Universidad Autónoma del Estado de Hidalgo
Instituto de Ciencias Básicas e Ingeniería
Licenciatura en Física y Tecnología Avanzada
Academia de Matemáticas y Física

Estadística Descriptiva

Palomares Maldonado Héctor Miguel

Índice

1. Variables Aleatorias

1.1. Distribución Binomial

1. Un evento tiene 55 % de probabilidad de ocurrir cada vez que alguien entra a determinado hospital. ¿Cuál es la probabilidad de que dicho evento suceda en 2 de 5 personas ingresadas hoy?
   Solución :
   sea p la probabilidad que suceda $p = 55\,\% = 0{,}55$, q la probabilidad de que no suceda $q = 45\,\% = 0{,}45$, n el total de experimentos y x el total de éxitos
   $$\left(\begin{array}{c} n \\ x \end{array} \right) p^x q^{n-x}$$
   sustituyendo valores
   $$\left(\begin{array}{c} 5 \\ 2 \end{array} \right) (0{,}55)^2 (0{,}45)^{5-2}$$
   $$\frac{5!}{2!(5-2)!}(0{,}55)^2(0{,}45)^3$$
   $$\left(\frac{5*4}{2*1}\right)(0{,}3025)(0{,}091125) = 0{,}27565$$
   La probabilidad buscada es : 0,27565

2. Se lanza una moneda seis veces. Calcular la probabilidad de que en dos lanzamientos de los seis caiga águila
   Solución:
   En este caso tenemos probabilidad que suceda $p = 1/2 = 0{,}5$, q la probabilidad de que no suceda $q = 1/2 = 0{,}5$, el total de experimentos $n = 6$ y el total de éxitos $x = 2$
   $$\left(\begin{array}{c} n \\ x \end{array} \right) p^x q^{n-x}$$
   sustituyendo valores
   $$\left(\begin{array}{c} 6 \\ 2 \end{array} \right) (0{,}5)^2 (0{,}5)^{6-2}$$
   $$\frac{6!}{2!(6-2)!}(0{,}25)(0{,}0625)$$
   $$\frac{6*5*4*3*2*1}{2*1*4*3*2*1}(0{,}25)(0{,}0625)$$
   $$\frac{6*5}{2*1}(0{,}25)(0{,}0625) = \frac{15}{64} \approx 0{,}234375$$
   La probabilidad Buscada es : 0,234375

1.2. Distribución Geometrica

1. Calcular la probabilidad de que al lanzar una moneda salga águila en el sexto lanzamiento
   Solución:
   Tenemos:

   $$p(y) = pq^{x-1}$$

   $x = 6$ el sexto lanzamiento
   $p = 1/2 = 0{,}5$ la probabilidad de que salga águila
   $q = 1/2 = 0{,}5$ la probabilidad de que salga no águila

   $$P(X = 6) = (0{,}5)(0{,}5)^{6-1} = (0{,}5)(0{,}5)^5 = 0{,}0156$$

   La probabilidad buscada es : 0,0156

2. Se lanza un dado hasta que aparece el número 6. ¿Cuál es la probabilidad de que el número de lanzamientos sean 3?
   Solución.
   En este problema el éxito es la aparición del número 6 y la probabilidad de que salga el número 6 al lanzar un dado es 1/6, por lo que $p = 1/6$ y q = 5/6. Como nos interesa calcular la probabilidad de que el 6 aparezca en el tercer lanzamiento, entonces:

   $$P(X = 3) = \left(\frac{1}{6}\right)(0{,}5)^{3-1} = \left(\frac{1}{6}\right)(0{,}5)^2 = 0{,}01157$$

   La probabilidad buscada es : 0,1157

3. Tres personas lanzan una moneda y el disparejo paga el café. Si los tres resultados son iguales, las monedas se lanzan nuevamente. Encontrar la probabilidad de que se necesiten menos de 4 intentos para saber quien paga el café.
   Solución.

   En este problema el éxito consiste en sacar el disparejo. Lo primero que debemos hacer para resolver el problema, es encontrar el espacio muestral correspondiente al lanzamiento de 3 monedas:

   sea c cara y + cruz

   $$S = \{(c, c, c)(c, c, +)(c, +, c)(+, c, c)(c, +, +)(+, c, +)(+, +, c)(+, +, +)\}$$

   El número de resultados en que aparece el disparejo es 6, por lo que $p = 6/8 = 0{,}75$ y $q = 0{,}25$. ya que la probabilidad es 1, $0{,}75 + x = 1$, Si queremos obtener la probabilidad de que se necesiten menos de 4 intentos para saber quien paga el café, entonces:

   $$P(X < 4) = P(X = 1) + P(X = 2) + P(X = 3)$$
   $$P(X < 4) = (0{,}75)(0{,}25)^{1-0} + (0{,}75)(0{,}25)^{2-1} + (0{,}75)(0{,}25)^{3-1} = 0{,}9844$$

   La probabilidad buscada es : 0,9844 0 98,44 %

1.3. Distribución Hipergeometrica

1. En una tienda de computo hay diez computadoras de las cuales tres están dañadas, si cinco clientes desean adquirir una computadora ¿cual es la probabilidad de tener el equipo dañado?
Solución:

$$p(X) = \frac{\begin{pmatrix} N-D \\ n-x \end{pmatrix}\begin{pmatrix} D \\ x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

donde:
$\begin{pmatrix} N-D \\ n-x \end{pmatrix}\begin{pmatrix} D \\ x \end{pmatrix}$ casos de éxito
$\begin{pmatrix} N \\ n \end{pmatrix}$ casos posibles

Tenemos:
$N = 10$ Numero de la población (total de equipos de computo)
$D = 3$ Numero de éxitos, equipos dañados, ya que queremos la probabilidad de estos
$n = 5$ tamaño de la muestra
$x = 1$ cada persona solo desea comprar una computadora
sustituyendo los valores en la función de probabilidad

$$p(X=1) = \frac{\begin{pmatrix} 10-3 \\ 5-1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix}}{\begin{pmatrix} 10 \\ 5 \end{pmatrix}}$$

$$p(X=1) = \frac{\dfrac{7!}{4!(7-4)!}\cdot\dfrac{3!}{(3-1)!}}{\dfrac{10!}{5!(10-5)!}}$$

$$p(X=1) = \frac{\dfrac{7*6*5}{2*1}}{\dfrac{10*9*8*7*6}{5*4*3*2*1}}$$

$$p(X=1) = \frac{105}{252} = \frac{5}{12} \approx 0{,}41\bar{6}$$

La probabilidad buscada es : 0,4166 o 41,66 %

2. Héctor llega a comprar un vaso con café camino a la universidad, tiene 20 opciones posibles de comercios para llegar a comprarlo, el precio medio del café es de \$8,50. Esta semana deberá de escoger 3 de estos comercios y cuando mucho debe gastar \$10 por café. sabe que un cuarto de dichos comercios rebasan los \$10 en el costo del café. Suponga que selecciona 3 de estos comercios. Calcule la probabilidad de que uno de los negocios rebase el costo

Solución:

$$p(X) = \frac{\begin{pmatrix} N-D \\ n-x \end{pmatrix}\begin{pmatrix} D \\ x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

donde:
$\begin{pmatrix} N-D \\ n-x \end{pmatrix}\begin{pmatrix} D \\ x \end{pmatrix}$ casos de éxito
$\begin{pmatrix} N \\ n \end{pmatrix}$ casos posibles

Tenemos:
$N = 20,\ D = 5,\ n = 3,\ x = 1$
sustituyendo los valores en la función de probabilidad

$$p(X) = \frac{\begin{pmatrix} 20-5 \\ 3-1 \end{pmatrix}\begin{pmatrix} 5 \\ 1 \end{pmatrix}}{\begin{pmatrix} 20 \\ 3 \end{pmatrix}}$$

$$p(X) = \frac{\dfrac{15!}{2!(15-2)!} \cdot \dfrac{5!}{(5-1)!}}{\dfrac{20!}{3!(20-3)!}} = \frac{\dfrac{15*14}{2*1} \cdot 5}{\dfrac{20*19*18}{3*2*1}} = \frac{5*105}{1140} = 0{,}4605$$

La probabilidad buscada es : 0,4605 o 46,05 %

1.4. Distribución de Poisson

1. Un cajero automático es utilizado cada 20 minutos por seis personas. Se desea saber cual es la probabilidad :

   a) que cajero sea utilizado por 5 personas en 20 minutos
   b) que el cajero sea utilizado por 10 personas en 20 minutos
   c) que el cajero sea utilizado por menos de 5 personas en 20 minutos

   Solución:
   la distribución de poisson esta dada por:

   $$p(x, \lambda) = \frac{\lambda^x e^\lambda}{x!}$$

   a)
   $$p(x = 5) = (5, 6) = \frac{6^5 e^{-6}}{5!}$$
   $$p(x = 5) = \frac{(7776)(0{,}0025)}{120}$$
   $$p(x = 5) = 0{,}162$$

   la probabilidad es de 0,162 o 16,2 %

   b)
   $$p(x = 10) = (10, 6) = \frac{6^{10} e^{-6}}{10!}$$
   $$p(x = 10) = \frac{(60, 466, 176)(0{,}0025)}{3628800}$$
   $$p(x = 5) = 0{,}0416$$

   la probabilidad es de 0,0416 o 4,16 %

   c)
   $$p(x \le 5) = p(x = 0) + p(x = 1) + p(x = 2) + p(x = 3) + p(x = 4) + p(x = 5)$$
   $$p(x \le 5) = \frac{6^0 e^{-6}}{0!} + \frac{6^1 e^{-6}}{1!} + \frac{6^2 e^{-6}}{2!} + \frac{6^3 e^{-6}}{3!} + \frac{6^4 e^{-6}}{4!} + \frac{6^5 e^{-6}}{5!}$$
   $$p(x \le 5) = 0{,}4483$$

   la probabilidad es de 0,4483 o 44,83 %

2. Los protones del Gran Colisionador recorren 27 km en un milisegundo en Ginebra, Si colisionan 4 protones por milisegundo. calcular la probabilidad de

   a) Tener 2 colisiones en un milisegundo

b) No obtener Ninguna Colisión

c) Obtener menos de 3 colisiones

a) Obtener mas de 3 colisiones

Solución:
De la distribución de Poisson Tenemos que la probabilidad esta dada por:

$$p(x, \lambda) = \frac{\lambda^x e^\lambda}{x!}$$

donde $\lambda = 4$colisiones/milisegundo

a)
$$p(x = 2) = \frac{4^2 e^{-4}}{2!}$$
$$p(x = 2) = \frac{16(0{,}01831)}{2} = 0{,}1465$$

La probabilidad para este inciso es de 0,1465 ó 14,65 %

b)
$$p(x = 0) = \frac{4^0 e^{-4}}{0!}$$
$$p(x = 0) = \frac{1(0{,}01831)}{1} = 0{,}01831$$

La probabilidad es de 0,01831 o 1,83 %

c)
$$p(x < 3) = p(x = 0) + p(x = 1) + p(x = 2)$$
$$p(x = 1) = \frac{4(e^{-4})}{1!} = 0{,}0732$$
$$0{,}01831 + 0{,}0732 + 0{,}1465 = 0{,}2380$$

La probabilidad

d) La probabilidad es 1 o del 100 %

1.5. Distibucion Normal

1. Una Ferreteria vende tubos de aluminio, la longitud de cada tuvo se distribuye de forma normal con media $15cm$ y una varianza de 2,25. Determine

   a La probabilidad de que una pieza exceda $18cm$

   b) La probabilidad de que los tubos estén entre $13cm$ y $17cm$

   Solución:

$$z = \frac{x - \mu}{\sigma}$$

donde σ es la desviación estándar

$$\sigma^2 = 2{,}25 \qquad \sigma = 1{,}5$$

$$Z = \frac{18cm - 15cm}{1{,}5cm} = 2$$

$$2{,}0 \rightarrow 0{,}4772$$

La distribución de probabilidad es:

$$0{,}5 - 4{,}772$$

$$0{,}0228$$

ya que la probabilidad es 1, para este caso solo nos fijamos en la mitad de la gráfica (campana de Gauss)

La probabilidad buscada es de 0,0228 ó 2,28 %

b)

$$Z = \frac{13cm - 15cm}{1{,}5cm} = -1{,}3\bar{3}$$

$$Z = \frac{17cm - 15cm}{1{,}5cm} = 1{,}3\bar{3}$$

$$1{,}33 \rightarrow 0{,}4082$$

la probabilidad entre que el tuvo esté entre $13cm$ y $17cm$ es $0{,}4032 * (2) = 0{,}8164$ ó 81,64 %

1.6. ejercicios

1. Se sabe que aproximadamente el 60 % de los hogares tiene dos o más televisores. Suponga que se lleva a cabo un muestreo de 15 casas y que x es el número de casas que tienen dos o más televisores

   a) ¿Cuál es la distribución de probabilidades para x?

   b) Encuentre $P(x \leq 8)$

   c) ¿Cuál es el valor más grande de c para el que $P(x \leq c) \leq 0.1$?

   Solución

   a)

   $$P(X)\begin{pmatrix} n \\ x \end{pmatrix} p^x q^{n-x}$$

   sustituyendo valores

   $$P(X) = \begin{pmatrix} 15 \\ x \end{pmatrix} = (0,6)^x (0{,}4)^{15-x}$$

   b)

   $$P(x \leq 8) = P(x=0)+P(x=1)+P(x=2)+P(x=3)+P(x=4)+P(x=5)+P(x=6)+P(x=7)+P(x=8)$$

   $$P(x \leq 8) = \frac{15!}{14!}(0,6)(0{,}4)^{14} + \frac{15!}{2!(13)!}(0,6)^2(0{,}4)^{13} + \frac{15!}{3!(12)!}(0,6)^3(0{,}4)^{12} + \frac{15!}{4!(11)!}(0,6)^4(0{,}4)^{11}$$

   $$+\frac{15!}{5!(10)!}(0,6)^5(0{,}4)^{10} + \frac{15!}{6!(9)!}(0,6)^6(0{,}4)^9 + \frac{15!}{7!(8)!}(0,6)^7(0{,}4)^8 + \frac{15!}{8!(7)!}(0,6)^8(0{,}4)^7$$

   $$P(x \leq 8) = 0{,}000024 + 0{,}000225 + 0{,}001648 + 0{,}007419 + 0{,}024485 + 0{,}061214 = 0{,}0950$$

   $$P(x \leq 8) = 0{,}0950$$

   c)

   $$p(x \leq c) = \leq 0{,}1$$

   $$\begin{pmatrix} 15 \\ c \end{pmatrix} = (0,6)^c (0{,}4)^{15-c} \leq 0{,}1$$

   $$\frac{15!}{c!(15-c)!}(0,6)^c (0{,}4)^{n-c}$$

2. Para verificar la exactitud de un pronosticador del tiempo, se revisaron los registros sólo para los días en que el pronosticador predijo lluvia "con de probabilidad 0.3". Una revisión de 25 de esos días indicó que llovió en 10 de los 25.

   a) ¿Cuál es el valor apropiado de la probabilidad de lluvia en uno de esos 25 días?

   b) ¿Cuáles son la media y la desviación estándar de x, el número de días que llovió?

   c) Calcule la puntuación z para el valor observado $x = 10$ ¿Discrepan estos datos con la previsión de lluvia con probabilidad de 0.3?

   Solución:

a) solución por el método de distribución binomial

$$p(x=1) = \left(\begin{array}{c} 15 \\ 1 \end{array} \right) (0{,}3)(0{,}7)^{15-1}$$

$$p(x=1) = \frac{15!}{14!}(0{,}3)(0{,}17)^{14} = (15)(0{,}3)(0{,}0067)$$

$$p(x=1) = 0{,}030519$$

La probabilidad de que llueva un dia es de 3,05 %

b) La media o valor esperado es

$$\mu(x) = \sum_{i=1}^{10} x_i p(x_i)$$

$$\mu = \frac{15!}{15!}(0{,}3)^0(0{,}7)^{15-0} + \frac{15!}{14!}(0{,}3)(0{,}7)^{15-1} + \frac{15!}{(2!)(13!)}(0{,}3)^2(0{,}7)^{15-2} + \frac{15!}{(3!)(12!)}(0{,}3)^3(0{,}7)^{15-3}$$

$$+\frac{15!}{(4!)(11!)}(0{,}3)^4(0{,}7)^{15-4}+\frac{15!}{(5!)(10!)}(0{,}3)^5(0{,}7)^{15-5}+\frac{15!}{(6!)(9!)}(0{,}3)^6(0{,}7)^{15-6}+\frac{15!}{(7!)(8!)}(0{,}3)^7(0{,}7)^{15-7}$$

$$+\frac{15!}{(8!)(7!)}(0{,}3)^8(0{,}7)^{15-8} + \frac{15!}{(!9)(6!)}(0{,}3)^9(0{,}7)^{15-9} + \frac{15!}{(10!)(5!)}(0{,}3)^{10}(0{,}7)^{15-10}$$

$$\mu(x) = 0{,}0047(0) + 0{,}0305(1) + 0{,}0915(2) + 0{,}1700(3) + 0{,}2186(4) + 0{,}2061(5)$$

$$+0{,}1472(6) + 0{,}0811(7) + 0{,}0347(8) + 0{,}0115(9) + 0{,}0029(10)$$

$$\mu(x) = 4{,}4894$$

Calculando $\mu(x^2)$

$$\mu(x^2) = 0{,}0305(1^2) + 0{,}0915(2^2) + 0{,}1700(3^2) + 0{,}2186(4^2) + 0{,}2061(5^2)$$

$$+0{,}1472(6^2) + 0{,}0811(7^2) + 0{,}0347(8^2) + 0{,}0115(9^2) + 0{,}0029(10^2)$$

$$\mu(x^2) = 23{,}2920$$

el valor esperado es:

$$Var(x) = \mu(x^2) - [\mu(x)]^1 = 23{,}2920 - 4{,}4894 = 3{,}13728764$$

La desviación es:

$$\sigma = \sqrt{Var(x)} = 1{,}771239013$$

c)

3. Una pieza de equipo electrónico contiene seis chips, de los cuales dos son defectuosos. Se eligen al azar tres chips para inspeccionarlos y se anota cuántos están defectuosos. Encuentre la distribución de probabilidad para x, el número de chips defectuosos.
Solución:
Solución por distribución Hipergeometrica

$$p(X) = \frac{\binom{N-D}{n-x}\binom{D}{x}}{\binom{N}{n}}$$

Tenemos:
$N = 6$ Numero total de chips
$D = 2$ numero de equipos defectuoso
$n = 3$ tamaño de la muestra
sustituyendo los valores correspondientes

$$p(x) = p(x=0) + p(x=1) + p(x=2)$$

$$p(X) = \frac{\binom{6-2}{3-0}\binom{2}{0}}{\binom{6}{3}} + \frac{\binom{6-2}{3-1}\binom{2}{1}}{\binom{6}{3}} + \frac{\binom{6-2}{3-2}\binom{2}{2}}{\binom{6}{3}}$$

$$p(x) = \frac{\frac{4!}{2!(4-2)}}{\frac{20!}{3!(20-3)!}} + \frac{\frac{4!}{2!(4-2)!} * 5}{\frac{20!}{3!(20-3)!}} + \frac{\frac{4!}{2!(4-2)!} * \frac{3!}{2!(3-2)!}}{\frac{20!}{3!(20-3)!}}$$

$$p(x) = 0{,}07142857 + 0{,}42857143 + 0{,}42857143$$

4. A un museo llegan 4 autobuses con 148 niños, distribuidos de la siguiente manera: 40, 33, 25 y 50. Se selecciona al azar a un niño y sea X el número de niños en el autobús donde iba el niño seleccionado. También se selecciona al azar a uno de los conductores y sea Y el número de niños en el autobús que conducía. ¿Cuál piensa que será mayor entre $E[X]$ y $E[Y]$?
Solución:

5. El director del CIMA entrevista a 11 estudiantes para ocupar cuatro vacantes de ayudante de profesor y ha programado seis entrevistas para el primer día y 5 para el segundo. Suponga que los estudiantes son entrevistados al azar. ¿Cuál es la probabilidad de que x de los mejores cuatro candidatos sean entrevistados el primer día? ¿Cuántos de los mejores candidatos pueden esperar ser entrevistados el primer día?

Solución : para esté problema utilizamos la distribución hipergeometrica

$$p(x) = \frac{\begin{pmatrix} N-D \\ n-x \end{pmatrix}\begin{pmatrix} D \\ x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

Tenemos:
$N = 11$ e numero total de alumnos, $D = 4$ Numero de éxitos, es decir los 4 alumnos seleccionados para el primer día, $n = 6$ y para el segundo n tamaño de la muestra, $x = 1$ ya que son entrevistados individualmente
sustituyendo los valores en la función de probabilidad para el primer día

$$p(X=1) = \frac{\begin{pmatrix} 11-4 \\ 6-1 \end{pmatrix}\begin{pmatrix} 4 \\ 1 \end{pmatrix}}{\begin{pmatrix} 11 \\ 6 \end{pmatrix}}$$

$$p(X=1) = \frac{\dfrac{7!}{5!(7-5)!}\cdot\dfrac{4!}{(4-1)!}}{\dfrac{11!}{6!(11-6)!}} = 0{,}1818$$

La probabilidad buscada es : 0,1818 o 18,18 %

x	0	1	2	3	4
$p(x)$	0.1515152	0.18181818	0.45454545	0.30303030	0.04545455

Probabilidad Acumulada

x	0	1	2	3	4
$p(x)$	0.01515152	0.19696970	0.65151515	0.95454545	1.00000000

para el segundo dia

$$p(X=1) = \frac{\begin{pmatrix} 11-4 \\ 5-1 \end{pmatrix}\begin{pmatrix} 4 \\ 1 \end{pmatrix}}{\begin{pmatrix} 11 \\ 5 \end{pmatrix}}$$

$$p(X=1) = \frac{\dfrac{7!}{5!(7-5)!}\cdot\dfrac{5!}{(5-1)!}}{\dfrac{11!}{5!(11-5)!}} = 0{,}0{,}30303030$$

La probabilidad buscada es : 0,30303030 o 30,30 %

x	0	1	2	3	4
$p(x)$	0.04545455	0.30303030	0.45454545	0.18181818	0.01515152

Probabilidad Acumulada

x	0	1	2	3	4
$p(x)$	0.04545455	0.34848485	0.80303030	0.98484848	1.00000000

6. El número x de personas que entran en una unidad de terapia intensiva en un hospital tiene una distribución de probabilidad de *Poisson* con media igual a cinco personas por día.

   a) ¿Cuál es la probabilidad de que el número de personas que entran a la unidad de terapia intensiva en un día sea dos?

   b) ¿Cuál es la probabilidad de que x sea mayor que diez?

   Solución:
   sea λ el parametro 5personnas/dia, para el inciso a $x = 2$

   a) La distribución de poisson se determina

   $$p(x = 2) = \frac{\lambda^x e^\lambda}{x!}$$

   sustituyendo valores

   $$p(x = 2) = \frac{5^2 e^{-5}}{2!}$$

   $$p(x = 2) = \frac{(25)(0{,}006737946999)}{2} =$$

   $$p(x = 2) = 0{,}08422433749$$

   la probabilidad del número de personas es 0,0842 ó 8,43 %

   b)

   $$p(x \leq 10) = p(x = 0) + p(x = 1) + p(x = 2) + p(x = 3) + p(x = 4) + p(x = 5) + p(x = 6)$$

   $$+p(x = 7) + p(x = 8) + p(x = 9) + p(x = 10)$$

   $p(x \leq 10) = 0{,}006737947 + 0{,}033689735 + 0{,}084224337 + 0{,}140373896 + 0{,}1754673700{,}175467370 + 0{,}146222808 + 0{,}104444863 + 0{,}065278039 + 0{,}036265577$

   $$p(x \leq 10) = 0{,}986304731$$

probabilidad de que x sea mayor que diez es $x = 0{,}986304731$ ó $x = 98{,}63\,\%$

7. La cantidad de casos de sarampión informados en el estado de Hidalgo durante el mes de mayo de 2004, tiene una distribución de *Poisson* con parámetro $\mu = 4.5$. ¿Cuál es la probabilidad de que se informen cuatro casos o más durante un mes?
   Solución:
   Tenemos la distribución de possion

   $$p(x) = \frac{\mu^x e^\mu}{x!}$$

   donde el parámetro es $\mu = 4{,}5$ y $x \leq 5$ de cuatro a 30 personas se obtienen las siguientes probabilidades

   $p(x \leq 4) = 1{,}898076 \times 10^{-1} + 1{,}708269 \times 10^{-1} + 1{,}281201 \times 10^{-1} + 8{,}236295^{-2} + 4{,}632916 \times 10^{-2} + 2{,}316458 \times 10^{-2} + 1{,}042406 \times 10^{-2} + 4{,}264389 \times 10^{-3} + 1{,}599146 \times 10^{-3} + 5{,}535504 \times 10^{-4} + 1{,}779269 \times 10^{-4} + 5{,}337808 \times 10^{-5} + 1{,}501258 \times 10^{-5} 3{,}973919 \times 10^{-6} + 9{,}934798 \times 10^{-7} + 2{,}352979 \times 10^{-7} + 5{,}294202 \times 10^{-8} + 1{,}134472 \times 10^{-8} + 2{,}320511 \times 10^{-9} + 4{,}540129 \times 10^{-10} + 8{,}512742 \times 10^{-11} + 1{,}532294 \times 10^{-11} + 2{,}652047 \times 10^{-12} + 4{,}420078 \times 10^{-12} + 7{,}103696 \times 10^{-14} + 1{,}102298 \times 10^{-14} + 1{,}653447 \times 10^{-15}$

   La probabilidad acumulada de 4 a 30 personas es la siguiente

0,5321036	0.7029304	0.8310506	0.9134135	0.9597427	0.9829073
0,9933313	0.9975957	0.9991949	0.9997484	0.9999263	0.9999797
0,9999947	0.9999987	0.9999997	0.9999999	1.0000000	1.0000000
1,0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000
1,0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

8. La estatura humana es una de las muchas variables aleatorias biológicas que tienen una distribución *normal*. Suponga que la estatura de los hombres tiene una media de 175.26 centímetros y una desviación estándar de 8.89 centímetros.

   a) ¿Qué porcentaje de hombres será más alto que 182.88 centímetros?
   Solución:
   Sabemos que la distribución Normal esta dada

$$p(x) = \frac{x-\mu}{\sigma}$$

   Donde $\mu = 175{,}26$ es la media y σ es la desviación estándar, sustituyendo datos Obtenemos

$$p(x = 182{,}88) = \frac{182{,}88cm - 175{,}26cm}{8{,}89cm} = 0{,}8571428571$$
$$p(x) = 0{,}8571428571 \rightarrow 0{,}03107932$$
$$0{,}5 - 0{,}3107932 = 0{,}46892068$$

   El porcentaje es de 46,98 %

9. Suponga que el número de veces que un humano adulto respira, por minuto y en reposo, tiene una distribución *normal*, con media igual a 16 y desviación estándar igual a 4. Si se elige al azar una persona, ¿cuál es la probabilidad de que su número de respiraciones por minuto sea mayor a 22?

   Solución
   Sabemos que la distribución Normal esta dada

$$p(x) = \frac{x-\mu}{\sigma}$$

   Donde $\mu = 175{,}26$ es la media y σ es la desviación estándar, sustituyendo datos Obtenemos

$$p(x = 182{,}88) = \frac{22 - 16}{4} = 1{,}5$$
$$p(x) = 1{,}5 \rightarrow 0{,}0323794$$
$$0{,}5 - 0{,}0323794 = 0{,}4676206$$

   La Probabilidad es 46,76 %

10. Sea X una variable aleatoria que describe el tiempo entre dos llegadas sucesivas a la ventanilla de atención de un banco. Si X tiene una distribución *exponencial* con $\lambda = 1$, calcule el tiempo esperado entre dos llegadas sucesivas. ¿Cual es el valor de $P(X \leq 4)$? ¿Cuál es el valor de $P(2 \leq X \leq 5)$?
    Solución:
    La distribución exponencial esta dada:

$$p(x) = \lambda \int e^{-\lambda x}$$

    sustituyendo valores, para $x \leq 4$

$$p(x \leq 4) = \lambda \int_4^\infty e^{-\lambda x} dx$$

$$p(x \leq 4) = \lambda \frac{e^{-\lambda x}}{-\lambda}$$
$$p(x \leq 4) = e^{-4}$$
$$p(x \leq 4) = 0{,}01831564$$

    para $P(2 \leq X \leq 5) = p(x \leq 5) - p(\leq 2)$ —

$$p(2 \leq x \leq 4) = \lambda \int_2^5 e^{-\lambda x} dx$$

$$p(2 \leq x \leq 4) = \lambda \frac{e^{-\lambda x}}{-\lambda}\Big|_2^5$$

$$p(2 \leq x \leq 4) = e^{-5} - e^{-2}$$

$$P(2 \leq X \leq 5) = 0{,}1285973$$

11. El tiempo que tarda un bibliotecario para localizar una ficha de la base de datos sobre libros prestados, tiene una distribución *exponencial* con un tiempo esperado de 20 segundos. ¿Cuál es el valor de $P(X \leq 30)$? ¿Cuál es el valor de $P(X \geq 20)$ ¿Cuál es el valor de $P(20 \leq X \leq 30)$?
Solución:

Valor de $p(x \leq 30)$

La distribución exponencial esta dada:

$$p(x) = \lambda \int e^{-\lambda x}$$

Valor de $p(x \leq 30)$

$$p(x \leq 30) = \lambda \int_0^{30} e^{-\lambda x} dx$$

$$p(x \leq 30) = \lambda \frac{e^{-\lambda x}}{-\lambda}$$

$$p(x \leq 30) = \frac{1}{20} \frac{e^{-(1/20)x}}{-1/20)}\Big|_0^{30}$$

$$p(x \leq 30) = 0{,}7768698$$

Valor de $p(x \leq 20)$

$$p(x \leq 20) = \lambda \int_0^{20} e^{-\lambda x} dx$$

$$p(x \leq 20) = \lambda \frac{e^{-\lambda x}}{-\lambda}$$

$$p(x \leq 20) = \frac{1}{20} \frac{e^{-\lambda x}}{-1/20}\Big|_0^{20}$$

$$p(x \leq 20) = 0{,}9999546$$

2. Estimadores

2.1. ejercicios

1. Sea $X_1, X_2, \ldots, X_n$ una muestra proveniente de la distribución cuya función de densidad es

$$f(x)\{ \begin{array}{cc} \mathrm{e}^{-(x-\theta)} & \mathrm{x} \leq 0 \\ 0 & enotrocaso \end{array}$$

Determine el estimador de máxima verosimilitud para θ.
solución:

$$\begin{aligned} e^{-(x_1-\theta)}e^{-(x_2-\theta)}\ldots e^{-(x_n-\theta)}) = e^{\sum x_i}e^{n\theta} \\ \ln(f(x)) = e^{\sum x_i}e^{n\theta} \\ = \ln e^{\sum x_i} + \ln e^{n\theta} \\ = \sum x_i \ln e^{+} n\theta \ln e \\ = \sum x_i + \theta \end{aligned}$$

por otra parte:

$$\sum x = n\bar{x} = n\mu$$

y

$$\mu(x) = \int x f(x) dx$$

$$\mu(x) = \int x e^{-(x-\theta)} dx$$

sea $u = -x + \theta$

$$\mu(x) = -\int (u-\theta)e^u du$$

$$\mu(x) = \theta \int e^u du - \int u e^u du$$

$$\mu(x) = \theta e^u - \left[u e^u - \int e^u du \right]$$

$$\mu(x) = \theta e^u - [u e^u - e^u]$$

$$\mu(x) = \theta e^u - [e^u(u-1)]$$

$$\mu(x) = \theta e^u + [e^u(1-u)$$

$$\mu(x) = \theta e^{-(x-\theta)} + e^{-(x-\theta)}\left[1 + (x-\theta)\right]$$

$$\frac{d \ln(\theta e^{-(x-\theta)} + e^{-(x-\theta)}\left[1 + (x-\theta)\right] + n\theta)}{d\theta} = 0$$

$$\frac{1}{\theta e^{-(x-\theta)} + e^{-(x-\theta)}\left[1 + (x-\theta)\right] + n\theta)} \left\{ [e^{-(1-\theta)} + \theta e^{-(x-\theta)}] + [e^{-(x-\theta)}(1 + x - \theta) - e^{-(x-\theta)} + n] \right\} = 0$$

$$\theta e^{-(x-\theta)} + e^{-(x-\theta)}\Big[1 + (x-\theta) + n\theta = e^{-(x-\theta)} + \theta e^{-(x-\theta)} + e^{-(x-\theta)}(1 + x - \theta) - e^{-(x-\theta)} + n$$

$$\theta e^{-(x-\theta)} + n\theta = e^{-(x-\theta)} + \theta e^{(x-\theta)} - e^{-(x-\theta)} + n$$

$$n\theta = n$$

$$n(\theta - 1) = 0$$

2. Si una variable aleatoria tiene por función de densidad

$$f(x:a) = \frac{2a}{1-a}x^{\frac{3a-1}{1-a}}$$

$0 \leq x \leq 1$; $a > 0$ hállese el estimador del parámetro a por el método de máxima verosimilitud, en muestras aleatorias simples de tamaño n

Solución:

La función de verosimilitud es:

$$L = K(x_1, x_2, x_3 \dots x_n) = \frac{2a}{1-a}x_1^{\frac{3a-1}{1-a}}, \frac{2a}{1-a}x_2^{\frac{3a-1}{1-a}}, \dots \frac{2a}{1-a}x_n^{\frac{3a-1}{1-a}}$$

$$L = \frac{2a}{1-a}(x_1 x_2 \dots x_n)^{\frac{3a-1}{1-a}}$$

tomando el logaritmo

$$\ln L = n \ln 2 + n \ln a - \ln(1-a) + \frac{3a-1}{1-a}\sum_{i=1}^{n} \ln x_i$$

Calculamos la derivada parcial de ln L respecto a a y la igualamos a cero, resolviendo la ecuación:

$$\frac{\partial \ln L}{\partial a} = 0$$

$$\frac{n}{a} + \frac{n}{1-a} + \frac{2}{(1-a)^2} \ln \sum_{i=0}^{n} x_i = 0$$

$$\frac{n(1-a) + na}{a(1-a)} + \frac{2}{(1-a)^2}\sum_{i=0}^{n} \ln x_i = 0$$

$$\frac{2}{(1-a)^2}\sum_{i=0} \ln x_i = -\frac{n}{a(1-a)}$$

$$\sum_{i=0}^{n} \ln x_i = -\frac{n(1-a)}{2a}$$

$$\sum_{i=0}^{n} \ln x_i = \frac{n(a-1)}{2a}$$

$$\sum_{i=0}^{n} \ln x_i = \frac{n}{2} - \frac{n}{2a}$$

$$\frac{n}{2a} = \frac{n}{2} - \sum_{i=0}^{n} \ln x_i$$

$$a = \frac{n}{n - 2\sum_{i=0}^{n} \ln x_i}$$

3. Una firma comercial encuesta a 100 individuos para conocer sus opiniones sobre la elección de dos productos alternativos A y B recientemente fabricados. El resultado de la encuesta arroja que el producto A lo han elegido 55 individuos y el producto B 45. Hallar un intervalo de confianza al 95 % para la proporción de individuos que eligen cada producto.
Solución:
Tenemos el siguiente estadístico

$$\frac{\hat{p}-p}{\sqrt{\dfrac{\hat{p}(1-p)}{n}}} \to N(0,1)$$

lo que nos lleva al intervalo de confianza para p definido por:

$$\hat{p} - Z_{a/2}\sqrt{\frac{\hat{p}(1-p)}{n}} \leq p \leq \hat{p} + Z_{a/2}\sqrt{\frac{\hat{p}(1-p)}{n}}$$

En nuestro caso, para el producto A tenemos el intervalo

$$\left[\frac{55}{100} - 1{,}96\sqrt{\frac{(0{,}45)(0{,}55)}{100}}, \frac{55}{100} + 1{,}96\sqrt{\frac{(0{,}45)(0{,}55)}{100}}\right] = [0{,}45, 0{,}65]$$

Para el producto B Tenemos el intervalo

$$\left[\frac{45}{100} - 1{,}96\sqrt{\frac{(0{,}55)(0{,}45)}{100}}, \frac{45}{100} + 1{,}96\sqrt{\frac{(0{,}55)(0{,}45)}{100}}\right] = [][0{,}35, 0{,}55]$$

Como conclusión podemos decir que hay una probabilidad del 95 % entre 0,45 y 0,65 de que el producto elegido sea el A, y hay una probabilidad entre 0,35 y 0,55 de que el producto elegido sea el B.

4. En una población de tamaño 64 se estudia una característica X medida sobre sus individuos de la que se sabe que su media es 1012 y su desviación típica es 25. Hallar intervalos de confianza para el valor medio de la característica X con coeficientes de confianza del 90 % y 95 %.
Solución:
En esta situación sabemos:

$$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \to N(1,0)$$

lo que nos lleva al intervalo de confianza para la media

$$\bar{x} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

para un coeficiente de confianza del 90 % se tiene

$$1-\alpha = 0{,}9$$

$$\alpha = 0{,}1$$

$$\frac{\alpha}{2} = \frac{0{,}1}{2} = 0,05$$

$$Z_{\alpha/2} = 1{,}645$$

el intervalo de confianza es

$$1012 \pm 1{,}645\frac{25}{\sqrt{1012}}$$

$$\left[1012 + 1{,}645\frac{25}{\sqrt{1012}}, 1012 - 1{,}645\frac{25}{\sqrt{1012}} = [1006{,}86, 1017{,}14]\right.$$

Podemos concluir entonces que hay una probabilidad del 90 % de que el valor medio de la característica esté entre 1006, 86 y 1017, 14. Para un coeficiente de confianza del 95 % se tiene que:

$$1-\alpha = 0{,}95$$

$$\alpha = 0{,}05$$
$$\frac{\alpha}{2} = 0{,}025$$
$$Z_{\alpha/2} = 1{,}96$$

El Intervalo de confianza es

$$1012 \pm 1{,}96 \frac{25}{\sqrt{1012}}$$
$$\left[1012 + 1{,}96 \frac{25}{\sqrt{1012}}, 1012 - 1{,}96 \frac{25}{\sqrt{1012}}\right] = [1005{,}875, 1018{,}125]$$

5. Se sabe que la longitud de los diámetros de los tornillos fabricados por una máquina siguen una distribución normal y se busca un intervalo en el cual se encuentre la variabilidad de las longitudes de los tornillos fabricados por la máquina con una probabilidad del 80 %. Construir dicho intervalo sabiendo que una muestra de 16 tornillos presenta una variabilidad cuantificada en 30.
   Solución:
   En este caso el intervalo de confianza para la varianza se basa en el siguiente estadístico:

$$\frac{(n-1)S^2}{\sigma^2} \to X^2_{n-1}$$

   lo que nos lleva al intervalo de confianza para la varianza definida por:

$$\left(\frac{(n-1)S^2}{X^2_{\alpha/2,n}}, \frac{(n-1)S^2}{X^2_{1-\alpha/2,n}}\right) = \left(\frac{n\sigma^2}{X^2_{\alpha/2,n}}, \frac{n\sigma^2}{X^2_{1-\alpha/2,n}}\right)$$

   para el 90 %

$$1 - \alpha = 0{,}8$$
$$\alpha = 0{,}2$$
$$\frac{\alpha}{2} = 0{,}1$$
$$X^2_{0,1,15} = 22{,}3$$
$$X^2_{0,9,15} = 8{,}55$$

   El intervalo de confianza es

$$= \left[\frac{(16)(30)}{22{,}3}, \frac{(16)(30)}{8{,}55}\right]$$
$$= (21{,}43, 56{,}14)$$

6. Los responsables municipales de la salud miden la radiactividad en el agua de una fuente natural en una zona abundante en granito. Realizadas 12 mediciones en diferentes fechas del año se observó una media de 3,6 picocurios con una desviación típica de 0,82. determinar, al 95 % y al 99 %, intervalos de confianza para la radiación media y para la varianza.
   Solución:
   Se tiene los siguientes datos

$$N = 12 \qquad \hat{x} = 3{,}6 \qquad s = 0{,}82$$

   Para determinar el intervalo de confianza para σ^2 tenemos en cuenta que

$$\frac{12S^2}{\sigma^2}$$

   Es una variable x^2 con 11 grados de libertad. Los percentiles necesarios para la determinación de los intervalos al 95 % y 99 % de confianza son

$$x^2_{11,0,005} = 2{,}603$$
$$x^2_{11,0,025} = 3{,}816$$

$$x^2_{11,0{,}975} = 21{,}92$$

$$x^2_{11,0{,}995} = 26{,}8$$

Así el intervalo de confianza al 95 % es

$$\left(\frac{12(0{,}82)^2}{21{,}92}, \frac{12(0{,}82)^2}{3{,}168} \right)$$

y para 99 %

$$\left(\frac{12(0{,}82)^2}{26{,}8}, \frac{12(0{,}82)^2}{2{,}603} \right)$$

tenemos

$$(0{,}3681, 2{,}1145) \text{ y } (0{,}3011, 3{,}0998)$$

El intervalo se determina teniendo en cuenta

$$\frac{\bar{x} - \mu}{S/\sqrt{n}}$$

donde $n = 11$ Es una variable t de Student con 11 grados de libertad. El intervalo, para un nivel de confianza $p = 1 - \alpha$, viene dado por

$$\bar{x} \pm t_{11,1-\alpha/2} \frac{S}{\sqrt{11}}$$

Siendo $t_{11,1-\alpha/2}$ el valor del percentil $100(1 - \alpha/2)$ para 11 grados de libertad

$$t_{11,0{,}975} = 2{,}201$$

$$t_{11,0{,}995} = 3{,}1058$$

y los intervalos de confianza son, respectivamente

$$3{,}6 \pm 2{,}201 \frac{0{,}82}{\sqrt{11}} = 3{,}6 \pm 0{,}544$$

y

$$3{,}6 \pm 3{,}1058 \frac{0{,}82}{\sqrt{11}} = 3{,}6 \pm 0{,}768$$

7. Determine el estimador de máxima verosimilitud para θ si $X_1, X_2 \dots X_n$ es una muestra con función de densidad

$$f(x) = \frac{1}{2}e^{-|x-\theta|} \text{ para } -\infty < x < \infty$$

Solución:

8. La variable X tiene una media $a\theta$ y varianza σ^2. La variable Y , sin relación con X, tiene una media $b\theta$ y varianza $3\sigma^2$. Se usa una observación de la variable $X+Y$ para estimar θ. Calcule $E[X+Y]$. ¿En qué casos $X+Y$ es un estimador insesgado de θ?
Solución:

$$E[X] = a\theta \qquad E[Y] = b\theta$$

$$E[X+Y] = E[X] + E[Y]$$

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

y para que el estimador sea insesgado se debe cumplir que

$$E[\hat{\theta}] = \theta$$

entonces

$$E[\hat{\theta}] = \theta(a+b)$$
$$a+b = 1$$

$$a = 1-b \quad \text{ó} \quad b = 1-a$$

$$\sigma^2 = E[(X-\mu)^2]$$
$$\sigma^2 = E[X^2] - 2\mu E[X] + \mu^2$$
$$\sigma^2 = E[X^2] - 2\mu^2 + \mu^2$$
$$\sigma^2 = E[X^2] - \mu^2$$

$$Var(X)^2 = E(X^2) - a^2\theta^2 \qquad E(X)^2 = \sigma^2 + a^2\theta^2$$
$$Var(Y)^2 = E(Y^2) - b^2\theta^2 \qquad E(Y)^2 = 3\sigma^2 + b^2\theta^2$$
$$Var(X+Y) = E[(X+Y)^2] - 2[E(X+Y)]^2$$
$$Var(X+Y) = E(X^2) + 2E(X)E(Y) + E(Y^2) - [\theta(a+b)^2]$$
$$Var(X+Y) = \sigma^2 + a^2\theta^2 + 2(a\theta)(b\theta) + 3\sigma^2 + b^2\theta^2 - a^2\theta^2 - 6ab\theta^2 - 3b^2\theta^2$$

$$Var(X+Y) = \sigma^2 + a^2\theta^2 + 2ab\theta^2 + 3\sigma^2 + b^2\sigma^2 - a^2\theta^2 - 2ab\theta^2 - b^2\theta^2$$

$$Var(X+Y) = \sigma^2 + 3\sigma^2 = 4\sigma^2$$

9. ¿Cuál es el nivel de confianza para el intervalo $\bar{x} \pm 2{,}81\sigma/\sqrt{n}$?¿Cuál es el nivel de confianza para el intervalo $\bar{x} \pm 1{,}44\sigma/\sqrt{n}$
   Solución:
   el nivel de confianza tenemos que es
   $$1 - \alpha$$
   y dado un intervalo
   $$\bar{x} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = \bar{x} \pm 2{,}81\frac{\sigma}{\sqrt{n}}$$
   para calcular α en **R** dado $Z_{\alpha/2}$ se utiliza el comando
   qmorm(q=2.81,lower.tail=F)
   0.002477075
   $$\alpha = 0{,}002477075$$
   el nivel de confianza es:
   $$1 - 0{,}002477075$$
   $$0{,}9975229$$
   qmorm(q=1.44,lower.tail=F)
   0.0749337
   $$\alpha = 0{,}0749337$$
   el nivel de confianza es:
   $$1 - 0{,}0749337$$
   $$0{,}9250663$$

10. En cuánto debe aumentarse el tamaño muestral n, si la longitud del intervalo de confianza se reduce a la mitad? Si el tamaño muestral aumenta en un factor de 25, ¿qué efecto tiene en la longitud del intervalo de confianza?
    Solución:
    para una istribucción normal
    $$Z_{\alpha/2} \leq \frac{\sigma}{\sqrt{n}} \leq Z_{\alpha/2}$$
    $$\frac{\sigma}{\sqrt{n}} = 2Z_{\alpha/2}$$
    debe aumentar n
    $$n = \frac{\sigma^2}{4Z^2_{\alpha/2}}$$
    $$n = 25\left(\frac{\sigma^2}{4Z^2_{\alpha/2}}\right)^2$$
    aumenta la longitud del intervalo

11. sea $\alpha_1 > 0$, $\alpha_2 > 0$ con $\alpha_1 + \alpha_2 = \alpha$. Entonces

$$P\left(-z_{\alpha_1} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < z_{\alpha_2}\right) = 1-\alpha$$

Utilice esta expresión para derivar una expresión general del intervalo de confianza de $100(1-\alpha)\,\%$ para μ.
Solución:

$$-z_{\alpha_1} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < z_{\alpha_2}$$

$$-z_{\alpha_1} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < z_{\alpha-\alpha_1}$$

$$-z_{\alpha_1}\frac{\sigma}{\sqrt{n}} < \bar{x}-\mu < z_{\alpha-\alpha_1}\frac{\sigma}{\sqrt{n}}$$

$$-z_{\alpha_1}\frac{\sigma}{\sqrt{n}} - \bar{x} < -\mu < z_{\alpha-\alpha_1}\frac{\sigma}{\sqrt{n}} - \bar{x}$$

$$-z_{\alpha_1}\frac{\sigma}{\sqrt{n}} - \bar{x} > \mu > z_{\alpha-\alpha_1}\frac{\sigma}{\sqrt{n}} - \bar{x}$$

12. En un experimento de electrólisis un grupo midió la cantidad de cobre precipitado de una solución saturada de sulfato de cobre durante un periodo de 30 minutos. 30 estudiantes calcularon una media muestral y una desviación estándar iguales a 0,145 y 0,0051 respectivamente. Encuentre el intervalo de confianza del 90 % para la cantidad de cobre precipitado de la solución durante un periodo de 30 minutos.
Solución:
El intervalo de confianza para la media

$$\bar{x} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

para un coeficiente de confianza del 90 %

$$1-\alpha = 0{,}9$$

$$\alpha = 0{,}1$$

$$\frac{\alpha}{2} = \frac{0{,}1}{2} = 0{,}05$$

por lo que

qmorm(q=0.05,lower.tail=F)
1.645

$$Z_{\alpha/2} = 1{,}645$$

el intervalo de confianza es

$$0{,}145 \pm 1{,}645\frac{0{,}0051}{\sqrt{30}}$$

$$\left(0{,}145 + 1{,}645\frac{0{,}0051}{\sqrt{30}}, 0{,}145 - 1{,}645\frac{0{,}0051}{\sqrt{30}}\right) = (0{,}1465317061, 0{,}1434682939)$$

podemos concluir que entonces que hay una probabilidad del 90 % de que el valor medio de la característica esté entre 0,1465317061 y 0,1434682939

13. Un estudio e1 una muestra de 1251 personas, mostró que sólo el 53 % consume las cantidades recomendadas de alimentos ricos en fibras y verduras, 51 % evita la grasa y 46 % evita el consumo excesivo de sal.

   a) Construya un intervalo de confianza del 90 % para la proporción de la población que come las cantidades recomendadas de alimentos con fibra y verduras.

   b) Construya un intervalo de confianza del 99 % para la proporción de la población que evitan la grasa en su dieta.

   Solución:

   a) El intervalo de confianza para la media

$$\bar{x} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

   para un coeficiente de confianza del 90 %

$$1-\alpha = 0{,}9$$

$$\alpha = 0{,}1$$

$$\frac{\alpha}{2} = \frac{0{,}1}{2} = 0{,}05$$

   por lo que
   qmorm(q=0.05 ,lower.tail=F)
   1.645

$$Z_{\alpha/2} = 1{,}645$$

   el intervalo de confianza es

$$\bar{x} \pm 1{,}645\frac{\sigma}{\sqrt{663{,}53}}$$

   b) para un coeficiente de confianza del 99 %

$$1-\alpha = 0{,}99$$

$$\alpha = 0{,}01$$

$$\frac{\alpha}{2} = \frac{0{,}1}{2} = 0{,}005$$

   por lo que
   qmorm(q=0.005,lower.tail=F)
   2.257829

$$Z_{\alpha/2} = 2{,}575829$$

   el intervalo de confianza es

$$\bar{x} \pm 2{,}575829\frac{\sigma}{\sqrt{638{,}01}}$$

14. Se realizó un experimento para comparar dos dietas, A y B, diseñadas para la reducción de peso. Se eligieron al azar dos grupos de 30 personas con sobrepeso. Un grupo tomó la dieta A y el otro la dieta B. Sus pérdidas de peso después de un mes se anotaron y se obtuvieron las siguientes medias y desviaciones estándar

Dieta A	Dieta B
$\bar{x}_A = 21{,}33$	$\bar{x}_B = 13{,}4$
$S_A = 2{,}6$	$S_B = 1{,}9$

Encuentre el intervalo de confianza del 90 % para la diferencia entre la pérdida de peso media con las dos dietas.

Solución

El intervalo de confianza para la media

$$\bar{x}_A \pm Z_{\alpha/2}\frac{S_A}{\sqrt{n}}$$

para un coeficiente de confianza del 90 %

$$1-\alpha = 0{,}9$$

$$\alpha = 0{,}1$$

$$\frac{\alpha}{2} = \frac{0{,}1}{2} = 0{,}05$$

por lo que

qmorm(q=0.05 ,lower.tail=F)
1.645

$$Z_{\alpha/2} = 1{,}645$$

el intervalo de confianza de la Dieta A es

$$21{,}33 \pm 1{,}645\frac{2{,}6}{\sqrt{30}}$$

$$\left(21{,}33 - 1{,}645\frac{21{,}33}{\sqrt{30}}, 21{,}33 + 1{,}645\frac{21{,}33}{\sqrt{30}}\right)$$

$$(20{,}54913, 22{,}11087)$$

el intervalo de confianza de la Dieta B es

$$21{,}33 \pm 1{,}645\frac{2{,}6}{\sqrt{30}}$$

$$\left(13{,}4 - 1{,}645\frac{1{,}9}{\sqrt{30}}, 13{,}4 + 1{,}645\frac{1{,}9}{\sqrt{30}}\right)$$

$$(12{,}82936, 13{,}97064)$$

3. Pruebas de Hipótesis

1. Los estudiantes de la UAEH gastan un promedio de $\$1,500$ mensuales una desviación estándar de $\$300$ se toma una muestra de 36 estudiantes de física y sus gastos fueron de $\$1,600$ utilizar un nivel de significancia del 5 % para ver si estos datos apoyan la idea de que los estudiantes de física gasten mas que el estudiante promedio

solución:
la prueba de hipotesis nula es

$$H_0 = \mu \geq 1500$$

y la prueba de hipótesis alterna es

$$H_a > 1500$$

El intervalo de confianza para la media

$$\bar{x}_A \pm Z_{\alpha/2} \frac{S_A}{\sqrt{n}}$$

para un coeficiente de confianza del 95 %

$$1 - \alpha = 0{,}95$$

$$\alpha = 0{,}05$$

$$\frac{\alpha}{2} = \frac{0{,}1}{2} = 0{,}025$$

por lo que

qmorm(q=0.025 ,lower.tail=F)
1.645

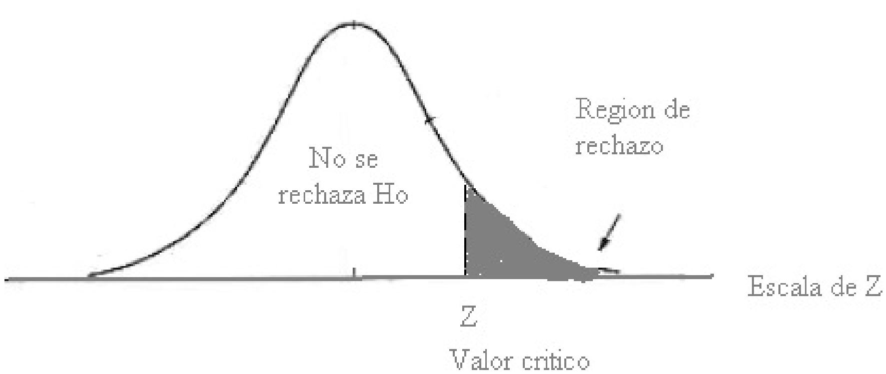
$$Z_{\alpha/2} = 1{,}959964$$

encontrando el estadístico de prueba

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$Z = \frac{1600 - 1500}{300/\sqrt{36}} = \frac{100}{50} = 2$$

por lo que el valor cae dentro de la región de la hipótesis alterna



2. Las siguientes observaciones de densidad se obtuvieron de 28 muestras de bosques con dos variedades A y B de árbol. Pruebe $H_o : \mu_1 = \mu_2$. Use una región de rechazo $\alpha = 0.05$

árbol A	0.76	0.74	0.75	0.72	0.73	0.71	0.73	0.75	0.76	0.75	0.69	0.73	0.74	0.74
	0.73	0.77	0.74	0.73	0.78	0.75	0.73	0.75	0.76	0.72	0.77	0.74	0.78	0.75
árbol B	0.68	0.67	0.72	0.69	0.66	0.71	0.69	0.71	0.66	0.70	0.72	0.68	0.70	0.69
	0.71	0.71	0.70	0.72	0.71	0.72	0.71	0.69	0.73	0.71	0.70	0.70	0.71	0.73

3. De las mujeres a las que se les diagnostica cáncer de mama, una tercera parte muere por esa enfermedad. Se selecciono una muestra de 200 mujeres que se examinaron periódicamente y finalmente se les diagnosticó la enfermedad. Sea p la tasa de supervivencia de un programa de detección oportuna. si 164 mujeres de la muestra sobreviven a la enfermedad, ¿se puede concluir que el programa de detección oportuna fue eficaz? haga una prueba de hipótesis con una región de rechazo de $\alpha = 0{,}05$
Solución:

$$H_0 \geq 1/3$$
$$H_a < 1/3$$
$$\alpha = 0{,}05$$

por lo que
qmorm(q=0.05 ,lower.tail=F)
1.645

$$Z_\alpha = 1{,}645$$

encontrando el estadístico de prueba

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$Z = \frac{\frac{164}{200} - 1/3}{\frac{(1/3)(2/3)}{200}} = \frac{100}{50} = 14{,}6$$

conclusión: el valor del estadistico de prueba H_a cae en la región de rechazo

4. Para que una línea aérea obtenga ganancias, necesita altas tasas de ocupación. Suponga que para que un vuelo resulte lucrativo debe contar con una ocupación promedio de 60 %. Al revisar la tasa de ocupación de 120 vuelos de las 10:00 de México a Monterrey, se encontró una ocupación promedio de 58 % con una desviación estándar de 11 %. Para una región de rechazo de $\alpha = 0.1$, verifique si hay suficiente evidencia para afirmar que el vuelo de las 10:00 no es lucrativo.
Solución

$> qnorm(p = 0{,}1)$
[1] -1.281552

$$H_0 = 0{,}6$$
$$H_0 \neq 0{,}6$$
$$Z_\alpha = -1{,}281552$$

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$
$$Z = \frac{72 - 69{,}6}{13{,}2/\sqrt{120}}$$
$$Z = 1{,}991718$$

conclusión: el valor del estadistico de prueba H_a cae en la región de rechazo

5. Aunque 85 % de la gente es diestra, una encuesta aplicada a 300 estudiantes de matemáticas demostró que 96 % eran diestros. ¿Es significativa esta diferencia, desde el punto de vista estadístico? Use una región de rechazo de $\alpha = 0.01$. Calcule el valor de p y explique su significado.
Solución:

$> qnorm(p = 0{,}01)$
$[1]\ -2{,}326348$

$$z_\alpha = -2{,}326348$$

$$Z = \frac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$$

$$Z = \frac{\dfrac{255}{300} - 0{,}96}{\sqrt{\dfrac{(0{,}96)(0{,}04)}{300}}}$$

$$Z = -9{,}722718$$

conclusión: el valor del estadistico de prueba H_a cae en la región de rechazo

6. Un fabricante de lavadoras automáticas ofrece un modelo en tres diferentes colores: A, B y C. De las primeras 1000 lavadoras que se vendieron, 400 eran de color A. ¿Concluiría usted que A es el color que prefieren los consumidores? Justifique su respuesta.
Solución:

$$H_0 = 1/3$$

$$H_a \neq 1/3$$

$$Z = \frac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$$

$$Z = \frac{\dfrac{1000}{400} - 0{,}6}{\sqrt{\dfrac{(0{,}6)(0{,}4)}{1000}}}$$

$$Z = 122{,}6445$$

7. Un epidemiólogo desea comparar dos vacunas para la rabia. Las personas que previamente habían recibido dichas vacunas se dividieron en dos grupos. El grupo 1 recibió una dosis de refuerzo de la vacuna 1 y el grupo 2 recibió una dosis de refuerzo de la vacuna 2. Las respuestas de los anticuerpos se registraron dos semanas después. Los datos se resumen en la siguiente tabla:

vacuna	tamaño de la muestra	$\bar{x}$	s
1	10	4.5	2.5
2	9	2.5	2.0

¿Indican estos datos que existe diferencia en la efectividad de las dos vacunas utilizadas? Use $\alpha = 0.05$
Solución:

$$z_p = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

$$z_p = \frac{2{,}5 - 4{,}5}{\sqrt{\frac{(2{,}5)^2}{10} + \frac{(2{,}0)^2}{9}}}$$

$$z_p = -1{,}933975$$

$> qnorm(p = 0{,}05)$
[1] $-1{,}644854$

$$z = -1{,}644854$$

se rechaza la prueba alternativa ya que $z_p < z$

8. Diez animales de laboratorio se sometieron a condiciones que simulaban una enfermedad. Se registró el número de latidos por minuto del corazón, antes y después del experimento, obteniéndose los siguientes resultados:

animal	1	2	3	4	5	6	7	8	9	10
antes	70	84	88	110	105	100	110	67	79	86
después	115	148	176	191	158	178	179	140	161	157

¿Proporcionan estos datos evidencia suficiente que indique que la condición experimental aumenta el número de latidos del corazón por minuto? Use $\alpha = 0.05$.
Solución:

$$H_0 : \mu_1 = \mu_2$$

$$H_a : \mu_1 \neq \mu_2$$

```
> datos1=read.delim('C:/Users/hec-p/Documents/6to_sem_LIFTA/Estadistica/archivos/animal.txt', header=F)
>antes=c(70, 84, 88, 110, 105, 100, 110, 67, 79, 86)
>despues=c(115, 148, 176, 191, 158, 178, 179, 140, 161, 157)
>t.test(antes, despues, paired=TRUE, alternative='less')
```

Paired t-test

data: antes and despues
t = $-16{,}632$, df = 9, p-value = $2{,}293 \times 10^{-08}$
alternative hypothesis: true difference in means is less than 0 95 percent confidence interval:
$-\infty, -62{,}64102$
sample estimates:
mean of the differences
$-70{,}4$
Como es de dos colas
$> qnorm(0{,}025)$
[1] $-1{,}959964$
Se observa que el p-valor es menor que $H_0 = -1{,}959964$, por lo que se rechaza la hipótesis nula

9. Se ha diseñado un aparato, semejante a una pistola, para reemplazar las jeringas en la administración de vacunas. Resultaría muy peligroso usar el aparato si la varianza fuera mayor a 0.01. Si en una muestra aleatoria de 24 inyecciones se obtuvo una desviación estándar de .08, ¿se descontinuará el uso del aparato si $\alpha = 0.1$? ¿Cuál será el nivel de significancia adecuado para este problema?
Solución:

$$X_p = \frac{(n-1)s^2}{\sigma^2}$$

$$X_p = \frac{(24-1)(0{,}08)^2}{0{,}01}$$

$$X_p = 14{,}72$$

> n=24 > $s = 0{,}08$
>sigma2=0.01
>chip= $(n-1) * s^2/$(sigma2)
>chialfa=qchisq($0{,}9, n-1$)
> *chip*
[1] 14,72
>chialfa
[1] 32,0069
Conclución, el estadístico de prueba no cae en la región de rechazo

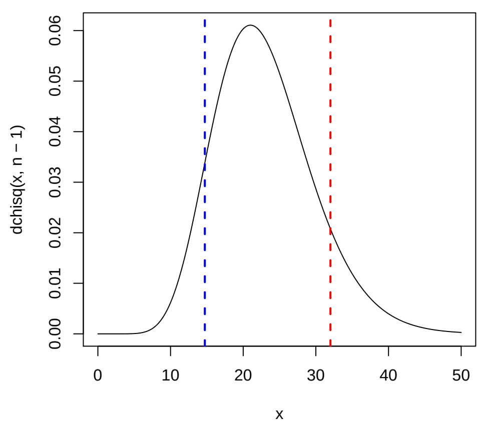


Figura 1: Gráficas de ejercicio 10.

10. Se cree que la cantidad de cera en cada lado de las bolsas de papel encerado está distribuida de forma equitativa. Sin embargo, hay alguna razón para pensar que hay una mayor variación en la cantidad de cera en la cara interior del papel que en la cara exterior. Se obtienen 25 muestras de la cantidad de cera en cada cara de esas bolsas y los datos se resumen en la tabla siguiente:

cara exterior	cara interior
$\bar{x}_1 = 0.948$	$\bar{x}_2 = 0.652$
$\sum x_{1i}^2 = 91$	$\sum x_{2i}^2 = 82$

Haga una prueba para determinar si la variabilidad de la cantidad de cera en la cara interior es mayor que en la cara exterior. Use $\alpha = 0.05$

Solución:

$$H_0 : \mu_1 = \mu_2$$

$$H_0 : \mu_1 \neq \mu_2$$

$$t_p = \frac{\bar{X}_1 - \bar{x}_2}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

donde

$$s^2 = \frac{(n_1 - 1)s_1 + (n_2 - 1)s_2}{n_1 + n_2 - 2}$$

en general

$$s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

$$s^2 = \frac{1}{n-1} \left(\sum x_i^2 - 2x_i\bar{x} + \bar{x}^2 \right)$$

$$s^2 = \frac{1}{n-1} \left(\sum x_i^2 - 2\bar{x} \sum x_i + \sum \bar{x}^2 \right)$$

por definición de media

$$\bar{x} = \sum x_i$$

entonces

$$s^2 = \frac{1}{n-1} \left(\sum x_i^2 - 2\bar{x} \sum x_i + n\bar{x}^2 \right)$$

$$s^2 = \frac{1}{n-1} \left(\sum x_i^2 - 2\bar{x}n\bar{x} + n\bar{x}^2 \right)$$

$$s^2 = \frac{1}{n-1} \left(\sum x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \right)$$

$$s^2 = \frac{1}{n-1} \left(\sum x_i^2 - n\bar{x}^2 \right)$$

sustituyendo los datos

$$s_1^2 = \frac{1}{25-1} (91 - 0{,}948)$$

$$s_1 = 3{,}752167$$

$$s_2^2 = \frac{1}{25-1} (82 - 0{,}652)$$

$$s_2 = 3{,}3895$$

$$s^2 = \frac{(25-1)(3{,}752167) + (25-1)(3{,}3895)}{25 + 25 - 2}$$

$$`2 = 0{,}1813335$$

$$s = 0{,}4258327$$

$$t_p = \frac{0{,}948 - 0{,}652}{0{,}4258327\sqrt{\frac{2}{25}}}$$

$$t_p = 5{,}771234$$

>qnorm(0.025) [1] -1.959964

conclusión : se no rechaza la prueba H_a

11. Una muestra aleatoria de 58 hombres con diabetes mellitus arrojó una media de índice de masa corporal de $\bar{x} = 25Kg/m^2$ y una desviación estándar de $s = 2.7Kg/m^2$. Para una $\alpha = 0.05$, pruebe si el índice de masa corporal medio representativo de la población de los hombres que contrajeron diabetes mellitus es igual a $24Kg/m^2$.

    Solución

$$Z_p = \frac{\bar{x} - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$

$$Z_p = \frac{25 - 24}{\sqrt{\frac{(2{,}7)^2}{58}}}$$

$$Z_p = 2{,}820657$$

$> qnorm(p = 0{,}025, lower.tail = F)$

[1] 1,959964

se rechaza la hipotesis H_a

12. La siguiente tabla compara niveles de carboxihemoglobina de un grupo de fumadores y un grupo de no fumadores. Si supone que la variabilidad en el nivel carboxihemoglobina es menor en los no fumadores que en los no fumadores, realice una prueba de hipótesis para $\alpha = 0.05$.

grupo	n	carboxihemoglobina [%]
fumadores	75	$\bar{x} = 4.1$ $s = 2.0$
no fumadores	121	$\bar{x} = 1.3$ $s = 1.3$

Solución

$$H_0 : f \leq nf$$

$$H_a : f < nf$$

$$Z_p = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$Z_p = \frac{4{,}1 - 1{,}3}{\sqrt{\frac{2}{75} + \frac{1{,}3}{121}}}$$

$$Z_p = -74{,}84536$$

$> qnorm(p = 0{,}05, lower.tail = F)$

[1] $-1{,}644854$

Conclusión. se rechaza la prueba alternativa, ya que cuyo valor es cae en la región de rechazo

13. Los siguientes datos indican la cantidad de glucosa en la sangre medida una y dos horas después de un desayuno.

Individuo	1	2	3	4	5	6	7
1 hora	73	58	67	93	33	18	147
2 horas	24	27	49	59	0	11	43

Pruebe que la concentración de glucosa en la sangre es mayor una hora después del desayuno. Use $\alpha = 0.05$.
Solución:

$$H_0 : \mu_2 \geq \mu_1$$
$$H_a : \mu < \mu_1$$

<hora1=c(73, 58, 67, 93, 33, 18, 147)
<hora2=c(24, 27, 49, 59, 0, 11, 43)
<t.test(hora2, hora1,paired=TRUE, alternative='less')
Paired t-test
data: hora2 and hora1
t = -3.3228, df = 6, p-value = 0.007974
alternative hypothesis: true difference in means is less than 0 95 percent confidence interval:
-Inf -16.37073
sample estimates:
mean of the differences
-39.42857
qnorm(0.05)
[1] -1.644854

$$H_0 = -1{,}644854$$

observemos que el $p - valor > H_0$ por lo que no se rechaza la hipótesis alternativa

14. Se aplicaron dos métodos para enseñar a leer a dos grupos de niños de primaria que se eligieron de manera aleatoria, y se realizó una comparación con una prueba de comprensión de la lectura al final del periodo de enseñanza. La siguiente tabla resume los resultados de la prueba. ¿Hay evidencia suficiente para afirmar que la media de los resultados de la prueba es diferente? ¿Qué se puede decir del nivel de significancia alcanzado?

	Método I	Método II
niños	11	14
$\overline{x}$	64	69
s	52	71

Solución

$$H_0 : \mu_1 = \mu_2$$
$$H_0 : \mu_1 \neq \mu_2$$
$$t_p = \frac{\bar{X}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

donde

$$s^2 = \frac{(n_1 - 1)s_1 + (n_2 - 1)s_2}{n_1 + n_2 - 2}$$
$$s^2 = \frac{(11 - 1)52 + (14 - 1)71}{11 + 14 - 2}$$
$$s^2 = 62{,}73913$$
$$s = 7{,}920804$$

$$t_p = \frac{64 - 69}{(7{,}920804)\sqrt{\frac{1}{11} + \frac{1}{14}}}$$
$$t_p = -1{,}566719$$

4. Momentos

1. Sea X una variable aleatoria con distribución binomial de parámetros n y p. Pruebe que

$$m(t) = (pe^t + q)^n.$$

solución:
tenemos la distribución binomial

$$f_x(X) = \binom{n}{X} p^X (1-p)^{n-X} \qquad X = 0, 1, \ldots, n$$

donde X representa el número de éxitos en n ensayos independientes, con probabilidad de éxito p constante en todos los ensayos obteniendo la función generadora

$$m(t) = E(e^{tX}) = \sum_X e^{tX} f_x(X)$$

sustituyendo la función de probabilidad

$$m(t) = E(e^{tX}) = \sum_X e^{tX} \binom{n}{x} p^X (1-p)^{n-X}$$

$$m(t) = E(e^{tX}) = \sum_X \binom{n}{x} (e^t p)^X (1-p)^{n-X}$$

por el binomio de newton

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$$

si

$$e^{\theta p} = a$$

$$1 - p = b$$

nos da la simplificación

$$m(t) = E(e^{tX}) = \sum_X \binom{n}{x} (e^t p)^X (1-p)^{n-X} = [e^\theta p + (1-p)]^n$$

$$q = 1 - p$$

se obtiene

$$m(t) = (pe^t + q)^n.$$

2. Sea X una variable aleatoria con distribución geométrica de parámetro p. Demuestre que

$$m(t) = \frac{pe^t}{1-qe^t}.$$

Solución:
Tenemos la distribución geométrica

$$f_x(X) = pq^{X-1}$$

y la función generadora

$$m(t) = E(e^{tX}) = \sum_{X=1}^{\infty} e^{tX} f_x(X)$$

sustituyendo la función de probabilidad

$$m(t) = E(e^{tX}) = \sum_{X=1}^{\infty} e^{tX} pq^{X-1}$$

$$m(t) = \sum_{X=1}^{\infty} e^{tX} pq^{X-1}$$

$$m(t) = p\sum_{X=1}^{\infty} e^{tX} q^{X-1}\frac{q}{q}$$

$$m(t) = \frac{p}{q}\sum_{X=1}^{\infty} e^{tX} q^{X}$$

$$m(t) = \frac{p}{q}\sum_{X=1}^{\infty} (e^{t}q)^{X}$$

$$m(t) = \frac{p}{q}\sum_{X=1}^{\infty} (e^{t}q)^{X}$$

$$m(t) = \frac{p}{q}\lim_{x\to\infty}\left(qe^t + (qe^t)^2 + (qe^t)^3 + \ldots + (qe^t)^x\right)$$

$$m(t) = \lim_{x\to\infty}\frac{p}{q}\left(\frac{(qe^t)^{x+1} - qe^{t\,n}}{qe^t - 1}\right)$$

$$m(t) = \frac{p}{q}\left(\frac{-qe^t}{qe^t-1}\right)$$

$$m(t) = \frac{p}{q}\left(\frac{qe^t}{1-qe^t}\right)$$

$$m(t) = p\left(\frac{e^t}{1-qe^t}\right)$$

$$m(t) = \frac{pe^t}{1-qe^t}$$

3. Sea
$$m(t) = \frac{1}{6}e^t + \frac{2}{6}e^{2t} + \frac{3}{6}e^{3t}.$$

Encuentre $E[X]$ y $Var(X)$.
Solución:
$\frac{dm(t)}{dt}\left(\frac{1}{6}e^t + \frac{2}{6}e^{2t} + \frac{3}{6}e^{3t}\right) = \frac{1}{6}e^t + \frac{4}{6}e^{2t} + \frac{9}{6}e^{3t} = \frac{1}{6}e^t + \frac{2}{3}e^{2t} + \frac{3}{2}e^{3t}$
$$E[X] = \frac{1}{6}e^t + \frac{2}{3}e^{2t} + \frac{3}{2}e^{3t}$$

$$\frac{d^2m(t)}{dt^2}\left(\frac{1}{6}e^t + \frac{2}{6}e^{2t} + \frac{3}{6}e^{3t}\right)$$

$$\frac{dm(t)}{dt}\left(\frac{1}{6}e^t + \frac{2}{3}e^{2t} + \frac{3}{2}e^{3t}\right)$$

$$Var(X) = \frac{1}{6}e^t + \frac{4}{3}e^{2t} + \frac{9}{2}e^{3t}$$

4. Sea X una variable aleatoria con función generadora de momentos $m(t)$.
Sea $Y = aX + b$. Demuestre que la función generadora de momentos de Y es
$$m_Y(t) = e^{bt}m(at).$$

Solución
$$f_x(x) = aX + b$$

y la función generadora

$$m(t) = E(e^{tX}) = \int_0^t e^{tX} f_x(X)$$

sustituyendo la función de probabilidad

$$m_Y(t) = \int_0^t e^{bX}(aX + b)db$$
$$m_Y(t) = e^{bX}at$$

5. Probabilidad Condicional y Teorema de Bayes

1. Sea A el evento de que una familia seleccionada tenga dos niños de cada sexo y sea B el evento de que la familia seleccionada tenga, al menos, una niña. Asuma que cada familia tiene 4 hijos. ¿Cuál es $P(A|B)$?
   Solución:
   $d_n = \{HHHH, HHHM, HHMH, HMHH, MHHH, HHMM, HMMH, MMHH, MHHM, MMMH, MMHM, M$
   al menos una niña
   $$d_n = \{HHHH, HHHM, HHMH, HMHH, MHHH, \}$$
   $$P(A/B) = P(A) = \frac{5}{15} = \frac{1}{3} = 0{,}3333$$

2. Sean $S = \{0; 1; 2; 3; 4; 5; 6; 7\}$, $A = \{0; 1; 6; 7\}$, $B = \{0; 3; 4; 7\}$ y $C = \{0; 2; 5; 7\}$. Muestre que A, B, C son independientes por pares, pero no son independientes en tripleta
   Solución:
   $$P(A/B) = P(A)$$
   $$P(A/C) = P(A)$$
   $$P(B/C) = P(B)$$
   $$P(B/A) = P(B)$$
   $$P(C/A) = P(C)$$
   $$P(C/B) = P(C)$$

3. Una urna contiene v bolas verdes y r bolas rojas. Se sacan dos bolas al azar, sin reemplazo, una después de otra. Cuál es la probabilidad de que la segunda bola sea verde, sabiendo que la primera fue verde?
   Solución:
   $$P(V_1, V_2) = P(V_1)P(V_2) = \frac{v}{v+r} \cdot \frac{v-1}{v+r-1}$$
   $$p(V_1, V_2) = \frac{v(v-1)}{(v+r)(v+r-1)} = \frac{v(v-1)}{v^2+vr-v+vr+r^2-r}$$
   $$p(V_1, V_2) = \frac{v(v-1)}{v^2+2vr+r^2-v-r} = \frac{v(v-1)}{(v+r)^2-(v-r)}$$
   la probabilidad es de

4. ¿Cuál es la probabilidad de tener tres cartas de corazones cuando se sacan tres cartas sin reemplazo de un mazo de 52 cartas?
   Solución:
   $$p(x) = p(c_1)p(c_2)p(c_3)$$
   $$p(x) = \frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} = \frac{1716}{132600} = \frac{11}{850}$$
   la probabilidad es de
   $$p(x)0{,}01294118$$

5. Un paciente es tratado simultáneamente con tres fármacos: A, B, C, los cuales no interactúan. La evidencia produce los siguientes resultados:

   $P(\text{responde a } A) = 0{,}2$

   $P(\text{responde a } B) = 0{,}4$

   $P(\text{responde a } C) = 0{,}4$

   $P(\text{cura}|\text{responde a } A) = 0{,}2$

   $P(\text{cura}|\text{responde a } B) = 0{,}5$

   $P(\text{cura}|\text{responde a } C) = 0{,}3$

   ¿Qué significa la suposición de que no interactúan entre sí los fármacos? ¿Cuál es la probabilidad de que el paciente que recibe los tres medicamentos sea curado? Si el paciente se cura, ¿cuál es la probabilidad de que la cura se deba al medicamento B?
   Solución:

   - Que son independientes
   - $P(A, B, C) = 1$
   - $$P(B) = \frac{(0{,}5)(0{,}4)}{(0{,}2)(0{,}2) + (0{,}4)(0{,}5) + (0{,}4)(0{,}3)} = 0{,}5555556$$

     la probabilidad que se deba a B es 0,5555556

6. Una prueba es 98 % eficiente para la detección de cierto tipo de cáncer y 95 % eficiente para quienes no tienen ese tipo de cáncer. Si el 25 % de quienes se practican la prueba tienen ese tipo de cáncer, ¿cuál es la probabilidad de que una persona tenga ese tipo de cáncer cuando la prueba dice que lo tiene?
   Solución
   aplicando el teorema de Bayes, Tenemos

   $$P(x) = \frac{(0{,}25)(0{,}95)}{(0{,}25)(0{,}95) + (0{,}25)(0{,}98)}$$

   $$P(x) = 0{,}507772$$