Ejercicios Resueltos TP4

Índice

1. Ejercicio 6

Enunciado

La duración en horas X de cierto dispositivo electrónico es una variable aleatoria continua con función densidad de probabilidad:

$$f_X(x) = \begin{cases} \dfrac{100}{x^2} & \text{si } x > 100 \\ \\ 0 & \text{si } x \leq 100 \end{cases}$$

a) ¿Cuál es la probabilidad de que un dispositivo dure menos de 200 hs si se sabe que el dispositivo aún funciona después de 150 hs de servicio?

b) ¿Cuál es la probabilidad de que si se instalan 3 de tales dispositivos en un sistema 2 fallen antes de las 150 horas de servicio y el restante después de las 150 hs?

c) ¿Cuál es el número máximo de dispositivos que se pueden poner en un sistema para que la probabilidad de que después de 150 horas de servicio todos ellos funcionen sea 0.4?

Resolución

Variable aleatoria

La variable aleatoria viene ya definida en el enunciado por:

$$X = \text{Duración (en hs.) de un dispositivo electrónico seleccionado al azar.}$$

Eventos y Datos

No se visualiza ningún otro evento de importancia en el enunciado que aquellos relativos a la distribución de la variable X. Sobre las probabilidades de X, tenemos la función de densidad $f_X(x)$.

1.1. Item a

Nos piden:

$$P(X < 200 | X > 150) = \frac{P(X < 200 \cap X > 150)}{P(X > 150)} = \frac{P(150 < X < 200)}{P(X > 150)}$$

Para calcular el denominador, debemos integrar la densidad sobre los valores de x considerados en la región, es decir, para todo $x > 150$:

$$\int_{150}^{+\infty} f_X(x)\, dx$$

Esta integral es impropia y puede causar dificultades, por lo que apelaremos al recurso del complemento:

$$P(X > 150) = 1 - P(X \leq 150) = 1 - F_X(150) = 1 - \int_{-\infty}^{150} f_X(x)\, dx$$

Como f_X es una función partida, dividimos la integral en dos partes, según los cambios de expresiones:

$$\int_{-\infty}^{150} f_X(x)\, dx = \int_{-\infty}^{100} \underbrace{f_X(x)}_{0}\, dx + \int_{100}^{150} \underbrace{f_X(x)}_{\frac{100}{x^2}}\, dx = \int_{100}^{150} 100x^{-2}\, dx = \left(-100x^{-1}\right)\Big|_{100}^{150} = -\frac{100}{150} - \left(-\frac{100}{100}\right) = \frac{1}{3}$$

Notemos que esta última integral coincide con $F_X(150)$, entonces podremos utilizar $F_X(150) = \dfrac{1}{3}$ como dato más adelante.

Ahora calcularemos el numerador, atendiendo algunas sutilezas, para enfatizar algunos conceptos. Sabiendo que la variable X es continua, sabemos que para todo valor $x \in \mathbb{R}$, $P(X = x) = 0$. Entonces, denotando M.E. como eventos mutuamente excluyentes:

$$P(150 < X < 200) = P(150 < X < 200) + \overbrace{P(X = 200)}^{0} \stackrel{\text{M.E.}}{=} P(150 < X < 200 \cup X = 200) = P(150 < X \leq 200)$$

Si bien no parece tener importancia desde lo operativo, esta sutileza esconde algunos conceptos importantes para comprender algunos objetivos de la materia. Notar que $150 < X < 200$ y $150 < X \leq 200$ son **distintos como eventos** (podemos pensarlos como subconjuntos del espacio muestral). Sin embargo, cuando se calculan las probabilidades de ambos, estos **números son iguales**. Se tiende a decir que "da lo mismo". Sin embargo, vale aclarar que se refiere a las **probabilidades** y no a los **eventos**.

Por otro lado, también por ser mutuamente excluyentes, se puede obtener la siguiente igualdad:

$$P(X \leq 200) = P(X \leq 150 \ \cup \ 150 < X \leq 200) \stackrel{\text{M.E.}}{=} P(X \leq 150) + P(150 < X \leq 200)$$

Luego, despejando $P(150 < X \leq 200)$:

$$P(150 < X \leq 200) = P(X \leq 200) - P(X \leq 150) = F_X(200) - F_X(150)$$

Ya sabemos que $F_X(150) = \dfrac{1}{3}$, sólo resta determinar $F_X(200)$:

$$F_X(200) = \int_{-\infty}^{200} f_X(x)\,dx = \int_{-\infty}^{100} \underbrace{f_X(x)}_{0}\,dx + \int_{100}^{200} \underbrace{f_X(x)}_{\frac{100}{x^2}}\,dx = \int_{100}^{200} 100x^{-2}\,dx = \left(-100x^{-1}\right)\Big|_{100}^{200} = -\frac{100}{200} - \left(-\frac{100}{100}\right) = \frac{1}{2}$$

Por último:

$$P(X < 200 | X > 150) = \frac{F_X(200) - F_X(150)}{1 - F_X(150)} = \frac{\frac{1}{2} - \frac{1}{3}}{1 - \frac{1}{3}} = \frac{\frac{1}{6}}{\frac{2}{3}} = \frac{1}{4}$$

Comentario integrales impropias:

Que en este caso hayamos esquivado la integral impropia no significa que siempre lo podamos hacer. Entonces, resolveremos las siguientes integrales impropias:

$$P(X > 150) = \int_{150}^{+\infty} f_X(x)\,dx = \lim_{M \to +\infty} \int_{150}^{M} \frac{100}{x^2}\,dx = \lim_{M \to +\infty} \left(-100x^{-1}\right)\Big|_{150}^{M} = \lim_{M \to +\infty} \underbrace{-\frac{100}{M}}_{\to 0} - \left(-\frac{100}{150}\right) = \frac{2}{3}$$

$$P(X > 200) = \int_{200}^{+\infty} f_X(x)\,dx = \lim_{M \to +\infty} \int_{200}^{M} \frac{100}{x^2}\,dx = \lim_{M \to +\infty} \left(-100x^{-1}\right)\Big|_{200}^{M} = \lim_{M \to +\infty} \underbrace{-\frac{100}{M}}_{\to 0} - \left(-\frac{100}{200}\right) = \frac{1}{2}$$

Notar además, que estos cálculos añaden consistencia a nuestros resultados, ya que:

- $P(X > 150) = 1 - F_X(150) = 1 - \frac{1}{3} = \frac{2}{3}$
- $P(X > 200) = 1 - F_X(200) = 1 - \frac{1}{2} = \frac{1}{2}$

1.2. Item b

En este ítem, ya que se analiza el funcionamiento de 3 dispositivos, debemos definir una nueva variable aleatoria:

$$Y_3 = \text{Cantidad de dispositivos que fallan antes de las 150 hs en la muestra de 3 dispositivos.}$$

Notar que esta variable es **discreta**, ya que cuenta una cantidad. Al ser una variable discreta, pensemos primero en su recorrido: Puede tener un valor mínimo de 0 y un valor máximo de 3. Es decir, $R_{Y_3} = \{0; 1; 2; 3\}$.

Nos piden $P(Y_3 = 2)$, por lo que nos vendría bien conocer la distribución de Y_3. Notemos que si consideramos que las 3 extracciones son independientes y la probabilidad de extraer un dispositivo de duración menor a 150 hs como un valor constante p en todas las extracciones, tenemos que $Y_3 \sim \text{Bi}(3, p)$.

En principio, no pareciera que esta distribución discreta Y_3 tuviera alguna relación con la distribución continua X definida en el ejercicio. Sin embargo, notemos que X influye sobre la probabilidad de "éxito":

$$p = P(X < 150) = P(X \leq 150) = F_X(150) = \frac{1}{3}$$

Es decir, $Y_3 \sim \text{Bi}(3; \frac{1}{3})$ y por lo tanto:

$$P(Y_3 = 2) = \binom{3}{2}\left(\frac{1}{3}\right)^2 \frac{2}{3} = \frac{2}{9}$$

Observación: Una pregunta frecuente es si se puede definir la siguiente variable:

$$Z_3 = \text{Cantidad de dispositivos que fallan } \textbf{después} \text{ de las 150 hs en la muestra de 3 dispositivos.}$$

Esto es completamente válido, sólo que cambia la distribución ($Z_3 \sim \text{Bi}(3, \frac{2}{3})$) el evento buscado ($Z_3 = 1$). Sin embargo, la probabilidad coincide:

$$P(Z_3 = 1) = \binom{3}{1}\frac{2}{3}\left(\frac{1}{3}\right)^2 = \frac{2}{9}$$

1.3. Item c

En este ítem, tenemos una variable similar a la anterior, sólo que debemos determinar el número de elementos en la muestra n:

$$Y_n = \text{Cantidad de dispositivos que fallan antes de las 150 hs en la muestra de n dispositivos.}$$

De manera similar a la anterior, podemos asumir que $Y_n \sim \text{Bi}(n, \frac{1}{3})$. Además, nos piden determinar n de forma que **todos** los dispositivos funcionen luego de las 150 hs, por lo que **ninguno** puede fallar antes de las 150 hs. Es decir, nos piden buscar n de forma que $P(Y_n = 0) = 0{,}4$.

Vale aclarar lo siguiente, n debe ser **entero**, y esta igualdad puede no cumplirse para valores enteros, por lo que habrá que plantear una desigualdad. En el contexto del problema, la probabilidad deseada representa una situación favorable, ya que se aseguraría que el sistema dura al menos 150 hs. Es decir, nos gustaría que esta probabilidad sea **alta**:

$$P(Y_n = 0) \geq 0{,}4 \Leftrightarrow \underbrace{\binom{n}{0}}_{1}\underbrace{\left(\frac{1}{3}\right)^0}_{1}\left(\frac{2}{3}\right)^n \geq 0{,}4 \Leftrightarrow \ln\left[\left(\frac{2}{3}\right)^n\right] \geq \ln(0{,}4) \Leftrightarrow n \cdot \ln\left(\frac{2}{3}\right) \geq \ln(0{,}4) \overset{\ln(\frac{2}{3})<0}{\Leftrightarrow} n \leq \frac{\ln(0{,}4)}{\ln(\frac{2}{3})} \approx 2{,}2598$$

Es decir, 2 es el número **máximo** de dispositivos para lograr que el sistema funcione con la probabilidad deseada luego de 150 hs. Tiene sentido que este número sea máximo porque mientras más dispositivos se conecten, más improbable es que funcionen **todos** luego de 150 hs.

Notemos además que $P(Y_2 = 0) = (\frac{2}{3})^2 \approx 0{,}444 > 0{,}4$ y $P(Y_3 = 0) = (\frac{2}{3})^3 \approx 0{,}296 < 0{,}4$, verificando así nuestro resultado.

2. Ejercicio 14

Enunciado

La duración de ciertos dispositivos es una variable aleatoria con distribución exponencial de parámetro λ (en 1/hora). La probabilidad de que en una muestra de 5 dispositivos haya por lo menos uno que dure más de 1200 hs. es 0.75.

a) Hallar la constante λ.

b) Suponga que estos dispositivos se conectan en serie de manera tal que al fallar por lo menos uno de ellos falla todo el sistema. Calcular la probabilidad de que el sistema funcione más de 1500 horas.

Resolución

Variables aleatorias

Notemos que en este enunciado se definen **dos** variables aleatorias:

- $X =$ duración de un dispositivo seleccionado al azar.
- $Y_5 =$ Cantidad de dispositivos que duran más de 1200 hs en una muestra de 5.

Eventos y Datos

Sabemos que $X \sim \mathcal{E}(\lambda)$ con λ desconocido. Sin embargo, al ser exponencial, sabemos que $F_X(x) = 1 - e^{-\lambda x}$ si $x > 0$ y $F_X(x) = 0$ en otro caso.

Notemos que otro dato relevante es $P(Y_5 \geq 1) = 0{,}75$. Además, como podemos asumir que el lote es grande y la probabilidad de extraer un dispositivo que dure más de 1200 hs se mantiene constante en cada extracción, se obtiene que $Y_5 \sim \text{Bi}(5; p)$, donde

$$p = P(X > 1200) = 1 - F_X(1200) = 1 - (1 - e^{-\lambda \cdot 1200}) = e^{-\lambda \cdot 1200}$$

2.1. Item a

Considerando el recorrido de Y_5, tenemos que $P(Y_5 \geq 1) = 1 - P(Y_5 = 0)$. Es decir,

$$P(Y_5 \geq 1) = 1 - P(Y_5 = 0) = 0{,}75 \Rightarrow P(Y_5 = 0) = 0{,}25 \Rightarrow \underbrace{\binom{5}{0}}_{1} \underbrace{\left(e^{-\lambda \cdot 1200}\right)^0}_{1} \left(1 - e^{-\lambda \cdot 1200}\right)^5 = 0{,}25$$

A partir de esta expresión podemos despejar λ:

$$\left(1 - e^{-\lambda \cdot 1200}\right)^5 = 0{,}25 \Rightarrow 1 - e^{-\lambda \cdot 1200} = \sqrt[5]{0{,}25} \Rightarrow e^{-\lambda \cdot 1200} = 1 - \sqrt[5]{0{,}25} \Rightarrow -\lambda \cdot 1200 = \ln(1 - \sqrt[5]{0{,}25})$$

Por último,

$$\lambda = -\frac{\ln(1 - \sqrt[5]{0{,}25})}{1200} \approx 0{,}0011819$$

2.2. Item b

Ahora que sabemos λ, tenemos caracterizada la distribución de X. Ahora, tenemos una variable similar a la anterior:

$$Z_5 = \text{ Cantidad de dispositivos que duran más de } \underline{1500 \text{ hs}} \text{ en una muestra de 5.}$$

La única diferencia respecto a Y_5 es el tiempo de duración. Con argumentos similares, $Z_5 \sim \text{Bi}(5, p)$, donde

$$p = P(X > 1500) = e^{-0{,}0011819 \cdot 1500}$$

Entonces, nos piden la probabilidad de que todos los dispositivos funcionen luego de 1500 hs. Es decir,

$$P(Z_5 = 5) = \underbrace{\binom{5}{5}}_{1} \left(e^{-0{,}0011819 \cdot 1500}\right)^5 \underbrace{\left(1 - e^{-0{,}0011819 \cdot 1500}\right)^0}_{1} \approx 0{,}00014135$$

3. Ejercicio 16

Enunciado

Se sabe que la duración X de una pieza responde a la distribución de Weibull cuya función de distribución es: $F_X(x) = 1 - \exp(-(\lambda x)^b)$ para $x > 0$ con b y λ constantes reales positivas. Si x se mide en miles de horas, $b = 2$ y $\lambda = 0.01$:

a) obtener la función de densidad de probabilidad de la duración de la pieza.

b) calcular la duración garantizada G_{90} con un 90 % de confiabilidad. Nota: Se cumple $P(X > G_{90}) = 0{,}90$.

c) Se realizó una prueba con un lote muy grande de piezas y se separaron 1000 con duración superior a 10000 horas; ¿cuántas unidades espera Ud. haya en este lote de 1000 cuyas duraciones totales sean superiores a 20000 horas?

Resolución

Variable aleatoria

La variable aleatoria viene definida en el enunciado:

$$X = \text{ Duración de una pieza elegida al azar en miles de hs.}$$

El hecho de que las unidades estén en miles de horas, nos dice que si para una pieza $X > 10$, entonces esa pieza dura más de 10000 hs.

Datos

Tenemos la distribución acumulada de la variable X. Como se dice que $b = 2$ y $\lambda = 0{,}01$, $F_X(x) = 1 - e^{-(0{,}01x)^2}$ si $x > 0$ y $F_X(x) = 0$ si $x \leq 0$.

3.1. Item a

Para obtener la densidad, podemos derivar la acumulada, por lo tanto, si $x > 0$:

$$f_X(x) = [F_X(x)]' = [1 - e^{-(0{,}01x)^2}]' = -e^{-(0{,}01x)^2} \cdot [-(0{,}01x)^2]' = -e^{-(0{,}01x)^2} \cdot (-2 \cdot 0{,}01x \cdot [0{,}01x]') = 0{,}0002 \cdot x \cdot e^{-(0{,}01x)^2}$$

Por otro lado, como $F_X(x) = 0$ si $x \leq 0$, entonces la derivada también es nula:

$$f_X(x) = \begin{cases} 0{,}0002 \cdot x \cdot e^{-(0{,}01x)^2} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

3.2. Item b

Nos piden hallar un valor G_{90} de forma que $P(X > G_{90}) = 0{,}90$. Para eso, utilizaremos la función de distribución:

$$P(X > G_{90}) = 0{,}9 \Rightarrow 1 - F_X(G_{90}) = 0{,}9 \Rightarrow 1 - (1 - e^{-(0{,}01G_{90})^2}) = 0{,}9 \Rightarrow e^{-(0{,}01G_{90})^2} = 0{,}9$$

Por lo tanto, para determinar G_{90}, utilizamos el logaritmo:

$$e^{-(0{,}01G_{90})^2} = 0{,}9 \Rightarrow -(0{,}01G_{90})^2 = \ln(0{,}9) \Rightarrow 0{,}01G_{90} = \sqrt{-\ln(0{,}9)} \Rightarrow G_{90} = \frac{\sqrt{-\ln(0{,}9)}}{0{,}01} \approx 32{,}459$$

Es decir, como G_{90} es 32.459 y está medido en miles de horas, se puede garantizar el funcionamiento de una pieza durante 32459 hs con una confianza del 90 %.

3.3. Item c

Aquí sabemos que las 1000 piezas seleccionadas tienen una duración mayor a 10000 hs (es decir, $X > 10$). Queremos estimar cuántas de estas cumplen $X > 20$. Podríamos pensar en las siguiente variable:

$$Y_{1000} = \text{Cantidad de piezas con duración mayor a 20000 hs entre las 1000 seleccionadas}$$

Podríamos pensar esta variable como una binomial, ya que se extraen las piezas de un lote muy grande y la probabilidad de extraer una pieza con una duración mayor a 20000 hs se mantiene con valores similares para las 1000 extracciones.

Sin embargo, cada una de estas sabemos que tienen una duración mayor a 10000 hs. Por lo que la probabilidad de "éxito" de esta binomial será **condicional**. Es decir, $Y_{1000} \sim \text{Bi}(1000, p)$ donde $p = P(X > 20|X > 10)$.

Calcularemos ahora dicha probabilidad p :

$$P(X > 20|X > 10) = \frac{P(X > 20 \cap X > 10)}{P(X > 10)} = \frac{P(X > 20)}{P(X > 10)} = \frac{1 - F_X(20)}{1 - F_X(10)} = \frac{e^{-(0{,}01 \cdot 20)^2}}{e^{-(0{,}01 \cdot 10)^2}} = \frac{e^{-0{,}04}}{e^{-0{,}01}} \approx 0{,}97045$$

Entonces, el número esperado de piezas con duración mayor a 20000 hs será:

$$E[Y_{1000}] = 1000 \cdot 0{,}97045 = 970{,}45$$

Como debe ser un número entero, redondearemos al más cercano y esperamos un número aproximado de 970 piezas con duración mayor a 20000 hs.

4. Ejercicio 23

Enunciado

El diámetro de contacto de la rosca de una unión se distribuye normalmente con media 10.02 mm y desvío standard 0.075 mm. Las especificaciones dadas para esa rosca son: 10.00 $\pm$ 0.025 mm. (la especificación es un intervalo de valores dentro del cual deberá estar el diámetro para que sea considerado aceptable).

a) ¿Cuál es la probabilidad de que la rosca esté fuera de la especificación dada?

b) Suponiendo que el diámetro de contacto esté centrado en el valor nominal de la especificación (10 mm), ¿cuál es la máxima desviación standard del diámetro aceptable que permitirá no más de una defectuosa entre mil producidas?

Resolución

Variable aleatoria

La variable aleatoria es:

$$D = \text{ Diámetro de contacto de una rosca seleccionada al azar.}$$

Eventos

Podemos considerar el evento:

$$E = \text{La rosca está dentro de los límites de especificación.}$$

Sin embargo, este evento lo podemos vincular con la variable D, ya que $E = 9{,}975 \leq D \leq 10{,}025$.

Datos

El dato que sabemos es que D tiene una distribución normal, cuyos parámetros también son dados: $D \sim N(10{,}02; 0{,}075)$.

4.1. Item a

Nos piden la probabilidad de que una rosca esté fuera de los límites de especificación. Es decir $P(\overline{E}) = 1 - P(E)$. Por lo tanto, podemos calcular $P(E)$ primero:

$$P(E) = P(9{,}975 \leq D \leq 10{,}025)$$

Como D tiene una distribución normal, sus probabilidades se pueden calcular en base a la distribución normal estándard, restando la media (10.02) y dividiendo por el desvío (0.075):

$$P(E) = P(9{,}975 \leq D \leq 10{,}025) = P\left(\frac{9{,}975 - 10{,}02}{0{,}075} \leq \frac{D - 10{,}02}{0{,}075} \leq \frac{10{,}025 - 10{,}02}{0{,}075}\right)$$

Esta última igualdad entre probabilidades, se debe a que como eventos deben ser iguales, es decir, siempre que se cumpla $9{,}975 \leq D \leq 10{,}025$ se cumple también $\frac{9{,}975-10{,}02}{0{,}075} \leq \frac{D-10{,}02}{0{,}075} \leq \frac{10{,}025-10{,}02}{0{,}075}$ y viceversa, ya que se aplican funciones crecientes y biyectivas en todos los miembros de la desigualdad. Si bien luego es algo que se incorpora, es importante saber los motivos por los que las probabilidades no se ven alteradas, ya que aplicar transformaciones erróneas puede derivar en un cálculo desacertado de probabilidades.

Por otro lado, como $D \sim N(10{,}02; 0{,}075) \Rightarrow \frac{D-10{,}02}{0{,}075} \sim N(0;1)$. Entonces:

$$P\left(\underbrace{\frac{9{,}975 - 10{,}02}{0{,}075}}_{-0{,}6} \leq \frac{D - 10{,}02}{0{,}075} \leq \underbrace{\frac{10{,}025 - 10{,}02}{0{,}075}}_{0{,}066}\right) = P\left(\frac{D - 10{,}02}{0{,}075} \leq 0{,}066\right) - P\left(\frac{D - 10{,}02}{0{,}075} \leq -0{,}6\right) = \Phi(0{,}066) - \Phi(-0{,}6)$$

De estos últimos valores, la primera se puede calcular por tabla y se corresponde con la siguiente área:

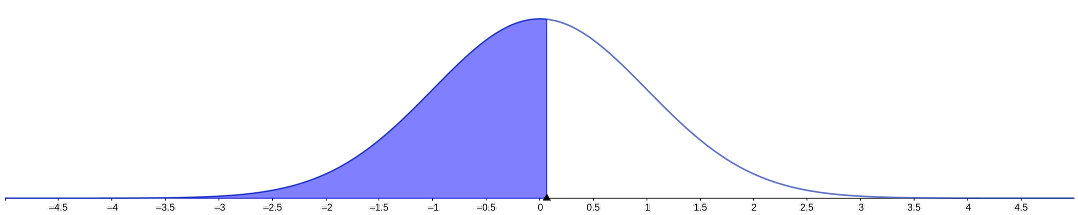

Figura 1: Área correspondiente a $\Phi(0{,}066) = 0{,}5239$

Para calcular $\Phi(-0{,}6)$, como la tabla no tiene valores negativos, nos valdremos de la simetría de la distribución normal:

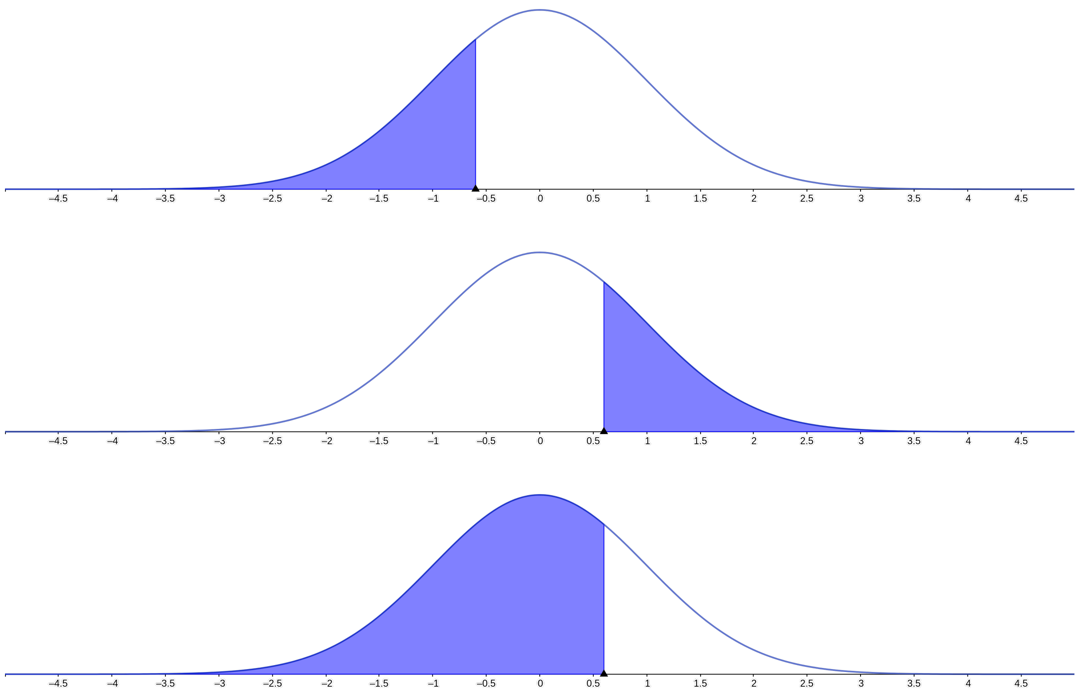

Figura 2: El primer gráfico corresponde al área que equivale a $\Phi(-0{,}6)$, que coincide con el área del segundo gráfico. Este valor se obtiene como el complemento de $\Phi(0{,}6)$, representado en el tercer gráfico. Es decir, se concluye que $\Phi(-0{,}6) = 1 - \Phi(0{,}6) = 1 - 0{,}7257 = 0{,}2743$.

Por lo tanto,

$$P(E) = \Phi(0{,}066) - \Phi(-0{,}6) = 0{,}5239 - 0{,}2743 = 0{,}2536 \Rightarrow P(\overline{E}) = 1 - 0{,}2536 = 0{,}7464$$

4.2. Item b

En este nuevo ítem, la media es conocida pero el desvío no. Por lo tanto, $D \sim N(10; \sigma)$.

Se pide σ de forma que $P(\overline{E}) = 0{,}001$. Esto equivale a pedir que $P(E) = 0{,}999$:

$$0{,}999 = P(E) = P\left(9{,}975 \leq D \leq 10{,}025\right) = P\left(\frac{9{,}975 - 10}{\sigma} \leq \frac{D - 10}{\sigma} \leq \frac{10{,}025 - 10}{\sigma}\right)$$

Notar que hemos restado y dividido por otros valores ya que en este ítem, la media es de 10 mm y el desvío desconocido. Con estos últimos cálculos, hemos estandarizado la variable D y por lo tanto:

$$0{,}999 = P(E) = \Phi\left(\frac{0{,}025}{\sigma}\right) - \Phi\left(-\frac{0{,}025}{\sigma}\right) \overset{\text{Simetría}}{=} \Phi\left(\frac{0{,}025}{\sigma}\right) - \left[1 - \Phi\left(\frac{0{,}025}{\sigma}\right)\right] = 2\Phi\left(\frac{0{,}025}{\sigma}\right) - 1$$

Para despejar σ, utilizaremos la última expresión:

$$0{,}999 = 2\Phi\left(\frac{0{,}025}{\sigma}\right) - 1 \Rightarrow \frac{0{,}999 + 1}{2} = \Phi\left(\frac{0{,}025}{\sigma}\right) \Rightarrow \Phi^{-1}(0{,}9995) = \frac{0{,}025}{\sigma} \Rightarrow$$

$$\Rightarrow \sigma = \frac{0{,}025}{\Phi^{-1}(0{,}9995)} \approx \frac{0{,}025}{3{,}0902} \approx 0{,}008090091$$

Observación: El valor del cuantil fue obtenido con el último valor de la tabla, porque era el mayor disponible. Pero si se utiliza un software, este valor puede calcularse de forma aún más precisa:

$$\Rightarrow \sigma = \frac{0{,}025}{\Phi^{-1}(0{,}9995)} \approx \frac{0{,}025}{3{,}2905} \approx 0{,}0075976$$

5. Ejercicio 31

Enunciado

La longitud de ciertos dispositivos mecánicos tiene la siguiente tolerancia: 58 $\pm$ 0.21 mm. Se han recibido tres partidas de distintos proveedores de 200, 300 y 500 cajas de 6 unidades cada una, que si bien no difieren en su valor medio 58 mm presentan distintas dispersiones: 0.102 mm, 0.12 mm y 0.146 mm. Se supone que la longitud de esos dispositivos es una variable aleatoria con distribución normal. En el almacén se han mezclado las cajas. Si se elige una caja al azar y se comprueba que exactamente 4 de las piezas están dentro de los límites de tolerancia, ¿cuál es la probabilidad de que la caja haya sido entregada por el proveedor número 2?

Resolución

Variables aleatorias

Podemos considerar las siguientes variables aleatorias:

- $X =$ la longitud de un dispositivo seleccionado al azar.
- $X_i =$ la longitud de un dispositivo del proveedor i seleccionado al azar. $(1 \leq i \leq 3)$
- $Y =$ Cantidad de dispositivos con longitud dentro de los límites de tolerancia entre los 6 de la caja seleccionada.
- $Y_i =$ Cantidad de dispositivos con longitud dentro de los límites de tolerancia entre los 6 de una caja seleccionada del proveedor i. $(1 \leq i \leq 3)$

Eventos

Definimos los siguientes eventos:

- $A_i =$ la caja seleccionada viene del proveedor i. $(1 \leq i \leq 3)$
- $T_i =$ Un dispositivo seleccionado al azar del proveedor i tiene una longitud dentro de los límites de tolerancia. $(1 \leq i \leq 3)$

Datos

Tenemos los siguientes datos:

- $P(A_1) = 0{,}2$
- $P(A_2) = 0{,}3$
- $P(A_3) = 0{,}5$
- $X_1 \sim N(58; 0{,}102)$
- $X_2 \sim N(58; 0{,}12)$
- $X_3 \sim N(58; 0{,}146)$

Además, podemos considerar que cada variable Y_i tiene una distribución binomial, con probabilidad de "éxito" $p_i = P(T_i)$. Es decir, $Y_i \sim \mathrm{Bi}(6; p_i)$.

Nos piden $P(A_2|Y = 4)$. Utilizando el teorema de Bayes:

$$P(A_2|Y=4) = \frac{P(Y=4|A_2)P(A_2)}{P(Y=4|A_1)P(A_1) + P(Y=4|A_2)P(A_2) + P(Y=4|A_3)P(A_3)}$$

Es decir,

$$P(A_2|Y=4) = \frac{P(Y_2=4)P(A_2)}{P(Y_1=4)P(A_1)+P(Y_2=4)P(A_2)+P(Y_3=4)P(A_3)}$$

Es decir, nos resta calcular $P(Y_i = 4)$ para $1 \leq i \leq 3$. Sabiendo que son binomiales, resta sacar la probabilidad $P(T_i)$.

- **Si** $i = 1$:

$$P(T_1) = P(58-0{,}21 \leq X_1 \leq 58+0{,}21) = \Phi\left(\frac{0{,}21}{0{,}102}\right) - \Phi\left(-\frac{0{,}21}{0{,}102}\right) = 2\Phi\left(\frac{0{,}21}{0{,}102}\right) - 1 \approx 0{,}96049$$

$$P(Y_1=4) = \binom{6}{4}(0{,}96049)^4(1-0{,}96049)^2 \approx 0{,}019929$$

- **Si** $i = 2$:

$$P(T_2) = P(58-0{,}21 \leq X_2 \leq 58+0{,}21) = \Phi\left(\frac{0{,}21}{0{,}12}\right) - \Phi\left(-\frac{0{,}21}{0{,}12}\right) = 2\Phi\left(\frac{0{,}21}{0{,}12}\right) - 1 \approx 0{,}91988$$

$$P(Y_2=4) = \binom{6}{4}(0{,}91988)^4(1-0{,}91988)^2 \approx 0{,}068944$$

- **Si** $i = 3$:

$$P(T_3) = P(58-0{,}21 \leq X_3 \leq 58+0{,}21) = \Phi\left(\frac{0{,}21}{0{,}146}\right) - \Phi\left(-\frac{0{,}21}{0{,}146}\right) = 2\Phi\left(\frac{0{,}21}{0{,}146}\right) - 1 \approx 0{,}84967$$

$$P(Y_3=4) = \binom{6}{4}(0{,}84967)^4(1-0{,}84967)^2 \approx 0{,}17668$$

Notar que a mayor dispersión, menor es la probabilidad de que un dispositivo esté dentro de los límites de tolerancia.

Por lo tanto,

$$P(A_2|Y=4) \approx \frac{0{,}068944 \cdot 0{,}3}{0{,}019929 \cdot 0{,}2 + 0{,}068944 \cdot 0{,}3 + 0{,}17668 \cdot 0{,}5} \approx 0{,}18302$$