Solutions Problem Set 1

Tatiana Rosá   Alejo Eyzaguirre

1st September 2021

Outline

1. Pregunta 1: Mixtura de Normales
   - Pregunta 1.A
   - Pregunta 1.B
   - Pregunta 1.C
2. Pregunta 2: Probabilidades, CDF's y PDF's
   - Pregunta 2.A
   - Pregunta 2.B
   - Pregunta 2.C
   - Pregunta 2.D
3. Pregunta 3: Media y Varianza
   - Pregunta 3.B

Pregunta 1: Mixtura de Normales

La familia de mixtura de normales (normal mixtures) con 2 componentes tiene CDF:

$$F(x) = p_1 \Phi(\frac{x - \mu_1}{\sigma_1}) + p_2 \Phi(\frac{x - \mu_2}{\sigma_2}) \tag{1}$$

Donde $p_1 \in [0, 1]$, $p_1 + p_2 = 1, \sigma_1, \sigma_2 < 0$ y μ_1 y μ_2 son escalares. La función Φ es la CDF de una normal estándar. Se le pide:

Pregunta 1.A

Compute la PDF de una V.A X que tiene como distribución una mixtura de normales de 2 componentes.

Solución: Derivamos la CDF y así obtenemos la PDF. Clave usar Regla de la Cadena.

$$\frac{\partial F(x)}{\partial x} = f(x) = \frac{p_1}{\sigma_1} \cdot \phi(\frac{x - \mu_1}{\sigma_1}) + \frac{p_2}{\sigma_2} \cdot \phi(\frac{x - \mu_2}{\sigma_2})$$

Pregunta 1.B

Compute la Esperanza y Varianza de X.

Solución: Usando la definición de Esperanza:

$$
\begin{aligned}
E(x) &= \int_{-\infty}^{+\infty} x \cdot \left(\frac{p_1}{\sigma_1}\phi(\frac{x-\mu_1}{\sigma_1}) + \frac{p_2}{\sigma_2}\phi(\frac{x-\mu_2}{\sigma_2}) \right) dx \\
&= p_1 \int_{-\infty}^{+\infty} \left(\frac{x}{\sigma_1}\phi(\frac{x-\mu_1}{\sigma_1}) \right) dx + p_2 \int_{-\infty}^{+\infty} \left(\frac{x}{\sigma_2}\phi(\frac{x-\mu_2}{\sigma_2}) \right) dx
\end{aligned}
$$

La primera integral es la esperanza de una V.A que distribuye $N(\mu_1, \sigma_1^2)$ por lo tanto, es igual a μ_1. La segunda integral es entonces μ_2.

$$E(x) = p_1\mu_1 + p_2\mu_2$$

Pregunta 1.B

Seguimos con la Varianza, la cual calcularemos con la fórmula $V(x) = E(x^2) - E^2[x]$.

$$E(x^2) = \int_{-\infty}^{+\infty} x^2 \cdot \left(\frac{p_1}{\sigma_1}\phi(\frac{x-\mu_1}{\sigma_1}) + \frac{p_2}{\sigma_2}\phi(\frac{x-\mu_2}{\sigma_2}) \right) dx$$
$$= p_1 \int_{-\infty}^{+\infty} \left(\frac{x^2}{\sigma_1}\phi(\frac{x-\mu_1}{\sigma_1}) \right) dx + p_2 \int_{-\infty}^{+\infty} \left(\frac{x^2}{\sigma_2}\phi(\frac{x-\mu_2}{\sigma_2}) \right) dx$$

El primer término es $E[Y^2]$ si Y es una V.A que distribuye $N(\mu_1, \sigma_1)$ por lo tanto, como sabemos que $E(Y^2) = V(Y) + E^2[Y]$, esta primera integral es equivalente a $\sigma_1^2 + \mu_1^2$. Análogamente, el segundo término es entonces $\sigma_2^2 + \mu_2^2$.

$$E(x^2) = p_1 \cdot (\mu_1^2 + \sigma_1^2) + p_2 \cdot (\mu_2^2 + \sigma_2^2)$$
$$Var(x) = p_1 \cdot (\mu_1^2 + \sigma_1^2) + p_2 \cdot (\mu_2^2 + \sigma_2^2) - (p_1\mu_1 + p_2\mu_2)^2$$

Pregunta 1.C

En el resto del ejercicio imponemos $\sigma_1 = \sigma_2 = 1$, $p_1 + p_2 = 1$ y $\mu_1 + \mu_2 = 0$. ¿Es la densidad simétrica? Grafique la PDF para $\mu_1 = 0.5, \mu_1 = 1, \mu_1 = 1.5$. Comente:

Solución: La densidad será simétrica si el coeficiente de asimetría (*skewness*) es cero. Es decir si $E[(\frac{x-\mu}{\sigma})^3] = 0$, o también si $E(x^3) - 3\mu\sigma^2 - \mu^3 = 0$.
Para nuestro caso particular, $\mu = 0,5\mu_1 + 0,5\mu_2 = 0$ (ver solución 1.A). Entonces simetría se da si y solo sí $E(x^3) = 0$. Revisemos:

$$E(x^3) = p_1 \int_{-\infty}^{+\infty} x^3 \cdot \left(\frac{1}{\sigma_1}\phi(\frac{x-\mu_1}{\sigma_1}) \right) dx + p_2 \int_{-\infty}^{+\infty} x^3 \cdot \left(\frac{1}{\sigma_2}\phi(\frac{x-\mu_2}{\sigma_2}) \right) dx$$

Donde el primer término es $E[Y^3]$ si Y es una V.A. que distribuye $N(\mu_1, \sigma_1^2)$. Como la densidad de la distribución de Y es normal, sabemos que el tercer momento es cero (densidad simétrica). Entonces, tenemos que la integral es equivalente a: $E(Y^3) = 3\mu_1\sigma_1^2 + \mu_1^3$ (ídem para la 2ª Integral).

Pregunta 1.C

Dicho lo anterior si reemplazamos con lo que tenemos hasta el momento:

$$\begin{aligned} E(x^3) &= p_1 \cdot (3\mu_1\sigma_1^2 + \mu_1^3) + p_2 \cdot (3\mu_2\sigma_2^2 + \mu_2^3) \\ &= 1.5\mu_1 + 0.5\mu_1^3 + 1.5\mu_2 + 0.5\mu_2^3 \\ &= 1.5\mu_1 + 0.5\mu_1^3 + 1.5(-\mu_1) + 0.5(-\mu_1)^3 \\ &= 0 \end{aligned}$$

Por lo tanto, la densidad es efectivamente simétrica.

Las gráficas se podían hacer a mano o vía Matlab.

Pregunta 2.A

Sea X $\sim N(0;9)$. Compute $p(4 < X < 4)$ de forma exacta (función) y de forma aproximada, usando la desigualdad de Chebychev.

Solución: Acá la clave es normalizar en la desigualdad así luego podemos revisar la tabla de la acumulada de la normal estándar.

$$
\begin{aligned}
Pr(-4 < X < 4) &= Pr(\frac{-4-0}{\sqrt{9}} < \frac{X-0}{\sqrt{9}} < \frac{4-0}{\sqrt{9}}) \\
&= \Phi(\frac{4}{3}) - \Phi(\frac{-4}{3}) \\
&\approx 0.82
\end{aligned}
$$

Ahora usando la Aproximación de Chebychev. Sabemos que:

$$
-4 < X < 4 \Leftrightarrow |X| < 4
$$
$$
Prob(-4 < X < 4) = Prob(|X| < 4) = 1 - Prob(|X - \mu_x| > 4)
$$

Y lo último es posible dado que por enunciado $\mu_x = 0$.

Pregunta 2.A

Ahora podemos usar Chebychev:

$$\begin{aligned} Prob(|x-\mu|>4) &= Prob((x-\mu)^2>16) \\ &\leq \frac{Var(X)}{4^2} = \frac{9}{4^2} \end{aligned}$$

Entonces, reordenando un poco:

$$\begin{aligned} Prob(-4<X<4) &= 1-Prob(|x-\mu|>4) \\ &\geq \frac{7}{16} \approx 0.4375 \end{aligned}$$

En este caso, la Desigualdad de Chebychev no nos aporta una gran aproximación.

Pregunta 2.B

Obtenga la función de distribución de $Y = X^2 - 7X + 10$ donde X es el resultado de tirar un dado. **Solución:** Calculemos el "recorrido" de Y.

Prob(X=x)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
X	1	2	3	4	5	6
Y	4	0	-2	-2	0	4

Entonces la función de distribución acumulada de Y sería:

$$F_y(y) = \begin{cases} 0 & \text{si } Y < -2 \\ \frac{1}{3} & \text{si } -2 \leq Y \leq 0 \\ \frac{2}{3} & \text{si } 0 \leq Y \leq 4 \\ 1 & \text{si } Y \geq 4 \end{cases}$$

Pregunta 2.C

Aplicando lo visto en clases (Slide 20 del PPT), tenemos que:

$$F_Y(Y) = F_x(g^{-1}(y)) \ , \forall \, y \in \mathcal{Y}$$

Hallamos entonces la inversa de la función g(x).

$$Y = \frac{X}{1+X}, \text{ cambiamos las X por Y.}$$
$$X = \frac{Y}{1+Y}, \text{ despejamos las Y.}$$
$$Y = \frac{X}{1-X}$$

Entonces,

$$F_Y(Y) = F_x(\frac{y}{1-y})$$

Pregunta 2.D

Sea X una variable aleatoria con CDF F(x) continua y estrictamente monotónica. Sea Y la variable aleatoria definida por Y=F(X) (i.e., la función que mapea X en Y es la CDF de X). Cual es la CDF de Y? Y la PDF? Grafique.

Solución: Primero obtenemos la CDF:

$$\begin{aligned} F_Y(y) &= Prob(Y \leq y) \\ &= Prob(F(x) \leq y) \text{ , aplico } F^{-1} \text{ que se que es monótona.} \\ &= Prob(F^{-1}(F(x)) \leq F^{-1}(y)) \\ &= Prob(X \leq F^{-1}(y)) \\ &= F(F^{-1}(y)) \\ &= y \end{aligned}$$

Note que la única forma de que esto se cumpla, es que $y \sim U(0,1)$. Si transformamos una variable aleatoria por su propia CDF, obtenemos una uniforme (0,1), que corresponde a la PDF.

Pregunta 3.B

Sea X una variable aleatoria con distribución U(-1,4). Obtenga la media y varianza de: $Y = X^2$.

Solución: Necesito $F_y(y)$. Sin embargo, acá la función no es monótona (para valores menores que cero es decreciente y para mayores es creciente). Note que es imposible que Y sea algo menor que cero $\forall\, X \in \mathbb{R}$, entonces $F_y(y) = Prob(x^2 \leq y) = 0$ para este caso. Para y $\geq$ 0, tenemos en cambio que:

$$F_y(y) = Prob(x^2 \leq y) = Prob(-\sqrt{y} \leq x \leq \sqrt{y}) = F_x(\sqrt{y}) - F_x(-\sqrt{y})$$

Como sabemos que $X \sim U(-1,4)$ entonces $F_x(x) = \frac{x+1}{4+1} = \frac{x+1}{5}$. Luego reemplazando en la expresión encontrada para $F_y(y)$, tendríamos que:

$$F_y(y) = \begin{cases} 0 & \text{si } y < 0 \\ \frac{2\sqrt{y}}{5} & \text{si } 0 \leq y \leq 16 \\ 1 & \text{si } y \geq 16 \end{cases}$$

Pregunta 3.B

Derivando la expresión obtenida en la expresión encontrada para $F_y(y)$ que es una CDF, obtendríamos entonces nuestra PDF, que nos permite encontrar la Esperanza y la Varianza solicitadas.

$$f_y(y) = \begin{cases} 0 & \text{si } y < 0 \\ \frac{1}{5\sqrt{y}} & \text{si } 0 \leq y \leq 16 \\ 0 & \text{si } y \geq 16 \end{cases}$$

Usando integrales y encontrando la $E(y)$ y la $E(y^2)$, se podían encontrar ambos momentos.

Gracias!

jeeyzaguirre@uc.cl