Modelos matemáticos Capítulo 1 Robert P. Dobrow

(Intro a los P. Estocásticos con R)

Ley de Probabilidad Total

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$

B_i son mutuamente excluyentes y su unión es igual a Ω.

Regla de Bayes

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

Más general...

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_j P(A|B_j)P(B_j)}$$

Distribución Condicional

Caso discreto

$$P(Y=y|X=x) = \frac{P(X=x, Y=y)}{P(X=x)}$$

también conocida como la función de masa de probabilidad condicional, es una función de y con x fijo.

Caso continuo

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

también conocida como la función de densidad condicional es una función de y donde x es fijo.

Las densidades condicionales sirven para calcular probabilidades condicionales de la siguiente manera

$$P(Y \in R | X=x) = \int_R f_{Y|X}(y|x)\,dy \quad ; \quad R \subseteq \mathbb{R}$$

Esperanza Condicional de Y dado X=x

$$E(Y|X=x) = \begin{cases} \sum_y y\,P(Y=y|X=x), & \text{para caso discreto} \\ \int_{-\infty}^{+\infty} y\, f_{Y|X}(y|x)\,dy, & \text{para caso continuo} \end{cases}$$

Donde $E(Y|X=x)$ es una función de x

Propiedades Esperanza Condicional

1) Linealidad $\rightarrow E(aY+bZ|X=x) = aE(Y|X=x) + bE(Z|X=x)$

2) Si g una función...

$$E(g(Y)|X=x) = \begin{cases} \sum_y g(y)P(Y=y|X=x), & \text{para caso discreto} \\ \int_{-\infty}^{+\infty} g(y)\, f_{Y|X}(y|x)\,dy, & \text{para caso continuo} \end{cases}$$

3) Independencia (X & Y independientes)

$$E(Y|X=x) = E(Y)$$

4) Si $Y = g(X)$ una función de X

$$E(Y|X=x) = g(x)$$

Ley de Esperanza Total

$$E(Y) = \sum_{i=1}^{k} E(Y|A_i)P(A_i)$$

A_i son eventos que particionan a Ω

⚠ Si X & Y están conjuntamente distribuidas

$$E(Y) = \sum_{x} E(Y|X=x)P(X=x)$$

Esperanza Condicional de Y dado X

$E(Y|X)$ tiene 3 propiedades definitorias

1) $E(Y|X)$ resulta ser una variable aleatoria
2) $E(Y|X)$ es una función de X
3) $E(Y|X) = E(Y|X=x)$ siempre que $X=x$. Esto puede ser visto de otra forma como:

   Si $E(Y|X=x) = g(x)$ $\forall x$,
   Entonces $E(Y|X) = g(X)$

Ley de Esperanza Total *

$$E(Y) = E(E(Y|X))$$

Suma Aleatoria de Variables Aleatorias

$$T = X_1 + \dots + X_N = \sum_{k=1}^{N} X_k$$

Donde $X_1, X_2, \dots$ es una secuencia i.i.d con media μ, y N es una var. aleatoria con media λ.

Esperanza de suma aleatoria de var. aleatorias.

Se tiene que

$\rightarrow$ $E(T|N=n) = n\mu$

y como $\forall n$ se cumple $E(T|N=n)$,
entonces $E(T|N) = N\mu$

Ahora por ley de esperanza total *

$$E(T) = E(E(T|N)) = E(N\mu) = \mu E(N) = \mu\lambda$$

Varianza Condicional

$$Var(Y|X=x) = E((Y-\mu_x)^2|X=x)$$

$$= \begin{cases} \sum_{y} (y-\mu_x)^2 P(Y=y|X=x), & \text{para caso discreto} \\ \int_{-\infty}^{+\infty} (y-\mu_x)^2 f_{Y|X}(y|x)dy, & \text{para caso continuo} \end{cases}$$

donde $\mu_x = E(Y|X=x)$

Propiedades de la Varianza Condicional

1) $Var(Y|X=x) = E(Y^2|X=x) - (E(Y|X=x))^2$

2) $Var(aY+b|X=x) = a^2 Var(Y|X=x)$

Ley de la Varianza Total

$$Var(Y) = E(Var(Y|X)) + Var(E(Y|X))$$

Varianza de suma aleatoria de variables aleatorias

Considerese que la secuencia $X_1, X_2, \dots$ es i.i.d y tiene media μ_X y varianza σ_X^2.

Además N una var. aleatoria con media μ_N y varianza σ_N^2; donde N es independiente de cada X_i.

Donde $T = X_1 + \dots + X_N$

$$\begin{aligned} Var(T) &= E(Var(T|N)) + Var(E(T|N)) \\ &= E(N\sigma_x^2) + Var(N\mu_x) \\ &= \sigma_x^2 E(N) + \mu_x Var(N) \\ &= \sigma_x^2 \mu_N + \mu_x \sigma_N^2 \end{aligned}$$

Generalizando y conjuntando los resultados de esperanza y varianza de sumas aleatorias de variables aleatorias. Sean

- $X_1, X_2, \dots$ secuencia i.i.d. con media común μ_X y varianza σ_X^2
- N una var. aleatoia positiva entero valuada independiente de cada X_i, y con media μ_N y varianza σ_N^2
- $T = \sum_{i=1}^{N} X_i$

Se tiene que

$$E(T) = \mu_X \mu_N$$

$$Var(T) = \sigma_X^2 \mu_N + \sigma_N^2 \mu_X$$

Modelos matemáticos Capítulo 2 Robert P. Dobrow

(Intro a los P. Estocásticos con R)

Cadena de Markov (Definición):

Una cadena de Markov es una secuencia de variables aleatorias $X_0, X_1, \ldots$ que toma valores de el conjunto de estados discreto S, teniendo la propiedad de que

$$P(X_{n+1}=j \mid X_0=x_0, \ldots, X_{n-1}=x_{n-1}, X_n=x_n) = P(X_{n+1}=j \mid X_n=i)$$

para todo $x_0, \ldots, x_{n-1}$; $i, j \in S$; $n \geq 0$

Homogeneidad del tiempo

A menos que se diga lo contrario, se asume que una cadena de Markov es homogénea en el tiempo lo cual genera que

$$P(X_{n+1}=j \mid X_n=i) = P(X_1=j \mid X_0=i)$$

Matriz de transición de n-pasos

Sea $X_0, X_1, \ldots$ una cadena de Markov con matriz de transición P. La matriz P^n es la matriz de transición de n-pasos de dicha cadena, si para $n \geq 0$

$$P^n_{ij} = P(X_n=j \mid X_0=i) \quad \forall i,j$$

⚠ Nota: P^0 es la matriz identidad

Distibución inicial

Para una cadena de Markov $X_0, X_1, \ldots$ la distibución de X_0 es la distibución inicial denotada por $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots)$. Donde $P(X_0=j) = \boldsymbol{\alpha}_j \quad \forall j$

Distibución de X_n

Sea la cadena de Markov $X_0, X_1, \ldots$ con matriz de transición P y distibución inicial $\boldsymbol{\alpha}$. Donde para $n \geq 0$, la distibución de X_n es $\boldsymbol{\alpha} P^n$. Esto es

$$P(X_n=j) = (\alpha P^n)_j \quad \forall j$$

Propiedad de Markov

Sea $X_0, X_1, \dots$ una cadena de Markov. Por lo que para todo $m<n$ se cumple

$$P(X_{n+1}=j \mid X_0=i_0, \dots, X_{n-m-1}=i_{n-m-1}, X_{n-m}=i)$$

$$= P(X_{n+1}=j \mid X_{n-m}=i)$$

$$= P(X_{m+1}=j \mid X_0=i) = P_{ij}^{m+1}$$

esto para todo $i, j, i_0, \dots, i_{n-m-1}$ y $n \geq 0$

Distribución conjunta

Sea $X_0, X_1, \dots$ una cadena de Markov con matriz de transición P y distribución inicial α.

Para todo $0 \leq n_1 < n_2 < \dots < n_{k-1} < n_k$ y estados $i_1, i_2, \dots, i_{k-1}, i_k$

$$P(X_{n_1}=i_1, X_{n_2}=i_2, \dots, X_{n_{k-1}}=i_{k-1}, X_{n_k}=i_k)$$

$$= (\alpha P^{n_1})_{i_1} (P^{n_2-n_1})_{i_1 i_2} \cdots (P^{n_k-n_{k-1}})_{i_{k-1} i_k}$$

Distribuciones útiles para los ejercicios Capítulos 1 y 2 Dobrow

Distribución binomial

$$X \sim Bin(n, p)$$

$$P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$E(X) = np \,, \quad Var(X) = np(1-p)$$

Modela el número de éxitos en una secuencia de n ensayos Bernoulli independientes entre sí, con una probabilidad fija p de ocurrencia de éxito.

Distribución Geométrica

$$X \sim Geo(p)$$

$$P(X=x) = p(1-p)^x$$

$$E(X) = \frac{1}{p} \,, \quad Var(X) = \frac{1-p}{p^2}$$

Modela el número de fracasos antes de obtener el primer éxito en una sucesión de ensayos Bernoulli independientes entre sí, con una probabilidad fija p de ocurrencia de éxito.

Distribución Uniforme Continua

$$X \sim U(a,b) \sim Unif(a,b)$$

$$f_X(x) = \frac{1}{b-a}$$

$$E(X) = \frac{a+b}{2} \,, \quad Var(X) = \frac{(b-a)^2}{12}$$

Modela cuando una variable aleatoria posee la misma probabilidad constante sobre el intervalo (a,b)

Distribución Poisson

$$X \sim Poisson(\lambda)$$

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$E(X) = \lambda \,, \quad Var(X) = \lambda$$

Modela la ocurrencia de un determinado número de eventos durante cierto periodo de tiempo, esto a partir de una frecuencia de ocurrencia media λ.