Tarea 5 - Análisis de datos
Giovanni Gamaliel López Padilla

Problema 3

La vida útil de un cierto tablet sigue una distribución normal con promedio 3 años y σ^2=0.9

a) **Calcula la probabilidad que funcionará más de cuatro años.**
Se tiene que $\mu = 3$ y $\sigma^2 = 0.9$. Entonces, al tratarse de una distribución normal, la probabilidad que la tablet funcione más de cuatro años es:

$$\begin{aligned} P(X > 4) &= 1 - P(x \leq 4) \\ &= 1 - \int_{-\infty}^{4} \mathcal{N}(3, 0.9) dx \\ &= 1 - 0.8540797 \\ P(X > 4) &= 0.1459203 \end{aligned}$$

lo anterior fue calculado con la función pnorm de R.

b) **Se quiere determinar la duración de la garantía. ¿Para cuántos meses máximos se debe dar la garantía para que la probabilidad que el tablet se descomponga antes del fin de la garantía sea no mayor que 0.25? La respuesta debe ser en términos de meses enteros.**
Se tiene que $\mu = 3$ y $\sigma^2 = 0.9$. Entonces, al tratarse de una distribución normal, lo que se quiere encontrar es el valor de x tal que:

$$\begin{aligned} P(X) &< 0.25 \\ \int_{-\infty}^{x} \mathcal{N}(3, 0.9) dx &= 0.25 \\ x &= 2.360123 \end{aligned}$$

Esto fue calculado con la función qnorm de R. Entonces se tiene que en 2.360123 años sería el límite para la gararantía. LLevando esta cantidad a números enteros, el máximo de meses es 28 meses.

Problema 5

El tiempo de ejecucion de algoritmo A es $\mathcal{N}$(2, 1) y de algoritmo B es $\mathcal{N}$(4, 2). Si los tiempos son independientes entre sí, calcula la probabilidad que en una corrida B es más rapido que A.

La probabilidad que el algoritmo B obtenga un tiempo menor que A es definida como $P(B < A)$. Sea C el conjunto de valores que toma la variable aleatoria X tal que:

$$C = \{x \in X : \mathcal{N}(x, 4, 2) < \mathcal{N}(x, 2, 1)\}$$

Encontrando el conjunto de C, se tiene lo siguiente:

$$\begin{aligned}
\frac{1}{\sigma_1\sqrt{2\pi}} exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right) &= \frac{1}{\sigma_2\sqrt{2\pi}} exp\left(-\frac{(x-\mu_2)^2}{2\sigma_2^2}\right) \\
-ln\left(\sigma_1\sqrt{2\pi}\right) - \frac{(x-\mu_1)^2}{2\sigma_1^2} &= -ln\left(\sigma_2\sqrt{2\pi}\right) - \frac{(x-\mu_2)^2}{2\sigma_2^2} \\
\frac{(x-\mu_2)^2}{2\sigma_2^2} - \frac{(x-\mu_1)^2}{2\sigma_1^2} &= ln\left(\sigma_1\sqrt{2\pi}\right) - ln\left(\sigma_2\sqrt{2\pi}\right) \\
\sigma_1^2(x-\mu_2)^2 - \sigma_2^2(x-\mu_1)^2 &= 2\sigma_1^2\sigma_2^2 ln\left(\frac{\sigma_1}{\sigma_2}\right) \\
(\sigma_1^2 - \sigma_2^2)x^2 + 2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2)x + \sigma_1^2\mu_2^2 - \sigma_2^2\mu_1^2 - 2\sigma_1^2\sigma_2^2 ln\left(\frac{\sigma_1}{\sigma_2}\right) &= 0 \qquad (1)
\end{aligned}$$

donde la ecuación 1 es una ecuación cuadratica. Tomando en cuenta los parametros $\mu_1 = 2$, $\sigma_1 = 1$, $\mu_2 = 4$ y $\sigma_2 = 2$, las soluciones son $x = \{-3.063706, 3.063706\}$. Como las distribuciones representan tiempos, entonces las solución negativa será despreciada. Por lo que la el conjunto C puede ser escrito como:

$$C = \{x \in X : 0 < x < 3.063706\}$$

Entonces

$$\begin{aligned}
P(B < A) &= P(0 < x < 3.063706) \qquad (2) \\
P(B < A) &= P(x < 3.063706) - P(x < 0) \qquad (3)
\end{aligned}$$

La probabilidad señalada en la ecuación 3 es la señalada en la figura 1.

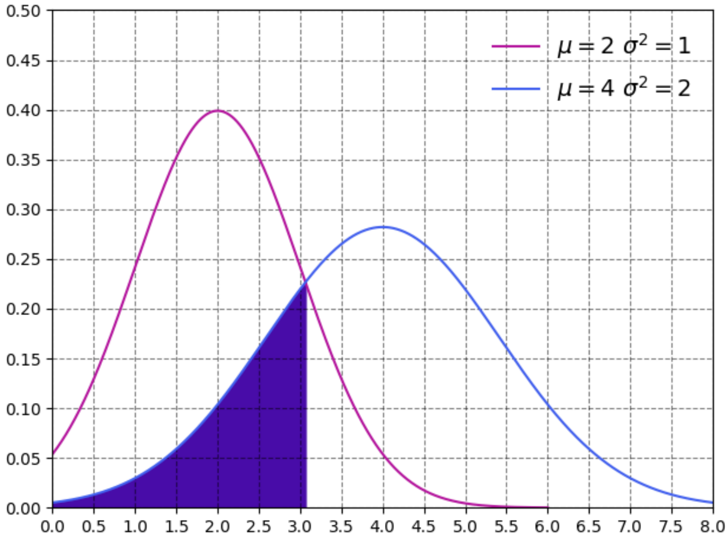

Figura 1: Área a integrar en la ecuación 1.

Por lo tanto

$$\begin{aligned}
P(B < A) &= P(x < 3.063706) - P(x < 0) \\
&= 0.3198396 - 0.02275013 \\
P(B < A) &= 0.2970894
\end{aligned}$$