

Probabilidad y Estadística (C)

December 5, 2021

Para las fechas de final de Diciembre 2021 el profesor Nicolas Saintier publicó listas de ejercicios a ser evaluados en cada fecha de examen. A continuación se resuelven los [35 ejercicios](#) para la 1er fecha de Diciembre. Si encontrás algún error por favor reportalo [acá](#).

Contents

| | |
|---|-----------|
| 1 Probabilidad: definición y enunciados | 2 |
| 1.1 Ejercicio 1 | 2 |
| 1.2 Ejercicio 2 | 3 |
| 2 Probabilidad condicional e independencia | 3 |
| 2.1 Ejercicio 3 | 3 |
| 2.2 Ejercicio 4 | 4 |
| 2.3 Ejercicio 5 | 5 |
| 3 Variables aleatorias | 7 |
| 3.1 Ejercicio 6 | 7 |
| 4 Vectores aleatorios | 9 |
| 4.1 Ejercicio 7 | 9 |
| 4.2 Ejercicio 8 | 10 |
| 5 Esperanza | 12 |
| 5.1 Ejercicio 9 | 12 |
| 5.2 Ejercicio 10 | 13 |
| 6 Esperanza condicional | 15 |
| 6.1 Ejercicio 11 | 15 |
| 6.2 Ejercicio 12 | 16 |
| 7 Generación de variables aleatorias | 16 |
| 7.1 Ejercicio 13 | 16 |
| 7.2 Ejercicio 14 | 17 |
| 8 Función característica | 18 |
| 8.1 Ejercicio 15 | 18 |
| 9 Convergencia de variables aleatorias | 18 |
| 9.1 Ejercicio 16 | 18 |

| | | |
|-----------|--------------------------------|-----------|
| 9.2 | Ejercicio 17 | 19 |
| 9.3 | Ejercicio 18 | 20 |
| 9.4 | Ejercicio 19 | 21 |
| 9.5 | Ejercicio 20 | 22 |
| 9.6 | Ejercicio 21 | 24 |
| 10 | Procesos de Poisson | 25 |
| 10.1 | Ejercicio 22 | 25 |
| 10.2 | Ejercicio 23 | 25 |
| 11 | Estimadores | 26 |
| 11.1 | Ejercicio 24 | 26 |
| 11.2 | Ejercicio 25 | 26 |
| 11.3 | Ejercicio 26 | 28 |
| 11.4 | Ejercicio 27 | 29 |
| 11.5 | Ejercicio 28 | 32 |
| 11.6 | Ejercicio 29 | 32 |
| 12 | Intervalos de confianza | 34 |
| 12.1 | Ejercicio 30 | 34 |
| 12.2 | Ejercicio 31 | 35 |
| 12.3 | Ejercicio 32 | 36 |
| 13 | Test de hipótesis | 38 |
| 13.1 | Ejercicio 33 | 38 |
| 13.2 | Ejercicio 34 | 39 |
| 13.3 | Ejercicio 35 | 40 |

1 Probabilidad: definición y enunciados

1.1 Ejercicio 1

Dar la definición de una medida de probabilidad P . Demuestre a partir de la definición que $P(A^c) = 1 - P(A)$ y que si $B \subset A$ entonces $P(A \setminus B) = P(A) - P(B)$.

Definiciones previas

- Dado un experimento aleatorio (cualquier proceso o acción repetible que genera una observación), llamamos **espacio muestral** al conjunto de todos los resultados posibles. Lo notamos S y éste puede ser finito o infinito numerable.
- Cualquier subconjunto del espacio muestral S lo llamamos **evento** o suceso.

Definición de una medida de probabilidad

Dado un espacio muestral S y un evento A , llamamos $P(A)$ a la probabilidad de que suceda el evento A . Esta medida debe satisfacer los siguientes axiomas:

- Axioma 1
 $P(A) \geq 0$ para todo evento $A \subset S$
La medida de probabilidad no puede ser negativa.

- Axioma 2

$$P(S) = 1$$

Como S es el conjunto de todos los resultados posibles, su probabilidad es 1 (el máximo posible en nuestra definición). Si no fuese 1 entonces estaría mal definido el espacio muestral, pues habría resultados posibles de nuestro experimento que no fueron incluidos en la definición de S .

- Axioma 3

Sean A_1, \dots, A_n una colección finita de eventos mutuamente excluyentes.

Es decir, $A_i \cap A_j = \emptyset \forall i \neq j$

$$\text{Entonces: } P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$$

Esto nos dice que la probabilidad de eventos mutuamente excluyentes (si sucede uno no sucede el otro) es la suma de sus probabilidades individuales. Por ejemplo, si tiramos un dado de 6 caras, 1 sola vez, la probabilidad de que salga un 2 o un 4 es la suma de la probabilidad de que salga un 2 y la de que salga un 4, pues no pueden suceder estos eventos al mismo tiempo.

Este axioma se extiende también para colecciones infinitas numerables.

Demostraciones

$$1 \stackrel{Ax2}{=} P(S) = P(A \cup A^c) \stackrel{Ax3}{=} P(A) + P(A^c) \implies P(A^c) = 1 - P(A)$$

$$B \subset A \implies A = (A \setminus B) \cup B$$

$$P(A) = P((A \setminus B) \cup B) \stackrel{Ax3}{=} P(A \setminus B) + P(B) \implies P(A \setminus B) = P(A) - P(B)$$

1.2 Ejercicio 2

Dar la definición de una medida de probabilidad P . Demuestre a partir de la definición que $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

$$P(B) = P((B \cap A) \cup (B \cap A^c)) \stackrel{Ax3}{=} P(B \cap A) + P(B \cap A^c) \implies P(B \cap A^c) = P(B) - P(B \cap A)$$

$$P(A \cup B) = P(A \cup (B \cap A^c)) \stackrel{Ax3}{=} P(A) + P(B \cap A^c) = P(A) + P(B) - P(B \cap A)$$

2 Probabilidad condicional e independencia

2.1 Ejercicio 3

En una muestra de 100 personas hay 13 enfermos no vacunados, 2 enfermos vacunados, 75 sanos vacunados, 10 sanos no vacunados. Elegimos una persona al azar y vemos que está enferma. ¿Cuál es la probabilidad que no se haya vacunado?

| | Vacunados | No vacunados | |
|----------|-----------|--------------|-----|
| Enfermos | 2 | 13 | 15 |
| Sanos | 75 | 10 | 85 |
| | 77 | 23 | 100 |

Definimos los siguientes eventos:

- E = La persona está enferma
- E^c = La persona no está enferma
- V = La persona está vacunada
- V^c = La persona no está vacunada

Nos piden calcular la probabilidad de que una persona no se haya vacunado sabiendo que está enferma. Esto lo calculamos mediante probabilidad condicional, y para las probabilidades puntuales de cada caso miramos la tabla y dividimos cantidad de personas en cada situación por la cantidad total de personas.

$$P(V^c|E) = \frac{P(V^c \cap E)}{P(E)} = \frac{\frac{13}{100}}{\frac{15}{100}} = 0.8\hat{6}$$

Utilizando probabilidad total (y así evitar calcular la tabla) considerando a V y V^c como una partición del espacio muestral:

$$P(V^c|E) = \frac{P(V^c \cap E)}{P(E \cap V) + P(E \cap V^c)} = \frac{\frac{13}{100}}{\frac{2}{100} + \frac{13}{100}} = 0.8\hat{6}$$

2.2 Ejercicio 4

Una urna tiene 4 bolas negras y 3 rojas. Sacamos tres bolas sin reposición. ¿Cuál es la probabilidad que la primera bola salga negra y la tercera salga roja?

Forma 1

Definimos los siguientes eventos:

- N_i = La bola extraída en la i -ésima extracción es negra
- R_i = La bola extraída en la i -ésima extracción es roja
- X_i = La bola extraída en la i -ésima extracción es negra o roja

Vamos calculando la probabilidad de obtener la bola deseada en cada extracción.

Primero vemos que $P(N_1) = \frac{4}{7}$ pues en la primer extracción hay 4 bolas negras del total de 7.

En la segunda extracción no nos importa qué bola sale, pero solo pueden pasar 2 cosas mutuamente excluyentes: sale una bola negra o sale una roja.

$P(N_2 \cap N_1) = P(N_2|N_1)P(N_1) = \frac{3}{6}P(N_1)$ pues en la segunda extracción quedan 3 bolas negras en las 6 bolas restantes en la urna.

$P(R_2 \cap N_1) = P(R_2|N_1)P(N_1) = \frac{3}{6}P(N_1)$ pues en la segunda extracción quedan aún las 3 bolas rojas en las 6 bolas restantes en la urna.

Finalmente, en la tercer extracción queremos ver la probabilidad de extraer una bola roja. Tenemos que calcular esto a partir de cada uno de los 2 resultados posibles de la segunda extracción (salió una bola negra o salió una roja).

$P(R_3 \cap N_2 \cap N_1) = P(R_3|N_2 \cap N_1)P(N_2 \cap N_1) = \frac{3}{5}P(N_2 \cap N_1)$ pues en la tercer extracción quedan aún las 3 bolas rojas en las 5 bolas restantes en la urna.

$P(R_3 \cap R_2 \cap N_1) = P(R_3|R_2 \cap N_1)P(R_2 \cap N_1) = \frac{2}{5}P(R_2 \cap N_1)$ pues en la tercer extracción quedan 2 bolas rojas en las 5 bolas restantes en la urna.

Entonces juntando los 2 escenarios posibles calculamos la probabilidad pedida:

$$P(R_3 \cap X_2 \cap N_1) = P(R_3 \cap N_2 \cap N_1) + P(R_3 \cap R_2 \cap N_1) = \frac{3}{5} * \frac{3}{6} * \frac{4}{7} + \frac{2}{5} * \frac{3}{6} * \frac{4}{7} = \frac{2}{7}$$

Forma 2

Definimos el espacio muestral S como el conjunto de todos los resultados posibles de extraer 3 bolas sin reposición.

$$\#S = 7 * 6 * 5 = 210$$

Definimos el evento A = la primer bola es negra y la tercera roja.

$$\#A = 4 * 3 * 3 + 4 * 3 * 2 = 60$$

La lógica para calcular $\#A$ es la misma que la forma 1, se divide en 2 casos (la segunda bola es negra o roja), y luego contamos la cantidad de bolas del color deseado en cada extracción.

$$\text{Finalmente, } P(A) = \frac{\#A}{\#S} = \frac{60}{210} = \frac{2}{7}$$

2.3 Ejercicio 5

Enuncie y demuestre la fórmula de la probabilidad total y el Teorema de Bayes.

Definiciones previas

- Sean A y B eventos cualesquiera del espacio muestral S tal que $P(B) > 0$. La probabilidad del evento A **condicional** a la ocurrencia del evento B es:

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

Una propiedad muy útil es la regla del producto: $P(A \cap B) = P(A|B)P(B)$

Si además $P(A) > 0 \implies P(A \cap B) = P(B|A)P(A)$

- Una colección de eventos A_1, \dots, A_n constituye una **partición** del espacio muestral S si se cumplen las siguientes condiciones:

$$- A_i \cap A_j = \emptyset \quad \forall i \neq j$$

Todos los eventos son disjuntos entre sí.

$$- P(A_i) > 0 \quad \forall i$$

Todos los eventos tienen probabilidad positiva.

$$- \bigcup_{i=1}^n A_i = S$$

Esto nos dice que al unir toda la colección de eventos A_1, \dots, A_n obtenemos el espacio muestral S en su totalidad.

Probabilidad total

Sea A_1, \dots, A_n una partición del espacio muestral S y sea B un evento cualquiera. Entonces:

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i)$$

Demostración:

$$B = B \cap S = B \cap \left(\bigcup_{i=1}^n A_i \right) = \bigcup_{i=1}^n (B \cap A_i)$$

Usando que $(B \cap A_i) \cap (B \cap A_j) = \emptyset \forall i \neq j$ y la regla del producto de la probabilidad condicional:

$$P(B) = P\left(\bigcup_{i=1}^n (B \cap A_i)\right) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B|A_i)P(A_i)$$

Teorema de Bayes

Sea A_1, \dots, A_n una partición del espacio muestral S y sea B un evento cualquiera tal que $P(B) > 0$. Entonces:

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^n P(B|A_i)P(A_i)}$$

El Teorema de Bayes nos permite calcular la probabilidad de un evento A sabiendo B en términos de la probabilidad del evento B sabiendo A. Es decir, nos permite “dar vuelta” el condicional.

Demostración:

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

En el numerador se usó la regla del producto de la probabilidad condicional, y en el denominador se usó el teorema de la probabilidad total.

3 Variables aleatorias

3.1 Ejercicio 6

Sea U una variable uniforme en $[0, 1]$. Para $u \in [0, 1]$ defina $h(u) = \max\{u, 1 - u\}$. Calcule la función de densidad de $X = h(U)$, su distribución acumulada, esperanza y varianza.

$$U \sim \mathcal{U}[0, 1]$$

$$f_U(u) = \frac{1}{1-0} I_{[0,1]}(u) = I_{[0,1]}(u)$$

$$F_U(u) = \frac{u-0}{1-0} = u \quad \forall 0 \leq u \leq 1$$

$$h(u) = \begin{cases} 1 - u & \text{si } 0 \leq u < 0.5 \\ u & \text{si } 0.5 \leq u \leq 1 \end{cases}$$

$$X = h(U) = \max\{U, 1 - U\}$$

$$F_X(x) = P(X \leq x) = P(\max\{U, 1 - U\} \leq x)$$

Observemos que:

$$\max\{U, 1 - U\} \leq x \iff u \leq x \quad \forall u \in \max\{U, 1 - U\}$$

Es decir, si el máximo de un conjunto es menor que x , entonces todos los elementos del conjunto son menores que x .

$$F_X(x) = P((1 - U \leq x) \cap (U \leq x)) = P((1 - x \leq U) \cap (U \leq x)) = P(1 - x \leq U \leq x)$$

Busquemos el nuevo intervalo de la distribución acumulada F_X :

$$\begin{aligned} 1 - x \leq U \leq x &\implies 1 - x \leq x \iff 0.5 \leq x \\ \implies F_X(x) &= 0 \quad \forall x < 0.5 \text{ pues } P(1 - x \leq U \leq x) = 0 \quad \forall x < 0.5 \end{aligned}$$

Como U es una uniforme estándar:

$$\begin{aligned} 0 \leq U \leq 1 &\implies 0 \leq 1 - x \wedge x \leq 1 \iff x \leq 1 \\ \implies F_X(x) &= 1 \quad \forall x > 1 \text{ pues } P(1 - x \leq U \leq x) = 1 \quad \forall x > 1 \end{aligned}$$

Por lo tanto, solo consideramos los $x \in [0.5, 1]$, y notemos que éste intervalo está incluido en el de la uniforme estándar.

$$F_X(x) = F_U(x) - F_U(1-x) = x - (1-x) = (2x-1) \quad \forall 0.5 \leq x \leq 1$$

Ahora derivamos para obtener la función de densidad:

$$f_X(x) = \frac{\partial}{\partial x} F_X(x) = 2I_{[0.5,1]}(x)$$

Por lo tanto podemos concluir que $X \sim \mathcal{U}[0.5, 1]$

Esperanza

Utilizando la fórmula para la esperanza de una uniforme:

$$E(X) = \frac{0.5+1}{2} = \frac{3}{4}$$

Por definición:

$$E(X) = \int_{0.5}^1 x f_X(x) dx = \int_{0.5}^1 2x dx = [x^2]_{0.5}^1 = 1^2 - 0.5^2 = \frac{3}{4}$$

Por definición usando $h(u)$:

$$\begin{aligned} E(X) &= E(h(U)) \\ &= \int_0^1 h(u) f_U(u) du \\ &= \left(\int_0^{0.5} 1-u du \right) + \left(\int_{0.5}^1 u du \right) \\ &= \left[u - \frac{u^2}{2} \right]_0^{0.5} + \left[\frac{u^2}{2} \right]_{0.5}^1 \\ &= \left(0.5 - \frac{0.5^2}{2} \right) + \left(\frac{1^2}{2} - \frac{0.5^2}{2} \right) \\ &= \frac{3}{4} \end{aligned}$$

Varianza

Utilizando la fórmula para la varianza de una uniforme:

$$V(X) = \frac{(1-0.5)^2}{12} = \frac{1}{48}$$

Por definición $E(X^2)$:

$$E(X^2) = \int_{0.5}^1 x^2 f_X(x) dx = \int_{0.5}^1 2x^2 dx = \left[\frac{2x^3}{3} \right]_{0.5}^1 = \frac{2}{3} - \frac{2 * 0.5^3}{3} = \frac{7}{12}$$

Por definición $E(X^2)$ usando $h(u)$:

$$\begin{aligned}
E(X^2) &= E(h(U)^2) \\
&= \int_0^1 h(u)^2 f_U(u) du \\
&= \left(\int_0^{0.5} (1-u)^2 du \right) + \left(\int_{0.5}^1 u^2 du \right) \\
&= \left(\int_0^{0.5} 1 - 2u + u^2 du \right) + \left(\int_{0.5}^1 u^2 du \right) \\
&= \left[u - u^2 + \frac{u^3}{3} \right]_0^{0.5} + \left[\frac{u^3}{3} \right]_{0.5}^1 \\
&= \left(0.5 - 0.5^2 + \frac{0.5^3}{3} \right) + \left(\frac{1^3}{3} - \frac{0.5^3}{3} \right) \\
&= \frac{7}{12}
\end{aligned}$$

$$V(X) = E(X^2) - E(X)^2 = \frac{7}{12} - \left(\frac{3}{4}\right)^2 = \frac{1}{48}$$

4 Vectores aleatorios

4.1 Ejercicio 7

Alicia y José acordaron encontrarse a las 8 de la noche para ir al cine. Como no son puntuales, se puede suponer que los tiempos X e Y en que cada uno de ellos llega son variables aleatorias con distribución uniforme entre las 8 y las 9. Además se supondrá que estos tiempos son independientes. Si ambos están dispuestos a esperar no más de 10 minutos al otro a partir del instante en que llegan, ¿cuál es la probabilidad de que se desencuentren?

Como X e Y son independientes $\implies f_{XY}(x, y) = f_X(x)f_Y(y) = I_{[8,9]}(x)I_{[8,9]}(y)$

Observación: 10 minutos = $\frac{1}{6}$ de hora

$P(|X - Y| \leq \frac{1}{6})$ es la probabilidad de encuentro

$P(|X - Y| > \frac{1}{6})$ es la probabilidad de **desencuentro**, lo que nos piden calcular

Necesitamos tomar el valor absoluto de la resta para manejar el caso de que Alicia llegue antes que José y viceversa. No obstante, como X e Y son uniformes en el mismo intervalo, las 2 regiones que se forman son simétricas y por lo tanto podemos calcular la probabilidad de 1 caso y luego multiplicar por 2.

Veamos primero cómo son las regiones de probabilidad de desencuentro ([ver en GeoGebra](#)).

Supongamos $X \geq Y$ (el caso del triángulo inferior):

$$|X - Y| > \frac{1}{6} \iff X > Y + \frac{1}{6}$$

Como despejamos X en función de Y , necesitamos ajustar el intervalo de integración de Y para integrar correctamente dentro del intervalo de la distribución. Recordemos que ambas variables

tienen distribución $U[8, 9]$, por lo tanto si el intervalo de integración de X es $[Y + \frac{1}{6}, 9]$, busquemos la cota superior del intervalo de integración de Y :

$$Y + \frac{1}{6} \leq 9 \iff Y \leq \frac{53}{6}$$

$$P(X > Y + \frac{1}{6}) = \int_8^{\frac{53}{6}} \int_{y+\frac{1}{6}}^9 f_{XY}(x, y) dx dy = \frac{25}{72}$$

Recordemos que este era solo uno de los casos posibles, así que debemos multiplicar por 2 para encontrar la probabilidad pedida.

$$P(|X - Y| > \frac{1}{6}) = 2P(X > Y + \frac{1}{6}) = \frac{25}{36}$$

4.2 Ejercicio 8

En un día una gallina pone N huevos donde $N \sim \mathcal{P}(\lambda)$. Independientemente de la cantidad de huevos puestos, de un huevo cualquiera nace un pollito con probabilidad p . Hallar la distribución de la cantidad de pollitos nacidos.

X = Cantidad de pollitos nacidos

Observemos que $(X|N = n) \sim Bi(n, p)$, es decir, la cantidad de pollitos que nacen condicionado a la cantidad de huevos puestos tiene una distribución Binomial.

El espacio muestral de N es infinito numerable, ya que la cantidad de huevos puestos puede ser en teoría infinito pero se puede contar con los números \mathbb{N}_0 . El conjunto de todos los valores de $n \geq 0$ genera una partición del espacio muestral, ya que la cantidad de huevos puestos son todos eventos disjuntos. Por lo tanto, para encontrar la distribución de X podemos usar el teorema de probabilidad total con una suma infinita.

$$P(X = k) = \sum_{n=0}^{\infty} P(X = k|N = n)P(N = n) = \sum_{n=0}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \frac{e^{-\lambda} \lambda^n}{n!}$$

Arrancamos la sumatoria desde k pues los términos anteriores son todos 0 por el combinatorio $\binom{n}{k}$

$$= \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \frac{e^{-\lambda} \lambda^n}{n!}$$

Aplicamos la fórmula del número combinatorio

$$= \sum_{n=k}^{\infty} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \frac{e^{-\lambda} \lambda^n}{n!}$$

Simplificamos los $n!$

$$= \sum_{n=k}^{\infty} \frac{p^k (1-p)^{n-k} e^{-\lambda} \lambda^n}{k!(n-k)!}$$

Hacemos aparecer un $n-k$ en la potencia de λ

$$= \sum_{n=k}^{\infty} \frac{p^k (1-p)^{n-k} e^{-\lambda} \lambda^{n-k+k}}{k!(n-k)!}$$

Sacamos afuera de la sumatoria todo lo que no depende de n

$$= \frac{p^k e^{-\lambda} \lambda^k}{k!} \sum_{n=k}^{\infty} \frac{(1-p)^{n-k} \lambda^{n-k}}{(n-k)!}$$

Juntamos el numerador de la sumatoria bajo una única potencia

$$= \frac{p^k e^{-\lambda} \lambda^k}{k!} \sum_{n=k}^{\infty} \frac{((1-p)\lambda)^{n-k}}{(n-k)!}$$

Ajustamos el rango de la sumatoria para que arranque en 0, en efecto le sumamos k a todos los n para compensar, y así se terminan cancelando con todos los k que estaban restando

$$= \frac{p^k e^{-\lambda} \lambda^k}{k!} \sum_{n=0}^{\infty} \frac{((1-p)\lambda)^n}{n!}$$

Utilizamos esta propiedad del número exponencial: $\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$

$$= \frac{p^k e^{-\lambda} \lambda^k}{k!} e^{(1-p)\lambda}$$

Agrupamos las potencias y encontramos la distribución

$$= \frac{e^{-\lambda+(1-p)\lambda} (\lambda p)^k}{k!} = \frac{e^{-\lambda p} (\lambda p)^k}{k!} \implies X \sim \mathcal{P}(\lambda p)$$

5 Esperanza

5.1 Ejercicio 9

Calcule la esperanza de una variable Binomial de parámetros n y p .

Sea $X \sim Bi(n, p)$

Por tener distribución Binomial, su función de densidad puntual está dada por:

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \forall k \in \mathbb{N}_0$$

Sea $R_X = \{0, \dots, n\}$ el rango de X (el conjunto de todos los posibles valores que puede tomar la v.a. X)

Utilizando la definición de la esperanza calculamos $E(X)$:

$$E(X) = \sum_{k \in R_X} k p_X(k) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k}$$

Reescribimos el rango de la sumatoria empezando en 1 (pues el primer término con $k = 0$ no aporta nada):

$$= \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k}$$

Aplicamos la fórmula del número combinatorio:

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

Cancelamos el k sacando un k del $k!$ que está en el denominador:

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

Sacamos un n y un p afuera de la sumatoria:

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

Volvemos a escribir el combinatorio pero con los nuevos valores:

$$= np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k}$$

Hacemos aparecer $(k-1)$ en la potencia:

$$= np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

Hacemos un cambio de variable: $j = k - 1$:

$$= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{(n-1)-j}$$

Reemplazamos la sumatoria por la fórmula del Binomio de Newton y encontramos la esperanza:

$$= np(p + (1-p))^{n-1} = np \implies E(X) = np$$

5.2 Ejercicio 10

Sea el vector aleatorio (X, Y) con distribución:

$$P(X = x, Y = y) = \begin{cases} \frac{1}{10} & x = 1, 2, 3, 4 \wedge y = 1, 2, 3, 4 \wedge y \leq x \\ 0 & \text{caso contrario} \end{cases}$$

Calcular el coeficiente de correlación $\rho(X, Y)$ entre X e Y y decidir si X e Y son o no independientes.

Primero calculamos todas las cosas que vamos a necesitar para resolver el ejercicio.

Calculamos las funciones de probabilidad marginal:

$$p_X(x) = \sum_{y \in R_Y} P(X = x, Y = y) = \sum_{y=1}^x \frac{1}{10} = \frac{x-1+1}{10} = \frac{x}{10}$$

$$p_Y(y) = \sum_{x \in R_X} P(X = x, Y = y) = \sum_{x=y}^4 \frac{1}{10} = \frac{4-y+1}{10} = \frac{5-y}{10}$$

Calculamos las esperanzas de X e Y :

$$E(X) = \sum_{x \in R_X} xp_X(x) = \sum_{x=1}^4 \frac{x^2}{10} = \frac{1}{10} + \frac{4}{10} + \frac{9}{10} + \frac{16}{10} = 3$$

$$E(Y) = \sum_{y \in R_Y} yp_Y(y) = \sum_{y=1}^4 \frac{y(5-y)}{10} = \frac{4}{10} + \frac{6}{10} + \frac{6}{10} + \frac{4}{10} = 2$$

Calculamos la esperanza de XY :

$$E(XY) = \sum_{x \in R_X} \sum_{y \in R_Y} xyP(X=x, Y=y) = \sum_{x=1}^4 \sum_{y=1}^x \frac{xy}{10} = \frac{1}{10} + \frac{2}{10} + \frac{4}{10} + \frac{3}{10} + \frac{6}{10} + \frac{9}{10} + \frac{4}{10} + \frac{8}{10} + \frac{12}{10} + \frac{16}{10} = 6.5$$

Calculamos las esperanzas de X^2 e Y^2 :

$$E(X^2) = \sum_{x \in R_X} x^2 P_X(x) = \sum_{x=1}^4 \frac{x^3}{10} = \frac{1}{10} + \frac{8}{10} + \frac{27}{10} + \frac{64}{10} = 10$$

$$E(Y^2) = \sum_{y \in R_Y} y^2 P_Y(y) = \sum_{y=1}^4 \frac{y^2(5-y)}{10} = \frac{4}{10} + \frac{12}{10} + \frac{18}{10} + \frac{16}{10} = 5$$

Calculamos las varianzas y los desvíos estándar de X e Y :

$$V(X) = E(X^2) - E(X)^2 = 10 - 3^2 = 1 \implies \sigma_X = \sqrt{V(X)} = 1$$

$$V(Y) = E(Y^2) - E(Y)^2 = 5 - 2^2 = 1 \implies \sigma_Y = \sqrt{V(Y)} = 1$$

Calculamos la covarianza:

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 6.5 - 3 * 2 = 0.5$$

Finalmente calculamos el coeficiente de correlación:

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{0.5}{1 * 1} = 0.5$$

Para decidir si X e Y son independientes, observemos primero que si fuesen independientes, entonces valdría que $E(XY) = E(X)E(Y) \implies \text{Cov}(X, Y) = 0$. Pero como $\text{Cov}(X, Y) = 0.5 \neq 0$ podemos afirmar entonces que **no** son independientes.

6 Esperanza condicional

6.1 Ejercicio 11

Se lanza un dado un equilibrado. Notamos X el número obtenido e $Y = 1$ si el número es par, $Y = 0$ sino. Calcule $E(X|Y = 1)$ y $E(Y|X \leq 4)$.

Intuitivamente podemos decir que $E(X|Y = 1) = 4$ pues al condicionar $Y = 1$ nos estamos quedando únicamente con los números pares: 2, 4, 6 y al ser un dado equilibrado, la probabilidad de obtener cualquiera de estos valores es la misma, por lo tanto la esperanza resulta el promedio de los valores posibles: $\frac{2+4+6}{3} = 4$.

Hagamos las cuentas. Primero armamos la tabla de probabilidad conjunta:

| Y/X | 1 | 2 | 3 | 4 | 5 | 6 | pY |
|-----|-----|-----|-----|-----|-----|-----|-----|
| 0 | 1/6 | 0 | 1/6 | 0 | 1/6 | 0 | 1/2 |
| 1 | 0 | 1/6 | 0 | 1/6 | 0 | 1/6 | 1/2 |
| pX | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1 |

Buscamos la función de probabilidad condicional de $X|Y = 1$:

$$p_{X|Y=1}(x) = \frac{p_{XY}(x, 1)}{p_Y(1)} = \frac{p_{XY}(x, 1)}{\frac{1}{2}} = 2p_{XY}(x, 1)$$

Ahora podemos calcular la esperanza:

$$E(X|Y = 1) = \sum_{x \in R_X} xp_{X|Y=1}(x) = \sum_{x=1}^6 2xp_{XY}(x, 1) = 0 + \frac{4}{6} + 0 + \frac{8}{6} + 0 + \frac{12}{6} = 4$$

Con la misma intuición que antes, podemos decir que $E(Y|X \leq 4) = 0.5$ pues al condicionar $X \leq 4$ nos estamos quedando únicamente con los números: 1, 2, 3, 4 de los cuales hay 2 pares y 2 impares. Como la v.a. Y asigna 0 o 1 dependiendo si es par o impar, al calcular el promedio obtenemos $\frac{0+1+0+1}{4} = \frac{1}{2}$.

Buscamos la función de probabilidad condicional de $Y|X \leq 4$:

$$p_{Y|X \leq 4}(y) = \frac{\sum_{x=1}^4 p_{XY}(x, y)}{\sum_{x=1}^4 p_X(x)}$$

El numerador depende de y por lo tanto no podemos simplificarlo aún, pero el denominador sí ya podemos calcularlo mirando la tabla:

$$\sum_{x=1}^4 p_X(x) = \frac{4}{6} = \frac{2}{3}$$

Por lo tanto:

$$p_{Y|X \leq 4}(y) = \frac{\sum_{x=1}^4 p_{XY}(x, y)}{\frac{2}{3}} = \frac{3}{2} \sum_{x=1}^4 p_{XY}(x, y)$$

Ahora podemos calcular la esperanza:

$$E(Y|X \leq 4) = \sum_{y \in R_Y} y p_{Y|X \leq 4}(y) = \sum_{y=0}^1 \left(y \frac{3}{2} \sum_{x=1}^4 p_{XY}(x, y) \right) = \frac{3}{2} \sum_{x=1}^4 p_{XY}(x, 1) = \frac{3}{2} \left(0 + \frac{1}{6} + 0 + \frac{1}{6} \right) = \frac{1}{2}$$

6.2 Ejercicio 12

Sea (X, Y) un vector aleatorio con función de densidad:

$$f_{XY}(x, y) = \lambda^2 e^{-\lambda y} I_{\{0 < x < y\}}$$

Qué distribución tiene $X|Y = y$ para $y > 0$?

Primero calculamos la función de densidad marginal de Y :

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx = \int_0^y \lambda^2 e^{-\lambda y} dx = \left[x \lambda^2 e^{-\lambda y} \right]_0^y = y \lambda^2 e^{-\lambda y} I_{\{0 < y\}}$$

Luego planteamos la función de densidad condicional:

$$f_{X|Y=y}(x, y) = \frac{f_{XY}(x, y)}{f_Y(y)} = \frac{\lambda^2 e^{-\lambda y} I_{\{0 < x < y\}}}{y \lambda^2 e^{-\lambda y} I_{\{0 < y\}}} = \frac{1}{y} I_{\{0 < x < y\}}$$

Como y está fijo $\implies (X|Y = y) \sim \mathcal{U}[0, y]$

7 Generación de variables aleatorias

7.1 Ejercicio 13

Sea F la función de distribución acumulada de una distribución de probabilidad f . Supongamos para simplificar que F es una biyección de \mathbb{R} en $(0, 1)$. Demuestre que si $U \sim \mathcal{U}[0, 1]$, entonces $Y := F^{-1}(U)$ tiene distribución f .

$$F_Y(t) = P(Y \leq t) = P(F^{-1}(U) \leq t)$$

Como F es una biyección en $(0, 1)$ vale que $F(F^{-1}(U)) = U$ siendo $U \sim \mathcal{U}[0, 1]$ y además al ser una función de distribución acumulada, es una función creciente y por lo tanto se mantiene la dirección de la desigualdad:

$$F_Y(t) = P(F(F^{-1}(U)) \leq F(t)) = P(U \leq F(t)) = F_U(F(t))$$

Derivamos para buscar la función de densidad. Observemos que por definición $F : \mathbb{R} \rightarrow [0, 1]$ y como U es una uniforme estándar, $f_U(u) = 1 \ \forall u \in [0, 1]$:

$$f_Y(t) = f_U(F(t))f(t) = 1f(t) = f(t)$$

7.2 Ejercicio 14

Sea $U \sim \mathcal{U}[0, 1]$. Demuestre que $\frac{-\log(1-U)}{\lambda} \sim \mathcal{E}(\lambda)$.

Asumimos que $\log = \ln$.

$$X = \frac{-\ln(1-U)}{\lambda}$$

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P\left(\frac{-\ln(1-U)}{\lambda} \leq x\right) \\ &= P(-\ln(1-U) \leq \lambda x) \\ &= P(\ln(1-U) \geq -\lambda x) \\ &= P(1-U \geq e^{-\lambda x}) \\ &= P(U \leq 1 - e^{-\lambda x}) \\ &= F_U(1 - e^{-\lambda x}) \end{aligned}$$

Busquemos el R_X :

$$R_U = [0, 1] \implies \begin{cases} \lim_{U \rightarrow 0} X = 0 \\ \lim_{U \rightarrow 1} X = +\infty \end{cases} \implies R_X = [0, +\infty]$$

Por lo tanto:

$$\begin{cases} \lim_{x \rightarrow 0} 1 - e^{-\lambda x} = 0 \\ \lim_{x \rightarrow +\infty} 1 - e^{-\lambda x} = 1 \end{cases} \implies 1 - e^{-\lambda x} \in [0, 1] \ \forall x \in R_X$$

Como $F_U(u) = u \ \forall 0 \leq u \leq 1$ por ser la acumulada de una variable con distribución uniforme estándar:

$$F_X(x) = F_U(1 - e^{-\lambda x}) = 1 - e^{-\lambda x}$$

Ahora derivamos para obtener la función de densidad de X :

$$f_X(x) = \frac{\partial}{\partial x} F_X(x) = \lambda e^{-\lambda x} \implies X \sim \mathcal{E}(\lambda)$$

8 Función característica

8.1 Ejercicio 15

Sea X_1, X_2, \dots v.a.i.i.d. distribución exponencial $\mathcal{E}(\lambda)$, y N una variable aleatoria geométrica independiente de los X_i . Calcular la función característica de $Z := \sum_{i=1}^N X_i$ y deducir la distribución de Z .

9 Convergencia de variables aleatorias

9.1 Ejercicio 16

Sea $U_n \sim \text{Uniforme}\{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\}$. Calcule el límite en distribución de U_n cuando $n \rightarrow \infty$.

Sea $R_{U_n} = \{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\}$ el rango de U_n .

Observemos que el conjunto infinito numerable R_{U_n} está contenido en el intervalo continuo $[0, 1]$:

$$\begin{aligned} \min(R_{U_n}) &= \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0 \\ \max(R_{U_n}) &= \frac{n}{n} = 1 \quad \forall n \neq 0 \end{aligned}$$

Sea $F : \mathbb{R} \rightarrow \mathbb{R}$ la función de distribución acumulada de U_n cuando $n \rightarrow \infty$. Por lo explicado anteriormente:

$$F(x) = P(U_n \leq x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x > 1 \end{cases}$$

Veamos qué pasa cuando $x \in [0, 1]$.

Como U_n es una v.a. discreta, su acumulada va a ser una sumatoria y por eso necesitamos encontrar una forma de enumerar todos los elementos del conjunto R_{U_n} .

$$\forall k \in \mathbb{N}, 1 \leq k \leq n, \frac{k}{n} \in R_{U_n} \implies \forall x \in \mathbb{R}, 0 \leq x \leq 1, \frac{k}{n} \leq x \iff k \leq [xn]$$

Como $\#R_{U_n} = n$ y U_n tiene distribución uniforme discreta, la probabilidad puntual está dada por:

$$P(U_n = \frac{k}{n}) = \frac{1}{n} \quad \forall 1 \leq k \leq n$$

Planteamos la función de distribución acumulada de U_n :

$$F(x) = P(U_n \leq x) = \sum_{k=1}^{[xn]} P(U_n = \frac{k}{n}) = \sum_{k=1}^{[xn]} \frac{1}{n} = ([xn] - 1 + 1) \frac{1}{n} = \frac{[xn]}{n}$$

Veamos qué pasa cuando $n \rightarrow \infty$.

$$F(x) = \frac{[xn]}{n} = \frac{[xn]}{n} \frac{xn}{xn} = \frac{[xn]}{xn} \frac{xn}{n} = \frac{[xn]}{xn} x \xrightarrow{n \rightarrow \infty} x$$

Recordemos que $\frac{[xn]}{xn} \xrightarrow{n \rightarrow \infty} 1$ pues:

$$xn - 1 \leq [xn] \leq xn \iff \frac{xn - 1}{xn} \leq \frac{[xn]}{xn} \leq \frac{xn}{xn} \iff 1 - \frac{1}{xn} \leq \frac{[xn]}{xn} \leq 1 \xrightarrow{n \rightarrow \infty} 1 \leq \frac{[xn]}{xn} \leq 1 \implies \frac{[xn]}{xn} = 1$$

Por lo tanto, $F(x) = x \forall x \in \mathbb{R}, 0 \leq x \leq 1$ cuando $n \rightarrow \infty$. Armemos entonces la función de distribución acumulada con todos los casos:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases} \implies U_n \sim \mathcal{U}[0, 1] \text{ cuando } n \rightarrow \infty$$

9.2 Ejercicio 17

Enunciar y demostrar las desigualdades de Markov y Tchebyshev.

Desigualdad de Markov

Sea X una v.a. cualquiera que solo toma valores no negativos. Entonces:

$$\forall \epsilon > 0, P(X \geq \epsilon) \leq \frac{E(X)}{\epsilon}$$

Demostración:

Dado $\epsilon > 0$, definimos una v.a.

$$I = \begin{cases} 1 & \text{si } X \geq \epsilon \\ 0 & \text{caso contrario} \end{cases}$$

Notemos que $I \leq \frac{X}{\epsilon}$ pues:

$$\begin{aligned}\epsilon \leq X &\implies 1 \leq \frac{X}{\epsilon} \wedge I = 1 \implies I \leq \frac{X}{\epsilon} \\ 0 \leq X < \epsilon &\implies 0 \leq \frac{X}{\epsilon} < 1 \wedge I = 0 \implies I \leq \frac{X}{\epsilon}\end{aligned}$$

Ya que I solo toma valores discretos 0 o 1.

Ahora tomamos esperanza de ambos lados:

$$I \leq \frac{X}{\epsilon} \implies E(I) \leq E\left(\frac{X}{\epsilon}\right) = \frac{E(X)}{\epsilon}$$

Finalmente veamos que I es un ensayo de Bernoulli con probabilidad $P(X \geq \epsilon)$. Por lo tanto:

$$E(I) = P(X \geq \epsilon) \leq \frac{E(X)}{\epsilon}$$

Desigualdad de Tchebyshev

Sea X una v.a. con $E(X) = \mu$ y $V(X) = \sigma^2 < \infty$. Entonces:

$$\forall \epsilon > 0, P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$$

Demostración:

Dado que $(X - \mu)^2$ es una v.a. no negativa, podemos aplicar la desigualdad de Markov:

$$\forall \epsilon > 0, P((X - \mu)^2 \geq \epsilon^2) \leq \frac{E((X - \mu)^2)}{\epsilon^2}$$

Teniendo en cuenta que $(X - \mu)^2 \geq \epsilon^2 \iff |X - \mu| \geq \epsilon$:

$$\forall \epsilon > 0, P(|X - \mu| \geq \epsilon) \leq \frac{E((X - \mu)^2)}{\epsilon^2} = \frac{V(X)}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}$$

9.3 Ejercicio 18

Enuncie y demuestre la Ley de los Grandes Números para variables aleatorias con segundo momento finito (admitiendo Tchebyshev).

Sean X_1, \dots, X_n v.a.i.i.d. (muestra aleatoria) con $E(X) = \mu$ y $V(X) = \sigma^2 < \infty$. Sea $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ el promedio muestral. Entonces:

$$\overline{X}_n \xrightarrow{p} \mu$$

Es decir:

$$\forall \epsilon > 0 \quad \lim_{n \rightarrow \infty} P\left(\left| \overline{X}_n - \mu \right| \geq \epsilon \right) = 0$$

Demostración:

Veamos la esperanza y varianza de \overline{X}_n :

$$\begin{aligned} E(\overline{X}_n) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) \stackrel{X_i \text{ i.i.d.}}{=} \frac{nE(X)}{n} = E(X) = \mu \\ V(\overline{X}_n) &= V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) \stackrel{X_i \text{ i.i.d.}}{=} \frac{nV(X)}{n^2} = \frac{V(X)}{n} = \frac{\sigma^2}{n} \end{aligned}$$

Notemos que $V(X) < \infty$ pues en el enunciado nos dicen que la v.a. tiene segundo momento finito:

$$E(X^2) < \infty \implies E(X) < \infty \implies V(X) = E(X^2) - E(X)^2 < \infty$$

Aplicamos entonces la desigualdad de Tchebyshev:

$$\forall \epsilon > 0, \quad P\left(\left| \overline{X}_n - E(\overline{X}_n) \right| \geq \epsilon \right) \leq \frac{V(\overline{X}_n)}{\epsilon^2} \iff P\left(\left| \overline{X}_n - \mu \right| \geq \epsilon \right) \leq \frac{\sigma^2}{n\epsilon^2}$$

Tomamos límite cuando $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} P\left(\left| \overline{X}_n - \mu \right| \geq \epsilon \right) \leq \lim_{n \rightarrow \infty} \frac{\sigma^2}{n\epsilon^2} = 0$$

Luego $\overline{X}_n \xrightarrow{p} \mu$ como queríamos probar.

9.4 Ejercicio 19

Sea X_n una variable Poisson(λn). Demuestre que $\frac{X_n}{n}$ converge en probabilidad a λ .

$$\begin{aligned} E(X_n) &= \lambda n \implies E\left(\frac{X_n}{n}\right) = \frac{E(X_n)}{n} = \frac{\lambda n}{n} = \lambda \\ V(X_n) &= \lambda n \implies V\left(\frac{X_n}{n}\right) = \frac{V(X_n)}{n^2} = \frac{\lambda n}{n^2} = \frac{\lambda}{n} \end{aligned}$$

Por Tchebyshev:

$$\forall \epsilon > 0, P\left(\left|\frac{X_n}{n} - E\left(\frac{X_n}{n}\right)\right| \geq \epsilon\right) \leq \frac{V\left(\frac{X_n}{n}\right)}{\epsilon^2} \implies P\left(\left|\frac{X_n}{n} - \lambda\right| \geq \epsilon\right) \leq \frac{\lambda}{n\epsilon^2}$$

Nuestra medida de probabilidad $\in [0, 1]$ por definición. Entonces, tomando límite cuando $n \rightarrow \infty$:

$$0 \leq P\left(\left|\frac{X_n}{n} - \lambda\right| \geq \epsilon\right) \leq \frac{\lambda}{n\epsilon^2} \xrightarrow{n \rightarrow \infty} 0 \leq P\left(\left|\frac{X_n}{n} - \lambda\right| \geq \epsilon\right) \leq 0 \implies \lim_{n \rightarrow \infty} P\left(\left|\frac{X_n}{n} - \lambda\right| \geq \epsilon\right) = 0$$

Lo cual es exactamente la definición de convergencia en probabilidad. Por lo tanto, $\frac{X_n}{n}$ converge en probabilidad a λ .

9.5 Ejercicio 20

Sean X_1, X_2, \dots v.a. i.i.d. con función de distribución exponencial de parámetro λ y sean Y_1, Y_2, \dots v.a. i.i.d. e independientes de las X_i con distribución uniforme en $(0, 1)$. Considerar las variables:

$$Z_n = \frac{1}{n} \sum_{i=1}^n I_{\{X_i \leq 1\}} I_{\{Y_i \leq 0.5\}}$$

Hallar la esperanza y la varianza de Z_n , y el valor del límite en probabilidad de Z_n .

Observemos que las indicadoras evalúan a 0 o 1 en función de una probabilidad de X_i e Y_i . Por lo tanto, son en sí mismas v.a. con distribución Bernoulli.

$$I_{\{X_i \leq 1\}} \sim \text{Bernoulli}(p_X) \text{ con } p_X = P(X_i \leq 1) = 1 - e^{-\lambda} \text{ pues } X_i \sim \mathcal{E}(\lambda)$$

$$I_{\{Y_i \leq 0.5\}} \sim \text{Bernoulli}(p_Y) \text{ con } p_Y = P(Y_i \leq 0.5) = 0.5 \text{ pues } Y_i \sim \mathcal{U}[0, 1]$$

Calculemos la esperanza:

$$E(Z_n) = E\left(\frac{1}{n} \sum_{i=1}^n I_{\{X_i \leq 1\}} I_{\{Y_i \leq 0.5\}}\right)$$

Por la linealidad de la esperanza:

$$E(Z_n) = \frac{1}{n} \sum_{i=1}^n E(I_{\{X_i \leq 1\}} I_{\{Y_i \leq 0.5\}})$$

Porque X_i e Y_i son independientes entre sí:

$$E(Z_n) = \frac{1}{n} \sum_{i=1}^n E(I_{\{X_i \leq 1\}})E(I_{\{Y_i \leq 0.5\}})$$

Porque X_i e Y_i son i.i.d.:

$$E(Z_n) = \frac{1}{n} n E(I_{\{X_i \leq 1\}})E(I_{\{Y_i \leq 0.5\}}) = E(I_{\{X_i \leq 1\}})E(I_{\{Y_i \leq 0.5\}})$$

Calculemos las esperanzas individuales. Recordemos que las indicadoras son v.a. Bernoulli, y la esperanza de una Bernoulli de parámetro p es p :

$$\begin{aligned} E(I_{\{X_i \leq 1\}}) &= p_X = 1 - e^{-\lambda} \\ E(I_{\{Y_i \leq 0.5\}}) &= p_Y = 0.5 \end{aligned}$$

Finalmente:

$$E(Z_n) = p_X p_Y = (1 - e^{-\lambda})0.5 = 0.5 - 0.5e^{-\lambda}$$

Ahora calculemos la varianza:

$$V(Z_n) = V\left(\frac{1}{n} \sum_{i=1}^n I_{\{X_i \leq 1\}} I_{\{Y_i \leq 0.5\}}\right) = \frac{1}{n^2} \sum_{i=1}^n V(I_{\{X_i \leq 1\}} I_{\{Y_i \leq 0.5\}}) = \frac{1}{n} V(I_{\{X_i \leq 1\}} I_{\{Y_i \leq 0.5\}})$$

La varianza de una Bernoulli de parámetro p es $p(1-p)$, y la varianza del producto de 2 Bernoulli independientes con parámetros p_X y p_Y es $p_X p_Y (1 - p_X p_Y)$. Por lo tanto:

$$\begin{aligned} V(Z_n) &= \frac{1}{n} (p_X p_Y (1 - p_X p_Y)) \\ &= \frac{1}{n} ((0.5 - 0.5e^{-\lambda})(1 - (0.5 - 0.5e^{-\lambda}))) \\ &= \frac{1}{n} ((0.5 - 0.5e^{-\lambda})(0.5 + 0.5e^{-\lambda})) \\ &= \frac{1}{n} (0.5^2 - (0.5e^{-\lambda})^2) \\ &= \frac{1}{n} (0.25 - 0.25e^{-2\lambda}) \end{aligned}$$

Dado que Z_n es el promedio de v.a.i.i.d. con varianza finita, por la Ley de los Grandes Números:

$$Z_n \xrightarrow{p} E(Z) = E(I_{\{X \leq 1\}} I_{\{Y \leq 0.5\}}) = 0.5 - 0.5e^{-\lambda}$$

Demostración auxiliar de que el producto de 2 Bernoulli independientes de parámetros p y q es una Bernoulli de parámetro pq .

Sean $A \sim \text{Bernoulli}(p)$, $B \sim \text{Bernoulli}(q)$ y $C = A * B$.

$$P(C = 1) = P(A = 1, B = 1) = P(A = 1)P(B = 1) = pq$$

$$\begin{aligned} P(C = 0) &= P(A = 0, B = 0) + P(A = 1, B = 0) + P(A = 0, B = 1) \\ &= P(A = 0)P(B = 0) + P(A = 1)P(B = 0) + P(A = 0)P(B = 1) \\ &= (1 - p)(1 - q) + p(1 - q) + (1 - p)q \\ &= 1 - q - p + pq + p - pq + q - pq \\ &= 1 - pq \end{aligned}$$

Por lo tanto $\implies C \sim \text{Bernoulli}(pq)$

9.6 Ejercicio 21

Enuncie el Teorema Central del Límite. Sea $Y_n \sim \mathcal{P}(\lambda n)$. Demuestre que $\frac{Y_n - n\lambda}{\sqrt{n\lambda}}$ converge en distribución a una variable $\mathcal{N}(0, 1)$ cuando $n \rightarrow +\infty$.

Sean X_1, \dots, X_n variables aleatorias independientes e idénticamente distribuidas, con $E(X_i) = \mu$ y $V(X_i) = \sigma^2 < \infty \forall i$. Entonces, si n es suficientemente grande:

$$\frac{T - n\mu}{\sqrt{n\sigma}} \xrightarrow{d} Z \sim \mathcal{N}(0, 1)$$

$$\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \xrightarrow{d} Z \sim \mathcal{N}(0, 1)$$

Siendo $T = \sum_{i=1}^n X_i$ y $\bar{X} = \frac{T}{n}$. Ambas expresiones son equivalentes.

Sean Y_1, \dots, Y_n v.a.i.i.d. tales que $Y_i \sim \mathcal{P}(\lambda) \forall i$. Si $Y_n = \sum_{i=1}^n Y_i \implies Y_n \sim \mathcal{P}(n\lambda)$.

Sabiendo que $E(Y_i) = V(Y_i) = \lambda \forall i$, por el TCL previamente enunciado resulta inmediato que:

$$\frac{Y_n - nE(Y_i)}{\sqrt{nV(Y_i)}} = \frac{Y_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow{d} Z \sim \mathcal{N}(0, 1)$$

Demostración auxiliar de que $Y_n \sim \mathcal{P}(n\lambda)$ utilizando la función característica:

$$Y_n = \sum_{i=1}^n Y_i \implies \Phi_{Y_n}(t) = \prod_{i=1}^n \Phi_{Y_i}(t) \stackrel{Y_i \text{ i.i.d.}}{=} (\Phi_{Y_1}(t))^n = \left(e^{\lambda(e^{it}-1)}\right)^n = e^{n\lambda(e^{it}-1)} = \Phi_{\mathcal{P}(n\lambda)}(t) \implies Y_n \sim \mathcal{P}(n\lambda)$$

10 Procesos de Poisson

10.1 Ejercicio 22

Los clientes llegan a un banco de acuerdo a un proceso de Poisson de parámetro 3 por minuto, cuál es la probabilidad que los tres primeros clientes lleguen antes de los 3 primeros minutos?

Sea $X \sim \mathcal{P}(9)$ un proceso de Poisson que modela la llegada de clientes con una tasa de 9 clientes cada 3 minutos.

Queremos ver la probabilidad de que $X \geq 3$, es decir, que para el minuto 3 ya hayan llegado al menos 3 clientes.

$$\begin{aligned} P(X \geq 3) &= 1 - P(X \leq 2) \\ &= 1 - \sum_{x=0}^2 P(X = x) \\ &= 1 - \sum_{x=0}^2 \frac{9^x e^{-9}}{x!} \\ &= 1 - e^{-9} \left(1 + 9 + \frac{81}{2}\right) \approx 0.9938 \end{aligned}$$

10.2 Ejercicio 23

Consideramos un proceso de Poisson de tasa λ . Calcule la probabilidad que ocurran simultáneamente los eventos A y B.

$$A = \{\text{hay 3 llegadas en el intervalo } [0, 3]\}$$

$$B = \{\text{hay 4 llegadas en el intervalo } (3, 4]\}$$

Sea $X \sim \mathcal{P}(\lambda)$ v.a. que mide la cantidad de llegadas por una unidad del intervalo.

Observemos que la probabilidad de ocurrencia de 2 eventos en intervalos disjuntos de un proceso de Poisson son independientes, es decir $P(A \cap B) = P(A)P(B)$.

Además, por la propiedad de “falta de memoria”, la probabilidad de llegada en cierto intervalo es la misma para cualquier intervalo de la misma longitud. Entonces la probabilidad en el intervalo $(3, 4]$ es la misma que en el intervalo $[0, 1]$.

Por lo tanto planteamos las siguientes v.a. para calcular la probabilidad de cada evento:

Sean $X_A \sim \mathcal{P}(3\lambda)$, $X_B \sim \mathcal{P}(\lambda)$ v.a. que miden la cantidad de llegadas en los intervalos $[0, 3]$ y $(3, 4]$ respectivamente.

$$P(A \cap B) = P(A)P(B) = P(X_A = 3)P(X_B = 4) = \frac{(3\lambda)^3 e^{-3\lambda}}{3!} \frac{\lambda^4 e^{-\lambda}}{4!} = \frac{3\lambda^7 e^{-4\lambda}}{16}$$

11 Estimadores

11.1 Ejercicio 24

Hallar el estimador de momentos de θ para la distribución Uniforme $[-\theta, \theta]$.

Sea X_1, \dots, X_n una muestra aleatoria de una distribución $\mathcal{U}[-\theta, \theta]$

Si bien hay un único parámetro a estimar, en este caso no podemos plantear una ecuación basada en el primer momento ya que éste no depende de θ :

$$\overline{X} = E(X) = \frac{-\theta + \theta}{2} = 0$$

Por lo tanto planteamos la ecuación basada en el segundo momento:

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

Calculemos $E(X^2)$:

$$V(X) = E(X^2) - E(X)^2 \iff E(X^2) = V(X) + E(X)^2 = \frac{(\theta - (-\theta))^2}{12} + \frac{-\theta + \theta}{2} = \frac{\theta^2}{3}$$

Luego:

$$\overline{X^2} = E(X^2) = \frac{\theta^2}{3} \iff \hat{\theta} = \sqrt{3\overline{X^2}}$$

11.2 Ejercicio 25

Hallar el estimador de máxima verosimilitud de p para la variable $X \sim \text{Bernoulli}(p)$.

Veamos la función de densidad de X :

$$f_X(k) = p^k (1-p)^{1-k} \quad k \in \{0, 1\}$$

Planteamos la función de máxima verosimilitud:

$$L(p) = f_X(x_1, \dots, x_n) = \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} = p^{\sum_{i=1}^n x_i} (1-p)^{\sum_{i=1}^n 1-x_i}$$

Le tomamos logaritmo para simplificar las cuentas. Esto vale porque la función logaritmo es monótona creciente y por lo tanto el argumento que maximiza $\ln(L(p))$ es el mismo para $L(p)$. Es decir, ambas funciones alcanzan su máximo en los mismos puntos.

$$\ln(L(p)) = \left(\sum_{i=1}^n x_i \right) \ln(p) + \left(\sum_{i=1}^n 1 - x_i \right) \ln(1-p)$$

Ahora derivamos para encontrar el punto crítico:

$$\begin{aligned} \frac{\partial}{\partial p} \ln(L(p)) &= \frac{1}{p} \left(\sum_{i=1}^n x_i \right) - \frac{1}{1-p} \left(\sum_{i=1}^n 1 - x_i \right) = 0 \\ \iff \frac{1}{p} \left(\sum_{i=1}^n x_i \right) &= \frac{1}{1-p} \left(\sum_{i=1}^n 1 - x_i \right) \\ \iff \frac{1}{p} \left(\sum_{i=1}^n x_i \right) &= \frac{1}{1-p} \left(n - \sum_{i=1}^n x_i \right) \\ \iff \frac{1-p}{p} &= \frac{n - \sum_{i=1}^n x_i}{\sum_{i=1}^n x_i} \\ \iff \frac{1}{p} - 1 &= \frac{n}{\sum_{i=1}^n x_i} - 1 \\ \iff \frac{1}{p} &= \frac{n}{\sum_{i=1}^n x_i} \\ \iff p &= \frac{\sum_{i=1}^n x_i}{n} \\ \iff p &= \bar{X} \end{aligned}$$

Encontramos un punto crítico en $p = \bar{X}$. Para confirmar que es el máximo absoluto, veamos los límites hacia los extremos del dominio de la función $\ln(L(p))$ el cual está definido en $p \in [0, 1]$.

$$\begin{aligned} \lim_{p \rightarrow 0} \ln(L(p)) &= \lim_{p \rightarrow 0} \left(\left(\sum_{i=1}^n x_i \right) \ln(p) + \left(\sum_{i=1}^n 1 - x_i \right) \ln(1-p) \right) = -\infty + 0 = -\infty \\ \lim_{p \rightarrow 1} \ln(L(p)) &= \lim_{p \rightarrow 1} \left(\left(\sum_{i=1}^n x_i \right) \ln(p) + \left(\sum_{i=1}^n 1 - x_i \right) \ln(1-p) \right) = 0 - \infty = -\infty \end{aligned}$$

Entonces en efecto $p = \bar{X}$ es el argumento que maximiza $L(p)$ ya que es el único punto crítico encontrado y ambos extremos del dominio tienden a $-\infty$. Así el estimador de máxima verosimilitud resulta:

$$\hat{p}_{MV} = \overline{X}$$

11.3 Ejercicio 26

Hallar el estimador de máxima verosimilitud del parámetro a de la siguiente distribución donde $a > 1$. Es consistente?

$$f(x) = \begin{cases} ax^{-(a+1)} & \text{si } x \geq 1 \\ 0 & \text{en otro caso} \end{cases}$$

Sea X_1, \dots, X_n una muestra aleatoria con distribución $f(x)$ y asumimos que son i.i.d.

Planteamos la función de máxima verosimilitud. Asumimos que $x_i \geq 1 \forall i$ pues sino $f(x_i) = 0$.

$$L(a) = f(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i) = \prod_{i=1}^n ax_i^{-(a+1)}$$

Le tomamos logaritmo para simplificar las cuentas. Esto vale porque la función logaritmo es monótona creciente y por lo tanto el argumento que maximiza $\ln(L(a))$ es el mismo para $L(a)$.

$$\begin{aligned} \ln(L(a)) &= \ln\left(\prod_{i=1}^n ax_i^{-(a+1)}\right) \\ &= \sum_{i=1}^n \ln(ax_i^{-(a+1)}) \\ &= \sum_{i=1}^n \left(\ln(a) + \ln(x_i^{-(a+1)})\right) \\ &= \left(\sum_{i=1}^n \ln(a)\right) + \left(\sum_{i=1}^n -(a+1)\ln(x_i)\right) \\ &= n\ln(a) - (a+1) \sum_{i=1}^n \ln(x_i) \end{aligned}$$

Derivamos e igualamos a 0 para buscar los puntos críticos:

$$\begin{aligned} \frac{\partial}{\partial a} \ln(L(a)) &= \frac{n}{a} - \sum_{i=1}^n \ln(x_i) = 0 \\ \iff \frac{n}{a} &= \sum_{i=1}^n \ln(x_i) = n\overline{\ln(X)} \\ \iff \hat{a} &= \frac{n}{n\overline{\ln(X)}} = \frac{1}{\overline{\ln(X)}} \end{aligned}$$

Verifiquemos que el argumento encontrado es el que maximiza la función mediante el criterio de la segunda derivada:

$$\frac{\partial^2}{\partial a^2} \ln(L(a)) = \frac{-n}{a^2}$$

Como $n > 0$ y $a^2 > 0$ la segunda derivada es siempre negativa, por lo tanto el argumento encontrado es el máximo.

Veamos si es consistente. Por definición, el estimador \hat{a} es consistente si:

$$\forall \epsilon > 0, P(|\hat{a} - a| > \epsilon) \xrightarrow{n \rightarrow \infty} 0$$

Por la Ley de los Grandes Números:

$$\hat{a} = \frac{1}{\ln(X)} \xrightarrow{n \rightarrow \infty} \frac{1}{E(\ln(X))}$$

Calculamos $E(\ln(X))$ (la integral se resuelve por partes):

$$E(\ln(X)) = \int_1^\infty \ln(x) f(x) dx = \int_1^\infty \ln(x) a x^{-(a+1)} dx = \frac{1}{a}$$

Por lo tanto:

$$\hat{a} \xrightarrow{n \rightarrow \infty} \frac{1}{\frac{1}{a}} = a$$

Finalmente, \hat{a} es consistente pues:

$$\forall \epsilon > 0, P(|\hat{a} - a| > \epsilon) \xrightarrow{n \rightarrow \infty} P(|a - a| > \epsilon) = P(0 > \epsilon) = 0$$

11.4 Ejercicio 27

Sea X_1, \dots, X_n una muestra de la distribución:

$$f(x; \theta) = e^{-(x-\theta)} I_{\{x \geq \theta\}}$$

Hallar el estimador de máxima verosimilitud de θ . Es consistente?

Planteamos la función de máxima verosimilitud:

$$\begin{aligned}
L(\theta) &= f(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i) = \prod_{i=1}^n e^{-(x_i-\theta)} I_{\{x_i \geq \theta\}} = \prod_{i=1}^n e^{-(x_i-\theta)} \prod_{i=1}^n I_{\{x_i \geq \theta\}} \\
&= e^{\sum_{i=1}^n -(x_i-\theta)} \prod_{i=1}^n I_{\{x_i \geq \theta\}} = e^{\sum_{i=1}^n (-x_i+\theta)} \prod_{i=1}^n I_{\{x_i \geq \theta\}} = e^{\sum_{i=1}^n -x_i} e^{\sum_{i=1}^n \theta} \prod_{i=1}^n I_{\{x_i \geq \theta\}} \\
&= e^{\sum_{i=1}^n -x_i} e^{n\theta} \prod_{i=1}^n I_{\{x_i \geq \theta\}}
\end{aligned}$$

Ahora queremos ver qué valores de θ maximizan $L(\theta)$. Analizamos cada término:

$e^{\sum_{i=1}^n -x_i}$ Es una constante, así que podemos ignorar este término.

$e^{n\theta} \xrightarrow{\theta \rightarrow \infty} \infty$ Esto nos dice que queremos el θ más grande posible.

Pero también necesitamos que la productoria de la indicadora de 1 para todos los X_i . Veamos cuándo pasa eso:

$$\prod_{i=1}^n I_{\{x_i \geq \theta\}} = 1 \iff \theta \leq X_i \forall 1 \leq i \leq n \iff \theta \leq \min_{1 \leq i \leq n} X_i$$

Por lo tanto el θ más grande que podemos elegir es el mínimo de todos los X_i y ese es el estimador de máxima verosimilitud:

$$\hat{\theta} = \min_{1 \leq i \leq n} X_i$$

Para ver la consistencia, veamos primero el sesgo del estimador y luego buscamos su varianza.

Sea $U = \min_{1 \leq i \leq n} X_i$

$$F_U(u) = P(U \leq u) = 1 - P(U \geq u) = 1 - P(\min_{1 \leq i \leq n} X_i \geq u)$$

Observación:

$$\min_{1 \leq i \leq n} X_i \geq u \implies X_i \geq u \forall 1 \leq i \leq n$$

Es decir, si el X_i mínimo es mayor que u , entonces todos los X_i son mayores que u .

$$\begin{aligned}
1 - P(\min_{1 \leq i \leq n} X_i \geq u) &= 1 - P(\cap_{i=1}^n X_i \geq u) \\
&= 1 - \prod_{i=1}^n P(X_i \geq u) = 1 - \prod_{i=1}^n (1 - P(X_i \leq u)) \\
&= 1 - \prod_{i=1}^n (1 - F_{X_i}(u)) = 1 - \prod_{i=1}^n (1 - F_X(u)) \\
&= 1 - (1 - F_X(u))^n
\end{aligned}$$

Para el desarrollo anterior usamos que todas las X_i son i.i.d. Recapitulando:

$$F_U(u) = 1 - (1 - F_X(u))^n$$

No sabemos cómo es F_X , calculemos eso antes de seguir:

$$F_X(t) = \int_{\theta}^t f(x)dx = \int_{\theta}^t e^{-(x-\theta)}dx = 1 - e^{\theta-t}$$

Ahora sí podemos seguir. Derivamos ambos lados de la igualdad para obtener f_U :

$$f_U(u) = \frac{\partial}{\partial u} F_U(u) = n(1 - F_X(u))^{n-1} f_X(u) = n(e^{\theta-u})^{n-1} e^{-(u-\theta)} = ne^{n(\theta-u)}$$

Finalmente podemos calcular la esperanza de $\hat{\theta}$:

$$E(\hat{\theta}) = E(U) = \int_{\theta}^{\infty} u f_U(u) du = \int_{\theta}^{\infty} u n e^{n(\theta-u)} du = \theta + \frac{1}{n}$$

Veamos que $\hat{\theta}$ es asintóticamente insesgado pues:

$$E(\hat{\theta}) = \theta + \frac{1}{n} \xrightarrow{n \rightarrow \infty} \theta$$

Calculemos $E(\hat{\theta}^2)$ para poder luego calcular la varianza:

$$E(\hat{\theta}^2) = E(U^2) = \int_{\theta}^{\infty} u^2 f_U(u) du = \int_{\theta}^{\infty} u^2 n e^{n(\theta-u)} du = \frac{\theta^2 n^2 + 2\theta n + 2}{n^2}$$

Ahora sí, calculemos la varianza de $\hat{\theta}$ y veamos si es consistente:

$$\begin{aligned}
V(\hat{\theta}) &= E(\hat{\theta}^2) - E(\hat{\theta})^2 = \frac{\theta^2 n^2 + 2\theta n + 2}{n^2} - \left(\theta + \frac{1}{n}\right)^2 \\
&= \theta^2 + \frac{2\theta}{n} + \frac{2}{n^2} - \theta^2 - \frac{2\theta}{n} - \frac{1}{n^2} \\
&= \frac{1}{n^2} \xrightarrow{n \rightarrow \infty} 0
\end{aligned}$$

Por lo tanto, como se cumplen estas 2 condiciones:

$$\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta \wedge \lim_{n \rightarrow \infty} V(\hat{\theta}) = 0 \implies \hat{\theta} \text{ es consistente}$$

11.5 Ejercicio 28

Defina el error cuadrático medio (ECM) de un estimador y demuestre que es la suma de la varianza mas el cuadrado del sesgo.

Sea $\hat{\theta}$ un estimador de θ . Su error cuadrático medio (ECM) es:

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

$$\begin{aligned}
V(\hat{\theta}) + \text{sesgo}(\hat{\theta})^2 &= E(\hat{\theta}^2) - E(\hat{\theta})^2 + (E(\hat{\theta}) - \theta)^2 \\
&= E(\hat{\theta}^2) - E(\hat{\theta})^2 + E(\hat{\theta})^2 - 2\theta E(\hat{\theta}) + \theta^2 \\
&= E(\hat{\theta}^2 - 2\theta\hat{\theta} + \theta^2) \\
&= E((\hat{\theta} - \theta)^2)
\end{aligned}$$

11.6 Ejercicio 29

Decir si los estimadores de momentos y de máxima verosimilitud de θ para $X \sim \mathcal{U}[0, \theta]$ son insesgados y/o asintóticamente insesgados.

Estimador de momentos

$$E(\hat{\theta}_M) = E(2\bar{X}) = 2E(\bar{X}) = 2E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{2}{n} \sum_{i=1}^n E(X_i) = \frac{2}{n} nE(X) = 2E(X) = 2\frac{0+\theta}{2} = \theta$$

$\implies \hat{\theta}_M$ es insesgado pues su esperanza coincide con el parámetro que estima.

Estimador de máxima verosimilitud

Recordemos que los estimadores son en sí mismos variables aleatorias. Entonces para calcular la esperanza de $\hat{\theta}_{MV}$ vamos a necesitar conocer su función de densidad. Y para eso primero busquemos la función acumulada y luego la derivamos.

Sea $U = \max_{1 \leq i \leq n} X_i$ una v.a. (podríamos llamarla $\hat{\theta}_{MV}$ también pero así quedá más claro para leer).

$$F_U(u) = P(U \leq u) = P(\max_{1 \leq i \leq n} X_i \leq u)$$

Observación:

$$\max_{1 \leq i \leq n} X_i \leq u \implies X_i \leq u \forall 1 \leq i \leq n$$

Es decir, si el X_i máximo es menor que u , entonces todos los X_i son menores que u :

$$P(\max_{1 \leq i \leq n} X_i \leq u) = P(\cap_{i=1}^n X_i \leq u) = \prod_{i=1}^n P(X_i \leq u) = \prod_{i=1}^n F_X(u) = F_X(u)^n$$

Para el desarrollo anterior usamos que todas las X_i son i.i.d. Recapitulando:

$$F_U(u) = F_X(u)^n$$

Como $X \sim \mathcal{U}(0, \theta)$, sabemos exactamente cómo es F_X y por lo tanto podemos plantear directamente F_U :

$$F_U(u) = \begin{cases} 0 & u < 0 \\ \left(\frac{u}{\theta}\right)^n & 0 \leq u \leq \theta \\ 1 & \theta < u \end{cases}$$

Muy bien, ahora que tenemos F_U podemos derivar para obtener f_U :

$$f_U(u) = \frac{\partial}{\partial u} F_U(u) = \frac{nu^{n-1}}{\theta^n} I_{[0, \theta]}(u)$$

Finalmente podemos calcular la esperanza de $\hat{\theta}_{MV}$:

$$\begin{aligned} E(\hat{\theta}_{MV}) &= E(U) = \int_0^\theta u f_U(u) du = \int_0^\theta u \frac{nu^{n-1}}{\theta^n} du = \frac{n}{\theta^n} \int_0^\theta u^n du = \frac{n}{\theta^n} \left[\frac{u^{n+1}}{n+1} \right]_0^\theta = \frac{n}{\theta^n} * \frac{\theta^{n+1}}{n+1} = \frac{n}{n+1} \theta \\ &\implies E(\hat{\theta}_{MV}) = \frac{n}{n+1} \theta \end{aligned}$$

$\hat{\theta}_{MV}$ es asintóticamente insesgado pues:

$$E(\hat{\theta}_{MV}) = \frac{n}{n+1} \theta \xrightarrow{n \rightarrow \infty} \theta$$

12 Intervalos de confianza

12.1 Ejercicio 30

Sean X_1, \dots, X_n v.a.i.i.d. con distribución Exponencial de parámetro λ . Halle un intervalo de confianza de nivel exactamente $1 - \alpha$ para el parámetro λ . Sugerencia: se podrá usar (después de probarlo) que si $X \sim \mathcal{E}(\lambda)$ entonces $2\lambda X \sim \mathcal{E}(\frac{1}{2})$ por lo que $2\lambda \sum_{i=1}^n X_i \sim \Gamma(n, \frac{1}{2}) = \mathcal{X}_{2n}^2$.

Sean $X \sim \mathcal{E}(\lambda)$, $Y = 2\lambda X$

Buscamos la distribución de Y :

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(2\lambda X \leq y) = P(X \leq \frac{y}{2\lambda}) = F_X(\frac{y}{2\lambda}) \\ f_Y(y) &= \frac{\partial}{\partial y} F_Y(y) = \frac{\partial}{\partial y} F_X(\frac{y}{2\lambda}) = f_X(\frac{y}{2\lambda}) \frac{1}{2\lambda} = \lambda e^{-\lambda \frac{y}{2\lambda}} \frac{1}{2\lambda} = \frac{1}{2} e^{-\frac{y}{2}} \\ &\implies Y = 2\lambda X \sim \mathcal{E}(\frac{1}{2}) \end{aligned}$$

Buscamos la distribución de $\sum_{j=1}^n Y_j$ usando la función característica:

$$\Phi_{\sum_{j=1}^n Y_j}(t) = \prod_{j=1}^n \Phi_{Y_j}(t) \stackrel{Y_j \text{ i.i.d.}}{=} \left(\Phi_Y(t) \right)^n = \left(\frac{\frac{1}{2}}{\frac{1}{2} - it} \right)^n = \Phi_{\Gamma(n, \frac{1}{2})}(t) \implies \sum_{j=1}^n Y_j \sim \Gamma(n, \frac{1}{2}) = \mathcal{X}_{2n}^2$$

Usamos como pivote a la siguiente función:

$$T(X_1, \dots, X_n, \lambda) = 2\lambda \sum_{i=1}^n X_i$$

Y usamos la siguiente notación para los cuantiles (consideramos el área hacia la izquierda):

$$P(\mathcal{X}_{2n}^2 \leq \mathcal{X}_{2n,q}^2) = q$$

Ahora podemos obtener un intervalo de confianza de nivel exacto $1 - \alpha$ para el parámetro λ :

$$\begin{aligned} P\left(\mathcal{X}_{2n,\alpha/2}^2 \leq 2\lambda \sum_{i=1}^n X_i \leq \mathcal{X}_{2n,1-\alpha/2}^2\right) &= 1 - \alpha \\ P\left(\frac{\mathcal{X}_{2n,\alpha/2}^2}{2 \sum_{i=1}^n X_i} \leq \lambda \leq \frac{\mathcal{X}_{2n,1-\alpha/2}^2}{2 \sum_{i=1}^n X_i}\right) &= 1 - \alpha \end{aligned}$$

El intervalo de confianza resulta:

$$\left[\frac{\mathcal{X}_{2n,\alpha/2}^2}{2 \sum_{i=1}^n X_i}, \frac{\mathcal{X}_{2n,1-\alpha/2}^2}{2 \sum_{i=1}^n X_i} \right]$$

12.2 Ejercicio 31

Construya un intervalo de confianza de nivel $1 - \alpha$ para la media de la distribución normal $\mathcal{N}(\mu, \sigma^2)$ con varianza conocida. Y si no se conoce la varianza?

Sea X_1, \dots, X_n una muestra aleatoria de una población $\mathcal{N}(\mu, \sigma^2)$ con la varianza σ^2 conocida.

$$\bar{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) \iff \sqrt{n} \frac{\bar{X} - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

Usamos como pivote a la siguiente función:

$$T(X_1, \dots, X_n, \mu) = \sqrt{n} \frac{\bar{X} - \mu}{\sigma}$$

Y usamos la siguiente notación para los cuantiles (consideramos el área hacia la izquierda):

$$P(Z \leq z_q) = q$$

Ahora podemos obtener un intervalo de confianza de nivel $1 - \alpha$ para la media μ con varianza σ^2 conocida:

$$\begin{aligned} P\left(z_{\alpha/2} \leq \sqrt{n} \frac{\bar{X} - \mu}{\sigma} \leq z_{1-\alpha/2}\right) &= 1 - \alpha \\ P\left(z_{\alpha/2} \frac{\sigma}{\sqrt{n}} - \bar{X} \leq -\mu \leq z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} - \bar{X}\right) &= 1 - \alpha \\ P\left(\bar{X} - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) &= 1 - \alpha \end{aligned}$$

Notemos que por la simetría de la distribución normal estándar:

$$P(Z \geq z_{1-q}) = q \implies z_{1-q} = -z_q$$

Por lo tanto podemos simplificar el intervalo para buscar un único cuantil:

$$P\left(\bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

El intervalo de confianza resulta:

$$\left[\bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

Si no se conoce la varianza, podemos reemplazar σ por el desvío muestral S :

$$\sqrt{n} \frac{\bar{X} - \mu}{S} \sim t_{n-1}$$

Siendo t_{n-1} la distribución t-Student con $n-1$ grados de libertad. Esta distribución se comporta de forma similar a una normal, tiene forma de campana pero con colas más pesadas. Cuando $n \rightarrow \infty$ aproxima una normal estándar.

Planteamos nuevamente el intervalo de confianza con este otro pivote y hacemos el mismo despeje. Consideramos la misma interpretación de los cuantiles.

$$P\left(t_{n-1, \alpha/2} \leq \sqrt{n} \frac{\bar{X} - \mu}{S} \leq t_{n-1, 1-\alpha/2}\right) = 1 - \alpha$$

$$P\left(\bar{X} + t_{n-1, \alpha/2} \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} - t_{n-1, \alpha/2} \frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

El intervalo de confianza resulta:

$$\left[\bar{X} + t_{n-1, \alpha/2} \frac{S}{\sqrt{n}}, \bar{X} - t_{n-1, \alpha/2} \frac{S}{\sqrt{n}}\right]$$

Si no se conoce la esperanza

12.3 Ejercicio 32

Construya un intervalo de confianza de nivel $1 - \alpha$ para la varianza de la distribución normal $\mathcal{N}(\mu, \sigma^2)$ con esperanza conocida. Y si no se conoce la esperanza?

Sea X_1, \dots, X_n una muestra aleatoria de una población $\mathcal{N}(\mu, \sigma^2)$ con la esperanza μ conocida.

$$\frac{X_i - \mu}{\sigma} \sim \mathcal{N}(0, 1) \quad \forall 1 \leq i \leq n \implies \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi_1^2 = \Gamma\left(\frac{1}{2}, \frac{1}{2}\right) \quad \forall 1 \leq i \leq n$$

Como además las v.a. son independientes:

$$\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi_n^2 = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$$

Usamos como pivote a la siguiente función:

$$T(X_1, \dots, X_n, \sigma^2) = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2}$$

Y usamos la siguiente notación para los cuantiles (consideramos el área hacia la izquierda):

$$P(\chi_n^2 \leq \chi_{n,q}^2) = q$$

Ahora podemos obtener un intervalo de confianza de nivel $1 - \alpha$ para la varianza σ^2 con esperanza μ conocida:

$$P\left(\chi_{n,\alpha/2}^2 \leq \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \leq \chi_{n,1-\alpha/2}^2\right) = 1 - \alpha$$

$$P\left(\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{n,1-\alpha/2}^2} \leq \sigma^2 \leq \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{n,\alpha/2}^2}\right) = 1 - \alpha$$

El intervalo de confianza resulta:

$$\left[\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{n,1-\alpha/2}^2}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{n,\alpha/2}^2} \right]$$

Si no se conoce la esperanza, podemos usar el siguiente pivote, siendo S^2 la varianza muestral:

$$T(X_1, \dots, X_n, \sigma^2) = \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

Ahora podemos obtener un intervalo de confianza de nivel $1 - \alpha$ para la varianza σ^2 con esperanza desconocida:

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

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

El intervalo de confianza resulta:

$$\left[\frac{(n-1)S^2}{\chi_{n-1,1-\alpha/2}^2}, \frac{(n-1)S^2}{\chi_{n-1,\alpha/2}^2} \right]$$

13 Test de hipótesis

13.1 Ejercicio 33

Sea $X \sim \mathcal{N}(\mu, \sigma^2)$ con $\sigma^2 = 9$ conocida. Se quiere testear la hipótesis $H_0 : \mu = 30$ contra $H_a : \mu > 30$. Se toma una muestra de tamaño 16 obteniendo un promedio muestral $\bar{x} = 31$. Calcule el P-valor y decida si se rechaza el test a nivel 0,01. Calcule el error de tipo 2 a nivel 0,01 si el valor verdadero de μ es $\mu = 32$.

$n = 16$ tamaño de la muestra

$\bar{X} = 31$ promedio muestral

$\mu_0 = 30$

$\sigma_0^2 = 9 \implies \sigma_0 = 3$

$H_0 : \mu = 30$ vs $H_a : \mu > 30$

$T = \sqrt{n} \frac{\bar{X} - \mu_0}{\sigma_0} \sim \mathcal{N}(0, 1)$

α el nivel del test

z_α el cuantil tal que $P(Z \geq z_\alpha) = \alpha \iff 1 - P(Z \leq z_\alpha) = \alpha \iff \Phi(z_\alpha) = 1 - \alpha \iff z_\alpha = \Phi^{-1}(1 - \alpha)$

$T \geq z_\alpha$ región de rechazo

P-Valor = $P(Z \geq T_{obs}) = 1 - P(Z \leq T_{obs}) = 1 - \Phi(T_{obs})$

$T_{obs} = \frac{4}{3}$

P-Valor = $1 - 0.9082 = 0.0918$

El P-Valor nos dice la probabilidad mínima necesaria para rechazar H_0 en función del valor de T observado. El P-Valor cambia por cada muestra observada. Entonces, sabiendo el P-Valor podemos decidir si rechazar o no H_0 simplemente comparando el nivel deseado con el P-Valor.

A nivel $\alpha = 0.01 < 0.0918 \implies$ no se rechaza H_0

Para calcular el error de tipo 2 (aceptar H_0 cuando en realidad es falsa) podemos usar la función de potencia:

$$\pi(\mu) = P_\mu(T \in R) = \begin{cases} \alpha(\mu) & \text{si } \mu \in H_0 \\ 1 - \beta(\mu) & \text{si } \mu \in H_1 \end{cases}$$

Siendo R la región de rechazo del test: $T \geq z_\alpha$, $\alpha(\mu)$ la probabilidad de error de tipo 1 y $\beta(\mu)$ la probabilidad de error de tipo 2.

Observemos que si $\mu = 32 \in H_1$ por lo tanto:

$$\pi(\mu) = 1 - \beta(\mu) \iff \beta(\mu) = 1 - \pi(\mu) = 1 - P_\mu(T \geq z_\alpha) = P_\mu(T \leq z_\alpha)$$

Entonces:

$$\begin{aligned}\beta(\mu) &= P_\mu(T \leq z_\alpha) \\ &= P_\mu\left(\sqrt{n} \frac{\bar{X} - \mu_0}{\sigma_0} \leq z_\alpha\right) \\ &= P_\mu\left(\sqrt{n} \frac{\bar{X} - \mu + \mu - \mu_0}{\sigma_0} \leq z_\alpha\right) \\ &= P_\mu\left(\sqrt{n} \frac{\bar{X} - \mu}{\sigma_0} + \sqrt{n} \frac{\mu - \mu_0}{\sigma_0} \leq z_\alpha\right) \\ &= P_\mu\left(Z \leq z_\alpha - \sqrt{n} \frac{\mu - \mu_0}{\sigma_0}\right) \\ &= \Phi\left(z_\alpha - \sqrt{n} \frac{\mu - \mu_0}{\sigma_0}\right)\end{aligned}$$

Ahora sí, calculemos el error de tipo 2 siendo $\alpha = 0.01$ y $\mu = 32$:

$$\begin{aligned}z_\alpha &= \Phi^{-1}(1 - \alpha) = \Phi^{-1}(0.99) = 2.33 \\ \beta &= \Phi\left(2.33 - \sqrt{16} \frac{32 - 30}{3}\right) = \Phi(-0.3366) = 1 - \Phi(0.3366) = 0.3707\end{aligned}$$

Nota: Los cálculos se hicieron con la tabla de la distribución Normal, si los haces con R te van a quedar números levemente distintos (con más precisión).

13.2 Ejercicio 34

Se toman 25 mediciones de la temperatura en cierto sector de un reactor obteniéndose un promedio muestral $\bar{x} = 249^\circ\text{C}$ y un desvío muestral $s = 2.8^\circ\text{C}$. A nivel $\alpha = 0.05$ decida si (a) la temperatura media en ese sector del reactor es menor que 250°C y (b) la varianza de la temperatura en ese sector del reactor es mayor que $(2^\circ\text{C})^2$. Suponga los datos normales.

Pasamos en limpio los datos del enunciado:

$$n = 25$$

$$\bar{X} = 249$$

$$s = 2.8$$

$$\alpha = 0.05$$

Primer test para la media con varianza desconocida

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu < \mu_0 \text{ con } \mu_0 = 250$$

Al no conocer la varianza, usamos el siguiente estadístico:

$$T = \sqrt{n} \frac{\bar{X} - \mu_0}{S} \sim t_{n-1} \text{ con } t_{n-1} \text{ la distribución t-Student con } n - 1 \text{ grados de libertad.}$$

La región de rechazo está dada por:

$$T \leq t_{n-1,\alpha}$$

Con $t_{n-1,\alpha}$ tal que $P(T \leq t_{n-1,\alpha}) = \alpha \iff t_{n-1,\alpha} \approx -1.7109$

Calculamos el valor de T observado:

$$T_{obs} = \sqrt{25} \frac{249-250}{2.8} \approx -1.7857$$

Por lo tanto:

$T_{obs} \leq t_{n-1,\alpha} \iff -1.7857 \leq -1.7109 \implies$ rechazamos H_0 en favor de H_1 y por lo tanto podemos afirmar con nivel 0.05 que la temperatura media en ese sector es menor que 250°C .

Segundo test para la varianza con media desconocida

$$H_0 : \sigma^2 = \sigma_0^2 \text{ vs } H_1 : \sigma^2 > \sigma_0^2 \text{ con } \sigma_0^2 = 2^2$$

Al no conocer la media, usamos el siguiente estadístico:

$$T = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi_{n-1}^2 \text{ distribución Chi cuadrado con } n-1 \text{ grados de libertad.}$$

La región de rechazo está dada por:

$$T \geq \chi_{n-1,1-\alpha}^2$$

$$\text{Con } \chi_{n-1,1-\alpha}^2 \text{ tal que } P(\chi_{n-1}^2 \geq \chi_{n-1,1-\alpha}^2) = \alpha \iff \chi_{n-1,1-\alpha}^2 \approx 36.415$$

Calculamos el valor de T observado:

$$T_{obs} = \frac{(25-1)2.8^2}{2^2} \approx 47.04$$

Por lo tanto:

$T_{obs} \geq \chi_{n-1,1-\alpha}^2 \iff 47.04 \geq 36.415 \implies$ rechazamos H_0 en favor de H_1 y por lo tanto podemos afirmar con nivel 0.05 que la varianza de la temperatura en ese sector es mayor que 4°C .

13.3 Ejercicio 35

Un sujeto acierta el color (rojo o negro) de 850 de 1600 cartas puestas al dorso. Queremos decidir si creemos que es adivino. Defina el test, el estadístico utilizado y calcule el P-valor para este test. ¿Qué decisión toma a nivel 1%? ¿Y a nivel 0.5%?

$n = 1600$ tamaño de la muestra, es decir, el total de cartas

$$\bar{X} = \frac{850}{1600} = 0.53125 \text{ la proporción muestral de las cartas adivinadas por el supuesto adivino}$$

$p_0 = 0.5$ la proporción ideal suponiendo equiprobabilidad de que las cartas sean rojas o negras

$$\sigma_0 = \sqrt{p_0(1-p_0)} = 0.5 \text{ el desvío ideal de } p_0$$

$$H_0 : p = p_0 \text{ vs } H_1 : p > p_0$$

$$T = \sqrt{n} \frac{\bar{X} - p_0}{\sigma_0} \sim \mathcal{N}(0, 1)$$

α el nivel del test

z_α el cuantil tal que $P(Z \geq z_\alpha) = \alpha$

$$T \geq z_\alpha \text{ región de rechazo}$$

$$\text{P-Valor} = P(Z \geq T_{obs}) = 1 - P(Z \leq T_{obs}) = 1 - \Phi(T_{obs})$$

$$T_{obs} = 2.5$$

$$\text{P-Valor} = 1 - 0.9938 = 0.0062$$

El P-Valor nos dice la probabilidad mínima necesaria para rechazar H_0 en función del valor de T observado. El P-Valor cambia por cada muestra observada. Entonces, sabiendo el P-Valor podemos decidir si rechazar o no H_0 simplemente comparando el nivel deseado con el P-Valor.

A nivel 1% $\implies \alpha = 0.01 > 0.0062 \implies$ se rechaza H_0

A nivel 0.5% $\implies \alpha = 0.005 < 0.0062 \implies$ no se rechaza H_0