Análisis de Datos 2016 - Práctica 3

1. Sea X una variable aleatoria con densidad dada por:

$$f(x) = \begin{cases} c(2-x) & si \quad 0 \leq x \leq 2 \\ 0 & cc \end{cases}$$

   (a) Determinar el valor de c y graficar $f(x)$
   (b) Obtener y graficar $F(x)$
   (c) Calcular $P(1 \leq X \leq 2)$
   (d) Calcule $E(X)$ y $V(X)$
   (e) Sea $Y = X^3$, calcular $E(Y)$
   (f) Calcular el 80-*percentil*. (el *cuantil*-0.80)

2. Un proveedor de queroseno tiene un tanque de 150 galones que se llena al empezar cada semana. Si Y denota la demanda semanal en cientos de galones, la función de disatribución acumulada de la demanda puede ser modelada por

$$F(y) = \begin{cases} 0 & si & y < 0 \\ y^2/2 & s & 0 < y < 1 \\ y - 1/2 & si & 1 \leq y < 1,5 \\ 1 & si & y > 1,5 \end{cases}$$

   (a) Grafique $F(y)$, calcule y grafique $f(y)$
   (b) Encuentre $P(0 \leq Y \leq 0,5)$.
   (c) Encuentre $P(0,5 \leq Y \leq 1,2)$.
   (d) Calcule $E(Y)$ y $V(Y)$
   (e) ¿Cuál es el valor de t (medido en cientos de galones) para que se verifique que el 80% de las semanas le sobren más de ese valor t?

3. Como una medición de inteligencia, a unos ratones se les toma el tiempo que tardan para pasar por un laberinto para llegar a una recompensa de alimento. El tiempo (en segundos) necesario para cualquier ratón es una variable aleatoria Y con una función de densidad dada por

$$f(x) = \begin{cases} b/x^2 & si \quad x \geq b \\ 0 & cc \end{cases}$$

   donde b es el tiempo mínimo posible necesario para recorrer el laberinto.

   (a) Demuestre que $f(x)$ tiene las propiedades de una función de densidad.
   (b) Encuentre $F(x)$.
   (c) Encuentre $P(Y > b + c)$ para una constante positiva c.
   (d) Si c y d son constantes positivas tales que $d > c$, encuentre $P(Y > b + d \mid Y > b + c)$

4. Una familia de funciones de densidad que se ha usado para modelizar la distribución del ingresos, es la familia Pareto. Esta tiene dos parámetros k y θ, ambos positivos.y la densidad es:

$$f(x) = \begin{cases} \frac{k\theta^k}{x^{k+1}} & si \quad x \geq \theta \\ 0 & si \quad x < \theta \end{cases}$$

   (a) Graficar y verificar que es una función de densidad
   (b) Calcular y graficar la función de distribución
   (c) Si $k > 1$ calcule $E(X)$

(d) ¿qué se puede decir de $E(X)$ si $k = 1$?

5. La demanda semanal de gas propano (en miles de galones) de una instalación particular es una variable aleatoria X con función de densidad de probabilidad

$$f(x) = \begin{cases} 2(1 - \frac{1}{x^2}) & si \quad 1 < x < 2 \\ 0 & cc \end{cases}$$

   (a) Calcule la función de distribución acumulativa de X.
   (b) Obtenga una expresión para el cuantil α. ¿Cuál es el valor de $\widetilde{\mu}$?
   (c) Calcule $E(X)$ y $V(X)$.
   (d) Si 1500 galones están en existencia al principio de la semana y no se espera ningún nuevo suministro durante la semana, ¿cuántos de los 1500 galones se espera que queden al final de la semana? [Sugerencia: Sea $h(x)$ =cantidad que queda cuando la demanda es x.]

6. Se cree que el tiempo X (min) para que un ayudante de laboratorio prepare el equipo para cierto experimento tiene una distribución uniforme (25 ; 35).

   (a) Determine la función de densidad de probabilidad de X y trace la curva de densidad de correspondiente.
   (b) ¿Cuál es la probabilidad de que el tiempo de preparación exceda de 33 min?
   (c) ¿Cuál es la probabilidad de que el tiempo de preparación esté dentro de dos min del tiempo medio? [Sugerencia: Identifique en la gráfica de $f(x)$.]
   (d) Con cualquier a de modo que $25 < a < a + 2 < 35$, ¿cuál es la probabilidad de que el tiempo de preparación esté entre a y $a + 2$ min?

7. La falla de una tarjeta de circuito que utiliza un sistema de cómputo interrumpe el trabajo hasta que se instala una nueva. El tiempo de entrega de la nueva tarjeta, Y, está uniformemente distribuido en el intervalo de uno a cinco días. El costo de la falla de una tarjeta y la interrupción incluye el costo fi jo c_0 de una nueva tarjeta y un costo que aumenta proporcionalmente con Y^2. Si C es el costo en que se incurre, $C = c_0 + c_1Y^2$.

   (a) Encuentre la probabilidad de que el tiempo de entrega exceda de dos días.
   (b) En términos de c_0 y c_1, encuentre el costo esperado asociado con una sola tarjeta de circuito que falle.

8. Sea $T \sim \epsilon(\lambda)$, verificar que para cualquier t_0 se cumple:

$$P(T \geq t + t_0 \mid T \geq t_0) = P(T \geq t)$$

   esta propiedad se suele llamar "ausencia de memoria" de la distribución exponencial.

9. El tiempo requerido para pasar por la inspección en los aeropuertos puede ser molesto para los pasajeros. En un aeorpuerto internacional, en los periodos pico, el tiempo medio de espera en la fila de inspección de seguridad es de 12, 1 min. Suponga que los tiempos para pasar por la inspección de seguridad tienen una distribución exponencial.

   (a) ¿Cuál es la probabilidad de que durante los periodos pico se requieran al menos 10 minutos para pasar la inspección de seguridad?
   (b) ¿De que durante los periodos pico se requieran más de 20 minutos para pasar la inspección de seguridad?
   (c) ¿De que durante los periodos pico se requieran entre 10 y 20 minutos para pasar la inspección de seguridad?

(d) Son las 8 de la mañana (periodo pico) y usted se acaba de formar en la fila para la inspección de seguridad. Para alcanzar su avión, tiene que estar en la puerta de arribo en no más de 30 minutos. Si necesitara 12 minutos una vez pasada la inspección de seguridad para llegar a la puerta de arribo, ¿cuál es la probabilidad de que pierda el avión?

10. El tiempo (en horas) que un gerente tarda en entrevistar a un solicitante de trabajo tiene una distribución exponencia con $\lambda = 1/2$. Los solicitantes son citados a intervalos de un cuarto de hora, empezando a las 8:00 a.m. y llegan exactamente a tiempo. Cuando el candidato citado a las 8:15 a.m. llega a la oficina del gerente, ¿cuál es la probabilidad de que tenga que esperar antes de ser entrevistado?

11. La magnitud de temblores registrados en una región de América del Norte puede modelarse como si tuviera una distribución exponencial con media $2,4$, según se mide en la escala de Richter. Encuentre la probabilidad de que un temblor que ocurra en esta región

    (a) sea mayor que $3,0$ en la escala de Richter.

    (b) caiga entre $2,0$ y $3,0$ en la escala de Richter.

    (c) De los siguientes diez temblores que afecten esta región, ¿cuál es la probabilidad de que al menos uno de ellos sea mayor que $5,0$ en la escala de Richter?

12. Sea X el tiempo (en minutos) entre dos llegadas sucesivas a un servicio de emergencias. Si X tiene distribución exponencial con $\lambda = 0,125$. Calcular:

    (a) El tiempo esperado entre dos llegadas sucesivas.

    (b) La probabilidad de que el tiempo entre dos llegadas sea menor de 10 minutos

    (c) La probabilidad de más de 5 llegadas en un período de 20 minutos

13. La variable Y tiene distribución normal típica. Calcular las probabilidades de (a) $Y \leq 2,23$, (b) $Y > 1,35$, (c) $-0,51 < Y < 1,54$

14. Una distribución normal tiene una media de 50 y una desviación estándar de 4.

    (a) Calcule la probabilidad de un valor localizado entre $44,0$ y $55,0$.

    (b) Calcule la probabilidad de un valor mayor que $55,0$.

    (c) Calcule la probabilidad de un valor localizado entre $52,0$ y $55,0$

15. El valor medio de radiación cósmica recibida por viajar en avión es de $4,35$ mrem y su desviación estándar de $0,59$. Si la variable sigue una distribución normal,

    (a) determine la probabilidad de que la radiación recibida esté entre $4,01$ y $4,99$ mrem.

    (b) sea de al menos $5,51$ mrem

16. Una fábrica produce tornillos, las especificaciones indican que el diámetro de los mismos debe estar entre $1,19$ y $1,21$ pulgadas. Si el proceso de producción es tal que el diámetro de los tornillos es una variable aleatoria con distribución normal con media $1,196$ y desviación estandar $0,005$. ¿Qué porcentaje de la producción no satisface las especificaciones?

17. Las ventas mensuales de un producto, tienen una distribución normal, con una media de 1200 y una desviación estándar de 225. Al fabricante le gustaría establecer niveles de inventario de manera que sólo haya a lo sumo 5% de probabilidad de que se agoten las existencias. ¿Dónde se deben establecer los niveles de inventario?

18. Se ha estudiado el volumen corpuscular medio eritrocitario (VCM) en pacientes con posible diagnóstico de anemia ferropénica como indicador de esta patología, el verdadero diagnóstico se establece con biopsia de médula osea. Para simplificar, suponemos que los valores de VCM en la población siguen ua distribución normal. En una población de pacientes con diagnóstico confirmado de anemia ferropénica se ha estimado $\mu = 73$ y $\sigma = 6,1$. En una población donde se ha descartado ese diagnóstico se ha estimado $\mu = 82$ y $\sigma = 5,9$. Si se utilza un valor del VCM inferior a 75, para diagnosticar anemia ferropénica.

(a) ¿Cuál es la sensibilidad de este método? Esto es: ¿cuál es probabilidad de diagnosticar correctamente un paciente que tiene anemia ferropénica? ¿Cuál es la probabilidad de un falso positivo?

(b) ¿Cuál es la especificidad? Esto es la probabilidad de clasificar correctamente a un paciente sin anemia ferropénica ¿Cuál es la probabilidad de un falso negativo?

(c) ¿Cómo cambiarían esos valores si el valor de corte fuera 80, y si fuera 85?