Probabilidad y estadística

Clase 2

Transformaciones de variables

Función de variable aleatoria

Método de la transformada inversa

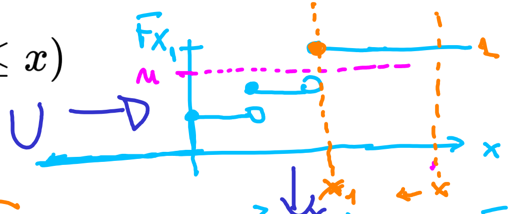
El objetivo es poder generar una variable aleatoria con una cierta distribución deseada a partir de una que conozco.

Es decir, quiero construir una variable aleatoria cuya función de distribución sea una F dada.

Método de la transformada inversa

Sea $F : \mathbb{R} \to [0,1]$ una función de distribución, existe una variable aleatoria

$$X / \boxed{F(x)} = \mathbb{P}(X \le x)$$



Definimos la inversa generalizada como:

$$\left\{ F_X^{-1}(u) = \min\{x \in \mathbb{R} : \underbrace{F_X(x)}_{\text{continuo}} \ge u\},\ u \in (0,1) \right\} = \{x : F_X(x) = u\} = F_X^{-1}(u)$$

Teorema: Si F es una función que cumple que

H)
- Es no decreciente ✓
- $\lim_{x\to-\infty} F(x) = 0$ y $\lim_{x\to\infty} F(x) = 1$ ✓
- Es continua a derecha ✓

$\Rightarrow$ F es una fun de distribución

Entonces, si defino $\boxed{X = F^{-1}(U)}$ con $U \sim \mathcal{U}(0,1)$ $\Rightarrow$ X es una v.a. con función de distribución F

genera muestras de X con fun de distr F

Ejercicio 1

$X \in \{1, \dots, 6\}$ y $P(X = i) = \frac{1}{6}$ $i = 1, \dots, 6$

Sea X el resultado de arrojar un $\overbrace{\text{dado equilibrado}}$. A partir de 1000 realizaciones de una v.a. $\underbrace{\text{uniforme en el intervalo (0,1)}}$, simular 1000 realizaciones de X.

$$\overset{U}{\downarrow}$$

$$F_X^{-1}(u) = \begin{cases} 1 & \text{si} \quad 0 \leq u < 1/6 \\ 2 & \text{si} \quad 1/6 \leq u < 2/6 \\ 3 & \dots \\ 6 & \text{si} \quad 5/6 \leq u < 1 \end{cases} \longrightarrow X$$

Ejercicio 2

$$\begin{cases} X: \text{"tiempo entre llamadas"} \\ f_X(x) = \frac{1}{5}\, e^{-1/5\, x}\ \mathbb{1}_{\{x \geq 0\}} \end{cases} \Rightarrow F_X(u) = \int_{-\infty}^{u} f_X(x)\,dx = \int_0^u \frac{1}{5} e^{-1/5\, x}\,dx = -e^{-1/5\, x}\Big|_0^u = \left(1 - e^{-1/5\, u}\right) \mathbb{1}_{\{u \geq 0\}}$$

$$F_X(x) = y \Rightarrow 1 - e^{-u/5} = y \Rightarrow 1 - y = e^{-u/5} \Rightarrow u = -5 \log(1-y)$$

El tiempo (en minutos) entre llamadas a un call center tiene una distribución exponencial de parámetro ⅕.

A partir de 1000 observaciones de una v.a. uniforme en el intervalo (0,1), simular 1000 valores de tiempos entre llamadas.

$$F_X^{-1}(y) = -5 \log(1 - y)$$

uniforme (0,1)

$\Big\{ X \,/\, F_X(x)$ sea su func. de distrib. i.e. $X \sim \mathcal{E}(1/5)$

Función de variable aleatoria

Motivación

$Y = g(X)$

conozco

En este caso, lo que conocemos es la relación entre variables aleatorias (transformación), pero sólo conocemos la distribución de una ellas.

Usos: en ML, se suele usar la transformación de variables para mejorar los resultados de ajustes y predicciones.
Algunas transformaciones más comunes son

- Log
- Exp
- Sqrt
- Inversa
- Binning

Usos en ML y DS

- Ingeniería de features: binning, log, exp, etc.
- Normalización
- Aumentación de datos. Ej: en imágenes se aplican tx como rotaciones, traslaciones, etc.
- Interpretación del modelo. Por ej. en RL cuando incluimos X2, X3, etc. como regresores.

$$Y = \alpha x_1 + \beta x_2$$

Definición

dato

¿ $F_Y(y)$?

Sea X una v.a. con función de distribución $F_X(x)$, y sea $Y = g(X)$ una función de la variable aleatoria X. El objetivo es hallar la función de Y.

Esto puede hacerse considerando que $\boxed{F_Y(y)} = P(Y \leq y) = P(g(X) \leq y)$ y desarrollando la probabilidad en términos de la v.a. X.

A este camino se lo llama **método de eventos equivalentes.**

Ejercicio 3

me hace integrar 1.

$$f_X(x) = \frac{1}{2}\,\mathbb{1}\{-1 \le x \le 1\}$$

$$1 = \int_{-\infty}^{\infty} f_X(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{2}\,\mathbb{1}\{-1 \le x \le 1\}\,dx = \int_{-1}^{1} \frac{1}{2}\,dx = \frac{1}{2}\cdot 2$$

Sea X~U(-1,1), y sea $Y=X^2$. Hallar la función de densidad de Y

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(|X| \le \sqrt{y}) = P(-\sqrt{y} \le X \le \sqrt{y})$$

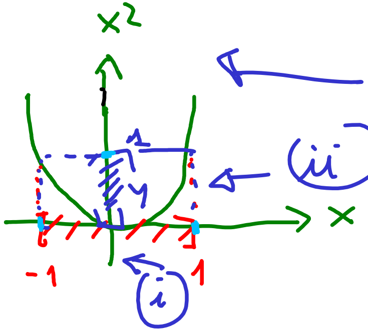
$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2}\,\mathbb{1}\{-1 \le x \le 1\}\,dx = \begin{cases} \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2}\,dx = \frac{2\sqrt{y}}{2} = \sqrt{y} \\ 0 \\ 1 \end{cases}$$

(ii) $0 \le y < 1 \rightarrow \mathbb{1}\{\ldots\} = 1$

$y \le 0$ (i)

$y \ge 1$ (iii)

$= P(Y \le 1)$



(iii)

(ii)

(i)

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{2\sqrt{y}}\,\mathbb{1}\{0 < y < 1\}$$

Ejercicio 4

En un equipo de fútbol, Juan y Esteban son los encargados de patear penales. En un torneo Juan pateó 3 penales, mientras que Esteban pateó 5. Si ambos tienen una probabilidad de 0.4 de acertar cada penal, ¿qué distribución sigue el total de penales acertados?

En un equipo de fútbol, Juan y Esteban son los encargados de patear penales. En un torneo Juan pateó 3 penales, mientras que Esteban pateó 5. Si ambos tienen una probabilidad de 0.4 de acertar cada penal, ¿qué distribución sigue el total de penales acertados?

X: # de penales que mete Juan $\rightarrow X \sim Bin(3; 0{,}4)$

Y: " Esteban $\rightarrow Y \sim Bin(5; 0{,}4)$

W: " Juan o Esteban $\rightarrow W = X + Y$

$$
\begin{aligned}
P_W(w) &= P(W=w) \quad , \quad 0 \le w \le 8 \\
&= \sum_{i=0}^{w} P(X=i ; Y=w-i) \\
&= \sum_{i=1}^{w} P(X=i) \cdot P(Y=w-i) \\
&= \sum_{i=1}^{w} \binom{3}{i} 0{,}4^{i} \, 0{,}6^{3-i} \binom{5}{w-i} 0{,}4^{w-i} \, 0{,}6^{5-w+i} \\
&= 0{,}4^{w} \, 0{,}6^{8-w} \underbrace{\sum_{i=1}^{w} \binom{3}{i}\binom{5}{w-i}}_{\binom{8}{w}}
\end{aligned}
$$

$\Rightarrow W \sim Bin(8; 0{,}4)$

$X = \sum_{i=1}^{3} B_i$ $\quad B_i \overset{iid}{\sim} Ber(0{,}4)$ $\quad \overset{ind}{e} \quad$ $Y = \sum_{j=1}^{5} B_j$ con $B_j \overset{iid}{\sim} Ber(0{,}4)$

$W = X + Y = \sum_{i=1}^{3} B_i + \sum_{j=1}^{5} B_j = \sum_{i=1}^{8} B_i \Rightarrow W \sim Bin(8; 0{,}4$

Ejercicio 5

$T \sim \mathcal{E}(1000)$

La cantidad de kilómetros que dura una cámara de bicicleta sigue una distribución exponencial de media 1000.

Hallar la distribución de la cantidad de kilómetros que se puede usar una bicicleta antes de tener que realizar la primera reparación.

$Y = \min(T_1, T_2)$ con $T_1, T_2 \sim \mathcal{E}(1000)$

$F_Y(y) = P(Y \leq y) = P(\min(T_1, T_2) \leq y) = 1 - P(\min(T_1, T_2) > y)$

$= 1 - P(T_1 > y \, ; \, T_2 > y) = \ldots$ indep

Método de transformaciones

continuas

- Sea X una v.a.c. con función de densidad $f_X(x)$,
- Sea $\boxed{Y=g(X)}$.
- g(x) es una función 1 a 1 (existe $g^{-1}(y)$)

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{d\,y} \right|$$

$F_Y(y) = P(Y \le y)$

$= P(g(x) \le y)$

$= P(x > \underbrace{g^{-1}(y)}) = 1 - F_X(g^{-1}(y))$

g es creciente

$= P(x \le \underbrace{g^{-1}(y)})$

g es decreciente

$= F_X(g^{-1}(y))$

$f_Y(y) = \frac{d}{dy} F_Y(y) = f_X(g^{-1}(y)) \cdot \frac{d\,g^{-1}(y)}{dy}$

Método del Jacobiano

Sean X_1 y X_2 son v.a **continuas** con función de densidad conjunta $f_{X_1,X_2}(x_1,x_2)$. Sean también h_1, h_2 dos func. tales que para todo (x_1,x_2) en el soporte de (X_1,X_2), $\boxed{y_1 = h_1(x_1,x_2)}$ y $\boxed{y_2 = h_2(x_1,x_2)}$ son una transformación uno a uno con **inversa** $x_1 = h_1^{-1}(y_1,y_2)$ y $x_2 = h_2^{-1}(y_1,y_2)$. Si las inversas tienen **derivadas parciales continuas** respecto de y_1 e y_2 y jacobiano J, entonces la densidad conjunta de Y_1, Y_2 será:

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(x_1,x_2)\Big|_{h_1^{-1}(y_1,y_2),h_2^{-1}(y_1,y_2)} |J| \quad \det(J)$$

$$J = \begin{pmatrix} \frac{dh_1^{-1}}{dy_1} & \frac{dh_1^{-1}}{dy_2} \\ \frac{dh_2^{-1}}{dy_1} & \frac{dh_2^{-1}}{dy_2} \end{pmatrix}$$

Otrz fonz:

$$f_{Y_1 Y_2}(y_1, y_2) = \frac{f_{X_1 X_2}(x_1, x_2)}{\det(\tilde{J})} \Bigg|_{\substack{x_1 = h_1^{-1}(y_1, y_2) \\ x_2 = h_2^{-1}(y_1, y_2)}}$$

$$\tilde{J} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} \end{bmatrix}$$

Ejercicio 6

$\to$ indep $\Rightarrow$ corr $= 0$

(para la Normal corr $= 0 \Rightarrow$ indep)

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{\mu}, \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\Sigma = I}\right)$$

Sean $\overbrace{X_1, X_2 \overset{i.i.d}{\sim} \mathcal{N}(0,1)}^{\text{|||}}$ y sean $\begin{pmatrix} U \\ V \end{pmatrix} = \overbrace{\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ sin(\theta) & \cos(\theta) \end{pmatrix}}^{\substack{\text{matriz de rotación} \\ R}} \begin{pmatrix} X \\ Y \end{pmatrix}$. Hallar $f_{U,V}(u,v)$

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}^{-1} \begin{pmatrix} U \\ V \end{pmatrix}}_{h^{-1}(u,v)} = \underbrace{\begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix}}_{R^{-1}} \begin{pmatrix} U \\ V \end{pmatrix}$$

$$J = \begin{pmatrix} \frac{\partial h_1^{-1}}{\partial u} & \frac{\partial h_1^{-1}}{\partial v} \\ \frac{\partial h_2^{-1}}{\partial u} & \frac{\partial h_2^{-1}}{\partial v} \end{pmatrix} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{pmatrix} \Rightarrow |J| = \cos^2\theta + \operatorname{sen}^2\theta = 1$$

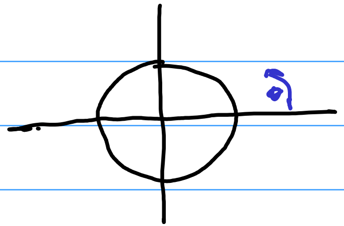
$$f_{U,V}(u,v) = f_{X,Y}(x,y) \Big|_{\binom{x}{y} = R^{-1}\binom{u}{v}} = \frac{1}{2\pi} e^{\binom{x}{y}^T I \binom{x}{y}} \Big|_{R^{-1}\binom{u}{v}} = \frac{1}{2\pi} e^{\binom{u}{v}^T (R^{-1})^T I R^{-1} \binom{u}{v}}$$

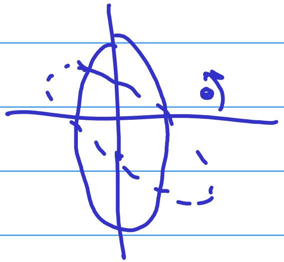
$$f_{U,V}(u,v) = \frac{1}{2\pi} e^{-\begin{pmatrix} u \\ v \end{pmatrix}^T \overbrace{(R^{-1})^T}^{=R} I R^{-1} \begin{pmatrix} u \\ v \end{pmatrix}}$$

$$\left\{ \begin{array}{l} R \text{ es ortonormal} \\ \Rightarrow R^T = R^{-1} \end{array} \right.$$

$$= \frac{1}{2\pi} e^{-\begin{pmatrix} u \\ v \end{pmatrix}^T R R^{-1} \begin{pmatrix} u \\ v \end{pmatrix}}$$

$$= \frac{1}{2\pi} e^{-\begin{pmatrix} u \\ v \end{pmatrix}^T I \begin{pmatrix} u \\ v \end{pmatrix}} \Rightarrow \begin{pmatrix} U \\ V \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, I \right)$$





$\sigma_2^2 > \sigma_1^2$

Ejercicio 7

Sean $X_1, X_2 \overset{i.i.d}{\sim} \mathcal{E}(\lambda)$ y sean $U = X_1 + X_2$ y $V = \frac{X_1}{X_1+X_2}$. Hallar $f_{U,V}(u,v)$. ¿Qué puede decir al respecto?

Sean $X_1, X_2 \overset{i.i.d}{\sim} \mathcal{E}(\lambda)$ y sean $U = \underbrace{X_1 + X_2}_{h_1}$ y $V = \underbrace{\frac{X_1}{X_1+X_2}}_{h_2}$. Hallar $f_{U,V}(u,v)$. ¿Qué puede decir al respecto?

$\Rightarrow U \sim \Gamma(2,\lambda)$ y $V \sim Unif(0,1)$ y son indep.

$$h_1(x_1,x_2) = x_1 + x_2 = u \quad \Rightarrow \quad x_2 = u - uv = u(1-v) = h_1^{-1}(u,v)$$

$$h_2(x_1,x_2) = \frac{x_1}{\underbrace{x_1+x_2}_{=u}} = v \quad \Rightarrow \quad x_1 = uv = h_2^{-1}(u,v)$$

$$J = \begin{pmatrix} 1-v & -u \\ v & u \end{pmatrix} \qquad |\det J| = |u(1-v) + uv| = |u| = u$$

u es suma de 2 v.a. positivas $\Rightarrow u \geq 0$

$$f_{UV}(u,v) = f_{X_1X_2}(x_1,x_2)\Big|_{\substack{x_1 = uv \\ x_2 = u(1-v)}} |\det J|$$

$$= \lambda e^{-\lambda x_1} \mathbb{1}\{x_1 \geq 0\} \lambda e^{-\lambda x_2} \mathbb{1}\{x_2 \geq 0\}\Big|_{\substack{x_1 = uv \\ x_2 = u(1-v)}} u$$

$$= \lambda^2 e^{-\lambda \overbrace{(x_1+x_2)}^{u}} u \, \mathbb{1}\{uv \geq 0\} \mathbb{1}\{u(1-v) \geq 0\}$$

$$U \sim \Gamma(2,\lambda) \quad = \boxed{\lambda^2 e^{-\lambda u} u^{1} \mathbb{1}\{u \geq 0\}} \times \boxed{\mathbb{1}\{0 < v < 1\}} \Rightarrow V \sim Unif(0,1)$$

$\nu = 2$

$\lambda^2 / \Gamma(2)$ pero $\Gamma(2) = 1! = 1$

Método del Jacobiano

Observación:

Si solamente me interesa una única función de las variables, por ejemplo $U_1 = h_1(Y_1, Y_2)$ puedo “inventarme” una $U_2 = h_2(Y_1, Y_2)$ y aplicar el método del Jacobiano para encontrar la función de densidad conjunta de ambas variables.

Finalmente, calculo la densidad marginal de la v.a. deseada U_1.

Método del jacobiano generalizado (BONUS)

Si $\underline{X}$ un vector aleatorio e $\underline{Y} = g(\underline{X})$ con g tal que $g|A_i = g_i : A_i \to B$ biyectiva, continua, con derivada continua, donde $A_1, \ldots, A_k$ es una partición del $sop(X)$, entonces

$$f_{\underline{Y}}(\underline{y}) = \sum_{i=1}^{k} \frac{f_{\underline{X}}(\underline{x})\mathbf{1}\{\underline{x} \in A_i\}}{|J_{g_i}(\underline{x})|} |\underline{x} = g_i^{-1}(\underline{y})$$

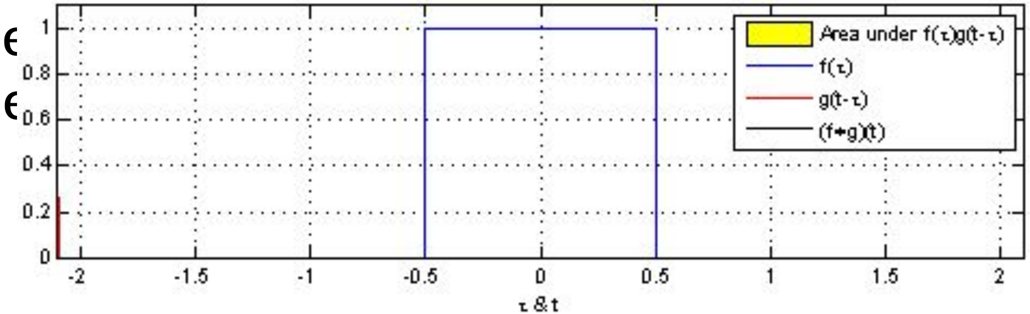
Ejercicio 2

Sea $X \sim U(-1,1)$ e $Y=X^2$. Hallar por el método del Jacobiano generalizado la función de densidad de Y.

Observación: Suma de v.a.

Sean X,Y dos v.a. con función de densidad $f_X(x)$ y $f_Y(y)$, la variable W = X+Y tiene densidad:

$$f_W(w) = f_X(x) * f_X(y)$$



Donde $*$ representa la **función convolución** definida por:

$$(f * g)(t) \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$$

f,g continuas

$$(f * g)(n) \sum_{k=-\infty}^{\infty} f(k)g(n-k)$$

f,g discretas

Función generadora de momentos

Habíamos definido a la función generadora de momentos de X como

$$M_X = \mathbb{E}[e^{tX}]$$

Teorema: Sean M_X y $M_Y$dos funciones generadoras de momentos de las v.a. X e Y, respectivamente. Si

$$M_X(t) = M_Y(t)$$

Para todo t, entonces X e Y tienen la misma distribución.

V.A Gaussianas: Proyección

Sea $\underline{X} \sim \mathcal{N}(\underline{\mu}, \Sigma)$ de dimensión n. Y sea $w \in \mathbb{R}^n$.
Definamos $Z = w^T \underline{X}$ la proyección de $\underline{X}$ en w

$$Z \sim \mathcal{N}(w^T \underline{\mu}, w^T \Sigma w)$$