ESTADÍSTICA Y ANÁLISIS DE DATOS

Práctica del Tema 7. Distribución Normal, teorema central del límite

1. Sea Z una variable aleatoria con distribución normal reducida $N(0,1)$. Calcular:

   a) $P(Z \leq 0{,}6)$
   b) $P(Z > 0{,}35)$
   c) $P(Z \leq -0{,}30)$
   d) $P(Z > -1{,}60)$
   e) $P(1{,}30 < Z < 1{,}40)$
   f) $P(|Z| \leq 2)$
   g) $P(|Z| > 2)$
   h) $P(|Z| \leq 3{,}275)$
   i) $P(1{,}35 \leq Z \leq 1{,}45)$
   j) $P(-0{,}33 < Z \leq 0{,}563)$

2. Sea Z una variable aleatoria con distribución normal reducida $N(0,1)$. Calcular el valor de a en los siguientes casos:

   a) $P(Z \leq a) = 0{,}8264$
   b) $P(Z \geq a) = 0{,}9$
   c) $P(Z \leq a) = 0{,}1314$
   d) $P(Z > a) = 0{,}3$
   e) $P(-0{,}5 < Z < a) = 0{,}6687$
   f) $P(|Z| \leq a) = 0{,}2$

3. Sea X una variable aleatoria con distribución normal reducida $N(0,1)$ e $Y = X^2$.

   a) Demuestra que $E(Y) = 1$.
   b) Demuestra que $Var(Y) = \sigma_Y^2 = 2$.

4. Sea X una variable aleatoria con distribución normal $N(5, \sigma^2 = 4)$.

   a) Calcular las siguientes probabilidades:
      1) $P(X > 6)$
      2) $P(X < 7)$
      3) $P(X > 0)$
      4) $P(|X| < 8)$

   b) Obtener el valor de k tal que:
      1) $P(X < k) = 0{,}5$
      2) $P(|X - 5| > k) = 0{,}05$

c) Escribir la función de densidad y la función característica de X

5. La cantidad de papel, en kg, que una persona utiliza al mes es una variable aleatoria con distribución $N(14, \sigma^2 = 50)$ y se supone que el consumo entre los diferentes meses es independiente.

   a) Calcular la probabilidad de que dicha persona utilice más de 20 kg en un mes.

   b) Calcular la probabilidad de que dicha persona utilice más de 150 kg en un año.

   c) Si se sabe que dicha persona recicla sólo el 40 % del papel que utiliza, calcular la probabilidad de que en un año recicle menos de 70 kg.

6. La producción de piezas por hora de una máquina tiene una distribución normal, $N(6, \sigma^2 = 0{,}81)$. Suponiendo que cada hora de trabajo la máquina gasta 30000 euros y que cada pieza supone un ingreso de 6000 euros, calcular:

   a) La distribución del beneficio de la máquina por hora de trabajo.

   b) Probabilidad de que en una hora cualquiera se pierda dinero.

   Si en una fábrica hay 8 máquinas de este tipo, que trabajan de forma independiente, calcular:

   c) La distribución del beneficio de la fábrica por hora de trabajo.

   d) Probabilidad de que la fábrica pierda dinero en una hora cualquiera.

7. El número de viajeros de un tren que llegan a una determinada estación tiene una distribución normal $N(100, \sigma^2 = 20)$. En dicha estación los viajeros pueden bajar del tren o continuar viaje, y pueden subir nuevos viajeros. El número de viajeros que bajan del tren en esta estación tiene una distribución $N(30, \sigma^2 = 12)$, y el número de pasajeros que suben, una $N(40, \sigma^2 = 9)$. Si todas las v.a. anteriores son independientes entre sí, hallar:

   a) La distribución del número de viajeros con los que parte el tren de dicha estación.

   b) Determinar qué número de asientos ha de tener el tren para que haya probabilidad 0,95 de que un viajero no tenga que viajar de pie a partir de dicha estación.

8. Una empresa produce sacos de patatas cuyo peso en kg., X, sigue una distribución $N(30, \sigma^2 = 9)$. La empresa vende mensualmente 1000 sacos de patatas. El coste de producción de cada kilo de patatas es de 1 euro (no hay otros costes). El precio de mercado de cada kilo de patatas es de 2 euros (no hay otros ingresos). Suponer independencia en los pesos de los sacos.

   a) Si elegimos un saco de patatas al azar, ¿cuál es la probabilidad de que pese más de 28 kilos?

   b) Si elegimos dos sacos de patatas al azar, ¿cuál es la probabilidad de que juntos pesen más de 56 kilos?

c) Obtén la distribución de los kilos de patatas vendidos mensualmente por la empresa.

d) Obtén la distribución de los ingresos mensuales de la empresa.

e) Obtén la distribución de los beneficios mensuales de la empresa.

f) ¿Cuál es la probabilidad de que la empresa obtenga unos beneficios mensuales inferiores a 25000 euros?

9. Consideremos la sucesión de v.a. continuas $\{X_n\}_{n\in N}$, con funciones de densidad respectivas:

$$f_n(x) = \begin{cases} 0 & x < 0 \\ ne^{-nx} & x \geq 0 \end{cases}$$

a) Demostrar que la sucesión $\{X_n\}$ converge en distribución a la v.a $X \equiv 0$.

b) Si la sucesión de funciones características es $\psi_n(u) = \frac{n}{n-iu}$, probar el mismo resultado teniendo en cuenta que la f.c. de $X \equiv 0$ es 1.

10. Sea una sucesión de v.a. independientes, $\{X_n\}_{n=1,2,3,\ldots}$, con distribución uniforme en el intervalo $\left[0, \frac{2n}{n+1}\right]$.

a) Obtener la función de distribución de $\{X_n\}$, denotada por $F_n(x)$.

b) Obtener la función generatriz de momentos de la v.a. X_n, denotada por $\alpha_n(u)$.

c) Estudiar la convergencia en distribución de la sucesión $\{X_n\}$.

11. Un granjero utiliza un sistema automático para llenar sacos de maíz garantizando 50 kg. por saco. El peso Y de los sacos es una v.a. con distribución normal de media 50,25 kg. y varianza 2,1 kg^2.

a) Obtener la probabilidad de que un saco elegido al azar no llegue al peso garantizado.

b) En un camión es posible transportar 60 sacos de este tipo. ¿Cuál es la probabilidad de que en un camión el granjero estafe al menos 5 kg. de maíz? (Existe independencia en el peso de los distintos sacos del camión).

12. Una cadena dedicada a la venta de artículos de perfumería tiene establecimientos repartidos por toda la península. El beneficio anual de los 80 establecimientos de las capitales se distribuye (en miles de euros) uniformemente en el intervalo $(5, 25)$. Del beneficio anual (también en miles de euros) de los 60 establecimientos de fuera de las capitales, sólo se sabe que sigue una distribución continua de media 10 y desviación típica 5. Calcula la probabilidad de que el beneficio total anual supere los 1700 miles de euros. (Considerar que los beneficios en los distintos establecimientos son v.a. independientes).

13. La demanda diaria de un cierto artículo, en miles de unidades, es aleatoria con función de densidad:

$$f(x) = \begin{cases} \frac{x}{18} & 0 < x \leq 2 \\ \frac{x^2}{21} & 2 < x \leq 4 \\ 0 & \text{en otro caso} \end{cases}$$

Calcular la probabilidad de que en un periodo de 189 días, la demanda sea superior a 560 miles de unidades. (Partimos de que la demanda entre días es independiente).

14. Cada vez que un individuo participa en un determinado juego, lo hace con probabilidad $\frac{1}{2}$ de ganar, de manera independiente en cada partida. Cada vez que gana, le dan 3 euros y cada vez que pierde tiene que pagar 3 euros. En una tarde en que juega una partida,

    a) Obtén la distribución de probabilidad de la v.a. $X =$*ganancia total obtenida del juego*.
    b) ¿Cuál es la probabilidad de que la ganancia de esa tarde esté comprendida entre -5 y 5 euros?
    c) Calcula la media y la varianza de la v.a. X.
    d) Si a lo largo de un año juega alrededor de 400 partidas, ¿cuál es la probabilidad de que las ganancias de ese año estén comprendidas entre -100 y 100 euros?

15. Sean tres v.a. independientes:
    $X_1 \in N(5, \sigma^2 = 4)$, $X_2 \in N(6, \sigma^2 = 8)$, $X_3 \in N(7, \sigma^2 = 8)$.

    Obtener:

    a) La media y la varianza de $Z = 3X_1 - 4X_2 + X_3 + 9$.
    b) La media y la varianza de $Y = 2X_1 + 2X_2$.
    c) La covarianza entre Z y Y.

    Obtener, haciendo uso de que cuando $X \in N(m, \sigma^2)$, su función característica es $\psi_X(u) = e^{ium - \frac{1}{2}u^2\sigma^2}$ y por lo tanto su función cumulativa es $\mu_X(u) = um + \frac{1}{2}u^2\sigma^2$.

    d) La distribución de la v.a. $Z = 3X_1 - 4X_2 + X_3 + 9$.
    e) $E(Z)$ y $Var(Z)$ a partir de su función característica o cumulativa.
    f) El intervalo de probabilidad 0,95 centrado en la media para Z.

Soluciones a los problemas. Práctica 7

1.
   a) 0,7257
   b) 0,361
   c) 0,382
   d) 0,9452
   e) 0,016
   f) 0,9544
   g) 0,0456
   h) 0,9990
   i) 0,016
   j) 0,345

2.
   a) $a = 0{,}94$
   b) $a = -1{,}28$
   c) $a = -1{,}12$
   d) $a = 0{,}525$
   e) $a = 2$
   f) $a = 0{,}255$

3.
   a) $E(Y) = E(X^2) = Var(X) = 1$
   b) $Var(Y) = E(Y^2) - [E(Y)]^2 = E(X^4) - 1.$
   $E(X^4) = \alpha_4(X) = 3.$
   Luego $Var(Y) = 2.$

4.
   a) $X \in N(5\,, \sigma^2 = 4)$

   1) $$\begin{aligned} P(X > 6) &= P\left(\frac{X-5}{2} > \frac{6-5}{2}\right) = \\ &= P(T > 0{,}5) = 0{,}3085 \end{aligned}$$

   2) $$\begin{aligned} P(X < 7) &= P\left(\frac{X-5}{2} < \frac{7-5}{2}\right) = \\ &= P(T < 1) = \\ &= 0{,}8413 \end{aligned}$$

   3) $$\begin{aligned} P(X > 0) &= P\left(\frac{X-5}{2} > \frac{0-5}{2}\right) = \\ &= P(T > -2{,}5) = \\ &= 0{,}9938 \end{aligned}$$

   4) $$\begin{aligned} P(|X| < 8) &= P\left(\frac{-8-5}{2} < \frac{X-5}{2} < \frac{8-5}{2}\right) = \\ &= P(-6{,}5 < T < 1{,}5) = \\ &= 0{,}9332 \end{aligned}$$

   b) 1) $$\begin{aligned} P(X < k) = 0{,}5 &\Rightarrow P\left(\frac{X-5}{2} < \frac{k-5}{2}\right) = 0{,}5 \\ &\Rightarrow \frac{k-5}{2} = 0 \Rightarrow k = 5 \end{aligned}$$

2) $P(|X-5|>k)=0{,}05 \;\Rightarrow\; 1-P(-k<X-5<k)=0{,}05$
$\Rightarrow\; P(\frac{-k}{2}<\frac{X-5}{2}<\frac{k}{2})=0{,}95$
$\Rightarrow\; P(\frac{-k}{2}<T<\frac{k}{2})=0{,}95$
$\Rightarrow\; \frac{k}{2}=1{,}96 \Rightarrow k=3{,}92$

c) función de densidad: $f_X(x)=\frac{1}{2\sqrt{2\pi}}\,e^{-\frac{1}{2}\left(\frac{x-5}{2}\right)^2}$
función característica: $\psi_X(u)=e^{5iu-2u^2}$

5. $X\in N(14,\sigma^2=50)$

a) $$\begin{aligned} P(X>20) &= P\left(\tfrac{X-14}{\sqrt{50}}>\tfrac{20-14}{\sqrt{50}}\right)= \\ &= P(T>0{,}848) \\ &= 0{,}1977 \end{aligned}$$

b) $Z=X_1+\ldots+X_{12}\in N(12\cdot 14=168\,,\,\sigma^2=12\cdot 50=600)$
$$\begin{aligned} P(Z>150) &= P\left(\tfrac{Z-168}{\sqrt{600}}>\tfrac{150-168}{\sqrt{600}}\right)= \\ &= P(T>-0{,}7348)\approx 1-P(T>0{,}73)= \\ &= =0{,}7673 \end{aligned}$$

c) $R=0{,}4\,Z\in N(168\cdot 0{,}4=67{,}2\,,\,\sigma^2=0{,}4^2\cdot 600=96)$
$$\begin{aligned} P(R<70) &= P\left(\tfrac{R-67{,}2}{\sqrt{96}}<\tfrac{70-67{,}2}{\sqrt{96}}\right)= \\ &= P(T<0{,}2857) \\ &= 0{,}6141 \end{aligned}$$

6. a) Y : v.a. Beneficio por hora y máquina. $Y=6000X-30000$
X : producción por hora y máquina.
$Y\in N(6000,29160000=5400^2)$

b) $P(Y<0)=0{,}13$

c) Y_i : v.a. Beneficio por hora de la máquina i. $Y_i=6000X_i-30000$
X_i : Producción por hora de la máquina i. $Z=Y_1+Y_2+...+Y_8\in N(48000,233280000)$

d) $P(Z<0)=0{,}0009\approx 0$

7. X : v.a. número de viajeros que llegan $X\in N(100,20)$.
Y : v.a. número de viajeros que bajan $Y\in N(30,12)$.
U : v.a. número de viajeros de suben $U\in N(40,9)$.

a) Z : v.a. número de viajeros con los que parte el tren.
$Z=X-Y+U\in N(110,41)$.

b) $P(Z<k)=0{,}95$. Luego $k=120,53$.
Como son asientos ha de ser un número natural, luego son necesarios 121 asientos.

8. $X\in N(30,9)$.

a)
$$\begin{aligned} P(X>28) &= P\left(\frac{X-30}{3}>\frac{28-30}{3}\right)= \\ &= P(T>-0{,}66)= \\ &= 0{,}7454 \end{aligned}$$

b) X_1 : peso saco 1, X_2 : peso saco 2, $X_1 + X_2 \in N(60, \sigma^2 = 18)$

$$\begin{aligned} P(X_1 + X_2 > 56) &= P\left(\frac{X_1 + X_2 - 60}{\sqrt{18}} > \frac{56 - 60}{\sqrt{18}}\right) = \\ &= P(T > -0{,}94) = \\ &= 0{,}8264 \end{aligned}$$

c) Distribución de los kilos vendidos mensualmente:
$V = \sum_{i=1}^{1000} X_i$, es decir son los kilos que pesan la suma de los mil sacos, donde $X_i \in N(30, \sigma^2 = 9)$. Luego $E(V) = 30000$ y $Var(V) = 9000$, de donde $V \in N(30000, \sigma^2 = 9000)$.

d) Distribución de los ingresos mensuales: $Z = 2V$, si cada kilo vendido reporta 2 euros de ingresos, donde $Z \in N(60000, \sigma^2 = 36000)$.

e) Distribución de los beneficios mensuales: $B = 2V - V = V$, por lo tanto la distribución del beneficio coincide con la de las ventas, $B \in N(30000, \sigma^2 = 9000)$.

f)

$$\begin{aligned} P(B < 25000) &= P\left(\frac{B - 30000}{\sqrt{9000}} < \frac{25000 - 30000}{\sqrt{9000}}\right) \\ &= P(T < -52{,}70) \approx 0 \end{aligned}$$

9. *a*) La función de distribución es: $F_n(x) = 0$ si $x < 0$ y $F_n(x) = 1 - e^{-nx}$, si $x \geq 0$. Tomando límites cuando n tiende a ∞ se obtiene $F(x) = 0$, si $x < 0$ y $F(x) = 1$, si $x \geq 0$, que es la función de distribución de la variable $X \equiv 0$. Así pues, la sucesión converge en distribución a 0.

   b) Tomando límites cuando n tiende a ∞, en $\psi_n(u)$, se obtiene $\psi(u) = 1$ que es la f.c. de la v.a. $X \equiv 0$. Es decir, la sucesión converge en distribución a 0.

10. Para cada n, $X_n \in U[0, \frac{2n}{n+1}]$, de donde su función de densidad es:
$f_n(x) = \frac{n+1}{2n}$, si $0 \leq x \leq \frac{2n}{n+1}$.

   a) Para cada n, la función de distribución de X_n es:
   $F_n(x) = P(X_n \leq x) = \int_0^x f_n(x)dx = \int_0^x \frac{n+1}{2n}dx = \frac{n+1}{2n}x$ si $0 \leq x \leq \frac{2n}{n+1}$.
   $F_n(x) = 0$, si $x \leq 0$ y $F_n(x) = 1$, si $x \geq \frac{2n}{n+1}$.

   b) Para cada n, la función generatriz de momentos de X_n es:

$$\begin{aligned} \alpha_n(u) &= E(e^{uX_n}) = \int_0^{\frac{2n}{n+1}} e^{ux} f_n(x)dx = \\ &= \int_0^{\frac{2n}{n+1}} e^{ux} \frac{n+1}{2n} dx = \frac{n+1}{2n} \frac{1}{u} [e^{\frac{2nu}{n+1}} - 1] \end{aligned}$$

   c) A partir del apartado *a*), vemos que

$$lim_{n \to \infty} F_n(x) = \frac{1}{2}x, \quad \text{si} \quad 0 \leq x \leq 2$$

mientras que $lim_{n\to\infty}F_n(x) = 0$, si $x \leq 0$ y
$lim_{n\to\infty}F_n(x) = 1$, si $x \geq 2$.
Así X_n converge en distribución a la v.a. $X \in U[0,2]$.
El mismo resultado puede verse a partir del apartado *b*), probando que

$$lim_{n\to\infty}\alpha_n(u) = \frac{1}{2u}[e^{2u} - 1]$$

es la función generatriz de momentos de una v.a uniforme en $[0,2]$.

11. $Y \in N(50{,}25, \sigma^2 = 2{,}1)$

    a) $P(Y < 50) = P(\frac{Y-50{,}25}{\sqrt{2{,}1}} < \frac{50-50{,}25}{\sqrt{2{,}1}}) = P(T < -0{,}1725) = 0{,}4325$

    b) $Z = Y_1 + ... + Y_{60}$ peso de los 60 sacos del camión, $Z \in N(3015, \sigma^2 = 126)$.
    Así, la probabilidad de que el granjero estafe al menos 5 kilos, es la probabilidad de que el peso real sea inferior a 2995 kilos.
    $P(Z < 2995) = P(\frac{Z-3015}{\sqrt{126}} < \frac{2995-3015}{\sqrt{126}}) = P(T < -1{,}78) = 0{,}0375$

12. X: beneficio de establecimiento en capital.
    $X \in U(5,25)$.
    $E(X) = 15$.
    $Var(X) = 33{,}3$.

    Y: beneficio de establecimiento fuera de la capital.
    $E(Y) = 10$,
    $Var(Y) = 25$.

    Hay 80 establecimientos en capitales y 60 fuera de las mismas. El beneficio total será:

    $B_T = \sum_{i=1}^{80} X_i + \sum_{i=1}^{60} Y_i$

    suma de 140 v.a. independientes. Por el Teorema Central del Límite: B_T tiene distribución aproximadamente normal, de media:

    $E(B_T) = 80\,E(X) + 60\,E(Y) = 1800$ y varianza:

    $Var(B_T) = 80\,\mathrm{Var}(X) + 60\,\mathrm{Var}(Y) = 4166{,}4$

    Luego $B_T \in N(1800, \sigma = 64{,}54)$.
    Entonces: $P(B_T \geq 1700) \approx 0{,}9382$

13. $E(X) = 3,005$,
    $Var(X) = 0,63$.
    Por el T.C.L. $Z = X1 + ... + X_{189} \approx N(567'945, \sigma^2 = 119,07)$.
    $P(Z > 560) \approx 0{,}7642$

14. *a*)

    $$X = \begin{cases} 3, & \text{con } \; p = 0{,}5 \\ -3, & \text{con } 1 - p = 0{,}5 \end{cases}$$

    b) $P(-5 \leq X \leq 5) = 1$.

c) $E(X) = 3 \cdot \frac{1}{2} - 3 \cdot \frac{1}{2} = 0$.
$Var(X) = E(X^2) - [E(X)]^2 = E(X^2) = 3^2 \cdot \frac{1}{2} + 3^2 \cdot \frac{1}{2} = 9$.

d) La ganancias tras 400 partidas viene dada por:
$G = X_1 + \cdots X_{400} \in N(0, \sigma^2 = 400 \cdot 9 = 3600)$
Luego:

$$\begin{aligned} P\left(G \in (-100, 100)\right) &= P\left(\frac{|G - 0|}{\sqrt{3600}} < \frac{100}{\sqrt{3600}}\right) = \\ &= 0{,}905 \end{aligned}$$

15. *a*) $E(Z) = 3 \cdot 5 - 4 \cdot 6 + 7 + 9 = 7$
$\mathrm{Var}(Z) = 9 \cdot 4 + 16 \cdot 8 + 8 = 172$

b) $E(Y) = 2 \cdot 5 + 2 \cdot 6 = 22$
$\mathrm{Var}(Y) = 4 \cdot 4 + 4 \cdot 8 = 48$

c)

$$\begin{aligned} \mathrm{Cov}(Z, Y) &= \mathrm{Cov}\left[(3X_1 - 4X_2 + X_3 + 9)\,,\,(2X_1 + 2X_2)\right] \\ &= 6\mathrm{Var}(X_1) + 6\mathrm{Cov}(X_1, X_2) - 8\mathrm{Cov}(X_2, X_1) - 8\mathrm{Var}(X_2) + \\ &\quad 2\mathrm{Cov}(X_3, X_1) + 2\mathrm{Cov}(X_3, X_2) = \\ &= 6 \cdot 4 - 8 \cdot 8 = -40 \end{aligned}$$

d) Resolvemos el problema utilizando la función cumulativa. Del enunciado, se deduce: $\mu_1(u) = 5u + 4\frac{u^2}{2}$, $\mu_2(u) = 6u + 8\frac{u^2}{2}$, $\mu_3(u) = 7u + 8\frac{u^2}{2}$. Además, del tema anterior, sabemos: $\mu_Z(u) = 9u + \mu_1(3u) + \mu_2(-4u) + \mu_3(u)$, luego,

$$\begin{aligned} \mu_Z(u) &= 9u + [15u + 36\frac{u^2}{2}] + [-24u + 128\frac{u^2}{2}] + [7u + 8\frac{u^2}{2}] = \\ &= 7u + 172\frac{u^2}{2} \end{aligned}$$

Como la función cumulativa de la v.a. Z tiene la misma forma que la de una $N(m, \sigma^2) = N(7, \sigma^2 = 172)$, deducimos que $Z \in N(7, \sigma^2 = 172)$.

e) Sabemos que $\mu'(0) = E(Z)$. $\mu'(u) = 7 + 172u$, luego $E(Z) = 7$. Además $\mu''(0) = Var(Z)$. $\mu''(u) = 172$, luego $Var(Z) = 172$.

f) Tenemos que obtener el valor de a: $P(|Z-7| < a) = P(-a < Z-7 < a) = 0{,}95$. Tipificando la variable Z, tenemos que,

$$\begin{aligned} P(-a < Z - 7 < a) &= P(\frac{-a}{\sqrt{172}} < \frac{Z-7}{\sqrt{172}} < \frac{a}{\sqrt{172}}) = \\ &= P(\frac{-a}{\sqrt{172}} < T < \frac{a}{\sqrt{172}}) = 2\Phi(\frac{a}{\sqrt{172}}) - 1 = 0{,}95 \end{aligned}$$

De donde, $\Phi(\frac{a}{\sqrt{172}}) = 0{,}975$ y mirando en tablas, $\frac{a}{\sqrt{172}} = 1{,}96$, luego, $a = 25{,}71$. El intervalo buscado es,
$(-a < Z - 7 < a) = (-a + 7 < Z < a + 7) = (-18{,}71 < Z < 32{,}71)$.