UFM UNIVERSIDAD FRANCISCO MARROQUÍN
VERITAS • LIBERTAS • JVSTITIA

FACULTAD de CIENCIAS ECONÓMICAS

Variables aleatorias continuas

Alan Reyes-Figueroa

Elements of Machine Learning (Aula 05b) 31.enero.2023

Promedio: Sea una variable aleatoria continua X con densidad f_X. La **esperanza** (*expectativa*, *valor esperado*) de X se define como

$$\mathbb{E}(X) = \int_{\mathbb{R}} t\, f_X(t)\, dt$$

en caso de que la integral exista.

Mediana: Una **mediana** de X es cualquier valor $t \in RR$ que satisface $F_X(t) = \frac{1}{2}$. Dicho de otra manera, son las preimágenes $F_X^{-1}(1/2)$.

Obs! F_X no siempre es invertible!! Denotamos $Q_X : [0,1] \to \overline{\mathbb{R}}$ a la *función de cuantiles*, la inversa generalizada de F_X:

$$Q_X(\alpha) = \inf\{x \in \mathbb{R} : \alpha \leq F_X(x)\}, \quad \text{para } 0 < \alpha < 1.$$

Moda: Una **moda** de la distribución de X es cualquier máximo local de f_X.

UFM UNIVERSIDAD FRANCISCO MARROQUÍN

Esperanza condicional

Recordemos que dadas X, Y v.a. continuas

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

Entonces

$$\mathbb{P}(a \leq X \leq b \mid Y = y) = \int_a^b f_{X|Y}(x \mid y)\,dx.$$

Como estamos condicionando a un evento con probabilidad cero, en realidad la ecuación anterior debe entenderse como un límite

$$\mathbb{P}(a \leq X \leq b \mid Y = y) = \lim_{\epsilon \to 0^+} \mathbb{P}(a \leq X \leq b \mid |Y - y| < \epsilon).$$

Definición

*Para las v.a. X y Y continuas, se define la **esperanza condicional** de X dado que Y como*

$$\mathbb{E}(X \mid Y) = \int_{\mathbb{R}} t\, f_{X|Y}(t)\,dt.$$

Proposición (Ley de la probabilidad total para esperanzas)
Sean X, Y v.a. continuas, entonces

$$\mathbb{E}(X \mid Y) = \int_{\mathbb{R}} \mathbb{E}(X \mid Y = y)\, f_Y(y)\, dy.$$

UFM UNIVERSIDAD FRANCISCO MARROQUÍN

Definición

*Sea X una v.a. continua. Definimos su **varianza** como:*

$$Var(X) = \mathbb{E}\left[(X - \mathbb{E}X)^2\right] = \int_{\mathbb{R}} (t - \mu_X)^2 f_X(t)\,dt,$$

en caso de que este valor esperado exista.

Propiedades:

- $Var(X) \geq 0$.
- $Var(aX) = a^2 Var(X)$.
- Si X_1, X_2 son independientes, entonces
$$Var(aX_2 + bX_2) = a^2 Var(X_1) + b^2 Var(X_2).$$

UFM UNIVERSIDAD FRANCISCO MARROQUÍN

Definición

*Dada dos v.a. X_1, X_2 continuas (definidas sobre el mismo espacio). Definimos su **covarianza** como:*

$$Cov(X_1, X_2) = \mathbb{E}\big[(X_1 - \mathbb{E}X_1)(X_2 - \mathbb{E}X_2)\big] = \int_{\mathbb{R}}\int_{\mathbb{R}} (s - \mu_X)(t - \mu_Y)\, f_X(s) f_Y(t)\, ds\, dt,$$

en caso de que este valor esperado exista.

Propiedades:

- $Cov(X_1, X_2) = Cov(X_2, X_1)$.
- $Cov(aX, bY) = abCov(X, Y)$.
- $Cov(aX, X) = aVar(X)$.
- Si X_1, X_2 son independientes, entonces $Cov(X_1, X_2) = 0$.

Definición

Sea X una v.a. continua. Definimos su ***entropía de Shannon*** *como:*

$$H(X) = -\int_{\mathbb{R}} f_X(t)\, \log f_X(t)\, dt.$$

En ocasiones esta es llamada *entropía diferencial*.

Comentario: No estoy seguro si existe un análogo a la entropía de Gini en el caso continuo.

Definición

Sean X, Y dos variables aleatorias, la ***entropía condicional*** *de X dado Y es*

$$H_Y(X) = \mathbb{E}H(X \mid Y) = -\int_{\mathbb{R}}\int_{\mathbb{R}} f_{X|Y}(s \mid t) \log f_{X|Y}(s \mid t) f_Y(t)\, ds\, dt.$$

Obs. No es simétrica: $H_Y(X) \neq H_X(Y)$.

Definición

Sean X, Y dos variables aleatorias, la ***información mutua*** *de X y Y está dada por*

$$I(X, Y) = H(X) - H_Y(X).$$

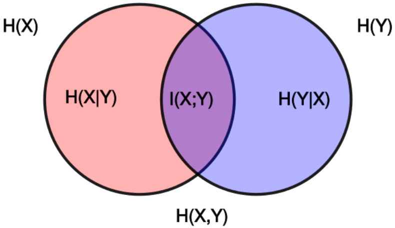
Proposición

$I(X, Y) = I(Y, X)$.

Definición

*La **entropía conjunta** de X y Y es*

$$H(X,Y) = -\int_{\mathbb{R}}\int_{\mathbb{R}} f_{X,Y}(s,t)\log f_{X,Y}(s,t)\,ds\,dt.$$



Vale el mismo diagrama que en el caso discreto.

Definición

*Sean P una distribución continua de probabilidad, con densidad $f_P(x)$. La **entropía** de P es*

$$H(P) = -\int_{\mathbb{R}} f_P(x) \log f_P(x)\, dx.$$

Definición

*Sean P, Q dos distribuciones discretas de probabilidad, la **entropía cruzada** (cross-entropy) de P y Q es*

$$H(P,Q) = -\int_{\mathbb{R}} f_P(x) \log f_Q(x)\, dx.$$

*Además, la **divergencia de Kullback-Leibler** de P y Q se define como*

$$\begin{aligned} D_{KL}(P \parallel Q) &= -\int_{\mathbb{R}} f_P(x) \log \frac{f_Q(x)}{f_P(x)}\, dx \\ &= -\int_{\mathbb{R}} f_P(x) \log f_Q(x)\, dx + \sum_x f_P(x) \log f_P(x)\, dx = H(P,Q) - H(P). \end{aligned}$$