

Curso : Probabilidad y Estadística
Sigla : EYP1113
Pauta : I2
Profesores : Ricardo Aravena C. (Sec 1 y Sec 3) y Ricardo Olea O. (Sec 2 y Sec 4)

Problema 1

A mediados de abril, Santiago se vio afectado por cortes parciales de agua debido a que la isoterma cero durante las lluvias estuvo entorno a los 3.000 metros, lo que implicó un escurrimiento de material en los afluentes que la empresa Aguas Andina utiliza para la captación de agua. Un investigador propone que la altura de las olas que produce el volumen de material que escurre a un afluente puede ser modelado a través de una distribución Gamma($k = 2, \nu = 0.8$). Cuando la altura de las olas supera los 2 metros, a las tomas de agua ingresa un mayor volumen de material, el cual no puede ser procesado para que sea apta para el consumo. Suponga que durante una lluvia bajo las condiciones climáticas que afectaron a Santiago a mediados de abril, se esperan cuatro olas debido al escurrimiento de material cada diez minutos y que el escurrimiento ocurre en distintos lugares de manera independiente según un proceso Poisson.

- (a) Suponga que durante un periodo de observación se produjeron 20 de estas olas, determine la probabilidad que en al menos tres de ellas ingrese un mayor volumen de material.
- (b) Si un encargado está observando este fenómeno, cuantas olas deberá observar hasta que por tercera vez se supera el volumen de material que impida un procesamiento adecuado para el consumo.
- (c) El encargado, desde que observa el fenómeno acaba de ver por tercera vez una ola producida por el escurrimiento de material a los afluentes. ¿Cuál es la probabilidad que la quinta ola ocurra dentro de los siguientes tres minutos?
- (d) Suponga que al producirse una ola que supera los 2 metros, automáticamente se cierra el paso del agua al canal principal durante 3 minutos (para una auto-limpieza y por tanto no es posible observar pulsos), determine la probabilidad que el siguiente pulso se observe después de 8 min del pulso inicial.

Solución

- (a) Tenemos que una ola supere los 2 metros en su altura Y ocurre con probabilidad:

$$\text{[0.3 Ptos]} \quad P(Y > 2) = \sum_{y=0}^{2-1} \frac{(0.8 \cdot 2)^y e^{-0.8 \cdot 2}}{y!} = 2.6 e^{-1.6} = 0.5249309 \quad \text{[0.3 Ptos]}$$

Si X corresponde al número de olas sobre los 2 metros, entonces

$$X \sim \text{Binomial}(n = 20, p = 0.5249309) \quad \text{[0.5 Ptos]}$$

Se pide

$$\text{[0.2 Ptos]} \quad P(X \geq 3) = 1 - P(X \leq 2) = 1 - p_X(0) - p_X(1) - p_X(2) = 0.9999125 \quad \text{[0.2 Ptos]}$$

- (b) Sea Z el número de olas observadas hasta que por tercera vez ocurre una ola que supera los 2 metros.

$$Z \sim \text{Bin-Neg}(k = 3, p = 0.5249309) \quad \text{[1.0 Ptos]}$$

Se pide $E(Z) = \frac{k}{p} = 5.715038$. [0.5 Ptos]

- (c) Tenemos que el tiempo entre olas se comporta como una variable aleatoria Exponencial($\nu = 0.4$) y entre r olas Gamma($k = r - 1, \nu = 0.4$).

Por carencia de memoria, se pide $P(W \leq 3)$ con $W \sim \text{Gamma}(k = 2, \nu = 0.4)$ **[1.0 Ptos]**

$$\textbf{[0.2 Ptos]} \quad P(W \leq 3) = 1 - \sum_{x=0}^{2-1} \frac{(3 \cdot 0.4)^x e^{-3 \cdot 0.4}}{x!} = 1 - 2.2 e^{-1.2} = 0.3373727 \quad \textbf{[0.3 Ptos]}$$

- (d) Se pide $P(T > 8)$ con $T \sim \text{Exponencial}(\nu = 0.4)$ trasladada en 3 minutos. **[0.5 Ptos]**

$$\textbf{[0.5 Ptos]} \quad P(T > 8) = e^{-0.4 \cdot (8-3)} = e^{-5 \cdot 0.4} = 0.1353353 \quad \textbf{[0.5 Ptos]}$$

Como alternativa el alumno podría citar la carencia de memoria y calcular $P(T > 5)$, asignar todo el puntaje.

+ 1 Punto Base

Problema 2

Recientemente el caudal de río Mapocho se metió por los túneles de obras en Costanera Norte, causando un colapso en la comuna de providencia. Suponga que el número de residentes (en miles) afectados por el área (en km^2) inundada es una variable aleatoria Poisson, cuyo valor esperado depende obviamente del área afectada por la salida del río y de la densidad poblacional de la comuna que es igual a 9 mil residentes por km^2 . Suponga que el caudal X que salió del río durante la inundación seguía una distribución Log-Normal con media igual a $100 \text{ m}^3/\text{seg}$ y c.o.v del 45 %, y el área afectada es igual a $X^2/100$. ¿Cuál sería el c.o.v. del número de residentes afectados incondicional al caudal?

Solución

Del enunciado tenemos que

$$\text{[1.0 Ptos]} \quad Y | X = x \sim \text{Poisson}(0.09 \cdot x^2) \quad \text{y} \quad X \sim \text{Log-Normal}(\lambda, \zeta) \quad \text{[0.5 Ptos]}$$

donde

$$\text{[0.5 Ptos]} \quad \zeta = \sqrt{\ln(1 + 0.45^2)} = 0.4294214 \quad \text{y} \quad \lambda = \ln(100) - 0.5 \zeta^2 = 4.512969 \quad \text{[0.5 Ptos]}$$

$$\text{Se pide } \delta_Y = \frac{\sigma_Y}{\mu_Y}. \quad \text{[0.5 Ptos]}$$

utilizando los teoremas de esperanza condicional se tiene que

$$\begin{aligned} \mu_Y &= E(Y) = E[E(Y | X)] \quad \text{[0.5 Ptos]} \\ &= E(0.09 \cdot X^2) = 0.09 \cdot E(X^2) = 0.09 \cdot \exp(2\lambda + 2\zeta^2) \quad \text{[0.5 Ptos]} \\ &= 1082.25 \quad \text{[0.3 Ptos]} \end{aligned}$$

y

$$\begin{aligned} \sigma_Y^2 &= \text{Var}(Y) = \text{Var}[E(Y | X)] + E[\text{Var}(Y | X)] \quad \text{[0.5 Ptos]} \\ &= \text{Var}(0.09 \cdot X^2) + E(0.09 \cdot X^2) \quad \text{[0.3 Ptos]} \\ &= 0.09^2 \cdot \text{Var}(X^2) + 0.09 \cdot E(X^2) \quad \text{[0.2 Ptos]} \\ &= 0.09^2 \cdot E(X^4) - 0.09^2 \cdot E^2(X^2) + 0.09 \cdot E(X^2) \quad \text{[0.2 Ptos]} \\ &= 0.09^2 \cdot \exp(4\lambda + 8\zeta^2) - 0.09^2 \cdot \exp(4\lambda + 4\zeta^2) + 0.09 \cdot \exp(2\lambda + 2\zeta^2) \quad \text{[0.2 Ptos]} \\ &= 1278855 \quad \text{[0.1 Ptos]} \end{aligned}$$

Por lo tanto

$$\delta_Y = \frac{\sqrt{1278855}}{1082.25} = 1.04492 \quad \text{[0.2 Ptos]}$$

+ 1 Punto Base

Problema 3

Durante abril los contribuyentes debieron realizar su declaración anual de impuestos. Para hacer más simple el trámite, el Servicio de Impuestos Internos (SII) propone una declaración lista para enviar. Un problema que usualmente ocurre, es que las empresas van ingresando nueva información al SII, que podría implicar que una propuesta enviada por un contribuyente sea rechazada por el SII. Si X representa la proporción de contribuyentes que envían la propuesta que el sistema ofrece e Y un valor en el intervalo $(-X, X)$, que indica como se distribuyen los contribuyentes: rechazados o aceptados. Por ejemplo un valor de Y igual a cero nos indicaría que la mitad fue rechazada y la otra aceptada. Si el comportamiento conjunto entre X e Y es Uniforme (constante) sobre el soporte conjunto, determine la covarianza entre X e Y . ¿Son independientes?

Solución

El soporte conjunto $\Theta_{X,Y} = \{(x, y) \in \mathbb{R}^2 \mid 0 < x < 1, -x < y < +x\}$ y del enunciado se tiene que

$$f_{X,Y}(x, y) = k, \quad \forall (x, y) \in \Theta_{X,Y} \quad [0.5 \text{ Ptos}]$$

Para determinar el valor de k , integramos sobre todo el soporte e igualamos a uno:

$$[0.5 \text{ Ptos}] \quad \int_0^1 \int_{-x}^{+x} k \, dy \, dx = k \int_0^1 2x \, dx = k \cdot x \Big|_0^1 = 1 \rightarrow k = 1 \quad [1.0 \text{ Ptos}]$$

Se pide

$$\text{Cov}(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y) \quad [0.5 \text{ Ptos}]$$

donde

$$E(X \cdot Y) = \int_0^1 \int_{-x}^{+x} x \cdot y \, dy \, dx = \int_0^1 x \cdot \frac{y^2}{2} \Big|_{-x}^{+x} dx = \int_0^1 x \cdot 0 \, dx = 0 \quad [0.5 \text{ Ptos}]$$

$$E(X) = \int_0^1 \int_{-x}^{+x} x \, dy \, dx = \int_0^1 x \cdot y \Big|_{-x}^{+x} dx = \int_0^1 2x^2 \, dx = \frac{2}{3} \quad [0.5 \text{ Ptos}]$$

$$E(Y) = \int_0^1 \int_{-x}^{+x} y \, dy \, dx = \int_0^1 \frac{y^2}{2} \Big|_{-x}^{+x} dx = \int_0^1 0 \, dx = 0 \quad [0.5 \text{ Ptos}]$$

Luego, $\text{Cov}(X, Y) = 0$. [0.5 Ptos]

Para chequear independencia, obtendremos f_X y f_Y , mediante el teorema de probabilidades totales:

$$\begin{aligned} f_X(x) &= \int_{-x}^{+x} 1 \, dy \\ &= 2x, \quad 0 < x < 1 \quad [0.5 \text{ Ptos}] \\ f_Y(y) &= \begin{cases} \int_y^1 1 \, dx & , 0 < y < 1 \\ \int_{-y}^1 1 \, dx & , -1 < y < 0 \end{cases} \\ &= \begin{cases} 1 - y & , 0 < y < 1 \\ 1 + y & , -1 < y < 0 \end{cases} \quad [0.5 \text{ Ptos}] \end{aligned}$$

Como $f_{X,Y}(x, y) \neq f_X(x) \cdot f_Y(y)$, entonces X e Y no son variables aleatorias independientes. [0.5 Ptos]

+ 1 Punto Base

Problema 4

El tiempo en minutos transcurridos entre sismos en el país se comportan según un modelo Exponencial(ν), cuyo valor esperado es igual a 144 minutos. A usted le piden simular una realización (valor) de una variable aleatoria Poisson que represente el número de sismos en t minutos.

- (a) Muestre que si $X \sim \text{Uniforme}(0, 1)$, entonces $Y = F^{-1}(X)$ distribuye Exponencial(ν), con $F(\cdot)$ igual a la función de probabilidad acumulada de una Exponencial(ν).
- (b) A partir de las siguientes realizaciones de X : 0.10, 0.26, 0.71, 0.51, 0.04, simule realizaciones exponenciales como las descritas en el enunciado. Según esta información, ¿que valor tomaría el número de sismos ocurridos en las primeras dos horas?

Solución

- (a) Tenemos que

$$Y = g(X) = \frac{-\ln(1-X)}{\nu} \rightarrow X = g^{-1}(Y) = 1 - \exp\{-\nu Y\} \quad [1.0 \text{ Ptos}]$$

$$\text{y } \Theta_Y = \mathbb{R}_0^+. \quad [0.5 \text{ Ptos}]$$

Como $g(\cdot)$ es una función invertible sobre Θ_X , entonces

$$f_Y(y) = f_X(1 - e^{-\nu y}) \left| \nu e^{-\nu y} \right| = 1 \cdot \nu e^{-\nu y} = \nu e^{-\nu y}, \quad y \geq 0 \quad [1.0 \text{ Ptos}]$$

La cual corresponde a una función de densidad Exponencial(ν). [0.5 Ptos]

- (a) Evaluando con $\nu = 1/144$ [1.0 Ptos], se tiene que:

$$Y_1 = 15.171914, \quad Y_2 = 43.359133, \quad Y_3 = 178.253907, \quad Y_4 = 102.722384, \quad Y_5 = 5.878367 \quad [1.0 \text{ Ptos}]$$

Acumulando, se observa que antes de dos horas (120 minutos), el número de sismos es dos. [1.0 Ptos]

+ 1 Punto Base

Formulario

Propiedades función $\Gamma(\cdot)$

$$(1) \quad \Gamma(k) = \int_0^\infty u^{k-1} e^{-u} du; \quad (2) \quad \Gamma(a+1) = a \Gamma(a);$$

$$(3) \quad \Gamma(n+1) = n!, \quad \text{si } n \in \mathbb{N}_0; \quad (4) \quad \Gamma(1/2) = \sqrt{\pi}$$

Propiedades función $B(\cdot, \cdot)$

$$(1) \quad B(q, r) = \int_0^1 x^{q-1} (1-x)^{r-1} dx; \quad (2) \quad B(q, r) = \frac{\Gamma(q) \Gamma(r)}{\Gamma(q+r)}$$

Igualdades

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}, \quad \sum_{k=x}^{\infty} \phi^k = \frac{\phi^x}{1-\phi} \quad \text{si } |\phi| < 1, \quad \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \exp(\lambda)$$

Propiedad distribución Gamma

$$\text{Si } T \sim \text{Gamma}(k, \nu), \text{ con } k \in \mathbb{N} \longrightarrow F_T(t) = 1 - \sum_{x=0}^{k-1} \frac{(\nu t)^x e^{-\nu t}}{x!}$$

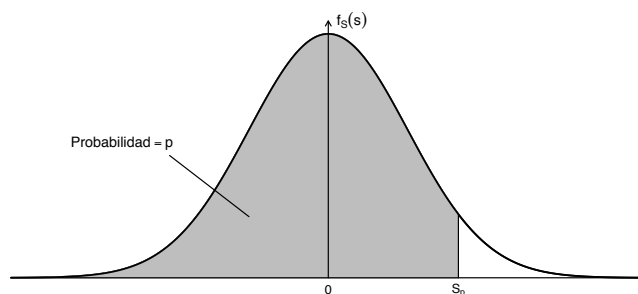
Transformación

Sea $Y = g(X)$ una función cualquiera, con k raíces:

$$f_Y(y) = \sum_{i=1}^k f_X(g_i^{-1}(y)) \cdot \left| \frac{d}{dy} g_i^{-1}(y) \right| \quad \text{o} \quad p_Y(y) = \sum_{i=1}^k p_X(g_i^{-1}(y))$$

Distribución	Densidad de Probabilidad	Θ_X	Parámetros	Esperanza y Varianza
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$	$x = 0, \dots, n$	n, p	$\mu_X = n p$ $\sigma_X^2 = n p (1-p)$ $M(t) = [p e^t + (1-p)]^n, \quad t \in \mathbb{R}$
Geométrica	$p (1-p)^{x-1}$	$x = 1, 2, \dots$	p	$\mu_X = 1/p$ $\sigma_X^2 = (1-p)/p^2$ $M(t) = p e^t / [1 - (1-p) e^t], \quad t < -\ln(1-p)$
Binomial-Negativa	$\binom{x-1}{r-1} p^r (1-p)^{x-r}$	$x = r, r+1, \dots$	r, p	$\mu_X = r/p$ $\sigma_X^2 = r(1-p)/p^2$ $M(t) = \{p e^t / [1 - (1-p) e^t]\}^r, \quad t < -\ln(1-p)$
Poisson	$\frac{(\nu t)^x e^{-\nu t}}{x!}$	$x = 0, 1, \dots$	ν	$\mu_X = \nu t$ $\sigma_X^2 = \nu t$ $M(t) = \exp \left[\lambda (e^t - 1) \right], \quad t \in \mathbb{R}$
Exponencial	$\nu e^{-\nu x}$	$x \geq 0$	ν	$\mu_X = 1/\nu$ $\sigma_X^2 = 1/\nu^2$ $M(t) = \nu / (\nu - t), \quad t < \nu$
Gamma	$\frac{\nu^k}{\Gamma(k)} x^{k-1} e^{-\nu x}$	$x \geq 0$	k, ν	$\mu_X = k/\nu$ $\sigma_X^2 = k/\nu^2$ $M(t) = [\nu / (\nu - t)]^k, \quad t < \nu$
Normal	$\frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right]$	$-\infty < x < \infty$	μ, σ	$\mu_X = \mu$ $\sigma_X^2 = \sigma^2$ $M(t) = \exp(\mu t + \sigma^2 t^2 / 2), \quad t \in \mathbb{R}$
Log-Normal	$\frac{1}{\sqrt{2\pi}(\zeta x)} \exp \left[-\frac{1}{2} \left(\frac{\ln x - \lambda}{\zeta} \right)^2 \right]$	$x \geq 0$	λ, ζ	$\mu_X = \exp \left(\lambda + \frac{1}{2} \zeta^2 \right)$ $\sigma_X^2 = \mu_X^2 (e^{\zeta^2} - 1)$ $E(X^r) = e^{r\lambda} M_Z(r\zeta), \text{ con } Z \sim \text{Normal}(0,1)$
Uniforme	$\frac{1}{(b-a)}$	$a \leq x \leq b$	a, b	$\mu_X = (a+b)/2$ $\sigma_X^2 = (b-a)^2/12$ $M(t) = [e^t b - e^t a] / [t(b-a)], \quad t \in \mathbb{R}$
Beta	$\frac{1}{B(q, r)} \frac{(x-a)^{q-1} (b-x)^{r-1}}{(b-a)^{q+r-1}}$	$a \leq x \leq b$	q, r	$\mu_X = a + \frac{q}{q+r} (b-a)$ $\sigma_X^2 = \frac{q r (b-a)^2}{(q+r)^2 (q+r+1)}$
Hipergeométrica	$\frac{\binom{m}{x} \binom{N-m}{n-x}}{\binom{N}{n}}$	$\max\{0, n+m-N\} \leq x \leq \min\{n, m\}$	N, m, n	$\mu_X = n \frac{m}{N}$ $\sigma_X^2 = \left(\frac{N-n}{N-1} \right) n \frac{m}{N} \left(1 - \frac{m}{N} \right)$

Tabla Percentiles Distribución Normal Estándar



S_p	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

Curso : Probabilidad y Estadística
Sigla : EYP1113
Bonus : I2
Profesores : Ricardo Aravena C. y Ricardo Olea O.

Nombre: **Sección:** **N° Lista:**

¿Son las siguientes afirmaciones verdaderas o falsas? Justifique en ambos casos, no basta con poner V o F.

Observación: respuesta correcta +0.05 puntos en nota I2, respuesta incorrecta -0.05 puntos en nota I2.

1. La mejor forma de evaluar visualmente el ajuste de un modelo de probabilidad continuo a un conjunto de datos, es graficar sobre el histograma de densidad la función de densidad.

Justificación FALSO, los gráficos de probabilidad detectan de mejor manera un buen ajuste.

2. En simulación de variables aleatorias, la distribución que obtenemos viene definida por la semilla que se fija.

Justificación FALSO, la distribución depende de la inversa de la acumulada que se utilice y no afectan el modelo, si los valores simulados.

3. El método de la transformada inversa, nos entrega la función de distribución acumulada asociada a los datos.

Justificación FALSO, a partir de la función de distribución acumulada nos entrega valores correspondiente a percentiles asociados a la distribución simulada.

4. Para simular variables aleatorias, se usan número aleatorios uniformes(0,1) por que son valores que siempre se encuentran en el soporte del modelo que se quiere simular.

Justificación FALSO, el soporte de una variable ealtoria no tiene por que considerar el intervalo (0,1), ejemplo una Beta en el intervalo (10,20).