Trabajo práctico 1: Combinatoria

1. a) ¿Cuántas banderas diferentes de 3 bandas horizontales distintas se puede formar con los colores: verde, rojo, azul, blanco, amarillo, negro?
   b) ¿Cuántas si se permite que la banda inferior sea del mismo color que la banda superior?

2. Cada pieza del dominó está marcada con dos números. Las piezas son simétricas de manera que la pareja de números no está ordenada ¿Cuántas piezas diferentes se pueden formar con los números $\mathbf{1, 2, \cdots, n}$?

3. Los números $\mathbf{1, 2, \cdots, n}$ se colocan en fila. Determinar:
   a) De cuántas formas puede realizarse.
   b) De cuántas si los números 1 y 2 deben aparecer en ese orden (primero el 1 y en algún lugar posterior el 2)
   c) De cuántas si los números 1,2 y 3 deben aparecer en su orden natural.

4. a) ¿Cuántos números de cuatro cifras pueden formarse?
   b) ¿Cuántos con todas sus cifras diferentes?
   c) ¿Cuántos con todas sus cifras diferentes y que terminen en 0?
   d) ¿Cuántos capicúas?

5. ¿De cuántas maneras diferentes pueden ordenarse las letras de la palabra ARBOL?

6. ¿De cuántas maneras puede ubicarse en fila a 10 chicos si uno de ellos en particular no puede estar a la cabeza?

7. Se desea disponer 4 libros de matemática distintos, 6 distintos de física y 3 distintos de química en un estante.
   a) ¿De cuántas maneras puede llevarse a cabo?
   b) ¿De cuántas si los libros de una misma materia deben quedar juntos?
   c) ¿De cuántas si los libros de química deben quedar juntos?
   d) ¿De cuántas si los libros de química deben quedar juntos al comienzo o al final del estante?

8. a) ¿Cuántas ordenaciones de la palabra ABRACADABRA existen?
   b) ¿Cuántas en las que todas las A aparecen juntas?
   c) ¿Cuántas que comiencen con ABRA?

9. a) ¿Cuántos números de cinco cifras puede formarse con los dígitos 1,2,3,4,5,6,7,8,9?
   b) ¿Cuántos si las cifras deben ser impares?
   c) ¿Cuántos si deben tener las dos primeras cifras pares?

10. ¿De cuántas maneras es posible elegir de entre 9 personas un comite de 5?

11. Dados cinco puntos del plano tales que ninguna terna de ellos yace sobre una recta, ¿cuántos triángulos pueden formarse con sus vértices en tres de los puntos dados?

12. Se elige un comité de cinco miembros entre 30 hombres y 20 mujeres.
    a) ¿De cuántas maneras puede hacerse la elección?
    b) ¿De cuántas si la cantidad de mujeres debe exceder a la cantidad de hombres?

13. Se distribuyen diez bolitas iguales en tres cajas distintas.
    a) ¿De cuántas formas puede hacerse?
    b) ¿De cuántas si la primera caja debe contener exactamente tres bolitas?
    c) ¿De cuántas si la primera caja debe contener al menos tres bolitas?

14. En una oficina de correos se vende estampillas de diez tipos.
    a) ¿De cuántas maneras pueden comprarse 12 estampillas?
    b) ¿De cuántas maneras pueden comprarse 8 estampillas diferentes?

15. ¿Cuántas palabras de cuatro consonantes distintas y tres vocales distintas se puede formar con 9 consonantes y 5 vocales?

16. Una caja contiene $\mathbf{90}$ bolillas blancas numeradas y $\mathbf{10}$ bolillas rojas numeradas. Averiguar de cuántas formas es posible escoger
    a) nueve bolillas de la urna.
    b) nueve bolillas de las cuales seis resulten blancas.

17. Una mano de poker consiste en 5 cartas (baraja francesa). Averiguar de cuántas formas distintas es posible obtener una mano de poker y de cuántas es posible obtener "full", es decir una mano con tres cartas de igual numeración y dos cartas de igual numeración. Por ejemplo, si una mano resulta en un dos de picas, un dos de tréboles, un dos de corazones, un ocho de corazones y un ocho de diamantes, se trataría de un full.

18. En el juego de bridge se reparten las 52 cartas de la baraja francesa entre cuatro jugadores, que designamos N,S,E y O. Averiguar:
    a) De cuántas formas es posible repartir las 52 cartas entre los 4 jugadores?
    b) De cuántas maneras es posible repartir las 52 cartas y que E reciba 13 cartas de diamantes?
    c) De cuántas maneras es posible repartir las 52 cartas de modo que uno de los cuatro jugadores reciba 13 cartas de diamantes?

19. Se lanza una moneda normal $\boldsymbol{n}$ veces. Designemos $\boldsymbol{A_n} =$ "no salen caras consecutivas". Así por ejemplo, si $\boldsymbol{n = 5}$ y saliera $\boldsymbol{CSCSS}$ entonces habría ocurrido $\boldsymbol{A}$, en tanto que si saliera $\boldsymbol{CSCCC}$ no habría ocurrido $\boldsymbol{A}$. Anotemos $\boldsymbol{f_n = \#(A_n)}$. Argumentar para justificar que:

$$\begin{cases} \boldsymbol{f_o} & = & \mathbf{1} \\ \boldsymbol{f_1} & = & \mathbf{2} \\ \boldsymbol{f_n} & = & \boldsymbol{f_{n-1} + f_{n-2}} \end{cases}$$

RESPUESTAS

1) a) $\mathbf{6 \cdot 5 \cdot 4}$;b) $\mathbf{6 \cdot 5 \cdot 5}$

2) $\binom{n}{2} = \frac{n(n-1)}{2}$

3) a) $\boldsymbol{n!}$;b) $\binom{n}{2}(n-2)!$;c) $\binom{n}{3}(n-3)!$

4) a) $\mathbf{9 \cdot 10^3}$;b) $\mathbf{9 \cdot 9 \cdot 8 \cdot 7}$;c) $\mathbf{9 \cdot 8 \cdot 7}$;d) $\mathbf{9 \cdot 10}$

5) $\mathbf{5!}$

6) $\mathbf{9 \cdot 9!}$

7) a) $\mathbf{(4+6+3)! = 13!}$;b) $\mathbf{4!6!3!3!}$;c) $\mathbf{11!3!}$;d) $\mathbf{2!3!10!}$

8) a) $\frac{11!}{5!2!2!1!1!}$;b) $\frac{7!}{2!2!1!1!1!}$;c) $\frac{7!}{3!1!1!1!1!}$

9) a) 9^5 ;b) 5^5 ;c) $4^2 \cdot 9^3$

10) $\binom{9}{5}$

11) $\binom{5}{3}$

12) a) $\binom{30+20}{5} = \binom{50}{5}$;b) $\sum_{i=0}^{2} \binom{30}{i}\binom{20}{5-i}$

13) a) $\frac{(10+3-1)!}{10!(3-1)!}$;b) $\frac{(7+2-1)!}{7!(2-1)!}$; c) $\frac{(7+3-1)!}{7!(3-1)!}$

14) a) $\frac{(12+10-1)!}{12!(10-1)!}$;b) $\binom{10}{8}$

15) $\binom{9}{4}\binom{5}{3}7!$

16) a) $\binom{90+10}{9} = \binom{100}{9}$;b) $\binom{90}{6}\binom{10}{3}$

17) $\binom{13}{1}\binom{4}{3}\binom{12}{1}\binom{4}{2}$; $\binom{52}{5}$

18) a) $\binom{52}{13,13,13,13} = \binom{52}{13}\binom{39}{13}\binom{26}{13}\binom{13}{13}$;b) $\binom{39}{13,13,13} = \binom{39}{13}\binom{26}{13}\binom{13}{13}$;c) $4\binom{39}{13}\binom{26}{13}\binom{13}{13}$

19) Separar en dos casos: Que la última ($\boldsymbol{n}$-ésima) salga cara (C) o que la última salga ceca (S). Si la última fue cara, la anteúltima fue....... y las $\boldsymbol{n-2}$ primeras pueden haber sido....... En cambio si la última fue ceca, las $\boldsymbol{n-1}$ anteúltimas pueden haber sido....... Se entiende?

Trabajo práctico 2: Espacios de probabilidad - Variables aleatorias

1. Determinar un espacio muestral apropiado para cada uno de los siguientes experimentos aleatorios. Decidir si el espacio muestral es finito, infinito numerable o infinito no numerable. Si es finito, justificar si es razonable suponer un modelo de equiprobabilidad:

   a) Una urna contiene dos bolas blancas y una bola negra. Se extraen dos bolas al azar y se anotan los colores.

   b) Se lanza un dado dos veces.

   c) Un llavero contiene n llaves, de las cuales sólo una abre la puerta indicada. Se prueban al azar una a una las llaves hasta encontrar la que abre, de la forma siguiente:

      c1) descartando las llaves que no van abriendo.

      c2) conservando las llaves que no van abriendo.

   d) Se mide el tiempo de vida en horas de una lamparita eléctrica de cierta marca particular.

   e) Se colocan al azar tres bolas diferentes en tres urnas distintas.

   f) Se colocan al azar tres bolas idénticas en tres urnas diferentes.

   g) Una caja contiene 30 lámparas de las cuales 3 tienen roto el filamento. Se prueban de a una hasta encontrar una defectuosa.

   h) Se arroja una moneda tantas veces como sea necesario hasta obtener cara.

   i) Se determina el peso de una rata de diez días de vida.

2. Verificar las siguientes igualdades mediante el uso de diagramas de Venn:

   a) $A \setminus B = A \setminus (A \cap B) = A \cap B^c$ b) $A = (A \cap B) \uplus (A \cap B^c)$

   c) $A \cup B = A \uplus (B \setminus A)$ d) $A \cup B = (A \setminus B) \uplus (A \cap B) \uplus (B \setminus A)$

3. Una compañía constructora trabaja en dos proyectos diferentes (que indicaremos 1 y 2). Sea A_i el evento que el proyecto i se termina en la fecha del contrato ($i = 1, 2$). Sabiendo que $P(A_1 \cup A_2) = 0.9$ y que $P(A_1 \cap A_2) = 0.5$, ¿cuál es la probabilidad de que exactamente un proyecto se termine en la fecha del contrato? Utilice un diagrame de Venn para visualizar la situación.

4. En un taller de máquinas de una escuela el 60% de todas las descomposturas de máquinas ocurre con tornos, el 15% con taladros y el resto por otras causas. Supongamos que cada descompostura es atribuible a una única causa. Sean:

   $A =$ "la siguiente descompostura es en torno"
   $B =$ "la siguiente descompostura es en taladro"

   Suponiendo que $P(A) = 0.6$ y $P(B) = 0.15$, calcular:

   a) $P(A^c)$ *b)* $P(A \cup B)$ *c)* $P(A^c \cap B^c)$

5. Sea (Ω, Σ, P) un espacio de probabilidad.

   a) Dados eventos $A, B \in \Sigma$ probar que:

   $$P(A \cap B) \geq P(A) + P(B) - 1$$

b) (Generalización la desigualdad anterior a K eventos) Sean $A_k \in \Sigma$ $(k = 1, \cdots, K)$. Demostrar que:

$$P\left(\bigcap_{k=1}^{K} A_k\right) \geq \sum_{k=1}^{K} P(A_k) - (K-1)$$

c) En relación al item anterior y dado $\alpha \in (0,1)$, mostrar que si $P(A_k) \geq 1 - \frac{\alpha}{K}$ para $k = 1, \cdots, K$ entonces $P\left(\bigcap_{k=1}^{K} A_k\right) \geq 1 - \alpha$

6. Se lanzan dos monedas y un dado.

a) Describir el espacio muestral.

b) Enumerar los siguientes eventos:

$A =$ "salen dos caras y un número par" $B =$ "sale un 2"
$C =$ "sale exactamente una cara y sale número primo"

c) Calcular $P(A), P(B)$ y $P(C)$

7. Un lote de 20 artículos contiene 15 artículos buenos y 5 defectuosos. Se extraen cuatro artículos al azar y sin reemplazo.

a) Describir el espacio muestral asociado a este experimento aleatorio.

b) Describir como subconjuntos del espacio muestral los eventos

$A =$ "el primer artículo seleccionado es defectuoso"
$B =$ "al menos un artículo seleccionado es defectuoso"
$C =$ "a lo sumo un artículo seleccionado es defectuoso"

8. Sea (Ω, Σ, P) un espacio de probabilidad y sea $\{A_n\}$ una sucesión de eventos. Probar:

a) $(\forall n \, , \, P(A_n) = 0) \quad \Rightarrow \quad P\left(\bigcup_{n=1}^{\infty} A_n\right) = 0$

b) $(\forall n \, , \, P(A_n) = 1) \quad \Rightarrow \quad P\left(\bigcap_{n=1}^{\infty} A_n\right) = 1$

9. Se lanzan cinco dados equilibrados. Determinar un espacio muestral apropiado para describir este experimento aleatorio y calcular la probabilidad de obtener:

a) Cinco números iguales ("Generala servida").

b) Cuatro números iguales y uno distinto ("Poker").

c) Tres de un número y dos de otro número ("Full").

10. Tres aficionados al basquet A,B y C se turnan en ese orden para arrojar la pelota al aro. El primero en encestar gana. Podemos definir el espacio muestral como:

$$\Omega = \{1, 01, 001, 0001, 00001, \cdots\}$$

a) Interpretar el espacio muestral.

b) Expresar por extensión los eventos siguientes:

$E_1 =$ "gana A en el tercer intento"
$E_2 =$ "gana B en el segundo intento"

c) Asignar probabilidades a los eventos elementales, sabiendo que A,B y C embocan al aro con probabilidades **0.3 , 0.7 , 0.6** respectivamente ¿Qué hipótesis debe realizarse?

11. En un ropero hay $\boldsymbol{n}$ pares distintos de zapatos. Se eligen al azar $\boldsymbol{2r}$ zapatos ($\boldsymbol{2r < n}$). Calcular la probabilidad de que:

    a) No se obtenga ningún par completo.

    b) Se obtenga exactamente un par completo.

    c) Se obtengan exactamente dos pares completos.

12. La ruta utilizada por cierto automovilista para ir al trabajo tiene dos cruces con semáforo. La probabilidad de que tenga que detenerse en el primer semáforo es **0.4**, la probabilidad de que deba detenerse en el segundo es **0.5** y la probabilidad de que deba detenerse por lo menos en uno de ellos es **0.6** ¿Cuál es la probabilidad de que deba detenerse en

    a) ambos semáforos?

    b) el primero pero no en el segundo?

    c) exactamente uno de ellos?

    Interpretar utilizando diagramas de Venn.

13. Una comisión de cinco personas va a ser seleccionada entre un grupo de 6 hombres y 9 mujeres.

    a) Describir un espacio muestral adecuado.

    b) Si la elección es al azar, ¿cuál es la probabilidad de que la comisión quede integrada por 3 hombres y 2 mujeres?

14. De un mazo de baraja española se extraen 3 cartas al azar y sin reemplazo.

    a) Describir un espacio muestral apropiado.

    b) Calcular la probabilidad de que todas las cartas obtenidas sean del mismo palo (”flor”).

15. De un bolillero que bontiene $\boldsymbol{b}$ bolillas blancas y $\boldsymbol{r}$ bolillas rojas, se extraen al azar, sucesivamente y sin reemplazo $\boldsymbol{n}$ bolillas ($\boldsymbol{1 \leq n \leq b + r}$).

    a) Fijado $\boldsymbol{k}$ tal que $\boldsymbol{1 \leq k \leq n}$, hallar la probabilidad de que la $\boldsymbol{k}$-ésima bolilla extraída sea:

       a1) blanca.

       a2) la primera blanca extraída.

    b) Tomando $\boldsymbol{k = n}$ sea $\boldsymbol{p_n}$ la probabilidad hallada en a2). Probar que si $\boldsymbol{b \to \infty}$ y $\boldsymbol{r \to \infty}$ pero de modo tal que la proporción de blancas $\boldsymbol{p \overset{\text{def}}{=} b/(b + r)}$ permanezca constante, entonces:

$$\boldsymbol{p_n \to (1 - p)^{n-1} p}$$

16. De un bolillero que contiene $\boldsymbol{b}$ bolillas blancas y $\boldsymbol{r}$ bolillas rojas, se extraen al azar, sucesivamente y sin reemplazo $\boldsymbol{n}$ bolillas ($\boldsymbol{1 \leq n \leq b + r}$). Hallar la probabilidad $\boldsymbol{p_k}$ de obteber exactamente $\boldsymbol{k}$ bolillas blancas. Probar que si $\boldsymbol{b \to \infty}$ y $\boldsymbol{r \to \infty}$ pero de modo tal que $\boldsymbol{p \overset{\text{def}}{=} b/(b + r)}$ permanezca constante, entonces:

$$\boldsymbol{p_k \to \binom{n}{k} p^k (1 - p)^{n-k}}$$

Nota: Este resultado se conoce como la *aproximación binomial a la distribución hipergeométrica.*

ALGUNAS RESPUESTAS

1) a) $\Omega = \{(x, y) \,:\, x, y \in \{B_1, B_2, N_1\} \,,\, x \neq y\}$. Finito: $\#(\Omega) = 6$. Equiprobable. Otra opción: $\Omega = \{A \,:\, A \subseteq \{B_1, B_2, N_1\} \,,\, \#(A) = 2\}$. Finito: $\#(\Omega) = 3$. Equiprobable.

b) $\Omega = \{(x, y) \,:\, x, y \in \{1, 2, 3, 4, 5, 6\}\}$. Finito: $\#(\Omega) = 6^2$. Equiprobable si los dados son equilibrados.

c) c1) $\Omega = \left\{ A, NA, NNA, \cdots, \underbrace{N \cdots N}_{n-1} A \right\}$. Finito: $\#(\Omega) = n$. Equiprobable:

$$P(\{A\}) = \frac{1}{n}$$
$$P(\{NA\}) = \frac{(n-1)\cdot 1}{n(n-1)} = \frac{1}{n}$$
$$P(\{NNA\}) = \frac{(n-1)(n-2)\cdot 1}{n(n-1)(n-2)} = \frac{1}{n}$$
$$\vdots$$
$$P\left(\left\{\underbrace{N \cdots N}_{r} A\right\}\right) = \frac{(n-1)(n-2)\cdots(n-r)\cdot 1}{n(n-1)(n-2)\cdots(n-r)} = \frac{1}{n}$$

c2) $\Omega = \{A, NA, NNA, NNNA, \cdots\}$. Infinito numerable.

d) $\Omega = [0, \infty)$. Infinito no numerable.

e) $\Omega = \{(x, y, z) \,:\, x, y, z \in \{1, 2, 3\}\}$. Acá la componente j de la terna ordenada (x, y, z) indica la urna en la que se deposita la bolita j, para $j = 1, 2, 3$. Por ejemplo: $(3, 3, 2)$ indica que las bolitas 1 y 2 están en la urna 3 y la bolita 3 en la urna 2. Este esp.muestral es finito: $\#(\Omega) = 3^3$. Equiprobable.

f) Una opción es numerar las bolitas y continuar utilizando el esp.muestral Ω de e). Otra opción es $\Omega^* = \{(x, y, z) \,:\, x, y, z \in \{0, 1, 2, 3\} \,,\, x + y + z = 3\}$. Acá la componente j de la terna ordenada (x, y, z) indica la cantidad de bolitas en la urna j, para $j = 1, 2, 3$. Este espacio muestral es finito pero no es equiprobable. Por ejemplo: Sean $\omega_1^* = (0, 0, 3)$, $\omega_2^* = (1, 0, 2)$. En Ω^* los eventos asociados a w_1^* y ω_2^* son simples. Pero vistos desde Ω uno de ellos pasa a ser compuesto:

$$\omega_1^* \equiv (3, 3, 3) \quad ; \quad \omega_2^* \equiv \{(1, 3, 3), (3, 1, 3), (3, 3, 1)\}$$

Entonces:

$$P^*(\omega_1^*) = P(\{(3, 3, 3)\}) = 1/3^3 = 1/27$$

$$P^*(\omega_2^*) = P(\{(1, 3, 3), (3, 1, 3), (3, 3, 1)\}) = 3/3^3 = 1/9$$

9) a) $1/6^4$; b) $25/6^4$; c) $50/6^4$

11) a) $\binom{n}{2r} 2^{2r} / \binom{2n}{2r}$; b) $\binom{n}{1}\binom{n-1}{2r-2} 2^{2r-2} / \binom{2n}{2r}$; c) $\binom{n}{2}\binom{n-2}{2r-4} 2^{2r-4} / \binom{2n}{2r}$

15) a1) $b/(b + r)$

a2) Si $k \leq r + 1$ la probabilidad es $\frac{r}{b+r} \cdot \frac{r-1}{b+r-1} \cdots \frac{r-k+2}{b+r-k+2} \cdot \frac{b}{b+r-k+1} = \frac{r!\, b\, (b+r-k)!}{(r-k+1)!\, (b+r)!}$

Trabajo práctico 3: Probabilidad Condicional e independencia

1. Probar que si $\boldsymbol{A}$ y $\boldsymbol{B}$ son eventos independientes entonces también lo son $\boldsymbol{A}$ y $\boldsymbol{B^c}$, $\boldsymbol{A^c}$ y $\boldsymbol{B}$, $\boldsymbol{A^c}$ y $\boldsymbol{B^c}$.

2. Sean $\boldsymbol{A}$ y $\boldsymbol{B}$ eventos asociados a un mismo experimento aleatorio. Supongamos que $\boldsymbol{P(A) = 0.4}$ y $\boldsymbol{P(A \cup B) = 0.7}$. Sea $\boldsymbol{p = P(B)}$

   a) ¿Para qué valores de $\boldsymbol{p}$ son $\boldsymbol{A}$ y $\boldsymbol{B}$ mutuamente excluyentes?
   b) ¿Para qué valores de $\boldsymbol{p}$ son $\boldsymbol{A}$ y $\boldsymbol{B}$ independientes?

3. Se selecciona al azar un individuo de la población de adultos que viven en EEUU. Sea A el evento que el individuo seleccionado tenga una estatura de más de 6 pies, y sea B el evento que el individuo seleccionado sea un jugador profesional de básquet ¿Cuál considera que será mayor: $\boldsymbol{P(A|B)}$ ó $\boldsymbol{P(B|A)}$? ¿Porqué?

4. En una fábrica de pernos, las máquinas A,B y C producen respectivamente el 25%, el 35% y el 40% de la producción total. En tal producción el 5%, el 4% y el 2% de los pernos producidos respectivamennte por A, B y C, son defectuosos. Se escoge al azar un perno de la producción total.

   a) Calcular la probabilidad de que el perno escogido resulte defectuoso ¿Cómo se denomina la propiedad utilizada?
   b) Si el perno extraído resulta defectuoso, hallar las probabilidades de que haya sido producido por cada una de las máquinas A, B y C ¿Cómo se denomina la propiedad utilizada?

5. Supongamos que 5 hombres de cada 100 y 25 mujeres de cada 10.000 sufren de daltonismo. Se escoge aleatoriamente una persona de una población donde hay la misma cantidad de hombres que de mujeres. Si dicha persona resulta daltónica, hallar la probabilidad de que sea hombre.

6. En una prueba de opciones múltiples ("multiple choice") un estudiante contesta una pregunta que ofrece cuatro posibles respuestas, de las cuales sólo una es correcta. Suponga que la probabilidad de que el estudiante sepa la respuesta correcta es $\boldsymbol{p}$ y la probabilidad de que conteste al azar es $\boldsymbol{q = 1 - p}$

   a) Calcular la probabilidad de que conteste correctamente la pregunta.
   b) Si contesta correctamente la pregunta, ¿cuál es la probabilidad de que realmente sepa la respuesta correcta?

7. La urna $\boldsymbol{\mathcal{U}_1}$ contiene **6** bolillas blancas y **2** bolillas rojas. La urna $\boldsymbol{\mathcal{U}_2}$ contiene **3** bolillas blancas y **8** bolillas rojas. Se elige al azar una bolilla de $\boldsymbol{\mathcal{U}_1}$ y se la coloca en $\boldsymbol{\mathcal{U}_2}$. A continuación se escoge una bolilla al azar de $\boldsymbol{\mathcal{U}_2}$. Hallar la probabilidad de que esta segunda bolilla sea blanca.

8. Una caja contiene 6 bolas rojas y 4 bolas verdes. Una segunda caja contiene 7 rojas y 3 verdes. Se escoge al azar una bola de la primera caja y se la coloca en la segunda caja. A continuación se elige al azar una bola de la segunda caja y se la deposita en la primera.

   a) ¿Cuál es la probabilidad de que se seleccione una bola roja de la primera caja y una bola roja de la segunda?
   b) Al finalizar el proceso de selección, ¿cuál es la probabilidad de que la primera caja no haya cambiado su composición con respecto al principio?

9. Una caja contiene 6 lapiceras buenas y 4 defectuosas. Las lapiceras se verifican extrayendo una al azar y probándola, y se repite este proceso hasta haber encontrado las 4 defectuosas. Hallar la probabilidad de que la cuarta lapicera defectuosa se encuentre en la

a) quinta extracción.

b) décima extracción.

10. Al contestar a una pregunta con respecto a un tema controversial, como por ejemplo "¿ alguna vez usted ha fumado marihuana?", mucha gente prefiere mentir (es decir, cuando la respuesta debiera ser afirmativa, contesta negativamente). Suponga que, frente a cierta pregunta controversial, el 80% de una población debe contestar "NO", y así lo hace. Del 20% que debería contestar "SÍ", el 70% miente y contesta "NO".

    a) Calcule la probabilidad de que una persona, elegida al azar entre tal población, conteste afirmativamente a la pregunta.

    b) ¿ Cuál es la probabilidad de que sea necesario entrevistar al menos a 25 personas para obtener tres respuestas afirmativas?

11. Un bolso contiene tres monedas, una de las cuales está acuñada con dos caras, mientras que las otras dos monedas son normales. Se escoge una moneda al azar y se lanza cuatro veces sucesivas. Si cada vez sale cara, ¿ cuál es la probabilidad de que ésta sea la moneda con dos caras?

12. Se arrojan diez dados equilibrados. Si al menos uno de ellos resulta as, ¿ cuál es la probabilidad de que resulten dos o más ases?

13. Se arrojan tres dados equilibrados. Si no salen números repetidos, ¿ cuál es la probabilidad de que entre ellos haya un as?

14. En el juego de *bridge* se reparten al azar las 52 cartas de la baraja francesa entre cuatro jugadores que desinamos N,S,E y O. Supongamos que N y S forman una pareja de juego y E y O forman otra pareja de juego. Supongamos que O no recibe ases. Calcular la probabilidades de que su compañero:

    a) no reciba ases.

    b) reciba dos o más ases.

15. Un avión se ha perdido y se presume que ha caído en una de tres regiones con igual probabilidad. Sea $\mathbf{1-\alpha_i}$ la probabilidad de que el avión sea hallado al ser buscado en la región $\boldsymbol{i}$ cuando realmente haya caído en dicha región ($\boldsymbol{i = 1, 2, 3}$). Las constantes $\boldsymbol{\alpha_i}$ se denominan probabilidades de "no avistaje" porque representan la probabilidad de no hallar al avión buscando sobre una región, cuando efectivamente se encuentra en ella. Dichas probabilidades son generalmente atribuibles a condiciones geográficas y/o ambientales de las regiones.

    a) Si se realiza una búsqueda exhaustiva (es decir en todas las regiones), ¿ cuál es la probabilidad de que el avión sea encontrado?

    b) ¿ Cuál es la probabilidad condicional de que el avión haya caído en la $\boldsymbol{i}$-ésima región ($\boldsymbol{i = 1, 2, 3}$) dado que no fue hallado en una búsqueda exhaustiva?

    c) ¿ Cuál es la probabilidad de que el avión haya caído en la región 1 sabiendo que al buscarlo en la región $\boldsymbol{j}$ ($\boldsymbol{j = 1, 2, 3}$) no fue hallado?

16. Se arroja un dado equilibrado todas las veces que sea necesario hasta que aparezca un as. Si suponemos que el as no sale en la primera tirada, ¿ cuál es la probabilidad de que sean necesarios más de tres lanzamientos?

17. Se arroja un dado equilibrado tantas veces como sea necesario hasta que aparezca un as. Suponiendo que son necesarias una cantidad par de tiradas, calcular la probabilidad de que salga as por primera vez en la segunda tirada.

18. La junta de un avión requiere 25 remaches. La junta se tiene que volver a hacer si cualquiera de los remaches es defectuoso. Suponga que cada uno de los remaches tiene la misma probabilidad de ser defectuoso, y que no depende de los demás.

    a) Si 20% de las juntas se tienen que volver a hacer, ¿cuál es la probabilidad de que un remache sea defectuoso?

    b) ¿Qué tan pequeña debe ser la probabilidad de que un remache sea defectuoso, para asegurar que sólo el 10% de las juntas se tengan que volver a realizar?

19. Cada una de las urnas $\mathcal{U}_1, \cdots, \mathcal{U}_n$ contiene b bolitas blancas y r bolitas rojas. Se elige al azar una bolita de $\mathcal{U}_1$ y se la pasa a $\mathcal{U}_2$, se elige entonces al azar una bolita de $\mathcal{U}_2$ y se la pasa a $\mathcal{U}_3$, y así sucesivamente hasta elegir al azar una bolita de $\mathcal{U}_{n-1}$ y pasarla a $\mathcal{U}_n$. Finalmente se escoge una bolita al azar de $\mathcal{U}_n$. Bajo la hipótesis de que la primera bolita que se pasó fue blanca, ¿cuál es la probablidad de que la última bolita extraida sea también blanca? ¿Qué sucede para $n \to \infty$? (*Sugerencia:* Sea $p_n = P($"n-ésimo pasaje es blanca"$)$ Expresar p_n en términos de p_{n-1}).

ALGUNAS RESPUESTAS

2) a) $\mathbf{0.3}$; b) $\mathbf{0.5}$

3) Es de esperar que $P(A|B) > P(B|A)$

4) a) $\mathbf{0.0345}$; b) $\mathbf{25/69}$; $\mathbf{28/69}$; $\mathbf{16/69}$

5) $\mathbf{20/21}$

6) a) $(3p+1)/4$; b) $4p/(3p+1)$

7) $\mathbf{5/16}$

8) a) $\mathbf{24/55}$; b) $\mathbf{32/55}$

9) a) $\binom{4}{3}/\binom{10}{4}$; b) $\binom{9}{3}/\binom{10}{4}$

10) a) $\mathbf{0.06}$; b) $1 - \sum_{k=3}^{24} \binom{k-1}{2} (0.06)^3(0.94)^{k-3}$

11) $\mathbf{8/9}$

12) $1 - \frac{10 \cdot 5^9}{6^{10} - 5^{10}}$

13) $\mathbf{0.5}$

14) a) $\binom{35}{13}/\binom{39}{13}$; b) $1 - \binom{35}{13}/\binom{39}{13} - \binom{4}{1}\binom{35}{12}/\binom{39}{13}$

15) a) $1 - \frac{\alpha_1+\alpha_2+\alpha_3}{3}$; b) $\frac{\alpha_i}{\alpha_1+\alpha_2+\alpha_3}$ $(i = 1,2,3)$; c) $\frac{\alpha_1}{2+\alpha_1}$; $\frac{1}{2+\alpha_2}$; $\frac{1}{2+\alpha_3}$

16) $(5/6)^2$

17) $1 - (5/6)^2$

18) a) $1 - \sqrt[25]{4/5}$; b) $1 - \sqrt[25]{9/10}$

19) ME CANSE!!!

Trabajo práctico 4: Variables aleatorias discretas

1. Se arroja en forma consecutiva una moneda normal tres veces. Sean:
   $\boldsymbol{X}$ = cantidad de cecas obtenidas
   $\boldsymbol{Y}$ = exceso de cecas sobre caras
   $\boldsymbol{Z}$ = proporción de caras obtenidas

   a) Determinar un espacio muestral adecuado para este experimento aleatorio. Discutir equiprobabilidad.

   b) Escribir por extensión los siguientes eventos y luego calcular la probabilidad de cada uno:

   $\{X = 0\}$ $\{X = 1\}$ $\{X = 2\}$ $\{X = 3\}$
   $\{Y = -3\}$ $\{Y = -1\}$ $\{Y = 1\}$ $\{Y = 3\}$
   $\{Z = 0\}$ $\{Z = 1/3\}$ $\{Z = 2/3\}$ $\{Z = 1\}$

   ¿Es **1** la suma de las probabilidades de los cuatro eventos definidos utilizando $\boldsymbol{X}$? ¿Porqué? Misma pregunta para los cuatro eventos definidos utilizando $\boldsymbol{Y}$ y luego para los cuatro eventos definidos utilizando $\boldsymbol{Z}$

   c) Determinar para cada variable aleatoria $\boldsymbol{X}$, $\boldsymbol{Y}$ y $\boldsymbol{Z}$ su rango o recorrido ¿Son v.a. discretas? ¿Porqué?

   d) Escribir por extensión los siguientes eventos y luego calcular la probabilidad de cada uno:

   $\{X \leq 0.5\}$ $\{X < 1\}$ $\{1 \leq X < 4\}$ $\{X \neq 2\}$
   $\{Y < -3\}$ $\{Y > -1\}$ $\{-1 \leq Y \leq 1\}$ $\{Y \geq 1.7\}$
   $\{Z \geq 3/5\}$ $\{Z \geq 1\}$ $\{Z \leq 0\}$ $\{Z \leq 1\}$

   ¿Es **1** la suma de las probabilidades de los eventos definidos utilizando $\boldsymbol{X}$? ¿Porqué? Misma pregunta para los cuatro eventos definidos utilizando $\boldsymbol{Y}$ y luego para los cuatro eventos definidos utilizando $\boldsymbol{Z}$

   e) Escribir por extensión y calcular la probabilidad del evento:

   $$\{X \geq 1\,,\, Y > -3\} \stackrel{\text{def}}{=} \{X \geq 1\} \cap \{Y > -3\}$$

2. En el ejercicio 1 calcular la fmp y la fda de cada una de las v.a. $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$ y representarlas gráficamente. Comprobar a partir de los gráficos las propiedades de una fda. A partir de los gráficos de las fmp, ¿cuál de las tres variables $\boldsymbol{X}, \boldsymbol{Y}\,\boldsymbol{Z}$ considera que posee la mayor varianza y cuál la menor? Verifique luego calculando dichas varianzas.

3. Una caja contiene 3 bolitas rojas y 4 bolitas verdes. Se extraen al azar tres bolitas. Sea $\boldsymbol{V}$ = cantidad de bolitas verdes extraídas

   a) Hallar la fmp y la fda de $\boldsymbol{V}$. Graficarlas.

   b) Hallar la probabilidad de extraer a lo sumo una bolita verde.

   c) Expresar con palabras el evento $\{V \geq 1\}$ y hallar la probabilidad de que ocurra.

4. Se lanza una moneda normal hasta que salga cara. Sea $\boldsymbol{X}$ = cantidad de lanzamientos necesarios

   a) Proponer un espacio muestral adecuado para este experimento aleatorio.

   b) Determinar el rango de $\boldsymbol{X}$ y hallar la fmp y la fda de $\boldsymbol{X}$

   c) Expresar por medio de $\boldsymbol{X}$ los eventos:

   $\boldsymbol{A}$ = "se necesitan al menos tres lanzamientos"
   $\boldsymbol{B}$ = "se necesitan a lo sumo tres lanzamientos"

$C =$ "se necesita un número impar de lanzamientos"

$D =$ "la cantidad de lanzamientos necesarios es múltiplo de cinco"

d) Calcular $P(X \leq 3)$, $P(1 < X \leq 4)$, $P(A)$, $P(B)$, $P(C)$, $P(D)$

5. Un lote de artículos contiene 20 buenos y 5 defectuosos. Se extraen al azar cuatro artículos. Sea X la cantidad de defectuosos hallados. Obtener la distribución de probabilidad de X en los siguientes casos:

   a) La extracción es con reposición.

   b) La extracción es sin reposición.

6. Se arroja un dado equilibrado hasta que sale un divisor de 5. Sea X la cantidad de lanzamientos necesarios.

   a) Hallar R_X ¿Es X una v.a. discreta?

   b) Hallar la fmp de X. Comprobar sus propiedades.

   c) Demostrar que dados $r, s \in \mathbb{N}$ se verifica:

   $$P(X > r + s | X > r) = P(X > s)$$

   Esta propiedad se conoce como "ausencia de memoria" ¿Considera apropiado este nombre?

7. Dadas las siguientes funciones determinar cuáles de ellas son la fmp de alguna variable aleatoria, digamos X. Para aquellas que lo sean calcular:

   $$P(1 \leq X \leq 4), \; P(X = 1), \; P(X \neq 2), \; P(X > 2)$$

   a) $p(x) = \begin{cases} 1/4 & \text{si } x = 1 \;\vee\; x = 4 \\ 1/2 & \text{si } x = 3 \\ 0 & \text{otros casos} \end{cases}$

   b) $p(x) = \begin{cases} 1/4 & \text{si } x = 1 \\ 1/2 & \text{si } x = 2 \\ 0 & \text{otros casos} \end{cases}$

   c) $p(x) = \begin{cases} (2x - x^2 + 7)/18 & \text{si } x = 1, 2, 3, 4 \\ 0 & \text{otros casos} \end{cases}$

   d) $p(x) = \begin{cases} x^2/30 & \text{si } x = 1, 2, 3, 4 \\ 0 & \text{otros casos} \end{cases}$

8. La fmp de X =cantidad de defectos importantes en un electrodoméstico seleccionado al azar y de cierto tipo, viene dada por:

x	0	1	2	3	4
$f_X(x)$	0.08	0.15	0.45	0.27	0.05

   Calcular:

   a) $E(X)$

   b) $V(X)$ utilizando directamente la definición.

   c) El desvío standard de X

   d) $V(X)$ utilizando la fórmula abreviada: $V(X) = E(X^2) - E^2(X)$

9. Un instructor de un grupo de escritura técnica ha solicitado que cierto informe sea entregado la semana siguiente, agregando la restricción que cualquier informe que exceda las cuatro páginas será rechazado. Sea Y el número de páginas del informe de cierto estudiante seleccionado al azar. Supongamos que la fmp de Y viene dada por:

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

a) Calcular $E(Y)$

b) Supongamos que el instructor demora $\sqrt{Y}$ minutos calificando un informe de Y páginas. Hallar el tiempo esperado en calificar un informe seleccionado al azar.

c) Calcular $V(Y)$ y σ_Y. Determinar además la probabilidad de que Y no se aleje de su media más de un desvío standard.

10. Un distribuidor de aparatos electrodomésticos vende tres modelos diferentes de congeladores verticales con capacidades de 13.5, 15.9 y 19.1 pies cúbicos de almacenamiento, respectivamente. Sea X la cantidad de pies cúbicos de almacenamiento en el congelador comprado por el siguiente cliente. Supongamos que X se distribuye como:

x	13.5	15.9	19.1
$f_X(x)$	0.2	0.5	0.3

a) Calcular $E(X)$, $E\left(X^2\right)$ y $V(X)$

b) Si el precio de un congelador que tiene una capacidad de X pies cúbicos es $C = 25X - 8.5$, ¿cuál es el precio esperado a pagar por el siguiente cliente que va a comprar un congelador?

c) Calcular la varianza del precio C definido en el item anterior.

d) Supongamos que mientras que la capacidad nominal de un congelador es X, la capacidad real es $h(X) = X - 0.01X^2$. Hallar la capacidad real esperada de un congelador adquirido por el siguiente cliente.

11. Sea X una variable aleatoria tipo Bernoulli, que toma los valores 0 y 1 con probabilidades $1-p$ y p respectivamente.

a) Calcular EX^2

b) Demostrar que $V(X) = p(1-p)$

c) Calcular EX^{79}

12. Supongamos que el número de plantas de cierto tipo que se encuentrar en una región rectangular (denominada "cuadrante" por los ecologistas) de cierta zona geográfica particular, es una v.a. X con la siguiente fmp:

$$f_X(x) = \begin{cases} c/x^3 & \text{si } x = 1, 2, 3, \cdots \\ 0 & \text{otros casos} \end{cases}$$

¿Existe $V(X)$?

13. Un llavero contiene n llaves, de las cuales sólo una abre determinada puerta. Se van probando las llaves una a una, descartando las que no abren, hasta encontrar la que abre dicha puerta. Sea X la cantidad de pruebas necesarias hasta lograr abrir la puerta (por ejemplo, si abriera en el tercer intento sería $X = 3$). Calcular $E(X)$ y $V(X)$

14. ¿Qué relación existe entre $V(X)$ y $V(-X)$? Ejemplifique y luego demuestre.

15. Demostrar que $V(aX + b) = a^2 V(X)$ siendo a, b constantes reales cualesquiera.

16. Al examinar los pozos de agua de cierto distrito buscando algunos tipos de impurezas que suelen encontrarse en el agua potable, se ha determinado que el 20% de los pozos de ese distrito no tenían ninguna impureza, el 40% de los pozos tenían impureza A y el 50% de los pozos tenían impureza B (Naturalmente algunos pozos tenían ambas). Se elige un pozo al azar de ese distrito y se define Y como la variable aleatoria que cuenta cuántos tipos de impurezas se encuentran en el pozo.

a) Encontrar la fmp de $\boldsymbol{Y}$
b) Calcular $\boldsymbol{E(Y)}$ y $\boldsymbol{V(Y)}$

17. Se distribuyen al azar 3 bolitas distintas en 3 cajas diferentes, permitiéndose más de una bolita en cada caja. Sea $\boldsymbol{N}$ la cantidad de cajas "ocupadas" (es decir con al menos una bolita).

    a) Determinar la fmp y la fda de $\boldsymbol{N}$
    b) Calcular $\boldsymbol{E(N)}$ y $\boldsymbol{V(N)}$

18. Se arrojan dos dados equilibrados. Sea $\boldsymbol{X}$ el mínimo de los números obtenidos.

    a) Determinar la fmp y la fda de $\boldsymbol{X}$
    b) Calcular $\boldsymbol{E(X)}$ y $\boldsymbol{V(X)}$

19. Un sistema para detectar incendios utiliza tres celdas sensibles al calor que actúan en forma independiente. La alarma se activa si al menos una de las celdas detecta un exceso de temperatura (alcanzar una temperatura de al menos $\mathbf{100°}\boldsymbol{C}$). Cada celda tiene una probabilidad $\boldsymbol{p = 0.8}$ de detectar exceso de temperatura. Sea $\boldsymbol{Y}$ la cantidad de celdas que activan la alarma cuando la temperatura alcanza los $\mathbf{100°}\boldsymbol{C}$. Encontrar:

    a) La fmp de $\boldsymbol{Y}$
    b) La probabilidad de que la alarma se active cuando la temperatura alcanza los $\mathbf{100°}\boldsymbol{C}$
    c) $\boldsymbol{E(Y)}$ y $\boldsymbol{V(Y)}$

20. La cantidad de solicitudes de asistencia recibidas por un servicio de remolque de vehículos es un proceso de Poisson con tasa $\mathbf{4}$ por hora.

    a) Calcular la probabilidad de que se reciban $\mathbf{10}$ solicitudes entre las 16 y las 17 h.
    b) Si los operadores de los remolques se toman un descanso de $\mathbf{30}$ min, ¿cuál es la probabilidad de que no se pierda ninguna llamada de asistencia durante ese lapso?
    c) Calcular la probabilidad de que se reciban $\mathbf{2}$ solicitudes entre las $\mathbf{13}$ y las $\mathbf{13:15}$ horas.

21. En cierto negocio los clientes llegan al mostrador de la caja de acuerdo a una distribución de Poisson con un promedio de siete clientes por hora. Calular la probabilidad de que en una hora dada:

    a) No lleguen más de tres clientes.
    b) Lleguen al menos dos clientes.
    c) Lleguen al menos cinco clientes.

22. El número de colonias de bacterias de cierto tipo presentes en una muestra de agua contaminada tiene una distribución de Poisson con una media de dos colonias de bacterias por $\text{cm}^{\mathbf{3}}$

    a) Si se toman en forma independientes cuatro muestras de $\mathbf{1}\,\text{cm}^{\mathbf{3}}$ de esta agua, encontrar la probabilidad de que al menos una de las muestras tenga una o más colonias de bacterias.
    b) ¿Cuántas muestras de $\mathbf{1}\,\text{cm}^{\mathbf{3}}$ de esta agua deben seleccionarse para tener una probabilidad de aproximadamente $\mathbf{0.95}$ de hallar al menos una colonia de bacterias?

23. Un explorador de petróleo perfora una serie de siete pozos en cierta área hasta encontrar uno productivo. La probabilidad de que tenga éxito en cada prueba es $\mathbf{0.2}$

    a) ¿Cuál es la probabilidad de que el primer pozo productivo sea el tercer pozo perforado?

b) ¿Cuál es la probabilidad de que el explorador no encuentre ningún pozo productivo, si debe suspender la búsqueda al terminar de explorar el décimo pozo perforado?

c) ¿Cuál es la cantidad esperada de pozos perforados hasta hallar el primero productivo?

RESPUESTAS

1. a) $\boldsymbol{\Omega = \{(x,y,z)\ :\ x,y,z \in \{0,1\}\}}$ donde por ejemplo $\boldsymbol{\omega = (1,0,1)}$ representa el resultado en el cual el primero y el tercer lanzamiento resultaron "cara" y el segundo lanzamiento resultó "ceca". Este espacio es equiprobable dado que la moneda es normal. De hecho: $\boldsymbol{P(\{(x,y,z)\}) = (1/2)^3}$ para cualquier $\boldsymbol{\omega = (x,y,z) \in \Omega}$

b)

$$\begin{array}{l|l} \{X=0\} = \{(1,1,1)\} & P(X=0) = 1/2^3 = 1/8 \\ \{X=1\} = \{(0,1,1),(1,0,1),(1,1,0)\} & P(X=1) = 3/2^3 = 3/8 \\ \{X=2\} = \{(0,0,1),(0,1,0),(1,0,0)\} & P(X=2) = 3/2^3 = 3/8 \\ \{X=3\} = \{(0,0,0)\} & P(X=3) = 1/2^3 = 1/8 \end{array}$$

La suma de las cuatro probabilidades es $\mathbf{1}$ Esto se debe a que $\boldsymbol{\Omega = \biguplus_{x \in R_X} \{X = x\}}$

$$\begin{array}{l|l} \{Y=-3\} = \{(1,1,1)\} & P(Y=0) = 1/2^3 = 1/8 \\ \{Y=-1\} = \{(0,1,1),(1,0,1),(1,1,0)\} & P(Y=1) = 3/2^3 = 3/8 \\ \{Y=1\} = \{(0,0,1),(0,1,0),(1,0,0)\} & P(Y=2) = 3/2^3 = 3/8 \\ \{Y=3\} = \{(0,0,0)\} & P(Y=3) = 1/2^3 = 1/8 \end{array}$$

La suma de las cuatro probabilidades es $\mathbf{1}$ Esto se debe a que $\boldsymbol{\Omega = \biguplus_{y \in R_Y} \{Y = y\}}$

$$\begin{array}{l|l} \{Z=0\} = \{(0,0,0)\} & P(Y=0) = 1/2^3 = 1/8 \\ \{Z=1/3\} = \{(0,0,1),(0,1,0),(1,0,0)\} & P(Y=1) = 3/2^3 = 3/8 \\ \{Z=2/3\} = \{(0,1,1),(1,0,1),(1,1,0)\} & P(Y=2) = 3/2^3 = 3/8 \\ \{Z=1\} = \{(1,1,1)\} & P(Y=3) = 1/2^3 = 1/8 \end{array}$$

La suma de las cuatro probabilidades es $\mathbf{1}$ Esto se debe a que $\boldsymbol{\Omega = \biguplus_{z \in R_Z} \{Z = z\}}$

c) $\boldsymbol{R_X = \{0,1,2,3\}}$, $\boldsymbol{R_Y = \{-3,-1,1,3\}}$, $\boldsymbol{R_Z = 0, 1/3, 2/3, 1}$ Las tres variables aleatorias son discretas pues sus rangos son a lo sumo numerables (de hecho son finitos).

d)

$$\begin{array}{l} \{X \leq 0.5\} = \{(1,1,1)\} \\ \{X < 1\} = \{(1,1,1)\} \\ \{1 \leq X < 4\} = \{(0,1,1),(1,0,1),(1,1,0),(0,0,1),(0,1,0),(1,0,0)\} \\ \{X \neq 2\} = \{(1,1,1),(0,1,1),(1,0,1),(1,1,0),(0,0,0)\} \\ P(X < 0.5) = 1/8 \\ P(X < 1) = 1/8 \\ P(1 \leq X < 4) = 3/4 \\ P(X \neq 2) = 5/8 \end{array}$$

La suma de las cuatro probabilidades NO es $\mathbf{1}$ Esto se debe a que los cuatro eventos no forman una partición de $\boldsymbol{\Omega}$ Por ejemplo no son dos a dos disjuntos

$$\begin{aligned}
&\{Y < -3\} = \emptyset \\
&\{Y > -1\} = \{(0,0,1),(0,1,0),(1,0,0),(0,0,0)\} \\
&\{-1 \leq Y \leq 1\} = \{(0,1,1),(1,0,1),(1,1,0),(0,0,1),(0,1,0),(1,0,0)\} \\
&\{Y \geq 1.7\} = \{(0,0,0)\} \\
&P(Y < -3) = 0 \\
&P(Y > -1) = 1/2 \\
&P(-1 \leq Y \leq 1) = 3/4 \\
&P(Y \geq 1.7) = 1/8
\end{aligned}$$

La suma de las cuatro probabilidades NO es $\mathbf{1}$ Esto se debe a que los cuatro eventos no forman una partición de $\boldsymbol{\Omega}$ Por ejemplo no son dos a dos disjuntos

$$\begin{aligned}
&\{Z \geq 3/5\} = \{(0,1,1),(1,0,1),(1,1,0),(1,1,1)\} \\
&\{Z \geq 1\} = \{(1,1,1)\} \\
&\{Z \leq 0\} = \{(0,0,0)\} \\
&\{Z \leq 1\} = \Omega \\
&P(Z \geq 3/5) = 1/2 \\
&P(Z \geq 1) = 1/8 \\
&P(Z \leq 0) = 1/8 \\
&P(Z \leq 1) = 1
\end{aligned}$$

La suma de las cuatro probabilidades NO es $\mathbf{1}$ Esto se debe a que los cuatro eventos no forman una partición de $\boldsymbol{\Omega}$ Por ejemplo no son dos a dos disjuntos.

e)

$$\begin{aligned}
&\{X \leq 1\,,\, Y > -3\} = \{(0,0,0),(0,1,1),(1,0,1),(1,1,0)\} \cap \\
&\ \cap \{(0,1,1),(1,0,1),(1,1,0),(0,0,1),(0,1,0),(1,0,0),(0,0,0)\} = \\
&= \{(0,0,0),(0,1,1),(1,0,1),(1,1,0)\}
\end{aligned}$$

Luego: $P(X \leq 1\,,\, Y > -3) = 1/2$

2. Las fmp de X, Y, Z ya fueron calculadas en 1.b) Resumimos los resultados en las siguientes tablas:

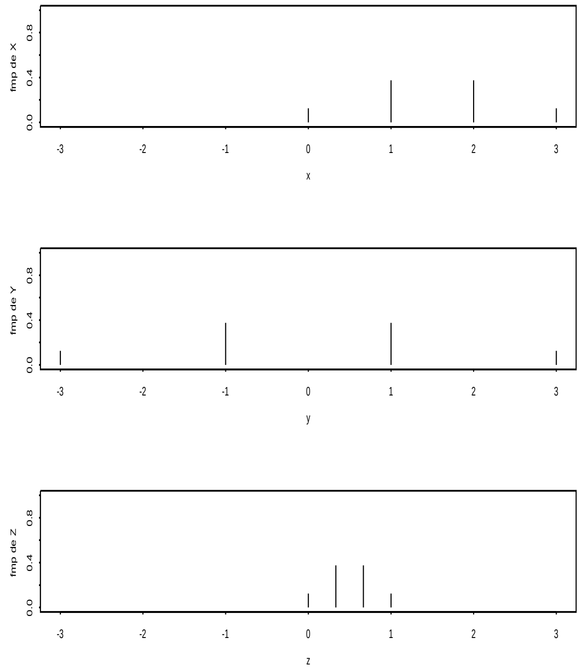
x	0	1	2	3
$p_X(x)$	$1/8$	$3/8$	$3/8$	$1/8$
y	-3	-1	1	3
$p_Y(y)$	$1/8$	$3/8$	$3/8$	$1/8$
z	0	$1/3$	$2/3$	1
$p_Z(z)$	$1/8$	$3/8$	$3/8$	$1/8$

Los gráficos indican que Z es la de menor varianza e Y la de mayor varianza. Resumimos los cálculos necesarios para verificarlo analíticamente:

$$E(X) = 0 \cdot 0.125 + 1 \cdot 0.375 + 2 \cdot 0.375 + 3 \cdot 0.125 = 1.5$$
$$E(X^2) = 0^2 \cdot 0.125 + 1^2 \cdot 0.375 + 2^2 \cdot 0.375 + 3^2 \cdot 0.125 = 3$$
$$V(X) = 3 - (1.5)^2 = 0.75$$

$$E(Y) = (-3) \cdot 0.125 + (-1) \cdot 0.375 + 1 \cdot 0.375 + 3 \cdot 0.125 = 0$$
$$E(Y^2) = (-3)^2 \cdot 0.125 + (-1)^2 \cdot 0.375 + 1^2 \cdot 0.375 + 3^2 \cdot 0.125 = 3$$
$$V(Y) = 3 - 0^2 = 3$$

$$E(Z) = 0 \cdot 0.125 + 1/3 \cdot 0.375 + 2/3 \cdot 0.375 + 1 \cdot 0.125 = 0.5$$
$$E(Z^2) = 0^2 \cdot 0.125 + (1/3)^2 \cdot 0.375 + (2/3)^2 \cdot 0.375 + 1^2 \cdot 0.125 = 1/3 \approx 0.333$$
$$V(Z) = 1/3 - (0.5)^2 = 1/12 \approx 0.083$$

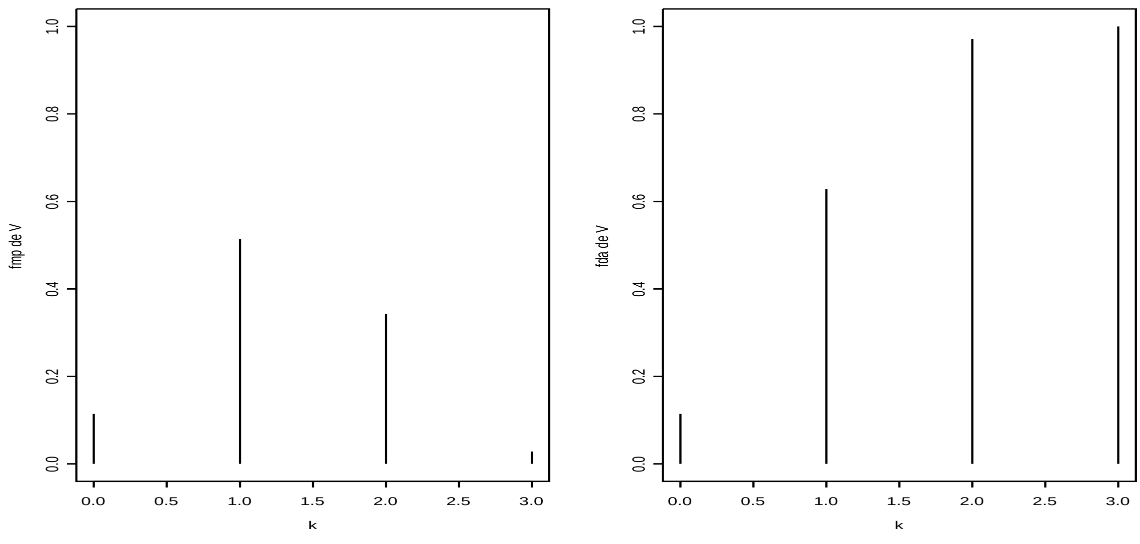


3. a) $\boldsymbol{R_V = \{0, 1, 2, 3\}}$

$$\boldsymbol{P(V = k) = \frac{\binom{3}{k}\binom{4}{3-k}}{\binom{7}{3}} \quad (k = 0, 1, 2, 3)}$$

De hecho: $\boldsymbol{V \sim \mathcal{H}(3, 3, 7)}$ Escribamos la fmp y la fda en forma de tabla:

$\boldsymbol{k}$	**0**	**1**	**2**	**3**
$\boldsymbol{p_V(k)}$	**4/35**	**18/35**	**12/35**	**1/35**
$\boldsymbol{F_V(k)}$	**4/35**	**22/35**	**34/35**	**1**



b) $P(V \le 1) = F_V(1) = 22/35 \approx 0.629$

c) $\{V \ge 1\}$ = "sale al menos una bolita verde"
$P(V \ge 1) = 1 - P(V < 1) = 1 - P(V = 0) = 1 - p_V(0) = 1 - 4/35 = 31/35 \approx 0.886$

4. a) $\Omega = \{1, 01, 001, 0001, \cdots\}$ donde $\underbrace{0\cdots0}_{k-1}1$ indica el resultado en el cual se arroja la moneda k veces, saliendo "ceca" en los primeros $k-1$ intentos y "cara" en el k-ésimo intento.

b) $R_X = \mathbb{N}$
La fmp de X viene dada por: $P(X = k) = (1/2)^k$ para $k \in \mathbb{N}$ Podemos expresar la fda de X de manera elegante y fácilmente calculable si tenemos en cuenta que:

$$F_X(k) \;=\; P(X \le k) = \sum_{j=1}^{k} p_X(j) = \sum_{j=1}^{k} (1/2)^j = 1 - (1/2)^k$$

donde hemos utilizado que:

$$ap^r + ap^{r+1} + \cdots + ap^{r+s} = ap^r\,\frac{1-p^{s+1}}{1-p}$$

Notemos de paso que $X \sim \mathcal{G}(1/2)$

c) y d)

$A = \{X \ge 3\}$	$P(A) = 1 - F_X(2) = (1/2)^2 = 1/4 = 0.25$
$B = \{X \le 3\}$	$P(B) = F_X(3) = 1 - (1/2)^3 = 7/8$
$C = \{X \in \{1, 3, 5, \cdots\}\}$	
$D = \{X \in \{5, 10, 15, 20, \cdots\}\}$	

$$P(D) \;=\; \sum_{s=1}^{\infty} p_X(5s) = \sum_{s=1}^{\infty} (1/2)^{5s} = (1/2)^5\,\frac{1}{1-(1/2)^5} = \frac{1}{31}$$

$P(X \le 3) = P(B) = 7/8$; $P(1 < X \le 4) = p_X(2) + p_X(3) + p_X(4) = 7/16$

5. $R_X = \{0, 1, 2, 3, 4\}$

a) Si las extraccines son con reposición entonces $X \sim \mathcal{B}i(4; 0.2)$ de modo que la fmp de X resulta: $p_X(k) = \binom{4}{k}(0.2)^k\,(0.8)^{4-k}$ Resumiéndola en una tabla:

k	0	1	2	3	4
$p_X(k)$	0.4096	0.4096	0.1536	0.0256	0.0016

b) Si las extraccines son sin reposición entonces $X \sim \mathcal{H}i(4, 5, 25)$ de modo que la fmp de X resulta: $p_X(k) = \frac{\binom{5}{k}\binom{20}{4-k}}{\binom{25}{4}}$ Resumiéndola en una tabla:

k	0	1	2	3	4
$p_X(k)$	0.3830	0.4506	0.1502	0.0158	0.0004

6. a) $R_X = \mathbb{N}$ Siendo R_X a lo sumo numerable resulta X variable aleatoria discreta.

b) Dado que $X \sim \mathcal{G}(0.5)$ se tiene: $p_X(k) = (0.5)^{k-1}(0.5) = (0.5)^k$ Claramente es $p_X(k) \ge 0$, $(k = 1, 2, \cdots)$ y además: $\sum_{k=1}^{\infty} p_X(k) = \sum_{k=1}^{\infty} (1/2)^k = \frac{1/2}{1-1/2} = 1$ La fda de X resulta: $F_X(k) = P(X \le k) = \sum_{j=1}^{k} (1/2)^j = \frac{1}{2}\,\frac{1-(1/2)^{k+1}}{1-1/2} = 1 - (1/2)^{k+1}$

c)

$$\begin{aligned} P(X > r+s | X > r) &= \frac{P(X>r+s\,,X>r)}{P(X>r)} = \frac{P(X>r+s)}{P(X>r)} = \frac{1-F_X(r+s)}{1-F_X(r)} = \\ &= \frac{(1/2)^{r+s}}{(1/2)^r} = (1/2)^s = 1 - F_X(s) = P(X > s) \end{aligned}$$

7. Para abreviar anotemos las dos propiedades básicas de una fmp como:

$$\begin{array}{ll} \text{(FMP1):} & \forall x\,,\ p(x) \geq 0 \\ \text{(FMP2):} & \sum\limits_{x \text{ tq } p(x)>0} p(x) = 1 \end{array}$$

a) Se verifica evidentemente (FMP1). Además: $\sum\limits_{x \text{ tq } p(x)>0} p(x) = p(1)+p(3)+p(4) = 1/4 + 1/4 + 1/2 = 1$ De manera que también se verifica (FMP2). Luego, $p(x)$ es la fmp de cierta variable aleatoria X Calculemos las probabilidades pedidas:

$$\begin{array}{c|c} P(1 \leq X \leq 4) = p(1) + p(3) + p(4) = 1 & P(X = 1) = p(1) = 1/4 \\ P(X \neq 2) = p(1) + p(3) + p(4) = 1 & P(X > 2) = p(3) + p(4) = 3/4 \end{array}$$

b) Se verifica evidentemente (FMP1). Además: $\sum\limits_{x \text{ tq } p(x)>0} p(x) = p(1)+p(2) = 1/4 + 1/2 = 3/4 \neq 1$ De manera que no se verifica (FMP2) por lo cual $p(x)$ no es una fmp.

c) En este caso no se verifica (FMP1) pues $p(4) = -\frac{1}{18} < 0$ Luego, $p(x)$ no es una fmp.

d) Evidentemente se verifica (FMP1). Además:

$$\sum_{x \text{ tq } p(x)>0} p(x) = \sum_{x=1}^{4} \frac{x^2}{30} = \frac{1}{30} \sum_{x=1}^{4} x^2 = \frac{1}{30}\,\frac{4 \cdot 5 \cdot 9}{6} = 1$$

Por lo tanto $p(x)$ es una fmp. Calculemos las probabilidades pedidas:

$$\begin{array}{c|c} P(1 \leq X \leq 4) = p(1) + p(2) + p(3) + p(4) = 1 & P(X = 1) = p(1) = \frac{1}{30} \\ P(X \neq 2) = p(1) + p(3) + p(4) = \frac{13}{15} & P(X > 2) = p(3) + p(4) = 5/6 \end{array}$$

8. Recordemos que las definiciones de esperanza y varianza de una v.a. discreta son:

$$E(X) = \sum_{x \in R_X} x \cdot p_X(x) \qquad ; \qquad V(X) = \sum_{x \in R_X} (x - E(X))^2 \cdot p_X(x)$$

donde la primera serie se supone absolutamente CV (la serie de los valores absolutos es finita) y la segunda serie se supone CV (finita).

a) $E(X) = 0 \cdot (0.08) + 1 \cdot (0.15) + 2 \cdot (0.45) + 3 \cdot (0.27) + 4 \cdot (0.05) = 2.06$

b) $V(X) = (0 - 2.06)^2 \cdot (0.08) + (1 - 2.06)^2 \cdot (0.15) + (2 - 2.06)^2 \cdot (0.45) + (3 - 2.06)^2 \cdot (0.27) + (4 - 2.06)^2 \cdot (0.05) = 0.9364$

c) $SD(X) \overset{\text{def}}{=} \sqrt{V(X)} = \sqrt{0.9364} \approx 0.9677$

d)

$$\begin{array}{l} E(X^2) = 0^2 \cdot (0.08) + 1^2 \cdot (0.15) + 2^2 \cdot (0.45) + 3^2 \cdot (0.27) + 4^2 \cdot (0.05) = 5.18 \\ V(X) = 5.18 - (2.06)^2 = 0.9364 \end{array}$$

9. Recordemos que para calcular la esperanza de una función $W = g(Y)$ de una v.a. X podemos proceder de dos formas alternativas:
Utilizando la fmp de Y:
$$E(W) \sum_{y \in R_Y} g(y)\, p_Y(y)$$
Hallando primero la fmp de W y luego calculando $E(W)$ por definición:
$$E(W) \sum_{w \in R_W} w\, p_W(w)$$
   a) $E(Y) = 1 \cdot (0.01) + 2 \cdot (0.19) + 3 \cdot (0.35) + 4 \cdot (0.45) = 3.24$
   b) $E(\sqrt{Y}) = \sqrt{1} \cdot (0.01) + \sqrt{2} \cdot (0.19) + \sqrt{3}1 \cdot (0.35) + \sqrt{4} \cdot (0.45) = 1.784918$
   c)
$$E(Y^2) = 1^2 \cdot (0.01) + 2^2 \cdot (0.19) + 3^2 \cdot (0.35) + 4^2 \cdot (0.45) = 11.12$$
$$V(Y) = E(Y^2) - E^2(Y) = (11.12)^2 - (3.24)^2 = 0.6224$$
El desvío standard de Y es $\sigma_Y = SD(Y) = \sqrt{V(Y)} = \sqrt{0.6224} \approx 0.7889$
$$\begin{aligned} &P\left(|Y - E(Y)| \leq SD(Y)\right) \approx P\left(|Y - 3.24| \leq 0.7889\right) = \\ &= P(-0.7889 \leq Y - 3.24 \leq 0.7889) = P(3.24 - 0.7889 \leq Y \leq 3.24 + 0.7889) = \\ &= P(3.24 - 0.7889 \leq Y \leq 3.24 + 0.7889) = P(2.4511 \leq Y \leq 4.0289) = \\ &= p_Y(3) + p_Y(4) = 0.35 + 0.45 = 0.8 \end{aligned}$$

10. a)
$$E(X) = (13.5)(0.2) + (15.9)(0.5) + (19.2)(0.3) = 16.38$$
$$E(X^2) = (13.5)^2(0.2) + (15.9)^2(0.5) + (19.2)^2(0.3) = 272.298$$
$$V(X) = 272.298 - (16.38)^2 = 3.9936$$
   b) $E(C) = E(25X - 8.5) = 25\, E(X) - 8.5 = 25(16.38) - 8.5 = 401$
   c) $V(C) = V(25X - 98.5) = (25)^2\, V(X) = (25)^2(3.9936) = 2496$
   d) $E(h(X)) = E(X - 0.01X^2) = E(X) - 0.01\, E(X^2) = 16.38 - (0.01)(272.298) \approx 13.66$

11. a) $E(X^2) = 0^2 \cdot (1-p) + 1^2 \cdot p = p$
   b) $V(X) = E(X^2) - E^2(X) = p - p^2 = p(1-p)$
   c) $E(X^{79}) = 0^{79} \cdot (1-p) + 1^{79} \cdot p = p$

12. La varianza existe sii existe $E(X^2)$ En este caso dicha esperanza no existe pues:
$$E(X^2) = \sum_{x=1}^{\infty} x^2 \frac{c}{x^3} = \sum_{x=1}^{\infty} \frac{c}{x} = \infty$$
dado que la "serie armónica" es divergente. Luego, la v.a. X no posee varianza (finita).

13. En el ejercicio 1.c1) se calculó indirectamente la fmp de X dando por resultado que X tiene distribución uniforme discreta en $\{1, 2, \cdots, n\}$ de manera que $P_X(k) = n^{-1}$ para $k = 1, \cdots, n$ Entonces:
$$E(X) = \sum_{k=1}^{n} k\, p_X(k) = \sum_{k=1}^{n} k\, \tfrac{1}{n} = \tfrac{1}{n} \sum_{k=1}^{n} k = \tfrac{1}{n} \tfrac{n(n+1)}{2} = \tfrac{n+1}{2}$$
$$E(X^2) = \sum_{k=1}^{n} k^2\, p_X(k) = \sum_{k=1}^{n} k^2\, \tfrac{1}{n} = \tfrac{1}{n} \sum_{k=1}^{n} k^2 = \tfrac{1}{n} \tfrac{n(n+1)(2n+1)}{6} = \tfrac{(n+1)(2n+1)}{6}$$
$$V(X) = E(X^2) - E^2(X) = \tfrac{(n+1)(2n+1)}{6} - \left(\tfrac{n+1}{2}\right)^2 = \tfrac{n^2-1}{12}$$

14. Ejemplifiquemos con el ejercicio 10 calculando $V(-X)$ Se tiene:

$$\begin{aligned} &E(-X) = (-13.5)(0.2) + (-15.9)(0.5) + (-19.1)(0.3) = \\ &= -\left((13.5)(0.2) + (15.9)(0.5) + (19.1)(0.3)\right) = -16.38 \\ &E\left((-X)^2\right) = (-13.5)^2(0.2) + (-15.9)^2(0.5) + (-19.1)^2(0.3) = 272.298 \\ &V(-X) = E\left((-X)^2\right) - E^(- X) = 272.298 - (-16.38)^2 = 3.9936 = V(X) \end{aligned}$$

En general:

$$V(-X) = E\left((-X)^2\right) - E^2\left(-X\right) = = E(X^2) - (-E(X))^2 = E(X^2) - E^2(X) = V(X)$$

15.

$$\begin{aligned} V(aX+b) &= E\left\{[(aX+b) - E(aX+b)]^2\right\} = E\left\{[(aX+b) - (aE(X)+b)]^2\right\} = \\ &= E\left\{[a(X - E(X))]^2\right\} = E\left\{a^2(X - E(X))^2\right\} = \\ &= a^2\, E\left\{(X - E(X))^2\right\} = a^2\, V(X) \end{aligned}$$

16. a) $R_Y = \{0,1,2\}$ Además:

$$\begin{aligned} &p_Y(0) = 0.2 = P\left((A \cup B)'\right) \text{ Luego: } P(A \cup B) = 1 - 0.2 = 0.8 \\ &p_Y(2) = P(AB) = P(A) + P(B) - P(A \cup B) = 0.4 + 0.5 - 0.8 = 0.1 \\ &p_Y(1) = P\left((A \cup B) \setminus (AB)\right) = P(A \cup B) - P(AB) = 0.8 - 0.1 = 0.7 \end{aligned}$$

Luego, la fmp de Y viene dada por:

y	0	1	2
$p_Y(y)$	0.2	0.7	0.1

b)

$$\begin{aligned} &E(Y) = 0 \cdot (0.2) + 1 \cdot (0.7) + 2 \cdot (0.1) = 0.9 \\ &E(Y^2) = 0^2 \cdot (0.2) + 1^2 \cdot (0.7) + 2^2 \cdot (0.1) = 1.1 \\ &V(Y) = 1.1 - (0.9)^2 = 0.3 \end{aligned}$$

17. a) $R_N = \{1,2,3\}$ Pensemos en el espacio muestral $\Omega = \{(x,y,z) \,:\, x,y,z \in \{1,2,3\}\}$ donde por ejemplo $\omega = (2,3,1)$ indica que la bolita 1 está en la caja 2, la bolita 2 en la caja 3 y la bolita 3 en la caja 1. Naturalmente este espacio es equiprobable y vale $\#(\Omega) = 3^3$ Entonces:

$$p_N(1) = \binom{3}{1}/3^3 = 1/9$$

$$p_N(2) = \frac{\binom{3}{2}\binom{2}{1}\frac{3!}{2!\,1!}}{3^3} = 2/3$$

$$p_N(3) = \frac{3!}{3^3} = 2/9$$

La fmp de N queda resumida en la tabla:

n	1	2	3
$p_N(n)$	1/9	2/3	2/9

La fda de N resulta:

n	1	2	3
$F_N(n)$	1/9	7/9	1

b)

$$\begin{aligned} &E(N) = 1 \cdot (1/9) + 2 \cdot (2/3) + 3 \cdot (2/9) = 19/9 \\ &E(N^2) = 1^2 \cdot (1/9) + 2^2 \cdot (2/3) + 3^2 \cdot (2/9) = 43/9 \\ &V(N) = 43/9 - (19/9)^2 = 26/81 \approx 0.321 \end{aligned}$$

18. a) $R_X = \{1, 2, 3, 4, 5, 6\}$

$$p_X(1) = \frac{6+5}{6^2} = 11/36 \quad p_X(2) = \frac{5+4}{6^2} = 1/4 \quad p_X(3) = \frac{4+3}{6^2} = 7/36$$
$$p_X(4) = \frac{3+2}{6^2} = 5/36 \quad p_X(5) = \frac{2+1}{6^2} = 1/12 \quad p_X(6) = \frac{1}{6^2} = 1/36$$

Resumimos la fmp y la fda en una tablita:

x	1	2	3	4	5	6
$p_X(x)$	$11/36$	$1/4$	$7/36$	$5/36$	$1/12$	$1/36$
$F_X(x)$	$11/36$	$5/9$	$3/4$	$8/9$	$35/36$	1

b)

$$E(X) = 1 \cdot \frac{11}{36} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{7}{36} + 4 \cdot \frac{5}{36} + 5 \cdot \frac{1}{12} + 6 \cdot \frac{1}{36} = 91/36 = 2.527778$$
$$E(X^2) = 1^2 \cdot \frac{11}{36} + 2^{\cdot} \frac{1}{4} + 3^2 \cdot \frac{7}{36} + 4^2 \cdot \frac{5}{36} + 5^2 \cdot \frac{1}{12} + 6^2 \cdot \frac{1}{36} = 301/36 = 8.361111$$
$$V(X) = 301/36 - (91/36)^2 = 2555/1296 = 1.971451$$

19. a) Podemos pensar en una sucesión de $n = 3$ ensayos de Bernoulli, con probabilidad de éxito en cada ensayo igual a $p = 0.8$ Esto se debe a que cada celda se activa en forma INDEPENDIENTE de las demás (cuando la temperatura supera los $100°C$) y al hecho que la probabilidad de que ello ocurra es la misma para las tres celdas. Entonces dado que Y cuenta la cantidad de éxitos (celdas activadas cuando la temperatura supera los $100°C$) se tiene: $Y \sim \mathcal{B}i(3, 0.8)$ Entonces: $p_Y(y) = \binom{3}{y}(0.8)^y(0.2)^{3-y}$ para $y = 0, 1, 2, 3$ Esto se resume en la siguiente tabla:

y	0	1	2	3
$p_Y(y)$	0.008	0.096	0.384	0.512

b) $P(Y \geq 1) = 1 - P(Y = 0) = 1 - p_Y(0) = 1 - 0.008 = 0.992$

c) Una opción para calcular $E(X)$ y $V(X)$ es utilizando la tabla de la fmp recién obtenida y procediendo como en ejercicios anteriores (Recomiendo hacerlo). Otra posibilidad es utilizar las fórmulas de la esperanza y la varianza de una binomial:

$$E(X) = np = 3(0.8) = 2.4 \quad ; \quad V(X) = npq = 3(0.8)(1 - 0.8) = 0.48$$

20. Tomando la unidad de tiempo como 1 hora, resulta $\lambda = 4$ Si $\{X_t\}_{t>0}$ sigue un proceso de Poisson de tasa λ, entonces por definición es $X_t \sim \mathcal{P}(\lambda t)$ Además recordemos que X_t mide la cantidad de ocurrencias del evento (solicitudes de asistencia de remolque) en un intervalo de duración t

a) $P(X_1 = 10) = e^{-4} \cdot \frac{4^{10}}{(10)!} = 0.005292477 \approx 0.0053$

b) $P(X_{0.5} = 0) = e^{-4(0.5)} \cdot \frac{(4(0.5))^0}{0!} = e^{-2} \approx 0.1353$

c) $P(X_{0.25} = 2) = e^{-4(0.25)} \cdot \frac{(4(0.25))^2}{2!} \approx 0.1839$

21. En este caso si fijamos la unidad de tiempo como 1 hora, se tiene $\lambda = 7$

a) $P(X_1 \leq 3) = e^{-7}\left(1 + \frac{7}{1!} + \frac{7^2}{2!} + \frac{7^3}{3!}\right) \approx 0.0818$

b) $P(X_1 \geq 2) = 1 - P(X_1 \leq 1) = 1 - e^{-7}\left(1 + \frac{7}{1!}\right) \approx 0.9927$

c) $P(X_1 \geq 5) = 1 - P(X_1 \leq 4) = 1 - e^{-7}\left(1 + \frac{7}{1!} + \frac{7^2}{2!} + \frac{7^3}{3!} + \frac{7^4}{4!}\right) \approx 0.8270$

22. Fijamos la unidad de volumen en $1\,\text{cm}^3$ Entonces $\lambda = 2$

a) Podemos pensar en una sucesión de $\boldsymbol{n = 4}$ ensayos de Bernoulli, con probabilidad de éxito en cada ensayo igual a $\boldsymbol{p}$, que determinaremos luego. Acá, cada ensayo consiste en ver si la muestra tiene o no al menos una colonia. Se trata efectivamente de ensayos de Bernoulli dado que las muestras de agua son independientes entre sí y además la probabilidad de éxito ("hallar al menos una colonia en la muestra") $\boldsymbol{p}$ es la misma para las cuatro muestras de agua (porque se supone que se toman las cuatro muestras de la misma agua). Por lo tanto, si $\boldsymbol{Y}$ cuenta cuántas de las $\mathbf{4}$ muestras presentan al menos una colonia, entonces $\boldsymbol{Y \sim \mathcal{B}i(4, p)}$ y lo que se pretende es calcular $\boldsymbol{P(Y \geq 1)}$
Ahora, para determinar $\boldsymbol{p}$ utilizamos la distribución de Poisson puesto que $\boldsymbol{p}$ representa la probabilidad de hallar al menos una colonia en la muestra de agua considerada. Es decir:

$$\boldsymbol{p = P(X_1 \geq 1) = 1 - P(X_1 = 0) = 1 - e^{-2} \cdot \frac{2^0}{0!} = 1 - e^{-2} \approx 0.8647}$$

Luego:

$$\boldsymbol{P(Y \geq 1) = 1 - P(Y = 0) = 1 - (1-p)^4 = 1 - \left(e^{-2}\right)^4 = 1 - e^{-8} \approx 0.9997}$$

b) En este caso se pretende averiguar la cantidad $\boldsymbol{n}$ de muestras a tomar, es decir la cantidad $\boldsymbol{n}$ de ensayos de Bernoulli necesarios. El requerimiento del enunciado se expresa como (donde $\boldsymbol{p = 1 - e^{-2}}$):

$$\boldsymbol{Y \sim \mathcal{B}i(n, p) \quad \wedge \quad P(Y \geq 1) \approx 0.95 \quad \Rightarrow \quad n = ?}$$

Es decir que debemos determinar $\boldsymbol{n}$ de modo que: $\boldsymbol{1 - (1-p)^n \approx 0.95}$ O sea: $\boldsymbol{1 - e^{-2n} \approx 0.95}$
Despejando la exponencial, tomando logaritmo natural y despejando n se deduce que: $\boldsymbol{n \approx -(0.5)\ \ln(0.05) \approx 1.5}$ Entonces basta tomar $\boldsymbol{n = 2}$ muestras de agua.

23. ME CANSÉ. Después se los hago!

Trabajo práctico 5: Variables aleatorias continuas

1. Sea X v.a. con fdp dada por:

$$f_X(x) = \begin{cases} c(2-x) & \text{si} \quad 0 \leq x \leq 2 \\ 0 & \text{CC} \end{cases}$$

   a) Determinar el valor de la constante c
   b) Hallar F_X
   c) Graficar f_X y F_X
   d) Utilizando F_X calcular $P(1 \leq X < 2)$ y $P(|X| > 1)$
   e) Calcular $P(1 \leq X < 2)$ como el área debajo de la fdp entre los valores que correspondan.
   f) Hallar $E(X)$, $V(X)$ y el tercer cuartil de la distribución de X

2. El período T de funcionamiento hasta su primera falla (en cientos de horas) para cierto transistor es una v.a. con fda dada por: $F_T(t) = (1 - e^{-t^2})\, I_{(0,\infty)}(t)$

   a) Mostrar que F_T tiene las propiedades de una fda.
   b) Obtener la fdp de T
   c) Calcular la probabilidad de que el transistor funcione al menos **200** horas antes de presentar su primera falla.
   d) Calcular $E(T)$, $V(T)$ y la mediana de la distribución de T

3. Para determinar el grado de inteligencia de un ratón se mide el tiempo que tarda en recorrer un laberinto para encontrar la comida (estímulo). El tiempo (en segundos) que emplea un ratón es una v.a. T con fdp dada por $f_T(t) = bt^{-2}\, I_{[b,\infty)}(t)$ siendo b el tiempo mínimo necesario para recorrer el laberinto.

   a) Demostrar que f_T tiene las propiedades de una fdp.
   b) Determinar F_T
   c) Calcular $P(T > b + c)$ siendo $c > 0$ una constante
   d) Calcular los cuartiles de la distribución de T
   e) Determinar si existen y en tal caso calcular $E(T)$ y $V(T)$

4. Si un paracaidista cae en un sitio aleatorio del segmento de recta que une los puntos A y B

   a) Encuentre la probabilidad de que caiga más cerca de A que de B
   b) Calcule la probabilidad de que caiga en un sitio cuya distancia a A sea más del triple de su distancia a B

5. La duración T en horas de una lámpara de cierto aparato electrónico es una v.a. con distribución exponencial de parámetro $\lambda = 0.001$

   a) Calcular la probabilidad de que la lámpara dure más de **500** horas.
   b) Calcular la probabilidad de que la duración de la lámpara se aparte de su duración esperada en menos que un desvío standard.
   c) ¿Cuántas horas mínimas de funcionamiento debe garantizar una compañía fabricante de lámparas para que haya un probabilidad de al menos **0.95** de que la lámpara dure al menos el mínimo garantizado?
   d) Demuestre que $P(T > 1500 | T > 700) = P(T > 800)$

6. Sea $\boldsymbol{Z}$ una v.a. normal standard. Utilizando tablas normales:
   a) Calcular las siguientes probabilidades:
      a1) $\boldsymbol{P(0 < Z < 1.2)}$
      a2) $\boldsymbol{P(-0.9 < Z \leq 0)}$
      a3) $\boldsymbol{P(0.3 \leq Z < 1.56)}$
      a4) $\boldsymbol{P(|Z| < 0.2)}$
      a5) $\boldsymbol{P(|Z| > 1.53)}$
   b) Determinar el valor $\boldsymbol{z}$ tal que:
      b1) $\boldsymbol{P(Z > z) = 0.5}$
      b2) $\boldsymbol{P(Z < z) = 0.8643}$
      b3) $\boldsymbol{P(|Z| < z) = 0.90}$
      b4) $\boldsymbol{P(|Z| > z) = 0.85}$
      b5) $\boldsymbol{P(Z < z) = 0.0054}$

7. Sea $\boldsymbol{X}$ es una v.a. con distribución normal de parámetros $\boldsymbol{\mu = 10}$ y $\boldsymbol{\sigma^2 = 36}$
   a) Calcular: $\boldsymbol{P(X > 5)}$, $\boldsymbol{P(4 < X < 16)}$, $\boldsymbol{P(X \leq 8)}$
   b) Determinar la constante $\boldsymbol{c}$ que verifica: $\boldsymbol{P(X \leq c) = 2\,P(X > c)}$

8. El diámetro de los pernos producidos por una fábrica tiene distribución normal con una media de **950** mm y un desvío standard de **10** mm
   a) ¿Cuál es el valor $\boldsymbol{c}$ tal que la probabilidad de que un perno elegido al azar tenga un diámetro menor que $\boldsymbol{c}$ sea de **0.8531** ?
   b) ¿Cuánto debería reducirse el desvío standard en la producción de pernos para que con la constante $\boldsymbol{c}$ determinada en el item anterior, la probabilidad de tener un diámetro menor que $\boldsymbol{c}$ ascienda a **0.9** ?

9. La resistencia a la ruptura de cierto tipo A de acero es una variable aleatoria con distribución normal con media **43** y desviación típica **4.4** Para el tipo B de acero la resistencia a la ruptura tiene distribución normal con media **44** y desviación típica **6.1** Si se prueba una muestra de cada tipo de acero, ¿para cuál de ellos es menor la probabilidad de que la resistencia sea a lo sumo **40** ?

10. Una fábrica produce tornillos. Las especificaciones indican que el diámetro de los mismos debe estar entre **1.19** y **1.21** pulgadas. Supongamos que el proceso de produción es tal que el diámetro de los tornillos es una variable aleatoria con distribución normal con media **1.196** y desvío standard **0.005**
    a) ¿Qué porcentaje de la producción no satisface las especificaciones?
    b) ¿Cuál debería ser la media para que el porcentaje que no satisfaga las especificaciones sea mínimo?

11. Sea $\boldsymbol{X}$ una v.a. con distribución uniforme en el intervalo $\boldsymbol{(0,1)}$ Sean $\boldsymbol{a}$, $\boldsymbol{b}$ constantes reales, con $\boldsymbol{a \neq 0}$ Obtener la fdp de la variable aleatoria $\boldsymbol{Y = aX + b}$

12. Sea $\boldsymbol{U \sim \mathcal{U}(0,1)}$ y sea $\boldsymbol{\lambda > 0}$ una constante. Demostrar que la variable aleatoria $\boldsymbol{X = -\frac{1}{\lambda}\ln(1-U)}$ posee distribución exponencial de parámetro $\boldsymbol{\lambda}$

13. Considere la fdp siguiente:

$$\boldsymbol{f(y)} = \begin{cases} \boldsymbol{e^{-(y-3)}} & \text{si} \quad \boldsymbol{y \geq 3} \\ \boldsymbol{0} & \text{caso contrario} \end{cases}$$

a) Obtener la fdp de la variable aleatoria $\boldsymbol{W}$ en cada uno de los casos siguientes:

a1) $\boldsymbol{W = e^{-Y}}$ a2) $\boldsymbol{W = Y - 3}$ a3) $\boldsymbol{W = Y/3}$

b) Calcular $\boldsymbol{E(W)}$ en los tres casos y de dos formas diferentes.

14. El tiempo $\boldsymbol{T}$ que tarda en realizarse cierta tarea en la construcción de una casa, es una variable aleatoria con distribución exponencial con media $\mathbf{10}$ horas El costo $\boldsymbol{C}$ para completar esta tarea está relacionado con $\boldsymbol{X}$ mediante la fórmula: $\boldsymbol{C = 100 + 40\,T + 3\,T^2}$
Hallar el valor esperado de $\boldsymbol{C}$

15. Se dice que una v.a. $\boldsymbol{X}$ tiene distribución simétrica respecto de $\boldsymbol{\theta}$ (una constante fija) sii se verifica:

$$\forall h > 0\,,\ P(X \leq \theta - h) = P(X \geq \theta + h)$$

a) Dar dos ejemplos de variables aleatorias con distribuciones simétricas, una discreta y otra continua.

b) Sea $\boldsymbol{X}$ una v.a.continua. Probar que las siguientes afirmaciones son equivalentes:

i) $\boldsymbol{X}$ posee distribución simétrica respecto de $\boldsymbol{\theta}$

ii) $\forall x < \theta\,,\ P(X \leq x) = P(X \geq 2\theta - x)$

iii) $\forall x < \theta\,,\ F_X(x) = 1 - F_X(2\theta - x)$

iv) $\forall x < \theta\,,\ f_X(x) = f_X(2\theta - x)$

v) $\forall x > 0\,,\ f_X(\theta - x) = f_X(\theta + x)$

EJERCICIOS ADICIONALES

16. La fracción de alcohol $\boldsymbol{X}$ presente en cierto compuesto puede considerarse una v.a.continua con fdp dada por: $\boldsymbol{f_X(x) = c(1 - x)I_{[0,1]}(x)}$

a) Determinar el valor de la constante $\boldsymbol{c}$

b) Supóngase que el precio de venta del compuesto depende del contenido $\boldsymbol{X}$ de alcohol del modo siguiente: Si $\boldsymbol{X < 1/3}$ el precio es **\$1**, si $\boldsymbol{1/3 \leq X \leq 2/3}$ el precio es **\$2** y si $\boldsymbol{X > 2/3}$ el precio es **\$3** Hallar la distribución del precio de venta de este compuesto.

17. El diámetro $\boldsymbol{D}$ (expresado en $\boldsymbol{dm}$) del tronco de cierta especie de árboles es una v.a. con función de densidad: $\boldsymbol{f_D(x) = kxI_{(0,10)}(x)}$

a) Hallar el valor de la constante $\boldsymbol{k}$

b) ¿Cuál es la probabilidad de que el diámetro de un árbol de esa especie elegido al azar mida entre **4** y **6** dm ?

c) Idem b) pero sabiendo que el diámetro mide más de **5** dm

d) En un área del bosque hay tres árboles de dicha especie. Calcular la probabilidad de que exactamente 2 de ellos tengan diámetro entre **4** y **6** dm

e) ¿Cuántos árboles habría que muestrear en el bosque para que la probabilidad de encontrar al menos uno cuyo diámetro mida entre **4** y **6** dm sea al menos **0.99** ?

18. Se supone que en una cierta población humana el índice cefálico $\boldsymbol{I}$ (ancho del cráneo expresado como porcentaje de su longitud) se distribuye normalmente entre los individuos. Hay un $\mathbf{58\%}$ con $\boldsymbol{I < 75}$, un $\mathbf{38\%}$ con $\boldsymbol{75 \leq I \leq 80}$ y un $\mathbf{4\%}$ con $\boldsymbol{I > 80}$ Determinar la función de densidad del índice cefálico y calcular $\boldsymbol{P(78 \leq I \leq 82)}$

19. Se dice que una v.a. posee distribución lognormal de parámetros μ, σ^2 sii su fdp viene dada por:
$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma\,x}\, e^{-\,(\ln x-\mu)^2/2\sigma^2}\, I(0,\infty)(x)$$
Comprobar que en tal caso se verifica: $Y = \ln X \sim \mathcal{N}(\mu, \sigma^2)$

20. Se dice que una v.a. X posee distribución Weibull de parámetros α, β siendo $\alpha > 0$ sii su fdp está dada por:
$$f_X(x) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left[(x/\alpha)^\beta\right]}\, I_{(0,\infty)}(x)$$
Mostrar que en tal caso: $Y = (X/\alpha)^\beta \sim \mathcal{E}(1)$

21. Sea $X \sim \mathcal{U}(0,1)$ Hallar la fda y la fdp de las siguientes variables aleatorias:
    a) $Y = X^\alpha$ siendo $\alpha \in \mathbb{R}$ con $\alpha \neq 0$
    b) $Y = \ln X$
    c) $Y = \frac{X}{X+1}$

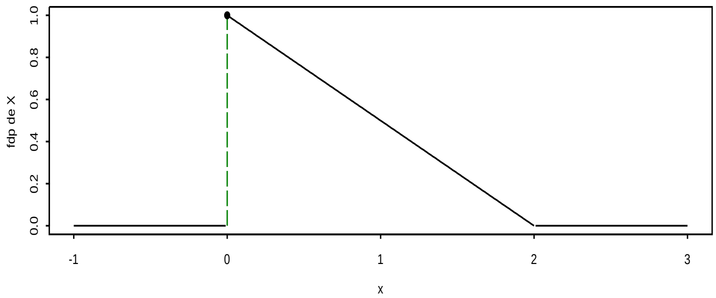
22. Sea $U \sim \mathcal{U}(0,1)$ Encontrar una función g tal que $Y = g(U)$ tenga distribución:
    a) $\mathcal{E}(1)$
    b) Doble exponencial, es decir con fdp dada por: $f_Y(y) = \frac{1}{2}\, e^{-\,|y|}$, $-\infty < y < \infty$
    c) Cauchy, es decir con fdp dada por: $f_Y(y) = \frac{1}{\pi}\,\frac{1}{1+y^2}$, $-\infty < y < \infty$
    d) $\mathcal{B}i(5, 1/3)$

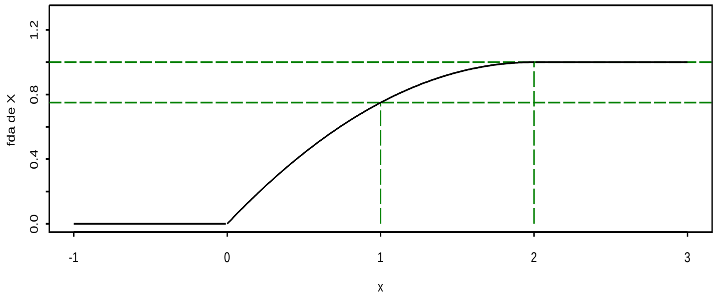
RESPUESTAS

1. a) Planteamos:
$$\int_0^2 c(2-x)\,dx = c\int_0^2 (2-x)\,dx = c\left(2x - \frac{x^2}{2}\right)\Big|_0^2 = 2c = 1$$
Luego $c = 1/2$ Entonces: $f_X(x) = \left(1 - \frac{x}{2}\right) I_{[0,2]}(x)$

   b) Para $0 \le x \le 2$ se tiene:
$$F_X(x) = \int_0^x \left(1 - \frac{x}{2}\right) dx = x - \frac{x^2}{4}$$
Entonces:
$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ x - \frac{x^2}{4} & \text{si} \quad 0 \le x < 2 \\ 1 & \text{si} x \ge 2 \end{cases}$$

   d)
$P(1 \le X \le 2) = F_X(2) - F_X(1) = 1 - 3/4 = 1/4 = 0.25$
$P\left(|X| > 1\right) = 1 - P\left(|X| \le 1\right) = 1 - P(-1 \le X \le 1) = 1 - F_X(1) + F_X(-1) =$
$= 1 - 3/4 = 1/4 = 0.25$





e)

$$P(1 \leq X < 2) = \int_1^2 \left(1 - \frac{x}{2}\right)\, dx = \left(x - \frac{x^2}{4}\right)\Big|_1^2 = 1 - 3/4 = 1/4$$

f)

$$\begin{aligned} E(X) &= \int_0^2 x\left(1 - \tfrac{x}{2}\right)\, dx = \left(\tfrac{x^2}{2} - \tfrac{x^3}{6}\right)\Big|_0^2 = 2 - 4/3 = 2/3 \\ E(X^2) &= \int_0^2 x^2\left(1 - \tfrac{x}{2}\right)\, dx = \left(\tfrac{x^3}{3} - \tfrac{x^4}{8}\right)\Big|_0^2 = 8/3 - 2 = 2/3 \\ V(X) &= 2/3 - (2/3)^2 = 2/9 \\ F(x_{0.75}) &= 0.75 \quad \Rightarrow \quad x_{0.75} - \tfrac{1}{4}x_{0.75}^2 = 0.75 \quad \Rightarrow \quad x_{0.75}^2 - 4x_{0.75} + 3 = 0 \quad \Rightarrow \end{aligned}$$

La cuadrática tiene raíces 1 y 3 Dado que $0 < x_{0.75} < 2$ debemos descartar la segunda raíz. Entonces: $x_{0.75} = 1$ es el tercer cuartil de la distribución de X

2. a) Claramente F_T es no decreciente y continua a derecha (de hecho es continua). Además:

$$\lim_{t\to\infty} F_T(t) = \lim_{t\to\infty}(1 - e^{-t^2}) = 1 \quad \text{y} \quad \lim_{t\to-\infty} F_T(t) = \lim_{t\to-\infty} 0 = 0$$

b) La fdp de T se obtiene derivando la fda:

$$f_T(t) = \begin{cases} 2te^{-t^2} & \text{si} \quad t > 0 \\ 0 & \text{otros casos} \end{cases}$$

c) $P(T \geq 2) = 1 - P(T < 2) = 1 - F_T(2) = 1 - (1 - e^{-2^2}) = e^{-4} \approx 0.018$

d) Utilizamos integración por partes:

$$\begin{aligned}
E(T) &= \int_0^\infty t f_T(t)\,dt = \int_0^\infty 2t^2 e^{-t^2}\,dt = -\,te^{-t^2}\Big|_0^\infty + \int_0^\infty e^{-t^2}\,dt = \\
&= \int_0^\infty e^{-t^2}\,dt = \int_0^\infty e^{-1/2\,s^2}\,\tfrac{ds}{\sqrt{2}} = \sqrt{\pi}\int_0^\infty e^{-1/2\,s^2}\,\tfrac{ds}{\sqrt{2\pi}} = \\
&= \tfrac{\sqrt{\pi}}{2} \\
E(T^2) &= \int_0^\infty t^2 f_T(t)\,dt = \int_0^\infty 2t^3 e^{-t^2}\,dt = -\,t^2e^{-t^2}\Big|_0^\infty + \int_0^\infty 2te^{-t^2}\,dt = \\
&= \int_0^\infty 2te^{-t^2}\,dt = -\,e^{-t^2}\Big|_0^\infty = 1
\end{aligned}$$

Luego: $V(T) = 1 - \left(\frac{\sqrt{\pi}}{2}\right)^2 = 1 - \frac{\pi}{4}$
La mediana de la distribución de T se encuentra hallando $t_{0.5}$ tal que $F_T(t_{0.5}) = 1/2$ Es decir: $1 - e^{-t_{0.5}^2} = 1/2$ Equivalentemente: $-\,t_{0.5}^2 = -\ln 2$ Por lo tanto: $t_{0.5} = (\ln 2)^{1/2} \approx 0.833$

3. a) Claramente f_T es no negativa. Además:

$$\int_{-\infty}^{\infty} f_T(t)\,dt = \int_b^\infty bt^{-2}\,dt = -\,\frac{b}{t}\Big|_b^\infty = \frac{b}{b} = 1$$

Luego, f_T es una fdp.

b) La fda se obtiene integrando la f_T (con límite superior de integración variable). En este caso se ve claramente que $F_T(t) = 0$ si $t \le b$ Para $t > b$ se tiene:

$$F_T(t) = \int_{-\infty}^{t} f_T(s)\,ds = \int_b^t \frac{b}{s^2}\,ds = -b\,\frac{1}{s}\Big|_b^t = b\left(\frac{1}{b} - \frac{1}{t}\right) = 1 - \frac{b}{t}$$

Por lo tanto: $F_T(t) = \begin{cases} 1 - bt^{-1} & \text{si} \quad t \ge b \\ 0 & \text{en otros casos} \end{cases}$

c) $P(T > b + c) = 1 - P(T \le b + c) = 1 - F_T(b + c) = 1 - \left(1 - \frac{b}{b+c}\right) = \frac{b}{b+c}$

d)
- Primer cuartil:

$$f_T(t_{0.25}) = 1/4 \Leftrightarrow 1 - bt_{0.25}^{-1} = 1/4 \Leftrightarrow t_{0.25} = \frac{4}{3}\,b$$

- Segundo cuartil (es decir, la mediana):

$$f_T(t_{0.5}) = 1/2 \Leftrightarrow 1 - bt_{0.5}^{-1} = 1/2 \Leftrightarrow t_{0.5} = 2b$$

- Tercer cuartil:

$$f_T(t_{0.75}) = 3/4 \Leftrightarrow 1 - bt_{0.75}^{-1} = 3/4 \Leftrightarrow t_{0.75} = 4b$$

e)

$$E(T) = \int_b^\infty bt^{-1}\,dt = b\,\ln t|_b^\infty = \infty$$

Luego: $E(T)$ no existe y por ende $V(T)$ tampoco.

4. Consideremos la recta que pasa por los puntos $\boldsymbol{A}$ y $\boldsymbol{B}$ Fijemos como origen de coordenadas en dicha recta al punto $\boldsymbol{A}$ y como sentido positivo el del vector dirigido de $\boldsymbol{A}$ hacia $\boldsymbol{B}$. Sea $\boldsymbol{X}$ la coordenada del punto donde cae el paracaidista. Como cae aleatoriamente podemos decir que $X \sim \mathcal{U}(0,d)$ siendo d la distancia entre A y B. Entonces: $f_X(x) = d^{-1} I_{(0,d)}(x)$ Por lo tanto:

   a) $P(X < d/2) = \int_0^{d/2} d^{-1}\,dx = \frac{d}{2d} = 1/2$

   b) $P(X > 3(d-X)) = P(4X > 3d) = P(X > 3d/4) = \int_{3d/4}^{d} d^{-1}\,dx = \frac{1}{d}\left(d - \frac{3d}{4}\right) = 1/4$

5. Como $T \sim \mathcal{E}(0.001)$ se tiene $F_T(t) = 1 - e^{-\,0.001t} I_{(0,\infty)}(t)$

   a) $P(T > 500) = 1 - F_T(500) = 1 - \left(1 - e^{-\,0.001\cdot 500}\right) = e^{-\,0.5} \approx 0.607$

   b) Para $T \sim \mathcal{E}(\lambda)$ se tiene $E(T) = 1/\lambda\,,\ V(T) = 1/\lambda^2$ Por lo tanto $SD(T) = 1/\lambda$ Entonces se pide calcular:

   $$P\left(\left|T - \tfrac{1}{0.001}\right| < \tfrac{1}{0.001}\right) = P\left(|T - 1000| < 1000\right) = P(-1000 < T - 1000 < 1000) =$$

   $$= P(0 < T < 2000) = F_T(2000) - F_T(0) = 1 - e^{-\,0.001\cdot 2000} = 1 - e^{-\,2} \approx 0.865$$

   c) Queremos hallar c tal que $P(T > c) \geq 0.95 \geq$ Es decir: $1 - F_T(c) \geq 0.95$ Se deduce que $F_T(c) = 1 - e^{-\,0.001c} \leq 0.05$ Por lo tanto: $0.001c \leq -\ln(0.95)$ de modo que $c \leq -1000\ \ln(0.95) \approx 51.293$ La compañía debe garantizar aproximadamente **52** horas de funcionamiento.

   d)
   $$P(T > 1500|T > 700) = \frac{P(T>1500, T>700)}{P(T>700)} = \frac{P(T>1500)}{P(T>700)} = \frac{1-F_T(1500)}{1-F_T(700)} =$$

   $$= \frac{e^{-\,0.001\cdot 1500}}{e^{-\,0.001\cdot 700}} = e^{-\,0.8} \approx 0.449$$

6. $Z \sim \mathcal{N}(0,1)$

   a)
   - $P(0 < Z < 1.2) = \Phi(1.2) - \Phi(0) \approx 0.8849 - 0.5000 = 0.3849$
   - $P(-0.9 < Z \leq 0) = \Phi(0) - \Phi(-0.9) = 0.5 - (1 - \Phi(0.9)) \approx 0.8159 - 0.5000 = 0.3159$
   - $P(0.3 \leq Z < 1.56) = \Phi(1.56) - \Phi(0.3) \approx 0.9406 - 0.6179 = 0.3227$
   - $P\left(|Z| < 0.2\right) = P(-0.2 < Z < 0.2) = 2\Phi(0.2) - 1 = 2\cdot 0.5793 - 1 = 0.1586$
   - $P\left(|Z| > 1.53\right) = 1 - P\left(|Z| < 1.53\right) = 1 - (2\Phi(1.53) - 1) = 2\left(1 - \Phi(1.53)\right) \approx 2(1 - 0.9370) = 0.126$

   b)
   - $0.5 = P(Z > z) = 1 - \Phi(z)$ Luego: $\Phi(z) = 0.5$ de modo que $z = z_{0.5} = 0$
   - $0.8643 = P(Z < z) = \Phi(z)$ Luego: $z = z_{0.8643} \approx 1.1$
   - $0.90 = P\left(|Z| < z\right) = 2\Phi(z) - 1$ Luego: $\Phi(z) = 0.95$ Entonces: $z = z_{0.95} \approx 1.65$
   - $0.85 = P\left(|Z| > z\right) = 2\left(1 - \Phi(z)\right)$ Luego: $\Phi(z) = 0.575$ Entonces: $z = z_{0.575} \approx 0.19$
   - $z = z_{0.0054} = -\,z_{1-0.0054} = -\,z_{0.9946} \approx -\,2.55$

7. $X \sim \mathcal{N}(10, 36)$

   a) $P(X > 5) = P\left(\frac{X-10}{6} > \frac{5-10}{6}\right) = 1 - \Phi(-\,5/6) = \Phi(5/6) \approx 0.7967$

   $P(4 < X < 16) = P\left(\frac{4-10}{6} < \frac{X-10}{6} < \frac{16-10}{6}\right) = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 \approx 0.6826$

   $P(X \leq 8) = P\left(\frac{X-10}{6} \leq \frac{8-10}{6}\right) = \Phi(-\,1/3) = 1 - \Phi(1/3) \approx 0.3707$

b)

$$P(X \leq c) = 2P(X > c) \quad \Leftrightarrow \quad P\left(\frac{X-10}{6} \leq \frac{c-10}{6}\right) = 2P\left(\frac{X-10}{6} > \frac{c-10}{6}\right) \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \Phi\left(\frac{c-10}{6}\right) = 2 - 2\Phi\left(\frac{c-10}{6}\right) \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \Phi\left(\frac{c-10}{6}\right) = 2/3 \quad \Leftrightarrow \quad \frac{c-10}{6} = z_{2/3}$$

Entonces se tiene: $(c-10)/6 \approx 0.44$ Por lo tanto: $c \approx 12.64$

8. Sea X el diámetro del perno en mm. Entonces $X \sim \mathcal{N}(950, 100)$

   a) $0.8531 = P(X < c) = P\left(\frac{X-950}{10} < \frac{c-950}{10}\right)$ Luego: $\Phi\left(\frac{c-950}{10}\right) = 0.8531$ Por lo tanto: $\frac{c-950}{10} = z_{0.8531} \approx 1.05$ de modo que $c \approx 960.5$
   b) Supongamos ahora que $X \sim \mathcal{N}(950, \sigma^2)$ Se tiene: $0.9 = P(X < 960.5) = P\left(\frac{X-950}{\sigma} < \frac{960.5-950}{\sigma}\right) = \Phi\left(\frac{10.5}{\sigma}\right)$ Entonces: $\frac{10.5}{\sigma} = z_{0.9} \approx 1.29$ de modo que $\sigma \approx \frac{10.5}{1.29} = 8.1395$ La reducción en el desvío standard debería ser aproximadamente del 18.6%

9. Sean X_A, X_B las resistencias a la ruptura de sendos tipos de acero. Se supone $X_A \sim \mathcal{N}(43, (4.4)^2)$ y $X_B \sim \mathcal{N}(44, (6.1)^2)$ Se tiene:

   $$P(X_A < 40) = P\left(\frac{X_A-43}{4.4} < \frac{40-43}{4.4}\right) = \Phi(-3/4.4) = 1 - \Phi(3/4.4) \approx 1 - 0.47 = 0.53$$
   $$P(X_B < 40) = P\left(\frac{X_B-44}{6.1} < \frac{40-44}{6.1}\right) = \Phi(-4/6.1) = 1 - \Phi(4/6.1) \approx 1 - 0.4 = 0.6$$

   La probabilidad de que la resistencia sea a lo sumo de 40 es menor para el acero tipo A, a pesar de que el acero tipo A posee una resistencia esperada menor a la del acero B. El resultado obedece al hecho que la varianza de la resistencia es adecuadamente menor en el acero A que en acero B.

10. Sea X el diámetro de los tornillos expresado en pulgadas. Se supone $X \sim \mathcal{N}(1.196, (0.005)^2)$

    a) $1 - P(1.19 \leq X \leq 1.21) = 1 - P\left(\frac{1.19-1.196}{0.005} \leq \frac{X-1.196}{0.005} \leq \frac{1.21-1.196}{0.005}\right) = 1 - [\Phi(2.8) - \Phi(-1.2)] = 2 - \Phi(2.8) - \Phi(1.2) \approx 0.1177$ El porcentaje pedido es entonces aproximadamente: 11.77%
    b) Suponemos ahora que $X \sim \mathcal{N}(\mu, (0.005)^2)$ Queremos hallar μ de modo que $1 - P(1.19 \leq X \leq 1.21)$ sea mínima. Pero: $1 - P(1.19 \leq X \leq 1.21) = 1 - P\left(\frac{1.19-\mu}{0.005} \leq \frac{X-\mu}{0.005} \leq \frac{1.21-\mu}{0.005}\right) = 1 - \left[\Phi\left(\frac{1.21-\mu}{0.005}\right) - \Phi\left(\frac{1.19-\mu}{0.005}\right)\right]$
    Debemos minimizar respecto de μ para lo cual derivamos e igualamos a cero (puntos críticos):
    $\frac{\phi\left(\frac{1.21-\mu}{0.005}\right)}{0.005} - \frac{\phi\left(\frac{1.19-\mu}{0.005}\right)}{0.005} = 0$ Esto se reduce a: $\phi(200(1.21-\mu)) = \phi(200(1.19-\mu))$ Es decir:
    $[200(1.21-\mu)]^2 = [200(1.19-\mu)]^2$ Por lo tanto: $1.21 - \mu = \mu - 1.19$ O sea: $\mu = 1.2$

11. $X \sim \mathcal{U}(0,1)$ Sea $Y = aX + b$ donde a, b son constantes y $a \neq 0$ Obtebgamos la fda de Y sin apelar al teorema de cambio de variables:

    - Caso $a > 0$

    $$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P(X \leq \frac{y-b}{a}) = F_X\left(\frac{y-b}{a}\right)$$

Observemos que $0 < \frac{y-b}{a} < 1 \quad \Leftrightarrow \quad b < y < b+a$ Luego, para $y \in (b, b+a)$ se obtiene por derivación:

$$f_Y(y) = \frac{1}{a}\, f_X\left(\frac{y-b}{a}\right) = 1/a$$

Es decir: $f_Y(y) = \frac{1}{a}\, I_{(b,b+a)}(y)$ O sea: $Y \sim \mathcal{U}(b, b+a)$

- Caso $a < 0$

$$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P(X \geq \frac{y-b}{a}) = 1 - F_X\left(\frac{y-b}{a}\right)$$

Observemos que $0 < \frac{y-b}{a} < 1 \quad \Leftrightarrow \quad b+a < y < b$ Luego, para $y \in (b+a, b)$ se obtiene por derivación:

$$f_Y(y) = -\,\frac{1}{a}\, f_X\left(\frac{y-b}{a}\right) = -\,1/a$$

Es decir: $f_Y(y) = -\,\frac{1}{a}\, I_{(b+a,b)}(y)$ O sea: $Y \sim \mathcal{U}(b+a, b)$

12. $U \sim \mathcal{U}(0,1)$, $X = -\lambda^{-1}\, \ln(1-U)$ Calculemos la fda de X
En este caso $X = g(U)$ donde $g(u) = -\lambda^{-1}\, \ln(1-u)$ Sean $S_U = (0,1)$ el soporte de f_U , $S_X = g\,(S_U) = (0, \infty)$ La función $(0,1) \xrightarrow{g} (0,\infty)$ es continua, estrictamente creciente y su inversa es diferenciable con continuidad: $(0,\infty) \xrightarrow{g^{-1}} (0,1)$

$$g^{-1}(x) = 1 - e^{-\,\lambda x} \qquad ; \qquad \tfrac{d}{dx}\left(g^{-1}(x)\right) = \lambda\, e^{-\,\lambda x}$$

Entonces por el teorema de cambio de variables resulta X variable aleatoria continua con fdp dada por:

$$f_X(x) = f_U(g^{-1}(x)) \left|\frac{d}{dx}\;\left(g^{-1}(x)\right)\right| I_{(0,\infty)}(x) = \lambda\, e^{-\,\lambda x} I_{(0,\infty)}(x)$$

Pero esta es la densidad de una $\mathcal{E}(\lambda)$ Luego: $X \sim \mathcal{E}(\lambda)$

13. a) a1) $W = e^{-\,Y}$ Notemos que $S_Y = (3,\infty)$ Pero: $y \geq 3 \quad \Leftrightarrow \quad -\,y \leq -3 \quad \Leftrightarrow \quad 0 < e^{-y} \leq e^{-3}$ Por lo tanto: $S_W = (0, e^{-3})$ Para $0 < w \leq e^{-3}$ se tiene:

$$\begin{aligned} F_W(w) &= P(W \leq w) = P(e^{-\,Y} \leq w) = P(-\,Y \leq \ln w) = \\ &= P(Y \geq -\,\ln w) = 1 - F_Y(-\,\ln w) \end{aligned}$$

Derivando se obtiene: $f_W(w) = \frac{1}{w}\, f_Y(-\,\ln w) = e^3\, I_{(0,e^{-3})}$ de manera que $W \sim \mathcal{U}(0, e^{-3})$

a2) $W = g(Y) = Y - 3$ En este caso $S_W = g\,((3,\infty)) = (0,\infty)$ Calculamos la fda de W para $w > 0$

$$F_W(w) = P(W \leq w) = P(Y - 3 \leq w) = P(Y \leq w+3) = F_Y(w+3)$$

Derivando se tiene: $f_W(w) = f_Y(w-3)$

$$f_W(w) = \begin{cases} e^{-\,((w+3)-3)} & \text{si} \quad w > 0 \\ 0 & \text{si} \quad w \leq 0 \end{cases} = \begin{cases} e^{-\,w} & \text{si} \quad w > 0 \\ 0 & \text{si} \quad w \leq 0 \end{cases}$$

Deducimos que $W \sim \mathcal{E}(1)$

a3) $W = g(y) = Y/3$ En este caso $g(y) = Y/3$ de modo que $S_W = g\left((3,\infty)\right) = (1,\infty)$ Para $w > 1$ se tiene:

$$F_W(w) = P(W \leq w) = P(Y/3 \leq w) = P(Y \leq 3w) = F_Y(3w)$$

Derivando se obtiene: $f_W(w) = 3\, f_Y(3w)$

$$f_W(w) = \begin{cases} 3\, e^{-3\,(w-1)} & \text{si} \quad w > 1 \\ 0 & \text{si} \quad w \leq 1 \end{cases}$$

b) Calculemos primero esperanzas por definición:

a1) $W \sim \mathcal{U}(0, e^{-3})$

$$E(W) = \int_0^{e^{-3}} e^3\, w\, dw = e^3 \int_0^{e^{-3}} w\, dw = \frac{e^3}{2}\ w^2\Big|_0^{e^{-3}} = \frac{1}{2}\, e^{-3}$$

a2) $W \sim \mathcal{E}(1)$

$$E(W) = \int_0^{\infty} w\, e^{-w}\, dw = -\ we^{-w}\Big|_0^{\infty} + \int_0^{\infty} e^{-w}\, dw = -\ e^{-w}\Big|_0^{\infty} = 1$$

a3)

$$E(W) = \int_1^{\infty} 3w\, e^{-3(w-1)}\, dw = \int_0^{\infty} 3(t+1)\, e^{-3t}\, dt =$$

$$= \int_0^{\infty} 3t\, e^{-3t}\, dt + \int_0^{\infty} 3\, e^{-3t}\, dt = -\ t\, e^{-3t}\Big|_0^{\infty} + \int_0^{\infty} e^{-3t}\, dt + \int_0^{\infty} 3\, e^{-3t}\, dt =$$

$$= -\ t\, e^{-3t}\Big|_0^{\infty} + \int_0^{\infty} e^{-3t}\, dt + \int_0^{\infty} 3\, e^{-3t}\, dt =$$

$$= -\tfrac{1}{3}\ e^{-3t}\Big|_0^{\infty} - e^{-3t}\Big|_0^{\infty} = 1/3 + 1 = 4/3$$

Ahora calculemos esperanzas sin recurrir a la fdp de W

a1) $W = g(Y)$ con $g(y) = e^{-y}$

$$E(W) = \int_3^{\infty} e^{-y}\, e^{-(y-3)}\, dy = e^3 \int_3^{\infty} e^{-2y}\, dy = -\tfrac{1}{2}\, e^3\ e^{-2y}\Big|_3^{\infty} = \tfrac{1}{2}\, e^3\, e^{-6} = \tfrac{1}{2}\, e^{-3}$$

a2) $W = g(Y)$ con $g(y) = y - 3$

$$E(W) = \int_3^{\infty} (y-3)\, e^{-(y-3)}\, dy = -\ (y-3)\, e^{-(y-3)}\Big|_3^{\infty} + \int_3^{\infty} e^{-(y-3)}\, dy =$$

$$= -\ e^{-(y-e)}\Big|_3^{\infty} = 1$$

Una tercera posibilidad para calcular $E(W)$ calculando $E(Y)$ y utilizando linealidad.

a3) $W = g(Y)$ con $g(y) = y/3$

$$E(W) = \int_3^{\infty} \tfrac{y}{3}\, e^{-(y-3)}\, dy = \tfrac{1}{3}\, e^3 \left\{-\ y\, e^{-y}\Big|_3^{\infty} + \int_3^{\infty} e^{-y}\, dy\right\} =$$

$$= \tfrac{1}{3}\, e^3 \left\{3\, e^{-3} - e^{-y}\Big|_3^{\infty}\right\} = \tfrac{1}{3}\, e^3 \left(3\, e^{-3} + e^{-3}\right) = 4/3$$

Una tercera posibilidad para calcular $E(W)$ calculando $E(Y)$ y utilizando linealidad.

14. $T \sim \mathcal{E}(\lambda)$ siendo $\lambda = 1/10$ Sabemos que para la exponencial se cumple $E(T) = \frac{1}{\lambda}$ y $V(T) = \frac{1}{\lambda^2}$ de modo que en este caso $E(T) = 10$, $V(T) = 100$ Por lo tanto:

$$E(C) = E(100 + 40T + 3T^2) = 100 + 40E(T) + 3E(T^2) =$$

$$= 100 + 40E(T) + 3\left(V(T) + E^2(T)\right) = 100 + 40 \cdot 10 + 3(100 + 10^2) = 1100$$

15. a) • Ejemplo discreto: Sea $X \sim \mathcal{B}i(n, 1/2)$ y sea $\theta = n/2$ Probemos que X posee distribución simétrica respecto de θ En primer lugar observemos que:

$$\sum_{k \geq \theta+h} \binom{n}{k} = \sum_{k \geq \theta+h} \binom{n}{n-k} = \sum_{j \leq \theta-h} \binom{n}{j}$$

donde hemos realizado un cambio de índice en la sumatoria: $j = n - k$ con lo cual se tienen las siguientes equivalencias que justifican la última igualdad entre sumatorias:

$$k \geq \theta + h \Leftrightarrow -k \leq -\theta - h \Leftrightarrow n - k \leq n - \theta - h \Leftrightarrow j \leq n - \theta - h \Leftrightarrow$$

$$\Leftrightarrow j \leq n - n/2 - h \Leftrightarrow j \leq n/2 - h \Leftrightarrow j \leq \theta - h$$

Por lo tanto:

$$P(X \geq \theta + h) = (1/2)^n \sum_{k \geq \theta+h} \binom{n}{k} = (1/2)^n \sum_{j \leq \theta-h} \binom{n}{j} = P(X \leq \theta - h)$$

• Ejemplo continuo: Sea $X \sim \mathcal{N}(\mu, \sigma^2)$ y sea $\theta = \mu$ Probemos que la distribución de X es simétrica respecto de μ

$$P(X \leq \mu - h) = P\left(\tfrac{X-\mu}{\sigma} \leq -\tfrac{h}{\sigma}\right) = \Phi(-h/\sigma) = 1 - \Phi(h/\sigma) =$$

$$= P\left(\tfrac{X-\mu}{\sigma} \geq \tfrac{h}{\sigma}\right) = P(X \geq \mu + h)$$

b) Comentario al margen: La condición de la definición de simetría respecto de θ no necesita suponer $h > 0$ En efecto: Si vale la condición de simetría entonces para $h < 0$ se tiene también

$$\begin{aligned} P(X \leq \theta - h) &= P(X \leq \theta + |h|) = 1 - P(X \geq \theta + |h|) = 1 - P(X \leq \theta - |h|) = \\ &= P(X \geq \theta - |h|) = P(X \geq \theta + h) \end{aligned}$$

La condición también es válida para $h = 0$ es pues:

$$\begin{aligned} P(X \leq \theta) &= P(X < \theta) = \lim_{h \to 0^+} P(X \leq \theta - h) = \lim_{h \to 0^+} P(X \geq \theta + h) = \\ &= P(X > \theta) = P(X \geq \theta) \end{aligned}$$

En resumen, la variable aleatoria continua X posee distribución simétrica respecto de θ sii se verifica:

$$\forall h \in \mathbb{R}\,,\ P(X \leq \theta - h) = P(X \geq \theta + h)$$

Ahora sí demostremos las equivalencias.

i)⇒ii) Supongamos simetría respecto de θ Tomando en esa definición $h = \theta - x$ se obtiene:

$$P(X \leq x) = P(X \leq \theta - h) = P(X \geq \theta + h) = P(X \geq 2\theta - x)$$

ii)⇒iii) Directamente utilizando la definición de fda.

iii)⇒iv) Directamente derivando respecto de x en ambos miembros.

iv)⇒v) Para $x > 0$ resulta $\theta - x < \theta$ Por lo tanto se tiene:

$$f_X(\theta - x) = f_X(2\theta - (\theta - x)) = f_X(\theta + x)$$

v)⇒i) Supongamos que vale la condición v). Fijemos $h > 0$ Se tiene:

$$P(X \leq \theta - h) = \int_{-\infty}^{\theta-h} f_X(t)\,dt = -\int_{\infty}^{h} f_X(\theta - x)\,dx = \int_{h}^{\infty} f_X(\theta - x)\,dx =$$

$$= \int_{h}^{\infty} f_X(\theta + x)\,dx = \int_{\theta+h}^{\infty} f_X(s)\,ds = P(X \geq \theta + h)$$

EJERCICIOS ADICIONALES

16. a) Se tiene:

$$1 = \int_0^1 c(1-x)\,dx = c\ \left(x - x^2/2\right)\Big|_0^1 = c/2$$

Por lo tanto: $c = 2$
Para abreviar posibles cálculos posteriores vamos a determinar la fda de X
Para $0 < x < 1$ se tiene:

$$F_X(x) = \int_0^x 2(1-t)\,dt = \left(2t - t^2\right)\Big|_0^x = 2x - x^2$$

de manera que:

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ 2x - x^2 & \text{si} \quad 0 \le x \le 1 \\ 1 & \text{si} \quad x > 1 \end{cases}$$

b) Observemos que C es variable aleatoria discreta con $R_C = \{1, 2, 3\}$ Corresponde entonces hallar su fmp. Se tiene:

$$\begin{aligned} p_C(1) &= P(C=1) = P(X < 1/3) = F_X(1/3) = 2/3 - (1/3)^2 = 5/9 \\ p_C(2) &= P(C=2) = P(1/3 \le X \le 2/3) = F_X(2/3) - F_X(1/3) = \\ &= \left[4/3 - (2/3)^2\right] - \left[2/3 - (1/3)^2\right] = 1/3 \\ p_C(3) &= P(C=3) = P(X > 2/3) = 1 - F_X(2/3) = 1 - \left[4/3 - (2/3)^2\right] = 1/9 \end{aligned}$$

Por lo tanto la fmp de C queda resumida en la siguiente tabla:

k	1	2	3
$p_C(k)$	$5/9$	$1/3$	$1/9$

17. a)

$$1 = \int_0^{10} k\,x\,dx = \frac{k}{2}\ x^2\Big|_0^{10} = 50\,k$$

Luego: $k = 1/50$ Por comodidad para cálculos posteriores, hallemos la fda de D Para $0 < x < 10$ se tiene:

$$F_D(x) = \int_0^x \frac{1}{50}\,t\,dt = \frac{1}{100}\ t^2\Big|_0^x = \frac{x^2}{100}$$

Luego:

$$F_D(x) = \begin{cases} 0 & \text{si} \quad x \le 0 \\ x^2/100 & \text{si} \quad 0 < x < 10 \\ 1 & \text{si} \quad x \ge 10 \end{cases}$$

b)

$$P(4 \le D \le 6) = F_D(6) - F_D(4) = 36/100 - 16/100 = 1/5 = 0.2$$

c)

$$\begin{aligned} P(4 \le D \le 6 | D > 5) &= \frac{P(4 \le D \le 6, D > 5)}{P(D > 5)} = \frac{P(5 < D \le 6)}{P(D > 5)} = \\ &= \frac{F_D(6) - F_D(5)}{1 - F_D(5)} = \frac{\frac{36}{100} - \frac{25}{100}}{1 - \frac{25}{100}} = \frac{11}{75} \approx 0.147 \end{aligned}$$

d) Sea $\boldsymbol{Y}$ la cantidad de árboles (en la muestra de 3) que tienen diámetro entre $\mathbf{4}$ y $\mathbf{6}$ dm. Entonces $\boldsymbol{Y \sim \mathcal{B}i(3,p)}$ siendo $\boldsymbol{p = 0.2}$ Por lo tanto:

$$P(Y=2) = \binom{3}{2}(0.2)^2 \cdot (0.8) \approx 0.096$$

e) Sea ahora $\boldsymbol{Y}$ la cantidad de árboles en una muestra de tamaño $\boldsymbol{n}$ que tienen diámetro entre $\mathbf{4}$ y $\mathbf{6}$ dm. Entonces $\boldsymbol{Y \sim \mathcal{B}i(n, 0.2)}$ El objetivo es hallar $\boldsymbol{n}$ de modo que se cumpla $\boldsymbol{P(Y \geq 1) \geq 0.99}$ Pero:

$$P(Y \geq 1) \geq 0.99 \Leftrightarrow 1 - P(Y=0) \geq 0.99 \Leftrightarrow P(Y=0) \leq 0.01 \Leftrightarrow$$

$$\Leftrightarrow (0.8)^n \leq 0.01 \Leftrightarrow n \ln(0.8) \leq \ln(0.01) \Leftrightarrow n \geq \tfrac{\ln(0.01)}{\ln(0.8)} \Leftrightarrow n \geq 20.64$$

Por lo tanto, habría que tomar una muestra de al menos $\boldsymbol{n = 21}$ árboles.

18. $\boldsymbol{I \sim \mathcal{N}(\mu, \sigma^2)}$ Se tiene:

$$\begin{aligned} 0.58 &= P(I < 75) = P\left(\tfrac{I-\mu}{\sigma} < \tfrac{75-\mu}{\sigma}\right) = \Phi\left(\tfrac{75-\mu}{\sigma}\right) \\ 0.04 &= P(I > 80) = P\left(\tfrac{I-\mu}{\sigma} > \tfrac{80-\mu}{\sigma}\right) = 1 - \Phi\left(\tfrac{80-\mu}{\sigma}\right) \end{aligned}$$

Por lo tanto:

$$\begin{cases} \tfrac{75-\mu}{\sigma} = z_{0.58} \\ \tfrac{80-\mu}{\sigma} = z_{0.96} \end{cases}$$

Resolviendo este sistema en las incógnitas $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$ se obtiene:

$$\mu = 80 - \tfrac{5\, z_{0.96}}{z_{0.96} - z_{0.58}} \approx 74.3548 \quad ; \quad \sigma = \tfrac{5}{z_{0.96} - z_{0.58}} \approx 3.2258$$

Luego:

$$P(78 \leq I \leq 82) \approx P\left(\tfrac{78-74.3548}{3.2258} \leq \tfrac{I-74.3548}{3.2258} \leq \tfrac{82-74.3548}{3.2258}\right) =$$

$$= \Phi\left(\tfrac{82-74.3548}{3.2258}\right) - \Phi\left(\tfrac{78-74.3548}{3.2258}\right) \approx \Phi(2.37) - \Phi(1.13) \approx 0.1203$$

19. $\boldsymbol{Y = \ln X}$ Calculemos su fda:

$$F_Y(y) = P(Y \leq y) = P(\ln X \leq y) = P(X \leq e^y) = F_X(e^y)$$

Derivando se obtiene la fdp:

$$f_Y(y) = f_X(e^y)\, e^y = \frac{e^y}{\sqrt{2\pi}\,\sigma\, e^y}\, e^{-1/2\,[\ln(e^y)-\mu]^2/\sigma^2} = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-(x-\mu)/2\sigma^2}$$

Vemos que la fdp corresponde a la de una distribución $\boldsymbol{\mathcal{N}(\mu, \sigma^2)}$

20. $\boldsymbol{Y = g(X)}$ siendo $\boldsymbol{g(x) = (x/\alpha)^\beta}$ Se tiene: $\boldsymbol{S_Y = g(S_X) = g((0,\infty)) = (0,\infty)}$ Calculemos la fda de $\boldsymbol{Y}$

$$F_Y(y) = P(Y \leq y) = P((X/\alpha)^\beta \leq y) = P(X \leq \alpha\, y^{1/\beta}) = F_X(\alpha\, y^{1/\beta})$$

Derivando se obtiene:

$$f_Y(y) = f_X(\alpha\, y^{1/\beta})\, \frac{\alpha}{\beta}\, y^{-1+1/\beta} = \frac{\beta}{\alpha}\, y^{(\beta-1)/\beta}\, e^{-y}\, \frac{\alpha}{\beta}\, y^{-1+1/\beta} = e^{-y}$$

Pero esta es precisamente la fdp de una $\boldsymbol{\mathcal{E}(1)}$

21. $\boldsymbol{X \sim \mathcal{U}(0,1)}$

a) $\boldsymbol{Y = g(X)}$ donde $\boldsymbol{g(x) = x^{\alpha} = e^{\alpha \ln x}}$

- Caso $\boldsymbol{\alpha > 0}$ En este caso $\boldsymbol{S_Y = g\left((0,1)\right) = (0,1)}$ Entonces para $\boldsymbol{0 < y < 1}$ es:

$$\boldsymbol{F_Y(y) = P(Y \le y) = P(X^{\alpha} \le y) = P(X \le y^{1/\alpha}) = F_X(y^{1/\alpha}) = y^{1/\alpha}}$$

Entonces:

$$\boldsymbol{F_Y(y) = \begin{cases} 0 & \text{si} \quad y \le 0 \\ y^{1/\alpha} & \text{si} \quad 0 < y < 1 \\ 1 & \text{si} \quad y \ge 1 \end{cases}}$$

Derivando se obtiene la fdp:

$$\boldsymbol{f_Y(y) = \frac{1}{\alpha}\, y^{-1+1/\alpha}\, I_{(0,1)}(y)}$$

- Caso $\boldsymbol{\alpha < 0}$ En este caso $\boldsymbol{S_Y = g\left((0,1)\right) = (1,\infty)}$ Entonces para $\boldsymbol{y > 1}$ es:

$$\boldsymbol{F_Y(y) = P(Y \le y) = P(X^{\alpha} \le y) = P(X \ge y^{1/\alpha}) = 1 - F_X(y^{1/\alpha}) = 1 - y^{1/\alpha}}$$

Entonces:

$$\boldsymbol{F_Y(y) = \begin{cases} 0 & \text{si} \quad y \le 1 \\ 1 - y^{1/\alpha} & \text{si} \quad y > 1 \end{cases}}$$

Derivando se obtiene la fdp:

$$\boldsymbol{f_Y(y) = -\,\frac{1}{\alpha}\, y^{-1+1/\alpha}\, I_{(1,\infty)}(y)}$$

b) $\boldsymbol{Y = g(X)}$ donde $\boldsymbol{g(x) = \ln x}$ Entonces: $\boldsymbol{S_Y = g\left((0,1)\right) = (-\infty, 0)}$ Para $\boldsymbol{y < 0}$ se tiene:

$$\boldsymbol{F_Y(y) = P(Y \le y) = P(\ln X \le y) = P(X \le e^{y}) = F_X(e^{y}) = e^{y}}$$

Luego:

$$\boldsymbol{F_Y(y) = \begin{cases} e^{y} & \text{si} \quad y < 0 \\ 1 & \text{si} \quad y \ge 0 \end{cases}}$$

Derivando obtenemos la fdp:

$$\boldsymbol{f_Y(y) = e^{y}\, I_{(-\infty,0)}(y)}$$

c) $\boldsymbol{Y = g(X)}$ donde $\boldsymbol{g(x) = \frac{x}{x+1}}$ Entonces: $\boldsymbol{S_Y = g\left((0,1)\right) = (0, 1/2)}$ Luego, para $\boldsymbol{0 < y < 1/2}$ se tiene:

$$\boldsymbol{F_Y(y) = P(Y \le y) = P\left(\frac{X}{X+1} \le y\right) = P\left(X \le \frac{y}{1-y}\right) = F_X\left(\frac{y}{1-y}\right) = \frac{y}{1-y}}$$

Por lo tanto:

$$\boldsymbol{F_Y(y) = \begin{cases} 0 & \text{si} \quad y \le 0 \\ y/(1-y) & \text{si} \quad 0 < y < 1/2 \\ 1 & \text{si} \quad y \ge 1/2 \end{cases}}$$

Derivando se obtiene la fdp:

$$\boldsymbol{f_Y(y) = \frac{1}{(1-y)^2}\, I_{(0,1/2)}(y)}$$

22. $\boldsymbol{U \sim \mathcal{U}(0,1)\,,\ Y = g(U)}$

a) Hallemos la fda de $\boldsymbol{Y \sim \mathcal{E}(1)}$

$$\boldsymbol{F_Y(y) = 1 - e^{-\,y}\, I_{(0,\infty)}(y)}$$

Entonces $\boldsymbol{(0,\infty) \xrightarrow{F_Y} (0,1)}$ es continua y estrictamente creciente. Por lo tanto $\boldsymbol{Y = F_Y^{-1}(U) \sim F_Y}$ Pero $\boldsymbol{F_Y^{-1}(u) = -\ \ln(1-u)}$ de modo que $\boldsymbol{Y = -\ \ln(1-U) \sim \mathcal{E}(1)}$

b) Hallemos la fda de $\boldsymbol{Y \sim \mathcal{DE}(1)}$ En la integral de la fdp hay que separar los casos positivo y negativo. Para resumir cuenterío les doy directamente el resultado:

$$\boldsymbol{F_Y(y) = \begin{cases} 1/2\, e^{y} & \text{si} \quad y \le 0 \\ 1 - 1/2\, e^{-\,y} & \text{si} \quad y > 0 \end{cases}}$$

Esta $\boldsymbol{\mathbb{R} \xrightarrow{F_Y} (0,1)}$ es continua y estrictamente creciente. Para calcular su inversa consideramos dos casos:

- Caso $\boldsymbol{1/2 \le u < 1}$ Entonces: $\boldsymbol{F_Y^{-1}(u) = \ln(2u)}$
- Caso $\boldsymbol{0 < u < 1/2}$ Entonces: $\boldsymbol{F_Y^{-1}(u) = -\ \ln(2-2u)}$

Por lo tanto:

$$\boldsymbol{Y = \begin{cases} -\ \ln(2-2U) & \text{si} \quad 0 < U < 1/2 \\ \ln(2U) & \text{si} \quad 1/2 \le U < 1 \end{cases}}$$

c) Hallemos la fda de $\boldsymbol{Y}$

$$\boldsymbol{F_Y(y) = \int_{-\,\infty}^{y} \frac{1}{\pi}\,\frac{dt}{1+t^2} = \frac{1}{\pi}\ \text{arctg}\ t|_{-\,\infty}^{y} = \frac{1}{2} + \frac{1}{\pi}\,\text{arctg}\ y}$$

Entonces $\boldsymbol{\mathbb{R} \xrightarrow{F_Y} (0,1)}$ es continua y estrictamente creciente. Su inversa viene dada por: $\boldsymbol{F_Y^{-1}(u) = \text{tg}\ [\pi(u-1/2)]}$ de manera que

$$\boldsymbol{Y = F_Y^{-1}(U) = \text{tg}\ \left[\pi\left(U - \frac{1}{2}\right)\right]}$$

d) ME CANSE (Además no es elemental!!!)

Trabajo práctico 5 (Anexo): Momentos de una distribución - Función generadora de momentos - Cuantiles de distribuciones continuas

1. En cada uno de los siguientes casos verifique la expresión dada para la fgm. Utilícela luego para hallar $E(X)$ y $V(X)$

   a) $X \sim \mathcal{P}(\lambda) \quad \Rightarrow \quad M_X(t) = e^{\lambda(e^t-1)}$

   b) $X \sim \mathcal{E}(\lambda) \quad \Rightarrow \quad M_X(t) = \left(1 - \frac{t}{\lambda}\right)^{-1}$ si $t < \lambda$

   c) $X \sim \mathcal{N}(\mu, \sigma^2) \quad \Rightarrow \quad M_X(t) = e^{\mu\, t + \frac{1}{2}\,\sigma^2\, t^2}$

   d) $X \sim \chi^2_n \quad \Rightarrow \quad M_X(t) = (1-2t)^{-\,n/2}$ si $t < 1/2$

2. Utilizando la fgm, calcular: $E(X^4)$ siendo $X \sim \mathcal{N}(0,1)$

3. Sea $M_X(t)$ la fgm de la distribución de la v.a. X Definamos $S(t) = \ln\left(M_X(t)\right)$ Mostrar que:

$$\left.\frac{d}{dt}\right|_{t=0} S(t) = E(X) \qquad \text{y} \qquad \left.\frac{d^2}{dt^2}\right|_{t=0} S(t) = V(X)$$

4. Sean μ_n , μ^*_n los momentos y momentos centrados (respectivamente) de orden n de la distribución de una v.a. X (supuestos existentes). Considere la fórmula:

$$\mu^*_n = \left.\left(\frac{d}{dt} - \mu_1\right)^{(n)}\right|_{t=0} M_X(t)$$

   donde $\mu_1 = E(X)$ Utilice dicha fórmula para calcular μ^*_2 y μ^*_3 cuando $X \sim \mathcal{N}(\mu, \sigma^2)$

5. Determinar la mediana de las siguientes distribuciones:

   a) $f_X(x) = 3x^2\, I_{(0,1)}(x)$ b) $f_X(x) = \frac{1}{\pi}\,\frac{1}{1+x^2}$

   En el caso a) determine también el α-cuantil para $\alpha \in (0,1)$

6. Sea μ^*_n el momento central de orden n de la distribución de la v.a.X Además de la esperanza y la varianza de X pueden resultar también de interés los siguientes parámetros asociados a la distribución de X

$$\alpha_3 = \frac{\mu^*_3}{(\mu^*_2)^{3/2}} \qquad ; \qquad \alpha_4 = \frac{\mu^*_4}{(\mu^*_2)^2}$$

   Estos parámetros se denominan respectivamente: *coeficiente de asimetría* y *curtosis*

   a) Mostrar que si X posee fdp simétrica alrededor de cierto centro $x = c$ entonces $\alpha_3 = 0$

   b) Calcular el coeficiente de asimetría para la distribución $\mathcal{E}(1)$ Observe que en este caso la distribución es asimétrica con colas pesadas a derecha.

   c) Determinar la curtosis para cada una de las distribuciones siguientes:

      (1) $X \sim \mathcal{N}(0,1)$

      (2) $X \sim \mathcal{U}(-1,1)$

      (3) $X \sim \mathcal{DE}(1)$ es decir: $f_X(x) = \frac{1}{2}\, e^{-\,|x|}$ para $-\infty < x < \infty$

RESPUESTAS

1. a) $X \sim \mathcal{P}(\lambda)$

$$\begin{aligned} M_X(t) &= E(e^{tX}) = \textstyle\sum_{k=0}^{\infty} e^{kt}\,\frac{\lambda^k}{k!}\,e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda\,e^t)^k}{k!} = \\ &= e^{-\lambda}\,e^{\lambda\,e^t} = e^{\lambda(e^t-1)} \end{aligned}$$

Derivando:

$$\begin{aligned} M_X(0) &= 1 \\ M'_X(t) &= M_X(t)\lambda e^t \quad\Rightarrow\quad E(X) = M'_X(0) = \lambda \\ M''_X(t) &= \lambda\left(M_X(t)\lambda e^{2t} + M_X(t)e^t\right) = \\ &= M_X(t)\lambda e^t\left(\lambda e^t + 1\right) \quad\Rightarrow\quad M''_X(0) = \lambda(\lambda+1) \end{aligned}$$

Entonces:

$$V(X) = E(X^2) - E^2(X) = \lambda(\lambda+1) - \lambda^2 = \lambda$$

b) $X \sim \mathcal{E}(\lambda)$

$$\begin{aligned} M_X(t) &= E(e^{tX}) = \textstyle\int_0^{\infty} e^{tx}\,\lambda e^{-\lambda x}\,dx = \lambda\int_0^{\infty} e^{-(\lambda-t)x}\,dx = \\ &= -\frac{\lambda}{\lambda-t}\,e^{-(\lambda-t)x}\Big|_0^{\infty} = \frac{\lambda}{\lambda-t} = \left(1-\frac{t}{\lambda}\right)^{-1} \quad \text{si } t<\lambda \end{aligned}$$

Derivando:

$$\begin{aligned} M_X(0) &= 1 \\ M'_X(t) &= \lambda^{-1}(1+t/\lambda)^{-2} \quad\Rightarrow\quad M'_X(0) = \lambda^{-1} \\ M''_X(t) &= 2\lambda^{-2}(1-t/\lambda)^{-3} \quad\Rightarrow\quad M''_X(0) = 2\lambda^{-2} \end{aligned}$$

Entonces:

$$V(X) = E(X^2) - E^2(X) = 2\lambda^{-2} - (\lambda^{-1})^2 = \frac{1}{\lambda^2}$$

c) Hagamos primeramente el caso $Z \sim \mathcal{N}(0,1)$

$$\begin{aligned} M_Z(t) &= E(e^{tZ}) = \textstyle\int_{-\infty}^{\infty} e^{tz}\,\frac{1}{\sqrt{2\pi}}\,e^{-z^2/2}\,dz = \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\,e^{-1/2(z^2-2tz)}\,dz = \\ &= \textstyle\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\,e^{-1/2(z-t)^2+\frac{1}{2}t^2}\,dz = e^{t^2/2}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\,e^{-1/2(z-t)^2}\,dz = e^{t^2/2} \end{aligned}$$

Derivando:

$$\begin{aligned} M_Z(0) &= 1 \\ M'_Z(t) &= tM_Z(t) \quad\Rightarrow\quad M'_Z(0) = 0 \\ M''_Z(t) &= M_Z(t) + t^2M_Z(t) = (1+t^2)M_Z(t) \quad\Rightarrow\quad M''_Z(0) = 1 \end{aligned}$$

Ahora, si $X \sim \mathcal{N}(\mu,\sigma^2)$ entonces $Z = (X-\mu)/\sigma \sim \mathcal{N}(0,1)$ de manera que:

$$M_X(t) = E\left(e^{tX}\right) = E\left(e^{t(\mu+\sigma Z)}\right) = E\left(e^{t\mu}\,e^{t\sigma Z}\right) = e^{t\mu}M_Z(\sigma t)$$

Derivando:

$$\begin{aligned} M_X(0) &= 1 \\ M'_X(t) &= e^{t\mu}\left(\mu M_Z(\sigma t) + \sigma M'_Z(\sigma t)\right) \quad\Rightarrow\quad E(X) = M'_X(0) = \mu \\ M''_X(t) &= e^{t\mu}\left(\mu^2 M_Z(\sigma t) + 2\mu\sigma M'_Z(\sigma t) + \sigma^2 M_Z(\sigma t)\right) \quad\Rightarrow \\ &\Rightarrow\quad E(X^2) = M''_X(0) = \mu^2+\sigma^2 \end{aligned}$$

Luego:

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

d) Más generalmente, supongamos $X \sim \Gamma(\alpha,\lambda)$ Si $t<\lambda$ se tiene:

$$\begin{aligned} M_X(t) &= E(e^tX) = \int_0^\infty e^{tx}\,\tfrac{\lambda^\alpha}{\Gamma(\alpha)}\,x^{\alpha-1}\,e^{-\lambda x}\,dx = \lambda^\alpha \int_0^\infty \tfrac{1}{\Gamma(\alpha)}\,x^{\alpha-1}\,e^{-(\lambda-t)x}\,dx = \\ &= \tfrac{\lambda^\alpha}{(\lambda-t)^\alpha}\int_0^\infty \tfrac{(\lambda-t)^\alpha}{\Gamma(\alpha)}\,x^{\alpha-1}\,e^{-(\lambda-t)x}\,dx = \tfrac{\lambda^\alpha}{(\lambda-t)^\alpha} = \left(1-\tfrac{t}{\lambda}\right)^{-\alpha} \end{aligned}$$

Derivando se tiene:

$$\begin{aligned} M_X(0) &= 1 \\ M'_X(t) &= \tfrac{\alpha}{\lambda}\left(1-\tfrac{t}{\lambda}\right)^{-\alpha-1} \;\Rightarrow\; E(X) = M'_X(0) = \tfrac{\alpha}{\lambda} \\ M''_X(t) &= \tfrac{\alpha(\alpha+1)}{\lambda^2}\left(1-\tfrac{t}{\lambda}\right)^{-\alpha-2} \;\Rightarrow\; E(X^2) = M''_X(0) = \tfrac{\alpha(\alpha+1)}{\lambda^2} \end{aligned}$$

Por lo tanto:

$$V(X) = E(X^2) - E^2(X) = \frac{\alpha(\alpha+1)}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2 = \frac{\alpha}{\lambda^2}$$

En particular cuando $X \sim \chi^2_n = \Gamma(n/2, 1/2)$ se tiene:

$$\begin{aligned} M_X(t) &= \left(1-\tfrac{t}{\lambda}\right)^{-n/2} \\ E(X) &= \tfrac{n/2}{1/2} = n \quad ; \quad V(X) = \tfrac{n/2}{1/4} = 2n \end{aligned}$$

2. $X \sim \mathcal{N}(0,1)$ de modo que $M_X(t) = e^{t^2/2}$ Derivando cuatro veces:

$$\begin{aligned} M_X(0) &= 1 \\ M'_X(t) &= tM_X(t) \;\Rightarrow\; M'_X(0) = 0 \\ M''_X(t) &= M_X(t) + t^2 M_X(t) = (1+t^2)M_X(t) \;\Rightarrow\; M''_X(0) = 1 \\ M^{(3)}_X(t) &= 2tM_X(t) + (1+t^2)tM_X(t) = (t^3+3t)M_X(t) \;\Rightarrow\; M^{(3)}_X(0) = 0 \\ M^{(4)}_X(t) &= (3t^2+3)M_X(t) + (t^3+3t)tM_X(t) = \\ &= (t^4+6t^2+3)M_X(t) \;\Rightarrow\; E(X^4) = M^{(4)}_X(0) = 3 \end{aligned}$$

3. $S(t) = \ln(M_X(t))$ de modo que por derivación:

$$\begin{aligned} S'(t) &= \tfrac{M'_X(t)}{M_X(t)} \;\Rightarrow\; S'(0) = \tfrac{M'_X(t)}{M_X(t)} = E(X) \\ S''(t) &= \tfrac{M''_X(t)\,M_X(t) - \left(M'_X(t)\right)^2}{M^2_X(t)} \;\Rightarrow\; S''(0) = E(X^2) - E^2(X) = V(X) \end{aligned}$$

4. Completemos el ejercicio 1.c) hasta el momento de orden 3

$$\begin{aligned} M^{(3)}_X(t) &= M_X(t)\left\{(\mu+\sigma^2 t)\left[(\mu+\sigma^2 t)^2 + \sigma^2\right] + 2(\mu+\sigma^2 t)\sigma^2\right\} \\ E(X^3) &= M^{(3)}_X(0) = \mu(\mu+\sigma^2) + 2\mu\sigma^2 = \mu^2 + 3\mu\sigma^2 \end{aligned}$$

Entonces:

$$\begin{aligned} \mu^*_2 &= \left(\tfrac{d}{dt}-\mu\right)^{(2)}\Big|_{t=0} M_X(t) = M''_X(0) - 2\mu M'_X(0) + \mu^2 = \\ &= \mu^2 + \sigma^2 - 2\mu^2 + \mu^2 = \sigma^2 \\ \mu^*_3 &= \left(\tfrac{d}{dt}-\mu\right)^{(3)}\Big|_{t=0} M_X(t) = M^{(3)}_X(0) - 3\mu M''_X(0) + 3\mu^2 M'_X(0) - \mu^3 = \\ &= \mu^2 + 3\mu\sigma^2 - 3\mu(\mu^2+\sigma^2) + 3\mu^3 - \mu^3 = 0 \end{aligned}$$

5. Para $\alpha \in (0,1)$ determinemos el α-cuantil de la distribución de X

   a)
   $$\int_0^{x_\alpha} 3t^2\,dt = \alpha \quad \Leftrightarrow \quad t^3\big|_0^{x_\alpha} = \alpha \quad \Leftrightarrow \quad x_\alpha^3 = \alpha \quad \Leftrightarrow \quad x_\alpha = \alpha^{1/3}$$
   En el caso de la mediana es $\alpha = 1/2$ de modo que $x_{0.5} = 1/\sqrt[3]{2} \approx 0.79$

   b)
   $$\int_{-\infty}^{x_\alpha} \frac{dt}{\pi(1+t^2)} = \alpha \quad \Leftrightarrow \quad \frac{1}{\pi}\ \mathrm{arctg}\ t\big|_{-\infty}^{x_\alpha} = \alpha \quad \Leftrightarrow \quad \frac{1}{\pi}\left(\mathrm{arctg}\ (x_\alpha) + \frac{\pi}{2}\right) = \alpha \quad \Leftrightarrow$$
   $$\Leftrightarrow \quad \mathrm{arctg}\ (x_\alpha) = \pi(\alpha - 1/2) \quad \Leftrightarrow \quad x_\alpha = \mathrm{tg}\ [\pi\,(\alpha - 1/2)]$$
   En el caso de la mediana es $\alpha = 1/2$ de modo que $x_{0.5} = 0$ En este ejemplo (distribución de Cauchy) es importante notar que la esperanza de X no existe pero la mediana sí existe.

6.

   a) Sea $n \in \mathbb{N}$, n impar Vamos a morstrar en primer lugar que si X está distribuida simétricamente respecto del punto $x = c$ y si $E\left((X-c)^n\right)$ existe, entonces necesariamente $E\left((X-c)^n\right) = 0$ Para ello utilizaremos la equivalencia $i) \Leftrightarrow v)$ demostrada en el ejercicio 15.b) del práctico 5. De modo que podemos suponer que:
   $$\forall x > 0\,,\ f_X(c-x) = f_X(c+x)$$
   Entonces se tiene:
   $$\int_{-\infty}^{c} (x-c)^n f_X(x)\,dx = -\int_{\infty}^{0} (-t)^n f_X(c-t)\,dt = -\int_0^{\infty} t^n f_X(c+t)\,dt$$
   $$\int_c^{\infty} (x-c)^n f_X(x)\,dx = \int_0^{\infty} t^n f_X(c+t)\,dt$$
   donde en la primera integral hemos efectuado la sustitución $t = c - x$ y la simetría de la distribución, mientras que en la segunda integral hemos realizado la sustitución $t = x - c$ Sumando ambas integrales se obtiene:
   $$E\left((X-c)^n\right) = \int_{-\infty}^{\infty} (x-c)^n f_X(x)\,dx = 0$$
   Observemos que en particular, para $n = 1$ se tiene: $E(X-c) = 0$ Es decir: $E(X) = c$ Por lo tanto:
   $$\mu_3^* = E\left((X-c)^3\right) = 0$$
   de manera que $\alpha_3 = 0$ para distribuciones simétricas con tercer momento finito.

   b) Utilizamos la fgm para obtener μ_2 , μ_3 , μ_2^* , μ_3^*
   $$\begin{array}{l|l|l} M_X(t) = (1-t)^{-1} & & \\ M_X'(t) = (1-t)^{-2} & \mu_1 = 1 & \\ M_X''(t) = 2(1-t)^{-3} & \mu_2 = 2 & \mu_2^* = 2 - 2 + 1 = 1 \\ M_X^{(3)}(t) = 6(1-t)^{-4} & \mu_3 = 6 & \mu_3^* = 6 - 3\cdot 2 + 3 - 1 = 2 \end{array}$$
   Entonces: $\alpha_3 = 2$

   c) Determinaremos por derivación sucesiva hasta μ_4^* Les doy los resultados de las derivadas en cada caso. Tengan en cuenta que cuando la distribución es simétrica (alrededor del cero) entonces los momentos coinciden con los momentos centrados.

(1) $X \sim \mathcal{N}(0,1)$

$$\begin{array}{l|l} M_X(t) = e^{t^2/2} & M_X(0) = 1 \\ M'_X(t) = tM_X(t) & \mu_1^* = 0 \\ M''_X(t) = (1+t^2)M_X(t) & \mu_2^* = 1 \\ M_X^{(3)}(t) = (t^3+3t)M_X(t) & \mu_3^* = 0 \\ M_X^{(4)}(t) = (t^4+6t^2+3)M_X(t) & \mu_4^* = 3 \end{array}$$

Entonces: $\alpha_4 = 3$

(2) $X \sim \mathcal{U}(-1,1)$ En este caso no recomiendo utilizar la fgm porque resulta más engorroso que calcular los momentos "a mano". De todas formas les doy los resultados por si se presentan dudas. El problema es que las sucesivas derivadas se van complicando (tediosas) y además para $t=0$ habría que calcularlas por definición.
Para $t \neq 0$ se tiene:

$$M_X(t) = E\left(e^{tX}\right) = \frac{1}{2}\int_{-1}^{1} e^{tx}\,dx = \frac{1}{t}\,\frac{e^t - e^{-t}}{2} = \frac{\operatorname{sh} t}{t}$$

Para $t=0$ la situación es diferente:

$$M_X(0) = E(e^0) = E(1) = 1$$

Entonces:

$$M_X(t) = \begin{cases} \frac{\operatorname{sh} t}{t} & \text{si} \quad t \neq 0 \\ 1 & \text{si} \quad t = 0 \end{cases}$$

Los valores de las sucesivas derivadas de $M_X(t)$ en el origen pueden calcularse por L'Hôpital, dando los siguientes resultados para $M_X(t)$

$$M'_X(t) = \begin{cases} \frac{t\operatorname{ch} t - \operatorname{sh} t}{t^2} & \text{si} \quad t \neq 0 \\ 0 & \text{si} \quad t = 0 \end{cases}$$

$$M''_X(t) = \begin{cases} \frac{t^2\operatorname{sh} t - 2t\operatorname{ch} t + 2\operatorname{sh} t}{t^3} & \text{si} \quad t \neq 0 \\ 1/3 & \text{si} \quad t = 0 \end{cases}$$

$$M_X^{(3)}(t) = \begin{cases} \frac{t^3\operatorname{ch} t - 3t^2\operatorname{sh} t + 6t\operatorname{ch} t - 6\operatorname{sh} t}{t^4} & \text{si} \quad t \neq 0 \\ 0 & \text{si} \quad t = 0 \end{cases}$$

$$M_X^{(4)}(t) = \begin{cases} \frac{t^4\operatorname{sh} t - 4t^3\operatorname{ch} t + 12t^2\operatorname{sh} t - 24t\operatorname{ch} t + 24\operatorname{sh} t}{t^5} & \text{si} \quad t \neq 0 \\ 1/5 & \text{si} \quad t = 0 \end{cases}$$

¿Se dan cuenta cuánto más complicado es hallar los momentos mediante la fgm en este caso particular? En fin, en este caso hemos obtenido:

$$\begin{array}{l} \mu_1 = 0 \\ \mu_2 = \mu_2^* = 1/3 \\ \mu_3 = \mu_3^* = 0 \\ \mu_4 = \mu_4^* = 1/5 \end{array}$$

Por lo tanto: $\alpha_4 = \frac{1/5}{(1/3)^2} = \frac{9}{5} = 1.8$

(3) $X \sim DE(1)$ Calculemos la fgm. Si $|t| < 1$ se tiene:

$$\begin{aligned} M_X(t) &= E\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tx}\,\tfrac{1}{2}\,e^{-|x|}\,dx = \tfrac{1}{2}\left[\int_{-\infty}^{0} e^{(1+t)x}\,dx + \int_0^{\infty} e^{-(1-t)x}\,dx\right] = \\ &= \tfrac{1}{2}\left[\tfrac{1}{1+t}\,e^{(1+t)x}\Big|_{-\infty}^{0} - \tfrac{1}{1-t}\,e^{-(1-t)x}\Big|_{0}^{\infty}\right] = \tfrac{1}{2}\left(\tfrac{1}{1+t} + \tfrac{1}{1-t}\right) = (1-t^2)^{-1} \end{aligned}$$

Luego:

$$\begin{aligned}
&M'_X(t) = 2t(1-t^2)^{-2} \\
&M''_X(t) = 2(1-t^2)^{-2} + 8t^2(1-t^2)^{-3} \\
&M^{(3)}_X(t) = 24t(1-t^2)^{-3} + 48t^3(1-t^2)^{-4} \\
&M^{(4)}_X(t) = 24(1-t^2)^{-3} + 288t^2(1-t^2)^{-4} + 384t^4(1-t^2)^{-5}
\end{aligned}$$

Por lo tanto:

$$\begin{aligned}
&\mu_1 = \mu_1^* = 0 \\
&\mu_2 = \mu_2^* = 2 \\
&\mu_3 = \mu_3^* = 0 \\
&\mu_4 = \mu_4^* = 24
\end{aligned}$$

De manera que: $\alpha_4 = 24/4 = 6$

Trabajo práctico 6: Vectores aleatorios

1. Una caja contiene 5 bolillas numeradas de 1 a 5. Se extraen de ella sucesivamente dos bolillas sin reposición. Sean X el mínimo número obtenido e Y el número obtenido en la primera extracción.

   a) Hallar la fmp de X y la fmp de Y

   b) Hallar la fmp conjunta de (X, Y) Calcular las fmp marginales de X e Y, verificando luego el resultado de a).

   c) Calcular $P(X > 3|Y = 3)$

   d) ¿Son X e Y independientes?

2. Supongamos que X e Y son variables aleatorias independientes, cada una con fmp dada por la siguiente tabla:

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

   a) Hallar la fmp conjunta

   b) Hallar la fmp de $Z = X + Y$

   c) Calcular $E(Z)$ de dos maneras distintas.

3. Cuando un automóvil es detenido por una patrulla se revisa el desgaste de cada neumático y se verifica cada faro delantero, para ver si está correctamente alineado. Sea X el número de faros delanteros que necesitan ajuste e Y el número de neumáticos defectuosos. Si X e Y son independientes con respectivas fmp dadas por:

x	0	1	2
$p_X(x)$	0.5	0.3	0.2

y	0	1	2	3	4
$p_Y(y)$	0.6	0.1	0.05	0.05	0.2

   a) Presentar la fmp conjunta de (X, Y) en una tabla a boble entrada.

   b) Calcular $P(X \leq 1, Y \leq 1)$ a partir de la fmp conjunta. Verificar que se cumple:

$$P(X \leq 1, Y \leq 1) = P(X \leq 1)P(Y \leq 1)$$

   c) ¿Cuál es la probabilidad de que no se detecte ninguna violación?

   d) Determinar si existen y en tal caso calcular $P(X + Y \leq 1)$

4. Si X e Y representan las proporciones de dos tipos diferentes de componentes en una mezcla de productos químicos utilizada como insecticida y suponiendo que (X, Y) posee densidad conjunta dada por:

$$f_{X,Y}(x, y) = \begin{cases} c & \text{si} \quad 0 \leq x, y \leq 1\,,\ x + y \leq 1 \\ 0 & \text{en caso contrario} \end{cases}$$

   a) Determinar el valor de la constante c

   b) Calcular $P(X \leq 3/4, Y \leq 3/4)$ y $P(X \leq 1/2, Y \leq 1/2)$

   c) Hallar las densidades marginales de X e Y ¿Son X e Y variables aleatorias independientes?

   d) Calcular $E(XY)$ ¿Qué se puede decir acerca de $cov(X, Y)$ y $\rho(X, Y)$?

5. Dos componentes de microcomputadora tienen la siguiente fdp conjunta para sus tiempos (en horas) de vida útiles X e Y

$$f_{X,Y}(x,y) = \begin{cases} cxe^{-x(1+y)} & \text{si} \quad x \geq 0, y \geq 0 \\ 0 & \text{en caso contrario} \end{cases}$$

   a) Determinar el valor de la constante c
   b) Calcular la probabilidad de que el primer componente dure más de 3 horas.
   c) Hallar las fdp marginales de X e Y ¿Son independientes las duraciones de los dos componentes?
   d) ¿Cuál es la probabilidad de que la duración de por lo menos uno de los componentes sea mayor que 3 horas?

6. Un restaurante sirve 3 comidas de precio fijo, que cuestan $\$\,7$, $\$\,9$ y $\$\,10$, Para una pareja seleccionada al azar que va a comer en ese restaurante sean X el costo de la comida del hombre e Y el costo de la comida de la mujer. La siguiente tabla resume la fmp conjunta del vector aleatorio (X, Y)

X\Y	7	9	10
7	0.05	0.05	0.10
9	0.05	0.10	0.35
10	0.00	0.20	0.10

   a) Calcular las fmp marginales de X e Y
   b) ¿Cuál es la probabilidad de que la comida del hombre y la mujer tengan un costo de a lo sumo $\$\,9$ cada una?
   c) ¿Son X e Y independientes?

7. Se supone que cada neumático delantero de cierto tipo de automóvil se va a llenar a una presión de 26 lb/pulg^2 Suponga que la presión real de aire de cada neumático es una variable aleatoria, X para el derecho e Y para el izquierdo, y que la fdp conjunta de (X, Y) viene dada por:

$$f_{X,Y}(x,y) = \begin{cases} k(x^2+y^2) & \text{si} \quad 20 \leq x \leq 30, 20 \leq y \leq 30 \\ 0 & \text{en caso contrario} \end{cases}$$

   a) ¿Cuál es el valor de la constante k ?
   b) ¿Cuál es la probabilidad de que ambos neumáticos tengan menos presión que la requerida?
   c) ¿Cuál es la probabilidad de que la diferencia de presiones entre los dos neumáticos sea a lo sumo de 2 lb/pulg^2 ?
   d) Determinar la distribución de la presión del neumático derecho.
   e) ¿Son X e Y independientes? ¿Cuánto vale $\rho(X, Y)$?

8. Un ingeniero mide la cantidad de un contaminante en muestras de aire recogidas sobre la chimenea de una central de energía eléctrica que funciona a carbón. Sea X la cantidad del contaminante por muestra recogida cuando no está funcionando cierto dispositivo de limpieza en la chimenea, y sea Y la cantidad de contaminante por muestra recogida bajo las mismas condiciones ambientales pero cuando el dispositivo de limpieza se encuentra en funcionamiento. Se observa que la fdp conjunta de (X, Y) se adecúa al siguiente modelo:

$$f_{X,Y}(x,y) = \begin{cases} k & \text{si} \quad 0 \leq x \leq 2, 0 \leq y \leq 1 \\ 0 & \text{en caso contrario} \end{cases}$$

a) Determinar el valor de la constante k

b) Calcular $P(X > 3Y)$, es decir la probabilidad de que el dispositivo de limpieza reduzca la cantidad de contaminante al menos en un tercio.

c) Si está funcionando el dispositivo limpiador, calcule la probabilidad de que la cantidad de contaminante en una muestra elegida al azar sea mayor que 0.5

d) ¿Son independientes las cantidades de contaminante por muestra con y sin el dispositivo limpiador?

e) La variable aleatoria $X - Y$ representa la cantidad en que se puede reducir el contaminante utilizando el dispositivo de limpieza. Obtener $E(X - Y)$ y $V(X - Y)$

9. Supongamos que la fdp conjunta de un vector aleatorio (X, Y) viene dada por:

$$f_{X,Y}(x, y) = \begin{cases} e^{-y} & \text{si} \quad x > 0, y > x \\ 0 & \text{en caso contrario} \end{cases}$$

a) Calcular las fdp marginales de X e Y

b) Determinar si X e Y son independientes.

c) Calcular $P(X > 1, Y > 1)$

d) Calcular $P(X > 3)$ de dos formas diferentes: Utilizando la fdp conjunta y luego utilizando la fdp marginal de X

e) Calcular $\rho(X, Y)$

RESPUESTAS

1.

(a)

$$R_X = \{1, 2, 3, 4\} \ , \ p_X(k) = \tfrac{5-k}{10}$$

$$R_Y = \{1, 2, 3, 4, 5\} \ , \ p_Y(s) = \tfrac{1}{5}$$

(b)

$X \setminus Y$	1	2	3	4	5	p_X
1	1/5	1/20	1/20	1/20	1/20	2/5
2	0	3/20	1/20	1/20	1/20	3/10
3	0	0	1/10	1/20	1/20	1/5
4	0	0	0	1/20	1/20	1/10
p_Y	1/5	1/5	1/5	1/5	1/5	1

(c)

$$P(X > 3|Y = 3) = \frac{P(X > 3, Y = 3)}{P(Y = 3)} = \frac{P(X = 4, Y = 3)}{P(Y = 3)} = 0$$

(d) X e Y no son independientes. Por ejemplo: $p_{X,Y}(2, 1) = 0 \neq \dfrac{3}{10}\cdot\dfrac{1}{5} = p_X(2)p_Y(1)$ Otra forma de verlo es usando (c). Si X e Y fueran independientes debería ser $P(X > 3|Y = 3) = P(X > 3)$ Esto no ocurre pues $P(X > 3) = 3/10$

2. (a)

$X \setminus Y$	1	2	3	4	p_X
1	0.04	0.06	0.08	0.02	0.2
2	0.06	0.09	0.12	0.03	0.3
3	0.08	0.12	0.16	0.04	0.4
4	0.02	0.03	0.04	0.01	0.1
p_Y	0.2	0.3	0.4	0.1	1

(b) Como X e Y son independientes, puede utilizarse la fórmula de convolución, que en este caso resulta:

$$p_Z(s) = \sum_{k=\text{máx}\{1,s-4\}}^{\text{mín}\{s-1,4\}} p_X(k)p_Y(s-k)$$

Sin embargo es más fácil observar que en nuestra tabla de la fmp conjunta, los pares (x, y) con $x + y = s$ (para s constante) son precisamente los que caen en las "diagonales" de la tabla. Por ejemplo, para hallar todas las combinaciones de valores de x e y que dan suma $s = 4$ simplemente miramos la diagonal indicada con asteriscos en la siguiente tablita:

$X \setminus Y$	1	2	3	4
1			*	
2		*		
3	*			
4				

Entonces por ejemplo para hallar $p_Z(4)$ debemos sumar los valores de la fmp conjunta precisamente sobre los lugares indicados por los asteriscos. De todo esto resulta finalmente:

Z	2	3	4	5	6	7	8
p_Z	0.04	0.12	0.25	0.28	0.22	0.08	0.01

(c) Una manera de calcular $E(Z)$ es por definición, dado que en (b) ya hallamos su fmp. Entonces:

$$E(Z) = \sum_{s=2}^{8} s p_Z(s) = 2(0.04) + 3(0.12) + 4(0.25) + 5(0.28) + 6(0.22) +$$

$$+7(0.08) + 8(0.01) = 4.8$$

Otra forma es calculando $E(X)$ por definición (que coincidirá con $E(Y)$ por ser X e Y idénticamente distribuidas) y luego utilizando linealidad de la esperanza: $E(Z) = E(X) + E(Y)$ En este caso:

$$E(X) = \sum_{s=1}^{4} k p_X(k) = 1(0.2) + 2(0.3) + 3(0.4) + 4(0.1) = 2.4$$

Luego: $E(Z) = 2.4 + 2.4 = 4.8$
Una tercera posibilidad sería utilizando la fórmula: $E(Z) = \sum_{k \in R_X} \sum_{s \in R_Y} (k+s) p_{X,Y}(k, s)$

En este caso el cálculo sería:

$$\begin{aligned}
\boldsymbol{E(Z)} &= \mathbf{(1+1)(0.04)+(1+2)(0.06)+(1+3)(0.08)+(1+4)(0.02)+} \\
&\mathbf{+(2+1)(0.06)+(2+2)(0.09)+(2+3)(0.12)+(2+4)(0.03)+} \\
&\mathbf{+(3+1)(0.08)+(3+2)(0.12)+(3+3)(0.16)+(3+4)(0.04)+} \\
&\mathbf{+(4+1)(0.02)+(4+2)(0.03)+(4+3)(0.04)+(4+4)(0.01)=0.48}
\end{aligned}$$

Trabajo práctico 7: Estimación puntual

1. Sean $X_1, \cdots, X_n$ variables aleatorias i.i.d., cada una con fdp f.
   (a) Sea $X_{(1)} = \min_{1 \le i \le n} X_i$ Determinar la fda de $X_{(1)}$
   (b) Sea $X_{(n)} = \max_{1 \le i \le n} X_i$ Determinar la fda de $X_{(n)}$
   (c) Los estadísticos definidos en (a) y (b) son dos ejemplos de los denominados *estadísticos de orden* de la muestra. Más generalmente, el k-ésimo estadístico de orden $X_{(k)}$ de la muestra se define como el valor muestral que ocupa la posición k cuando la muestra se ordena de menor a mayor, es decir: $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ es la muestra $X_1, X_2, \cdots, X_n$ pero ordenada en forma creciente. Determinar la fda del k-ésimo estadístico de orden.
   (d) Calcular los estadísticos de orden de la muestra:

   5.7 6.3 0.4 -18.8 -18.7 -9.9 -14.9 -12.8 -0.9 1.4

2. Sea $X_1, \cdots, X_n$ una m.a. de la distribución $\mathcal{U}(0, \theta)$. Determinar:
   (a) La fdp de $X_{(1)}$, su esperanza y su varianza.
   (b) La fdp de $X_{(n)}$, su esperanza y su varianza. Si se estima θ mediante $X_{(n)}$, mostrar que tal estimador es sesgado. Obtener a partir de él un estimador insesgado de θ.
   (c) Si se estima θ mediante $2\bar{X}$, mostrar que la estimación resulta insesgada.
   (d) ¿Cuál de los dos estimadores de θ es preferible, el hallado en (b) o el hallado en (c)? Justificar la elección.

3. Sea $t_1, t_2, \cdots, t_n$ una muestra aleatoria de tamaño n de una variable aleatoria discreta T con fmp dada por:

t	0	1	2
$p_T(t)$	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

   donde $0 < \theta < 1$

   (a) Hallar un estimador de θ utilizando el método de los momentos.
   (b) ¿Es insesgado de θ el estimador hallado en (a)?
   (c) Calcular el ECM del estimador hallado en (a) ¿Es consistente dicho estimador?
   (d) Dada la muestra: $0, 0, 0, 0, 1, 1, 1, 1, 2, 2$, ¿cuál es la estimación de θ basada en dicha muestra?

4. Sea X el tiempo relativo empleado en un examen por un estudiante elegido aleatoriamente. Supongamos que X posee fdp perteneciente a la siguiente familia uniparamétrica, donde el parámetro θ verifica $\theta > -1$

$$f_X(x) = \begin{cases} (1+\theta)x^\theta & \text{si} \quad 0 < x < 1 \\ 0 & \text{cc} \end{cases}$$

   Los siguientes son los valores observados de una muestra aleatoria de los tiempos de 10 estudiantes:

   0.92 0.79 0.90 0.65 0.86 0.47 0.73 0.97 0.94 0.77

   (a) Obtener un estimador de θ por el método de los momentos.
   (b) Obtener el EMV de θ

(c) Calcular las estimaciones de $\boldsymbol{\theta}$ para los valores muestrales registrados.

5. Sea $X_1, \cdots, X_n$ una muestra aleatoria de una distribución $\mathcal{P}(\lambda)$
   (a) Deducir el EMV $\hat{\lambda}$ de λ
   (b) Calcular $E_\lambda(\hat{\lambda})$ y $V_\lambda(\hat{\lambda})$
   (c) Demostrar que $\hat{\lambda}$ es un estimador consistente de λ
   (d) En una fábrica de explosivos pueden ocurrir cierto número de inflamaciones al azar. Sea X el número de inflamaciones por día. Suponiendo que X posee distribución de Poisson de parámetro λ, utilizar los datos muestrales siguientes para estimar λ

nro. k de inflamaciones por día:	**0**	**1**	**2**	**3**	**4**	**5**	**6**
nro. de días con k inflamaciones:	**75**	**90**	**54**	**22**	**6**	**2**	**1**

6. La tabla siguiente contiene **24** determinaciones de la temperatura de fusión X del plomo, en grados centígrados:

330.0	**328.6**	**342.4**	**334.0**	**337.5**	**341.0**	**343.3**	**329.5**
322.0	**331.0**	**340.4**	**326.5**	**327.3**	**340.0**	**331.0**	**332.3**
345.0	**342.0**	**329.7**	**325.8**	**322.6**	**333.0**	**341.0**	**340.0**

   Suponiendo normalidad con media μ y varianza σ^2:
   (a) Calcular los valores de los EMV de μ y σ^2 para esta muestra.
   (b) Utilizando (a) hallar el EMV de $P(X > 345)$

7. Supongamos que cierto mecanismo tiene una probabilidad p de fallar. Sea Y el número de días hasta que el mecanismo falle. Se toma una muestra aleatoria $Y_1, \cdots, Y_n$ de la distribución de Y.
   (a) Modelizar la situación y hallar el EMV de p
   (b) Hallar el EMV de la probabilidad de que el mecanismo dure más de **6** días.
   (c) Para **10** de tales mecanismos se registró el número de días hasta la primera falla, resultando los siguientes valores: **5 , 7 , 3 , 6 , 7 , 7 , 5 , 4 , 5 , 6**
   Calcular las estimaciones de (a) y (b) para esta muestra.

8. Sea X una variable aleatoria con $E(X) = \mu$ y $V(X) = \sigma^2$. Sea $X_1, \cdots, X_n$ una muestra aleatoria de la distribución de X. Hay muchos otros estimadores de σ^2 que son insesgados y no coinciden con s^2. Por ejemplo: $\hat{\sigma^2} = C \sum\limits_{i=1}^{n-1} (X_{i+1} - X_i)^2$ es uno de tales estimadores, para un valor apropiado de la constante C. Determinar dicho valor.

9. Supongamos que el tiempo de vida (en horas) de un instrumento electrónico posee la siguiente fdp:

$$f_T(t) = \begin{cases} \beta e^{-\beta(t-t_o)} & \text{si} \quad t > t_o \\ 0 & \text{cc} \end{cases}$$

   donde t_o es un valor fijo, $t_o > 0$. Supongamos que se prueban n instrumentos y se anotan los tiempos $T_1, \cdots, T_n$ hasta que ocurre la falla.
   (a) Determinar el EMV de β suponiendo t_o conocido.
   (b) Determinar el EMV de t_o suponiendo β conocido.

10. Supongamos que $X \sim \mathcal{U}(-\theta, \theta)$. Hallar el EMV de θ basado en una muestra aleatoria $X_1, \cdots, X_n$. Calcular el estimador hallado para la siguiente muestra aleatoria:

-2.04 ;	-0.79 ;	-1.37 ;	-0.11 ;	0.10 ;	1.97 ;	2.86 ;	-0.90 ;	2.77 ;	2.08
1.59 ;	-1.46 ;	1.11 ;	-1.39 ;	-1.64 ;	-0.02 ;	0.64 ;	0.28 ;	0.79 ;	2.03

11. Una variable aleatoria X posee distribución $\mathcal{N}(\mu, 1)$. Se realizan 20 observaciones independientes de X, pero en lugar de anotar su valor sólo observamos si X es negativa o no. Suponiendo que el evento $\{X < 0\}$ ocurrió exactamente 14 veces, utilizar esta información para obtener el EMV de μ.

12. Sea $X_1, \cdots, X_n$ una m.a. de una distribución $\Gamma(\alpha, \lambda)$, donde α se supone conocido. Obtener el EMV de λ basado en la muestra.

13. Sea $X_1, \cdots, X_n$ una m.a. de una distribución con fdp dada por

$$f(x) = \alpha\lambda x^{\alpha-1} e^{-\lambda x^{\alpha}} I_{(0,\infty)}(x)$$

donde α se supone conocido. Obtener el EMV de λ basado en la muestra.

14. Sea $Y_1, \cdots, Y_n$ una m.a. de una población con media μ y varianza σ^2. Considere los tres estimadores siguientes para μ:

$$\hat{\mu}_1 = \frac{Y_1 + Y_2}{2} \quad ; \quad \hat{\mu}_2 = \frac{Y_1}{4} + \frac{Y_2 + \cdots + Y_{n-1}}{2(n-2)} + \frac{Y_n}{4} \quad ; \quad \hat{\mu}_3 = \bar{Y}$$

(a) Demostrar que los tres estimadores son insesgados para μ

(b) Determinar la eficiencia relativa de $\hat{\mu}_3$ respecto de $\hat{\mu}_1$ y luego respecto de $\hat{\mu}_2$

(c) ¿Cuáles de los tres estimadores son consistentes para μ?

15. Los datos siguientes representan la cantidad de minutos que debe esperar una persona el colectivo que lo lleva a su trabajo, habiéndose realizado observaciones durante 15 días laborables:

$$10 \quad 1 \quad 13 \quad 9 \quad 5 \quad 9 \quad 2 \quad 10 \quad 3 \quad 8 \quad 6 \quad 17 \quad 2 \quad 10 \quad 15$$

La muestra puede considerarse proveniente de una distribución $\mathcal{U}(0, \theta)$

(a) Calcular estimadores de θ por método de los momentos y por método de máxima verosimilitud. Hallar las estimaciones correspondientes a los valores muestrales dados.

(b) ¿Son insesgados los estimadores obtenidos en (a)? ¿Son consistentes?

(c) Compare los ECM de los estimadores hallados en (a).

USO DEL R PARA CÁLCULOS

Para ingresar los datos muestrales por ejemplo en el ejercicio 15 se los asigna (operador de asignación `<-`) a un vector de datos, que llamaremos por ejemplo "tiempos", del modo siguiente:

```
> tiempos<-c(10,1,13,9,5,9,10,3,8,6,17,2,10,15)
> tiempos
 [1] 10  1 13  9  5  9 10  3  8  6 17  2 10 15
```

Para averiguar cuántos datos ingresamos utilizamos el comando `length()` que devuelve la longitud del vector puesto como argumento entre `()`

```
> n<-length(tiempos)
> n
[1] 14
```

Si deseamos ordenar los datos en forma creciente utilizamos el comando `sort()`

```
> sort(tiempos)
 [1]  1  2  3  5  6  8  9  9 10 10 10 13 15 17
```

Para calcular la media muestral $\bar{x}$ de los datos se usa el comando `mean()`. El comando `sum()` calcula para un vector de datos $x = (x_1, \cdots, x_n)$, la suma de sus componentes $x_1 + \cdots + x_n$. De este modo es posible verificar que `mean(x)` da efectivamente la media muestral, comparándola con `sum(x)/length(x)`

```
> media<-mean(tiempos)
> media;sum(tiempos)/n
[1] 8.428571
[1] 8.428571
```

Para calcular la varianza muestral s^2 utilizamos el comando `var()`. Para que se convenzan que esto produce el resultado correcto, expliquemos un poco la forma en que R opera con vectores. Las operaciones entre vectores en el R son **componente a componente**. Ejemplifiquemos suma (+), producto (*) por escalares y producto (*) entre vectores, realizando la suma entre `tiempos` y `otros`, siendo éstos los datos: $\mathbf{-3, 3, 1, 0, 6, 1, 1, 1, -8, -1, 0, 0, 2, 0}$

```
> otros<-c(-3,3,1,0,6,1,1,1,-8,-1,0,0,2,0)
> otros
 [1] -3  3  1  0  6  1  1  1 -8 -1  0  0  2  0
> tiempos
 [1] 10  1 13  9  5  9 10  3  8  6 17  2 10 15
> otros
 [1] -3  3  1  0  6  1  1  1 -8 -1  0  0  2  0
> tiempos+otros
 [1]  7  4 14  9 11 10 11  4  0  5 17  2 12 15
 > tiempos*otros
 [1] -30   3  13   0  30   9  10   3 -64  -6   0   0  20   0
> 3*tiempos
 [1] 30  3 39 27 15 27 30  9 24 18 51  6 30 45
```

Entonces podemos calcular la varianza muestral s^2 como `var(tiempos)` y verificarla calculando `sum((tiempos-media)^2)/(n-1)`

```
> var(tiempos);sum((tiempos-media)^2)/(n-1)
[1] 22.26374
[1] 22.26374
```

El desvío standard muestral se calcula como:

```
> desvio<-sqrt(var(tiempos))
> desvio
[1] 4.718446
```

RESPUESTAS

1. Sea F la fda de cada X_i (la misma para todas pues son idénticamente distribuidas)

(a) y (b) Tengamos en cuenta que:

$$\left\{\max_{1\leq i\leq n} X_i \leq u\right\} = \bigcap_{i=1}^{n} \{X_i \leq u\} \quad \text{y} \quad \left\{\min_{1\leq i\leq n} X_i > u\right\} = \bigcap_{i=1}^{n} \{X_i > u\}$$

Entonces:

$$\begin{aligned}
F_{X_{(n)}}(u) &= P(X_{(1)} \leq u) = P\left(\bigcap_{i=1}^{n} \{X_i \leq u\}\right) = \prod_{i=1}^{n} P(X_i \leq u) = (F(u))^n \\
F_{X_{(1)}}(u) &= P(X_{(n)} \leq u) = 1 - P(X_{(n)} > u) = 1 - P\left(\bigcap_{i=1}^{n} \{X_i > u\}\right) = \\
&= 1 - \prod_{i=1}^{n} P(X_i > u) = 1 - (1 - F(u))^n
\end{aligned}$$

(c) Definamos

$$Y_i = I_{(-\infty,u]}(X_i) = \begin{cases} 1 & \text{si} \quad X_i \leq u \\ 0 & \text{si} \quad X_i > u \end{cases}$$

Puesto que $X_1, \cdots, X_n$ son independientes e Y_i depende sólo de X_i, entonces $Y_1, \cdots, Y_n$ son independientes. Además son equidistribuidas, por serlo las X_i. De hecho, si anotamos $p = F(u)$, $q = 1 - F(u)$ entonces $Y_1, \cdots, Y_n$ son iid $\sim \mathcal{B}i(1,p)$
Sea $S = \sum_{i=1}^{n} Y_i$ de modo que $Y \sim \mathcal{B}i(n,p)$ La variable aleatoria S cuenta cuántas de las X_i's son menores o iguales que u
Notemos que el evento $\left\{X_{(k)} \leq u\right\}$ ocurre sii al menos k de las X_i's son menores o iguales que u sii $S \geq k$ Por lo tanto:

$$F_{X_{(k)}}(u) = P(X_{(k)} \leq u) = P(S \geq k) = \sum_{i=k}^{n} \binom{n}{i} p^i q^{n-i} = \sum_{i=k}^{n} \binom{n}{i} [F(u)]^i [1 - F(u)]^{n-i}$$

(d) Utilizamos el comando `sort` del SPlus:

```
x_c(5.7,6.3,0.4,-18.8,-18.7,-9.9,-14.9,-12.8,-0.9,1.4)
sort(x)
 [1] -18.8 -18.7 -14.9 -12.8  -9.9  -0.9   0.4   1.4   5.7   6.3
```

Así por ejemplo $x_{(4)} = -12.8$

2. (a) Usando el ejercicio 1 obtenemos $F_{X_{(1)}}(u) = 1 - (1 - F(u))^n$ de manera que derivando se obtiene: $f_{X_{(1)}}(u) = n\,(1 - F(u))^{n-1} f(u)$ En el caso de la uniforme esto es:

$$f_{X_{(1)}}(u) = \frac{n}{\theta}\left(1 - \frac{u}{\theta}\right)^{n-1} I_{(0,\theta)(u)}$$

El cálculo de $E(X_{(1)})$ y $E(X_{(1)}^2)$ requiere integrar. Utilizaremos el cambio de variables $t = 1 - \frac{u}{\theta}$:

$$\begin{aligned}
E(X_{(1)}) &= \tfrac{n}{\theta} \int_0^{\theta} u \left(1 - \tfrac{u}{\theta}\right)^{n-1} du = n\theta \int_0^1 (1-t) t^{n-1}\, dt = \frac{\theta}{n+1} \\
E(X_{(1)}^2) &= \tfrac{n}{\theta} \int_0^{\theta} u^2 \left(1 - \tfrac{u}{\theta}\right)^{n-1} du = n\theta^2 \int_0^1 (1-t^2) t^{n-1}\, dt = \frac{2\theta^2}{(n+1)(n+2)}
\end{aligned}$$

Por lo tanto:

$$E(X_{(1)}) = \frac{\theta}{n+1} \quad ; \quad V(X_{(1)}) = \frac{n\theta^2}{(n+1)^2(n+2)}$$

(b) Usando el ejercicio 1 obtenemos $F_{X_{(n)}}(u) = (F(u))^n$ de manera que derivando se obtiene: $f_{X_{(n)}}(u) = n\,(F(u))^{n-1}\,f(u)$ En el caso de la uniforme esto es:

$$f_{X_{(n)}}(u) = \frac{n}{\theta}\left(\frac{u}{\theta}\right)^{n-1} I_{(0,\theta)(u)}$$

Calculamos $E(X_{(n)})$ y $E(X^2_{(n)})$ haciendo el mismo cambio de variables que en (a), lo cual da por resultado:

$$E(X_{(n)}) = \frac{n\theta}{n+1} \quad ; \quad V(X_{(n)}) = \frac{n\theta^2}{(n+1)^2(n+2)}$$

Si estimamos $\hat{\theta} = X_{(n)}$ este estimador resulta sesgado pues $E_\theta(\hat{\theta}) = \frac{n}{n+1}\,\theta \neq \theta$ Pero entonces $\hat{\theta}^* = \frac{n+1}{n}\,\hat{\theta}$ resulta insesgado puesto que:

$$E_\theta(\hat{\theta}^*) = E_\theta\left(\frac{n+1}{n}\,\hat{\theta}\right) = \frac{n+1}{n}\,E_\theta(\hat{\theta}) = \frac{n+1}{n}\,\frac{n}{n+1}\,\theta = \theta\ ,\ \forall\theta > 0$$

(c) Definimos $\hat{\theta}^{**} = 2\bar{X}$ Entonces:

$$E_\theta(\hat{\theta}^{**}) = E_\theta(2\bar{X}) = 2\,E_\theta\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{2}{n}\sum_{i=1}^{n} E_\theta(X_i) = \frac{2}{n}\cdot\frac{n\theta}{2} = \theta\ ,\ \forall\theta > 0$$

de manera que $\hat{\theta}^{**}$ es insesgado para θ.

(d) Disponemos de dos estimadores insesgados $\hat{\theta}^*$ y $\hat{\theta}^{**}$ de θ Naturalmente elegimos entre ellos el de menor varianza. Esto nos conduce a comparar las varianzas de ambos estimadores. Calculemos ante todo la varianza de $\hat{\theta}^{**}$

$$V(\hat{\theta}^{**}) = V(2\bar{X}) = 4\,V(\bar{X}) = 4\,\frac{V(X_1)}{n} = 4\cdot\frac{\theta^2}{12n} = \frac{\theta^2}{3n}$$

Entonces:

$$\frac{V(\hat{\theta}^{**})}{V(\hat{\theta}^*)} = \frac{(n+1)^2(n+2)}{3n^2}$$

Desarrollando se obtiene:

$$\frac{V(\hat{\theta}^{**})}{V(\hat{\theta}^*)} > 1 \quad \text{sii} \quad n^3 + n^2 + 5n + 2 > 0$$

Dado que esta desigualdad se verifica para todo n natural, concluimos que: $V(\hat{\theta}^{**}) > V(\hat{\theta}^*)$ para todo n natural. Por lo tanto elegimos como estimador de θ el estimador $\hat{\theta}^* = \max\limits_{1\leq i\leq n} X_i$ En realidad se puede demostrar que $\hat{\theta}^*$ es IMVU (insesgado de mínima varianza uniformemente) de θ, pero la demostración excede los contenidos de nuestro curso.

3. Notemos que la distribución común de los datos es $\mathcal{B}i(2, 1-\theta)$ Sea $T \sim \mathcal{B}i(2, 1-\theta)$

(a) Teniendo en cuenta que $E_\theta(T) = 2(1-\theta)$ y planteando momentos de primer orden, se tiene:

$$2(1-\theta) = \bar{T}$$

Despejando resulta: $\hat{\theta} = 1 - \frac{\bar{T}}{2}$

(b) Para todo $\theta \in (0,1)$ se tiene:

$$E_\theta(\hat{\theta}) = E_\theta\left(1 - \frac{\bar{T}}{2}\right) = 1 - \frac{1}{2}\,E_\theta(\bar{T}) = 1 - \frac{1}{2}\,E_\theta(T_1) = 1 - \frac{1}{2}\cdot 2(1-\theta) = \theta$$

Luego $\hat{\theta}$ es insesgado de θ

(c) Siendo $\hat{\theta}$ insesgado, su ECM coincide con su varianza, que calculamos:

$$E_\theta(\hat{\theta}^2) = E_\theta\left[\left(1 - \frac{\bar{T}}{2}\right)^2\right] = E_\theta\left(1 - \bar{T} + \frac{\bar{T}^2}{4}\right) = 1 - E_\theta(\bar{T}) + \frac{1}{4}\,E_\theta(\bar{T}^2)$$

Pero:

$$\begin{aligned} E_\theta(\bar{T}^2) &= V_\theta(\bar{T}) + E_\theta^2(\bar{T}) = \tfrac{1}{n}\,V_\theta(T_1) + E_\theta^2(T_1) = \\ &= \frac{2\theta(1-\theta)}{n} + 4(1-\theta)^2 \end{aligned}$$

Reemplazando se deduce que:

$$E_\theta(\hat{\theta}^2) = 1 - 2(1-\theta) + \frac{\theta(1-\theta)}{2n} + (1-\theta)^2$$

Luego:

$$ECM_\theta(\hat{\theta}) = V_\theta(\hat{\theta}) = 1 - 2(1-\theta) + \tfrac{\theta(1-\theta)}{2n} + (1-\theta)^2 - \theta^2 = \frac{\theta(1-\theta)}{2n}$$

Como para todo $\theta \in (0,1)$ es $ECM_\theta(\hat{\theta}) = \lim\limits_{n\to\infty} \tfrac{\theta(1-\theta)}{2n} = 0$ entonces $\hat{\theta}$ es un estimador consistente de θ.

(d) Con SPlus:

```
> x_c(0,0,0,0,1,1,1,1,2,2)
> 1-mean(x)/2
[1] 0.6
```

de manera que la estimación de θ a partir de las observaciones es: $\hat{\theta} = 0.6$

4. (a) Calculemos $E_\theta(X)$

$$E_\theta(X) = \int_0^1 (1+\theta)x^{\theta+1}\,dx = (1+\theta)\left.\frac{x^{\theta+2}}{\theta+2}\right|_0^1 = \frac{1+\theta}{2+\theta}$$

La ecuación de momentos de primer orden es $\frac{1+\theta}{2+\theta} = \bar{X}$ Tiene como solución:

$$\hat{\theta} = \frac{2\bar{X} - 1}{1 - \bar{X}}$$

(b) La función de verosimilitud de la muestra es:

$$L(x_1,\cdots,x_n;\theta) = \prod_{i=1}^{n}(1+\theta)x_i^\theta I_{(0,1)}(x_i) = (1+\theta)^n\left(\prod_{i=1}^{n} x_i\right)^\theta h(x_1,\cdots,x_n)$$

donde $h(x_1, \cdots, x_n) = \prod_{i=1}^{n} I_{(0,1)}(x_i)$ no depende de θ
Tomando logaritmo se tiene:

$$\ln\left(L(x_1, \cdots, x_n; \theta)\right) = n \ln(1+\theta) + \theta \sum_{i=1}^{n} x_i$$

Derivando respecto de θ e igualando a cero:

$$\frac{n}{1+\theta} + \sum_{i=1}^{n} \ln x_i = 0$$

ecuación que tiene como solución:

$$\hat{\theta}^* = -1 - \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \ln X_i}$$

(c) Con SPlus:

```
> x_c(0.92,0.7,0.90,0.65,0.86,0.47,0.73,0.97,0.94,0.77)
> theta.hat_(2*mean(x)-1)/(1-mean(x))
> hat_(2*mean(x)-1)/(1-mean(x))
> n_length(x)
> hat.star_-1-n/sum(log(x))
> hat
[1] 2.784689
> hat.star
[1] 2.920868
```

5. (a) La función de verosimilitud es:

$$L(x_1, \cdots, x_n; \lambda) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

Tomando logaritmo se obtiene:

$$\ln(L(x_1, \cdots, x_n; \lambda)) = -n\lambda + \left(\sum_{i=1}^{n} x_i\right) \ln \lambda - \sum_{i=1}^{n} \ln(x_i!)$$

Derivando respecto de λ e igualando a cero se obtiene: $\hat{\lambda} = \bar{X}$

(b) Se tiene:

$$E_\lambda(\hat{\lambda}) = E_\lambda(\bar{X}) = E_\lambda(X_1) = \lambda$$

$$V_\lambda(\hat{\lambda}) = V_\lambda(\bar{X}) = \frac{V_\lambda(X_1)}{n} = \frac{\lambda}{n}$$

(c) De (b) se deduce que $\hat{\lambda}$ es insesgado de λ. Entonces:

$$ECM_\lambda(\hat{\lambda}) = V_\lambda(\hat{\lambda}) = \frac{\lambda}{n}$$

Luego: $\lim_{n \to \infty} ECM_\lambda(\hat{\lambda}) = 0$ lo que muestra que $\hat{\lambda}$ es consistente para λ

(d) Con SPlus:

```
> k_0:6
> x_c(75,90,54,22,6,2,1)
> n_sum(x)
> n
[1] 250
> x.bar_sum(x*k)/n
> x.bar
[1] 1.216
```

6. (a) Como vimos en la teórica, los EMV de μ y de σ^2 son:

$$\hat{\mu} = \bar{X} \qquad ; \qquad \hat{\sigma^2} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Con el SPlus:

```
> x1_c(330.0,328.6,342.4,334.0,337.5,341.0,343.3,329.5)
> x2_c(322.0,331.0,340.4,326.5,327.3,340.0,331.0,332.3)
> x3_c(345.0,342.0,329.7,325.8,322.6,333.0,341.0,340.0)
> x_c(x1,x2,x3)
> n_length(x)
> n
[1] 24
> mu.hat_mean(x)
> mu.hat
[1] 333.9958
> sigmacuad.hat_var(x,unbiased=F)
> sigmacuad.hat
[1] 46.39123
```

(b) Se tiene:

$$P_{(\mu,\sigma^2)}(X > 345) = P_{(\mu,\sigma^2)}\left(\frac{X-\mu}{\sigma} > \frac{345-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{345-\mu}{\sigma}\right)$$

Es decir, la probabilidad pedida es una función de (μ, σ^2). Entonces para hallar su EMV se reemplazan μ y σ^2 por $\hat{\mu}$ y $\hat{\sigma}^2$ respectivamente. Luego:

$$\hat{P}_{(\mu,\sigma^2)}(X > 345) = 1 - \Phi\left(\frac{345 - \mu\hat{\mu}}{\sqrt{\hat{\sigma^2}}}\right)$$

La estimación de esta probabilidad a partir de los datos muestrales es:

$$\hat{P}_{(\mu,\sigma^2)}(X > 345) = 1 - \Phi\left(\frac{345 - 333.9958}{\sqrt{46.39123}}\right)$$

Con SPlus:

```
> proba.hat_1-pnorm((345-mu.hat)/sqrt(sigmacuad.hat))
> proba.hat
[1] 0.05308826
```

7. (a) La variable aleatoria Y tiene distribución geométrica de parámetro p, de modo que:

$$p_Y(y) = (1-p)^{y-1}p \quad , y \in \mathbb{N}$$

La función de verosimilitud es:

$$L(y_1, \cdots, y_n; p) = \prod_{i=1}^{n} (1-p)^{y_i - 1} p = p^n (1-p)^{-n + \sum\limits_{i=1}^{n} y_i}$$

Tomando logaritmo:

$$\ln(L(y_1, \cdots, y_n; p)) = n \ln p + \left(\sum_{i=1}^{n} y_i - n \right) \ln(1-p)$$

Derivando respecto de p se obtiene:

$$\frac{n}{p} - \left(\sum_{i=1}^{n} y_i - n \right) \frac{1}{1-p}$$

Igualando a cero y despejando se obtiene como EMV de p el siguiente: $\hat{p} = \dfrac{1}{\bar{Y}}$

(b) La probabilidad de que el mecanismo dure más de 6 días es una función paramétrica:

$$P_p(Y > 6) = 1 - P_p(Y \leq 6) = 1 - \sum_{i=1}^{6} q^{i-1} p = 1 - p \cdot \frac{1-q^6}{p} = (1-p)^6$$

Entonces:

$$\hat{P}_p(Y > 6) = (1-\hat{p})^6 = \left(1 - \frac{1}{\bar{Y}} \right)^6$$

(c) Con SPlus:

```
> x_c(5,7,3,6,7,7,5,4,5,6)
> n_length(x)
> n
[1] 10
> p.hat_1/mean(x)
> p.hat
[1] 0.1818182
> prob6.hat_(1-p.hat)^6
> prob6.hat
[1] 0.2999846
```

8. Se tiene

$$\begin{aligned}
E(\hat{\sigma^2}) &= E\left(C \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2 \right) = C \sum_{i=1}^{n-1} E\left[(X_{i+1} - X_i)^2 \right] = \\
&= C \sum_{i=1}^{n-1} E\left[(X_{i+1}^2 - 2X_i X_{i+1} + X_i^2)^2 \right] = C \sum_{i=1}^{n-1} \left[E(X_{i+1}^2) - 2E(X_i X_{i+1}) + E(X_i^2)^2 \right] = \\
&= C \sum_{i=1}^{n-1} \left[\sigma^2 + \mu^2 - 2E(X_i)E(X_{i+1}) + \sigma^2 + \mu^2 \right] = \\
&= C \sum_{i=1}^{n-1} \left[2\sigma^2 + 2\mu^2 - 2\mu^2) \right] = 2C \sum_{i=1}^{n-1} \sigma^2 = 2(n-1)C\sigma^2
\end{aligned}$$

Por lo tanto $\hat{\sigma^2}$ es insesgado de σ^2 sii $\forall \sigma^2 > 0$, $2(n-1)C\sigma^2 = \sigma^2$
Es decir sii $2(n-1)C = 1$ sii $C = \frac{1}{2(n-1)}$ De manera que otro estimador insesgado de σ^2 resulta ser:

$$\hat{\sigma^2} = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2$$

Comentario: Calculemos la varianza de s^2 Vimos en la teórica que:

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

Pero en la primera parte del curso calculamos esperanza y varianza de una chi cuadrado con k grados de libertad: Si $W \sim \chi^2_k$ entonces $E(W) = k$ y $V(W) = 2k$ Por lo tanto:

$$V\left(\frac{(n-1)s^2}{\sigma^2}\right) = 2(n-1)$$

Se deduce que: $\dfrac{(n-1)^2 V(s^2)}{\sigma^4} = 2(n-1)$ de modo que:

$$V(s^2) = \frac{2}{n-1}\,\sigma^4$$

Esto muestra que s^2 es consistente para σ^2 (por lo menos bajo normalidad)

9. La función de verosimilitud es:

$$\begin{aligned} L &= \prod_{i=1}^{n} \beta e^{-\beta(t_i - t_o)} I_{(t_o,\infty)}(t_i) = \beta^n e^{-\beta \sum\limits_{i=1}^{n}(t_i - t_o)} I_{(t_o,\infty)}(t_{(1)}) = \\ &= \beta^n e^{-n\beta(\bar{t} - t_o)} I_{(t_o,\infty)}(t_{(1)}) \end{aligned}$$

   (a) Omitiendo en L la indicadora (que no depende de β) y tomando logaritmo se obtiene:

   $$\ln(L(t_1, \cdots, t_n; \beta)) = n \ln \beta - n\beta(\bar{t} - t_o)$$

   Derivamos respecto de β e igualamos a cero, lo cual da: $\hat{\beta} = \dfrac{1}{\bar{T} - t_o}$

   (b) El factor β^n en L no depende de t_o Reescribimos L como:

   $$L(t_1, \cdots, t_n; t_o) = \beta^n e^{-n\beta(\bar{t} - t_o)} I_{(-\infty, t_{(1)})}(t_o)$$

   Esta función (como función de t_o) es estrictamente creciente en $(-\infty, t_{(1)})$ y nula de ahí en adelante. Por lo tanto: $\hat{t_o} = T_{(1)} = \min\limits_{1 \le i \le n} T_i$

10. La función de verosimilitud es:

$$\begin{aligned} L(x_1, \cdots, x_n; \theta) &= \prod_{i=1}^{n} \tfrac{1}{2\theta} I_{(-\theta,\theta)}(x_i) = \left(\tfrac{1}{2\theta}\right)^n I_{(-x_{(1)},\infty)}(\theta) I_{(x_{(n)},\infty)}(\theta) = \\ &= \left(\tfrac{1}{2\theta}\right)^n I_{(\max\{-x_{(1)}, x_{(n)}\},\infty)}(\theta) = \left(\tfrac{1}{2\theta}\right)^n I_{(\max\limits_{1\le i\le n} |X_i|,\infty)}(\theta) \end{aligned}$$

Esta función es estrictamente decreciente en θ a partir de $\max\limits_{1 \le i \le n} |X_i|$ y nula antes, de manera que el EMV de θ resulta:

$$\hat{\theta} = \max_{1 \le i \le n} |X_i| = |X|_{(i)}$$

Es decir: $\hat{\theta}$ es el máximo módulo de los datos observados. Para los datos observados la estimación resulta:

```
> x1 <- c(-2.04, -0.79, -1.37, -0.11, 0.1, 1.97, 2.86, -0.9, 2.77, 2.08)
> x2 <- c(1.59, -1.46, 1.11, -1.39, -1.64, -0.02, 0.64, 0.28, 0.79, 2.03)
> x <- c(x1, x2)
> n <- length(x)
> n
[1] 20
> abs.ord <- sort(abs(x))
> abs.ord
 [1] 0.02 0.10 0.11 0.28 0.64 0.79 0.79 0.90 1.11 1.37 1.39 1.46 1.59 1.64 1.97
[16] 2.03 2.04 2.08 2.77 2.86
> theta.hat <- max(abs.ord)
> theta.hat
[1] 2.86
```

11. Sean $Y_i = I_{(-\infty,0)}(X_i)$ Entonces las Y_i son las observaciones. Se tiene: $Y_1, \cdots, Y_n$ es muestra aleatoria de la distribución $\mathcal{B}i(1,p)$ siendo $p = P_\mu(X < 0) = P(X - \mu < -\mu) = \Phi(-\mu)$ La función de verosimilitud es:

$$L(y_1, \cdots, y_n; \mu) = \prod_{i=1}^{n} p^{y_i}(1-p)^{1-y_i} = p^{\sum\limits_{i=1}^{n} y_i}(1-p)^{n-\sum\limits_{i=1}^{n} y_i}$$

Tomando logaritmo:

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

Derivando respecto de μ e igualando a cero, resulta:

$$-\left(\sum_{i=1}^{n} y_i\right)\frac{\varphi(-\mu)}{\Phi(-\mu)} + \left(n - \sum_{i=1}^{n} y_i\right)\frac{\varphi(-\mu)}{1-\Phi(-\mu)} = 0$$

Despejando se obtiene $\Phi(-\mu) = \bar{y}$ Por lo tanto:

$$\hat{\mu} = -\Phi^{-1}(\bar{Y})$$

Para la muestra observada resulta $\bar{y} = \frac{14}{20}$ de manera que:

```
> mu.hat <-  - pnorm(14/20)
> mu.hat
[1] -0.7580363
```

12. La función de verosimilitud en este caso es:

$$L(x_1, \cdots .x_n; \lambda) = \prod_{i=1}^{n} \frac{\lambda^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\lambda x_i} I_{(0,\infty)}(x_i) = \frac{\lambda^{n\alpha}}{(\Gamma(\alpha))^n}\left(\prod_{i=1}^{n} x_i\right) e^{-\sum\limits_{i=1}^{n} x_i} I_{(0,\infty)}(x_{(1)})$$

Tomando logaritmo se obtiene:

$$\ln(L) = n\alpha\ln\lambda - n\ln(\Gamma(\alpha)) + (\alpha-1)\sum_{i=1}^{n}\ln x_i - \lambda\sum_{i=1}^{n} x_i$$

Derivando respecto de λ e igualando a cero se tiene:

$$\frac{n\alpha}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

cuya solución es $\hat{\lambda} = \dfrac{\alpha}{\bar{X}}$

13. La función de verosimilitud es:

$$\begin{aligned} L(x_1,\cdots,x_n;\lambda,\alpha) &= \prod_{i=1}^{n} \lambda\alpha x_i^{\alpha-1} e^{-\lambda x_i^\alpha} I_{(0,\infty)}(x_i) = \\ &= \lambda^n \alpha^n \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} e^{-\lambda \sum\limits_{i=1}^{n} x_i^\alpha} I_{(0,\infty)}(\min_{1\le i\le n} x_i) \end{aligned}$$

Tomando logaritmo:

$$\ln L = n\ln\alpha + n\ln\lambda + (\alpha-1)\sum_{i=1}^{n}\ln x_i - \lambda\sum_{i=1}^{n} x_i^\alpha$$

Derivando respecto de λ e igualando a cero se tiene:

$$\frac{n}{\lambda} - \sum_{i=1}^{n} x_i^\alpha = 0$$

cuya solución es $\hat{\lambda} = \dfrac{1}{\frac{1}{n}\sum\limits_{i=1}^{n} X_i^\alpha}$

14. (a) Se tiene:

$$\begin{aligned} E(\hat{\mu}_1) &= E\left(\tfrac{Y_1+Y_2}{2}\right) = \tfrac{1}{2}(E(Y_1)+E(Y_2)) = \tfrac{1}{2}(\mu+\mu) = \mu \\ E(\hat{\mu}_2) &= E\left(\tfrac{Y_1}{4} + \tfrac{1}{2(n-2)}\sum_{i=2}^{n-1} Y_i + \tfrac{Y_n}{4}\right) = \tfrac{E(Y_1)}{4} + \tfrac{1}{2(n-2)}\sum_{i=2}^{n-1} E(Y_i) + \tfrac{E(Y_n)}{4} = \\ &= \tfrac{\mu}{4} + \tfrac{1}{2(n-2)}(n-2)\mu + \tfrac{\mu}{4} = \mu \\ E(\hat{\mu}_3) &= E(\bar{Y}) = E\left(\tfrac{1}{n}\sum_{i=1}^{n} Y_i\right) = \tfrac{1}{n}\sum_{i=1}^{n}\mu = \tfrac{n\mu}{n} = \mu \end{aligned}$$

(b) Dados dos estimadores $\hat{\theta}_1$, $\hat{\theta}_2$ insesgados de θ, se define la eficiencia relativa de $\hat{\theta}_2$ respecto de $\hat{\theta}_1$ como:

$$\text{eff}_{\,2,1} = \frac{Var(\hat{\theta}_1)}{Var(\hat{\theta}_2)}$$

La eficiencia relativa es pues una manera de comparar la performance relativa de dos estimadores insesgados para un parámetro, mediante la comparación de sus varianzas (que miden en cada caso la precisión de cada estimación). Se tiene: eff $_{2,1} > 1$ si y sólo si $Var(\hat{\theta}_2) < Var(\hat{\theta}_1)$ Cuando eff $_{2,1} > 1$ se dice que $\hat{\theta}_2$ es más eficiente que $\hat{\theta}_1$ dado que produce estimaciones más precisas que $\hat{\theta}_1$

Calculemos en este caso las varianzas de los tres estimadores:

$$\begin{aligned} V(\hat{\mu}_1) &= V\left(\tfrac{Y_1+Y_2}{2}\right) = \tfrac{1}{4}(V(Y_1)+V(Y_2)) = \tfrac{\sigma^2+\sigma^2}{4} = \tfrac{\sigma^2}{2} \\ V(\hat{\mu}_2) &= V\left(\tfrac{Y_1}{4} + \tfrac{1}{2(n-2)}\sum_{i=2}^{n-1} Y_i + \tfrac{Y_n}{4}\right) = \tfrac{V(Y_1)}{16} + \tfrac{1}{4(n-2)^2}\sum_{i=2}^{n-1} V(Y_i) + \tfrac{V(Y_n)}{16} = \\ &= \tfrac{\sigma^2}{16} + \tfrac{(n-2)\sigma^2}{4(n-2)^2} + \tfrac{\sigma^2}{16} = \tfrac{n\sigma^2}{8(n-2)} \\ V(\hat{\mu}_3) &= V(\bar{Y}) = \tfrac{\sigma^2}{n} \end{aligned}$$

Por lo tanto:

$$\text{eff}_{\,3,1} = \tfrac{n}{2} \quad ; \quad \text{eff}_{\,3,2} = \tfrac{n^2}{8(n-2)} \quad ; \quad \text{eff}_{\,2,1} = \tfrac{4(n-2)}{n}$$

Por lo tanto, para $n \geq 5$ el más eficiente de los tres es $\hat{\mu}_3 = \bar{Y}$ Observemos que para $n = 4$ es $\hat{\mu}_2 = \hat{\mu}_3$

(c) Evidentemente $\hat{\mu}_1$ es inconsistente para μ puesto que como no depende de n resulta imposible que tienda en probabilidad a la constante μ También resulta $\hat{\mu}_2$ inconsistente para μ puesto que su varianza no tiende a cero: $Var(\hat{\mu}_2) \underset{n\to\infty}{\longrightarrow} \sigma^2/8$ Naturalmente $\hat{\mu}_3 = \bar{Y}$ es consistente para μ

15. (a) Por método de los momentos de primer orden:

$$\frac{\theta}{2} = E_\theta(X_1) = \bar{X}$$

de modo que $\hat{\theta} = 2\bar{X}$ Por método de máxima verosimilitud:

$$\begin{aligned} L(x_1, \cdots, x_n; \theta) &= \prod_{i=1}^{n} \frac{1}{\theta} I_{(0,\theta)}(x_i) = \frac{1}{\theta^n} I_{(-\,\theta)}(\max_{1\leq i\leq n} x_i) I_{(0,\infty)}(\min_{1\leq i\leq n} x_i) = \\ &= \frac{1}{\theta^n} I_{(\max\limits_{1\leq i\leq n} x_i, \infty)}(\theta) I_{(0,\infty)}(\min_{1\leq i\leq n} x_i) \end{aligned}$$

de modo que: $\hat{\theta}_{EMV} = \max\limits_{1\leq i\leq n} X_i$

Para la muestra dada se tienen las siguientes estimaciones:

```
x_c(10,1,13,9,5,9,2,10,3,8,6,17,2,10,15)
hat.mom_2*mean(x)
hat.mom
[1] 16
hat.mv_max(x)
hat.mv
[1] 17
```

(b) Se tiene:

$$\begin{aligned} E_\theta(\hat{\theta}) &= E_\theta(2\bar{X}) = 2E_\theta(X_1) = 2 \cdot \frac{\theta}{2} = \theta \\ E_\theta(\hat{\theta}_{EMV}) &= E_\theta(\max_{1\leq i\leq n} X_i) = \frac{n\theta}{n+1} \end{aligned}$$

donde para el segundo resultado usamos el ejercicio 2. Entonces $\hat{\theta}$ es insesgado pero $\hat{\theta}_{EMV}$ es sesgado. En cuanto a las varianzas, se tiene:

$$\begin{aligned} V_\theta(\hat{\theta}) &= V_\theta(2\bar{X}) = 4 \cdot \frac{V(X_1)}{n} = \frac{4\theta^2}{12n} = \frac{\theta^2}{3n} \underset{n\to\infty}{\longrightarrow} 0 \\ V_\theta(\hat{\theta}_{EMV}) &= V_\theta(\max_{1\leq i\leq n} X_i) = \frac{n\theta^2}{(n+1)^2(n+2)} \underset{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

Además el sesgo de $\hat{\theta}_{EMV}$ es:

$$b_\theta(\hat{\theta}_{EMV}) = E_\theta(\hat{\theta}_{EMV}) - \theta = \frac{n\theta}{n+1} - \theta = -\frac{\theta}{n+1} \underset{n\to\infty}{\longrightarrow} 0$$

tiende a cero para $n \to \infty$ Por lo tanto, ambos estimadores son consistentes de θ

(c) Se tiene:

$$\boldsymbol{ECM_\theta(\hat{\theta}) = \frac{\theta^2}{3n} \quad ; \quad ECM_\theta(\hat{\theta}_{EMV}) = \frac{2\theta^2}{(n+1)(n+2)}}$$

Entonces:

$$\boldsymbol{\frac{ECM_\theta(\hat{\theta})}{ECM_\theta(\hat{\theta}_{EMV})} = \frac{(n+1)(n+2)}{6n} > 1} \quad \text{sii} \quad \boldsymbol{n \geq 3}$$

Trabajo práctico 8: Intervalos de confianza

1. Interesa conocer el nivel medio de hemoglobina de la población de niños menores de 6 años intoxicados con plomo. Se supone que la distribución del nivel de hemoglobina en esta población tiene distribución normal con desvío standard $\mathbf{0.85}$ gramos cada $\mathbf{100}$ mililitros.

   (a) Si se desea tener una estimación con un IC del $\mathbf{95\%}$ cuya longitud no sea mayor que $\mathbf{0.7}$, ¿cuál será el tamaño muestral necesario?

   (b) Se tiene una muestra de $\mathbf{26}$ niños que han estado expuestos a altos niveles de plomo. Para estos niños el nivel promedio de hemoglobina fue de $\mathbf{10.6}$ gramos cada $\mathbf{100}$ mililitros. Construir un IC de $\mathbf{95\%}$ de confianza para la media poblacional $\boldsymbol{\mu}$ del nivel de hemoglobina.

   (c) Calcular un IC de nivel $\mathbf{95\%}$ para $\boldsymbol{e^{-\mu}}$

2. La tabla siguiente presenta $\mathbf{24}$ determinaciones de la temperatura de fusión del plomo, en grados centígrados:

330.0	328.6	342.4	334.0	337.5	341.0	343.3	329.5	322.0	331.0	340.4	326.5
327.3	340.0	331.0	332.3	345.0	342.0	329.7	325.8	322.6	333.0	341.0	340.0

   Suponiendo normalidad:

   (a) Calcular un IC de nivel $\mathbf{0.95}$ para la media poblacional $\boldsymbol{\mu}$

   (b) ¿De qué tamaño debería tomarse una segunda muestra para que la longitud del IC para $\boldsymbol{\mu}$ al $\mathbf{95\%}$ fuera a lo sumo $\mathbf{4.0}$?

   (c) Calcular un IC de nivel $\mathbf{0.95}$ para el desvío standard poblacional.

3. A continuación se listan los valores del consumo diario de calorías, en kilocalorías por kilogramo, de una muestra aleatoria de $\mathbf{35}$ adolescentes saludables:

20.7	22.4	23.1	23.8	24.5	25.3	25.7
26.3	26.7	30.2	30.6	31.6	32.1	32.9
33.2	33.7	34.2	34.6	35.0	35.4	35.4
36.1	36.6	36.8	37.0	37.1	37.4	37.4
38.1	38.3	38.6	39.0	39.2	40.0	40.8

   Construir un IC del $\mathbf{95\%}$ para la media del consumo diario de calorías para la población de adolescentes En realidad, ¿es exacto o aproximado el nivel?

4. Se tiene una caja de $\mathbf{10000}$ tornillos, de los cuales existe cierta proporción $\boldsymbol{p}$ desconocida de defectuosos.

   (a) Se extraen $\mathbf{50}$ tornillos al azar y se encuentra que $\mathbf{4}$ de ellos son defectuosos. Con estos datos construir un IC de nivel de confianza $\mathbf{0.90}$ para $\boldsymbol{p}$ ¿Es exacto el nivel del IC? ¿Porqué?

   (b) ¿Cuál debe ser el tamaño muestral si se deseara que la longitud del IC de (a) fuera a lo sumo $\mathbf{0.08}$?

5. Arrojar $\mathbf{30}$ veces un dado, anotar los números obtenidos y utilizar la muestra para calcular un IC de nivel $\mathbf{0.95}$ para la probabilidad de obtener con una tirada de dicho dado un $\mathbf{5}$ o un $\mathbf{6}$.

6. En un estudio de angina de pecho en ratas, se dividieron aleatoriamente $\mathbf{18}$ animales infectados en dos grupos de $\mathbf{9}$ individuos cada uno. A uno de los grupos se le suministró un placebo y al otro un fármaco experimental. Después de un ejercicio controlado sobre una rueda de andar, se determinó el tiempo de recuperación de cada rata. Se piensa que el fármaco reducirá el tiempo

de recuperación. Se dispone de la información siguiente:

Placebo	$n_1 = 9$	$\bar{x}_1 = 329$ seg	$S_1 = 45$ seg
Fármaco	$n_2 = 9$	$\bar{x}_2 = 283$ seg	$S_2 = 43$ seg

Suponiendo normalidad y varianzas iguales, calcular un IC de nivel $\mathbf{0.99}$ para la diferencia de medias de tiempos de recuperación sin y con el fármaco.

7. En un estudio de características corporales de las gaviotas de pico anillado, la variable aleatoria considerada es la longitud del pico. Se dispone de los siguientes datos para una muestra aleatoria:

Hembras	$n_1 = 51$	$\bar{x}_1 = 59.9$ mm	$S_1 = 1.9$ mm
Machos	$n_2 = 41$	$\bar{x}_2 = 65.2$ mm	$S_2 = 2.0$ mm

Si se desconoce la distribución de la variable aleatoria bajo estudio, ¿cómo calcularía un IC para la diferencia de medias entre las longitudes de los picos de hembras y machos? ($\alpha = 0.05$) ¿Es el IC de nivel exacto o aproximado?

8. Un fabricante de baterías para automóvil afirma que sus baterías durarán en promedio tres años, con una varianza de un año. Cinco de estas baterías presentaron las siguientes duraciones en años: $\mathbf{1.9}$, $\mathbf{2.4}$, $\mathbf{3.0}$, $\mathbf{3.5}$, $\mathbf{4.2}$ Construir un IC del $\mathbf{95\%}$ para la varianza poblacional de la duración de las baterías y decida si es aceptable a tal nivel de confianza la afirmación del fabricante. Suponer que la duración de las baterías es normal.

9. Un experimento compara las economías en combustible de camiones compactos a diesel equipados en forma similar. Supongamos que se utilizaron $\mathbf{12}$ camiones Volkswagen (VW) y $\mathbf{10}$ camiones Toyota en pruebas de velocidad constante de $\mathbf{90}$ kilómetros por hora. Los $\mathbf{12}$ VW promediaron $\mathbf{16}$ kilómetros por litro con una desviación standard de $\mathbf{1.0}$ kilómetros por litro y los $\mathbf{10}$ Toyota promediaron $\mathbf{11}$ kilómetros por litro, con un desvío standard de $\mathbf{0.8}$ kilómetros por litro. Supongamos que las distancias por litro para cada modelo de camión están normalmente distribuídas.

   (a) Construir un IC de nivel $\mathbf{98\%}$ para el cociente entre los desvíos standard poblacionales de las distancias por litro entre ambos modelos de camiones.

   (b) Obtener un IC del $\mathbf{90\%}$ para la diferencia entre los kilómetros promedio por litro de estos dos camiones compactos.

10. Los siguientes datos representan la duración en horas de una muestra aleatoria de $\mathbf{36}$ lamparitas eléctricas:

1009	1085	1123	1235	1249	1263	1368	1397	1406	1425	1437	1483
1519	1541	1543	1620	1625	1638	1720	1757	1783	1796	1828	1834
1292	1488	1658	1871	1352	1499	1673	1881	1359	1509	1682	1502

Supongamos que la duración en horas de las lamparitas tiene distribución exponencial de parámetro $\boldsymbol{\lambda}$

   (a) Hallar un IC de nivel exacto $\mathbf{95\%}$ para la duración media real de las lamparitas.

   (b) Hallar un IC de nivel exacto $\mathbf{95\%}$ para la mediana poblacional de la duración de las lamparitas.

11. En una fábrica de explosivos pueden ocurrir cierta cantidad de inflamaciones al azar. Sea $\boldsymbol{X}$ el número de inflamaciones por día. Supongamos que $\boldsymbol{X \sim \mathcal{P}(\lambda)}$. Una muestra aleatoria de la distribución de $\boldsymbol{X}$ exhibe los siguientes resultados:

k: # inflamaciones por día	0	1	2	3	4	5	6
# días con k inflamaciones	75	90	54	22	6	2	1

Calcular un IC de nivel asintótico **0.90** para $\boldsymbol{\lambda}$.

12. El gobierno otorga fondos para los departamentos de agricultura de nueve universidades, para probar las capacidades de rendimiento de dos nuevas variedades $\boldsymbol{A}$ y $\boldsymbol{B}$ de trigo. Dentro de cada universidad ambas variedades se plantan en parcelas de igual área y similares características ambientales. El rendimiento en $\boldsymbol{kg}$ por parcela se presenta en la siguiente tabla:

	Universidad								
Variedad	1	2	3	4	5	6	7	8	9
A	38	23	35	41	44	29	37	31	38
B	45	25	31	38	50	33	36	40	43

Determinar un IC del **95%** para la diferencia media entre los rendiemientos de las dos variedades. Supóngase que las diferencias de rendimiento se distribuyen en forma normal.

13. Un fabricante de reproductores de discos compactos utiliza un amplio conjunto de pruebas para evaluar el funcionamiento de su producto. Todos los reproductores deben pasar exitosamente por todas las pruebas antes de llevarlos a la venta. Una muestra aleatoria de **500** reproductores tiene como resultado **15** que fallan al menos en una prueba. Encontrar un IC de nivel **0.90** para la verdadera proporción de reproductores que son admisibles para ser llevados a la venta.

14. En una muestra aleatoria de **1000** casas de cierta ciudad, se encuentra que **228** se calientan en base a petróleo.

    (a) Hallar un IC del **99%** para la proporción de casas en dicha siudad que se calientan en base a petróleo.

    (b) ¿Qué tan grande se requiere que sea la muestra si deseamos tener una confianza del **99%** de que la proporción muestral esté dentro del **0.05** de la verdadera proporción de casas que se calientan en base a petróleo?

15. Un genetista se interesa en la proporción de hombres y de mujeres en cierta población que padecen determinado trastorno sanguíneo menor. En una muestra aleatoria de **1000** hombres se encuentra que **250** lo padecen, mientras que en una de **1000** mujeres se encuentra que **275** lo padecen. Calcular un IC de nivel aproximado **95%** para la diferencia entre la proporción real de hombres y mujeres que padecen el trastorno sanguíneo.

16. Supongamos que se tienen dos variedades A y B de trigo.

    (a) Se dispone de $\boldsymbol{n = 8}$ parcelas de cultivo, cada una dividida en dos porciones similares. En cada porción se cultiva una variedad. Supongamos además que la diferencia entre los rendimientos de las dos variedades cultivadas en una misma parcela es una variable aleatoria normal. Se observan los siguientes resultados:

	Parcela							
Variedad	1	2	3	4	5	6	7	8
A	5.8	8.3	6.1	7.5	9.0	4.5	7.1	7.9
B	12.3	7.5	9.9	11.4	11.3	14.6	10.7	8.1

    Hallar un IC de nivel **0.95** para la media de la diferencia de rendimientos entre A y B.

    (b) Supongamos ahora que se dispone de **16** parcelas, **8** de ellas para cultivarlas con A y las otras **8** para cultivarlas con B. Suponiendo normalidad, hallar un IC de nivel **0.95** para la diferencia de los rendimientos medios de A y B, con los mismos datos que la tabla dada en (a). Realice previamente un boxplot para cada variedad y determine si es razonable la suposición de igualdad de varianzas entre los rendimientos de las dos variedades.

(c) ¿El IC para la media de las diferencias es el mismo que el IC para la diferencia de las medias?

Optativo: Si se sabe que la proporción $\boldsymbol{p}$ de familias con uno o más miembros que practiquen tenis es a lo sumo $\mathbf{1/4}$, mostrar que si se considera una muestra aleatoria de tamaño $\boldsymbol{n}$ y definiendo:

$$\boldsymbol{X_i} = \begin{cases} \mathbf{1} & \text{si} \quad \text{al menos un miembro de la flía i practica tenis} \\ \mathbf{0} & \qquad\ \text{en caso contrario} \end{cases} \qquad\qquad (\mathbf{1} \leq \boldsymbol{i} \leq \boldsymbol{n})$$

entonces un IC de nivel asintótico $\mathbf{1} - \boldsymbol{\alpha}$ para $\boldsymbol{p}$ está dado por:

$$\left[\bar{\boldsymbol{X}} - \frac{\sqrt{\mathbf{3}}\, \boldsymbol{z_{\alpha/2}}}{\mathbf{4}\,\sqrt{\boldsymbol{n}}} \, , \, \bar{\boldsymbol{X}} + \frac{\sqrt{\mathbf{3}}\, \boldsymbol{z_{\alpha/2}}}{\mathbf{4}\,\sqrt{\boldsymbol{n}}} \right]$$

Trabajo práctico 9: Tests de Hipótesis

1. Denotemos por $\boldsymbol{p}$ la proporción de los votantes de cierto municipio que están a favor del candidato A sobre el candidato B como postulante a intendente. Considerar las hipótesis:

$$\boldsymbol{H_o : p = 0.5} \quad \text{vs} \quad \boldsymbol{H_1 : p \neq 0.5}$$

Para una muestra aleatoria de **25** votantes, sea $\boldsymbol{X}$ la cantidad que están a favor de A.

   (a) ¿Cuál es la distribución del estadístico de prueba $\boldsymbol{X}$ cuando $\boldsymbol{H_o}$ es verdadera?

   (b) En el contexto de la situación de este problema, describa los errores de tipo I y de tipo II.

   (c) Averiguar los niveles de los tests correspondientes a las siguientes regiones críticas y calcular luego, para cada uno de ellos, las probabilidades de error de tipo II cuando $\boldsymbol{p = 0.3}$ y cuando $\boldsymbol{p = 0.7}$ ¿Cuál de las regiones críticas considera que es más adecuada?

$$\boldsymbol{R_1 = \{x : x \leq 7 \quad \vee \quad x \geq 18\}}$$
$$\boldsymbol{R_2 = \{x : x \leq 8\}}$$
$$\boldsymbol{R_3 = \{x : x \geq 17\}}$$

   (d) Utilizando la región que Ud. seleccionó, ¿qué concluye si **6** de los **25** entrevistados favoreció a A?

2. Se pregunta a una muestra aleatoria de **400** votantes en cierta ciudad si están a favor de un impuesto adicional del **4%** sobre la venta de gasolina, para proporcionar fondos que se necesitan con urgencia para reparación de calles. Si más de **220** pero menos de **260** favorecen el impuesto, concluiremos que el **60%** de los votantes lo apoyan.

   (a) Encontrar la probabilidad de cometer un error de tipo I si realmente el **60%** de los votantes está a favor del impuesto.

   (b) ¿Cuál es la probabilidad de cometer un error de tipo II al utilizar esta regla de decisión si en realidad hay un **48%** de los votantes que está a favor del impuesto?

   (c) Mostrar que si se toma la nueva región de aceptación como $\boldsymbol{214 \leq X \leq 266}$, donde $\boldsymbol{X}$ es la cantidad en la muestra a favor del impuesto, el nivel del test disminuye pero a costa de disminuir la potencia en (b).

3. En un informe de investigación de R.H.Weindruch de la Escuela de Medicina de la UCLA, se afirma que los ratones con una vida promedio de **32** meses vivirán hasta alrededor de **40** meses de edad cuando **40%** de las calorías en su comida es reemplazada por vitaminas y proteínas ¿Hay alguna evidencia significativa al nivel **0.025** para inclinarse a creer que $\boldsymbol{\mu < 40}$, si de **64** ratones que se someten a esta dieta resulta un promedio de **38** meses de sobrevivencia con un desvío standard de **5.8** meses?

4. Se desea que una mezcla de cemento Portland tenga una resistencia a la compresión mayor que **1300** KN/m$^{\boldsymbol{2}}$. La mezcla no se utilizará a menos que una evidencia experimental indique de manera concluyente que se ha satisfecho la especificación de resistencia. Se mide la tensión de **20** muestras y se obtiene una resitencia promedio de **1329.2** KN/m$^{\boldsymbol{2}}$. Suponiendo normalidad y $\boldsymbol{\sigma = 60}$

   (a) ¿Cuáles son las hipótesis nula y alternativa adecuadas? ¿En qué estadístico basaría un test para estas hipótesis? ¿Cuál es la distribución nula del estadístico?

   (b) Si se utiliza como región de rechazo $\boldsymbol{\bar{x} \geq 1331.26}$, ¿cuál es la probabilidad de un error de tipo I?

(a) (c) ¿Cuál es la distribución del estadístico del test cuando $\boldsymbol{\mu} = \mathbf{1350}$? Utilizando la región de rechazo dada en (b), ¿cuál es la probabilidad de que la mezcla sea considerada no satisfactoria cuando de hecho $\boldsymbol{\mu} = \mathbf{1350}$?

(d) Si se quiere estar seguro de que la probabilidad de un error de tipo I sea **0.05**, ¿qué región crítica utilizaría? ¿Debería o no utilizarse la mezcla a este nivel de significación?

(e) Utilizando el mismo nivel de significación que en (d), ¿cuál es la probabilidad de decidir utilizar la mezcla si $\boldsymbol{\mu} = \mathbf{1325}$?

(f) Sea $\boldsymbol{Z} = \sqrt{\mathbf{20}}\ \frac{\bar{\boldsymbol{X}} - \mathbf{1300}}{\mathbf{60}}$ el estadístico estandarizado ¿Cómo escribiría la misma región de rechazo de (b) a partir de valores de $\boldsymbol{Z}$?

5. Se ha determinado el punto de fusión de cada ua de las **16** muestras de cierta marca de aceite vegetal hidrogenado, resultando $\bar{\boldsymbol{x}} = \mathbf{94.32}$ Supongamos que la distribución del punto de fusión es normal con $\boldsymbol{\sigma} = \mathbf{1.20}$

   (a) Pruebe $\boldsymbol{H_o} : \boldsymbol{\mu} = \mathbf{95}$ vs $\boldsymbol{H_1} : \boldsymbol{\mu} \neq \mathbf{95}$ a nivel **0.01**

   (b) Calcular la probabilidad de error de tipo II cuando $\boldsymbol{\mu} = \mathbf{94}$ trabajando al mismo nivel que en (a).

   (c) Hallar el mínimo tamaño muestral $\boldsymbol{n}$ de modo que la probabilidad del error de tipo II descrito en (b) sea **0.1**, trabajando al mismo nivel que en (a).

6. Se seleccionó una muestra de **12** detectores de radón de cierto tipo y cada una se expuso a **100** pCi/L de radón. Las lecturas resultantes fueron:

105.6	**90.9**	**91.2**	**96.9**	**96.5**	**91.3**
100.1	**105.0**	**99.6**	**107.7**	**103.3**	**92.4**

   (a) ¿Sugiere esta infromación que la lectura media difiere de **100**? Establecer y testear las hipótesis que considere apropiadas, trabajando a nivel **0.05**

   (b) Supongamos que antes del experimento se supuso $\boldsymbol{\sigma} = \mathbf{7.5}$ ¿Cuántas determinaciones hubieran sido apropiadas para obtener una potencia de **0.90** ante la alternativa $\boldsymbol{\mu} = \mathbf{95}$?

7. Sobre Ud. recae la responsabilidad de autorizar la puesta a la venta de cierto jugo comercial cuyos envases estipulan un contenido mínimo de **10** litros. Para decidir acerca de la veracidad de tal afirmación Ud. toma una muestra de tamaño **10** obteniendo los siguientes contenidos (en litros):

10.2	9.7	10.1	10.3	10.1	9.8	9.9	10.4	10.3	9.8

   (a) Si Ud. se pone del lado del consumidor, ¿qué hipótesis nula versus qué hipótesis alternativa testearía?

   (b) Para las hipótesis planteadas en (a) y suponiendo normalidad, ¿son significativos al **1%** los datos muestrales? ¿Qué decisión toma en base a la muestra?

8. El mínimo de consultas recibidas por un estadístico durante una semana de **5** días de trabajo, tiene una distribución de Poisson.

   (a) Si el mínimo de consultas efectuadas en **36** semanas es **160**, ¿se puede afirmar que el verdadero promedio semanal no excede a **4** ? (Halle el $\boldsymbol{p}$-valor)

   (b) Calcular la función de potencia del test (Utilizar aproximación normal).

9. Se lleva a cabo un estudio para ver si el aumento de la concentración del sustrato tiene un efecto apreciable sobre la velocidad de una reacción química. Con una concentración de **1.5** moles por litro la reacción se realizó **15** veces, dando una velocidad promedio de **7.5** micromoles cada

30 minutos y un desvío standard de **1.5** Con un concentración de **2.0** moles por litro, se realizan **12** reacciones, que dan una velocidad promedio de **8.8** micromoles cada **30** minutos, con un desvío standard de **1.2** Suponiendo que ambas poblaciones se distribuyen en forma aproximadamente normal con varianzas iguales, ¿hay evidencia significativa ($\boldsymbol{\alpha} = \mathbf{0.01}$) para creer que el aumento en la concentración del sustrato ocasiona un incremento en la velocidad media por más de **0.5** micromoles cada **30** minutos?

10. En un estudio se compararon los niveles de ácido ascórbico en plasma en mujeres embarazadas fumadoras contra no fumadoras. Se seleccionaron para el estudio **32** mujeres en los últimos tres meses de embarazo, libres de padecimientos importantes y con edades entre **15** y **32** años. Antes de tomar las muestras de **20** ml de sangre, se pidió a las participantes concurrir al laboratorio en ayunas, no consumir complementos vitamínicos y evitar comidas con alto contenido de ácido ascórbico. De las muestras de sangre se determinaron los siguientes valores de ácido ascórbico en miligramos cada **100** ml

Fumadoras:	0.97	0.72	1.00	0.81	0.62	1.32	1.24	0.99	0.90	0.74	0.88	0.94
	1.16	0.86	0.85	0.58	0.57	0.64	0.98	1.09	0.92	0.78	1.24	1.18
No fumadoras:	0.48	0.71	0.98	0.68	1.18	1.36	0.78	1.64				

Suponiendo normalidad de ambas poblaciones:

(a) Realice boxplots comparativos de las dos poblaciones. ¿Le parece razonable suponer igualdad de varianzas?

(b) Realizar un test para igualdad de varianzas.

(c) ¿Son concluyentes los datos para decidir que existe una diferencia significativa entre los niveles de ácido ascórbico entre fumadoras y no fumadoras? Determine el $\boldsymbol{p}$-valor.

11. (*Muestras apareadas*) En un estudio se examinó la influencia del fármaco *succinylcholine* sobre los niveles de circulación de andrógenos en la sangre. Se obtuvieron muestras de sengre de venados inmediatamente luego de una inyección intramuscular de la droga. Luego de **30** minutos se extrajeron nuevas muestras de sangre de los venados. Los niveles de andrógenos al momento de la captura y **30** minutos después, medidos en nanogramos por mililitro fueron los siguientes:

Venado	1	2	3	4	5	6	7	
Nivel antes	**2.76**	**5.18**	**2.68**	**3.05**	**4.10**	**7.05**	**6.60**	
Nivel después	**7.02**	**3.10**	**5.44**	**3.99**	**5.21**	**10.26**	**13.91**	
Venado	8	9	10	11	12	13	14	15
Nivel antes	**4.79**	**7.39**	**7.30**	**11.78**	**3.90**	**26.00**	**67.48**	**17.04**
Nivel después	**18.53**	**7.91**	**4.85**	**11.10**	**3.74**	**94.03**	**94.03**	**41.70**

Suponiendo que los niveles de andrógenos en sangre en el momento de inyectar y pasados los **30** minutos, se distribuyen normalmente. Probar a un nivel de significación **0.05** si los niveles de andrógenos aumentan luego de **30** minutos de inyectada la droga.

12. Se comparan dos tipos de instrumentos (A y B) para medir la cantidad de monóxido de azufre en la atmósfera en un experimento acerca de la contaminación del aire. Se desea determinar si ambos tipos de instrumentos conducen a mediciones con la misma precisión (variabilidad). Se registraron las lecturas siguientes para los dos tipos de instrumentos:

A	**0.86**	**0.82**	**0.75**	**0.61**	**0.89**	**0.64**	**0.81**	**0.68**	**0.65**
B	**0.87**	**0.74**	**0.63**	**0.55**	**0.76**	**0.70**	**0.69**	**0.57**	**0.53**

Suponiendo que las poblaciones se distribuyen normalmente, plantear y testear las hipótesis de interés en este caso. Determine el $\boldsymbol{p}$-valor del test considerado.

Trabajo práctico 10
Estimadores de mínimos cuadrados. Regresión lineal simple

1. Hallar los EMC de $\alpha_1 , \cdots , \alpha_I$ para el siguientes modelo:

$$Y_{ij} = \alpha_i + \epsilon_i \quad (1 \leq j \leq J \,,\, 1 \leq i \leq I)$$

2. En un experimento se estudió la relación entre los años (x) y el número de palabras del vocabulario (y) para una muestra de 10 niños. Los datos son los siguientes:

x	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0
y	100	250	400	850	1210	1530	1840	2060	2300	2500

   (a) Realizar un diagrama de dispersión de los datos.
   (b) Hallar la recta de mínimos cuadrados y graficarla en el scatter plot de (a)
   (c) Estimar el verdadero promedio de palabras para todos los niños de 4 años en la población muestreada.
   (d) Interpretar el significado de $\hat{\beta}_1$ en el contexto de este problema.
   (e) ¿Le parece razonable estimar con la recta ajustada en (b) el promedio poblacional de la cantidad de palabras para un niño de 1 año?
   (f) ¿Es razonable que y esté linealmente relacionada con x para valores de x entre 1.5 y 21?

3. Los datos siguientes corresponden a mediciones experimentales de la presión de gas extraído (x) y el tiempo de extracción en minutos (y)

x	40	130	155	160	260	275	325	370	420	480
y	2.5	3.0	3.1	3.3	3.7	4.1	4.3	4.8	5.0	5.4

   (a) Estimar σ y el desvío standard de $\hat{\beta}_1$
   (b) Los investigadores hacían la suposición siguiente: $\beta_1 = 0.0060$ ¿Está Ud. de acuerdo con esta suposición? Usar $\alpha = 0.10$
   (c) Hallar un IC de nivel 0.95 para la pendiente de la verdadera recta de regresión.
   (d) Si la presión de gas extraído es de 300, ¿Ud. cree que el tiempo esperado de extracción es a lo sumo de 4 minutos? (Usar $\alpha = 0.01$). Calcular el p-valor.
   (e) Hallar un intervalo de predicción de nivel 0.90 para un futuro tiempo de extracción a una presión de 200

4. Utilizando los datos del ejercicio 2

   (a) Supongamos que al cumplir un año más, el vocabulario esperado aumenta en no más de 500 palabras (para niños cuyas edades se encuentran dentro del rango estudiado) ¿Está Ud. de acuerdo con esta suposición? (Utilizar $\alpha = 0.01$)
   (b) Interpretar en palabras la cantidad $\mu_{Y.3} = \beta_o + 3\beta_1$ Hallar un estimador puntual $\hat{\mu}_{Y.3}$ de $\mu_{Y.3}$
   (c) Determinar el desvío standard de $\hat{\mu}_{Y.3}$
   (d) Hallar un IC de nivel 0.95 para $\mu_{Y.3}$
   (e) ¿Cuál IC sería más corto, uno para $\mu_{Y.3}$ o uno para $\mu_{Y.5}$? Responder sin calcular los intervalos en cuestión.

(f) Testear $\boldsymbol{H_o : \mu_{Y.3} = 1000}$ vs $\boldsymbol{H_1 : \mu_{Y.3} < 1000}$ Espresar con palabras lo que se está testeando.

5. En un establecimiento industrial se ha descubierto que la humedad relativa ($\boldsymbol{x}$) de la metria prima ejerce cierto efecto sobre la densidad ($\boldsymbol{y}$) del producto terminado. Para precisar la forma exacta de esta regresión se controló la humedad y se midió la densidad para una muestra aleatoria de **23** observaciones:

x	**1.3**	**1.3**	**2.0**	**2.0**	**2.7**	**3.3**	**3.7**	**3.7**	**4.0**	**4.0**	**4.0**	
y	**2.3**	**1.8**	**2.8**	**1.5**	**2.2**	**1.8**	**3.7**	**1.7**	**2.8**	**2.9**	**2.2**	
x	**4.7**	**4.7**	**4.7**	**5.0**	**5.3**	**5.3**	**5.3**	**5.7**	**5.7**	**6.0**	**6.3**	**6.7**
y	**5.4**	**3.2**	**1.9**	**1.8**	**3.5**	**2.8**	**2.1**	**3.4**	**3.2**	**3.0**	**3.0**	**3.9**

(a) Ajustar a los datos un modelo de regresión lineal simple.

(b) Estudiar los residuos mediante:

- Un diagrama de residuos vs predichos
- Un diagrama de residuos vs $\boldsymbol{x_i}$

6. En un laboratorio se quiere estudiar la influencia de un antibiótico sobre el crecimiento de un cultivo microbiano en tubos de ensayo. Para esto se asigna a cada tubo un columen constante de cultivo y se le adiciona una cantidad de antibiótico de las tres disponibles (**0.2 ; 0, 4 ; 0.6** unidades), efectuándose seis verificaciones para cada nivel de dosis. Después de cuatro horas de cultivo se mide la densidad óptica de cada tubo mediante un fotómetro eléctrico. En la tabla siguiente ($\boldsymbol{y}$ es la lectura sobre el galvanómetro del aparato) se dan los resultados obtenidos:

x	**y**					
0.2	**13**	**15**	**10**	**17**	**10**	**12**
0.4	**60**	**62**	**60**	**60**	**60**	**65**
0.6	**94**	**93**	**92**	**97**	**90**	**92**

(a) Ajustar un modelo lineal: $\boldsymbol{Y = \beta_o + \beta_1 x + \epsilon}$

(b) Hallar un IC de nivel **0.95** para $\boldsymbol{\beta_1}$

(c) Analizar los resuduos del ajuste.

(d) Ajustar un nuevo modelo: $\boldsymbol{Y = \alpha + \beta \ln x + \epsilon^*}$

(e) Analizar los residuos del nuevo ajuste.

7. En un experimento se registró el peso ($\boldsymbol{x}$ en gramos) y la edad ($\boldsymbol{y}$ en días) de embriones de pollo.

y	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
x	**0.029**	**0.052**	**0.079**	**0.125**	**0.181**	**0.261**	**0.425**	**0.738**	**1.130**	**1.882**	**2.812**

(a) Realizar un diagrama de dispersión de los datos ¿Considera apropiado un modelo de regresión lineal simple?

(b) Ajustar el modelo: $\boldsymbol{Y = \alpha\, e^{\beta\, x}\, \epsilon}$

(c) Estudiar los residuos para detectar si el modelo exponencial propuesto es razonable.

(d) Testear $\boldsymbol{H_o : \beta = 0.5}$ vs $\boldsymbol{H_1 : \beta \neq 0.5}$