Sea X una variable aleatoria con la siguiente función de probabilidad:*

$$p_k = P[X = k] = \begin{cases} c[N - k + 1] & k = 1, \ldots, N \\ 0 & \textit{en otro caso} \end{cases}$$

con $0 < c < 1$.

1. *Determinar c en función de N, sabiendo que $\sum_{k=1}^{N} k = \frac{N(N+1)}{2}$.*

La variable aleatoria puede emplearse para modelar la popularidad de las N páginas web de un servidor, $X = \{$página web a la que accede un usuario$\}$ (Están ordenadas por popularidad). Así, se asume que la página más popular tendrá una probabilidad de acceso al servidor de cN, la segunda de $c(N-1)$, y, en general, la página en el puesto k del ranking tiene una probabilidad de acceso $c(N-k+1)$.

Para acelerar el acceso, el servidor guarda las x páginas más populares en un tipo especial de memoria (memoria caché).

2. *Calcular, en función de N y x, la probabilidad de accecer a la memoria caché.*

3. *Suponiendo que tenemos $N = 260$ páginas web en el servidor, calcular x para que el 80 % de los accesos sean de páginas en la memoria caché.*

Resultado

1. $c = \frac{2}{N(1+N)}$
2. La probabilidad pedida es la función de distribución de X. $F(x) = P[X \leq x] = \frac{2}{N(N+1)}\left(xN - \frac{x(x+1)}{2} + x\right)$
3. $x = 144$

Problema 2.6. *** La duración en minutos de las llamadas telefónicas a una centralita, T, tiene como función de densidad*

$$f(t) = \begin{cases} \alpha e^{-kt} & si\ t \geq 0 \\ 0 & en\ otro\ caso \end{cases}$$

con $k > 0$

1. *Hallar α para que f sea función de densidad.*
2. *Hallar la función de distribución de T.*
3. *Hallar $E[T]$.*
4. *Hallar $V[T]$.*
5. *Determinar la probabilidad de que una llamada dure más de tres minutos.*
6. *Determinar la probabilidad de que una llamada dure entre tres y seis minutos.*

Resultado

1. $\alpha = k$
2. $F(t) = \begin{cases} 0 & \text{si } t < 0 \\ 1 - e^{-kt} & \text{si } t \geq 0 \end{cases}$
3. $E[T] = \frac{1}{k}$
4. $V[T] = \frac{1}{k^2}$
5. $P[T \geq 3] = e^{-3k}$
6. $P[3 \leq T \leq 6] = e^{-3k} - e^{-6k}$

Problema 2.7. **** Dada la siguiente función dependiendo del parámetro $\theta > 0$:*

$$f(x) = \frac{kx^{1/2}}{\theta^{3/2}} e^{-\left(\frac{x}{\theta}\right)^{3/2}} \quad x \geq 0$$

1. *Hallar k para que f sea función de densidad de una variable aleatoria X.*

2. *Hallar la función de distribución de X.*

3. *Hallar la esperanza de X, sabiendo que $\Gamma(p) = \int_0^{+\infty} x^{p-1}e^{-x}dx$*

Resultado

1. $k = \frac{3}{2}$
2. $F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\left(\frac{x}{\theta}\right)^{3/2}} & x \geq 0 \end{cases}$
3. $E[X] = \theta\Gamma\left(\frac{5}{3}\right)$

Problema 2.8. **** Dada la siguiente función:*

$$f(x) = \begin{cases} \theta\left(\frac{1}{x}\right)^{1-\theta} & 0 < x < 1 \\ 0 & \textit{en caso contrario} \end{cases}$$

1. *Demostrar que $f(x)$ es función de densidad de una variable aleatoria X.*

2. *Hallar la función de distribución de X.*

3. *Hallar la esperanza de X.*

Resultado

1. Ver teoría.
2. $F(x) = \begin{cases} 0 & x \leq 0 \\ x^\theta & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$
3. $E[X] = \frac{\theta}{1+\theta}$

Problema 2.9. **** Dada la siguiente función:*

$$f(x) = \begin{cases} \theta(1-x)^{\theta-1} & 0 < x < 1 \\ 0 & \textit{e.o.c.} \end{cases}$$

1. *Demostrar que $f(x)$ es función de densidad de una variable aleatoria X (se dice que X sigue la distribución beta con parámetros $(1,\theta)$, $X \sim \beta(1,\theta)$, con $\theta > 0$).*
2. *Hallar la función de distribución de X.*

Resultado

1. Ver teoría.
2. $F(x) = \begin{cases} 0 & x \leq 0 \\ 1-(1-x)^\theta & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$

Problema 2.10. *** En un sistema de comunicaciones a través de un determinado canal la probabilidad de que un mensaje llegue correctamente de un extremo a otro del canal es p. Cada vez que el receptor recibe un mensaje, envía una señal de confirmación al emisor (la cual llega correctamente al emisor con probabilidad p). Si el emisor no recibe esa confirmación, vuelve a mandar el mensaje.*

1. *Modelar la variable aleatoria $N = \{$número de veces que el emisor manda un mismo mensaje$\}$?*
2. *Determinar el valor mínimo de p para que el número esperado de mensajes repetidos sea menor o igual que 3. Recordar que $\sum_{k=1}^{\infty} k\rho^{k-1} = \frac{1}{(1-\rho)^2}$ para $|\rho| < 1$.*
3. *Se pretende que el número de mensajes repetidos sea menor o igual que 1 con probabilidad de 0.95. ¿Cuál es el valor mínimo de p para conseguir este objetivo?*

Resultado

1. $p_k = (1-p^2)^k p^2$
2. $p \geq 0.5$
3. $p = 0.881$

Capítulo 3

Variables aleatorias multidimensionales

3.1. Variable Aleatoria n-dimensional

Definición (Variable Aleatoria n-Dimensional): Sea un experimento aleatorio donde Ω es el conjunto de resultados, y ω un resultado.
Dado un conjunto de variables aleatorias $X_1(\omega), \ldots, X_n(\omega)$, el vector:

$$\mathbf{X}(\omega) = (X_1(\omega), \ldots, X_n(\omega))$$

es una **variable aleatoria n-dimensional**, o **vector aleatorio**.

En particular, dadas dos variables aleatorias $X, Y : \Omega \to \mathbb{R}$, $(X, Y) : \Omega \to \mathbb{R}^2$ es una **variable aleatoria bidimensional**.

Teorema: Dada una variable aleatoria n-dimensional $\mathbf{X}$, se tiene que cualquier función $g(\mathbf{X}) : \mathbb{R}^n \to \mathbb{R}$ es una variable aleatoria real.

Al igual que en el caso de una dimensión, si tenemos una función de una colección de variables aleatorias, parece intuitivo que ésta también lo sea. No obstante, la función g debe cumplir determinadas condiciones, por lo que el enunciado anterior no es completamente riguroso. Sin embargo, a efectos de las aplicaciones que nos interesan dentro del ámbito de la asignatura, asumiremos que cualquier función de $\mathbb{R}^n$ en $\mathbb{R}$ las cumplirá.

Por simplicidad en la notación, en los siguientes apartados de este tema nos referiremos al caso de variables aleatorias bidimensionales, si bien la generalización a n dimensiones es inmediata.

3.2. Funciones de distribución conjunta y marginales

Definición (Función de distribución conjunta): Se define **función de distribución conjunta** como la función

$$F(x, y) = P[X \leq x; Y \leq y]$$

con $P[X \leq x; Y \leq y] = P(X^{-1}((-\infty, x]) \cap Y^{-1}((-\infty, y]))$

Nótese que la función de distribución conjunta expresa la probabilidad de que las variables aleatorias X e Y tomen de forma simultánea valores no superiores a x e y.

La función de distribución conjunta verifica:

1. $0 \leq F \leq 1$

$$F(+\infty, +\infty) = lim_{(x,y)\to(+\infty,+\infty)} F(x, y) = 1$$

$$F(-\infty, y) = lim_{x\to-\infty} F(x, y) = 0$$

$$F(x, -\infty) = lim_{y\to-\infty} F(x, y) = 0$$

2. F es monótona creciente para x e y.

3. F es continua por la derecha para x e y.

Definición (Funciones de distribución marginales): Las funciones de distribución de X e Y, $F_X(x)$ y $F_Y(y)$ se denominan **funciones de distribución marginales**.

Se verifica que

$$F_X(x) = F(x, +\infty) = lim_{y\to+\infty} F(x, y)$$

$$F_Y(y) = F(+\infty, y) = lim_{x\to+\infty} F(x, y)$$

3.3. Clasificación de variables aleatorias n-dimensionales

3.3.1. Variable aleatoria bidimensional discreta

Si X e Y son dos variables aleatorias discretas, entonces (X, Y) es una **variable aleatoria bidimensional discreta**. En este caso consideraremos que el conjunto numerable de valores que toma X es $\{x_k\}_{k\in K}$, K conjunto numerable, y el de Y es $\{y_l\}_{l\in L}$, L conjunto numerable. Las funciones de probabilidad asociadas a cada una de ellas son $p_k^X = P[X = x_k]$ y $p_l^Y = P[Y = y_l]$ respectivamente.

Si (X, Y) es una variable aleatoria bidimensional discreta, entonces

$$p_{kl} = P[X = x_k; Y = y_l], \; k \in K, l \in L$$

se denomina **función de probabilidad conjunta** de (X, Y).

En este caso la función de distribución conjunta viene dada por

$$F(x, y) = P[X \leq x; Y \leq y] = \sum_{x_k \leq x} \sum_{y_l \leq y} p_{kl}$$

Las **funciones de probabilidad marginales** de X e Y son respectivamente

$$p_k^X = P[X = x_k] = \sum_{\forall y_l} p_{kl}, \ k \in K$$

$$p_l^Y = P[Y = y_l] = \sum_{\forall x_k} p_{kl}, \ l \in L$$

3.3.2. Variable aleatoria bidimensional continua

(X, Y) es una **variable aleatoria bidimensional continua** si existe $f : \mathbb{R}^2 \to \mathbb{R}$ integrable y no negativa tal que

$$F(x, y) = P[X \leq x; Y \leq y] = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v) du dv$$

Diremos que f es la **función de densidad conjunta** de (X, Y).

Denotamos por $f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$ y $f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$ a las **funciones de densidad marginales** de X e Y respectivamente.

3.4. Características de la variable aleatoria n-dimensional

3.4.1. Esperanza de una función de variable aleatoria bidimensional

Dada $g : \mathbb{R}^2 \to \mathbb{R}$ una función bajo las condiciones expuestas en 3.1:

- Si (X, Y) es una variable aleatoria bidimensional discreta, entonces

  $$E[g(X, Y)] = \sum_{\forall x_k} \sum_{\forall y_l} g(x_k, y_l) p_{kl}$$

  siempre que la serie sea absolutamente convergente. Como en el caso unidimensional, supondremos que todas las variables aleatorias lo cumplirán.

- Si (X, Y) es una variable aleatoria bidimensional continua, entonces

  $$E[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy$$

  siempre que la integral sea absolutamente convergente. Como en el caso unidimensional, supondremos que todas las variables aleatorias lo cumplirán.

Se verifica que:

- $E[ag(X,Y) + bh(X,Y)] = aE[g(X,Y)] + bE[h(X,Y)]$, con $a, b \in \mathbb{R}$
- En particular, $E[X+Y] = E[X] + E[Y]$
- La expresión anterior se puede generalizar para la suma de n variables aleatorias $X_1, \ldots, X_n$: $E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E[X_i]$

3.4.2. Covarianza

Definición (Covarianza): Se define la **covarianza matemática** de una variable aleatoria (X,Y) como $Cov[X,Y] = E[(X - E[X])(Y - E[Y])]$.
Si (X,Y) es una variable aleatoria bidimensional discreta, entonces

$$Cov[X,Y] = \sum_{\forall x_k} \sum_{\forall y_l} (x_k - E[X])(y_l - E[Y]) p_{kl}$$

Si (X,Y) es una variable aleatoria bidimensional continua, entonces

$$Cov[X,Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - E[X])(y - E[Y]) f(x,y) dxdy$$

La covarianza verifica las siguientes propiedades:

1. $Cov[X,X] = V[X]$
2. $Cov[X,Y] = E[XY] - E[X]E[Y]$
3. $Cov[X,Y] = Cov[Y,X]$
4. $V[X+Y] = V[X] + V[Y] + 2Cov[X,Y]$
5. De la propiedad anterior, se puede ver que, en general, para n variables aleatorias $X_1, \ldots, X_n$
$$V\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n V[X_i] + 2\sum_{i<j} Cov[X_i, X_j]$$
6. $Cov[X,Y]^2 \leq V[X] \cdot V[Y]$

En ocasiones resulta de interés representar la covarianza de forma normalizada respecto a las desviaciones típicas de las variables aleatorias. A dicho ratio se le denomina **coeficiente de correlacion** $\rho(X,Y)$:

$$\rho(X,Y) = \frac{Cov[X,Y]}{\sigma_X \cdot \sigma_Y} = \frac{Cov[X,Y]}{\sqrt{V[X] \cdot V[Y]}}$$

A la vista de la propiedad anterior de la covarianza, se puede ver directamente que el coeficiente de correlación está entre -1 y 1.

3.5. Distribuciones condicionadas

Definición (Función de probabilidad condicionada): Dadas dos VAD X e Y, se define la **función de probabilidad de X condicionada a $Y = y_l$**, para $y_l \in \mathbb{R}$ tal que $P[Y = y_l] > 0$ como:

$$P[X = x_k | Y = y_l] = \frac{P[X = x_k; Y = y_l]}{P[Y = y_l]} = \frac{p_{kl}}{p_l}, \ \forall k \in K$$

Nótese que la función de probabilidad anterior toma valores **en el rango** de X (es decir, para todos los x_k) **para un valor fijo** de Y ($Y = y_l$). Es fácil ver que $\sum_{\forall x_k} P[X = x_k | Y = y_l] = 1$.

Análogamente, la **función de probabilidad de Y condicionada a $X = x_k$**, con $x_k \in \mathbb{R}$ tal que $P[X = x_k] > 0$ es:

$$P[Y = y_l | X = x_k] = \frac{P[X = x_k; Y = y_l]}{P[X = x_k]} = \frac{p_{kl}}{p_k}, \ \forall l \in L$$

Definición (Función de densidad condicionada): Dadas dos VAC X e Y, se define la **función de densidad** de X **condicionada** a $Y = y$ como:

$$f_{[X|Y=y]}(x) = \frac{f(x,y)}{f_Y(y)}, \ \forall x \in \mathbb{R}, \ y \in \mathbb{R}/f_Y(y) > 0$$

Análogamente, la función de densidad de Y condicionada a $X = x$ viene dada por:

$$f_{[Y|X=x]}(y) = \frac{f(x,y)}{f_X(x)}, \ \forall y \in \mathbb{R}, \ x \in \mathbb{R}/f_X(x) > 0$$

3.6. Independencia de variables aleatorias

Definición (Independencia de variables aleatorias): Dos variables aleatorias X e Y son **independientes** si para todo $I_1, I_2 \subseteq \mathbb{R}$ se verifica que

$$P(X^{-1}(I_1) \cap Y^{-1}(I_2)) = P(X^{-1}(I_1))P(Y^{-1}(I_2))$$

que es equivalente a que

$$P[X \in I_1; Y \in I_2] = P[X \in I_1]P[Y \in I_2]$$

Las variables aleatorias independientes cumplen las siguientes propiedades:

1. Dada una variable aleatoria bidimensional (X, Y), con función de distribución conjunta F, y F_X y F_Y las funciones de distribuciones marginales de X e Y respectivamente, se verifica que X e Y son independientes si y sólo si

$$F(x,y) = F_X(x)F_Y(y), \ \forall x, y \in \mathbb{R}$$

2. En el caso de variable aleatoria bidimensional discreta, se verifica que X e Y son independientes si y sólo si

$$P[X = x_k; Y = y_l] = P[X = x_k]P[Y = y_l],\ \forall x_k, y_l$$

3. En el caso de variable aleatoria bidimensional continua, se verifica que X e Y son independientes si y sólo si

$$f(x,y) = f_X(x)f_Y(y),\ \forall x, y \in \mathbb{R}$$

4. Dadas dos variables aleatorias X e Y independientes, entonces

$$E[XY] = E[X]E[Y]$$

Puede demostrarse que esta propiedad es condición necesaria de independencia, pero no suficiente.

5. Dadas dos variables aleatorias X e Y independientes, entonces, en general, se verifica que

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$

En particular, se tiene que para X e Y independientes, entonces

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$$

6. Dadas dos variables aleatorias X e Y independientes, entonces

$$V[X+Y] = V[X] + V[Y]$$

La independencia se puede generalizar para n variables aleatorias. Diremos que n **variables aleatorias** $X_1, \ldots, X_n$ son **mutuamente independientes** si

$$F(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i),\ \forall x_1, \ldots, x_n \in \mathbb{R}$$

Finalmente, diremos que n variables aleatorias $X_1, \ldots, X_n$ son **independientes e idénticamente distribuidas** (iid) si son mutuamente independientes y están idénticamente distribuidas.

3.7. Ejercicios Tema 3

Ejercicio 3.1. *Dadas X e Y dos variables aleatorias independientes, con $P[X+Y=m] \neq 0$, demostrar que*

$$P[X = k|X+Y=m] = \frac{P[Y=m-k]P[X=k]}{P[X+Y=m]}$$

Ejercicio 3.2. *Dada dos variables aleatorias X e Y independientes, demostrar que*

$$V[XY] = V[X]V[Y] + V[X]E[Y]^2 + V[Y]E[X]^2$$

Para ello, demostrar las siguientes igualdades:

1. $V[XY] = E[X^2]E[Y^2] - E^2[X]E^2[Y]$

2. $V[X]V[Y] + V[X]E[Y]^2 + V[Y]E[X]^2 = E[X^2]E[Y^2] - E^2[X]E^2[Y]$

Ejercicio 3.3. *Sea (X, Y) una variable aleatoria bidimensional con la siguiente función de probabilidad conjunta:*

	Y		
X	1	2	3
1	0.1	0.3	0
2	0	0	0.2
3	0.1	0.1	0
4	0	0.2	x

1. *Determinar el valor de x.*

2. *Determinar la esperanza y la varianza de X e Y.*

3. *Hallar $P[X > 2; Y > 2]$ y $P[X < 2; Y > 3]$*

4. *Función de probabilidad de X condicionada a $Y = 3$*

5. *Hallar $P[Y = 2|X = 1]$*

6. *¿Son X e Y independientes?*

7. *Hallar $Cov[X, Y]$.*

Ejercicio 3.4. *Sea (X, Y) una variable aleatoria bidimensional discreta, cuya función de probabilidad conjunta está dada en la siguiente tabla:*

$X\backslash Y$	1	2	3
-1	$t/2$	$2z$	$2t$
0	z	$3z$	$4z$
1	$t/4$	z	t

1. *Determinar z y t de forma que X e Y sean independientes.*

2. *Hallar la $Cov[X, Y]$ en función de t.*

3. *Se dice que X e Y están incorreladas si $Cov[X, Y] = 0$. Hallar los valores de t para los que X e Y están incorreladas.*

4. *Con los resultados obtenidos, comprobar que si X e Y son independientes entonces están incorreladas (Nota: el recíproco no es cierto en general).*

5. *Determinar la distribución de Y condicionada a $X = 1$ para cualquier valor de z y t.*

Ejercicio 3.5. *Sea la variable aleatoria bidimensional continua (X, Y) con función de densidad conjunta:*

$$f(x,y) = \begin{cases} 1 & (x,y) \in D \\ 0 & \textit{en otro caso} \end{cases}$$

con $D = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 2; 0 \leq y \leq 1; y \leq -\frac{x}{2} + 1\}$

1. *Demostrar que f es función de densidad.*
2. *Hallar las funciones de densidad marginales.*
3. *Hallar las funciones de densidad condicionadas.*

Ejercicio 3.6. *Sea (X, Y) una variable aleatoria bidimensional continua con función de densidad conjunta:*

$$f(x,y) = \begin{cases} 0 & \textit{si } (x,y) \textit{ está fuera del triángulo } T \\ k & \textit{si } (x,y) \textit{ está en el triángulo } T \end{cases}$$

Con T triángulo rectángulo que tiene como vértices los puntos $(0,0), (0,1), (1,1)$ y tiene el ángulo recto en el vértice $(0,1)$

1. *Hallar k para que f sea función de densidad conjunta.*
2. *Hallar las funciones de densidad marginales.*
3. *Hallar las funciones de distribución marginales.*
4. *Hallar las funciones de densidad condicionadas.*

Ejercicio 3.7. **** Sea (X, Y) una variable aleatoria bidimensional continua con función de densidad conjunta:*

$$f(x,y) = \begin{cases} 0 & \textit{si } (x,y) \textit{ está fuera del triángulo } T \\ k(x+y) & \textit{si } (x,y) \textit{ está en el triángulo } T \end{cases}$$

Con T triángulo rectángulo que tiene como vértices los puntos $(0,0), (0,1), (1,0)$ y tiene el ángulo recto en el vértice $(0,0)$

1. *Hallar k para que f sea función de densidad conjunta.*
2. *Hallar las funciones de densidad marginales.*
3. *Hallar las funciones de distribución marginales.*
4. *Hallar las funciones de distribución condicionadas.*

3.8. Problemas Tema 3

Problema 3.1. **** Tenemos tres monedas trucadas. Las probabilidades de cara de cada una son 0.5, 0.4 y 0.3 respectivamente. Sea* $X = \{$*número de caras (c) en las primeras dos monedas*$\}$ *e* $Y = \{$*número de cruces (x) en las últimas dos monedas*$\}$.

1. *Hallar las probabilidades de cada resultado posible.*
2. *Hallar la función de probabilidad conjunta de la variable bidimensional* (X, Y).
3. *Hallar la función de probabilidad marginal de* X.
4. *Hallar la función de probabilidad marginal de* Y.
5. *Hallar la función de distribución conjunta de* (X, Y).
6. *Hallar la función de distribución marginal de* X *y la de* Y.
7. *Hallar la función de probabilidad de* X *condicionada a* Y.
8. *Hallar la covarianza de* X *e* Y.
9. *¿Son* X *e* Y *independientes?*

Resultado

Resultado	Probabilidad	X	Y
$\{c, c, c\}$	$0.5 \cdot 0.4 \cdot 0.3 = 0.06$	2	0
$\{c, c, x\}$	$0.5 \cdot 0.4 \cdot 0.7 = 0.14$	2	1
$\{c, x, c\}$	$0.5 \cdot 0.6 \cdot 0.3 = 0.09$	1	1
$\{c, x, x\}$	$0.5 \cdot 0.6 \cdot 0.7 = 0.21$	1	2
$\{x, c, c\}$	$0.5 \cdot 0.4 \cdot 0.3 = 0.06$	1	0
$\{x, c, x\}$	$0.5 \cdot 0.4 \cdot 0.7 = 0.14$	1	1
$\{x, x, c\}$	$0.5 \cdot 0.6 \cdot 0.3 = 0.09$	0	1
$\{x, x, x\}$	$0.5 \cdot 0.6 \cdot 0.7 = 0.21$	0	2

1.

2.

	Y		
X	0	1	2
0	0.00	0.09	0.21
1	0.06	0.23	0.21
2	0.06	0.14	0.00

3. $P[X = 0] = 0.3 \quad P[X = 1] = 0.5 \quad P[X = 2] = 0.2$

4. $P[Y = 0] = 0.12 \quad P[Y = 1] = 0.46 \quad P[Y = 2] = 0.42$

5.

$$F(x, y) = P[X \le x, Y \le y] = \begin{cases} 0 & 0 \le x < 1, 0 \le y < 1; x \in \mathbb{R}, y < 0; x < 0, y \in \mathbb{R} \\ 0.06 & 1 \le x < 2, 0 \le y < 1 \\ 0.12 & x \ge 2, 0 \le y < 1 \\ 0.09 & 0 \le x < 1, 1 \le y < 2 \\ 0.38 & 1 \le x < 2, 1 \le y < 2 \\ 0.58 & x \ge 2, 1 \le y < 2 \\ 0.30 & 0 \le x < 1, y \ge 2 \\ 0.8 & 1 \le x < 2, y \ge 2 \\ 1 & x \ge 2, y \ge 2 \end{cases}$$

6. $F_X(x) = \begin{cases} 0 & x < 0 \\ 0.3 & 0 \le x < 1 \\ 0.8 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ 0.12 & 0 \le y < 1 \\ 0.58 & 1 \le y < 2 \\ 1 & y \ge 2 \end{cases}$$

7. $$\begin{array}{ccc} P[X = 0|Y = 0] = 0 & P[X = 1|Y = 0] = 0.5 & P[X = 2|Y = 0] = 0.5 \\ P[X = 0|Y = 1] = 0.1957 & P[X = 1|Y = 1] = 0.5 & P[X = 2|Y = 1] = 0.3043 \\ P[X = 0|Y = 2] = 0.5 & P[X = 1|Y = 2] = 0.5 & P[X = 2|Y = 2] = 0 \end{array}$$

8. $Cov[X, Y] = -0.24$

9. No son independientes. Ver teoría.

Problema 3.2. *Una fábrica realiza un control de calidad. Se toma una muestra de N piezas. Sea d el número de piezas defectuosas que hay en la muestra. Se extraen dos piezas al azar y se comprueba si son defectuosas o no.*

1. *Si hay reemplazamiento:*

a) *Hallar la probabilidad de que la primera pieza sea buena y la segunda defectuosa.*

b) *Hallar la probabilidad de que la segunda pieza sea buena si la primera ha sido defectuosa.*

c) *Hallar la función de probabilidad del número de piezas defectuosas en la extracción.*

d) *Hallar la esperanza del número de piezas defectuosas en la extracción.*

e) *Hallar la varianza del número de piezas defectuosas en la extracción.*

2. *Si no hay reemplazamiento:*

f) *Hallar la probabilidad de que la primera pieza sea buena y la segunda defectuosa.*

g) *Hallar la probabilidad de que la primera pieza sea buena si la segunda es defectuosa.*

h) *Hallar la función de probabilidad conjunta de las variables aleatorias que dan el resultado de la primera y la segunda pieza en la extracción.*

i) *Hallar las funciones de probabilidad marginales de las variables aleatorias que dan el resultado de la primera y la segunda pieza en la extracción.*

j) *Hallar la función de probabilidad del número de piezas defectuosas en la extracción.*

k) *Hallar la esperanza del número de piezas defectuosas en la extracción.*

l) *Hallar la varianza del número de piezas defectuosas en la extracción.*

Resultado

Sean

$$X = \begin{cases} 0 & \text{primera pieza buena} \\ 1 & \text{primera pieza defectuosa} \end{cases}$$

$$Y = \begin{cases} 0 & \text{segunda pieza buena} \\ 1 & \text{segunda pieza defectuosa} \end{cases}$$

1. Primer caso:

   a) $P[X = 0; Y = 1] = \frac{N-d}{N}\frac{d}{N}$

   b) $P[Y = 0|X = 1] = \frac{N-d}{N}$

   c) $Z = X + Y$={Número de piezas defectuosas}

   $P[Z = 0] = \frac{(N-d)^2}{N^2} \quad P[Z = 1] = \frac{2d(N-d)}{N^2} \quad P[Z = 2] = \frac{d^2}{N^2}$

   d) $E[Z] = \frac{2d}{N}$

   e) $V[Z] = \frac{2d(N-d)}{N^2}$

2. Segundo caso:

   f) $P[X = 0; Y = 1] = \frac{d}{N-1}\frac{N-d}{N}$

   g) $P[X = 0|Y = 1] = \frac{N-d}{N-1}$

h) $P[X=0;Y=0]=\frac{N-d-1}{N-1}\frac{N-d}{N}$ $\quad P[X=0;Y=1]=\frac{d}{N-1}\frac{N-d}{N}$

$P[X=1;Y=0]=\frac{N-d}{N-1}\frac{d}{N}$ $\quad P[X=1;Y=1]=\frac{d-1}{N-1}\frac{d}{N}$

i) $P[X=0]=P[Y=0]=\frac{N-d}{N}$; $P[X=1]=P[Y=1]=\frac{d}{N}$

j) $P[Z=0]=\frac{(N-d)(N-d-1)}{N(N-1)}$ $\quad P[Z=1]=\frac{2d(N-d)}{N(N-1)}$ $\quad P[Z=2]=\frac{d(d-1)}{N(N-1)}$

k) $E[Z]=2\frac{d}{N}$

l) $V[Z]=\frac{2d(N+d-2)}{N(N-1)}-4\frac{d^2}{N^2}$

Problema 3.3. *Sea X una variable aleatoria discreta que toma valores $\{-1,0,1\}$ con probabilidades $p_X(x)=P[X=x]=ax^2+b$.*

1. *Sabiendo que la probabilidad de que X sea uno es $\frac{1}{2}$, hallar los valores de a y b.*
2. *Hallar $E[X]$ y $V[X]$.*
3. *Hallar $\varphi(t)$. ¿Cuánto vale $E[X^n]$ si n es par? ¿Y si n es impar?*

Sea Y otra variable aleatoria discreta que toma valores $\{1,4,9\}$, de forma que la función de probabilidad conjunta de (X,Y) es

$$p_{XY}(x,y)=P[X=x;Y=y]=\begin{cases} cx^2\sqrt{y} & x=-1,0,1;\ y=1,4,9 \\ 0 & \textit{en otro caso} \end{cases}$$

4. *Hallar la función de probabilidad de Y.*
5. *Hallar $E[\sqrt{Y}]$.*
6. *Hallar $E[XY]$. ¿Cuánto vale $E[X^nY]$ para n impar?*
7. *Hallar $Cov[X,Y]$.*
8. *Hallar $p_{X|Y}(x)$.*

Resultado

1. $p_X(x)=\frac{1}{2}x^2$
2. $E[X]=0$; $V[X]=1$
3. $\varphi(t)=\frac{1}{2}(e^{-t}+e^{t})$; $E[X^n]=0$ cuando n es impar y $E[X^n]=1$ cuando es par.
4. $p_Y(y)=\frac{1}{6}\sqrt{y}$
5. $E[\sqrt{Y}]=\frac{7}{3}$

6. $E[XY] = 0$
7. $Cov[X, Y] = 0$
8. $p_{X|Y=y_0}(x) = \frac{1}{2}x^2$ cuando $y_0 \in \{1, 4, 9\}$

Problema 3.4. ****** Se tienen 9 piezas, 5 de ellas son buenas y 4 defectuosas. Se realizan dos extracciones sin reemplazamiento. Si se definen dos variables aleatorias $X = \{resultado\ de\ la\ primera\ extracción\}$, e $Y = \{resultado\ de\ la\ segunda\ extracción\}$, se pide:*

1. *Función de probabilidad de X.*
2. *Funciones de probabilidad condicionadas de $Y|X$.*
3. *Función de probabilidad de Y.*
4. *Función de probabilidad conjunta de la variable bidimensional (X, Y).*
5. *Funciones de probabilidad condicionadas de $X|Y$.*

Definiendo los valores que toman ambas variables como 0 si la pieza es buena y 1 si es defectuosa, calcular:

6. *Varianzas de X e Y.*
7. *Covarianza de la variable $(X, Y) : Cov(X, Y)$.*

Resultado

1. $X = \{\text{Resultado de la primera extracción}\} = \begin{cases} 0 & \text{pieza buena} \\ 1 & \text{pieza defectuosa} \end{cases}$

   $P[X = 0] = 0.556$; $P[X = 1] = 0.444$

2. $Y = \{\text{Resultado de la segunda extracción}\} = \begin{cases} 0 & \text{pieza buena} \\ 1 & \text{pieza defectuosa} \end{cases}.$

   $P[Y = 0 \mid X = 0] = 0.5 \qquad P[Y = 1 \mid X = 0] = 0.5$
   $P[Y = 0 \mid X = 1] = 0.625 \qquad P[Y = 1 \mid X = 1] = 0.375$

3. $P[Y = 0] = 0.556 \qquad P[Y = 1] = 0.444$

4. Las probabilidades conjuntas son:

   $P[Y = 0; X = 0] = 0.278 \qquad P[Y = 1; X = 0] = 0.278$
   $P[Y = 0; X = 1] = 0.278 \qquad P[Y = 1; X = 1] = 0.167$

5. Las probabilidades condicionadas son:

$$P[X = 0 \mid Y = 0] = 0.5 \qquad P[X = 1 \mid Y = 0] = 0.5$$
$$P[X = 0 \mid Y = 1] = 0.6247 \quad P[X = 1 \mid Y = 1] = 0.3753$$

6. $E[X] = E[Y] = 0.444$; $V[X] = V[Y] = 0.247$

7. $Cov[X, Y] = -0.03$

Problema 3.5. ****** Sea (X, Y) una variable aleatoria bidimensional continua.*

1. *Hallar k para que la función $f(x, y)$ sea función de densidad conjunta de (X, Y).*

$$f(x,y) = \begin{cases} 0 & (x,y) \in A \\ kx^2y & (x,y) \in A \end{cases}$$

   Con $A = \{(x, y) : -1 \le x \le 1,\ 0 \le x^2 < y < 1\}$

2. *Hallar las funciones de densidad marginales de X y de Y.*

3. *Hallar las funciones de distribución marginales de X e Y.*

4. *Hallar $E[X]$ y $V[Y]$.*

5. *Hallar $E[X^n]$.*

6. *Hallar $E[\sqrt{Y}]$.*

7. *Hallar $E[XY]$ y $Cov[X, Y]$. ¿Son X e Y independientes? Razonar la respuesta.*

8. *Hallar $E[X^nY]$.*

9. *Hallar $P[X \le x; Y \le y]$ con $-1 \le x \le 1$, $0 \le x^2 < y < 1$*

10. *Hallar $P[X \le x | Y \le y]$ con $-1 \le x \le 1$, $0 \le x^2 < y < 1$*

Resultado

1. Función de densidad conjunta:

$$f(x,y) = \begin{cases} 0 & (x,y) \in A \\ \frac{21}{4}x^2y & (x,y) \in A \end{cases}$$

2. Funciones de densidad marginales:

$$f_X(x) = \begin{cases} 0 & x < -1 \text{ ó } x > 1 \\ \frac{21}{8}x^2(1-x^4) & -1 \le x \le 1 \end{cases}$$

$$f_Y(y) = \begin{cases} 0 & y \leq 0 \text{ ó } y \geq 1 \\ \frac{7}{2}y^{5/2} & 0 < y < 1 \end{cases}$$

3. Funciones de distribución marginales

$$F_X(x) = \begin{cases} 0 & x < -1 \\ \frac{7x^3 - 3x^7}{8} + \frac{1}{2} & -1 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

$$F_Y(y) = \begin{cases} 0 & y \leq 0 \\ y^{7/2} & 0 < y < 1 \\ 1 & y \geq 1 \end{cases}$$

4. $E[X] = 0$, $V[Y] = \frac{28}{891}$

5. Momento de orden n de X:

$$E[X^n] = \begin{cases} 0 & n \text{ impar} \\ \frac{21}{2(n+3)(n+5)} & n \text{ par} \end{cases}$$

6. $E[\sqrt{Y}] = \frac{7}{8}$

7. $E[XY] = 0$, $Cov[X,Y] = 0$. No, ya que $f(x,y) \neq f_X(x) \cdot f_Y(y)$

8. La esperanza pedida es

$$E[X^n Y] = \begin{cases} 0 & n \text{ impar} \\ \frac{21}{(n+3)(n+9)} & n \text{ par} \end{cases}$$

9. $(x,y) \in A$ $P[X \leq x; Y \leq y] = \frac{7}{8}(y^2 - x^4)(x^3 + y^{3/2})$

10. $(x,y) \in A$ $P[X \leq x | Y \leq y] = \frac{\frac{7}{8}(y^2 - x^4)(x^3 + y^{3/2})}{y^{7/2}}$

Capítulo 4

Introducción a las muestras. Estimación puntual

4.1. Muestra Aleatoria Simple

Definición (Muestra Aleatoria Simple): Dada una variable aleatoria X, se denomina **muestra aleatoria simple** (*mas*) de tamaño n a la colección $(X_1, \ldots, X_n)$ resultante de n observaciones independientes de X.

Es conveniente precisar que una *mas* define una colección de resultados *a priori* (es decir, antes de realizar o conocer los experimentos). Si se tratase de resultados *a posteriori*, se trataría de un vector numérico. En ese caso, se denominaría **realización muestral** y lo denotaremos por $(x_1, \ldots, x_n)$.

Una consecuencia importante de este desconocimiento a priori de los resultados es que una *mas* es un vector aleatorio n-dimensional constituido por n variables independientes e idénticamente distribuidas.

Puesto que una *mas* es una variable aleatoria n-dimensional, tendrá asociada una función de distribución conjunta. Más específicamente, si F es la función de distribución de X, entonces la **función de distribución de la *mas*** es la función de distribución conjunta del vector aleatorio $(X_1, \ldots, X_n)$. Como X_i son i.i.d. $i = 1, \ldots, n$ se tiene que

$$F_{(X_1,\ldots,X_n)}(x_1, \ldots, x_n) = \prod_{i=1}^{n} F(x_i)$$

Si X es discreta, entonces

$$P[X_1 = x_1; \ldots; X_n = x_n] = \prod_{i=1}^{n} P[X_i = x_i]$$

Si X es continua, entonces

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i)$$

4.2. Funciones de una mas: Estadísticos

Definición (Estadístico): Sea una *mas* $(X_1, \ldots, X_n)$ siguiendo la distribución de X. La función $T(X_1, \ldots, X_n) : \mathbb{R}^n \to \mathbb{R}$, se denomina **estadístico**.

Nótese que un estadístico es una función de una variable aleatoria, por lo tanto es una variable aleatoria. Para una realización muestral $(x_1, \ldots, x_n)$, el estadístico tomará un valor numérico $T(x_1, \ldots, x_n)$.

4.2.1. Algunos estadísticos de interés

Definición (Media muestral): Sea $(X_1, \ldots, X_n)$ una *mas*. Se define la **media muestral** como

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Para una realización muestral $(x_1, \ldots, x_n)$ de la *mas* se tiene que la media muestral es $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$.

Definición (Mediana): Dados los datos de una muestra ordenados (por ejemplo de menor a mayor), $(x_{[1]}, \ldots, x_{[N]})$, la **mediana**, Me, es el valor que ocupa la posición central.

- Si N es impar, Me es el valor que ocupa la posición $\frac{N+1}{2}$

$$Me = x_{[(N+1)/2]}$$

- Si N es par, Me es la media de los datos que ocupan las posiciones $\frac{N}{2}$ y $\frac{N}{2} + 1$

$$Me = \frac{x_{[N/2]} + x_{[N/2+1]}}{2}$$

Definición (Moda): La **moda** es el valor que se repite más veces en la muestra.

En el caso en el que haya varios valores que se repiten el mismo número de veces, todos ellos (o cualquiera de ellos) es la moda.

Definición (Percentil): En una muestra ordenada, el percentil $Perc(E)$ deja por debajo el $E\,\%$ de los datos, y por el encima el $100 - E\,\%$.

4.3. Estimación puntual

Definición (Parámetro): Un parámetro es un valor numérico asociado a una distribución.

Definición (Estimador): Un estadístico $T(X_1, \ldots, X_n)$ es un **estimador** del parámetro θ si se utiliza para obtener valores aproximados de θ. Se notará como $\hat{\theta}$.

Nótese que un estimador es una variable aleatoria (ya que, como cualquier estadístico, es una función de la *mas*, que es un vector aleatorio). Cuando se tiene una realización muestral concreta, entonces $\hat{\theta} = T(x_1, \ldots, x_n)$ es una **estimación** del parámetro θ.

4.3.1. Propiedades de los estimadores

Definición (Sesgo): Se define el **sesgo** del estimador $\hat{\theta}$ como la función:

$$b(\hat{\theta}) = E[\hat{\theta} - \theta] = E[\hat{\theta}] - \theta$$

Si $b(\hat{\theta}) = 0$ se dice que $\hat{\theta}$ es un **estimador insesgado** de θ. En caso contrario, es un **estimador sesgado**.

Definición (Varianza de un estimador): La **varianza** de $\hat{\theta}$ es

$$V[\hat{\theta}] = E[(\hat{\theta} - E[\hat{\theta}])^2]$$

Definición (Error cuadrático medio): Sea θ un parámetro de una distribución, y $\hat{\theta}$ un estimador de θ. Se define el **Error Cuadrático Medio** de $\hat{\theta}$ como

$$ECM[\hat{\theta}] = E[(\hat{\theta} - \theta)^2]$$

Nótese lo siguiente:

- Dados $\hat{\theta}_1, \hat{\theta}_2$ estimadores de θ, se dice que $\hat{\theta}_1$ es **mejor** que $\hat{\theta}_2$ si el error cuadrático medio de $\hat{\theta}_1$ es menor que el de $\hat{\theta}_2$, es decir, si

$$ECM[\hat{\theta}_1] = E[(\hat{\theta}_1 - \theta)^2] \leq ECM[\hat{\theta}_2] = E[(\hat{\theta}_2 - \theta)^2]$$

- Existe una relación entre ECM, sesgo y varianza:

$$ECM[\hat{\theta}] = V[\hat{\theta}] + b^2(\hat{\theta})$$

4.4. Estimación de la esperanza y la varianza

4.4.1. Estimación de la esperanza

Para la estimación de la esperanza (que por simplicidad denotaremos por μ) nos serán de utilidad las siguientes propiedades de la media muestral $\overline{X}$ de una *mas*:

- $E[\overline{X}] = \mu$
- $V[\overline{X}] = \frac{\sigma^2}{n}$, siendo $\sigma^2 = V[X]$

4.4.2. Estimación de la varianza

Para la estimación de la varianza es preciso definir el siguiente estadístico:

Definición (Varianza muestral): Sea $(X_1, \ldots, X_n)$ una *mas*. Se define la **varianza muestral** como

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$$

Para una realización muestral $(x_1, \ldots, x_n)$ de la *mas* se tiene que la varianza muestral es $s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}$.

Si σ^2 es la varianza de la variable aleatoria, se tiene que $E[S^2] = \sigma^2$.

4.5. Dependencia entre variables

4.5.1. Estimación de la covarianza

Para la estimación de la covarianza es preciso definir el siguiente estadístico:

Definición (Covarianza muestral): Sea $(X_1, Y_1), \ldots, (X_n, Y_n)$ una *mas* que sigue al vector aleatorio (X, Y). Se define la **covarianza muestral** como

$$S_{XY} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{n-1}$$

Para unas realizaciones muestrales $(x_1, \ldots, x_n)$ y $(y_1, \ldots, y_n)$ de la *mas* se tiene que la covarianza muestral es $s_{XY} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{x})}{n-1}$.

Si $Cov[X, Y]$ es la covarianza de las variable aleatorias, se tiene que $E[S_{XY}] = Cov[X, Y]$.

4.6. Ejercicios Tema 4

Ejercicio 4.1. *Sea X una variable aleatoria cualquiera con media μ y varianza σ^2. Sea $(X_1, \ldots, X_n)$ una mas de X. ¿Cuál de estas expresiones no es un estadístico? Justificar por qué:*

1	$T_1 = X_1$	6	$Min = \min\{X_i\}$	11	Me	16	$T_8 = X_n - X_{n-1}$
2	$T_2 = \mu$	7	$T_4 = \min\{\mu, \sigma^2\}$	12	$Perc(50)$	17	$T_9 = \sum_{i=1}^{n-1} X_i$
3	$\overline{X} = \frac{\sum_{i=1}^n X_i}{n}$	8	$S^2 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n-1}$	13	$Perc(25)$	18	$T_{10} = e^{\mu}$
4	$T_3 = \sigma^2$	9	$T_5 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n}$	14	Mo	19	$Perc(75)$
5	$Max = \max\{X_i\}$	10	$T_6 = \mu + 3\sigma$	15	$T_7 = n(n-1)$	20	$Perc(90)$

Ejercicio 4.2. *Sea $(X_1, \ldots, X_n)$ una mas de X, donde $E[X] = \mu$ y $V[X] = \sigma^2$. Se desea estimar μ.*

1. *Hallar el sesgo, la varianza y el error cuadrático medio de $\hat{\mu}_1 = \overline{X}$*
2. *Idem para $\hat{\mu}_2 = X_1$*
3. *¿Cuál de los dos es mejor estimador de μ? ¿Por qué?*

Ejercicio 4.3. *Sea $(X_1, \ldots, X_n)$ una mas de X, donde $E[X] = \mu$ y $V[X] = \sigma^2$. Se desea estimar σ^2. Sean los estadísticos:*

$$\hat{\sigma}_1^2 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n-1}$$
$$\hat{\sigma}_2^2 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n}$$

1. *Hallar $E[\hat{\sigma}_1^2]$.*
2. *Hallar el sesgo de $\hat{\sigma}_1^2$.*
3. *Hallar $E[\hat{\sigma}_2^2]$.*
4. *Hallar el sesgo de $\hat{\sigma}_2^2 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n}$*
5. *¿Con qué estadístico se estimaría σ^2? ¿Por qué?*

Ejercicio 4.4. *Sea X una variable aleatoria tal que $E[X] = Np$ y $V[X] = Npq$ con $(q = 1 - p)$. Se desea estimar el parámetro $\theta = Np$, y para ello se toma una mas de X: $(X_1, \ldots, X_n)$.*

1. *Hallar el sesgo, la varianza y el error cuadrático medio de $\hat{\theta}_1 = \overline{X}$*
2. *Idem para $\hat{\theta}_2 = X_{n-1}$*
3. *¿Cuál de los dos es mejor estimador de θ?*
4. *Si se usa $\hat{\theta} = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n}$ como estimador de θ, demostrar que $b(\hat{\theta}) = -pq$.*

Ejercicio 4.5. *Sea X una variable aleatoria cualquiera con media μ y varianza σ^2. Sea $(X_1, \ldots, X_n)$ una mas de X.*

1. *Hallar $E[\overline{X}]$.*
2. *Hallar $V[\overline{X}]$.*

3. *¿Qué ocurre con la varianza cuando n crece?*

4. *Supongamos que estamos estudiando la variable X dada en el enunciado, y no conocemos μ. Tomamos dos muestras:*

   a) *n=100 y obtenemos como media muestral de dichas observaciones $\overline{x}_1 = 13.4$.*

   b) *n=10000 y obtenemos $\overline{x}_2 = 13.7$*

   Teniendo en cuenta el apartado anterior, ¿qué valor usaría para estimar μ y por qué?

Ejercicio 4.6. *Sea $(X_1, \ldots, X_n)$ una mas de una variable aleatoria X con función de distribución $F(x)$. Sean $X_{min} = \min_{i=1,\ldots,n}\{X_1, \ldots, X_n\}$, $X_{max} = \max_{i=1,\ldots,n}\{X_1, \ldots, X_n\}$. Demostrar:*

1. $P[X_{min} > x] = (1 - F(x))^n$

2. $P[X_{min} \leq x] = 1 - (1 - F(x))^n$

3. $P[X_{max} \leq x] = (F(x))^n$

4. $P[X_{max} > x] = 1 - (F(x))^n$

4.7. Problemas Tema 4

Problema 4.1. *** El tiempo de vida T de unos componentes electrónicos se distribuye según la siguiente función de densidad de parámetro $\theta > 0$:*

$$f(x) = \frac{e^{-\frac{x}{\theta}}}{\theta}, x > 0$$

Sea $(T_1, \ldots, T_n)$ una mas de T. Dados los siguientes estimadores de θ :

$$\hat{\theta}_1 = \frac{\sum_{i=1}^n T_i}{n-1}$$

$$\hat{\theta}_2 = \frac{\sum_{i=1}^n T_i}{n+1}$$

1. *Calcular el sesgo de ambos estimadores.*

2. *¿Cuál tiene menor varianza?*

3. *¿Cuál de los estimadores es mejor?*

4. *Sea $\hat{\theta}_3 = \bar{T}$, ¿cuál de los estimadores $\hat{\theta}_1$, $\hat{\theta}_2$ y $\hat{\theta}_3$ es el mejor?*

Resultado

1. $b(\hat{\theta}) = \frac{\theta}{n-1}$, $\hat{\theta}_2 = \frac{-\theta}{n+1}$
2. $V[\hat{\theta}_1] = \frac{n\theta^2}{(n-1)^2}$, $V[\hat{\theta}_2] = \frac{n\theta^2}{(n+1)^2}$
3. $ECM[\hat{\theta}_1] = \frac{(n+1)\theta^2}{(n-1)^2}$, $ECM[\hat{\theta}_2] = \frac{\theta^2}{n+1}$
4. $ECM[\hat{\theta}_3] = \frac{\theta^2}{n}$

Resultado

1. $\theta = \left(\frac{2}{k}\right)^{1/2}$
2. $E[\hat{\theta}_1] = \frac{2n}{2n+1}\theta$, $V[\hat{\theta}_1] = \left(\frac{2n}{2n+2} - \left(\frac{2n}{2n+1}\right)^2\right)\theta^2$
3. $\hat{\theta}_2 = \frac{3}{2}\bar{X}$, $E[\hat{\theta}_2] = \theta$, $V[\hat{\theta}_2] = \frac{\theta^2}{8n}$
4. $ECM(\hat{\theta}_1) = 0.01515 \cdot \theta^2$, $ECM(\hat{\theta}_2) = 0.025 \cdot \theta^2$

Problema 4.2. *El número de accidentes al mes en un determinado tramo de carretera X se distribuye de acuerdo a la siguiente función de probabilidad:*

$$P[X = k] = e^{-\lambda}\frac{\lambda^k}{k!}$$

con $k = 0, 1, \ldots$

Sabiendo que la función generatriz de momentos de dicha distribución es $\varphi(t) = e^{-\lambda(1-e^t)}$, se pide:

1. *Determinar el valor esperado de X.*
2. *Se han recogido los siguientes datos del número mensual de accidentes en ese tramo:*

Mes	1	2	3	4	5	6	7	8	9	10	11	12
No. accidentes	4	1	9	6	9	5	6	1	7	4	5	5

A la vista de estos datos se pide:

a) *Estimar un valor numérico para el número mensual esperado de accidentes.*

b) *Estimar la probabilidad de que en un mes no haya accidentes.*

c) *Estimar la probabilidad de que en un mes el número de accidentes sea menor que 3.*

Resultado

1. λ.
2. *a*) 5.167

   b) 0.0057

   c) 0.1113

Problema 4.3. *** Sea X una variable aleatoria discreta de parámetro λ cuya función generatriz de momentos es:*

$$\varphi(t) = e^{-\lambda(1-e^t)}$$

Sea (X_1, X_2, X_3) una mas de X. Determinar el error cuadrático medio de los siguientes estimadores de λ: $\hat{\lambda}_1 = \frac{X_1+X_2+X_3}{5}$ y $\hat{\lambda}_2 = \frac{2X_1+3X_2+3X_3}{10}$

Resultado

$ECM(\hat{\lambda}_1) = \frac{\lambda}{25}(3+4\lambda)$ $ECM(\hat{\lambda}_2) = \frac{1}{50}\lambda(11+2\lambda)$

Problema 4.4. ***** Se tiene (X_1, X_2) una mas de una característica modelada con la variable aleatoria X, siendo $E[X] = \mu$ y $V[X] = \sigma^2$. Se quiere emplear el estadístico $T = \alpha X_1 + (1-\alpha)X_2$ $(0 < \alpha < 1)$ para estimar μ. Se pide:*

1. *Determinar α para que T sea insesgado.*
2. *Determinar α para que T tenga varianza mínima.*

Resultado

1. $\forall\ \alpha$
2. $\alpha = 1/2$

Problema 4.5. ****** Sea X una variable aleatoria que mide en cm^2 los desperdicios generados en la fabricación de un lote de planchas metálicas. Se sabe que la función de densidad de dicha variable viene dada por la siguiente expresión:*

$$f(x) = kx, \text{ con } x \in [0, \theta]$$

1. *Determine el máximo área desperdiciada, θ, que puede generarse en función del parámetro desconocido k.*

Como θ es desconocido, se pretende estimar a través de una muestra aleatoria simple de tamaño n $(X_1, X_2, \ldots, X_n)$.

2. *Sea $\hat{\theta}_1 = max\{X_1, \ldots, X_n\}$ un estimador para θ. Si dicho estadístico tiene la siguiente función de densidad:*
$$g(\hat{\theta}_1) = n \cdot k^n \cdot \frac{\hat{\theta}_1^{\,2n-1}}{2^{n-1}}, \text{ con } \hat{\theta}_1 \in [0, \theta]$$
calcule la esperanza y la varianza de dicho estimador.

3. *Sabiendo que la media muestral $\overline{X}$ es en general un buen estimador de la esperanza de la variable aleatoria X, proponga un estimador de θ, $\hat{\theta}_2$, que se calcule utilizando dicho estadístico. Determine la esperanza y la varianza del estimador $\hat{\theta}_2$.*

4. *¿Cuál de los estimadores de θ propuestos en el ejercicio es mejor para una muestra de tamaño $n = 5$?*

5. *Sea la siguiente mas tomada del proceso de fabricación en la que se dan los cm^2 de desperdicio de 5 lotes de planchas metálicas: (3.25,5.32,8.23,4.25,6.35). Dar de forma justificada una buena estimación de θ.*

6. *Hallar la probabilidad aproximada de que se desperdicien menos de 8 cm^2 al fabricar un lote.*

Resultado

1. $\theta = \left(\frac{2}{k}\right)^{1/2}$
2. $E[\hat{\theta}_1] = \frac{2n}{2n+1}\theta$, $V[\hat{\theta}_1] = \left(\frac{2n}{2n+2} - \left(\frac{2n}{2n+1}\right)^2\right)\theta^2$
3. $\hat{\theta}_2 = \frac{3}{2}\bar{X}$, $E[\hat{\theta}_2] = \theta$, $V[\hat{\theta}_2] = \frac{\theta^2}{8n}$
4. $ECM(\hat{\theta}_1) = 0.01515 \cdot \theta^2$, $ECM(\hat{\theta}_2) = 0.025 \cdot \theta^2$
5. 8.23
6. 0.9448

Problema 4.6. ****** La siguiente distribución, que depende de los parámetros α y θ, tiene la particularidad de que no posee función generatriz de momentos:*
$$f(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, \; x \geq k$$

1. *Hallar k, el valor mínimo que puede tomar la variable aleatoria.*

2. *Hallar la función de distribución de X.*

3. *Hallar el valor esperado de X y su varianza.*

4. *Sea $Y = \ln \frac{X}{\theta}$, hallar $E[Y]$ y $V[Y]$.*

5. *Demostrar que $P[Y \leq y] = 1 - e^{-\alpha y}$ $y > 0$ (Nota: Se recomienda usar el resultado del apartado 2).*

Resultado

1. $k = \theta$,
   $alpha > 0$

2. $$F(x) = 1 - \left(\frac{\theta}{x}\right)^{\alpha}, \ \alpha > 0$$

3. $E[X] = \frac{\alpha\theta}{\alpha-1}$,
   $alpha > 1$, $V[X] = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}, \ \alpha > 2$

4. $E[Y] = \frac{1}{\alpha}, \ \alpha > 0$ $V[Y] = \frac{1}{\alpha^2}, \alpha > 0$

5. $P[Y \leq y] = 1 - e^{-\alpha y}$ $y > 0$

Capítulo 5

Distribuciones de variables aleatorias discretas

5.1. Uniforme discreta en n puntos

Definición (Distribución Uniforme): Una variable aleatoria $X = \{1, 2, \ldots, n\}$ sigue una **distribución uniforme** en n puntos si:

$$p_k = P[X = k] = \frac{1}{n}, k = 1, \ldots, n \tag{5.1}$$

lo que se denota $X \sim U(n)$

El experimento aleatorio asociado a esta distribución es cualquier experimento con n resultados equiprobables.

En la figura 5.1 se muestra la función de probabilidad de la distribución uniforme discreta para el caso general $U(n)$.

5.1.1. Propiedades

1. $\sum_{\forall x_k} p_k = 1$

2. $\varphi(t) = \frac{e^{t(n+1)} - e^t}{n(e^t - 1)}$, función que no está bien definida en $t = 0$, por lo que no es práctico para calcular los momentos de X.

3. $E[X] = \frac{n+1}{2}$

4. $V[X] = \frac{n^2-1}{12}$

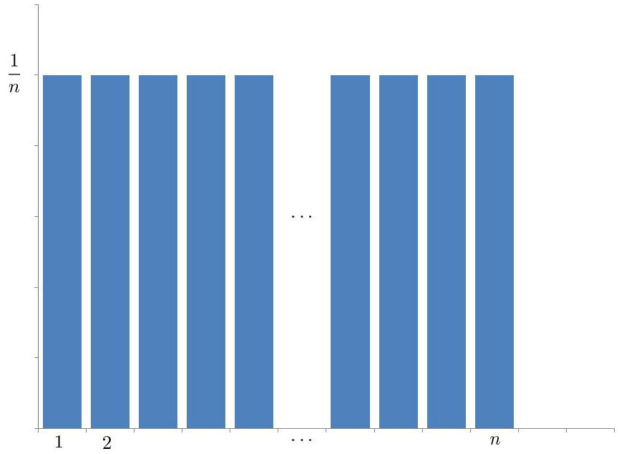

Figura 5.1: Función de probabilidad de la distribución Uniforme Discreta, $U(n)$

5.2. Concentrada en dos puntos/Bernoulli

Definición (Distribución concentrada en dos puntos): Una variable aleatoria $X = \{x_1, x_2\}$ sigue una distribución concentrada en dos puntos si:

$$p_1 = P[X = x_1] = 1 - p \text{ y } p_2 = P[X = x_2] = p, \text{ con } p \in (0,1).$$

Si $x_1 = 0$ y $x_2 = 1$ diremos que X sigue una **distribución de Bernoulli**, $X \sim Be(p)$. En este caso, la función de probabilidad se puede expresar de la siguiente forma compacta:

$$P[X = k] = p^k q^{1-k} \text{ con } k \in \{0,1\}, \text{ y } q = 1 - p$$

Esta distribución se asocia a cualquier experimento aleatorio con dos resultados posibles (éxito y fracaso), siendo p la probabilidad de éxito (asignado al valor 1, en el caso de la distribución de Bernoulli).

5.2.1. Propiedades

1. $\sum_{\forall x_k} p_k = 1$

2. $\varphi(t) = q + pe^t$

3. $E[X] = p$

4. $V[X] = pq$

5.3. Binomial

Definición (Distribución Binomial): Una variable aleatoria $X = \{0, 1, 2, \ldots, n\}$ sigue una distribución binomial $Bi(n, p)$, con $p \in (0, 1)$ y $q = 1 - p$, si:

$$p_k = P[X = k] = \binom{n}{k} p^k q^{n-k}, \; k = 0, \ldots, n \tag{5.2}$$

donde

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

El experimento asociado resulta de contar el número de éxitos al repetir n veces un subexperimento que tiene dos resultados posibles (éxito y fracaso) con probabilidades p y q, respectivamente.

En la figura 5.2 se muestra la función de probabilidad de la distribución binomial en el caso $Bi(50, 0.2)$.

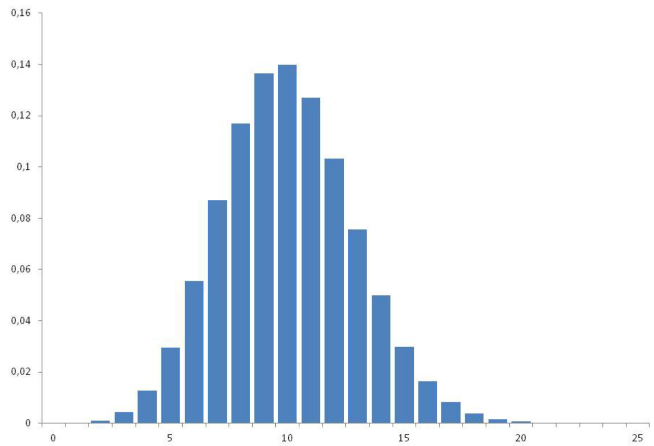

Figura 5.2: Función de probabilidad de la distribución Binomial, $Bi(50, 0.2)$

5.3.1. Propiedades

1. $\sum_{\forall x_k} p_k = 1$

2. $\varphi(t) = (q + pe^t)^n$

3. $E[X] = np$

4. $V[X] = npq$

5. Dadas $X_1, X_2, \ldots, X_n$ variables aleatorias discretas independientes que siguen una distribución de

Bernoulli, $X_i \sim Be(p)$, se tiene que

$$\sum_{i=1}^{n} X_i \sim Bi(n,p)$$

6. Reproductividad respecto a n: Dadas $X_1, X_2, \ldots, X_n$ variables aleatorias discretas independientes, cada una siguiendo una distribución binomial, $X_i \sim Bi(n_i, p)$, se tiene que

$$\sum_{i=1}^{n} X_i \sim Bi\left(\sum_{i=1}^{n} n_i, p\right)$$

5.4. Geométrica

Definición (Distribución Geométrica): Una variable aleatoria discreta $X = \{0, 1, 2, \ldots, \infty\}$ sigue una distribución geométrica $Ge(p)$, con $p \in (0,1)$ y $q = 1-p$, si:

$$p_k = P[X = k] = q^k p, \ k = 0, 1, 2, \ldots$$

El experimento asociado consiste en repetir un subexperimento aleatorio con dos resultados posibles (éxito y fracaso, con probabilidades p y q, respectivamente) hasta que se produzca el primer éxito, y contar el número de fracasos antes del primer éxito.

En la figura 5.3 se muestra la función de probabilidad de la distribución Geométrica en el caso $Ge(0.17)$.

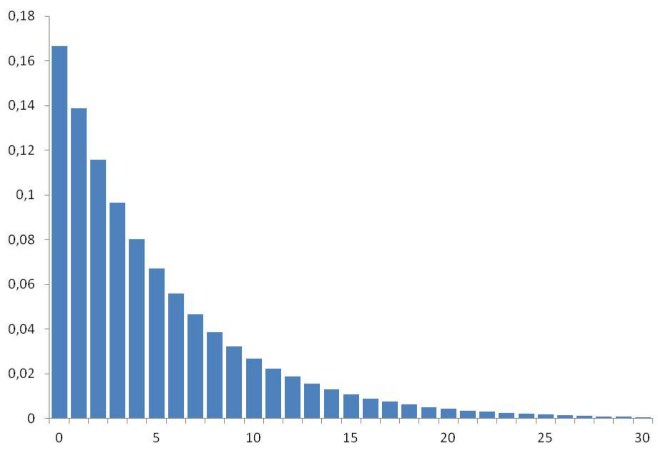

Figura 5.3: Función de probabilidad de la distribución Geométrica, $Ge(0.17)$

5.4.1. Propiedades

1. $\sum_{\forall x_k} p_k = 1$

2. $\varphi(t) = \frac{p}{1-qe^t}, \ \forall t \in \mathbb{R}/|qe^t| < 1$

3. $E[X] = \frac{q}{p}$

4. $V[X] = \frac{q}{p^2}$

5.5. Binomial Negativa

Definición (Distribución Binomial Negativa): Una variable aleatoria discreta $X = \{0, 1, 2, \ldots, \infty\}$ sigue una distribución binomial negativa $BN(r, p)$, con $p \in (0, 1)$, $q = 1 - p$ y $r \in \{1, 2, \ldots\}$, si:

$$p_k = P[X = k] = \binom{k + r - 1}{k} p^r q^k,\ \ k = 0, 1, 2, \ldots$$

En el experimento aleatorio asociado, se trata de realizar un número de subexperimentos en los que los posibles resultados son éxito (con probabilidad p) y fracaso (con probabilidad $q = 1 - p$), y contar el número de fracasos antes del r-ésimo éxito.

En la figura 5.4 se muestra la función de probabilidad de la distribución Binomial Negativa en el caso $BN(3, 0.17)$.

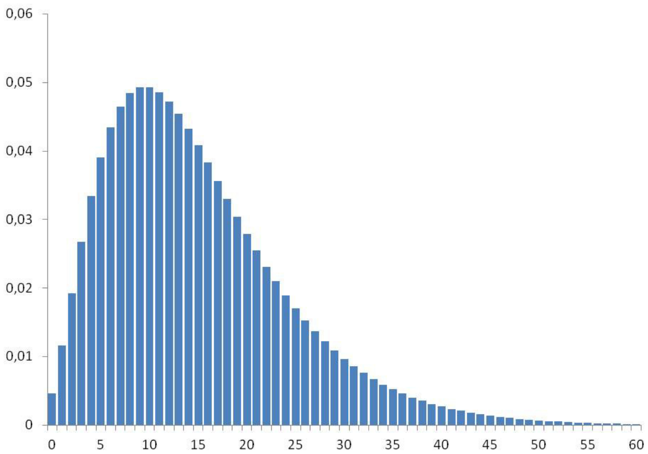

Figura 5.4: Función de probabilidad de la distribución Binomial Negativa, caso $BN(3, 0.17)$

5.5.1. Propiedades

1. $\sum_{\forall x_k} p_k = 1$

2. $\varphi(t) = \left(\frac{p}{1 - qe^t}\right)^r$, $\forall t \in \mathbb{R} / |qe^t| < 1$.

3. $E[X] = \frac{rq}{p}$

4. $V[X] = \frac{rq}{p^2}$

5. Si $r = 1 \Rightarrow X \sim BN(1, p) \equiv Ge(p)$

6. Dadas $X_1, X_2, \ldots, X_r$ variables aleatorias independientes que siguen una distribución Geométrica de parámetro p, $X_i \sim Ge(p)$, se tiene que

$$\sum_{i=1}^{r} X_i \sim BN(r, p)$$

7. Reproductividad respecto a r: Dadas $X_1, X_2, \ldots, X_n$ variables aleatorias independientes, cada una siguiendo una distribución binomial negativa, $X_i \sim BN(r_i, p)$, se tiene que

$$\sum_{i=1}^{n} X_i \sim BN\left(\sum_{i=1}^{n} r_i, p\right)$$

5.6. Poisson

Definición (Distribución de Poisson): Una variable aleatoria discreta $X = \{0, 1, 2, \ldots, \infty\}$ sigue una distribución de Poisson $\mathcal{P}o(\lambda)$, con $\lambda > 0$, si:

$$p_k = P[X = k] = e^{-\lambda}\frac{\lambda^k}{k!},\ k = 0, 1, 2 \ldots \tag{5.3}$$

El experimento asociado consiste en contar el número de ocurrencias de un determinado fenómeno en un intervalo dado (de tiempo o de espacio, habitualmente), en el que se cumplen las siguientes hipótesis:

1. El número esperado de ocurrencias en el intervalo es constante λ, y es directamente proporcional a la amplitud de dicho intervalo (regularidad).

2. La probabilidad de un número determinado de ocurrencias en un intervalo no depende del número de ocurrencias observadas en el intervalo anterior (independencia).

3. El número potencial de ocurrencias en el intervalo es teóricamente ilimitado.

En la figura 5.5 se muestra la función de probabilidad de la distribución de Poisson en el caso $\mathcal{P}o(8)$.

5.6.1. Propiedades

1. $\sum_{\forall x_k} p_k = 1$

2. $\varphi(t) = e^{-\lambda(1-e^t)}$

3. $E[X] = \lambda$

4. $V[X] = \lambda$

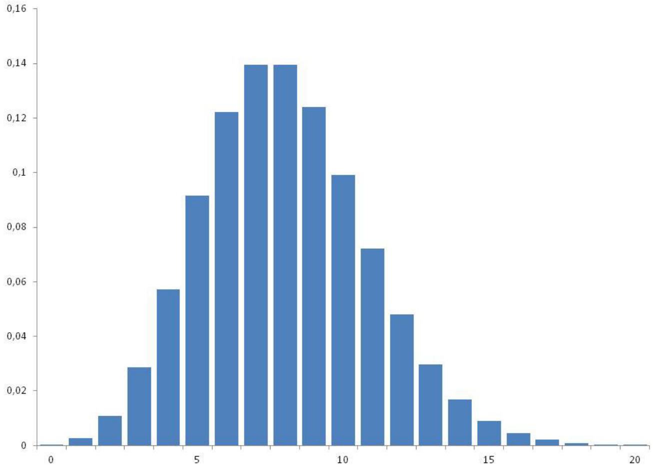

Figura 5.5: Función de probabilidad de la distribución de Poisson, $\mathcal{P}o(8)$

5. Reproductividad respecto a λ: Dadas $X_1, X_2, \ldots, X_n$ variables aleatorias discretas independientes que siguen una distribución de Poisson, $X_i \sim \mathcal{P}o(\lambda_i)$, se tiene que

$$\sum_{i=1}^{n} X_i \sim \mathcal{P}o\left(\sum_{i=1}^{n} \lambda_i\right)$$

6. $X \sim Bi(n, p)$ se puede aproximar a una Poisson si n es suficientemente grande y p suficientemente pequeño.

5.7. Ejercicios Tema 5

Ejercicio 5.1. *Si X es una variable aleatoria discreta que toma valores en $\mathbb{N}$, tal que $\forall n \in \mathbb{N}$*

$$P[X > n+1 | X > n] = P[X > 0]$$

Denotando $P[X > 0] = q_0$.

1. *Demostrar que $P[X > n+1 | X > n] = \frac{P[X > n+1]}{P[X > n]}$*

2. *Demostrar que $P[X > n] = q_0^{n+1}$*

3. *Hallar $P[X = n]$ a partir de los apartados anteriores.*

4. *A partir de los apartados anteriores, demostrar que $X \sim Ge(1 - q_0)$*

Ejercicio 5.2. *Sean $X, Y \sim Bi(n, p)$ independientes. Demostrar que $P[X = k \mid X + Y = m]$ es hipergeométrica con parámetros $N_1 = n$, $N_2 = n$ y $N = m$.*

Nota: Se dice que una variable aleatoria H sigue una distribución hipergeométrica de parámetros N_1, N_2 y N, $H \sim \mathcal{H}(N_1, N_2, N)$, si

$$P[H = k] = \frac{\binom{N_1}{k}\binom{N_2}{N-k}}{\binom{N_1+N_2}{N}}$$

con $\max\{0, N - N_2\} \leq k \leq \min\{N, N_1\}$, *es decir,* $k \geq 0$, $k \leq N$, $k \leq N_1$ *y* $N - k \leq N_2$.

Ejercicio 5.3. *1. Si $X \sim Bi(n,p)$, demostrar que se cumple $p_{k+1} = \frac{n-k}{k+1}\frac{p}{q}p_k$, donde $p_k = P[X = k]$.*

2. Si $X \sim Po(\lambda)$, demostrar que se cumple $p_{k+1} = \frac{\lambda}{k+1}p_k$, donde $p_k = P[X = k]$.

Ejercicio 5.4. *** Se tiene que el número medio de accidentes que ocurren en una semana en un tramo particular de una autovía de mucho tráfico es 3.*

1. *¿Cuál es la probabilidad de que haya un accidente esta semana?*
2. *¿Y de que no haya accidentes?*
3. *Calcular la probabilidad de que haya al menos un accidente esta semana.*
4. *¿Cuál es el número esperado de accidentes por semana?*
5. *¿Cuántos accidentes se esperan en un día? ¿y en dos semanas?*

Ejercicio 5.5. *** El número de veces que falla un instrumento de prueba debido a las partículas contaminantes de un producto se puede modelar como una variable aleatoria que sigue la distribución de Poisson de media 0.02 fallos por hora.*

1. *¿Cuál es la probabilidad de que el instrumento no falle en una jornada de 8 horas?*
2. *¿Cuál es la probabilidad de que haya al menos un fallo en un periodo de 24 horas?*

Ejercicio 5.6. **** Una empresa desea minimizar el número de artículos defectuosos enviados a los clientes. La empresa suele embalar los pedidos en lotes de 20 unidades. Para conseguir el objetivo propuesto decide establecer un plan de inspección consistente en tomar una muestra de cinco unidades de cada lote y rechazarlo si se detecta más de un artículo defectuoso. Si se rechaza el lote, deben inspeccionarse todos los artículos del mismo para poder reutilizar los buenos en otros lotes. Si un lote contiene 4 artículos defectuosos, ¿cuál es la probabilidad de que sea aceptado?*

Ejercicio 5.7. ***** Sea X una variable aleatoria discreta que toma los valores $\{1, 2, 3, 4\}$ con idéntica probabilidad. Sea la variable aleatoria discreta*

$$Y = \begin{cases} \frac{1}{X+1} & si\ X \leq 2 \\ \frac{1}{X-1} & si\ X > 2 \end{cases}$$

1. *Hallar la función de probabilidad de Y. ¿Qué distribución sigue?*

2. *Hallar la probabilidad de que* $Y = \frac{1}{3}$ *y* $X = 2$.

3. *Hallar la probabilidad de que* $X = 4$ *si* $Y = \frac{1}{3}$.

Ejercicio 5.8. *** Un examen contiene preguntas de cuatro temas distintos de una asignatura. Se pide:*

1. *Si la probabilidad de que el alumno A se haya estudiado un tema es* $p = \frac{10}{15}$*, ¿cuál es la probabilidad de que se haya estudiado los cuatro temas del examen?*

2. *Si el alumno B ha estudiado 10 temas del temario de 15 temas, ¿cuál es la probabilidad de que se haya estudiado los cuatro temas del examen?*

3. *¿Hay diferencias entre las probabilidades de los apartados anteriores? Justificar.*

5.8. Problemas Tema 5

Problema 5.1. *** Los dvd's que producen una compañía son defectuosos con una probabilidad de 0.01 independientemente uno de otro. La compañía vende los dvd's en paquetes de 10 y ofrece una garantía con devolución del dinero si más de un dvd es defectuoso.*

1. *Modelar la variable aleatoria para el número de dvd's defectuosos por paquete.*

2. *¿Cuál es la probabilidad de que un paquete sea devuelto?*

3. *Modelar la variable aleatoria para el número de paquetes defectuosos por cada* n *paquetes comprados.*

4. *Si una persona compra tres paquetes, ¿cuál es la probabilidad de que exactamente uno de ellos sea devuelto?*

5. *Hallar el número esperado de paquetes defectuosos en la compra de veinte paquetes.*

Resultado

1. $Bi(10, 0.01)$
2. 0.00426
3. $Bi(n, 0.00426)$
4. 0.01267
5. 0.0852

Problema 5.2. **** Se desea examinar a dos alumnos. Cada uno de ellos realiza una prueba diferente cada día. Si la prueba no es superada, el alumno tiene que realizar otra con un temario diferente. El examen es superado en el momento en que el alumno supera una de las pruebas. El alumno 1 estudia 1/3 del temario objeto de examen (con la garantía de que si se le pregunta va a superar la prueba) y el alumno 2 estudia 1/5. Los alumnos estudian por separado y no tienen posibilidad de copiar.*

Sean

$X_1 = \{$*número de días hasta que el alumno 1 pase el examen*$\}$

$X_2 = \{$*número de días hasta que el alumno 2 pase el examen*$\}$

1. *¿Qué distribuciones siguen X_1 y X_2?*
2. *Definimos $X = \{$número de días que transcurrirán hasta que al menos uno de ellos supere el examen$\}$. Expresar X en función de X_1 y X_2.*
3. *Calcular $P[\min\{X_1, X_2\} > x]$, con $x \in \mathbb{N}$, sabiendo que $\sum_{k=0}^{n} r^k = \frac{r^{n+1}-1}{r-1}$.*
4. *Calcular la función de distribución de X.*
5. *Calcular la función de probabilidad de X.*
6. *¿Qué distribución sigue X?*
7. *¿Cuál es el número esperado de días que tienen que pasar antes de que al menos uno de los dos alumnos pase el examen?*

Resultado

1. $X_1 \sim Ge(1/3)$ y $X_2 \sim Ge(1/5)$
2. $X = \min\{X_1, X_2\}$
3. $P[\min\{X_1, X_2\} > x] = \left(\frac{8}{15}\right)^{x+1}$. con $x \in \mathbb{N}$.
4. $F(x) = 1 - \left(\frac{8}{15}\right)^{Ent(x+1)}$ con $Ent(x) =$ parte entera de x.
5. $P[X = k] = \left(\frac{8}{15}\right)^{k} \frac{7}{15}$
6. $X \sim Ge\left(\frac{7}{15}\right)$
7. 1.1428

Problema 5.3. *** Un sistema de comunicaciones tiene n componentes, cada uno de los cuales funcionan de forma independiente con probabilidad p. El sistema total funcionará siempre que al menos la mitad de los componentes funcionen.*

1. *Modelar el número de componentes que funcionan en el sistema.*

2. *Hallar la probabilidad de que un sistema con cinco componentes sea efectivo.*

3. *La probabilidad de que un sistema con tres componentes sea efectivo.*

4. *¿Para qué valores de p es más probable que funcione un sistema con 5 componentes que otro con 3 componentes?*

Resultado

1. $Bi(n, p)$.
2. $p^3(6p^2 - 15p + 10)$
3. $p^2(3 - 2p)$
4. $p > \frac{1}{2}$

Problema 5.4. *** Una compañía de seguros con un gran número de clientes (que podemos asumir infinito), tiene un número medio de indemnizaciones diarias de 5. Se puede suponer que el número de indemnizaciones en diferentes días son independientes.*

1. *¿Qué proporción de días habrá menos de 3 indemnizaciones?*

2. *Hallar la probabilidad de que haya 4 indemnizaciones en un día.*

3. *Suponiendo que una semana tiene 5 días, se desean contar el número de días en los que hay 4 indemnizaciones. Modelar esta variable aleatoria.*

4. *¿Cuál es la probabilidad de que haya 4 indemnizaciones en exactamente 3 de los próximos 5 días?*

Resultado

1. 0.1246
2. 0.1755
3. $Bi(5, 0.1755)$
4. 0.0367

Problema 5.5. *** En un test de 100 preguntas hay cuatro opciones de respuesta de las que una y sólo una es correcta.*

1. *Si en todas las preguntas se marca una opción al azar, ¿cuál es el número esperado de respuestas correctas?*

2. *¿Cuál es la desviación típica?*

3. *Si para aprobar hace falta contestar correctamente 60 preguntas, ¿cuál es la probabilidad de aprobar? (utilizar una hoja de cálculo para contestar).*

4. *Hallar la probabilidad de aprobar (tener 60 preguntas correctas) si el test sólo tiene 2 opciones. (Utilizar una hoja de cálculo para contestar).*

Resultado

1. 25
2. 4.33
3. 0
4. 0.028

Problema 5.6. ***** Suponemos que una tarifa de llamadas de una determinada compañía de telefonía móvil es de* 20 *unidades monetarias por mes con 5 minutos en llamadas incluidos, y cada minuto adicional cuesta 1 unidad monetaria. Si el número de minutos de uso del móvil por mes, M, sigue una distribución geométrica con* $p = \frac{1}{5}$ *minutos, se pide:*

1. *Modelar la variable aleatoria coste del teléfono al mes, C, en función de M.*

2. *Escribir la función de probabilidad de la variable aleatoria anterior. Nota: La suma de la serie geométrica es*
$$\sum_{k=k_0}^{n} r^k = \frac{r^{k_0} - r^{n+1}}{1-r}$$

3. *¿Cuál es la esperanza de C? Nota1: Si $|r| < 1$, entonces*
$$\sum_{k=0}^{+\infty} k \cdot r^k = \frac{r}{(1-r)^2} \ y \ \sum_{k=0}^{+\infty} r^k = \frac{1}{1-r}$$
Nota2: Tener en cuenta que en la suma hay que hacer un cambio de variable

4. *Una compañía nueva ofrece una tarifa de 15 unidades monetarias al mes más 2 unidades monetarias por cada minuto de llamada. ¿Qué tarifa nos conviene más?*

Resultado

1. $C = g(M) = \begin{cases} 20 & M \leq 5 \\ 15 + M & M > 5 \end{cases}$ donde $M \sim Ge(p)$ con $p = \frac{1}{5}$

2. $P[C = c] = \begin{cases} 0 & c < 20 \\ 0.738 & c = 20 \\ \left(\frac{4}{5}\right)^{(c-15)} \cdot \frac{1}{5} & c > 20 \end{cases}$

3. 21.31

4. La de la primera compañía, ya que el valor esperado de la nueva es de 23

Problema 5.7. **** En un control de calidad se extrae una muestra de Y artículos, con Y un número elegido al azar entre 1 y 30. La probabilidad de que un artículo sea defectuoso es p.*

1. *Modelar la variable aleatoria que representa el número de artículos defectuosos en una muestra. ¿Qué distribución sigue?*

2. *Hallar la función de probabilidad conjunta de (X, Y).*

3. *Si la fracción defectuosa del proceso es $p = \frac{1}{1000}$:*

   a) *¿Cuál es la probabilidad de que la muestra sea de tamaño 20 y tenga 2 artículos defectuosos?*

   b) *¿Cuál es la probabilidad de que haya menos de un artículo defectuoso si la muestra tiene tamaño 20?*

   c) *¿Cuál es la probabilidad de que no haya ningún defectuoso en la muestra? Nota:*

   $$\sum_{k=1}^{n} r^k = \frac{r - r^{n+1}}{1 - r}$$

   d) *¿Cuál es la probabilidad de que la muestra tenga tamaño 25 si no hay ningún defectuoso en la muestra?*

Resultado

1. $X|Y = \{$número de artículos defectuosos en una muestra de tamaño $Y\}$. Para $Y = y_l$ se tiene que $X|Y = y_l \sim Bi(y_l, p)$.

2. $P[X = x_k; Y = y_l] = \frac{1}{30}\binom{y_l}{x_k} p^{x_k}(1-p)^{y_l - x_k}$ $y_l \in \{1, \ldots, 30\}$, $x_k \in \{0, \ldots, y_l\}$

3. Caso $p = 0.001$:

   a) $6.22{\cdot}10^{-6}$

   b) 0.9802

c) 0.9846

d) 0.033

Problema 5.8. **** Se ha realizado un examen en una clase de 50 alumnos y han aprobado 27 de ellos. El día de entrega de los exámenes sólo acuden a clase 10 alumnos. Hallar la probabilidad de que la mitad de los que han ido han aprobado.*

Resultado

0.2644

Problema 5.9. *Una fábrica realiza un control de calidad. Se toma una muestra de N piezas. Sea d el número de piezas defectuosas que hay en la muestra. Se extraen dos piezas al azar y se comprueba si son defectuosas o no.*

1. *Si hay reemplazamiento:*
   - a) *Hallar la función de probabilidad del número de piezas defectuosas. ¿Qué distribución sigue la variable aleatoria Z={Número de piezas defectuosas}?*
   - b) *Hallar la esperanza del número de piezas defectuosas en la extracción.*
   - c) *Hallar la varianza del número de piezas defectuosas en la extracción.*
2. *Si no hay reemplazamiento:*
   - d) *Hallar la función de probabilidad del número de piezas defectuosas en la extracción.*
   - e) *Demostrar que la función de probabilidad del número de piezas defectuosas es la función de probabilidad de una distribución hipergeométrica:*

$$p_k = \frac{\binom{N-d}{2-k}\binom{d}{k}}{\binom{N}{2}}$$

   - f) *Hallar la esperanza del número de piezas defectuosas en la extracción.*
   - g) *Hallar la varianza del número de piezas defectuosas en la extracción.*
3. *Comparar los resultados de esperanza y varianza con los obtenidos en el problema 3.2.*

Resultado

1. En el caso de que haya reemplazamiento:

a) $Z \sim Bi(2, \frac{d}{N})$

$$p_k = \binom{2}{k}\left(\frac{d}{N}\right)^k\left(\frac{N-d}{N}\right)^{2-k}$$

b) $E[Z] = 2\frac{d}{N}$.

c) $V[Z] = 2\frac{d}{N}\left(1-\frac{d}{N}\right) = 2\frac{N-d}{N}\frac{d}{N}$.

2. Si no hay reemplazamiento:

d)

$$p_0 = P[Z=0] = \frac{(N-d)(N-d-1)}{N(N-1)}$$
$$p_1 = P[Z=1] = \frac{2d(N-d)}{N(N-1)}$$
$$p_2 = P[Z=2] = \frac{d(d-1)}{N(N-1)}$$

e)

$$p_0 = \frac{\binom{N-d}{2}\binom{d}{0}}{\binom{N}{2}} = \frac{(N-d)(N-d-1)}{N(N-1)}$$
$$p_1 = \frac{\binom{N-d}{1}\binom{d}{1}}{\binom{N}{2}} = \frac{2d(N-d)}{N(N-1)}$$
$$p_2 = \frac{\binom{N-d}{0}\binom{d}{2}}{\binom{N}{2}} = \frac{d(d-1)}{N(N-1)}$$

f)

$$E[Z] = \frac{2d}{N}$$

g)

$$V[Z] = \frac{2d(N+d-2)}{N(N-1)} - \frac{4d^2}{N^2}$$

3. Coinciden con las halladas en ese problema.

Problema 5.10. ****** Un aerotaxi cubre diariamente la ruta Madrid-Barcelona (ida por la mañana y vuelta por la tarde). El número de pasajeros a la ida varía entre 1 y 5 con idéntica probabilidad. La probabilidad de que un pasajero vuelva en el día es de 0.6. Se pide:*

1. *Función de probabilidad de I número de pasajeros a la ida.*
2. *Función de probabilidad de V número de pasajeros a la vuelta condicionada a I*
3. *Determinar la probabilidad de que vuelvan 2 pasajeros si han viajado 5 a la ida.*
4. *Determinar la probabilidad de que vaya el pasaje lleno a la ida si a la vuelta hay 2 pasajeros.*

Resultado

1. $P[I = i] = \frac{1}{5}$, $i = 1, \ldots, 5$
2. $V = v \mid I = i \sim Bi(i, 0.6)$

   $P[V = v \mid I = i] = \binom{i}{v}(0.6)^v(0.4)^{(i-v)}$ para $v = 0, 1, \ldots, i$
3. 0.2304
4. 0.1684

Problema 5.11. ***** Un examen tipo test consta de n preguntas, cada una de ellas con cuatro opciones. La respuesta correcta a cada pregunta puede consistir en marcar ninguna, una, dos, tres o cuatro de las opciones posibles. Se pide:*

1. *Modelar la variable aleatoria X número de preguntas con respuesta correcta si en cada pregunta se marca un número al azar de las opciones.*
2. *Calcular la probabilidad de aprobar con el método anterior si el examen consta de 10 preguntas y se aprueba con más de dos preguntas acertadas.*
3. *Si cada respuesta correcta suma un punto y cada respuesta incorrecta restan medio punto, modelar (en función de X) la variable A puntuación obtenida si en un examen de n preguntas se marca en cada una de ellas un número al azar de las opciones.*
4. *Hallar la puntuación esperada para el caso definido en el apartado anterior.*

Resultado

1. $X \sim Bi(n, p)$, $p = 1/16$
2. 0.021
3. $A = \frac{3X-n}{2}$
4. $-\frac{13n}{32}$

Problema 5.12. ****** El número máximo de cadenas de televisión interesadas en una determinada competición internacional de fútbol a doble partido (ida y vuelta) es 4. En concreto, la variable aleatoria X que denota el número de cadenas interesadas en retransmitir el partido de vuelta se puede modelar como una binomial cuya media depende del resultado del equipo A en el partido de ida. Si el equipo A ha perdido en la ida, la media es de 1 cadena. Si ha empatado, de 2 cadenas, y si ha ganado, de 3 cadenas. Atendiendo al historial previo, el equipo A tiene una probabilidad de ganar de 0.6, de empatar de 0.1, y de perder de 0.3. Se pide:*

1. *Probabilidad de que el número de cadenas sea superior a 2, sabiendo que el equipo A ganó en la ida.*

2. *Probabilidad de que el número de cadenas sea superior a 2.*

3. *Probabilidad de que el equipo A pierda en la ida y el número de cadenas sea superior a 3.*

4. *Probabilidad de que el equipo A haya perdido en la ida sabiendo que en la vuelta el número de cadenas fue superior a 3.*

5. *Suponiendo que los ingresos del equipo (I) son proporcionales a X, es decir: $I = mX$, siendo m una constante, determinar los ingresos esperados debidos al partido de vuelta.*

Resultado

1. 0.7383
2. 0.4895
3. 0.00117
4. 0.00593
5. 2.3 m

Problema 5.13. *** Un lote de N componentes electrónicos es sometido a un control de calidad en el que se extrae una muestra de n componentes al azar, $n < N$. Sabemos que un lote dado tiene $d > 2$ componentes defectuosos, por lo tanto habrá $N - d$ componentes buenos.*

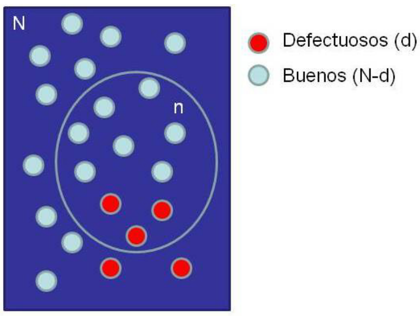

Figura 5.6: Distribución de los componentes electrónicos. Problema 5.13

1. *¿Cuál es la probabilidad de que haya x componentes defectuosos en cualquier muestra?*

2. *Para que el lote se considere aceptable puede haber como máximo dos componentes defectuosos en la muestra. Si $N = 100$, $n = 50$ y $d = 5$ ¿cuál es la probabilidad de que un lote cualquiera supere el control de calidad?*

Resultado

1. $P[X = x] = \frac{\binom{d}{x}\cdot\binom{N-d}{n-x}}{\binom{N}{n}}$
2. 0.5

Problema 5.14. ****

Un lote de componentes electrónicos es sometido a un control de calidad en el que se extrae una muestra de n componentes al azar. Si la fracción defectuosa de este proceso es $\frac{d}{N}$.

1. *¿Cuál es la probabilidad de que haya x componentes defectuosos en cualquier muestra?*
2. *Para que el lote se considere aceptable puede haber como máximo dos componentes defectuosos en la muestra. Si $N = 100$, $n = 50$ y $d = 5$ ¿cuál es la probabilidad de que un lote cualquiera supere el control de calidad?*

Resultado

1. $P[X = x] = \binom{n}{x}\left(\frac{d}{N}\right)^{x}\left(1 - \frac{d}{N}\right)^{n-x}$
2. 0.54

Problema 5.15. ***** Un jugador tiene c unidades monetarias que apuesta en un juego en el que en cada partida gana o pierde con idéntica probabilidad. Participa en N partidas independientes. Si gana dobla y si pierde divide por la mitad toda su fortuna.*

1. *Modelar la variable aleatoria del número de veces ganadas.*
2. *Modelar la variable aleatoria para la fortuna que gana el jugador en función de la variable anterior. Pista:*
   - *Calcular la fortuna al jugar tres veces en las que gana 2 partidas.*
   - *Calcular la fortuna al jugar diez veces en las que gana 4 partidas.*
   - *Calcular la fortuna al jugar N veces en las que gana X partidas.*
3. *Hallar el valor esperado de la fortuna del jugador.*

Problema 5.16. *Una línea de autobús tiene 12 paradas, de las que 4 se consideran de gran afluencia al estar situadas cerca de centros comerciales. Se asume que el número de pasajeros en una paradas sigue una distribución de Poisson. Adicionalmente, se sabe que la probabilidad de que no haya ningún pasajero esperando en una parada considerada de gran afluencia es de 0.1, mientras que asciende a 0.4 en el resto de paradas. Se pide:*

1. *La probabilidad de que haya exactamente dos pasajeros esperando en una parada de gran afluencia.*
2. *La probabilidad de que haya exactamente dos pasajeros esperando en una parada cualquiera.*
3. *Si el autobús se aproxima a una parada y hay tres pasajeros esperando, la probabilidad de que dicha parada sea de gran afluencia.*
4. *La probabilidad de que haya más de un pasajero esperando en dos de las paradas de gran afluencia de la línea.*
5. *La probabilidad de que haya más de un pasajero esperando en la tercera parada de gran afluencia de la línea si en las dos paradas de gran afluencia anteriores no lo ha habido.*

Resultado

1. 0.2651
2. 0.2003
3. 0.6653
4. 0.2935
5. 0.0730

Capítulo 6

Distribuciones de variables aleatorias continuas

6.1. Uniforme continua

Definición (Distribución Uniforme Continua): Una variable aleatoria X sigue una distribución uniforme en el intervalo $[a, b] \subseteq \mathbb{R}$, $X \sim U[a, b]$, si su función de densidad $f(x)$ tiene la siguiente expresión:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{e.c.c.} \end{cases} \tag{6.1}$$

El experimento asociado viene dado por cualquier experimento en el que la variable aleatoria toma valores reales en el intervalo $[a, b]$ y todos los subintervalos de la misma longitud son igualmente probables.

6.1.1. Propiedades

1. $\int_{-\infty}^{+\infty} f(x)dx = 1$

2. La función de distribución viene dada por:

$$F(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a < x \leq b \\ 1 & \text{si } x > b \end{cases}$$

3. $\varphi(t) = \frac{e^{bt} - e^{at}}{t(b-a)}$, función que no está bien definida en $t = 0$, por lo que no es útil a la hora de calcular los momentos de X.

4. $E[X] = \frac{a+b}{2}$

5. $V[X] = \frac{(b-a)^2}{12}$

En la figura 6.1 se muestra la función de densidad y la función de distribución Uniforme para el caso general $U[a,b]$.

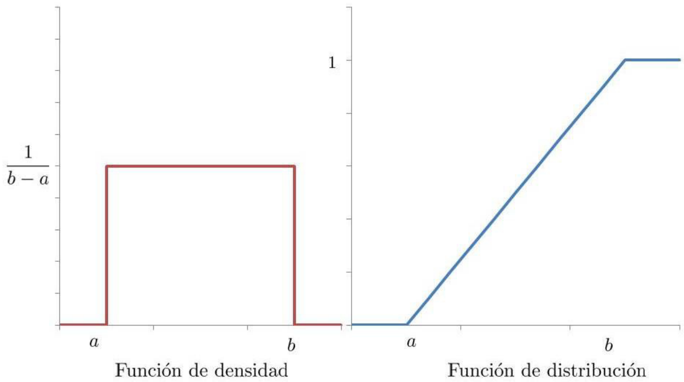

Figura 6.1: Función de densidad y función de distribución Uniforme, $U[a,b]$

6.2. Exponencial

Definición (Distribución Exponencial): Una variable aleatoria X sigue una distribución exponencial de parámetro $\lambda > 0$, $X \sim Exp(\lambda)$, si su función de densidad $f(x)$ tiene la siguiente expresión:

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \lambda e^{-\lambda x} & \text{si } x > 0 \end{cases} \tag{6.2}$$

Respecto al experimento asociado, el tiempo de espera entre ocurrencias de un experimento con distribución de Poisson sigue una distribución exponencial bajo ciertas condiciones que implican la independencia entre las ocurrencias en intervalos de tiempo que no se solapan.

6.2.1. Propiedades

1. $\int_{-\infty}^{+\infty} f(x)dx = 1$

2. La función de distribución viene dada por:

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-\lambda x} & x > 0 \end{cases}$$

3. $\varphi(t) = \frac{\lambda}{\lambda - t}$, con $t < \lambda$

4. $E[X] = \frac{1}{\lambda}$

5. $V[X] = \frac{1}{\lambda^2}$

6. Si $X_1, \ldots, X_n$ tal que $X_i \sim Exp(\lambda_i)$, $i = 1, \ldots, n$ independientes, entonces

$$min_{1 \leq i \leq n}\{X_i\} \sim Exp\left(\sum_{i=1}^{n} \lambda_i\right)$$

La representación gráfica de las funciones de densidad y de distribución Exponencial para el caso $Exp(0.26)$ se muestra en la figura 6.2.

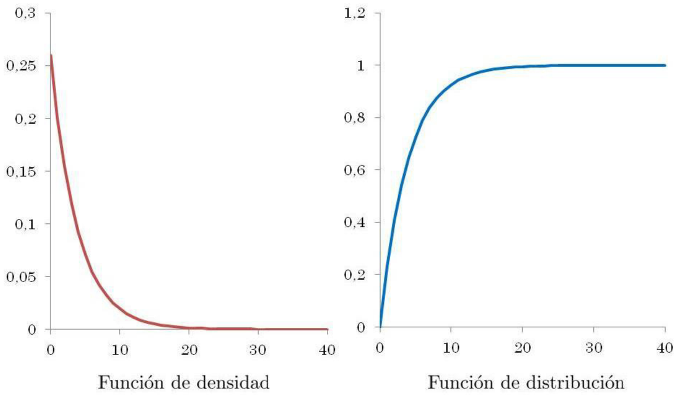

Figura 6.2: Función de densidad y función de distribución Exponencial, $Exp(0.26)$

6.3. Gamma

Definición (Distribución Gamma): Una variable aleatoria X sigue una distribución Gamma de parámetros $p > 0$ y $a > 0$, $X \sim Ga(p, a)$, si

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1} & x > 0 \end{cases}$$

donde $\Gamma(p)$ es la función gamma:

$$\Gamma(p) = \int_0^{+\infty} e^{-x} x^{p-1} dx$$

que, como es sabido, tiene –entre otras– las siguientes propiedades:

$$\begin{cases} \Gamma(p) = (p-1)\Gamma(p-1) \\ \Gamma(1) = 1 \\ \Gamma(n) = (n-1)! \\ \Gamma(\frac{1}{2}) = \sqrt{\pi} \end{cases}$$

La distribución Gamma permite modelar distintos fenómenos de interés en la ingeniería. En concreto, tiene mucho interés el caso particular en el que p toma valores naturales. Así, la distribución Gamma con $p = n \in \mathbb{N}$ y $a = \lambda > 0$ se denomina **distribución Erlang**.

Definición (Distribución Erlang): Una variable aleatoria X sigue una distribución Erlang de parámetros $n > 0$ y $\lambda > 0$, $X \sim Erlang(n, \lambda)$. Su función de densidad es

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \frac{\lambda^n}{(n-1)!} e^{-\lambda x} x^{n-1} & x > 0 \end{cases}$$

En este caso el experimento asociado es el tiempo de espera hasta la n-ésima ocurrencia de un determinado fenómeno aleatorio cuyo número de ocurrencias (bajo las mismas condiciones mencionadas en la distribución exponencial) sigue una Poisson de parámetro λ. Como se verá en las propiedades, es fácil demostrar (usando la función generatriz de momentos) que la suma de n variables aleatorias independientes de distribución $Exp(\lambda)$ sigue una distribución Gamma.

6.3.1. Propiedades

1. La función de distribución toma la siguiente forma (nótese que no es integrable en general).

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{a^p}{\Gamma(p)} \int_0^x e^{-at} t^{p-1} dt & x > 0 \end{cases}$$

   No obstante, sí es integrable para el caso particular de la $Erlang(n, \lambda)$:

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-\lambda x} \sum_{i=0}^{n-1} \frac{(\lambda x)^i}{i!} & x > 0 \end{cases}$$

2. $\varphi(t) = \left(\frac{a}{a-t}\right)^p$, $t < a$

3. $E[X] = \frac{p}{a}$

4. $V[X] = \frac{p}{a^2}$

5. Reproductividad con respecto al parámetro p: Si $X_1, \ldots, X_n$ son variables aleatorias independientes, con $X_i \sim Ga(p_i, a)$, $i = 1, \ldots, n$, entonces

$$\sum_{i=1}^{n} X_i \sim Ga(\sum_{i=1}^{n} p_i, a)$$

6. Para $p = 1$ se tiene que $X \sim Ga(1, \lambda) \equiv Exp(\lambda)$.

7. Si $X_1, \dots, X_n \sim Exp(\lambda)$ iid, entonces
$$\sum_{i=1}^{n} X_i \sim Ga(n, \lambda) \equiv Erlang(n, \lambda)$$

La representación gráfica de las funciones de densidad y de distribución Gamma para el caso $Ga(3, 0.1)$ se muestra en la figura 6.3.

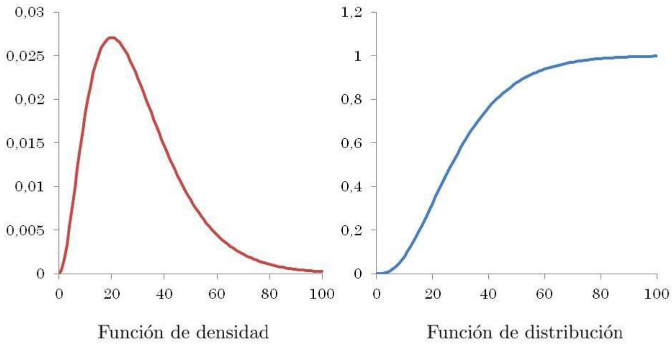

Figura 6.3: Función de densidad y función de distribución Gamma, $Ga(3, 0.1)$

6.4. Normal

Definición (Distribución Normal): Una variable aleatoria X sigue una distribución normal o de Gauss, $X \sim N(\mu, \sigma^2)$ si
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \; x \in \mathbb{R}$$

Teniendo en cuenta las propiedades de la función de densidad, existen multitud de experimentos en diversas ramas de la ciencia (Sociología, Medicina, Ingeniería, etc) que siguen esta distribución. Además, es también ampliamente utilizada debido a que, bajo ciertas condiciones, la suma de variables aleatorias independientes sigue una distribución normal.

6.4.1. Propiedades

1. La función de densidad verifica:

   a) $\lim\limits_{x \to +\infty} f(x) = 0$

   b) $\lim\limits_{x \to -\infty} f(x) = 0$

   c) $f(\mu + x) = f(\mu - x)$, es decir, es simétrica con respecto a μ.

   d) La representación de la función de densidad es la campana de Gauss (ver figura 6.4).

2. Su función de distribución es:
$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

Nótese que la función de densidad de la distribución normal no es integrable en el intervalo $(-\infty, x]$.

3. $\varphi(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

4. $E[X] = \mu$

5. $V[X] = \sigma^2$

6. Reproductividad respecto los parámetros μ y σ^2: Si $X_1, \ldots, X_n$ son variables aleatorias independientes, con $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \ldots, n$, entonces

$$S_n = \sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

7. Si $X_1, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas $N(\mu, \sigma^2)$, entonces

$$S_n = \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$$

8. Si X_1, X_2 son variables aleatorias independientes, con $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$, entonces

$$X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

9. En general, $X \sim N(\mu, \sigma^2) \Leftrightarrow \frac{X-a}{b} \sim N\left(\frac{\mu-a}{b}, \frac{\sigma^2}{b^2}\right)$, con $a \in \mathbb{R}$ y $b > 0$.

   Como consecuencia de esta propiedad tenemos que:

   - $X \sim N(\mu, \sigma^2) \Leftrightarrow \frac{X-\mu}{\sigma} \sim N(0, 1)$
   - $X \sim N(\mu, \sigma^2) \Leftrightarrow X - \mu \sim N(0, \sigma^2)$
   - $X \sim N(\mu, \sigma^2) \Leftrightarrow X/\sigma \sim N(\mu/\sigma, 1)$
   - $Z \sim N(0, 1) \Leftrightarrow \mu + \sigma Z \sim N(\mu, \sigma^2)$
   - Si $X_1, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas $N(\mu, \sigma^2)$, entonces $\frac{S_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n} \sim N(\mu, \sigma^2/n)$

La representación gráfica de las funciones de densidad y de distribución Normal para el caso $N(10, 10)$ se muestra en la figura 6.4.

6.4.2. Tipificación

Definición (Tipificación): Sea $X \sim N(\mu, \sigma^2)$. Definimos

$$Z = \frac{X - \mu}{\sigma}$$

como **tipificación** de X o X **tipificada**.

Z tiene las siguientes propiedades:

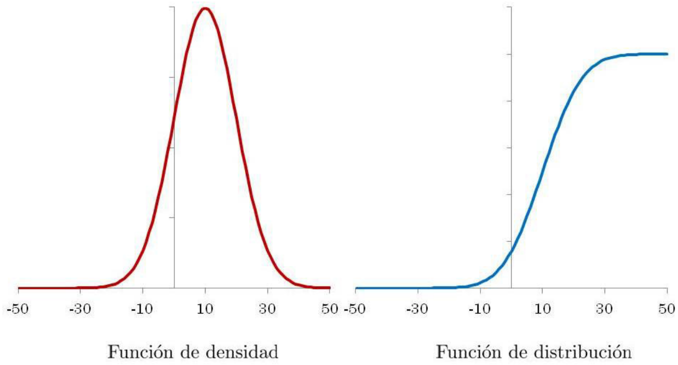

Figura 6.4: Función de densidad y función de distribución Normal, $N(10, 10)$

- $Z \sim N(0,1)$

- La función de densidad de Z es

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \ z \in \mathbb{R} \tag{6.3}$$

- La función de distribución en el punto z, $\Phi(z) = P[Z \leq z]$, es el área que encierra la curva de la función de densidad entre $-\infty$ y z (ver figura 6.5).

- Al ser la función de densidad de la normal simétrica respecto de μ, en este caso particular ($\mu = 0$) tenemos que $\phi(z) = \phi(-z)$. Por lo tanto, por la simetría de $\phi(z)$ se tiene que $\Phi(-z) = 1 - \Phi(z)$, $\forall z \in \mathbb{R}$, es decir, $P[Z \leq -z] = 1 - P[Z \leq z] = P[Z \geq z]$ (ver figura 6.6).

- Se tiene que para $X \sim N(\mu, \sigma^2)$ $P[a \leq X \leq b] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$

  para cualquier par de valores $a, b \in \mathbb{R}$, $a < b$.

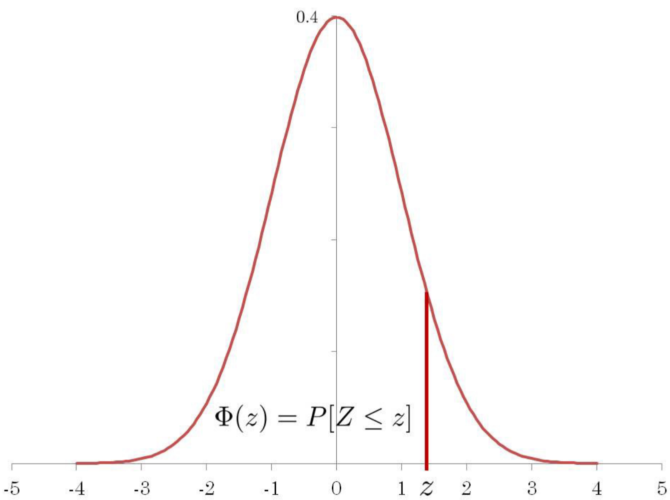

Figura 6.5: $\Phi(z) = P[Z \leq z]$

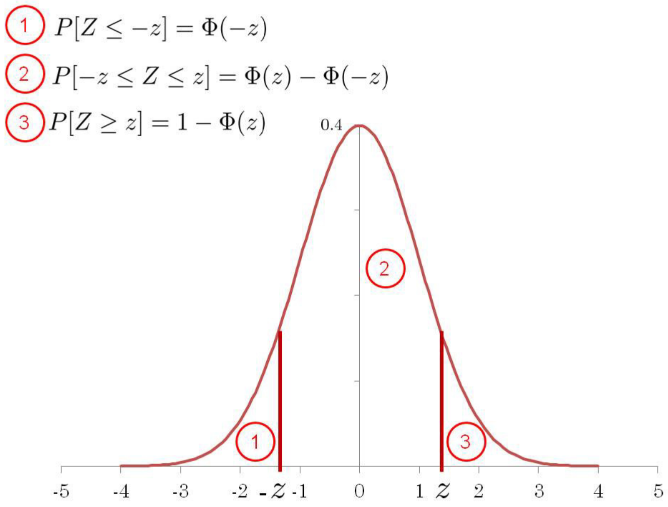

Figura 6.6: $\Phi(-z) = P[Z \le -z] = P[Z \ge z] = 1 - P[Z \le z] = 1 - \Phi(z)$

6.4.3. Teorema Central del Límite

Teorema Central del Límite: Sean $X_1, X_2, \ldots$ variables aleatorias independientes e idénticamente distribuidas con esperanza μ y varianza σ^2. Sea $S_n = \sum_{i=1}^n X_i$ $n = 1, 2, \ldots$. Entonces, para cada $x \in \mathbb{R}$ se tiene que

$$\lim_{n \to +\infty} P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right] = \Phi(x)$$

6.5. Ejercicios Tema 6

Ejercicio 6.1. *Sea X una variable aleatoria continua con distribución uniforme en el intervalo* $(-5, 5)$.

1. *¿Cuál es la función de densidad de X?*
2. *¿Cuál es la función de distribución de X?*
3. *¿Cuál es la esperanza de X?*
4. *¿Cuál es la esperanza de X^5?*
5. *¿Y la esperanza de X^4?*
6. *¿Cuál es la esperanza de e^X?*

Ejercicio 6.2. *Sean X_1, X_2 iid tal que $X_1, X_2 \sim Exp(\lambda)$. Demostrar que*

$$min\{X_1, X_2\} \sim Exp\,(2\lambda)$$

Ejercicio 6.3. *Si* $X \sim Exp(\lambda)$ *entonces:*

$$P[X > r + s | X > r] = P[X > s] \ \forall r, s \in \mathbb{R}, \ r, s > 0 \tag{6.4}$$

Ejercicio 6.4. *Sea* $X \sim N(\mu, \sigma^2)$*. Se pide:*

1. *Hallar* $P[|X - \mu| \leq k\sigma]$ *para* $k = 1, 2, 3$.
2. *Hallar* $P[(X - \mu)^2 \leq \sigma^2]$

Ejercicio 6.5. *Sea* $Z \sim N(0, 1)$*. Sea* $z \geq 0$*. Escribir las siguientes probabilidades en función de* $\Phi(z)$*:*

1. $P[Z \geq z]$
2. $P[Z \leq -z]$
3. $P[-z \leq Z \leq z]$
4. $P[Z \geq -z]$

Ejercicio 6.6. *Sea* X *una variable aleatoria con función generatriz de momentos:*

$$\varphi(t) = (1 - 2t)^{-\alpha/2}$$

¿Se verifica la reproductividad sobre el parámetro α*?*

Ejercicio 6.7. *Sea* $X \sim Exp(\lambda)$*. Se pide:*

1. *Demostrar que* $\frac{X}{c} \sim Exp(c\lambda)$ *para cualquier constante* $c > 0$.
2. *Si* $(X_1, \ldots, X_n)$ *es una mas de tamaño* n *que sigue a* X*, demostrar que* $\bar{X} \sim Erlang(n, \lambda n)$.

Ejercicio 6.8. **** Sea* T *el tiempo (en horas) de reparación de una avería de un determinado sistema. Se sabe que el logaritmo neperiano de* T *sigue una distribución* $N(\mu, \sigma^2)$*. Sea* $X = \ln T \sim N(\mu, \sigma^2) \Rightarrow T = e^X$*. Se tiene que la función de densidad de* T *es*

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} e^{-\frac{(\ln t - \mu)^2}{2\sigma^2}}, \ t > 0$$

y se dice que T *sigue una distribución log-normal,* $T \sim LogN(\mu, \sigma^2)$.

1. *Hallar la probabilidad de que* T *sea menor o igual que un tiempo dado* t*,* $F_T(t)$*, en función de* $\Phi(z)$*, función de distribución* $N(0, 1)$.
2. *Hallar la esperanza de* T.
3. *Si* $\mu = 0.5$ *y* $\sigma = 1$ *en horas, hallar el tiempo esperado de avería y la probabilidad de que una avería se repare en menos de 45 minutos.*

Ejercicio 6.9. *Sea* $X \sim Exp(1)$.

1. *Demostrar que* $\varphi_X(t) = \frac{1}{1-t}$, $t < 1$.
2. *Dado* $n \in \mathbb{N}$, *demostrar que* $E[X^n] = n!$ *a partir de la función generatriz de momentos de* X.
3. *Dado* $\alpha \in \mathbb{R}$, *demostrar que* $E[X^\alpha] = \Gamma(\alpha+1)$.
4. *Dadas* $X, Y \sim Exp(1)$ *independientes, demostrar que* $\varphi_{X+Y}(t) = \varphi_X(t) \cdot \varphi_Y(t)$.
5. *Dadas* $X_1, \dots, X_n \sim Exp(1)$ *independientes, hallar la función generatriz de momentos de la variable aleatoria* $S = \sum_{i=1}^n X_i$.
6. *Demostrar que* $E[S] = V[S] = n$.
7. *Demostrar que* $S \sim Erlang(n, 1)$

Nota: la función gamma viene dada por la expresión:

$$\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx$$

Ejercicio 6.10. *Considere la siguiente función:*

$$f(x) = \frac{x}{a^2} e^{-\frac{x}{a}}, x > 0$$

1. *Demostrar que la función anterior es una función de densidad de una variable aleatoria* X, *si* $a > 0$, *haciendo uso de las propiedades de la función de densidad.*
2. *Calcular de forma aproximada la probabilidad de que el estimador del parámetro* a, $\hat{a} = \frac{\bar{X}}{2}$ *tome valores por encima de* $\frac{3a}{4}$, *cuando se extrae una mas de tamaño 32.*

6.6. Problemas Tema 6

Problema 6.1. *** La temperatura máxima T medida en grados Fahrenheit, en un día de Julio en Nueva Jersey, sigue una distribución normal de media 85 y varianza 100. Hallar*

1. *La probabilidad de que la temperatura sea mayor que 100 ^oF.*
2. *La probabilidad de que la temperatura sea menor que 60 ^oF.*
3. *La probabilidad de que la temperatura esté entre 70 y 100 ^oF.*

Resultado

1. 0.06681
2. 0.00621
3. 0.86638

Problema 6.2. **** Un telefonista atiende dos teléfonos de una centralita con distintos números. El tiempo entre llamadas del teléfono 1 sigue una distribución exponencial de media 3 minutos, y el teléfono 2 otra distribución exponencial de media 5 minutos. Suponiendo que no hay llamadas en espera, hallar la probabilidad de que entre dos llamadas que atiende el telefonista haya más de 2 minutos.*

Resultado

0.3441

Problema 6.3. **** En un sistema se quieren instalar n componentes iguales de forma que cuando uno falle, el siguiente se active para mantener el sistema en pleno funcionamiento. Se supone que el tiempo de vida de cada componente sigue una distribución exponencial de parámetro λ. El tiempo de vida medio de un componente se estima en 1000 horas.*

1. *¿Cuántos componentes debemos poner como mínimo en el sistema para que el tiempo medio de vida sea un año?*
2. *Si el número de componentes es 6, ¿Cuál es la probabilidad de que el sistema dure al menos un año?*

Resultado

1. 9 ó más.
2. 0.13

Problema 6.4. **** En una fábrica se dispone de tres máquinas que producen el mismo tipo de componente electrónico. Las máquinas están dispuestas en paralelo. El tiempo de proceso de cada una de ellas sigue una distribución exponencial, siendo el tiempo medio de proceso de la primera máquina 60 segundos, el de la segunda máquina 90 segundos y el de la tercera 180 segundos por pieza.*

Aguas abajo de dichas máquinas se encuentra un operario encargado del control de calidad. La disposición de las máquinas y el operario se muestran en la figura 6.7.

Se pide

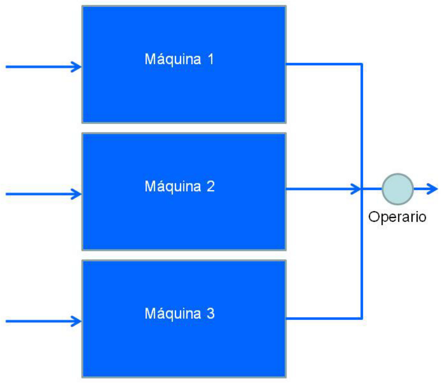

Figura 6.7: Disposición de las máquinas en la fábrica. Problema 6.4

1. *Tiempo medio (en minutos) entre llegadas de componentes al operario.*

2. *Probabilidad de que el tiempo entre piezas que le llegan al operario sea mayor que 30 segundos.*

Las máquinas no paran durante el turno de ocho horas del operario, pero por falta de espacio, nada más que salen de las máquinas los componentes son embalados tras el control de calidad. Por lo tanto hay un operario suplente para los descansos reglamentarios del operario principal del control de calidad. Suponiendo que el tiempo medio de descanso es de 30 minutos.

3. *¿Cuántas piezas se espera que procese el operario suplente en cada suplencia?*

4. *Hallar la probabilidad aproximada de que el operario suplente trabaje más tiempo que el tiempo esperado de suplencia.*

Resultado

1. 30 segundos
2. 0.3678
3. 60
4. 0.5

Problema 6.5. **** Cierto tipo de bombilla tiene una duración media de 1500 horas y una desviación típica de 150 horas. Se supone que las duraciones de las bombillas se distribuyen normalmente. Si ponemos tres bombillas en una misma habitación:*

1. *¿Cuál es la probabilidad de que se tenga luz al menos 1800 horas?*

2. *¿Y como mucho 1200 horas?*

Resultado

1. 0.06671
2. $1.177 \cdot 10^{-5}$

Problema 6.6. *** Un tubo fluorescente tiene una vida cuya duración en horas se distribuye como* $N(280, \sigma^2)$ *medida en horas. ¿Cuál debe ser el valor máximo que debe alcanzar* σ *si se quiere que el tubo fluorescente tenga una probabilidad no inferior a 0.8 de durar entre 240 y 320 horas?*

Resultado

31.13

Problema 6.7. *** Una compañía de seguros tiene 25000 pólizas de coches. El gasto en indemnizaciones anuales por cada póliza es una variable aleatoria con media 320 euros y desviación 540 euros. Aproximar la probabilidad de que el número de indemnizaciones en un año exceda los 8.3 millones de euros.*

Resultado

0.00022

Problema 6.8. **** Si lanzamos un dado n veces,*

1. *Definir una variable aleatoria que proporcione la puntuación total tras n lanzamientos.*

2. *Si lanzamos el dado 70 veces,*

   a) *Determinar el conjunto de valores que toma la variable aleatoria definida anteriormente.*

   b) *Hallar su esperanza.*

   c) *Hallar la probabilidad de obtener más de 200 puntos.*

Resultado

1. X_i = {Puntuación en el lanzamiento i-ésimo del dado} y S_n = {Suma de la puntuación obtenida en n lanzamientos del dado}, con $S_n = \sum_{i=1}^{n} X_i$, con $X_i \sim U(6)$. Nótese que la función de probabilidad de S_n no es trivial.

2. Para $n = 70$:

   a) $S_{70} = \{70, \ldots, 420\}$

   b) $E[S_{70}] = 245$

   c) 0.99918

Problema 6.9. ***** Un micro-cohete construido por los alumnos en una práctica de la ETSI tiene una probabilidad de 0.4 de fallar en el lanzamiento, caso en el que se considera que no recorre ninguna distancia. En caso de que no haya fallo en el lanzamiento, la distancia recorrida es $100 + X$ metros, donde X se distribuye según una exponencial de media 10. Determinar:*

1. *Modelar la variable aleatoria distancia D en función de X.*

2. *Función de distribución de la distancia recorrida.*

3. *Probabilidad de que la distancia recorrida esté entre 150 y 180 metros.*

4. *Asumiendo que la ocurrencia de fallo en el lanzamiento es independiente de X, hallar el valor de la distancia esperada.*

Resultado

1. $D = Y(100 + X)$, con $X \sim Exp(1/10)$ e $Y \sim Be(0.6)$.

2. $F(d) = \begin{cases} 0 & \text{si } d < 0 \\ 0.4 & \text{si } 0 \leq d \leq 100 \\ 0.4 + 0.6(1 - e^{-\frac{1}{10}(d-100)}) & d \geq 100 \end{cases}$

3. 0.0038

4. 66

Problema 6.10. ***** Una instrucción de un autómata industrial contiene 16 bits, y debe ser transmitida bit a bit en un entorno con interferencias. El valor de las interferencias puede ser aproximado como una variable aleatoria normal de valores $\mu = 4$ y $\sigma^2 = 1$. Si en el momento de transmitir el bit el valor de la interferencia es superior a 8, entonces el bit no es interpretado correctamente por el autómata. Se pide:*

1. *Obtener la probabilidad de que un bit sea interpretado correctamente.*

2. *Obtener la probabilidad de que una instrucción sea interpretada correctamente.*

3. *Obtener la probabilidad de que un programa de n instrucciones sea interpretado correctamente.*

Resultado

1. 0.99997
2. 0.9995
3. $(0.9995)^n$

Problema 6.11. ****** La probabilidad de que el AVE Madrid-Sevilla llegue puntual es del 95 %. El número de pasajeros en ese AVE se distribuye siguiendo una Poisson de parámetro $\lambda = 280$. Se pide:*

1. *Modelar la variable aleatoria T número de pasajeros que llegan en un AVE Madrid-Sevilla puntualmente.*

2. *Determinar el valor esperado de T empleando el modelo del apartado anterior.*

3. *Determinar el valor esperado de S el número total de viajeros que llegan puntualmente en una semana (7 días), sabiendo que hay 20 trenes AVE Madrid-Sevilla al día.*

4. *Calcular la varianza de T, haciendo uso de la siguiente propiedad: Si X, Y independientes, entonces $V[XY] = V[X]V[Y] + V[X]E[Y]^2 + V[Y]E[X]^2$.*

5. *Determinar la probabilidad de que S sea mayor que 37500.*

Resultado

1. $T = Z \cdot N$, siendo Z una variable aleatoria que vale 1 si el AVE llega a tiempo, y 0 en caso contrario, y N el número de viajeros que viaja en el AVE. Z sigue, por tanto, una Bernoulli de parámetro $p = 0.95$ y N (tal y como indica el enunciado) una Poisson de parámetro $\lambda = 280$.
2. $E[T] = 266$
3. $E[S] = 37240$
4. $V[T] = 3990$
5. 0.36317

Problema 6.12. ***** Un supermercado dispone de dos cajas para atender a sus clientes. El tiempo en que los clientes son atendidos en las cajas se considera una variable aleatoria exponencial. En la primera caja, los clientes son atendidos en 2 minutos de media, y en la segunda caja en 3 minutos de media. Si existe una cola de clientes diferente para cada caja:*

1. *Calcular la probabilidad de que en la primera caja, el tiempo en el que un cliente es atendido sea menor de 3 minutos y 30 segundos.*
2. *Calcular la probabilidad de que la caja 2 atienda a menos de 5 clientes en 6 minutos.*

Resultado

1. 0.8262
2. 0.947

Problema 6.13. ****** La humedad relativa semanal se puede modelar como una variable aleatoria de distribución normal de media 55 y desviación típica 20. Se asume además que la humedad relativa de cada semana es independiente de la de las demás. La densidad de frutos (número de frutos/árbol) de una determinada planta depende de la humedad relativa que se haya producido durante cada semana: si dicha humedad relativa es superior a 60, la densidad crece un 20 %, mientras que en caso contrario, decrece en un 10 %. La densidad de frutos inicial es D, y en la primera semana ésta se verá modificada según la humedad. Se pide:*

1. *Probabilidad de que, en una semana dada, la humedad relativa sea superior a 70.*
2. *Expresión de la variable aleatoria D_{k+1} densidad de frutos en la semana $k+1$ en función de la densidad de frutos en la semana anterior D_k*
3. *La densidad de frutos esperada tras 4 semanas.*
4. *El número de semanas que se debe esperar para que la densidad de frutos media sea al menos un 20 % superior a la densidad inicial D.*

Resultado

1. 0.22663
2. $D_k = D_{k-1}X_k$ para $k \geq 1$ con X_k variable concentrada en dos puntos (0.9 y 1.2) con probabilidades $p = 0.5987$ y $1 - p$ respectivamente, y $D_0 = D$.

3. $1.084D$
4. A partir de 10 semanas

Problema 6.14. ****** Se tiene un examen tipo test de n preguntas en el que cada pregunta tiene cuatro opciones. Cada pregunta acertada suma un punto, mientras que cada pregunta incorrecta resta α puntos $(0 < \alpha < 1)$. Se pide:*

1. *Modelar (en función de α y n) la variable aleatoria N puntuación obtenida si se contestasen todas las preguntas al azar suponiendo que cada pregunta solo tiene una y solo una opción correcta.*
2. *Calcular la probabilidad de sacar 10 puntos en el caso anterior si hay 10 preguntas y $\alpha = 0.25$.*
3. *Calcular de forma aproximada la probabilidad de sacar 30 puntos o más en el caso del primer apartado si hay 100 preguntas y $\alpha = 0$.*

Resultado

1. $N = (1+\alpha)X - n\alpha$ con $X \sim Bi(n, 0.25)$
2. $9.5 \cdot 10^{-7}$
3. 0.12507

Problema 6.15. ****** En cierto país, la humedad relativa de cada día viene dada por la siguiente expresión:*

$$Y = (1.5 + X)H_0$$

donde H_0 es un valor de referencia de la humedad relativa, y X es una variable aleatoria que toma valores en el intervalo $[-1, 1]$ con la siguiente función de densidad: $f(x) = A(x+1)$.

1. *Calcular el valor de la constante A, la función de distribución de X, su esperanza matemática y su varianza.*

Los valores aceptables de humedad relativa para el ser humano están entre la mitad y el doble del valor de referencia H_0. Si el valor de humedad relativa de un día en concreto es independiente del valor que haya tomado en días anteriores, se pide:

2. *Calcular la probabilidad de que un día cualquiera tenga una humedad relativa aceptable.*

3. *Modele la variable aleatoria que determina el número de días con valores aceptables de humedad relativa antes del primer día con valor inaceptable.*

4. *Modele la variable aleatoria que determina el número de días con valores aceptables antes del tercer día inaceptable.*

5. *Modele la variable aleatoria que determina el número máximo de días consecutivos con valores de humedad aceptables antes del tercer día inaceptable. Calcule la probabilidad de que dicha variable sea mayor que 2.*

6. *Calcule de forma aproximada la probabilidad de que existan entre 5 y 100 días de humedad aceptable en los próximos 160 días.*

Resultado

1. $A = 1/2$; $F(x) = \frac{1}{4}(x^2 + 2x + 1)$ con $-1 \leq x \leq 1$; $E[X] = \frac{1}{3}$; $V[X] = \frac{2}{9}$
2. 0.5625
3. $W \sim Ge(7/16)$
4. $T \sim BN(3, 7/16)$
5. $F = \max\{G_1, G_2, G_3\}$, con $G_i \sim Ge(7/16)$; $P[F > 2] = 0.4445$
6. 0.94408

Problema 6.16. ****** Una hormigonera tiene, en media, una avería por cada 1000 horas de funcionamiento. Esta hormigonera se utiliza en las obras de manera que, en un día (considerado como 10 horas de trabajo) concreto, o bien se arranca al principio del día y se mantiene en funcionamiento las 10 horas del día, o bien no se arranca porque ese día no es necesario. Se pide:*

1. *Probabilidad de que la hormigonera no se averíe en un día de funcionamiento.*

2. *Probabilidad de que transcurran al menos 5 días consecutivos de funcionamiento sin averías.*

3. *Determinar el número esperado de averías en una obra de 40 días de duración en la que todos los días se arranca la hormigonera. Desviación típica del número de averías en dicha obra.*

4. *En una obra determinada, la probabilidad de que en un día concreto sea necesario arrancar la hormigonera es de 0.15, con independencia de si se arrancó o no el día anterior. Si la obra dura 20 días, determinar el número esperado de averías de la hormigonera.*

Resultado

1. 0.99
2. 0.9512
3. 0.4 y 0.6325
4. 0.03

Problema 6.17. ****** Una empresa fabrica bombillas que vende en lotes de* 100 *bombillas. La política de la empresa es tal que, si al menos una bombilla del lote es defectuosa, se considera defectuoso todo el lote. Las bombillas se fabrican en dos máquinas de distinta antigüedad, por lo que tienen distintas tasas de defectuosos. En la primera máquina, la probabilidad de que una bombilla sea defectuosa es de* 0.001*, mientras que en la segunda es de* 0.01*. Se pide:*

1. *Si el 50 % de las bombillas de un lote han sido fabricadas en la primera máquina, determinar la probabilidad de que el lote no sea defectuoso.*
2. *Si el 50 % de las bombillas de un lote han sido fabricadas en la primera máquina, determinar la probabilidad de que haya una única bombilla defectuosa en un lote.*
3. *Si el 50 % de las bombillas de un lote han sido fabricadas en la primera máquina, determinar de forma aproximada la probabilidad de que haya más de 200 lotes defectuosos en un envío con 500 lotes.*
4. *Suponiendo que cada lote no defectuoso genera unos beneficios de* 100 *€mientras que cada lote defectuoso genera unas pérdidas de* 100 *€, determinar el porcentaje mínimo de bombillas de un lote que deberían ser fabricados en la primera máquina para que los beneficios esperados por lote sean positivos.*

Resultado

1. 0.576
2. 0.319
3. 0.8665
4. 0.344

Problema 6.18. ****** El servidor del FBI recibe ataques cibernéticos que son detectados por su firewall. La media de ataques por hora es de 120. Con el objeto de estudiar si es necesario un aumento de la seguridad del servidor, se decide esperar n ataques y medir el tiempo que pasa entre uno y otro. Suponiendo que el tiempo que pasa entre dos ataques no determina cuándo aparecerá el siguiente, se pide:*

1. *Probabilidad de que el menor de los tiempos entre un ataque y el siguiente esté entre 1 y 2 minutos, en función de n.*

2. *Probabilidad de que el mayor tiempo entre un ataque y el siguiente sea inferior a 3 minutos, en función de n.*

3. *Utilizando el resultado del apartado anterior, determinar la función de densidad de la variable máximo tiempo (en minutos) entre un ataque y el siguiente, en función de n.*

4. *Función de densidad de la variable aleatoria tiempo medio entre un ataque y el siguiente, en función de n.*

5. *Si $n = 50$, determine de forma aproximada la probabilidad de que el tiempo medio entre un ataque y el siguiente esté entre medio minuto y 0.75 minutos.*

Resultado

1. $e^{-2(n-1)} - e^{-4(n-1)}$
2. $(1 - e^{-6})^{n-1}$
3. $f(t) = (n-1)\left(1 - e^{-2t}\right)^{n-2} \cdot 2e^{-2t}, t > 0$
4. $f(\overline{x}) = \frac{(2(n-1))^{n-1}}{\Gamma(n-1)} \overline{x}^{n-2} e^{-2(n-1)\overline{x}}$
5. 0.49979

Problema 6.19. ****** Un comerciante se plantea montar un negocio especializado en venta de placas solares para uso doméstico. El empresario contaría con tres tipos diferentes de placas solares para poner a la venta. Se sabe que, en el caso de emprender el negocio, las ventas de la placa 1 al día, en locales similares, siguen aproximadamente una distribución $N(20, 25)$ con beneficio para el empresario de 2.5 euros por cada placa de este tipo vendida; las ventas de la placa 2 cada día una $N(10, 20)$ con beneficio 6 euros por unidad; mientras que la placa 3 tiene un número de ventas cada día de $N(25, 50)$ con beneficio de 3.5 euros. Se deben de tener en cuenta los gastos fijos independientes mensuales de 2000 euros debido al alquiler de local y gastos de mantenimiento. Se pide:*

1. *Modelar (definiendo también cómo se distribuyen) las ventas al mes del producto 1, 2 y 3 considerando que un mes está formado por 22 días laborables. Calcular, adicionalmente, la probabilidad de que se vendan más de 450 unidades del producto 1 y la probabilidad de que en total se vendan más de 1150 productos.*

2. *Modelar cómo sería el beneficio del comerciante al mes, suponiendo que las ventas siguen la misma distribución que las de los locales similares.*

3. *Definir la distribución que sigue el beneficio y calcular la probabilidad de que sea mayor que 2200 euros.*

Se decide emprender el negocio en la ciudad de Sevilla.

4. *Se realiza un estudio dando como resultado que existen 45.000 personas potencialmente interesadas en el producto. Si se contrata una campaña de una compañía de comunicación que nos asegura que la mitad de ellos será conocedor de nuestro negocio y tendrán una probabilidad del 1 % de visitarnos en un día, ¿Cuál es la probabilidad aproximada de que nos visiten más de 250 personas cada día?*

Resultado

1. 0.9049
2. Sea B={Beneficio mensual del comerciante}, g={Gasto fijo mensual}= 2000 € y $b_1 = 2.5$ € ; $b_2 = 6$ €; $b_3 = 3.5$ € los beneficios unitarios de las placas 1, 2 y 3 respectivamente.

$$B = VM_1 \cdot b_1 + VM_2 \cdot b_2 + VM_3 \cdot b_3 - g$$

3. $B \sim N(2345, 32752.5)$. 0.78814
4. 0.047

Problema 6.20. ****** El número medio de accesos al servidor web de la Universidad de Sevilla en un segundo es constante e igual a λ.*

1. *Modelar la variable aleatoria N = {número de accesos al servidor en un intervalo de un segundo}*
2. *Con la entrada de los nuevos grados se desea aumentar la capacidad del servidor poniendo otro en paralelo que contenga la información de dichos nuevos grados. Se estima que el número esperado de accesos por segundo al nuevo servidor es $\lambda_2 = 4$. El proveedor oferta tres tipos de servidores con diferentes memorias. El número máximo de accesos por segundo para que cada memoria no se bloquee es:*

Memoria 1	*Memoria 2*	*Memoria 3*
5	*7*	*10*

El precio se incrementa según las prestaciones. Si se desea que el servidor se bloquee con una probabilidad menor del 6 %, hallar la memoria de menor coste que cumpla con este requisito.

3. *Ahora la Universidad dispone de los dos servidores, uno con la información de las antiguas titulaciones y el número de accesos esperado por segundo $\lambda_1 = 8$, y otro con la información de los nuevos grados y el número de accesos esperado por segundo $\lambda_2 = 4$.*

Para registrar la información de los usuarios que acceden a ambos servidores, la Universidad los ha conectado a un servidor principal que funciona como base de datos. Como dicho servidor principal atenderá los accesos de ambos servidores, se desea decidir la velocidad del procesador que se le implementará (en Hertzios= $\frac{1}{segundos}$*), midiendo la frecuencia esperada de los accesos al servidor principal.*

a) *Modelar la variable T que mide el tiempo (en segundos) entre accesos al servidor principal.*

b) *¿Qué distribución sigue la variable aleatoria $F = \{frecuencia\ de\ accesos\ al\ servidor\ principal\}$ y cuál es la frecuencia esperada?*

Resultado

1. $N \sim \mathcal{P}o(\lambda)$
2. Memoria 2.
3. Nuevo caso:
   a) $T \sim Exp(12)$
   b) 12 Hertzios.

Problema 6.21. *En un hospital los pacientes que son operados en la unidad de cirugía son enviados a planta tras la operación para su hospitalización. La hospitalización es el tiempo que está el paciente ingresado desde el día después de la operación hasta que el cirujano le da el alta. El cirujano pasa por planta para ver a sus pacientes todos los días a partir del día de la operación, y decidir si les da el alta o no. La probabilidad de que el cirujano dé el alta al paciente es $p = 0.2$.*

1. *Hallar la probabilidad de tener el alta el mismo día de la operación.*
2. *Hallar la probabilidad de tener el alta tras 8 días hospitalizado.*
3. *Si en un día en planta hay 24 pacientes hospitalizados, hallar la probabilidad de darle el alta a 4 de ellos.*
4. *Unos de los mayores costes del hospital es la hospitalización, c unidades monetarias por día. ¿Cuál es el coste medio de hospitalización por paciente?*
5. *Al final de una semana se han dado 37 altas en la unidad. Modelar la variable aleatoria del coste total de las 37 hospitalizaciones. Escribir la expresión de la probabilidad exacta de que el coste total de esas 37 hospitalización sea menor o igual que una cantidad d.*
6. *Calcular la probabilidad aproximada de que el coste total de la hospitalización de los 37 pacientes sea menor que 100c unidades monetarias.*

Resultado

1. 0.2
2. 0.042
3. 0.196
4. $4c$ unidades monetarias por paciente.
5. $S \sim BN(37, 0.2)$. $P\left[S \leq \frac{d}{c}\right] = \sum_{k=0}^{Ent(d/c)} \binom{k+37-1}{k} (0.8)^k (0.2)^{37}$
6. 0.0392

Problema 6.22. *Una marca de bombillas experimentales LED tiene un tiempo de vida X que sigue una distribución exponencial de media 2 años en condiciones ideales. Este tipo de bombillas es sensible a la temperatura ambiente del entorno donde se utiliza, de forma que la vida de la bombilla se modifica en un factor $k > 0$ dependiendo del entorno en donde se utilice de la siguiente forma:*

$$Y = kX$$

Suponiendo que las bombillas se fabrican de forma independiente, se pide:

1. *Determine la función de distribución de la variable Y, a partir de la función de generatriz de momentos de X. ¿Qué distribución sigue?*
2. *Si en una habitación con un factor de temperatura k =1, se activan 5 bombillas de forma simultánea e ininterrumpida.*
   a) *Calcule la probabilidad de que al cabo de 3 años sólo una bombilla se haya fundido.*
   b) *Calcule la probabilidad de que la habitación tenga luz durante más de 6.5 años de forma ininterrumpida.*
3. *Suponga que en esta ocasión, en una habitación con un factor de temperatura k, se dispone un dispositivo con tres bombillas de forma que inicialmente hay una bombilla funcionando y dos de repuesto sin funcionar, y al fallar una bombilla activa, se pone en funcionamiento una de las bombillas de repuesto, hasta agotar el número de bombillas.*

   Si k=1.25, calcule la probabilidad de que, con un dispositivo análogo, la habitación tenga luz durante más de 6.5 años de forma ininterrumpida.

Resultado

1. $Y \sim Exp(\frac{1}{2k})$
2. Caso $k = 1$
   a) 0.0096
   b) 0.1795
3. 0.5184

Problema 6.23. *En una determinada zona, las precipitaciones en un día (mm. por metro cuadrado, m^2) se distribuyen según una normal de media 15 y varianza 25. En dicha zona se está construyendo una presa que puede soportar hasta 5 días consecutivos en los que las precipitaciones diarias sean superiores a 25 mm. por m^2. Más allá de ese límite, existe riesgo de rotura de la presa. Se pide:*

1. *Probabilidad de que en un día cualquiera las precipitaciones superen los 25 mm. por m^2.*
2. *Probabilidad de que la presa esté en riesgo de rotura, asumiendo que la cantidad de lluvia en un día es independiente de la del resto.*

También hay riesgo rotura de la presa en el caso en el que haya 50 días de lluvia de más de 20 mm. por m^2 de forma intermitente.

3. *Modelar la variable aleatoria del número de dias que llueve menos de 20 mm. por m^2 antes del quincuagésimo día de lluvias de más de 20 mm. por m^2.*
4. *Calcular la probabilidad aproximada de que haya riesgo de rotura en un periodo de exactamente 200 días.*
5. *Calcular aproximadamente la probabilidad de que haya riesgo de rotura de la presa en menos de 200 días.*

Resultado

1. 0.02275
2. $\simeq 0$
3. $Y \sim BN(50, 0.15866)$
4. 0.00021
5. 0.00248

Problema 6.24. *Una torre de perforación petrolífera está dispuesta a excavar un pozo de 3000 metros de profundidad 24 horas al día. Se sabe que, en terrenos parecidos, torres de perforación de similares características tienen una media de 50 fallos de operación por cada 3000 metros perforados. Si cada fallo es independiente del anterior, se pide:*

1. *Calcular la probabilidad de que se produzcan entre 49 y 52 fallos durante la excavación.*
2. *Calcular la probabilidad de que ocurra exactamente un fallo antes de los 40 metros de perforación.*
3. *Calcular la probabilidad de que se perfore menos de 50 metros entre dos fallos consecutivos.*

Asumiendo que el número de fallos hasta perforar los 3000 metros será justamente el esperado, se pide:

4. *Calcular la probabilidad de que la mínima distancia excavada entre fallos sea mayor de 4 metros.*
5. *Sabiendo que el tiempo en horas de reparación de un fallo se comporta como una $Exp(0.5)$, que el tiempo en horas de perforación entre fallos se comporta como una $Exp(0.125)$, y que ambos son independientes entre sí:*
    a) *Modelar las variables aleatorias que miden los tiempos totales de reparación y de perforación desde el inicio de la excavación hasta el último fallo.*
    b) *Calcular de forma aproximada la probabilidad de que el último fallo se repare antes de que pasen 20 días.*

Resultado

1. 0.22096
2. 0.342278
3. 0.5654
4. 0.03567
5. a) $TR = \{$tiempo total de reparación$\}$, $TR \sim Ga(50, 0.5)$. $TP = \{$tiempo total de perforación$\}$, $TP \sim Ga(50, 0.125)$
    b) 0.36693

Problema 6.25. *El tiempo que los autobuses de la línea circular C-1 tardan en completar una vuelta se distribuye según una exponencial. Cada día se completan 10 vueltas, considerándose que una vuelta es completada con retraso si se realiza en más de 80 minutos. Se sabe que, en un día lluvioso (10 % de los días), la probabilidad de que al menos una vuelta se complete con retraso es del 80 %, mientras que en caso contrario, dicha probabilidad baja al 20 %. Se pide:*

1. *Probabilidad de que se complete una (y solo una) vuelta con retraso en un día lluvioso.*
2. *Probabilidad de que una vuelta se complete en menos de 60 minutos en un día lluvioso.*
3. *Probabilidad de que se complete una vuelta en menos de 60 minutos en un día cualquiera.*
4. *Probabilidad de que, en un día lluvioso, la vuelta más lenta se complete en más de 90 minutos.*
5. *Probabilidad de que la duración de menos del 20 % de las vueltas en una semana cualquiera (7 días) esté por debajo de su valor esperado.*

Resultado

1. 0.3493
2. 0.76
3. 0.9265
4. 0.7132
5. ≈ 0

Problema 6.26. ****** El tiempo que tarda un automóvil de Fórmula 1 en dar una vuelta a un determinado circuito se asume que se distribuye según una exponencial que no depende del tiempo que haya tardado en dar la vuelta anterior. Al realizar varias series de entrenamiento de 4 vueltas cada una, se tiene que la vuelta más rápida es superior a 2 minutos en el 10 % de las series. Se pide:*

1. *Probabilidad de que tarde entre 1 y 2 minutos en la vuelta más rápida de una serie. ¿Y en una vuelta cualquiera?.*
2. *Probabilidad de que, en una serie, la vuelta más rápida sea la última de la serie.*
3. *Probabilidad de que haya tardado más de 2 minutos en dos de las cuatro vueltas de una serie.*
4. *Si el automóvil da vueltas al circuito durante 5 minutos seguidos, la probabilidad de que haya dado las cuatro vueltas de la serie.*
5. *Número mínimo de vueltas que tendría que dar el automóvil para que haya al menos un 90 % de probabilidad de que la vuelta más rápida de todas sea menor de 1.5 minutos.*

Resultado

1. 0.2166 y 0.1875

2. 0.25
3. 0.36344
4. 0.0423
5. 6

Problema 6.27. ****** El número de ingresos en un determinado hospital se distribuye como una Poisson cuya media es λ ingresos/año. Si la probabilidad de que un ingresado fallezca es* 0.05*:*

1. *Determinar, en función de λ, la función de distribución de la variable aleatoria $X :=$ \{ Tiempo (años) que transcurre sin que haya ningún ingreso en el hospital\}.*
2. *Determinar, en función de λ, la función de densidad de la variable aleatoria $Y :=$ \{ Tiempo medio (años) que transcurre en producirse cinco ingresos en años\}.*
3. *Calcular de forma aproximada λ, usando de forma justificada el TCL, si la probabilidad de que el número de ingresos al mes sea mayor que* 96 *es* 0.121.
4. *Determinar qué distribución sigue la variable aleatoria W=\{número de fallecimientos al año en el hospital, sabiendo que ha habido N ingresos en el año\}.*
5. *Determinar la función de probabilidad de la variable aleatoria M=\{número de fallecimientos al año en el hospital\}. ¿Qué distribución sigue M?*

   Nota: Para este apartado use
   $$\sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x$$

   Resultado

1. $F(x) = 1 - e^{-\lambda x}$
2. $f(y) = \frac{(5\lambda)^5}{4!} y^4 e^{-5\lambda y}, y > 0.$
3. 1022.31
4. $P[W = m] = P[M = m | I = N] = \binom{N}{m} 0.05^m 0.95^{N-m}$, $m = 0, \ldots, N$
5. $M \sim Po(0.05\lambda)$

Problema 6.28. ****** Una máquina procesa una pieza tras otra sin interrupción. El tiempo que tarda en procesar cada pieza se considera distribuido según una exponencial de media 25 segundos. Se pide:*

1. *Probabilidad de que se puedan procesar tres o más piezas en un minuto.*
2. *Probabilidad de que se tarde más de un minuto en procesar dos piezas.*
3. *Probabilidad de que, en un conjunto de cuatro piezas, ninguna de ellas haya tardado más de 22 segundos en procesarse.*
4. *Probabilidad de que, en un conjunto de cuatro piezas, haya dos piezas en las que se haya tardado más de 28 segundos en procesarse cada una.*
5. *Probabilidad aproximada de que se tarde menos de una hora en procesar 150 piezas.*

Resultado

1. 0.4303
2. 0.3084
3. 0.1173
4. 0.2900
5. 0.30854

Capítulo 7

Estimación Confidencial

7.1. Concepto de intervalo de confianza

Definición (Intervalo de confianza): Sea $(X_1, \ldots, X_n)$ una *mas* siguiendo la distribución de X cuya función de distribución depende de un parámetro real θ. Se define un **intervalo de confianza** (IC) para θ con **nivel de confianza** $1 - \alpha$, $\alpha \in (0, 1)$, como el intervalo:

$$IC_{1-\alpha}(\theta) = [T_1(X_1, \ldots, X_n), T_2(X_1, \ldots, X_n)]$$

con $T_1(X_1, \ldots, X_n)$ y $T_2(X_1, \ldots, X_n)$ tales que

$$P[T_1(X_1, \ldots, X_n) \leq \theta \leq T_2(X_1, \ldots, X_n)] \geq 1 - \alpha$$

El intervalo de confianza es un *intervalo aleatorio* (determinado por los extremos de dos variables aleatorias función de la *mas*) y, por tanto, para cada realización muestral $(x_1 \ldots, x_n)$ se obtendrá un intervalo numérico (figura 7.1). En ese sentido, la expresión $P[T_1(X_1, \ldots, X_n) \leq \theta \leq T_2(X_1, \ldots, X_n)] \geq 1-\alpha$ debe interpretarse de la siguiente forma: con una probabilidad superior a $1 - \alpha$, el intervalo que genera una determinada realización muestral contiene a θ. No obstante, por abuso del lenguaje, es frecuente utilizar la expresión 'θ está contenida en el intervalo de confianza'.

Para un determinado nivel de confianza y una determinada realización muestral se pueden obtener múltiples intervalos de confianza, no todos ellos necesariamente de la misma amplitud (ver figura 7.2). De entre ellos, nos interesará el de longitud mínima.

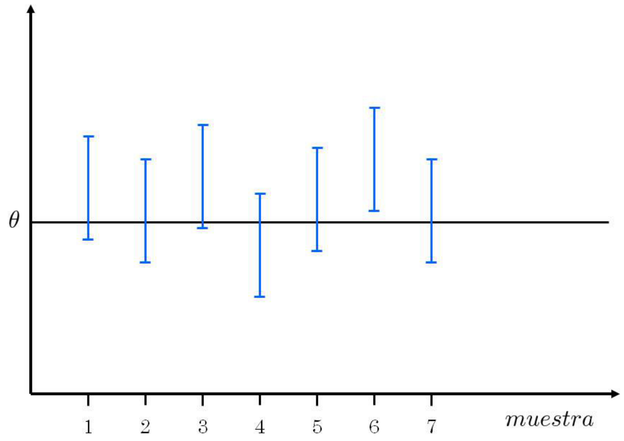

Figura 7.1: Intervalos de confianza para θ para varias muestras

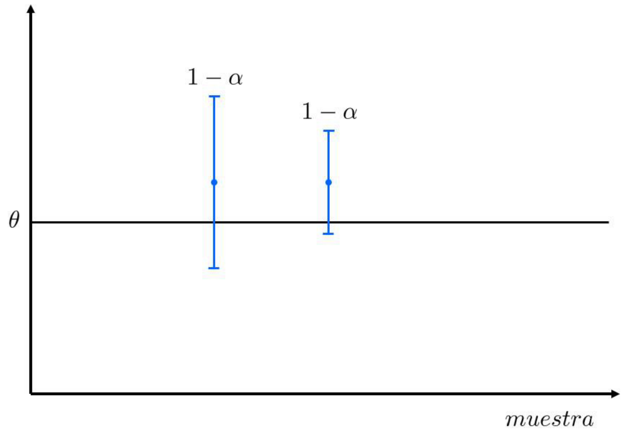

Figura 7.2: Intervalos de confianza para θ con un mismo nivel de confianza, $1-\alpha$

7.2. Introducción al contraste de hipótesis

Sea X una variable aleatoria cuya distribución depende del parámetro θ desconocido, para la que se quiere comprobar que dicho parámetro toma el valor θ_0, es decir, se desea saber si $\theta = \theta_0$. Se tiene que, con un nivel de confianza $1-\alpha$, el intervalo $IC_{1-\alpha}(\theta)$ debería contener a θ_0.

Si se obtiene una realización muestral y dicha inclusión no se verifica, diremos que **se rechaza la hipótesis** $\theta = \theta_0$. Por el contrario, si se verifica, diremos que **no existen evidencias para rechazar la hipótesis** y, por lo tanto, podremos asumir que $\theta = \theta_0$.

Más formalmente, en el contraste de hipótesis se formularán dos hipótesis:

- H_0 o **hipótesis nula**, que en nuestro caso asignará un valor a un parámetro de la distribución ($\theta = \theta_0$).

- H_1 o **hipótesis alternativa**, y que consiste en otra hipótesis referente al valor del parámetro (en general serán valores distintos a θ_0).

Para ser rigurosos, el contraste de hipótesis se debe enunciar de forma conservadora según hemos visto anteriormente: si el parámetro no está en el intervalo de confianza rechazamos H_0, pero si lo está, simplemente se dirá que no se puede (a la luz de los datos) rechazar H_0.

7.3. Intervalo de confianza para cualquier variable aleatoria

7.3.1. IC para la media con varianza conocida

IC para la media con varianza conocida: Sea una variable aleatoria X con esperanza μ y varianza σ^2 conocida. Para un nivel de confianza $1-\alpha$, el siguiente intervalo:

$$IC_{1-\alpha}(\mu) = \left[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

es un intervalo de confianza de μ cuando $n \to \infty$, donde $z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$

Respecto al teorema anterior, cabe hacer los siguientes comentarios:

- El intervalo requiere que $n \to \infty$ para poder aplicar el TCL. Por ello, la bondad de esta aproximación crecerá con n, y por lo tanto será interesante tomar *mas* de elevado tamaño.

- Si la distribución que sigue X (y por lo tanto los componentes de la *mas*) es normal con esperanza μ y varianza σ^2, el IC es exacto para cualquier n.

- La determinación del intervalo de confianza anterior requiere conocer σ, hipótesis que, como se ha comentado, no es realista en muchas aplicaciones de interés.

Dada X una variable aleatoria de esperanza μ desconocida y varianza σ^2 conocida, una aplicación de este IC es realizar un contraste de hipótesis para μ. Si suponemos que X tiene la esperanza μ_0, tenemos que, con un nivel de confianza $1-\alpha$, el intervalo $\left[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$ debería contener a μ_0 con una probabilidad $1-\alpha$. Si se obtiene una realización muestral y dicha inclusión no se verifica, diremos que **se rechaza la hipótesis** $\mu = \mu_0$. Por el contrario, si se verifica, diremos que **no existen evidencias para rechazar la hipótesis** y, por lo tanto, podremos asumir que $\mu = \mu_0$.

7.4. Suposición de normalidad. Distribuciones asociadas a la normal

En el apartado anterior, se ha determinado una aproximación a un intervalo de confianza, que en el caso particular en el que la variable aleatoria es normal el intervalo es exacto. En lo que sigue, a efectos de estimación confidencial, asumiremos que las variables aleatorias siguen una distribución normal debido a que:

1. Numerosos fenómenos naturales siguen la distribución normal, como por ejemplo:
   - La cinemática (velocidad, energía, momento) de los gases en equilibrio termodinámico sigue una distribución normal (ecuaciones de Maxwell-Boltzmann).
   - En procesamiento de señales, la función de impulso (delta de Dirac) es un caso particular de distribución normal.
2. Muchos fenómenos aleatorios que siguen otras distribuciones pueden aproximarse con gran precisión a la distribución normal. En general, y debido al TCL, podremos asumir que cuando la variable objeto de estudio sea el resultado de la suma de un número suficientemente grande de efectos que actúan de forma independiente, esta variable sigue una distribución normal.
3. El 'buen comportamiento' de la distribución normal (a efectos de poder encontrar ICs fácilmente para su esperanza y varianza) hace que, en primera instancia, sea conveniente que consideremos que las variables aleatorias son normales y, en todo caso, verificar si se alejan o no de dicha distribución.

Por tanto, los intervalos de confianza que se van a estudiar serán los asociados a los parámetros de una variable aleatoria normal, lo que simplifica notablemente la obtención de los intervalos de confianza. No obstante, antes de enunciarlos, deberemos introducir una serie de distribuciones basadas en la distribución normal que nos servirán para obtenerlos.

7.4.1. Distribuciones basadas en la distribución normal

Definición (Distribución Chi cuadrado de Pearson): Sean $Z_1, \ldots, Z_n$ variables aleatorias independientes e idénticamente distribuidas $Z_i \sim N(0,1)$, $i = 1, \ldots, n$. Se tiene que $J_n = Z_1^2 + \cdots + Z_n^2$ sigue una **distribución chi cuadrado de Pearson** con n grados de libertad, y se nota por $J_n \sim \chi_n^2$.

La figura 7.3 muestra la función de densidad de la distribución chi cuadrado para diferentes casos de n.

Para una variable aleatoria $J_n \sim \chi_n^2$, notaremos $j_{n,\alpha}$ como aquel valor que verifica que $P[J \leq j_{n,\alpha}] = 1 - \alpha$. Análogamente $P[J > j_{n,\alpha}] = \alpha$ (ver figura 7.4).

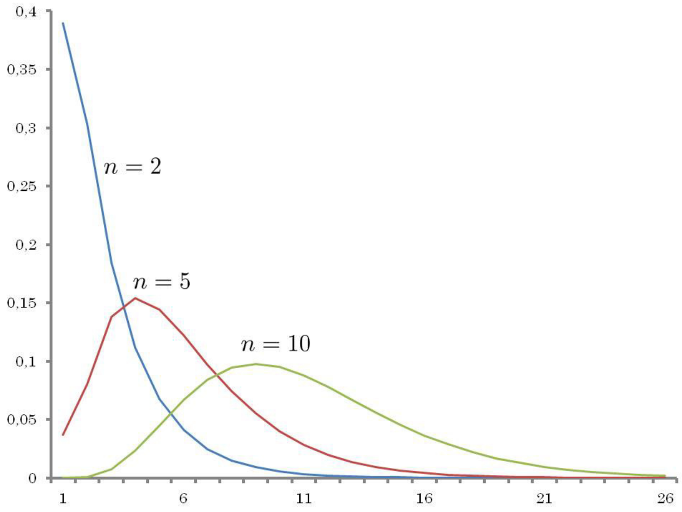

Figura 7.3: Función de densidad de la distribución χ^2 para $n = 2, 5, 10$

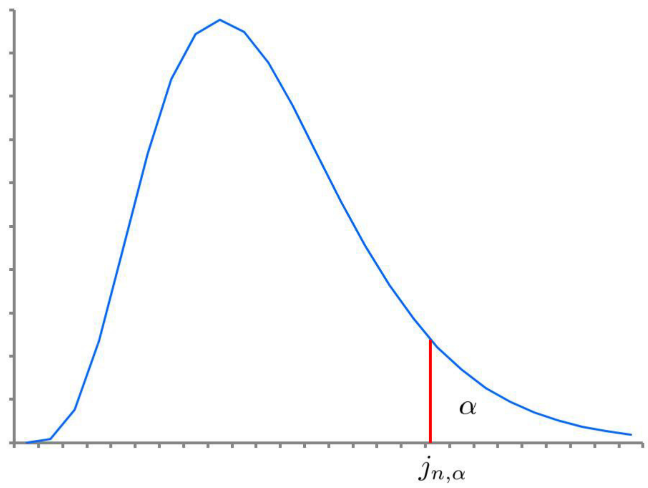

Figura 7.4: Distribución χ^2: $P[J > j_{n,\alpha}] = \alpha$

Al ser una suma de variables aleatorias normales al cuadrado, los valores de su función de distribución se obtienen de forma tabular.

Definición (Distribución t de Student): Sea $Z \sim N(0,1)$ y $J \sim \chi^2_n$ independientes. La variable aleatoria $T_n = \frac{Z}{\sqrt{J/n}}$ sigue una **distribución t de Student** de parámetro n, y se nota por $T_n \sim t_n$.

Para una variable aleatoria $T_n \sim t_n$, denotaremos por $t_{n,\alpha}$ al valor que verifica que $P[T_n \leq t_{n,\alpha}] = 1 - \alpha$. Por lo tanto $P[T_n > t_{n,\alpha}] = \alpha$.

Los valores de la función de distribución deben ser calculados a partir de los valores de la normal y de la χ^2, por lo que se presentan tabulados. No obstante, como su función de densidad es simétrica respecto del origen, se tiene que $P[T_n \leq -t] = 1 - P[T_n \leq t]$, lo que permite reducir la dimensión de la tabla.

Definición (Distribución $\mathcal{F}$ de Snedecor): Sean $J_1 \sim \chi^2_n$ y $J_2 \sim \chi^2_m$ independientes. La variable aleatoria $F_{n,m} = \frac{J_1/n}{J_2/m}$ sigue una **distribución $\mathcal{F}$ de Snedecor** con m y n grados de libertad, y se nota por $F_{n,m} \sim \mathcal{F}_{n,m}$.

Para una variable aleatoria $F_{n,m} \sim \mathcal{F}_{n,m}$, denotaremos por $f_{n,m;\alpha}$ al valor que verifica que $P[F_{n,m} \leq f_{n,m;\alpha}] = 1 - \alpha$. Por lo tanto $P[F_{n,m} > f_{n,m,\alpha}] = \alpha$ (ver la figura 7.5).

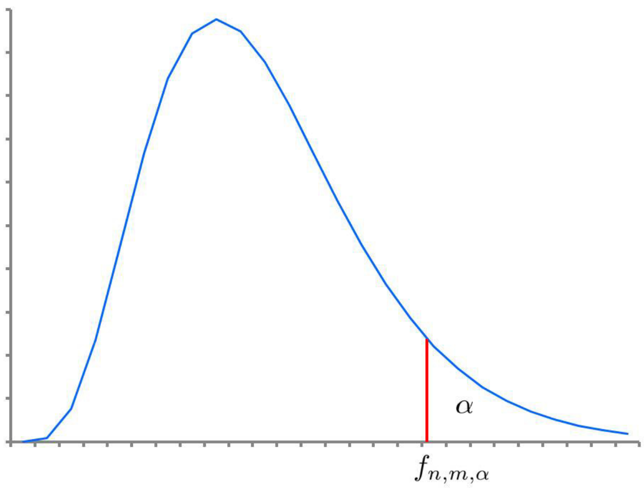

Figura 7.5: Distribución $\mathcal{F}$ de Snedecor $P[F_{n,m} > f_{n,m,\alpha}] = \alpha$

Los valores de la función de distribución deben ser calculados a partir de los de la χ^2, por lo que se presentan tabulados. No obstante, se puede reducir la dimensión de la tabla teniendo en cuenta que $F' = \frac{1}{F} = \frac{J_2/m}{J_1/n}$, y que por lo tanto $F' \sim \mathcal{F}_{m,n}$. Como consecuencia de lo anterior, se tiene que:

$$f_{m,n,1-\alpha} = \frac{1}{f_{n,m;\alpha}}$$

ya que

$$1 - \alpha = P[F \leq f_{n,m;\alpha}] = P\left[\frac{1}{F} \geq \frac{1}{f_{n,m;\alpha}}\right] = P\left[F' \geq \frac{1}{f_{n,m;\alpha}}\right]$$

Como $F' = \frac{1}{F} \sim \mathcal{F}_{m,n}$, se tiene que $f_{m,n,1-\alpha}$ es el valor que verifica que $1 - \alpha = P\left[F' \geq f_{m,n,1-\alpha}\right]$, por lo que $f_{m,n,1-\alpha} = \frac{1}{f_{n,m;\alpha}}$.

7.4.2. Teorema de Fisher

Teorema de Fisher: Sea $(X_1, \ldots, X_n)$ una *mas* siguiendo la distribución de $X \sim N(\mu, \sigma^2)$, se tiene que $\overline{X}$ y S^2 son independientes.

Como consecuencia del teorema de Fisher, se tiene que:

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

7.5. Intervalos de confianza para una variable aleatoria normal

En este apartado utilizaremos los resultados obtenidos en el apartado anterior para construir tanto intervalos de confianza para σ como para μ en el caso de σ desconocida (ya que el caso de μ con σ conocida ya ha sido tratado en general).

7.5.1. IC para la varianza

IC para la varianza: Sea una variable aleatoria $X \sim N(\mu, \sigma^2)$. Para un nivel de confianza $1-\alpha$, el siguiente intervalo:

$$IC_{1-\alpha}(\sigma^2) = \left[\frac{(n-1)S^2}{j_{n-1,\alpha/2}}, \frac{(n-1)S^2}{j_{n-1,1-\alpha/2}}\right]$$

es un intervalo de confianza de σ^2.

Este intervalo también puede aplicarse al contraste de hipótesis. En este caso, la hipótesis nula será que $\sigma^2 = \sigma_0^2$, y se verificará, con un nivel de confianza de $1-\alpha$, si la realización muestral del IC anterior incluye a σ_0^2.

7.5.2. IC para la media con varianza desconocida

IC para la media con varianza desconocida: Sea una variable aleatoria $X \sim N(\mu, \sigma^2)$, con σ^2 desconocida. Para un nivel de confianza $1-\alpha$, el siguiente intervalo:

$$IC_{1-\alpha}(\mu) = \left[\overline{X} - t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}, \overline{X} + t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}\right]$$

es un intervalo de confianza de μ.

Al igual que en casos anteriores, este IC nos permite realizar un contraste de hipótesis sobre la media, sin tener que aproximar σ mediante s.

7.5.3. IC para la predicción de una observación futura

IC para la predicción de una observación futura con varianza conocida: Sea una variable aleatoria X con esperanza μ y varianza σ^2 conocida. Para un nivel de confianza $1-\alpha$, el siguiente intervalo:

$$IC_{1-\alpha}(X_{n+1}) = \left[\overline{X} - z_{\alpha/2}\sigma\sqrt{1+1/n}, \overline{X} + z_{\alpha/2}\sigma\sqrt{1+1/n}\right]$$

es un intervalo de confianza de X_{n+1}, donde $z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$

7.5.4. IC de una cola

IC de una cola para la media con varianza conocida: Sea una variable aleatoria X con esperanza μ y varianza σ^2 conocida. Para un nivel de confianza $1-\alpha$, los siguientes intervalos:

$$IC^i_{1-\alpha}(\mu) = \left[-\infty, \overline{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right]$$

$$IC^s_{1-\alpha}(\mu) = \left[\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}, +\infty\right]$$

son dos intervalos de confianza de μ, donde $z_\alpha = \Phi^{-1}(1-\alpha)$

Este IC nos permite entonces realizar un contraste de hipótesis sobre la media, sabiendo el valor σ, utilizando sólo uno de los extremos del intervalo. Sin embargo, el contraste de hipótesis para este intervalo de confianza cambia ligeramente respecto a los contrastes vistos con anterioridad. Por un lado, suponiendo un nivel de confianza $1-\alpha$, el intervalo $IC^i_{1-\alpha}(\mu)$ del enunciado debería contener a μ_0 con una probabilidad $1-\alpha$. Si se obtiene una realización muestral y dicha inclusión se verifica, diremos que **no existen evidencias para rechazar la hipótesis** y, por lo tanto, podremos asumir que $\mu < \mu_0$. Por el contrario, si no se verifica, diremos que **se rechaza la hipótesis** y, por tanto, $\mu \geq \mu_0$ ya que sabemos que μ además de ser distinto a μ_0, debe ser inferior.

Procediendo de manera análoga para el intervalo $IC^i_{1-\alpha}(\mu)$, tenemos que la hipótesis nula es $\mu \leq \mu_0$ y la alternativa $\mu > \mu_0$.

Siguiendo un procedimiento análogo, es posible demostrar los siguientes intervalos de una cola:

IC de una cola para la media con varianza desconocida: Sea una variable aleatoria $X \sim N(\mu, \sigma^2)$, con σ^2 desconocida. Para un nivel de confianza $1-\alpha$, los siguientes intervalos:

$$IC_{1-\alpha}(\mu) = \left[-\infty, \overline{X} + t_{n-1,\alpha} \frac{S}{\sqrt{n}}\right]$$

$$IC_{1-\alpha}(\mu) = \left[\overline{X} - t_{n-1,\alpha} \frac{S}{\sqrt{n}}, +\infty\right]$$

son dos intervalos de confianza de μ.

IC de una cola para la varianza: Sea una variable aleatoria $X \sim N(\mu, \sigma^2)$. Para un nivel de confianza $1-\alpha$, los siguientes intervalos:

$$IC_{1-\alpha}(\sigma^2) = \left[-\infty, \frac{(n-1)S^2}{j_{n-1,1-\alpha}}\right]$$

$$IC_{1-\alpha}(\sigma^2) = \left[\frac{(n-1)S^2}{j_{n-1,\alpha}}, +\infty\right]$$

son intervalos de confianza de σ^2.

Como puede verse, no es necesario memorizar estos intervalos si se tienen los correspondientes de dos colas, ya que se obtienen sustituyendo $\alpha/2$ por α en el límite (inferior o superior) correspondiente.

7.6. Intervalos de confianza para dos variables aleatorias normales

7.6.1. IC para la diferencia de medias

IC para la diferencia de medias con varianzas conocidas: Dado un nivel de confianza $1-\alpha$, se tiene que el intervalo:

$$IC_{1-\alpha}(\mu_X - \mu_Y) = \left[(\overline{X} - \overline{Y}) - z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}, (\overline{X} - \overline{Y}) + z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right]$$

es un IC con nivel de confianza $1-\alpha$ para $\mu_X - \mu_Y$.

Con ayuda de este IC para $\mu_X - \mu_Y$ se puede realizar un contraste de hipótesis sobre la igualdad de medias de las variables aleatorias X e Y: $\mu_X = \mu_Y$, cuando las varianzas sean conocidas. Este contraste es equivalente a $H_0 : \mu_X - \mu_Y = 0$, por lo que no existirán evidencias para rechazar H_0 con nivel $1-\alpha$ si $0 \in IC_{1-\alpha}(\mu_x - \mu_Y)$, y se rechazará dicha hipótesis en caso contrario.

Este IC requiere el conocimiento de σ_X y σ_Y. O bien, en un razonamiento análogo al de una variable aleatoria, intentar construir un IC que no requiera σ_X y σ_Y. Como veremos, eso sólo será factible asumiendo que $\sigma_X = \sigma_Y$:

IC para la diferencia de medias con varianzas desconocidas e iguales: Dado un nivel de confianza $1-\alpha$, se tiene que el intervalo:

$$IC_{1-\alpha}(\mu_X - \mu_Y) = \left[(\overline{X} - \overline{Y}) - t_{n+m-2,\alpha/2}s_p\sqrt{\frac{1}{n} + \frac{1}{m}}, (\overline{X} - \overline{Y}) + t_{n+m-2,\alpha/2}s_p\sqrt{\frac{1}{n} + \frac{1}{m}}\right]$$

es un intervalo de confianza con nivel de confianza $1-\alpha$ para $\mu_X - \mu_Y$, con

$$s_p = \sqrt{\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}$$

Con este IC para $\mu_X - \mu_Y$ se realizará el contraste de hipótesis $\mu_X = \mu_Y$, cuando las varianzas sean desconocidas pero iguales. Es inmediato obtener los intervalos de confianza de una cola para realizar el contraste de hipótesis $\mu_X \leq \mu_Y$.

7.6.2. IC para el cociente de varianzas

<u>IC para el cociente de varianzas:</u> Dado un nivel de confianza $1-\alpha$, se tiene que el intervalo:

$$IC_{1-\alpha}(\sigma_X^2/\sigma_Y^2) = \left[\frac{S_X^2}{S_Y^2 f_{n-1,m-1,\alpha/2}}, \frac{S_X^2}{S_Y^2 f_{n-1,m-1,1-\alpha/2}}\right]$$

es un intervalo de confianza con nivel de confianza $1-\alpha$ para σ_X^2/σ_Y^2

Con ayuda de este IC para $\frac{\sigma_X^2}{\sigma_Y^2}$ se puede realizar un contraste de hipótesis sobre la igualdad de varianzas de las variables aleatorias X e Y: $\sigma_X^2 = \sigma_Y^2$. Este contraste es equivalente a $H_0 : \frac{\sigma_X^2}{\sigma_Y^2} = 1$, por lo que, no existirán evidencias para rechazar H_0 con nivel $1-\alpha$ si $1 \in IC_{1-\alpha}(\sigma_X^2/\sigma_Y^2)$, y se rechazará dicha hipótesis en caso contrario.

7.7. Ejercicios Tema 7

Ejercicio 7.1. *Sea* $Z \sim N(0,1)$.

1. *Sabiendo que* $\Phi(z_\alpha) = P[Z \le z_\alpha] = 1-\alpha$, *demostrar que* $P[-z_{\alpha/2} \le Z \le z_{\alpha/2}] = 1-\alpha$.

2. *Hallar:*

   a) $\Phi(2.33)$

   b) $\Phi(-2.33)$

   c) $\Phi(1.96)$

   d) $\Phi(-1.96)$

   e) $\Phi(1.29)$

   f) $\Phi(-1.29)$

   g) $z_{0.01}$ *y* $z_{0.99}$

   h) $z_{0.025}$ *y* $z_{0.975}$

   i) $z_{0.1}$ *y* $z_{0.9}$

   j) $z_{\alpha/2}$ *y* $z_{1-\alpha/2}$ *si* $1-\alpha = 0.99$

   k) $z_{\alpha/2}$ *y* $z_{1-\alpha/2}$ *si* $1-\alpha = 0.975$

   l) $z_{\alpha/2}$ *y* $z_{1-\alpha/2}$ *si* $1-\alpha = 0.9$

Ejercicio 7.2. 1. *Sean* $X_1, \ldots, X_n$ *variables aleatorias iid* $N(\mu, \sigma^2)$, *demostrar que:*

   a) $\overline{X} \sim N(\mu, \sigma^2/n)$

   b) $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

2. *Sean $X_1, \ldots, X_n$ variables aleatorias iid de distribución desconocida tal que $E[X_i] = \mu$ y $V[X_i] = \sigma^2$. Si consideramos n suficientemente grande ¿qué distribución aproximadamente siguen las siguientes variables aleatorias y por qué?*

    a) $\overline{X}$

    b) $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$

3. *¿Qué diferencia hay entre los apartados anteriores?*

Ejercicio 7.3. *Demostrar:*

1. *Si $Z \sim N(0,1) \Rightarrow Z^2 \sim \chi_1^2$*

2. *Si $J_n \sim \chi_n^2$, $\varphi(t) = \left(\frac{1}{1-2t}\right)^{\frac{n}{2}}$*

3. *Si $J_n \sim \chi_n^2$, $E[J_n] = n$ y $V[J_n] = 2n$*

4. *Si $Z_1, Z_2 \sim N(0,1)$ independientes $\Rightarrow \frac{Z_1}{Z_2} \sim t_1$*

5. *Sea $T \sim t_n$, $n > 1$. Si sabemos que la función de densidad de T es simétrica respecto del origen, ¿cuánto vale $E[T]$?. Si $E[T^2] = \frac{n}{n-2}$, ¿cuánto vale $V[T]$?*

6. *Si $F \sim \mathcal{F}_{n,m} \Rightarrow \frac{1}{F} \sim \mathcal{F}_{m,n}$.*

7. *Si $F \sim \mathcal{F}_{1,n} \Rightarrow F = T^2$, con $T \sim t_n$*

Ejercicio 7.4. *Usando la tabla de la distribución χ^2, hallar:*

1. $j_{3,0.9}$

2. $j_{26,0.1}$

3. $j_{12,0.02}$

4. $j_{15,\alpha/2}$ *y* $j_{15,1-\alpha/2}$ *si* $1-\alpha = 0.99$

5. $j_{4,\alpha/2}$ *y* $j_{4,\alpha/2}$ *si* $1-\alpha = 0.95$

6. $j_{30,\alpha/2}$ *y* $j_{30,\alpha/2}$ *si* $1-\alpha = 0.9$

Ejercicio 7.5. *Usando la tabla de la distribución t_n, hallar:*

1. $t_{3,0.125}$

2. $t_{3,0.875}$

3. $t_{12,0.03}$

4. $t_{12,0.97}$

5. $t_{15,\alpha/2}$ *y* $t_{15,1-\alpha/2}$ *si* $1-\alpha = 0.99$

6. $t_{4,\alpha/2}$ *y* $t_{4,1-\alpha/2}$ *si* $1-\alpha = 0.95$

7. $t_{30,\alpha/2}$ *y* $t_{30,1-\alpha/2}$ *si* $1-\alpha = 0.9$

Ejercicio 7.6. *Usando la tabla de la distribución* $\mathcal{F}_{n,m}$*, hallar:*

1. $f_{3,9,0.1}$

2. $f_{3,9,0.9}$

3. $f_{12,15,0.05}$

4. $f_{12,15,0.95}$

5. $f_{10,10,0.025}$

6. $f_{10,10,0.975}$

7. $f_{15,20,\alpha/2}$ *y* $f_{15,20,1-\alpha/2}$ *si* $1-\alpha = 0.8$

8. $f_{4,11,\alpha/2}$ *y* $f_{4,11,1-\alpha/2}$ *si* $1-\alpha = 0.95$

9. $f_{30,100,\alpha/2}$ *y* $f_{30,100,1-\alpha/2}$ *si* $1-\alpha = 0.98$

Ejercicio 7.7.

1. *Sea* $X \sim N(\mu, \sigma^2)$*. Sea* $X_1, \ldots, X_n$ *una mas de* X*, y* $Y = \frac{X-b}{a}$*, con* $a > 1$*,* $b > 0$*. Utilizando estos datos, resolver los siguientes apartados.*

   a) *Suponiendo* μ *y* σ *desconocidas, hallar un intervalo de confianza de* $\mu_Y = E[Y]$ *y otro para* $\sigma_Y^2 = V[Y]$ *con nivel de confianza* $1-\alpha$*.*

   b) *Hallar* $P[Y \leq y]$ *en función de* $\Phi(z)$ *función de distribución de* $Z \sim N(0,1)$*.*

   c) *Para estimar la media de* Y *se toman los estadísticos* $\hat{\mu}_Y^1 = \bar{Y}$ *y* $\hat{\mu}_Y^2 = a\bar{Y}$*. ¿Cuál es mejor estimador?*

   d) *Hallar la función generatriz de momentos de* Y*. Justificar qué distribución sigue* Y*.*

2. *Sean* $X \sim Be(p)$ *y* $Z \sim Be(p)$*, con* Z *y* X *independientes. Sea* $Y = \frac{1}{X+1}$ *y* $S = X + Z$*.*

   a) *Hallar la función de distribución de* Y*. ¿Es* Y *continua o discreta?. Hallar* $Cov[X,Y]$ *y* $Cov[Z,Y]$*.*

   b) *Hallar la función de probabilidad y la función de distribución de* S*.*

   c) *Hallar* $P\left[S = 1 | Y = \frac{1}{2}\right]$ *y* $P[Y = 1 | S = 2]$*.*

Ejercicio 7.8. *Sea* $X_n \sim Erlang\left(\frac{n}{2}, \frac{1}{2}\right)$*, n par.*

1. *¿Cuál es la función generatriz de momentos de* X_n*? ¿Cuáles son su esperanza y su varianza?*

Sean $Z_1, \ldots, Z_n$ variables aleatorias independientes e idénticamente distribuidas $N(0,1)$.

2. *¿Qué distribución sigue la variable aleatoria $J_n = Z_1^2 + \cdots + Z_n^2$?*
3. *Demostrar que la función generatriz de momentos de J_n es*

$$\varphi(t) = \left(\frac{1}{1-2t}\right)^{n/2}$$

   Nota: usar el cambio de variable $x^2 = (1-2t)z^2$

4. *A partir de la función generatriz de momentos, hallar la esperanza y la varianza de J_n.*
5. *¿Cuál es la relación entre X_n y J_n? Justifíquese.*
6. *Para $n = 5$, usando las tablas dadas para el examen, hallar $P[X_5 > 8.62]$*
7. *Para $n = 2$, ¿qué distribución sigue X_2? Sin usar las tablas, hallar $P[X_2 \leq 2]$*

Ejercicio 7.9. *** Sea X la cantidad de agua de lluvia caída en un mes, en cm^3 por metro cuadrado. Se ha medido X en los últimos nueve meses, obteniéndose que $\bar{x} = 407.56$. Del histórico de la estación meteorológica se tiene que la desviación típica de X es $\sigma = 17$.*

1. *Hallar el intervalo de confianza aproximado con nivel de confianza del 95 % para la media.*
2. *Si suponemos normalidad de la variable aleatoria X, con varianza desconocida, hallar el intervalo de confianza para σ^2 con nivel de confianza del 80 % si $s^2 = 294.78$ es la varianza muestral obtenida en la muestra de los nueve meses.*
3. *Suponiendo las condiciones del apartado anterior, hallar el intervalo de confianza con nivel de confianza del 95 % para μ.*

Resultado

1. $IC_{0.95}(\mu) = [396.4533, 418.6667]$
2. $IC_{0.8}(\sigma^2) = [176.49, 675.81]$
3. $IC_{0.95}(\mu) = [394.362, 420.758]$

Ejercicio 7.10. *** Para medir la contaminación del agua, se han tomado 16 medidas independientes de la cantidad de mercurio en microgramos por litro (X), obteniéndose que $\bar{x} = 51.3$.*

1. *Suponiendo que X tiene una desviación típica $\sigma = 2.5$, contrastar para $\alpha = 0.05$ si la media es igual a 50 microgramos por litro.*

2. *Contrastar para* $\alpha = 0.05$ *si la media es igual a 50 microgramos por litro suponiendo normalidad y que* $s = 2.5$.

3. *Razonar los resultados de los apartados anteriores.*

Resultado

1. $IC_{0.95}(\mu) = [50.075, 52.525]$
2. $IC_{0.95}(\mu) = [49.968, 52.632]$
3. La suposición de normalidad influye notablemente en las conclusiones

Ejercicio 7.11. *** En dos tiendas de una misma cadena se está estudiando el uso de una determinada tarjeta de crédito por parte de los clientes. La desviación típica del número de veces que se usa dicha tarjeta en la primera tienda es 21.7945 y 18.7083 en la segunda. Se han tomado el número de veces de uso en cada tienda durante 11 días resultado:*

Tienda 1	*99*	*94*	*84*	*125*	*129*	*109*	*60*	*75*	*90*	*113*	*132*
Tienda 2	*120*	*121*	*96*	*120*	*139*	*126*	*75*	*96*	*106*	*136*	*137*

Si se supone independencia y normalidad:

1. *¿Se puede suponer con una probabilidad de error del 10 % que las medias del número de uso de la tarjeta en ambas tiendas es la misma?*

2. *Realizar una estimación confidencial al* 90 % *sobre el cociente de varianzas del número de uso de la tarjeta.*

3. *¿Se puede suponer que las varianzas son iguales? Chequear en caso afirmativo si las medias del número de usos de la tarjeta en las dos tiendas es la misma con el mismo nivel de confianza anterior.*

Resultado

1. $IC_{0.9}(\mu_X - \mu_Y) = [-28.977, -0.485]$
2. $IC_{1-\alpha}(\sigma_X^2/\sigma_Y^2) = [0.45, 3.989]$
3. $IC_{1-\alpha}(\mu_X - \mu_Y) = [-30.27, 0.8154]$

Ejercicio 7.12. ****

La autonomía de cierto tipo de linterna sigue una distribución normal con desviación típica de 1.5 horas. El intervalo de confianza al $95\,\%$ *para la media es* $[28.85, 30.81]$. *Otro intervalo de confianza para la media, de nivel* $1-\alpha$, *resulta ser* $[29.01, 30.65]$. *Si ambos han sido estimados con la misma muestra, hallar:*

1. *El tamaño* n *de la muestra utilizada.*

2. *El nivel de confianza* α.

Resultado

1. $n = 9$

2. $\alpha = 0.1$

Ejercicio 7.13. *** La energía solar consumida entre los años 1989 y 2004 (en trillones de unidades del sistema inglés de medida), se muestra en la siguiente tabla:*

55.291	*66.458*	*70.237*	*65.454*
59.718	*68.548*	*69.787*	*64.391*
62.688	*69.857*	*68.793*	*63.62*
63.886	*70.833*	*66.388*	*63.287*

Suponiendo normalidad:

1. *Construir un intervalo de confianza del* $95\,\%$ *de la media de energía solar consumida si se tiene que la varianza es de 25.*

2. *Construir un intervalo de confianza del* $90\,\%$ *de la varianza de energía solar consumida.*

3. *Si se toma otra muestra* $(y_1, \ldots, y_{10})$, *obteniéndose una media muestral* $\bar{y} = 62.354$ *y varianza muestral* $s_Y^2 = 20.368$, *hallar el intervalo de confianza para la diferencia de medias con un nivel de confianza del* $90\,\%$.

4. *Teniendo en cuenta el resultado del apartado anterior, ¿se puede suponer que las medias de las dos muestras son iguales? Razonar la respuesta describiendo el contraste de hipótesis que se ha llevado a cabo para tomar la decisión.*

Resultado

1. $IC_{0.95}(\mu) = [63.12725, 68.02725]$
2. $IC_{0.90}(\sigma^2) = [10.715, 36.886]$
3. $IC_{0.9}(\sigma_X^2/\sigma_Y^2) = [0.2916, 2.2686]$ $IC_{0.9}(\mu_X - \mu_Y) = [0.2331, 6.2134]$
4. Ver teoría.

Ejercicio 7.14. **** El contenido medio en sodio que se indica en una caja de cereales de 300 gramos es de 130 miligramos. Se realiza un análisis en una muestra y se toman los siguientes datos:*

131.15	130.14	131.15	130.69	129.29	130.69
130.91	128.71	130.91	129.54	129.00	129.54
129.64	129.39	129.64	128.77	130.42	128.77
130.72	129.53	130.72	128.33	130.12	128.33
128.24	129.78	128.24	129.65	130.92	129.65

Suponiendo normalidad e independencia:

1. *Se desea comprobar si la indicación de las cajas es correcta con una probabilidad de error de tomar la decisión incorrecta del* 15 %.
2. *¿Y si se desea comprobar lo mismo que en el apartado anterior con una probabilidad de error de tomar la decisión incorrecta del* 5 %*?*
3. *Para la mitad de los datos, tomando las tres primeras columnas, hallar el intervalo de confianza con un nivel de confianza del* 95 %.
4. *Comparar las amplitudes de los intervalos hallados en los tres apartados anteriores. ¿Cuál es el más ajustado? ¿Qué relación se observa entre la amplitud de los intervalos, el tamaño muestral y el nivel de confianza de los intervalos?*
5. *¿Existen evidencias para no afirmar que la varianza es 1? Estudiarlo con un nivel de confianza del* 80 %.

Resultado

1. $IC_{0.85}(\mu) = [129.3979, 130.1074]$
2. $IC_{0.95}(\mu) = [129.238, 130.267]$
3. $IC_{0.95}(\mu) = [129.379, 130.479]$
4. Cuanto mayor es el nivel de confianza el intervalo es mayor. Además, cuanto mayor es el tamaño de la muestra mayor es la fiabilidad del intervalo y es más pequeño.

5. $IC_{0.8}(\sigma^2) = [0.641, 1.267]$

Resultado

1. $IC_{0.9}(\sigma_X^2) = [0.002053, 0.01985]$
2. $IC_{0.9}(\sigma_X^2/\sigma_Y^2) = [0.6652, 32.4340]$ $IC_{0.9}(\mu_X - \mu_Y) = [-0.0667, 0.0677]$

Ejercicio 7.15. **** Dos vigilantes están calculando el tiempo de la ronda por una ruta predeterminada del edificio que vigilan. Los datos tomados en minutos por cada uno de ellos son los siguientes:*

Vigilante Juan	*25.06*	*25.14*	*25.10*	*25.03*	*25.20*	*25.03*
Vigilante Miguel	*25.08*	*25.06*	*25.11*	*25.12*		

1. *Calcular el intervalo de confianza para la varianza de los tiempos del vigilante Juan con un nivel de confianza del* 90 %.
2. *Considerando* $\alpha = 0.1$ *, decir si existen diferencias significativas entre los tiempos de ronda de cada vigilante.*

Ejercicio 7.16. ***** Responda a las siguientes cuestiones:*

1. *Sea* $X \sim N(\mu, \sigma^2)$ *con* $\mu = 5$ *y* $\sigma^2 = 4$*. Hallar* $F_Y(1.5)$ *si* $Y = |X|$.
2. *Si* $X \sim \chi_n^2$*, determinar* $E[X]$.
3. *Se tiene una población (en la que se asume que la varianza es conocida) de la que se extrae una m.a.s. de tamaño* n *y se determina un intervalo de confianza para la media de nivel* $1 - \alpha$*. Si se extrae otra m.a.s. de tamaño* m *y se obtiene que la longitud del nuevo intervalo de confianza (de nivel* $1 - \alpha$*) para la media es el doble que la del primero, ¿cuál es la relación entre* n *y* m*?*

Resultado

1. $F_Y(1.5) = 0.03948$
2. $E[X] = n$
3. $4m = n$

Ejercicio 7.17. ***** Se tiene una variable aleatoria que sigue una distribución normal de parámetros* μ *y* σ^2*. Asumiendo que tanto* μ *como* σ^2 *son desconocidos, se ha extraído una mas de tamaño* $n = 16$*, de la que se sabe que con un nivel de confianza del 90 % (en ambos casos)* σ^2 *se encuentra en el intervalo* $[3.8, A]$ *y* μ *se encuentra en el intervalo* $[1.5, B]$*. Se pide:*

1. *Determinar el valor de* A*.*
2. *Determinar el valor de* B*.*

Resultado

1. $A = 13.0816$
2. $B = 3.7058$

Ejercicio 7.18. ***** Se realizan 10 mediciones en dos instrumentos de medida A y B con el objeto de averiguar cuál de ellos realiza las medidas con mayor exactitud. Para ello se utiliza un patrón de 1 centímetro, obteniéndose las siguientes medias y desviaciones típicas muestrales en cada instrumento:* $\bar{a} = 1.001$ *cm.;* $s_a = 0.01$ *cm.* $\bar{b} = 0.997$ *cm.;* $s_b = 0.02$ *cm. Determinar los intervalos de confianza al 90 % para:*

1. *La media del instrumento A,* μ_a
2. *La media del instrumento B,* μ_b
3. *La varianza del instrumento A,* σ_a^2
4. *La varianza del instrumento B,* σ_b^2
5. *El cociente de varianzas de ambos instrumentos,* $\frac{\sigma_a^2}{\sigma_b^2}$

Resultado

1. $I_{0.9}(\mu_a) = [0.995203, 1.006797]$
2. $I_{0.9}(\mu_b) = [0.9854, 1.0086]$
3. $I_{0.9}(\sigma_a^2) = [0.00005319, 0.0002707]$
4. $I_{0.9}(\sigma_b^2) = [0.000213, 0.010827]$
5. $I_{0.9}\left(\frac{\sigma_a^2}{\sigma_b^2}\right) = [0.07864, 0.79475]$

Ejercicio 7.19. ****** Sea X variable aleatoria con la función de densidad:*

$$f(x) = ae^{bx}, x \in [0, 1], a, b > 0$$

La función generatriz de momentos de X es

$$\varphi(t) = \frac{e^{(t+1)} - 1}{(e-1)(t+1)}$$

Se extrae una m.a.s de tamaño 30 de X, $(X_1, X_2, ..., X_{30})$. Se pide:

1. *Calcular la varianza de la variable $Y = X_{12} - X_{15} + 3X_{20}$.*
2. *Calcular la probabilidad de que la media muestral de la m.a.s. sea menor que 0.6.*
3. *Determinar los parámetros a y b de la función de densidad dada.*
4. *Probabilidad de que X_4 sea menor que 0.5 sabiendo que X_2 ha tomado el valor 0.8.*
5. *Probabilidad de que los únicos resultados muestrales de la m.a.s. mayores que 0.5 sean los cuatro primeros.*
6. *Probabilidad de que cuatro resultados muestrales de la m.a.s. sean mayores que 0.5.*
7. *Si todos los días se confeccionara un intervalo de confianza de forma independiente para estimar el parámetro a de nivel de confianza $(1 - \alpha)$, calcular dicho nivel de confianza para que el número medio de días consecutivos en los que el intervalo contenga al parámetro sea 49.*

Resultado

1. 0.8726
2. 0.63683
3. $a = \frac{1}{e-1}$, $b = 1$
4. 0.3775
5. $1.5 \cdot 10^{-12}$
6. $4.11 \cdot 10^{-8}$
7. 0.98

7.8. Problemas Tema 7

Problema 7.1. ***** Se sabe que la altura de los alumnos de la Universidad de Sevilla y de la Universidad de Málaga se distribuyen según sendas normales $X \sim N(\mu_X, \sigma_X^2)$ e $Y \sim N(\mu_Y, \sigma_Y^2)$, respectivamente.*

1. *Si se eligen al azar 9 alumnos de la Universidad de Sevilla, determinar las siguientes cuestiones:*

   a) *En el caso en el que conozca la desviación típica de la población, $\sigma_X = 6$ cm, razonar cuál es el intervalo de 6 cm de longitud que con mayor probabilidad contiene a la media . ¿Cuál es dicha probabilidad máxima?*

   b) *En el caso en el que se desconozca la desviación típica, determinar el valor esperado de la longitud del intervalo de confianza del 90 % para la varianza, en función de σ_X^2.*

2. *Suponga ahora que se eligen en esta ocasión 6 alumnos al azar de la Universidad de Sevilla diferentes del apartado anterior, resultando las siguientes alturas en cm:*

   $$(170.2; 160.4; 165.3; 157.8; 178.6; 170.3)$$

   Además se escogen 5 alumnos al azar de la Universidad de Málaga cuyas alturas en cm. son:

   $$(180.1; 162.3; 157.1; 163.5; 168.9)$$

   Considerando las varianzas σ_X^2 y σ_Y^2 desconocidas, calcule un intervalo de confianza del 90 % para el cociente entre varianzas de ambas variables aleatorias.

3. *¿Se puede suponer, con un nivel de confianza del 90 %, que las medias de las alturas son iguales?*

Resultado

1. $n = 9$

   a) $1 - \alpha = 0.8664$

   b) $E[L] = 2.41173\sigma_X^2$

2. $IC_{0.90}(\sigma_X^2/\sigma_Y^2) = [0.11996, 3.89638]$

3. $IC_{0.9}(\mu_X - \mu_Y) = [-8.2852, 9.725]$

Problema 7.2. ****** Se pretende analizar la idoneidad de una pista de aterrizaje para un determinado tipo de avión. Se asume que la distancia recorrida por dicho avión durante el aterrizaje sigue una distribución* *normal de parámetros μ y σ^2. Se han realizado una serie de experimentos midiendo las distancias recorridas por el avión, obtenido los siguientes valores (en km):*

$$[2.00, 2.04, 2.00, 1.97, 1.98, 2.01, 1.99, 2.00, 2.02, 1.99]$$

Se pide:

1. *Calcular los intervalos de confianza de los parámetros μ y σ^2 con un nivel de confianza del 95%.*
2. *Se pretende que la distancia recorrida por el 95 % de los aviones no sobrepase el valor r. Razone cuál sería dicho valor r a partir del resultado anterior, en el caso más desfavorable posible.*
3. *Si se añaden dos nuevas mediciones de la distancia recorrida con los valores $[2.00, 1.98]$, contrastar (al 95 %) las estimaciones de los parámetros μ y σ^2 realizadas anteriormente.*

Resultado

1. $IC_{0.95}(\mu) = [1.985, 2.014]$ $IC_{0.95}(\sigma^2) = [1.893 \cdot 10^{-4}, 1.333 \cdot 10^{-3}]$
2. $r = 2.074$
3. $IC_{0.95}(\mu) = [1.9863, 2.0104]$ $IC_{0.95}(\sigma^2) = [1.81 \cdot 10^{-4}, 10.396 \cdot 10^{-4}]$

Problema 7.3. ****** Sea X una variable aleatoria que representa el número de piezas tipo A elaboradas por día por una máquina M_1. La variable aleatoria tiene la siguiente función de probabilidad:*

$$p_k = P[X = x_k] = 2^{x_k} b^{x_k} \cdot \frac{1}{e^{2b}} \cdot \frac{1}{x_k!}, \quad x_k = 0, 1, \ldots$$

Sea $\hat{b}_1 = \frac{\overline{X}}{2}$ un estimador del parámetro b. Se pide:

1. *Para una mas $(X_1, \ldots, X_4)$, se desea comparar $\hat{b}_1$ con el estimador $\hat{b}_2 = \frac{X_1+2\cdot X_2+3\cdot X_3+4\cdot X_4}{20}$. ¿Cuál de los dos estimadores elegirías como estimación del parámetro? Nota: La función generatriz de momentos de X es*

$$\varphi(t) = \frac{1}{e^{2b}} e^{2be^t}$$

2. *Obtener las estimaciones del parámetro b utilizando $\hat{b}_1$ y $\hat{b}_2$ para la siguiente realización muestral de X:*

$$[230, 240, 235, 234]$$

Las piezas tipo A deben ser ensambladas con unas piezas tipo B, que se producen en la máquina M_2, produciendo Y piezas cada día. Se desea que las tasas de producción de ambas maquinas sea similar. Para comprobarlo, se toma la siguiente muestra de Y:

$$[226, 238, 235, 232, 233, 230, 240, 233]$$

3. *Suponiendo que las distribuciones de X e Y se pueden aproximar a normales con varianzas iguales, calcular el nivel de confianza para no rechazar que las medias de ambas variables difieren menos de 5.*

Resultado

1. $ECM(\hat{b}_1) = \frac{b}{8} = 0.125 \cdot b$ $ECM(\hat{b}_2) = 0.15 \cdot b$
2. $\hat{b}_1 = 117.375$ $\hat{b}_2 = 117.55$
3. $1 - \alpha \geq 0.9$

Problema 7.4. ****** En la Escuela cada día hay programadas cuatro clases de Estadística de idéntica duración. Durante diez días, se ha registrado la duración total de las clases de cada día:*

Día	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Duración (min)	*208*	*209*	*211*	*210*	*208*	*212*	*218*	*213*	*217*	*214*

Se asume que la duración total de las clases en el día sigue una distribución Erlang.

1. *A la vista de los datos, estimar el valor esperado y desviación típica de la duración total de las clases en un día.*
2. *Hallar la probabilidad de que, en un día cualquiera, una clase dure más de 55 minutos.*
3. *En un cuatrimestre hay 14 semanas con 5 días de clase cada una. Hallar la probabilidad de que, en un cuatrimestre, se hayan perdido más de 4 horas de clase si la duración oficial de éstas es de 55 min.*
4. *¿Es posible asegurar con un 99 % de probabilidad que el número esperado de minutos perdidos en un día será inferior a 3?*

Resultado

1. $\hat{\mu} = 212$, $\hat{\sigma} = 3.5276$
2. 0.354
3. 0.64
4. $IC_{0.99}(\mu)^s = [208.8526, +\infty]$, y se rechaza la hipótesis al estar 217 en el intervalo.

Problema 7.5. ****** Se sabe que el número medio de clientes por minuto atendidos en una caja cualquiera de un determinado hipermercado es 4 si ese día hay promociones especiales, y 2 si ese día no hay. Sabiendo que hay promociones especiales el 10 % de los días que se abre en un año, se pide:*

1. *La probabilidad de que un día sin promociones especiales se atiendan al menos 3 clientes por minuto en una caja.*
2. *La probabilidad de que en un día cualquiera se atiendan al menos 3 clientes por minuto en una caja.*
3. *La probabilidad de que en un día cualquiera se atiendan exactamente n clientes por minuto en una caja.*
4. *Valor esperado y varianza del número de clientes atendidos por minuto en una caja en un día cualquiera.*
5. *Probabilidad de que el número total de clientes atendidos por minuto en las 40 cajas del hipermercado en un día cualquiera sea menor que 100.*
6. *Se ha observado para distintos días el valor del número total de clientes atendidos por minuto en las 40 cajas del hipermercado, encontrando los siguientes valores:*

$$[120, 142, 67, 98, 112, 79, 131, 161]$$

¿Con qué nivel de confianza se puede afirmar que el valor esperado del apartado anterior es aceptable?

Resultado

1. 0.3233
2. 0.3671
3. $e^{-4}\frac{4^n}{n!}0.1 + e^{-2}\frac{2^n}{n!}0.9$
4. $E[N] = 2.2$, $V[N] = 2.6$
5. 0.881
6. 94 %

Problema 7.6. *La irradiación solar diaria que recibe cada metro cuadrado de una determinada zona se puede modelar mediante una distribución exponencial de media M Kilowatios-hora (Kwh). Si el día está despejado, $M = 4$, mientras que en caso contrario $M = 3.5$. En dicha zona, la probabilidad de que un día esté despejado es del 60 %. Se pide:*

1. *Probabilidad de que la irradiación solar que recibe un metro cuadrado de la zona en un día despejado sea inferior a 3.5 Kwh.*

2. *Probabilidad de que la irradiación solar que recibe un metro cuadrado de la zona en un día cualquiera sea inferior a 3.5 Kwh.*

3. *Modelar la variable aleatoria I irradiación solar diaria que recibe cada metro cuadrado de la zona. Determinar el valor esperado de I.*

4. *Para saber si la irradiación ha aumentado en los últimos años, los expertos quieren comparar una muestra de datos tomados en la actualidad con los tomados en 2007. Ante los datos dados en la siguiente tabla, justificar si dicho aumento ha ocurrido o no con un nivel de confianza del 95 %.*

2007	0.3	3.1	3.3	0.9	1.3	3.0	0.5	2.7	3.5	2.2	2.4	1.2	0.9	0.9		
Este año	0.2	1.0	1.1	1.2	2.1	0.5	2.9	1.3	4.5	3.4	0.6	3.4	3.7	1.6	0.1	3.5

5. *Asumiendo los datos de la actualidad, determinar la probabilidad de que, en un año (365 días), la irradiación solar que recibe un metro cuadrado de la zona sea superior a 755 Kwh.*

Resultado

1. 0.5831
2. 0.6027
3. 3.8
4. $IC_{0.95}(\sigma_X^2/\sigma_Y^2) = [0.2153, 1.9231]$, $IC_{0.95}(\mu_X - \mu_Y)^s = [-\infty, 0.7320]$
5. 0.04648

Problema 7.7. *El propietario de una pizzería A está interesado en analizar la competencia que le hace la pizzería B del barrio. El propietario se plantea, en primer lugar, determinar cuál de las dos pizzerías ofrece una mejor atención a los clientes. Para ello toma la siguiente muestra de 15 personas:*

$$[0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 1, 0]$$

dónde 1 representa que el cliente está más satisfecho con la pizzería A y 0 lo contrario.

1. *Dar una estimación del parámetro p de la distribución Bernoulli.*

2. *Hallar la probabilidad de que en una selección de 1000 clientes de forma independiente, más de la mitad estén satisfechos con la pizzería A.*

A continuación, quiere analizar los tiempos de preparación de sus pizzas en comparación con los de B. Para ello, compra 10 pizzas de B por internet, tomando una muestra aleatoria simple del tiempo de entrega (preparación más envío):

$$[30, 22, 26, 33, 42, 40, 38, 36, 37, 32]$$

Se sabe que el tiempo de envío de las pizzas de ambas pizzerías sigue una distribución $N(10, 16)$. Suponiendo que el tiempo de preparación en la pizzería B sigue una distribución $N(\mu, \sigma^2)$, se pide:

3. *Calcular los intervalos de confianza de μ y σ^2 con un grado de confianza del 95 %.*
4. *Calcular de forma aproximada la probabilidad de que el tiempo de preparación en la pizzería B sea mayor que el de la pizzería A, estimando el tiempo de preparación de la pizzería B, y suponiendo que para A sigue una distribución $N(20, 30)$.*

Resultado

1. $\hat{p} = \frac{8}{15}$
2. 0.98257
3. $IC_{0.95}(\mu) = [19.098, 28.102]$, $IC_{0.95}(\sigma^2) = [2.735, 115.981]$
4. 0.68793

Problema 7.8. *En la ETSI hay 4000 alumnos, cada uno de ellos con una probabilidad p de comer en la cafetería. Tradicionalmente el pago de las comidas se realizaba directamente a uno de los camareros de la caja. Recientemente cambió el sistema de pagos (introduciendo máquinas expendedoras de tickets) con el objetivo de reducir los tiempos de espera en la cola de la comida. Pasados ya cuatro meses de los cambios realizados, se tiene la impresión de que ha aumentado el número esperado de alumnos que comen en la cafetería. Para contrastar este hecho se realiza una muestra para 15 días del número de alumnos que comen en la cafetería dando los siguientes resultados:*

$$[240, 260, 256, 244, 242, 251, 250, 257, 255, 279, 255, 237, 249, 255, 266]$$

Adicionalmente, se disponía antiguamente de la siguiente muestra del número de alumnos antes de realizar los cambios:

$$[230, 242, 238, 264, 242, 251, 239, 228, 255, 243]$$

Se pide:

1. *Analizar si se ha producido un incremento en el número esperado de alumnos que come en la cafetería con un grado de confianza del 95 %. Justificar, en caso necesario, las hipótesis que tenga que realizar para tomar la decisión.*

2. *A partir de la muestra tomada después de los cambios en el sistema de pagos, calcule de manera aproximada la probabilidad de que el número de personas que comen en un día sea mayor de 200.*

Resultado

1. $IC_{0.95}\left(\sigma_x^2/\sigma_y^2\right) = [0.247, 3.016]$ $IC_{0.95}\left(\mu_X - \mu_Y\right) = [0.763, 18.97]$
2. 0.99972

Problema 7.9. *******

El número de billetes de AVE que se venden al día por Internet se distribuye como una Poisson cuya media es de 270 billetes/día.

1. *Sabiendo que la probabilidad de que un pasajero con billete pierdan su tren es del 1 %, se pide:*
   a) *Si un día se venden N billetes, calcular, en función de N, la probabilidad de que 2 pasajeros de los que compraron dichos billetes pierdan sus trenes.*
   b) *Probabilidad de que en un día se hayan vendido N billetes y un pasajero de los que compraron dichos billetes pierda su tren, en función de N.*
   c) *Calcular la probabilidad de que, en un día, ninguno de los pasajeros pierda su tren. Nota:*
   $$\sum_{i=0}^{\infty} \frac{x^i}{i!} = e^x$$
2. *Calcular la probabilidad de que entre las 8:00 y las 16:00 horas, el número de billetes vendidos por Internet sea menor que b. Además, teniendo en cuenta la reproductividad de la Poisson, calcular usando el TCL dicha probabilidad para $b = 100$.*
3. *Para la siguiente m.a.s. correspondiente a los billetes comprados al día por Internet durante una determinada semana $(268, 270, 272, 269, 261, 270, 264)$, calcular un intervalo de confianza de nivel de confianza 0.98 para la media de billetes comprados al día por Internet, demostrando que se puede asumir normalidad.*
4. *Sabiendo que el número medio de billetes de clase turista que se vende consecutivamente es de 2.8, calcular el valor esperado y la varianza del número de viajeros de clase turista si se han vendido 200 billetes.*

Resultado

1. a) $\binom{N}{2}0.01^2(1-0.01)^{N-2}$
   b) $0.01(0.99)^{N-1}\frac{(270)^N}{(N-1)!}e^{-270}$
   c) 0.0672
2. 0.85304
3. $[263.12, 272.3001]$
4. 147.368 y 38.77

Problema 7.10. ****** La demanda diaria de leche fresca (miles de litros) en un supermercado se distribuye según una normal de esperanza 100 y varianza 4. La leche fresca no vendida al final del día debe tirarse, por lo que cada día debe reponerse leche fresca nueva. Se pide:*

1. *Cantidad mínima de leche fresca a reponer diariamente para que no falte leche en al menos el 99 % de los días.*
2. *Número de días (en % sobre días del año) en los que sobrará leche si se reponen 101 miles de litros diarios.*
3. *Si se reponen 102 miles de litros diarios, probabilidad de que falte leche en más de un día de la semana (6 días laborables).*
4. *Si se reponen 101 miles de litros diarios, probabilidad de que al menos sobren 1000 litros en un día en el que ha sobrado leche.*
5. *Si se reponen 101 miles de litros diarios, probabilidad de que al menos sobren 1000 litros en un día cualquiera.*
6. *Se ha registrado la demanda de leche los días anteriores a un festivo, obteniendo los siguientes valores:* $(102, 95, 104, 98, 95, 102, 100, 103)$. *Con una confianza del 99 %, ¿podría afirmarse que los parámetros de la distribución normal de la demanda son distintos para estos días de los que se han considerado en el enunciado?*

Resultado

1. 104.66
2. 69.146 % de los días del año.
3. 0.2441
4. 0.7231
5. 0.5

6. No existen evidencias para aceptar que ninguno de los parámetros de la demanda son distintos para esos días.

Problema 7.11. ****** En un centro médico que abre 10 horas al día hay cuatro consultas de distintas especialidades. Se asume que el número medio de pacientes por unidad de tiempo que llegan a cada consulta es constante y el mismo para cada consulta. Sabiendo que la probabilidad de que no llegue ningún paciente en una hora al centro médico es de 0.005, se pide:*

1. *Probabilidad de que en una hora hayan llegado 4 pacientes al centro médico (a cualquiera de las especialidades).*
2. *Probabilidad de que el número de pacientes que han llegado a una de las consultas en un día sea de 12.*
3. *Probabilidad de que en una consulta pasen 2 horas sin que llegue ningún paciente.*
4. *Probabilidad de que en al menos dos de las consultas hayan llegado 10 pacientes en un día.*
5. *Se intentan validar los datos del enunciado mediante una muestra de 6 días en la que se ha medido el número de pacientes que llegan al centro médico, obteniéndose los siguientes valores:* $(56, 53, 56, 64, 51, 59)$. *De acuerdo a los datos de la muestra, ¿Cuál es el nivel de confianza más alto con el que se puede asumir el número medio de pacientes a la hora que se deduce del enunciado? Justificar que se puede asumir normalidad para aplicar el intervalo de confianza correspondiente.*

Resultado

1. 0.164
2. 0.1081
3. 0.071
4. 0.0353
5. 89 %.

Problema 7.12. ****** Se asume que el error de los radares (medido en ciertas unidades) se distribuye como una normal* $N(\mu, 40)$, *donde* μ *depende de la antigüedad del radar. Para radares instalados en 2005 o antes se puede asumir* $\mu = 30$, *mientras que para averiguar el valor de* μ *para los instalados con posterioridad se ha tomado la siguiente m.a.s.: (35.2, 23.6, 39.2, 19.4,27.4,12.4)*

Se pide:

1. *Estimación del error medio de los radares instalados con posterioridad a 2005 y su intervalo de confianza al 95 %.*

2. *Nivel de confianza mínimo con el que puede asumirse que* $\sigma^2 \geq 40$ *para estos radares.*

Asumiendo la estimación de la esperanza dada por la muestra anterior:

3. *Si se sabe que en una determinada zona hay una probabilidad del 50 % de que un radar cualquiera tenga un error menor que 28, determinar el porcentaje de radares de cada tipo (de 2005 y anteriores, o posteriores a 2005) que hay instalados en dicha zona.*

4. *Si un radar tiene un error menor que 25, determinar la probabilidad de que haya sido instalado en la zona con posterioridad a 2005.*

Resultado

1. $IC_{0.95}(\mu) = [21.1393, 31.2607]$
2. 97.5 %.
3. 53.24 % de radares de tipo moderno, 46.76 % de radares antiguos.
4. 0.6924

Capítulo 8

Problemas adicionales

Problema 8.1. *** Entran 3 personas en un edificio de 5 plantas y se montan en el ascensor. Cada una va saliendo del ascensor en las distintas plantas.*

1. *Hallar la probabilidad de que todas las personas salgan del ascensor en la misma planta.*
2. *Supongamos que todas las personas están en el ascensor, y van a pulsar una a una la planta a la que desean ir. Hallar la probabilidad de que todas las personas elijan plantas diferentes.*

Resultado

1. 0.04
2. 0.48

Problema 8.2. *Sea* $(X_1, \ldots, X_n)$ *una mas de* X*, cuya función de densidad es:*

$$f(x) = \lambda e^{-\lambda x}, x > 0$$

Se desea estimar $\theta = \frac{1}{\lambda}$

1. *Hallar el sesgo, la varianza y el error cuadrático medio de* $\hat{\theta}_1 = \overline{X}$
2. *Idem para* $\hat{\theta}_2 = \text{mín}\{X_i\}$
3. *Idem para* $\hat{\theta}_3 = \sum_{i=1}^{n} X_i$
4. *¿Cuál de los tres es mejor estimador de* θ*?*

Problema 8.3. *** Sea X una variable aleatoria que sigue una distribución de Poisson de parámetro λ. Sea (X_1, X_2, X_3) una mas. Determinar cuál es el mejor de los siguientes estimadores: $\hat{\lambda}_1 = \frac{X_1+X_2+X_3}{5}$ y $\hat{\lambda}_2 = \frac{2X_1+3X_2+3X_3}{10}$*

Problema 8.4. ***** Considere la siguiente función:*

$$f(x) = \frac{x}{a^2} e^{-\frac{x}{a}}, x > 0$$

1. *Demostrar que la función anterior es una función de densidad de una variable aleatoria X, si $a > 0$, haciendo uso de las propiedades de la función de densidad.*

2. *Calcular de forma aproximada la probabilidad de que el estimador del parámetro a, $\hat{a} = \frac{\bar{X}}{2}$ tome valores por encima de $\frac{3a}{4}$, cuando se extrae una mas de tamaño 32.*

Problema 8.5. *Se asume que las notas de un examen estarán distribuidas según una variable aleatoria X que tiene la siguiente función de densidad:*

$$f(x) = \begin{cases} \frac{1}{12} & 0 \leq x < 3 \\ \frac{1}{8} & 3 \leq x < 7 \\ k & 7 \leq x < 10 \end{cases}$$

donde k es una constante. Se pide:

1. *El valor de k.*

2. *Calcular y representar gráficamente la función de distribución de X.*

3. *La media y la varianza de X.*

4. *Suponiendo que a este examen se presentan 50 alumnos:*

   a) *Calcular de forma aproximada la probabilidad de que la nota media de los exámenes sea mayor que 6.*

   b) *Calcular la función de distribución de la nota más alta ¿Cuál es la probabilidad de que esa nota sea un sobresaliente?*

   c) *Calcular la función de distribución de la nota más baja ¿Cuál es la probabilidad de que esa nota sea un aprobado?*

Resultado

1. $k = \frac{1}{12}$

2. $F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{12} & 0 \leq x < 3 \\ \frac{x-1}{8} & 3 \leq x < 7 \\ \frac{x+2}{12} & 7 \leq x < 10 \\ 1 & x \geq 10 \end{cases}$

   La gráfica viene dada en la Figura 1.

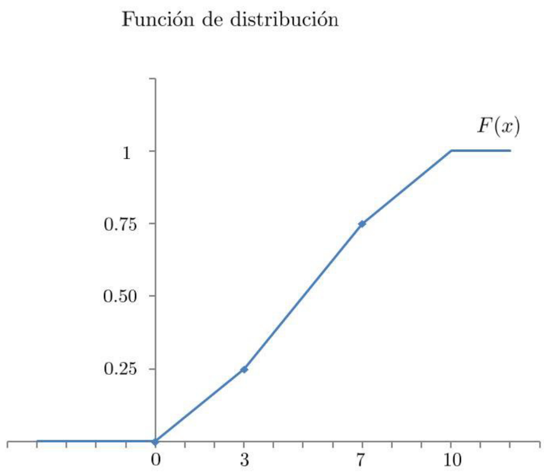

Figura 8.1: Función de distribución. Problema **??**

3. $E[X] = 5$, $V[X] = \frac{43}{6}$

4. Para una muestra con $n = 50$

   a) 0.00415

   b) $F_{MAX}(x) = \begin{cases} 0 & x < 0 \\ \left(\frac{x}{12}\right)^{50} & 0 \leq x < 3 \\ \left(\frac{x-1}{8}\right)^{50} & 3 \leq x < 7 \\ \left(\frac{x+2}{12}\right)^{50} & 7 \leq x < 10 \\ 1 & x \geq \end{cases}$

   c) 0.9871

   d) $F_{MIN}(x) = \begin{cases} 0 & x < 0 \\ 1 - \left(1 - \frac{x}{12}\right)^{50} & 0 \leq x < 3 \\ 1 - \left(\frac{9-x}{8}\right)^{50} & 3 \leq x < 7 \\ 1 - \left(\frac{10-x}{12}\right)^{50} & 7 \leq x < 10 \\ 1 & x \geq 10 \end{cases}$

e) $\simeq 0$

Problema 8.6. *Sea $X \sim N(\mu, \sigma^2)$. Sea $X_1, \ldots, X_n$ una mas de X, y $Y = aX + b$, con $a \in (0,1)$, $b > 0$. Utilizando estos datos, resolver los siguientes apartados.*

1. *Suponiendo μ y σ desconocidas, hallar un intervalo de confianza de $\mu_Y = E[Y]$ y otro para $\sigma_Y^2 = V[Y]$ con nivel de confianza $1-\alpha$.*
2. *Hallar $P[Y \leq y]$ en función de $\Phi(z)$ función de distribución de $Z \sim N(0,1)$.*
3. *Para estimar la media de Y se toman los estadísticos $\hat{\mu}_Y^1 = \bar{Y}$ y $\hat{\mu}_Y^2 = \frac{\bar{Y}}{a}$. ¿Cuál es mejor estimador?*
4. *Hallar la función generatriz de momentos de Y. Justificar qué distribución sigue Y.*

Sean $X \sim U[5]$ y $Z \sim U[5]$, con Z y X independientes. Sea $Y = \frac{1}{X}$ y $S = X + Z$.

1. *Hallar la función de distribución de Y. ¿Es Y continua o discreta?. Hallar $Cov[X,Y]$ y $Cov[Z,Y]$.*
2. *Hallar la función de probabilidad y la función de distribución de S.*
3. *Hallar $P\left[S = 7 | Y = \frac{1}{3}\right]$ y $P\left[Y = \frac{1}{2} | S = 5\right]$.*

Resultado

1. $IC(\mu_Y)_{1-\alpha} = \left[a\overline{X} + b - t_{n-1,\alpha/2}\frac{aS_X}{\sqrt{n}}, a\overline{X} + b + t_{n-1,\alpha/2}\frac{aS_X}{\sqrt{n}}\right], IC(\sigma_Y^2)_{1-\alpha} = \left[\frac{(n-1)(aS_X)^2}{j_{n-1,\alpha/2}}, \frac{(n-1)(aS_X)^2}{j_{n-1,1-\alpha/2}}\right]$
2. $\Phi\left(\frac{y-b-a\mu}{a\sigma}\right)$
3. μ_Y^1
4. $\varphi_Y(t) = e^{(\mu a+b)t+\frac{1}{2}(\sigma a)^2 t^2}$. $Y \sim N(a\mu + b, (a\sigma)^2)$.

   1.
$$F(y) = P[Y \leq y] = \begin{cases} 0 & y < 1/5 \\ 1/5 & 1/5 \leq y < 1/4 \\ 2/5 & 1/4 \leq y < 1/3 \\ 3/5 & 1/3 \leq y < 1/2 \\ 4/5 & 1/2 \leq y < 1 \\ 1 & y \geq 1 \end{cases}$$

   Y es una variable aleatoria discreta, concentrada en cinco puntos equiprobables.

   $Cov[X,Y] = -\frac{117}{20}$, $Cov[Y,Z] = 0$.

2.

$$F_S(s) = P[S \leq s] = \begin{cases} 0 & s < 2 \\ 1/25 & 2 \leq s < 3 \\ 3/25 & 3 \leq s < 4 \\ 6/25 & 4 \leq s < 5 \\ 10/25 & 5 \leq s < 6 \\ 15/25 & 6 \leq s < 7 \\ 19/25 & 7 \leq s < 8 \\ 22/25 & 8 \leq s < 9 \\ 24/25 & 9 \leq s < 10 \\ 1 & s \geq 10 \end{cases}$$

3. $P\left[S = 7 | Y = \frac{1}{3}\right] = \frac{1}{5}$, $P\left[Y = \frac{1}{2} | S = 5\right] = \frac{1}{4}$

Capítulo 9

Complemento Matemáticas

9.1. Sumas de series

Teorema: Serie Aritmética

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

Teorema: Sumas Geométricas Finitas

$$\sum_{k=0}^{n} p^k = \frac{1-p^{n+1}}{1-p}$$

$$\sum_{k=1}^{n} p^k = \frac{p-p^{n+1}}{1-p}$$

Teorema: Sumas Geométricas Infinitas

$$\sum_{k=0}^{\infty} p^k = \frac{1}{1-p},\ |p| < 1$$

$$\sum_{k=1}^{\infty} p^k = \frac{p}{1-p},\ |p| < 1$$

Teorema: Suma Aritmético-Geométrica orden 1

$$\sum_{k=0}^{\infty} kp^k = \frac{p}{(1-p)^2},\ |p| < 1$$

Teorema: Suma Aritmético-Geométrica orden 2

$$\sum_{k=0}^{\infty} k^2p^k = \frac{p(1+p)}{(1-p)^3},\ |p| < 1$$

Teorema: Suma Binomial Negativa

$$\sum_{k=0}^{\infty} \binom{k+r-1}{k} p^k = \left(\frac{1}{1-p}\right)^k, \ |p| < 1$$

Teorema: Binomio de Newton Sean $x, y \in \mathbb{R}$

$$\sum_{k=1}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n$$

Teorema: Suma número e

$$\sum_{k=0}^{\infty} \frac{1}{k!} = e$$

Teorema: Suma e^x

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$$

9.2. Límites

$$\lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x = e$$
$$\lim_{x\to\infty} \left(1 + \frac{a}{x}\right)^x = e^a$$

9.3. Función Gamma

$$\Gamma(p) = \int_0^{+\infty} e^{-x} x^{p-1} dx$$

Propiedades:

1. $$\Gamma(p) = (p-1)\Gamma(p-1)$$

2. $$\Gamma(1) = 1$$

3. $$\Gamma(n) = (n-1)!$$

4. $$\Gamma(1/2) = \sqrt{\pi}$$

Capítulo 10

Formulario de examen y tablas

Distribuciones Discretas					
Distribución	**Notación**	p_k	$\varphi(t)$	$E[X]$	$V[X]$
Uniforme Discreta	$U(n)$	$\frac{1}{n}$	$\frac{e^{t(n+1)}-e^t}{n(e^t-1)}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Bernoulli	$Be(p)$	$p^k q^{1-k}$	$q+pe^t$	p	pq
Binomial	$Bi(n,p)$	$\binom{n}{k}p^k q^{n-k}$	$(q+pe^t)^n$	np	npq
Geométrica	$Ge(p)$	$q^k p$	$\frac{p}{1-qe^t}$	$\frac{q}{p}$	$\frac{q}{p^2}$
Binomial Negativa	$BN(r,p)$	$\binom{k+r-1}{k}p^r q^k$	$\left(\frac{p}{1-qe^t}\right)^r$	$\frac{rq}{p}$	$\frac{rq}{p^2}$
Poisson	$\mathcal{P}o(\lambda)$	$e^{-\lambda}\frac{\lambda^k}{k!}$	$e^{-\lambda(1-e^t)}$	λ	λ

Distribuciones Continuas						
Distribución	**Notación**	$f(x)$	$F(x)$	$\varphi(t)$	$E[X]$	$V[X]$
Uniforme Continua	$U[a,b]$	$\frac{1}{b-a}$	$\frac{x-a}{b-a}$	$\frac{e^{bt}-e^{at}}{t(b-a)}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$
Exponencial	$Exp(\lambda)$	$\lambda e^{-\lambda x}$	$1-e^{-\lambda x}$	$\frac{\lambda}{\lambda-t}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma	$Ga(p,a)$	$\frac{a^p}{\Gamma(p)}e^{-ax}x^{p-1}$	$-$	$\left(\frac{a}{a-t}\right)^p$	$\frac{p}{a}$	$\frac{p}{a^2}$
Erlang	$Erlang(n,\lambda)$	$\frac{\lambda^n}{(n-1)!}e^{-\lambda x}x^{n-1}$	$1-e^{-\lambda x}\sum_{i=0}^{n-1}\frac{(\lambda x)^i}{i!}$	$\left(\frac{\lambda}{\lambda-t}\right)^n$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$
Normal	$N(\mu,\sigma^2)$	$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$-$	$e^{\mu t+\frac{1}{2}\sigma^2 t^2}$	μ	σ^2

Intervalos de Confianza (X)	
Parámetro	$IC_{1-\alpha}$
μ	$\left[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$
σ^2	$\left[\frac{(n-1)S^2}{j_{n-1,\alpha/2}}, \frac{(n-1)S^2}{j_{n-1,1-\alpha/2}}\right]$
μ	$\left[\overline{X} - t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}, \overline{X} + t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}\right]$

Intervalos de Confianza (X e Y)	
Parámetro	$IC_{1-\alpha}$
$\mu_X - \mu_Y$	$\left[(\overline{X} - \overline{Y}) - z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}, (\overline{X} - \overline{Y}) + z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right]$
$\mu_X - \mu_Y$	$\left[(\overline{X} - \overline{Y}) - t_{n+m-2,\alpha/2} s_p\sqrt{\frac{1}{n} + \frac{1}{m}}, (\overline{X} - \overline{Y}) + t_{n+m-2,\alpha/2} s_p\sqrt{\frac{1}{n} + \frac{1}{m}}\right]$
σ_X^2/σ_Y^2	$\left[\frac{S_X^2}{S_Y^2 f_{n-1,m-1,\alpha/2}}, \frac{S_X^2}{S_Y^2 f_{n-1,m-1,1-\alpha/2}}\right]$

$$s_p = \sqrt{\frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2}}$$

Distribución Normal

Z~N(0,1) P[Z≤z_α]=1-α

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,99980	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,99990	0,99990	0,99990	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997
4,0	0,99997	0,99997	0,99997	0,99997	0,99997	0,99997	0,99998	0,99998	0,99998	0,99998

Distribución t-Student

$T \sim t_n \quad P[T>t_{n,\alpha}]=\alpha$

n	**0,125**	**0,100**	**0,075**	**0,070**	**0,065**	**0,060**	**0,055**	**0,050**	**0,045**	**0,040**	**0,035**	**0,030**	**0,025**	**0,020**	**0,015**	**0,010**	**0,005**
1	2,4142	3,0777	4,1653	4,4737	4,8288	5,2422	5,7297	6,3138	7,0264	7,9158	9,0579	10,5789	12,7062	15,8945	21,2049	31,8205	63,6567
2	1,6036	1,8856	2,2819	2,3834	2,4954	2,6202	2,7604	2,9200	3,1040	3,3198	3,5782	3,8964	4,3027	4,8487	5,6428	6,9646	9,9248
3	1,4226	1,6377	1,9243	1,9950	2,0719	2,1562	2,2494	2,3534	2,4708	2,6054	2,7626	2,9505	3,1824	3,4819	3,8960	4,5407	5,8409
4	1,3444	1,5332	1,7782	1,8375	1,9016	1,9712	2,0475	2,1318	2,2261	2,3329	2,4559	2,6008	2,7764	2,9985	3,2976	3,7469	4,6041
5	1,3009	1,4759	1,6994	1,7529	1,8104	1,8727	1,9405	2,0150	2,0978	2,1910	2,2974	2,4216	2,5706	2,7565	3,0029	3,3649	4,0321
6	1,2733	1,4398	1,6502	1,7002	1,7538	1,8117	1,8744	1,9432	2,0192	2,1043	2,2011	2,3133	2,4469	2,6122	2,8289	3,1427	3,7074
7	1,2543	1,4149	1,6166	1,6643	1,7153	1,7702	1,8297	1,8946	1,9662	2,0460	2,1365	2,2409	2,3646	2,5168	2,7146	2,9980	3,4995
8	1,2403	1,3968	1,5922	1,6383	1,6874	1,7402	1,7973	1,8595	1,9280	2,0042	2,0902	2,1892	2,3060	2,4490	2,6338	2,8965	3,3554
9	1,2297	1,3830	1,5737	1,6185	1,6663	1,7176	1,7729	1,8331	1,8992	1,9727	2,0554	2,1504	2,2622	2,3984	2,5738	2,8214	3,2498
10	1,2213	1,3722	1,5592	1,6031	1,6498	1,6998	1,7538	1,8125	1,8768	1,9481	2,0283	2,1202	2,2281	2,3593	2,5275	2,7638	3,1693
11	1,2145	1,3634	1,5476	1,5906	1,6365	1,6856	1,7385	1,7959	1,8588	1,9284	2,0067	2,0961	2,2010	2,3281	2,4907	2,7181	3,1058
12	1,2089	1,3562	1,5380	1,5804	1,6256	1,6739	1,7259	1,7823	1,8440	1,9123	1,9889	2,0764	2,1788	2,3027	2,4607	2,6810	3,0545
13	1,2041	1,3502	1,5299	1,5718	1,6164	1,6641	1,7154	1,7709	1,8317	1,8989	1,9742	2,0600	2,1604	2,2816	2,4358	2,6503	3,0123
14	1,2001	1,3450	1,5231	1,5646	1,6087	1,6558	1,7064	1,7613	1,8213	1,8875	1,9617	2,0462	2,1448	2,2638	2,4149	2,6245	2,9768
15	1,1967	1,3406	1,5172	1,5583	1,6020	1,6487	1,6988	1,7531	1,8123	1,8777	1,9509	2,0343	2,1314	2,2485	2,3970	2,6025	2,9467
16	1,1937	1,3368	1,5121	1,5529	1,5962	1,6425	1,6921	1,7459	1,8046	1,8693	1,9417	2,0240	2,1199	2,2354	2,3815	2,5835	2,9208
17	1,1910	1,3334	1,5077	1,5482	1,5911	1,6370	1,6863	1,7396	1,7978	1,8619	1,9335	2,0150	2,1098	2,2238	2,3681	2,5669	2,8982
18	1,1887	1,3304	1,5037	1,5439	1,5867	1,6322	1,6812	1,7341	1,7918	1,8553	1,9264	2,0071	2,1009	2,2137	2,3562	2,5524	2,8784
19	1,1866	1,3277	1,5002	1,5402	1,5827	1,6280	1,6766	1,7291	1,7864	1,8495	1,9200	2,0000	2,0930	2,2047	2,3456	2,5395	2,8609
20	1,1848	1,3253	1,4970	1,5369	1,5791	1,6242	1,6725	1,7247	1,7816	1,8443	1,9143	1,9937	2,0860	2,1967	2,3362	2,5280	2,8453
21	1,1831	1,3232	1,4942	1,5338	1,5759	1,6207	1,6688	1,7207	1,7773	1,8397	1,9092	1,9880	2,0796	2,1894	2,3278	2,5176	2,8314
22	1,1815	1,3212	1,4916	1,5311	1,5730	1,6176	1,6655	1,7171	1,7734	1,8354	1,9045	1,9829	2,0739	2,1829	2,3202	2,5083	2,8188
23	1,1802	1,3195	1,4893	1,5286	1,5703	1,6148	1,6624	1,7139	1,7699	1,8316	1,9003	1,9782	2,0687	2,1770	2,3132	2,4999	2,8073
24	1,1789	1,3178	1,4871	1,5263	1,5679	1,6122	1,6596	1,7109	1,7667	1,8281	1,8965	1,9740	2,0639	2,1715	2,3069	2,4922	2,7969
25	1,1777	1,3163	1,4852	1,5242	1,5657	1,6098	1,6571	1,7081	1,7637	1,8248	1,8929	1,9701	2,0595	2,1666	2,3011	2,4851	2,7874
26	1,1766	1,3150	1,4834	1,5223	1,5636	1,6076	1,6547	1,7056	1,7610	1,8219	1,8897	1,9665	2,0555	2,1620	2,2958	2,4786	2,7787
27	1,1756	1,3137	1,4817	1,5205	1,5617	1,6056	1,6526	1,7033	1,7585	1,8191	1,8867	1,9632	2,0518	2,1578	2,2909	2,4727	2,7707
28	1,1747	1,3125	1,4801	1,5189	1,5600	1,6037	1,6506	1,7011	1,7561	1,8166	1,8839	1,9601	2,0484	2,1539	2,2864	2,4671	2,7633
29	1,1739	1,3114	1,4787	1,5174	1,5583	1,6020	1,6487	1,6991	1,7540	1,8142	1,8813	1,9573	2,0452	2,1503	2,2822	2,4620	2,7564
30	1,1731	1,3104	1,4774	1,5159	1,5568	1,6004	1,6470	1,6973	1,7520	1,8120	1,8789	1,9546	2,0423	2,1470	2,2783	2,4573	2,7500

Distribución Chi Cuadrado

$J \sim \chi 2 \quad P[J>j_{n,\alpha}]=\alpha$

n	0,995	0,990	0,985	0,980	0,975	0,950	0,925	0,900	0,875	0,125	0,100	0,075	0,050	0,025	0,020	0,015	0,010	0,005
1	0,0000	0,0002	0,0004	0,0006	0,0010	0,0039	0,0089	0,0158	0,0247	2,3535	2,7055	3,1701	3,8415	5,0239	5,4119	5,9165	6,6349	7,8794
2	0,0100	0,0201	0,0302	0,0404	0,0506	0,1026	0,1559	0,2107	0,2671	4,1589	4,6052	5,1805	5,9915	7,3778	7,8240	8,3994	9,2103	10,5966
3	0,0717	0,1148	0,1516	0,1848	0,2158	0,3518	0,4720	0,5844	0,6924	5,7394	6,2514	6,9046	7,8147	9,3484	9,8374	10,4650	11,3449	12,8382
4	0,2070	0,2971	0,3682	0,4294	0,4844	0,7107	0,8969	1,0636	1,2188	7,2140	7,7794	8,4963	9,4877	11,1433	11,6678	12,3391	13,2767	14,8603
5	0,4117	0,5543	0,6618	0,7519	0,8312	1,1455	1,3937	1,6103	1,8082	8,6248	9,2364	10,0083	11,0705	12,8325	13,3882	14,0978	15,0863	16,7496
6	0,6757	0,8721	1,0160	1,1344	1,2373	1,6354	1,9415	2,2041	2,4411	9,9917	10,6446	11,4659	12,5916	14,4494	15,0332	15,7774	16,8119	18,5476
7	0,9893	1,2390	1,4184	1,5643	1,6899	2,1673	2,5277	2,8331	3,1063	11,3264	12,0170	12,8834	14,0671	16,0128	16,6224	17,3984	18,4753	20,2777
8	1,3444	1,6465	1,8603	2,0325	2,1797	2,7326	3,1440	3,4895	3,7965	12,6361	13,3616	14,2697	15,5073	17,5345	18,1682	18,9739	20,0902	21,9550
9	1,7349	2,0879	2,3349	2,5324	2,7004	3,3251	3,7847	4,1682	4,5070	13,9255	14,6837	15,6309	16,9190	19,0228	19,6790	20,5125	21,6660	23,5894
10	2,1559	2,5582	2,8372	3,0591	3,2470	3,9403	4,4459	4,8652	5,2341	15,1982	15,9872	16,9714	18,3070	20,4832	21,1608	22,0206	23,2093	25,1882
11	2,6032	3,0535	3,3634	3,6087	3,8157	4,5748	5,1243	5,5778	5,9754	16,4568	17,2750	18,2942	19,6751	21,9200	22,6179	23,5028	24,7250	26,7568
12	3,0738	3,5706	3,9104	4,1783	4,4038	5,2260	5,8175	6,3038	6,7288	17,7033	18,5493	19,6020	21,0261	23,3367	24,0540	24,9628	26,2170	28,2995
13	3,5650	4,1069	4,4757	4,7654	5,0088	5,8919	6,5238	7,0415	7,4929	18,9392	19,8119	20,8966	22,3620	24,7356	25,4715	26,4034	27,6882	29,8195
14	4,0747	4,6604	5,0572	5,3682	5,6287	6,5706	7,2415	7,7895	8,2662	20,1658	21,0641	22,1795	23,6848	26,1189	26,8728	27,8268	29,1412	31,3193
15	4,6009	5,2293	5,6534	5,9849	6,2621	7,2609	7,9695	8,5468	9,0479	21,3841	22,3071	23,4522	24,9958	27,4884	28,2595	29,2349	30,5779	32,8013
16	5,1422	5,8122	6,2628	6,6142	6,9077	7,9616	8,7067	9,3122	9,8370	22,5949	23,5418	24,7155	26,2962	28,8454	29,6332	30,6292	31,9999	34,2672
17	5,6972	6,4078	6,8842	7,2550	7,5642	8,6718	9,4522	10,0852	10,6329	23,7990	24,7690	25,9705	27,5871	30,1910	30,9950	32,0112	33,4087	35,7185
18	6,2648	7,0149	7,5165	7,9062	8,2307	9,3905	10,2053	10,8649	11,4349	24,9970	25,9894	27,2178	28,8693	31,5264	32,3462	33,3817	34,8053	37,1565
19	6,8440	7,6327	8,1588	8,5670	8,9065	10,1170	10,9653	11,6509	12,2425	26,1893	27,2036	28,4581	30,1435	32,8523	33,6874	34,7420	36,1909	38,5823
20	7,4338	8,2604	8,8105	9,2367	9,5908	10,8508	11,7317	12,4426	13,0553	27,3765	28,4120	29,6920	31,4104	34,1696	35,0196	36,0926	37,5662	39,9968
21	8,0337	8,8972	9,4708	9,9146	10,2829	11,5913	12,5041	13,2396	13,8728	28,5589	29,6151	30,9200	32,6706	35,4789	36,3434	37,4345	38,9322	41,4011
22	8,6427	9,5425	10,1390	10,6000	10,9823	12,3380	13,2819	14,0415	14,6948	29,7369	30,8133	32,1424	33,9244	36,7807	37,6595	38,7681	40,2894	42,7957
23	9,2604	10,1957	10,8147	11,2926	11,6886	13,0905	14,0648	14,8480	15,5209	30,9108	32,0069	33,3597	35,1725	38,0756	38,9683	40,0941	41,6384	44,1813
24	9,8862	10,8564	11,4974	11,9918	12,4012	13,8484	14,8525	15,6587	16,3508	32,0809	33,1962	34,5723	36,4150	39,3641	40,2704	41,4130	42,9798	45,5585
25	10,5197	11,5240	12,1867	12,6973	13,1197	14,6114	15,6447	16,4734	17,1843	33,2473	34,3816	35,7803	37,6525	40,6465	41,5661	42,7252	44,3141	46,9279
26	11,1602	12,1981	12,8821	13,4086	13,8439	15,3792	16,4410	17,2919	18,0212	34,4104	35,5632	36,9841	38,8851	41,9232	42,8558	44,0311	45,6417	48,2899
27	11,8076	12,8785	13,5833	14,1254	14,5734	16,1514	17,2414	18,1139	18,8613	35,5703	36,7412	38,1840	40,1133	43,1945	44,1400	45,3311	46,9629	49,6449
28	12,4613	13,5647	14,2900	14,8475	15,3079	16,9279	18,0454	18,9392	19,7044	36,7272	37,9159	39,3801	41,3371	44,4608	45,4188	46,6256	48,2782	50,9934
29	13,1211	14,2565	15,0019	15,5745	16,0471	17,7084	18,8530	19,7677	20,5503	37,8812	39,0875	40,5727	42,5570	45,7223	46,6927	47,9147	49,5879	52,3356
30	13,7867	14,9535	15,7188	16,3062	16,7908	18,4927	19,6639	20,5992	21,3989	39,0326	40,2560	41,7619	43,7730	46,9792	47,9618	49,1989	50,8922	53,6720

Distribución F-Snedecor

0,1 $F \sim \mathcal{F}_{n,m}$ $P[F>f_{n,m,0.1}] = 0.1$

n \ m	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	24	25	29	30	99	100
2	9,000	5,462	4,325	3,780	3,463	3,257	3,113	3,006	2,924	2,860	2,807	2,763	2,726	2,695	2,668	2,645	2,624	2,606	2,589	2,538	2,528	2,495	2,489	2,357	2,356
3	9,162	5,391	4,191	3,619	3,289	3,074	2,924	2,813	2,728	2,660	2,606	2,560	2,522	2,490	2,462	2,437	2,416	2,397	2,380	2,327	2,317	2,283	2,276	2,140	2,139
4	9,243	5,343	4,107	3,520	3,181	2,961	2,806	2,693	2,605	2,536	2,480	2,434	2,395	2,361	2,333	2,308	2,286	2,266	2,249	2,195	2,184	2,149	2,142	2,003	2,002
5	9,293	5,309	4,051	3,453	3,108	2,883	2,726	2,611	2,522	2,451	2,394	2,347	2,307	2,273	2,244	2,218	2,196	2,176	2,158	2,103	2,092	2,057	2,049	1,906	1,906
6	9,326	5,285	4,010	3,405	3,055	2,827	2,668	2,551	2,461	2,389	2,331	2,283	2,243	2,208	2,178	2,152	2,130	2,109	2,091	2,035	2,024	1,988	1,980	1,835	1,834
7	9,349	5,266	3,979	3,368	3,014	2,785	2,624	2,505	2,414	2,342	2,283	2,234	2,193	2,158	2,128	2,102	2,079	2,058	2,040	1,983	1,971	1,935	1,927	1,778	1,778
8	9,367	5,252	3,955	3,339	2,983	2,752	2,589	2,469	2,377	2,304	2,245	2,195	2,154	2,119	2,088	2,061	2,038	2,017	1,999	1,941	1,929	1,892	1,884	1,733	1,732
9	9,381	5,240	3,936	3,316	2,958	2,725	2,561	2,440	2,347	2,274	2,214	2,164	2,122	2,086	2,055	2,028	2,005	1,984	1,965	1,906	1,895	1,857	1,849	1,696	1,695
10	9,392	5,230	3,920	3,297	2,937	2,703	2,538	2,416	2,323	2,248	2,188	2,138	2,095	2,059	2,028	2,001	1,977	1,956	1,937	1,877	1,866	1,827	1,819	1,664	1,663
11	9,401	5,222	3,907	3,282	2,920	2,684	2,519	2,396	2,302	2,227	2,166	2,116	2,073	2,037	2,005	1,978	1,954	1,932	1,913	1,853	1,841	1,802	1,794	1,637	1,636
12	9,408	5,216	3,896	3,268	2,905	2,668	2,502	2,379	2,284	2,209	2,147	2,097	2,054	2,017	1,985	1,958	1,933	1,912	1,892	1,832	1,820	1,781	1,773	1,613	1,612
13	9,415	5,210	3,886	3,257	2,892	2,654	2,488	2,364	2,269	2,193	2,131	2,080	2,037	2,000	1,968	1,940	1,916	1,894	1,875	1,814	1,802	1,762	1,754	1,592	1,592
14	9,420	5,205	3,878	3,247	2,881	2,643	2,475	2,351	2,255	2,179	2,117	2,066	2,022	1,985	1,953	1,925	1,900	1,878	1,859	1,797	1,785	1,745	1,737	1,574	1,573
15	9,425	5,200	3,870	3,238	2,871	2,632	2,464	2,340	2,244	2,167	2,105	2,053	2,010	1,972	1,940	1,912	1,887	1,865	1,845	1,783	1,771	1,731	1,722	1,557	1,557
16	9,429	5,196	3,864	3,230	2,863	2,623	2,455	2,329	2,233	2,156	2,094	2,042	1,998	1,961	1,928	1,900	1,875	1,852	1,833	1,770	1,758	1,717	1,709	1,542	1,542
17	9,433	5,193	3,858	3,223	2,855	2,615	2,446	2,320	2,224	2,147	2,084	2,032	1,988	1,950	1,917	1,889	1,864	1,841	1,821	1,759	1,746	1,705	1,697	1,529	1,528
18	9,436	5,190	3,853	3,217	2,848	2,607	2,438	2,312	2,215	2,138	2,075	2,023	1,978	1,941	1,908	1,879	1,854	1,831	1,811	1,748	1,736	1,695	1,686	1,517	1,516
19	9,439	5,187	3,849	3,212	2,842	2,601	2,431	2,305	2,208	2,130	2,067	2,014	1,970	1,932	1,899	1,870	1,845	1,822	1,802	1,739	1,726	1,685	1,676	1,505	1,505
20	9,441	5,184	3,844	3,207	2,836	2,595	2,425	2,298	2,201	2,123	2,060	2,007	1,962	1,924	1,891	1,862	1,837	1,814	1,794	1,730	1,718	1,676	1,667	1,495	1,494
24	9,450	5,176	3,831	3,191	2,818	2,575	2,404	2,277	2,178	2,100	2,036	1,983	1,938	1,899	1,866	1,836	1,810	1,787	1,767	1,702	1,689	1,647	1,638	1,461	1,460
25	9,451	5,175	3,828	3,187	2,815	2,571	2,400	2,272	2,174	2,095	2,031	1,978	1,933	1,894	1,860	1,831	1,805	1,782	1,761	1,696	1,683	1,640	1,632	1,454	1,453
29	9,457	5,169	3,819	3,176	2,803	2,558	2,386	2,258	2,159	2,080	2,015	1,961	1,916	1,876	1,843	1,813	1,787	1,763	1,742	1,676	1,663	1,620	1,611	1,429	1,428
30	9,458	5,168	3,817	3,174	2,800	2,555	2,383	2,255	2,155	2,076	2,011	1,958	1,912	1,873	1,839	1,809	1,783	1,759	1,738	1,672	1,659	1,616	1,606	1,423	1,423
99	9,481	5,144	3,778	3,127	2,746	2,497	2,321	2,189	2,087	2,005	1,938	1,882	1,834	1,793	1,757	1,726	1,698	1,673	1,651	1,579	1,565	1,517	1,507	1,295	1,294
100	9,481	5,144	3,778	3,126	2,746	2,497	2,321	2,189	2,087	2,005	1,938	1,882	1,834	1,793	1,757	1,726	1,698	1,673	1,650	1,579	1,565	1,517	1,507	1,294	1,293

Distribución F-Snedecor

$F \sim F_{n,m} \quad P[F > f_{n,m,0.05}] = 0.05$

0,05

n \ m	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	24	25	29	30	99	100
2	19,000	9,552	6,944	5,786	5,143	4,737	4,459	4,256	4,103	3,982	3,885	3,806	3,739	3,682	3,634	3,592	3,555	3,522	3,493	3,403	3,385	3,328	3,316	3,088	3,087
3	19,164	9,277	6,591	5,409	4,757	4,347	4,066	3,863	3,708	3,587	3,490	3,411	3,344	3,287	3,239	3,197	3,160	3,127	3,098	3,009	2,991	2,934	2,922	2,696	2,696
4	19,247	9,117	6,388	5,192	4,534	4,120	3,838	3,633	3,478	3,357	3,259	3,179	3,112	3,056	3,007	2,965	2,928	2,895	2,866	2,776	2,759	2,701	2,690	2,464	2,463
5	19,296	9,013	6,256	5,050	4,387	3,972	3,687	3,482	3,326	3,204	3,106	3,025	2,958	2,901	2,852	2,810	2,773	2,740	2,711	2,621	2,603	2,545	2,534	2,306	2,305
6	19,330	8,941	6,163	4,950	4,284	3,866	3,581	3,374	3,217	3,095	2,996	2,915	2,848	2,790	2,741	2,699	2,661	2,628	2,599	2,508	2,490	2,432	2,421	2,192	2,191
7	19,353	8,887	6,094	4,876	4,207	3,787	3,500	3,293	3,135	3,012	2,913	2,832	2,764	2,707	2,657	2,614	2,577	2,544	2,514	2,423	2,405	2,346	2,334	2,103	2,103
8	19,371	8,845	6,041	4,818	4,147	3,726	3,438	3,230	3,072	2,948	2,849	2,767	2,699	2,641	2,591	2,548	2,510	2,477	2,447	2,355	2,337	2,278	2,266	2,033	2,032
9	19,385	8,812	5,999	4,772	4,099	3,677	3,388	3,179	3,020	2,896	2,796	2,714	2,646	2,588	2,538	2,494	2,456	2,423	2,393	2,300	2,282	2,223	2,211	1,976	1,975
10	19,396	8,786	5,964	4,735	4,060	3,637	3,347	3,137	2,978	2,854	2,753	2,671	2,602	2,544	2,494	2,450	2,412	2,378	2,348	2,255	2,236	2,177	2,165	1,928	1,927
11	19,405	8,763	5,936	4,704	4,027	3,603	3,313	3,102	2,943	2,818	2,717	2,635	2,565	2,507	2,456	2,413	2,374	2,340	2,310	2,216	2,198	2,138	2,126	1,887	1,886
12	19,413	8,745	5,912	4,678	4,000	3,575	3,284	3,073	2,913	2,788	2,687	2,604	2,534	2,475	2,425	2,381	2,342	2,308	2,278	2,183	2,165	2,104	2,092	1,851	1,850
13	19,419	8,729	5,891	4,655	3,976	3,550	3,259	3,048	2,887	2,761	2,660	2,577	2,507	2,448	2,397	2,353	2,314	2,280	2,250	2,155	2,136	2,075	2,063	1,820	1,819
14	19,424	8,715	5,873	4,636	3,956	3,529	3,237	3,025	2,865	2,739	2,637	2,554	2,484	2,424	2,373	2,329	2,290	2,256	2,225	2,130	2,111	2,050	2,037	1,793	1,792
15	19,429	8,703	5,858	4,619	3,938	3,511	3,218	3,006	2,845	2,719	2,617	2,533	2,463	2,403	2,352	2,308	2,269	2,234	2,203	2,108	2,089	2,027	2,015	1,769	1,768
16	19,433	8,692	5,844	4,604	3,922	3,494	3,202	2,989	2,828	2,701	2,599	2,515	2,445	2,385	2,333	2,289	2,250	2,215	2,184	2,088	2,069	2,007	1,995	1,747	1,746
17	19,437	8,683	5,832	4,590	3,908	3,480	3,187	2,974	2,812	2,685	2,583	2,499	2,428	2,368	2,317	2,272	2,233	2,198	2,167	2,070	2,051	1,989	1,976	1,727	1,726
18	19,440	8,675	5,821	4,579	3,896	3,467	3,173	2,960	2,798	2,671	2,568	2,484	2,413	2,353	2,302	2,257	2,217	2,182	2,151	2,054	2,035	1,973	1,960	1,709	1,708
19	19,443	8,667	5,811	4,568	3,884	3,455	3,161	2,948	2,785	2,658	2,555	2,471	2,400	2,340	2,288	2,243	2,203	2,168	2,137	2,040	2,021	1,958	1,945	1,693	1,691
20	19,446	8,660	5,803	4,558	3,874	3,445	3,150	2,936	2,774	2,646	2,544	2,459	2,388	2,328	2,276	2,230	2,191	2,155	2,124	2,027	2,007	1,945	1,932	1,678	1,676
24	19,454	8,639	5,774	4,527	3,841	3,410	3,115	2,900	2,737	2,609	2,505	2,420	2,349	2,288	2,235	2,190	2,150	2,114	2,082	1,984	1,964	1,901	1,887	1,628	1,627
25	19,456	8,634	5,769	4,521	3,835	3,404	3,108	2,893	2,730	2,601	2,498	2,412	2,341	2,280	2,227	2,181	2,141	2,106	2,074	1,975	1,955	1,891	1,878	1,617	1,616
29	19,461	8,620	5,750	4,500	3,813	3,381	3,084	2,869	2,705	2,576	2,472	2,386	2,314	2,253	2,200	2,154	2,113	2,077	2,045	1,945	1,926	1,861	1,847	1,582	1,581
30	19,462	8,617	5,746	4,496	3,808	3,376	3,079	2,864	2,700	2,570	2,466	2,380	2,308	2,247	2,194	2,148	2,107	2,071	2,039	1,939	1,919	1,854	1,841	1,574	1,573
99	19,486	8,554	5,664	4,405	3,712	3,275	2,975	2,756	2,589	2,457	2,350	2,262	2,188	2,124	2,069	2,021	1,979	1,941	1,907	1,801	1,780	1,710	1,696	1,394	1,393
100	19,486	8,554	5,664	4,405	3,712	3,275	2,975	2,756	2,588	2,457	2,350	2,261	2,187	2,123	2,068	2,020	1,978	1,940	1,907	1,800	1,779	1,710	1,695	1,393	1,392

Distribución F-Snedecor

0,025 $F \sim \mathcal{F}_{n,m}$ $P[F>f_{n,m,0.025}] = 0.025$

n \ m	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	24	25	29	30	99	100
2	39,000	16,044	10,649	8,434	7,260	6,542	6,059	5,715	5,456	5,256	5,096	4,965	4,857	4,765	4,687	4,619	4,560	4,508	4,461	4,319	4,291	4,201	4,182	3,830	3,828
3	39,165	15,439	9,979	7,764	6,599	5,890	5,416	5,078	4,826	4,630	4,474	4,347	4,242	4,153	4,077	4,011	3,954	3,903	3,859	3,721	3,694	3,607	3,589	3,251	3,250
4	39,248	15,101	9,605	7,388	6,227	5,523	5,053	4,718	4,468	4,275	4,121	3,996	3,892	3,804	3,729	3,665	3,608	3,559	3,515	3,379	3,353	3,267	3,250	2,918	2,917
5	39,298	14,885	9,364	7,146	5,988	5,285	4,817	4,484	4,236	4,044	3,891	3,767	3,663	3,576	3,502	3,438	3,382	3,333	3,289	3,155	3,129	3,044	3,026	2,697	2,696
6	39,331	14,735	9,197	6,978	5,820	5,119	4,652	4,320	4,072	3,881	3,728	3,604	3,501	3,415	3,341	3,277	3,221	3,172	3,128	2,995	2,969	2,884	2,867	2,539	2,537
7	39,355	14,624	9,074	6,853	5,695	4,995	4,529	4,197	3,950	3,759	3,607	3,483	3,380	3,293	3,219	3,156	3,100	3,051	3,007	2,874	2,848	2,763	2,746	2,418	2,417
8	39,373	14,540	8,980	6,757	5,600	4,899	4,433	4,102	3,855	3,664	3,512	3,388	3,285	3,199	3,125	3,061	3,005	2,956	2,913	2,779	2,753	2,669	2,651	2,323	2,321
9	39,387	14,473	8,905	6,681	5,523	4,823	4,357	4,026	3,779	3,588	3,436	3,312	3,209	3,123	3,049	2,985	2,929	2,880	2,837	2,703	2,677	2,592	2,575	2,245	2,244
10	39,398	14,419	8,844	6,619	5,461	4,761	4,295	3,964	3,717	3,526	3,374	3,250	3,147	3,060	2,986	2,922	2,866	2,817	2,774	2,640	2,613	2,529	2,511	2,181	2,179
11	39,407	14,374	8,794	6,568	5,410	4,709	4,243	3,912	3,665	3,474	3,321	3,197	3,095	3,008	2,934	2,870	2,814	2,765	2,721	2,586	2,560	2,475	2,458	2,126	2,124
12	39,415	14,337	8,751	6,525	5,366	4,666	4,200	3,868	3,621	3,430	3,277	3,153	3,050	2,963	2,889	2,825	2,769	2,720	2,676	2,541	2,515	2,430	2,412	2,079	2,077
13	39,421	14,304	8,715	6,488	5,329	4,628	4,162	3,831	3,583	3,392	3,239	3,115	3,012	2,925	2,851	2,786	2,730	2,681	2,637	2,502	2,476	2,390	2,372	2,038	2,036
14	39,427	14,277	8,684	6,456	5,297	4,596	4,130	3,798	3,550	3,359	3,206	3,082	2,979	2,891	2,817	2,753	2,696	2,647	2,603	2,468	2,441	2,355	2,338	2,001	2,000
15	39,431	14,253	8,657	6,428	5,269	4,568	4,101	3,769	3,522	3,330	3,177	3,053	2,949	2,862	2,788	2,723	2,667	2,617	2,573	2,437	2,411	2,325	2,307	1,969	1,968
16	39,435	14,232	8,633	6,403	5,244	4,543	4,076	3,744	3,496	3,304	3,152	3,027	2,923	2,836	2,761	2,697	2,640	2,591	2,547	2,411	2,384	2,298	2,280	1,941	1,939
17	39,439	14,213	8,611	6,381	5,222	4,521	4,054	3,722	3,474	3,282	3,129	3,004	2,900	2,813	2,738	2,673	2,617	2,567	2,523	2,386	2,360	2,273	2,255	1,915	1,913
18	39,442	14,196	8,592	6,362	5,202	4,501	4,034	3,701	3,453	3,261	3,108	2,983	2,879	2,792	2,717	2,652	2,596	2,546	2,501	2,365	2,338	2,251	2,233	1,891	1,890
19	39,445	14,181	8,575	6,344	5,184	4,483	4,016	3,683	3,435	3,243	3,090	2,965	2,861	2,773	2,698	2,633	2,576	2,526	2,482	2,345	2,318	2,231	2,213	1,870	1,868
20	39,448	14,167	8,560	6,329	5,168	4,467	3,999	3,667	3,419	3,226	3,073	2,948	2,844	2,756	2,681	2,616	2,559	2,509	2,464	2,327	2,300	2,213	2,195	1,850	1,849
24	39,456	14,124	8,511	6,278	5,117	4,415	3,947	3,614	3,365	3,173	3,019	2,893	2,789	2,701	2,625	2,560	2,503	2,452	2,408	2,269	2,242	2,154	2,136	1,785	1,784
25	39,458	14,115	8,501	6,268	5,107	4,405	3,937	3,604	3,355	3,162	3,008	2,882	2,778	2,689	2,614	2,548	2,491	2,441	2,396	2,257	2,230	2,142	2,124	1,772	1,770
29	39,463	14,087	8,468	6,234	5,072	4,370	3,901	3,568	3,319	3,125	2,971	2,845	2,740	2,652	2,576	2,510	2,453	2,402	2,357	2,217	2,190	2,101	2,083	1,726	1,725
30	39,465	14,081	8,461	6,227	5,065	4,362	3,894	3,560	3,311	3,118	2,963	2,837	2,732	2,644	2,568	2,502	2,445	2,394	2,349	2,209	2,182	2,092	2,074	1,716	1,715
99	39,488	13,957	8,320	6,081	4,916	4,211	3,740	3,404	3,152	2,957	2,800	2,672	2,565	2,475	2,397	2,329	2,270	2,218	2,171	2,025	1,996	1,902	1,882	1,486	1,484
100	39,488	13,956	8,319	6,080	4,915	4,210	3,739	3,403	3,152	2,956	2,800	2,671	2,565	2,474	2,396	2,329	2,269	2,217	2,170	2,024	1,996	1,901	1,882	1,485	1,483

Distribución F-Snedecor

0,01 $F \sim \mathcal{F}_{n,m}$ $P[F>f_{n,m,0.01}] = 0.01$

n \ m	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	24	25	29	30	99	100
2	99,000	30,817	18,000	13,274	10,925	9,547	8,649	8,022	7,559	7,206	6,927	6,701	6,515	6,359	6,226	6,112	6,013	5,926	5,849	5,614	5,568	5,420	5,390	4,826	4,824
3	99,166	29,457	16,694	12,060	9,780	8,451	7,591	6,992	6,552	6,217	5,953	5,739	5,564	5,417	5,292	5,185	5,092	5,010	4,938	4,718	4,675	4,538	4,510	3,986	3,984
4	99,249	28,710	15,977	11,392	9,148	7,847	7,006	6,422	5,994	5,668	5,412	5,205	5,035	4,893	4,773	4,669	4,579	4,500	4,431	4,218	4,177	4,045	4,018	3,515	3,513
5	99,299	28,237	15,522	10,967	8,746	7,460	6,632	6,057	5,636	5,316	5,064	4,862	4,695	4,556	4,437	4,336	4,248	4,171	4,103	3,895	3,855	3,725	3,699	3,208	3,206
6	99,333	27,911	15,207	10,672	8,466	7,191	6,371	5,802	5,386	5,069	4,821	4,620	4,456	4,318	4,202	4,102	4,015	3,939	3,871	3,667	3,627	3,499	3,473	2,990	2,988
7	99,356	27,672	14,976	10,456	8,260	6,993	6,178	5,613	5,200	4,886	4,640	4,441	4,278	4,142	4,026	3,927	3,841	3,765	3,699	3,496	3,457	3,330	3,304	2,825	2,823
8	99,374	27,489	14,799	10,289	8,102	6,840	6,029	5,467	5,057	4,744	4,499	4,302	4,140	4,004	3,890	3,791	3,705	3,631	3,564	3,363	3,324	3,198	3,173	2,696	2,694
9	99,388	27,345	14,659	10,158	7,976	6,719	5,911	5,351	4,942	4,632	4,388	4,191	4,030	3,895	3,780	3,682	3,597	3,523	3,457	3,256	3,217	3,092	3,067	2,592	2,590
10	99,399	27,229	14,546	10,051	7,874	6,620	5,814	5,257	4,849	4,539	4,296	4,100	3,939	3,805	3,691	3,593	3,508	3,434	3,368	3,168	3,129	3,005	2,979	2,505	2,503
11	99,408	27,133	14,452	9,963	7,790	6,538	5,734	5,178	4,772	4,462	4,220	4,025	3,864	3,730	3,616	3,519	3,434	3,360	3,294	3,094	3,056	2,931	2,906	2,432	2,430
12	99,416	27,052	14,374	9,888	7,718	6,469	5,667	5,111	4,706	4,397	4,155	3,960	3,800	3,666	3,553	3,455	3,371	3,297	3,231	3,032	2,993	2,868	2,843	2,369	2,368
13	99,422	26,983	14,307	9,825	7,657	6,410	5,609	5,055	4,650	4,342	4,100	3,905	3,745	3,612	3,498	3,401	3,316	3,242	3,177	2,977	2,939	2,814	2,789	2,315	2,313
14	99,428	26,924	14,249	9,770	7,605	6,359	5,559	5,005	4,601	4,293	4,052	3,857	3,698	3,564	3,451	3,353	3,269	3,195	3,130	2,930	2,892	2,767	2,742	2,267	2,265
15	99,433	26,872	14,198	9,722	7,559	6,314	5,515	4,962	4,558	4,251	4,010	3,815	3,656	3,522	3,409	3,312	3,227	3,153	3,088	2,889	2,850	2,726	2,700	2,225	2,223
16	99,437	26,827	14,154	9,680	7,519	6,275	5,477	4,924	4,520	4,213	3,972	3,778	3,619	3,485	3,372	3,275	3,190	3,116	3,051	2,852	2,813	2,689	2,663	2,187	2,185
17	99,440	26,787	14,115	9,643	7,483	6,240	5,442	4,890	4,487	4,180	3,939	3,745	3,586	3,452	3,339	3,242	3,158	3,084	3,018	2,819	2,780	2,656	2,630	2,153	2,151
18	99,444	26,751	14,080	9,610	7,451	6,209	5,412	4,860	4,457	4,150	3,909	3,716	3,556	3,423	3,310	3,212	3,128	3,054	2,989	2,789	2,751	2,626	2,600	2,122	2,120
19	99,447	26,719	14,048	9,580	7,422	6,181	5,384	4,833	4,430	4,123	3,883	3,689	3,529	3,396	3,283	3,186	3,101	3,027	2,962	2,762	2,724	2,599	2,573	2,094	2,092
20	99,449	26,690	14,020	9,553	7,396	6,155	5,359	4,808	4,405	4,099	3,858	3,665	3,505	3,372	3,259	3,162	3,077	3,003	2,938	2,738	2,699	2,574	2,549	2,069	2,067
24	99,458	26,598	13,929	9,466	7,313	6,074	5,279	4,729	4,327	4,021	3,780	3,587	3,427	3,294	3,181	3,084	2,999	2,925	2,859	2,659	2,620	2,495	2,469	1,985	1,983
25	99,459	26,579	13,911	9,449	7,296	6,058	5,263	4,713	4,311	4,005	3,765	3,571	3,412	3,278	3,165	3,068	2,983	2,909	2,843	2,643	2,604	2,478	2,453	1,967	1,965
29	99,465	26,517	13,850	9,391	7,240	6,003	5,209	4,660	4,258	3,952	3,712	3,518	3,359	3,225	3,112	3,014	2,930	2,855	2,790	2,589	2,550	2,423	2,398	1,908	1,906
30	99,466	26,505	13,838	9,379	7,229	5,992	5,198	4,649	4,247	3,941	3,701	3,507	3,348	3,214	3,101	3,003	2,919	2,844	2,778	2,577	2,538	2,412	2,386	1,895	1,893
99	99,489	26,241	13,578	9,131	6,988	5,756	4,964	4,416	4,015	3,709	3,468	3,273	3,113	2,978	2,864	2,765	2,679	2,603	2,536	2,330	2,290	2,159	2,132	1,601	1,599
100	99,489	26,240	13,577	9,130	6,987	5,755	4,963	4,415	4,014	3,708	3,467	3,272	3,112	2,977	2,863	2,764	2,678	2,602	2,535	2,329	2,289	2,158	2,131	1,600	1,598