Juan Pablo Naranjo
201730006
Gabriel Mateo Mejía
201718064
Maria Fernanda Martinez
201716576

CASO DE ESTUDIO 4
Procesos Estocásticos - IELE 4010

15-09-2021

Punto 1

Sean $X_1, \ldots, X_{200}$ variables aleatorias discretas iid (independientes e idénticamente distribuidas) en donde: $X_i = 0$ con probabilidad 0.3, $X_i = 1$ con probabilidad 0.4 y $X_i = 2$ con probabilidad 0.3. Adicionalmente, sea $S = \sum_{i=1}^{200} X_i$.

a)

Se procede a encontrar una cota superior para la probabilidad de que $S \geq 210$ usando la desigualdad de Markov. Se parte de definir la media μ y varianza σ^2 de cada una de las variables aleatorias X_i. Se sabe que

$$p_{X_i}(x_i) = \begin{cases} 0{,}3 & , \quad x_i = 0 \\ 0{,}4 & , \quad x_i = 1 \\ 0{,}3 & , \quad x_i = 2 \end{cases}$$

Por lo tanto, la media $\mu = E[X_i]$ de cada una de las variables aleatorias X_i está dada por

$$\begin{aligned} \mu &= E[X_i] = (0{,}3)(0) + (0{,}4)(1) + (0{,}3)(2) \\ \mu &= 1 \end{aligned} \tag{1}$$

Por otro lado, la varianza $\sigma^2 = \text{var}(X_i)$ de cada una de las variables aleatorias X_i está dada por

$$\begin{aligned} \sigma^2 &= \text{var}(X_i) = E[X_i^2] - (E[X_i])^2 \\ \sigma^2 &= (0{,}3)(0^2) + (0{,}4)(1^2) + (0{,}3)(2^2) - (1)^2 \\ \sigma^2 &= 0{,}6 \end{aligned} \tag{2}$$

Con estos valores en mente, y sabiendo que $E[S] = 200\mu$ y $\text{var}(S) = 200\sigma^2$, ya se puede utilizar la desigualdad de Markov. En términos generales, esta nos dice que, para una variable aleatoria X,

$$P(X \geq a) \leq \frac{E[X]}{a}, a > 0 \tag{3}$$

En este caso, la variable aleatoria que nos interesa analizar es S, razón por la cual se puede utilizar la anterior desigualdad para decir que

$$P(S \geq 210) \leq \frac{200\mu}{210} \tag{4}$$

Reemplazando el valor de μ en la anterior desigualdad, se obtiene que

$$P(S \geq 210) \leq 0{,}9524 \tag{5}$$

y por lo tanto, 0,9524 sería una cota superior para la probabilidad de que $S \geq 210$.

b)

Se procede a utilizar la desigualdad de Chebyshev para hallar una cota superior para la probabilidad de que $S \geq 210$. En términos generales, esta desigualdad nos dice que, para una variable aleatoria X,

$$P(|X - E[X]| \geq \alpha) \leq \frac{\sigma_X^2}{\alpha^2} \tag{6}$$

Para empezar a resolver la desigualdad, se parte del hecho de que $P(S \geq 210)$ también se puede escribir como $P(S - 200 \geq 10)$. Escribir la probabilidad de esta forma nos permite verla de la misma forma en que está expresada la desigualdad de Chebyshev. Entonces, reemplazando los valores de nuestra variable aleatoria S, y sabiendo que $\mu_S = 200\mu = 200$ y que $\sigma_S^2 = 200\sigma^2 = 120$ se tiene que

$$\begin{aligned} P(|S - \mu_S| \geq 10) &\leq \frac{\sigma_S^2}{10^2} \\ P(|S - 200| \geq 10) &\leq \frac{120}{100} \\ P(|S - 200| \geq 10) &\leq 1{,}2 \end{aligned} \tag{7}$$

Ahora se determina qué probabilidad es mayor entre $P(|S - 200| \geq 10)$ y $P(S - 200 \geq 10)$. Para hacer esto, se empieza por reafirmar que $P(S - 200 \geq 10)$ tiene en cuenta todos los valores de S mayores o iguales a 210. Por otro lado, $P(|S - 200| \geq 10)$ abarca todos los valores mayores o iguales a 210, pero también abarca todos los valores menores o iguales a 190. Por esta razón, es claro que como esta última abarca una cantidad mayor de valores que $P(S - 200 \geq 10)$, se puede afirmar que

$$P(S - 200 \geq 10) \leq P(|S - 200| \geq 10)$$

Por lo tanto, como de la desigualdad de Chebyshev se encontró que $P(|S - 200| \geq 10) \leq 1{,}2$, esto le pone otra cota superior a $P(S \geq 210)$:

$$P(S \geq 210) = P(S - 200 \geq 10) \leq 1{,}2 \tag{8}$$

Sin embargo, la cota superior encontrada en este caso partiendo de la desigualdad de Chebyshev no es útil, pues nos está diciendo algo que se sabe siempre de antemano para cualquier probabilidad, pues toda probabilidad siembre es ≥ 0 y ≤ 1. Además, usando la desigualdad de Markov en el literal anterior ya se había encontrado una cota superior que es menor a la encontrada en este literal.

c)

Ahora se usa el teorema del límite central para aproximar la probabilidad de que $S \geq 210$. Es importante recordar que este teorema nos dice que si se tiene una colección de variables aleatorias X_i que son iid, con media μ y varianza σ^2 y se toma un grupo de estas que es suficientemente grande (en general para $n \geq 30$, con n el número de variables aleatorias en el grupo seleccionado), entonces la variable aleatoria $Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$ converge en distribución a una variable aleatoria W, que sigue una distribución normal estándar, es decir, con media cero y varianza 1 ($W \sim N(0, 1)$), en donde $S_n = \sum_{i=1}^{n} X_i$.

Aplicando el teorema para aproximar $P(S \geq 210)$, y recordando que en este caso $S = S_n = S_{200} = \sum_{i=1}^{200} X_i$, se sabe que

$$
\begin{aligned}
P(S \geq 210) &= P\left(\frac{S - 200\mu}{\sigma\sqrt{200}} \geq \frac{210 - 200\mu}{\sigma\sqrt{200}}\right)\\
P(S \geq 210) &= P\left(Y_n \geq \frac{210 - 200(1)}{\sqrt{(0{,}6)(200)}}\right)\\
P(S \geq 210) &= P\left(Y_n \geq 0{,}9129\right)
\end{aligned} \tag{9}
$$

Como se tiene un número de variables aleatorias suficientemente alto (200), en este punto ya se puede asumir que la variable aleatoria $Y_n = Y_{200}$ sigue una distribución normal con media cero y varianza 1. Por lo tanto, se puede aproximar $P(S \geq 210)$ como

$$
\begin{aligned}
P(S \geq 210) &= 1 - P(S \leq 210)\\
P(S \geq 210) &= 1 - P(Y_n \leq 0{,}9129)\\
P(S \geq 210) &= 1 - \Phi(0{,}9129)
\end{aligned} \tag{10}
$$

en donde $\Phi(.)$ es la CDF (función de distribución acumulada) de una variable aleatoria normal estándar. Finalmente,

$$
P(S \geq 210) \approx 0{,}1806 \tag{11}
$$

Comparando con los resultados obtenidos en los literales a) y b), se puede decir que la información proporcionada por las desigualdades de estos literales es cierta en ambos casos, y se pudo corroborar con el resultado obtenido en este literal.

d)

Ahora se vuelve a usar al desigualdad de Chebyshev, pero esta vez para hallar una cota superior para $P(S \geq 200 + \epsilon)$, con $\epsilon > 0$. Se pretende encontrar los valores de ϵ para los cuales esta cota comienza a ser útil.

Poniendo la información conocida dentro de la desigualdad de Chebyshev, se sabe que:

$$
\begin{aligned}
P(|S - \mu_S| \geq \alpha) &\leq \frac{\sigma_S^2}{\alpha^2}\\
P(|S - 200| \geq \alpha) &\leq \frac{120}{\alpha^2}
\end{aligned} \tag{12}
$$

De la misma manera en que se supo en el literal b) que $P(S - 200 \geq 10) \leq P(|S - 200| \geq 10)$, en este literal se usa la misma lógica, pero esta vez para determinar que

$$
P(S \geq 200 + \epsilon) = P(S - 200 \geq \epsilon) \leq P(|S - 200| \geq \epsilon)
$$

De la desigualdad de Chebyshev se sabe que

$$
P(|S - 200| \geq \epsilon) \leq \frac{120}{\epsilon^2}
$$

De aquí se sabe que la desigualdad de Chebyshev será útil siempre y cuando se cumpla que

$$\begin{aligned}\frac{120}{\epsilon^2} &\leq 1 \\ 120 &\leq \epsilon^2 \\ \epsilon &\geq 10{,}95\end{aligned} \tag{13}$$

Punto 2

a)

Se tiene que Z= X+Y donde X y Y son variables independientes con PMF's $fx(x)$ y $f_y(y)$, respectivamente.Como son independientes, se cumple que:

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \tag{14}$$

$$y = z - x \tag{15}$$

Por lo anterior, $F_Z(z) = \int_{-\infty}^{z} f_Z(\tau)\, d\tau$ se puede denotar como una doble integral de $f_{X,Y}(x,y)$ sobre la recta que se ve en la Figura 1 del plano (x,y)

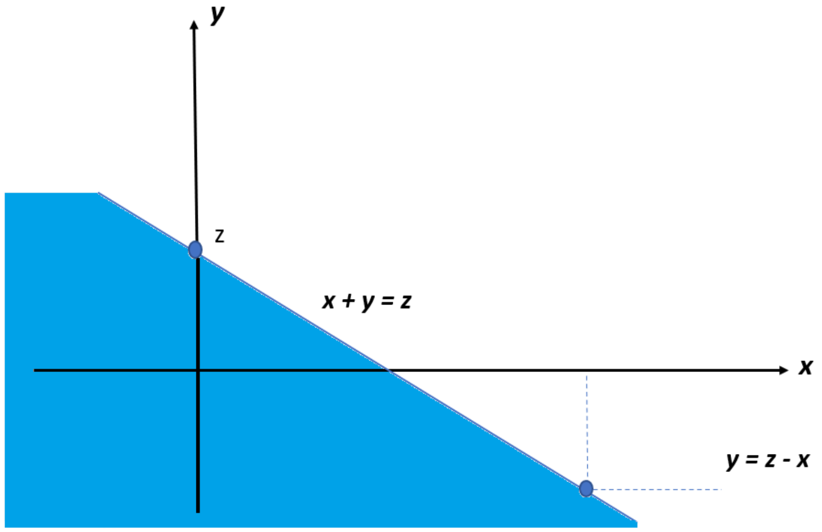

Figura 1

Donde se plantea el área de la siguiente manera:

$$F_Z(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_{X,Y}(x,y)\, dydx$$

Ahora se reescribe teniendo en cuenta que $f_{X,Y}(x,y) = f_X(x)f_Y(y)$:

$$F_Z(z) = \int_{-\infty}^{z-x} \left(\int_{-\infty}^{\infty} f_X(x) f_Y(y) dx \right) dy$$

$$F_Z(z) = \int_{-\infty}^{z-x} g(y)\, dy = G(z-x) - G(-\infty)$$

De esta manera, $f_Z(z) = g(z-x) = \int_{-\infty}^{\infty} f_X(x) f_Y(y)\, dx$ y haciendo el cambio de variable $x = u$ se tiene:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(u) f_Y(z-u)\, du$$

Como se conoce, la expresión descrita anteriormente corresponde a la definición de una convolución $f_Z = f_X * f_Y$

b)

Se conoce que X y Y son variables aleatorias uniformes, independientes, e idénticamente distribuidas, siguiendo una distribución uniforme entre 0 y 1. Por lo tanto, $f_X(x)$ y $f_Y(y)$ están descritas por las gráficas mostradas en las Figuras 2 y 3, respectivamente.

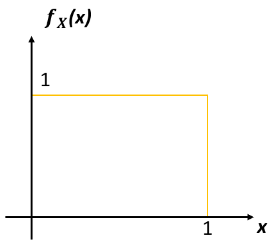

Figura 2: $f_X(x)$

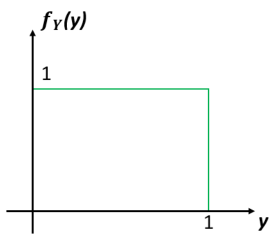

Figura 3: $f_Y(y)$

Ahora, en el proceso de convolución se voltea una de las gráficas y se empieza a desplazar como se muestra en la Figura 4:

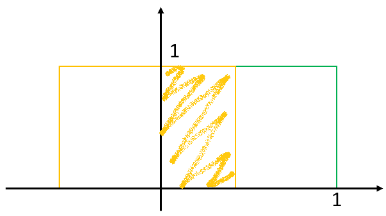

Figura 4: Convolución

Finalmente, se obtiene el resultado de la Figura 5

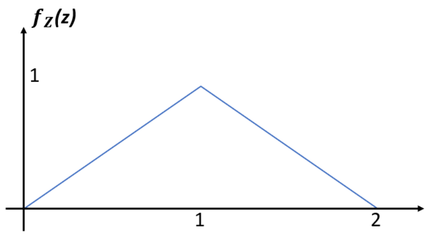

Figura 5: Resultado final

Punto 3

Notese que:

$$M_Z(s) = E\left[e^{sX}e^{sY}\right] \tag{16}$$

Puede verse como un valor esperado de función en general $E[g(X,Y)]$ con $g(X,Y) = e^{sX}e^{sY}$ que debe ser calculado como sigue:

$$M_Z(s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\ f_{X,Y}(x,y)\ dx\ dy \tag{17}$$

Aprovechándose del hecho de que X y Y son independientes ($f_{X,Y}(x,y) = f_X(x)f_Y(y)$) y de que las integrales son separables, se puede escribir:

$$M_Z(s) = \left[\int_{-\infty}^{\infty} e^{sx}\ f_X(x)\ dx\right]\left[\int_{-\infty}^{\infty} e^{sy}\ f_Y(y)\ dy\right] \tag{18}$$

De donde es directo el reconocimiento de la definición de las funciones generadoras de momentos de X y Y. Por lo cual finalmente es posible escribir:

$$M_Z(s) = M_X(s)M_Y(s) \tag{19}$$

Además de lo anterior, es posible ver la conexión con el literal **2. a)** en el hecho que un camino alternativo para calcular esta función de generación de momentos podría ser el siguiente:

$$M_Z(s) = \int_{-\infty}^{\infty} e^{sZ} f_Z(z)\, dz \tag{20}$$

Donde $z = x + y$ y $f_Z(z) = f_X * f_Y$ ya fue calculado justamente en el iteral mencionado. Esta idea es mas directa pero requiere el reemplazo de variables de la convolución.

Punto 4

Definimos T como la variable aleatoria continua del tiempo gastado en el banco y como C la variable aleatoria discreta del cajero escogido por el cliente. Usando las funciones de probabilidad conjuntas tenemos:

$$f_{T|C=1}(t \mid 1) \quad = \quad \lambda_1 e^{-\lambda_1 t} = \frac{f_{T,C}(t,1)}{f_C(1)} = \frac{f_{T,C}(t,1)}{2/3} \tag{21}$$

$$f_{T|C=2}(t \mid 2) \quad = \quad \lambda_2 e^{-\lambda_2 t} = \frac{f_{T,C}(t,2)}{f_C(2)} = \frac{f_{T,C}(t,2)}{1/3} \tag{22}$$

Y por tanto:

$$f_{T,C}(t,1) = (2/3)\, \lambda_1 e^{-\lambda_1 t} \tag{23}$$

$$f_{T,C}(t,2) = (1/3)\, \lambda_2 e^{-\lambda_2 t} \tag{24}$$

Sabiendo que solo hay 2 cajeros es posible encontrar la función de probabilidad marginal de T sumando sobre los cajeros como sigue:

$$f_T(t) = (2/3)\, \lambda_1 e^{-\lambda_1 t} + (1/3)\, \lambda_2 e^{-\lambda_2 t} \tag{25}$$

De donde la función generadora de momentos ya puede encontrarse directamente resolviendo la integral de la definición:

$$M_T(s) = \int_{-\infty}^{\infty} \left[(2/3)\,\lambda_1 e^{-\lambda_1 t} + (1/3)\,\lambda_2 e^{-\lambda_2 t}\right] e^{st} dt \quad (26)$$

$$M_T(s) = \left[\int_{-\infty}^{\infty} (2/3)\,\lambda_1 e^{-\lambda_1 t} e^{st} dt\right] + \left[\int_{-\infty}^{\infty} (1/3)\,\lambda_2 e^{-\lambda_2 t} e^{st} dt\right] \quad (27)$$

$$M_T(s) = (2/3)\,\frac{\lambda_1}{\lambda_1 - s} + (1/3)\,\frac{\lambda_2}{\lambda_2 - s} \quad (28)$$

$$M_T(s) = (2/3)\,\frac{2}{2 - s} + (1/3)\,\frac{6}{6 - s} \quad (29)$$

$$M_T(s) = \frac{(4/3)}{2 - s} + \frac{2}{6 - s} \quad (30)$$

Punto 5

Sean X_i con $i = 1, \ldots, n$ variables aleatorias iid, con distribución uniforme entre 2 y 4.

Sea $Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$, con $S_n = \sum_{i=1}^{n} X_i$, $\mu = E[X_i]$ y $\sigma^2 = \mathrm{var}(X_i)$. Como se sabe que cada variable X_i sigue una distribución uniforme entre 2 y 4, la media y la varianza se pueden calcular teóricamente como

$$\begin{gathered} \mu = E[X_i] = \frac{2+4}{2} = 3 \\ \sigma^2 = \mathrm{var}(X_i) = \frac{(4-2)^2}{12} = \frac{1}{3} \end{gathered} \quad (31)$$

Partiendo de estos resultados, se generaron 5000 realizaciones de la variable aleatoria X_i con $i = 1, \ldots, n$, y n el número de variables aleatorias iid con distribución uniforme entre 2 y 4.

Para cada una de las variables aleatorias X_i generadas se calculó el Y_n correspondiente y se graficó su histograma para diferentes valores de n. Los valores de n probados fueron $1, 2, 3, 5, 10, 100, 1000$ y 100000. En la figura 6 se pueden observar los histogramas de la variable aleatoria Y_n para estos valores de n.

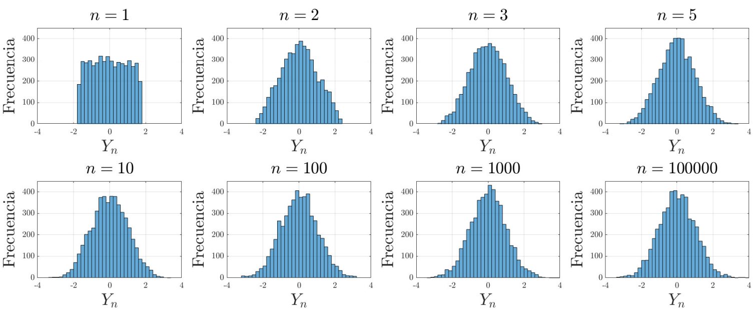

Figura 6: Histogramas de la variable aleatoria Y_n para diferentes valores de n

Como se puede notar, a medida que el valor de n aumentó, el histograma de Y_n empezó a tomar una forma cada vez más cercana al histograma de una variable aleatoria que sigue una distribución normal estándar (media igual a cero y varianza igual a 1). Se puede notar que el histograma de la variable aleatoria Y_n empezó a converger rápido al histograma normal estándar, pues a partir de $n = 2$ ya empezó a tomar una forma similar a la de la campana de Gauss característica de la distribución normal (estándar). También se puede notar que aproximadamente alrededor de $n = 100$, la forma del histograma de Y_n ya empezó a ser prácticamente idéntica a la del histograma normal estándar, y para ns mayores a este fue aún más similar.

Se puede concluir que el experimento realizado en este punto retornó una demostración clara del teorema del límite central, pues a medida que se sumaron más variables aleatorias independientes que además eran idénticamente distribuidas para generar un S_n específico, la variable aleatoria $Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$ se acercó cada vez más a seguir una distribución normal estándar, que es exactamente lo que indica el teorema. En este caso se observó una convergencia relativamente rápida a la distribución normal estándar, y esto se puede deber al hecho de que la forma de la distribución uniforme que caracterizaba a nuestras variables aleatorias X_i no tiene una forma sustancialmente distinta a la de la campana de Gauss, en comparación con otras distribuciones como la distribución exponencial, por ejemplo. Es posible que para una colección de variables aleatorias iid que sigan una distribución cuya forma es muy distinta a la de una campana de Gauss, la convergencia en distribución de la variable aleatoria Y_n a una distribución normal estándar sea más "lenta" en este caso, en el sentido en que se deberá sumar un número mayor de variables aleatorias X_i antes de que Y_n empiece a converger en distribución a una distribución normal estándar.