Probabilidad y Estadística

Clase 3

Esperanza condicional

Función de regresión

Def: Sean X, Y dos v.a. discretas, se llama función de regresión a

$$\varphi(x) = \mathbb{E}[Y|X = x] = \sum_{y \in R_y} y p_{Y|X=x}(y), \quad \forall x \in R_X$$

Def: Sean X, Y dos v.a. continuas, se llama función de regresión a

$$\varphi(x) = \mathbb{E}[Y|X = x] = \int_{y \in R_y} y f_{Y|X=x}(y), \quad \forall x \in R_X$$

Observar que es función de x

Ejercicio 1

Siguiendo con los ejercicios de la clase pasada calcular la función de regresión de X|Y=y o Y|X=x según corresponda

1. La probabilidad de acertar a un blanco es ⅕. Se realizan 10 tiros independientes y se cuenta la cantidad de aciertos. Sean X la cantidad de aciertos en los 10 tiros, e Y la cantidad de aciertos en el primer tiro. Hallar la distribución de X|Y=y y Y|X=x.

2. Sean X,Y dos v.a. conjuntamente uniformes en el triángulo de vértices (0,0), (2,2), (0,2). Hallar la función de densidad de Y|X=x.

   Sean X, Y dos v.a. con función de densidad conjunta

$$f_{X,Y}(x,y) = \frac{e^{-x/2y}}{4y}\mathbf{1}\{0 < x, 1 < y < 3\}$$

   Hallar la función de densidad de X|Y=y

Esperanza condicional

Def: La variable aleatoria esperanza condicional de Y dada X se define como

$$\mathbb{E}[Y|X] = \varphi(X).$$

Además $\varphi(X)$ satisface que $\mathbb{E}[(Y - \varphi(X))\, t(X)] = 0$ para toda función medible $t : R_X \to \mathbb{R}$, tal que $\mathbb{E}[t(X)] < \infty$.

¿A qué nos recuerda esto?

$\mathbb{E}[Y|X]$ es el mejor predictor de Y basado en X (i.e. es la proyección ortogonal de Y en el espacio de funciones de X)

Propiedades

1. $\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]]$

2. $\mathbb{E}[r(X)s(Y)|X] = r(X)\mathbb{E}[s(Y)|X]$, para r, s tal que $r(X)s(X)$, $r(X)$ y $s(Y)$ tienen esperanza finita

3. $\mathbb{E}[aY_1 + bY_2|X] = a\mathbb{E}[Y_1|X] + b\mathbb{E}[Y_2|X]$

4. $\mathbb{E}[Y|X] = \mathbb{E}[Y]]$ si X y Y son independientes

Ejercicio 2

1. Siguiendo con los ejemplos del Ej. 1 calcular la esperanza condicional de $X|Y$ o $Y|X$ según corresponda
2. Para el último ejemplo, calcular $E[X]$

Varianza condicional

Def: Dada $\tau(x) = var(Y|X = x)$, se define la varianza condicional como

$$\mathbb{V}(Y|X) = \tau(X) = \mathbb{E}[Y^2|X] - \mathbb{E}[Y|X]^2$$

Propiedad: (Pitágoras)

$$var(Y) = \mathbb{E}[\mathbb{V}(Y|X)] + var(\mathbb{E}[Y|X])$$

Ejercicio 3

1. Siguiendo con los ejemplos del Ej. 1 calcular la varianza condicional de X|Y o Y|X según corresponda
2. Para el último ejemplo, calcular var[X]

Algunos comportamientos límite

LGN + TCL

Ley (débil) de los grandes números

Sea $\underline{X} = (X_1, \ldots, X_n)$ una muestra aleatoria, con Xi i.i.d con media μ y varianza σ^2. Para $\epsilon > 0$ se tiene que

$$P\{|\frac{\Sigma_t X_t}{N} - \mu| > \epsilon\} \to 0 \text{ cuando } N \to \infty$$

Es decir, a medida que crece la cantidad de muestras, el promedio tiende a la media real de la distribución.

Ejercicio 1

Simular la LGN usando Python para:

1. $X_1, \ldots, X_n \overset{i.i.d}{\sim} \mathcal{N}(5, 9)$
2. $X_1, \ldots, X_n \overset{i.i.d}{\sim} \mathcal{U}(2, 4)$
3. $X_1, \ldots, X_n \overset{i.i.d}{\sim} \mathcal{B}in(10, 1/4)$

Teorema central del límite

Sea $\underline{X} = (X_1, \ldots, X_n)$ una muestra aleatoria, con Xi i.i.d con media μ y varianza σ^2. Luego,

$$\mathbb{P}\left(\frac{\sum_{i1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \leq z\right) \overset{n\to\infty}{\to} \Phi(z), \quad \Phi(z) = \mathbb{P}(Z \leq z), Z \sim \mathcal{N}(0,1)$$

En general decimos que para n finitos diremos que

$$\mathbb{P}\left(\frac{\sum_{i1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \leq z\right) \approx \Phi(z),$$

$$\mathbb{P}\left(\sqrt{n}\frac{\bar{X}_i - \mu}{\sigma} \leq z\right) \approx \Phi(z),$$

Ejercicio 2

Mostrar Notebook

Estadística

¿Qué es la estadística?

Ahora lo que tenemos son observaciones (realizaciones) de las variables aleatorias.

El objetivo es poder hacer algún tipo de inferencia a partir de los valores observados.

Muestra aleatoria

Supongamos que tenemos un experimento aleatorio relacionado con una v.a. X. La v.a. X representa un observable del experimento aleatorio.

Los valores de X son la población de estudio, y el objetivo es saber como se comporta esa población.

Una muestra aleatoria de tamaño n, es una sucesión de n v.a **independientes** $\underline{X}_n = (X_1, \ldots, X_n)$, tal que $X_i \sim X$

Estimador

Def: Un estimador para una cierta magnitud θ (desconocida) de la distribución de cierta población es una función $\delta(\underline{X})$ de la muestra aleatoria, que devuelve un valor aproximado de θ.

Error cuadrático medio

Def: El error cuadrático medio (ECM) como $\mathbb{E}[(\delta(\underline{X}_n) - \theta)^2]$

Def: Un estimador $\delta^*(\underline{X})$es óptimo si

$ECM(\delta^*(\underline{X})) \leq ECM(\delta(\underline{X}))$ para todo $\delta(\underline{X})$.

Ejercicio 3

Se desea estimar la media de una variable con distribución N(5,9) a partir del promedio de n realizaciones. Hallar el ECM para distintos valores de n.

Estimadores de mínimos cuadrados

Estimador de mínimos cuadrados

Es común querer estimar el valor de una v.a. X a partir de una medición Y. Ejemplo: Y es una versión ruidosa de X.

Buscamos un estimador $\hat{X}$ de X tal que tenga mínimo error cuadrático medio

$$ECM = \mathbb{E}[(X - \hat{X})^2]$$

Observar que se corresponde con la distnacia asociada al p.i. canónico para v.a.

Estimador de mínimos cuadrados

En otras palabras, queremos $\hat{X} = g^*(Y)$ tal que

$$\mathbb{E}[(X - \hat{X})^2] \leq \mathbb{E}[(X - g(Y))^2] \quad \forall\, g(Y) \text{ (medible).}$$

¿Quién era $\hat{X}$? $\qquad \hat{X} = \mathbb{E}[X|Y]$

Idea de demostración: [Ejercicio]

1. Probar que el mejor estimador constante es $\mathbb{E}[X]$
2. Probar que el mejor estimador condicional es $\mathbb{E}[X|Y = y]$.
3. Dejar que Y tome todos los valores posibles (i.e. reemplazo y por Y), recupero la esperanza condicional.

Mínimos cuadrados: caso lineal

A veces obtener $\mathbb{E}[X|Y]$ puede ser muy complicado, entonces nos restringimos a los estimadores lineales.

Buscamos a, b tq $\mathbb{E}[(X-(aY+b))^2]$ sea mínima.

Resulta que $a = \frac{cov(X,Y)}{var(Y)}$ y $b = \frac{cov(X,Y)}{var(Y)}\mathbb{E}[Y] + \mathbb{E}[X]$

Ejercicio 4

Sea $X \sim U(0,1)$ e $Y = X^2$. Hallar la mejor aproximación lineal de Y basada en X. Comparar con la mejor estimación de Y basada en X.

Regresión lineal

Tengo las observaciones $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$, observaciones de dos v.a. X e Y, y queremos hallar la mejor relación lineal $Y = aX + b$.

Definiendo $A = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}$, tenemos que

$[a, b]^T = (A^T A)^{-1} A^T y$ Nuevamente, la demostración la vimos en AM.