

Probabilidad y Estadística

Clase 2

1. Repaso de Python

(Notebook)

Distribución normal multivariada

Función de densidad conjunta

Sea $\underline{X} = [X_1, \dots, X_n]^T$ un vector aleatorio continuo, diremos que X tiene distribución normal multivariada, si su función de densidad conjunta es de la forma

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{-\frac{1}{2} (\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu})}$$

Donde $\underline{\mu}$ se corresponde con la media de la v.a. y Σ Es la matriz de covarianza .

Notación: $\underline{X} \sim \mathcal{N}(\underline{\mu}, \Sigma)$

Distribuciones marginales

$$\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma) \Rightarrow X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$$

$$\boldsymbol{\mu} = [\mu_1, \dots, \mu_n]^T$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & cov(X_1, X_2) & \dots & cov(X_1, X_n) \\ cov(X_2, X_1) & \sigma_2^2 & \dots & cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_n, X_1) & cov(X_n, X_2) & \dots & \sigma_n^2 \end{bmatrix}$$

Distribuciones condicionales

$$\text{Sea } \mathbf{X} = [X_1, X_2]^T \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$$

$$X_1|X_2 = x_2 \sim \mathcal{N}\left(\mu_1 + \frac{\text{cov}(X_1, X_2)}{\sigma_2^2}(x_2 - \mu_2), \left(1 - \frac{\text{cov}(X_1, X_2)^2}{\sigma_1^2 \sigma_2^2}\right)\sigma_1^2\right)$$

$$X_2|X_1 = x_1 \sim \mathcal{N}\left(\mu_2 + \frac{\text{cov}(X_1, X_2)}{\sigma_1^2}(x_1 - \mu_1), \left(1 - \frac{\text{cov}(X_1, X_2)^2}{\sigma_1^2 \sigma_2^2}\right)\sigma_2^2\right)$$


Transformaciones de variables

Método de la transformada inversa

El objetivo es poder generar una variable aleatoria con una cierta distribución deseada a partir de una que “tengo a mano”.

Es decir, si tengo una función de distribución F , y quiero construir una variable aleatoria cuya función de distribución coincida con F .

Primero un ejercicio:

2.13  Sea X una variable aleatoria continua con función de distribución $F_X(x) = \mathbf{P}(X \leq x)$ estrictamente creciente en el soporte de X . Hallar expresiones para las funciones de distribución de las siguientes variables aleatorias:

- (a) $T = G(X)$, donde $G : \mathbb{R} \rightarrow \mathbb{R}$ es una función continua estrictamente creciente;
- (b) $U = F_X(X)$;
- (c) $V = G^{-1}(F_X(X))$, donde $G : \mathbb{R} \rightarrow (0, 1)$ es una función continua estrictamente creciente.

Método de la transformada inversa

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

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

Definimos la inversa generalizada como:

$$F_X^{-1}(u) = \min\{x \in \mathbb{R} : F_X(x) \geq u\}, u \in (0, 1)$$

Teorema: Si F es una función que cumple que

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

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

Función de variable aleatoria

Motivación

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

Definición

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 hacer considerando que $F_y(y) = P(Y < y) = P(g(X) < y)$, y desarrollando la probabilidad en términos de la v.a. X . A este camino se lo llama **método de sucesos equivalentes**.

Método del Jacobiano

Método de transformaciones bivariadas Suponga que Y_1 e Y_2 son variables aleatorias continuas con función de densidad conjunta $f_{Y_1, Y_2}(y_1, y_2)$ y que para todo (y_1, y_2) tal que $f_{Y_1, Y_2}(y_1, y_2) > 0$, $u_1 = h_1(y_1, y_2)$ y $u_2 = h_2(y_1, y_2)$ es una transformación uno a uno de (y_1, y_2) y (u_1, u_2)

con inversa $y_1 = h_1^{-1}(u_1, u_2)$ y $y_2 = h_2^{-1}(u_1, u_2)$.

Si las inversas tienen derivadas parciales continuas respecto a u_1 y u_2 y jacobiano J , entonces la densidad conjunta de U_1, U_2 será

$$f_{U_1, U_2}(u_1, u_2) = \frac{f_{Y_1, Y_2}(y_1, y_2)}{|J|} |h_1^{-1}, h_2^{-1}|$$

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

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

Método del Jacobiano

Observación:

Si únicamente 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 aplico 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 $g|_{A_i} = g_i : A_i \rightarrow B$ biyectiva, continua, con derivada continua, donde A_1, \dots, A_k es una partición del $\text{supp}(\underline{X})$, entonces

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

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

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

Donde $*$ representa la función convolución definida por:

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau \quad f, g \text{ continuas}$$

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

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.

Bibliografía y ejercicios

Ejercicio 1

Demostrar las propiedades de distribución condicionar y distribuciones marginales para un vector aleatorio de dimensión 2 con distribución normal bivariada.

Sugerencia: Recordar que

$$f_{X,Y}(x,y) = f_X(x)f_{Y|X=x}(y) = f_Y(y)f_{X|Y=y}(x)$$

Ejercicio 2

Sea F una función de distribución de la forma:

$$F(x) = 1 - e^{-\sqrt{x/60}}$$

Hallar una función h tal que:

$$X = h(U) \sim F, \quad \text{con} \quad U \sim \mathcal{U}(0, 1)$$

Ejercicio 3

Se arroja un dado equilibrado dos veces, sean X e Y los resultados de cada tiro, respectivamente. Hallar la función de probabilidad de $W=X+Y$

Sugerencia: hacer una tabla con los valores posibles de X e Y analizar que ocurre con W en cada caso.

Bibliografía

“Mathematical Statistics with Applications”, Dennis D. Wackerly, William Mendenhall III, Richard L. Scheaffer.