RESOLUCIÓN DE LA AUTOEVALUACIÓN

1. Una variable aleatoria puede tomar los siguientes valores: {2,3,7,8,10}. Sabiendo que p(2)=p(3)=p(7)=1/4 y p(8)=p(10)=1/8, calcule los valores de asimetría y curtosis, indicando la forma de la distribución.

$$\gamma_1 = \frac{1}{\sigma^3} E(X - E(X))^3 = \frac{\mu_3}{\sigma^3}$$

$$E(X) = \sum_{i=1}^{n} x_i p(x_i)$$

$$E(X) = 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} + 7 \cdot \frac{1}{4} + 8 \cdot \frac{1}{8} + 10 \cdot \frac{1}{8} = 5,25$$

$$\mu_3 = E(X - E(X))^3 = \sum_{i=1}^{n} (x_i - E(X))^3 p(x_i) =$$

$$\mu_3 = (2-5,25)^3 \cdot \frac{1}{4} + (3-5,25)^3 \cdot \frac{1}{4} + (7-5,25)^3 \cdot \frac{1}{4} + (8-5,25)^3 \cdot \frac{1}{8} + (10-5,25)^3 \cdot \frac{1}{8} = 5,90625$$

$$\mu_2 = \sigma^2 = E(X - E(X))^2 = \sum_{i=1}^{n} (x_i - E(X))^2 p(x_i) =$$

$$\sigma^2 = (2-5,25)^2 \cdot \frac{1}{4} + (3-5,25)^2 \cdot \frac{1}{4} + (7-5,25)^2 \cdot \frac{1}{4} + (8-5,25)^2 \cdot \frac{1}{8} + (10-5,25)^2 \cdot \frac{1}{8} = 8,4375$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{8,4375}$$

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{5,90625}{\left(\sqrt{8,4375}\right)^3} = \boxed{0,24099}$$

El valor es muy próximo a 0, por lo que puede considerarse que la distribución es simétrica, aunque presente una ligera asimetría positiva.

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$$

$$\mu_4 = E(X - E(X))^4 = \sum_{i=1}^{n} (x_i - E(X))^4 \, p(x_i) =$$

$$\mu_4 = (2-5,25)^4 \cdot \frac{1}{4} + (3-5,25)^4 \cdot \frac{1}{4} + (7-5,25)^4 \cdot \frac{1}{4} + (8-5,25)^4 \cdot \frac{1}{8} + (10-5,25)^4 \cdot \frac{1}{8} = 107,4258$$

$$\mu_2 = \sigma^2 = E(X - E(X))^2 = \sum_{i=1}^{n} (x_i - E(X))^2 \, p(x_i) =$$

$$\sigma^2 = (2-5.25)^2 \cdot \frac{1}{4} + (3-5,25)^2 \cdot \frac{1}{4} + (7-5,25)^2 \cdot \frac{1}{4} + (8-5,25)^2 \cdot \frac{1}{8} + (10-5,25)^2 \cdot \frac{1}{8} = 8,4375$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{8,4375}$$

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3 = \frac{107,4258}{\left(\sqrt{8,4375}\right)^4} - 3 = \boxed{-1,4910}$$

El valor de la curtosis es negativo, por lo que puede considerarse que la distribución es platicúrtica.

2. Sea la esperanza de vida del radioisótopo una variable aleatoria con la siguiente función de densidad:

$$\mathrm{f}(x)=\begin{cases} ke^{-2x} & x>0 \\ 0 & x\leq 0 \end{cases}$$

   a) Calcule el valor de k para que $\mathrm{f}(x)$ sea una función de densidad.

   b) Obtenga la función de distribución de la variable aleatoria continua.

   c) Calcule la función característica de la distribución.

   d) Calcule la función generadora de momentos de la distribución.

   e) Calcule la media y la varianza de la distribución utilizando la función generadora de momentos.

a) Para que $\mathrm{f}(x)$ sea una función de densidad se deben satisfacer las siguientes dos condiciones:

1) $\mathrm{f}(x)\geq 0 \quad \forall x\in\mathbb{R}$

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

Para satisfacer la primera condición: $k\geq 0$

Segunda condición:

$$\int_{-\infty}^{+\infty}\mathrm{f}(x)dx=1;\quad \int_{-\infty}^{+\infty}\mathrm{f}(x)dx=\int_{-\infty}^{0}0dx+\int_{0}^{+\infty}ke^{-2x}dx=\lim_{t\to+\infty}k\int_{0}^{t}e^{-2x}dx;$$

$$\lim_{t\to+\infty}k\left(-\frac{1}{2}e^{-2x}\right)\Big|_{0}^{t}=k\cdot\left(\frac{1}{2}\right)=\frac{k}{2}\Rightarrow\frac{k}{2}=1;\quad \boxed{k=2}$$

Por lo tanto,

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

b) Función de distribución:

$x\leq 0$:

$$\int_{-\infty}^{x}\mathrm{f}(t)dt=\int_{-\infty}^{x}0dt=0\ dx$$

$x>0$:

$$\int_{-\infty}^{0}\mathrm{f}(t)dt+\int_{0}^{x}\mathrm{f}(t)dt=\int_{-\infty}^{0}0dt+\int_{0}^{x}2e^{-2t}dt=0+\left(-e^{-2t}\right)\Big|_{0}^{x}=-e^{-2x}+1=1-e^{-2x}$$

Con lo cual,

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

c)

$$\Psi(t)=E\left(e^{itx}\right)=\int_{-\infty}^{+\infty} e^{itx}\,\mathrm{f}(x)dx=\int_{-\infty}^{0} e^{itx}\,0dx+\int_{0}^{+\infty} e^{itx}\cdot 2e^{-2x}dx=2\lim_{a\to+\infty}\int_{0}^{a} e^{x(it-2)}dx=2\lim_{a\to+\infty}\left.\frac{e^{-x(2-it)}}{it-2}\right|_{0}^{a}=\boxed{\frac{2}{2-it}}$$

d)

$$\alpha(w)=E\left(e^{wx}\right)=\int_{-\infty}^{+\infty} e^{wx}\,\mathrm{f}(x)dx=\int_{-\infty}^{0} e^{wx}\,0dx+\int_{0}^{+\infty} e^{wx}\cdot 2e^{-2x}dx=2\lim_{a\to+\infty}\int_{0}^{a} e^{x(w-2)}dx=2\lim_{a\to+\infty}\left.\frac{e^{-x(2-w)}}{w-2}\right|_{0}^{a}=\boxed{\frac{2}{2-w}}$$

e)

Como la media es el momento de primer orden:

$$E(X)=\left.\frac{dE(e^{wx})}{dw}\right|_{w=0}=\left.\frac{2}{(2-w)^2}\right|_{w=0}=\boxed{\frac{1}{2}}$$

La varianza es el momento de segundo orden centrado en la media, por lo que:

$$\sigma^2=\alpha_2-\left(E(X)\right)^2=\left.\frac{d^2E(e^{wx})}{dw^2}\right|_{w=0}-\left(\frac{1}{2}\right)^2=\left.\frac{4}{(2-w)^3}\right|_{w=0}-\left(\frac{1}{2}\right)^2=\boxed{\frac{1}{4}}$$

3. Se están realizando diferentes ensayos con lindano en un laboratorio químico suizo en el que se trata de medir la resistencia de un determinado tipo de bacteria al lindano. Se sabe que el 0,5% de estas bacterias no tienen resistencia al lindano. Si en un medio dado hay una muestra de 3000 bacterias, calcule:

   a) La probabilidad de que el número de bacterias que no tienen resistencia al lindano sea mayor que 2 y menor que 8.

   b) La probabilidad de que haya 10 bacterias que no tengan resistencia al lindano.

a) En primer lugar, se han de definir la variable aleatoria y la distribución que sigue.

X: "Número de bacterias sin resistencia al lindano". $X \sim B(n = 3000, p = 0,005)$

En este caso, siendo $n = 3000 > 30$ y $p = 0,005 < 0,1$ la distribución binomial puede aproximarse por una distribución de Poisson.

$X \sim B(n = 3000, p = 0,005) \cong P(n \cdot p = 15)$

Por tanto,

$$P(2 < X < 8) = P(X \leq 7) - P(X \leq 2) = P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7)$$

$$P(2 < X < 8) = \left[e^{-15} \left(\frac{15^3}{3!} + \frac{15^4}{4!} + \frac{15^5}{5!} + \frac{15^6}{6!} + \frac{15^7}{7!} \right) \right] = \boxed{0,0180}$$

b) En particular, la probabilidad de que haya exactamente 10 bacterias que no presenten resistencia al lindano se calcula de la siguiente forma:

$$P(X = 10) = \frac{e^{-15} \cdot 15^{10}}{10!} = \boxed{0,0486}$$

4. Una empresa que busca gente para emplear está seleccionando candidatos y candidatas. El técnico de recursos humanos realiza un estudio a las y los candidatos. La media de la puntuación obtenida en el estudio es de 60 puntos y la desviación típica de 10 puntos.

   Para que las y los candidatos pasen a la siguiente fase de la selección es necesario obtener una nota entre 55 y 80 puntos. Sabiendo que la puntuación obtenida por las y los candidatos sigue una distribución normal,

   a) ¿Qué porcentaje de candidaturas pasará a la siguiente fase de selección?

   b) ¿Qué porcentaje de candidaturas ha obtenido más de 90 puntos?

   c) ¿Qué porcentaje de candidaturas ha obtenido menos de 50 puntos?

   d) ¿Cuáles son los dos valores centrales de las puntuaciones que tenemos que fijar para que entre ellos esté el 50% de las candidaturas?

En primer lugar, se definirá la variable aleatoria y su distribución.

X: "Puntuación obtenida por la o el candidato en el estudio".
$E(X)=60$ y $\sigma=10$. $X \sim N(60;10)$

a)

$55 \le X \le 80 \rightarrow$ La candidatura pasará a la siguiente fase de selección

$$P(55 \le X \le 80) = P(X \le 80) - P(X \le 55)$$

$$Z = \frac{x-60}{10}$$

$$P(55 \le \mathrm{X} \le 80) = P(Z \le \frac{80-60}{10}) - P(Z \le \frac{55-60}{10}) = P(Z \le 2) - P(Z \le -0,5)$$

$P(Z \le 2) = 0,9772$ Este valor se encuentra en la tabla de distribución normal estandarizada.

$P\left(Z \le -0,5\right)$ En la tabla no hay valores negativos por lo que se ha de aplicar la simetría.

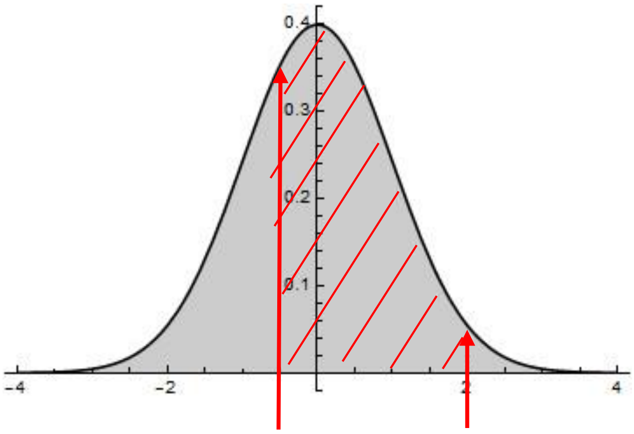
$$P(Z \le -0,5) = P(Z \ge 0,5) = 1 - P(Z \le 0,5)$$

$P(Z \le 0,5) = 0,6915$ Este valor se encuentra en la tabla de distribución normal estandarizada.

$$P(Z \le -0,5) = 1 - P(Z \le 0,5) = 1 - 0,6915 = 0,3085$$

$$P(55 \le X \le 80) = P(Z \le 2) - P(Z \le -0,5) = 0.9772 - 0,3085 = \boxed{0,6687}$$

Por tanto, el 66,87% de las candidaturas pasarán a la siguiente fase.



El área marcada en rojo es la probabilidad pedida.

Distribución normal estandarizada

b)

$$P(X > 90) = 1 - P(X \leq 90) = 1 - P(Z \leq \frac{90-60}{10}) = 1 - P(Z \leq 3)$$

$P(Z \leq 3) = 0,9987$ Este valor se encuentra en la tabla de distribución normal estandarizada.

$$P(X > 90) = 1 - P(Z \leq 3) = 1 - 0,9987 = 0,0013$$

Por lo tanto, el 0,13% de las candidaturas ha obtenido más de 90 puntos.

c)

$$P(X < 50) = P(Z < \frac{50-60}{10}) = P(Z < -1) = P(Z \leq -1)$$

$P(Z \leq -1)$ En la tabla no hay valores negativos por lo que se ha de aplicar la simetría.

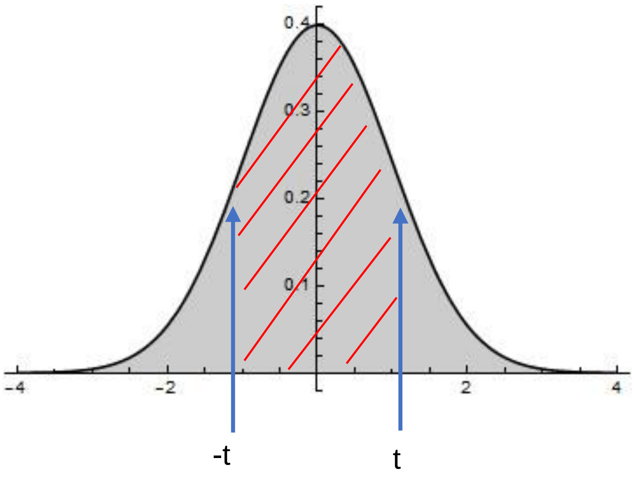
$$P(Z \leq -1) = P(Z \geq 1) = 1 - P(Z \leq 1)$$

$P(Z \leq 1) = 0,8413$ Este valor se encuentra en la tabla de distribución normal estandarizada.

$$P(Z \leq -1) = 1 - P(Z \leq 1) = 1 - 0,8413 = \boxed{0,1587}$$

Por tanto, el 15,87% de las candidaturas ha obtenido menos de 50 puntos.

e)



Distribución normal estandarizada

En este apartado se han de calcular los valores centrales *b* y *a*, o μ-d y μ+d (para utilizar la distribución normal estandarizada después del cambio de variable –t y t), donde cada uno de ellos está a la misma distancia de la media. El requisito que exige el enunciado corresponde con el área marcada en rojo que abarca el 50% de la distribución

Se utilizará la distribución normal estandarizada.

$$P(b \le X \le a) = P(\mu - d \le X \le \mu + d) = P(X \le a) - P(X \le b) = P(X \le \mu + d) - P(X \le \mu - d) = 0,5$$

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

Se empleará la simetría una y otra vez:

$$P(b \le \mathrm{X} \le a) = P(X \le \mu + d) - P(X \le \mu - d) = P(Z \le t) - P(Z \le -t) =$$

$$P(b \le \mathrm{X} \le a) = P(X \le \mu + d) - P(X \le \mu - d) = P(Z \le \frac{\mu + d - \mu}{\sigma}) - P(Z \le \frac{\mu - d - \mu}{\sigma}) =$$

$$P(b \le \mathrm{X} \le a) = P(Z \le \frac{d}{\sigma}) - P(Z \le \frac{-d}{\sigma}) = P(Z \le \frac{d}{10}) - P(Z \le \frac{-d}{10}) = P(Z \le \frac{d}{10}) - P(Z \ge \frac{d}{10})$$

$$P(b \le \mathrm{X} \le a) = P(Z \le \frac{d}{10}) - \left[1 - P(Z \le \frac{d}{10})\right] = -1 + 2P(Z \le \frac{d}{10}) = 0,5 \Rightarrow P(Z \le \frac{d}{10}) = 0,75$$

En la tabla de distribución normal estandarizada se puede observar para qué valor $\frac{d}{10}$ la probabilidad es de 0,75.

$$P\left(Z \le \frac{d}{10}\right) = 0,75 \Rightarrow \frac{d}{10} = 0,68 \Rightarrow d = 6,8$$

$b = \mu - d \Rightarrow b = 53,2$ puntos

$a = \mu + d \Rightarrow a = 66,8$ puntos

Por tanto, los valores centrales de las puntuaciones para que entre ellos esté el 50% de las candidaturas, b y a, son 53,2 y 66,8 puntos respectivamente.

5. Un sistema electrónico complejo requiere un componente específico en su montaje. Estos componentes se compran en lotes grandes y se testean secuencialmente hasta encontrar el primer componente defectuoso. El porcentaje de componentes defectuosos es del 5%. Determine usando R Studio:

   a) La probabilidad de que para encontrar el primer componente defectuoso haya que testear como mucho 4 componentes.

   b) La probabilidad de que para encontrar el primer componente defectuoso haya que testear al menos 4 y como máximo 10 componentes

En primer lugar, se definirán la variable aleatoria y su distribución.

X: "Número de componentes testeados hasta encontrar el primer componente defectuoso (sin tener en cuenta el componente defectuoso)".
$X \sim G(p = 0.05)$

a) $P(X \leq 3)$

$> \text{pgeom}(3, 0.05)$

$[1]\ 0.1854938$

$P(X \leq 3) = \boxed{0.1854938}$

b) $P(3 \leq X \leq 9)$

$> \text{pgeom}(9, 0.05) - \text{pgeom}(2, 0.05)$

$[1]\ 0.2586381$

$P(4 \leq X \leq 10) = \boxed{0.2586381}$