

Proba y Machine Learning 2024

MW

Contents

Chapter 1

Soluciones _____ Page 2

Chapter 1

Soluciones

Ejercicio

1) Sea $X \sim N(0, 1)$. Probar que para todo $a < 1$ existe $C_a < \infty$ tal que para todo $x > 0$,

$$P(X > x) \leq C_a \frac{e^{-ax^2/2}}{\sqrt{2\pi}}$$

Se puede probar mas aun, de hecho

$$P(X > x) \leq \frac{1}{x} \frac{e^{-x^2/2}}{\sqrt{2\pi}}.$$

Solucion ejercicio 1:

$$\begin{aligned} P(X > x) &= \int_x^{+\infty} e^{-t^2/2} dt \leq \int_x^{+\infty} (1 + t^{-2}) e^{-t^2/2} dt \\ \int_x^{+\infty} (1 + t^{-2}) e^{-t^2/2} dt &= \int_x^{+\infty} e^{-t^2/2} dt - [t^{-1} e^{-t^2/2}]_x^{+\infty} - \int_x^{+\infty} e^{-t^2/2} dt = \frac{e^{-x^2/2}}{x} \end{aligned}$$

Se puede ver que $\frac{e^{-x^2/2}}{x} \leq \frac{e^{-ax^2/2}}{x}, a < 1$. Entonces:

$$P(X > x) \leq \frac{e^{-ax^2/2}}{x}$$

Si $x > \sqrt{2\pi}$, esta claro que la desigualdad vale para $C_a \geq 1$. Si $x \in I = [0, \sqrt{2\pi}]$, habria que buscar C_a tal que la desigualdad valga para todos los x en este intervalo. Para hallarlo, calculamos $\arg \min_{x \in I} \frac{e^{-ax^2/2}}{\sqrt{2\pi}}$:

$$\arg \min_{x \in I} \frac{e^{-ax^2/2}}{\sqrt{2\pi}} = \sqrt{2\pi}$$

Sabiendo que $P(X > x)$ vale a lo sumo $\frac{1}{2}$, entonces:

$$\begin{aligned} \frac{1}{2} &\leq C_a \frac{e^{-a(\sqrt{2\pi})^2/2}}{\sqrt{2\pi}} \\ C_a &\geq \frac{\sqrt{2\pi}}{2} e^{-a(\sqrt{2\pi})^2/2} \end{aligned}$$

Concluimos que con $C_a = \max\{1, \frac{\sqrt{2\pi}}{2} e^{-a(\sqrt{2\pi})^2/2}\}$ se satisface la cota pedida.

Ejercicio

2) Calcular

a) $\int_{-\infty}^{+\infty} e^{-x^2/2} dx$

b) $\int_{-\infty}^{+\infty} e^{-\frac{x_1^2 + \dots + x_n^2}{2}} dx$

Solucion ejercicio 2:

a)

$$\left(\int_{-\infty}^{+\infty} e^{-x^2/2} dx\right)^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-x^2/2} e^{-y^2/2} dx dy$$

Usamos coordenadas polares:

$$\begin{aligned} & \int_0^{+\infty} \int_0^{2\pi} r e^{-r^2/2} d\theta dr \\ & 2\pi \int_0^{+\infty} r e^{-r^2/2} dr = 2\pi \end{aligned}$$

Entonces:

$$\left(\int_{-\infty}^{+\infty} e^{-x^2/2} dx\right)^2 = 2\pi \implies \int_{-\infty}^{+\infty} e^{-x^2/2} dx = \sqrt{2\pi}$$

☺

Ejercicio

3) Sea $Z = (Z_1, \dots, Z_d) \sim N(0, \mathbb{I}_d)$ y $X = (X_1, \dots, X_d) = \frac{1}{d}(Z_1, \dots, Z_d)$.

a) Hallar la distribución de X .

b) Sea $\{v_1, \dots, v_d\}$ una base ortonormal de \mathbb{R}^d . Probar que $E[\langle X, v_i \rangle^2] = \frac{1}{d}, \forall 1 \leq i \leq d$.

c) Probar que para todo i y para todo $\epsilon > 0$

$$P(|X \cdot v_i| > \epsilon) \leq \frac{1}{d\epsilon^2}$$

Solucion ejercicio 3:

a)

Por Teorema de Cambio de Variable:

$$f_X(x_1, \dots, x_d) = f_Z(dx_1, \dots, dx_d) |d^d|,$$

donde d^d es el Jacobiano de $g(x_1, \dots, x_d) = (dx_1, \dots, dx_d) = (z_1, \dots, z_d)$. Entonces:

$$f_X(x_1, \dots, x_d) = \frac{|d^d|}{\pi^{d/2}} e^{-d^2 \langle x, x \rangle / 2},$$

la función de densidad de una Normal Multivariada con $\mu = \vec{0}$ y $\Sigma = \frac{1}{d^2} \mathbb{I}$.

b)

Antes de empezar, es muy importante darse cuenta de que para la Normal Multivariada, $Cov(X_i, X_j) = 0 \implies X_i \perp X_j$ (cosa que no necesariamente es cierta para otras distribuciones). Esto se puede demostrar viendo que la convolucion se expresa como factores de cada variable X_i por separado, que no dependen de las otras X_j . Ahora si paso a la demo:

$$E[\langle v, X \rangle^2] = E\left[\sum_{i=1}^d v_i X_i \sum_{j=1}^d v_j X_j\right] = \sum_{i=1}^d \sum_{j=1}^d E[v_i v_j X_i X_j]$$

Es facil de ver que cuando $i \neq j$, $E[v_i v_j X_i X_j] = 0$ (por independencia de las variables). Por lo tanto, consideramos unicamente el caso en el que i y j son iguales:

$$\sum_{i=1}^d E[v_i^2 X_i^2] = \sum_{i=1}^d v_i^2 E[X_i^2] = \sum_{i=1}^d v_i^2 (E^2[X_i] + Var[X_i]) = \frac{1}{d^2} \sum_{i=1}^d v_i^2 = \frac{1}{d^2}$$

Deberia dar $\frac{1}{d}$. Preguntar.

c)

Usamos el corolario de Markov:

$$P(|X \cdot v_i| > \epsilon) \leq \frac{E[\langle X, v_i \rangle^2]}{\epsilon^2} = \frac{1}{d\epsilon^2}$$

⊗

Ejercicio

4) Sean X_1, \dots, X_n iid con distribucion Normal estandar.

a) Hallar la distribucion de X_1^2

b) Hallar la distribucion de $X_1^2 + \dots + X_n^2$

Solucion ejercicio 4:

$$P(X_1^2 < t) = P(|X_1| < \sqrt{t}) = P(-\sqrt{t} < X_1 < \sqrt{t})$$

Si llamamos φ a la funcion de densidad de X_1 , entonces:

$$P(X_1^2 < t) = \int_{-\sqrt{t}}^{\sqrt{t}} \varphi(x) dx = 2 \int_0^{\sqrt{t}} \varphi(x) dx$$

Denotemos a la funcion de distribucion acumulada con Φ :

$$P(X_1^2 < t) = 2\Phi(\sqrt{t}) - 1$$

Entonces:

$$f_{X_1^2}(t) = \frac{\varphi(\sqrt{t})}{\sqrt{t}} = \frac{1}{\sqrt{2\pi t}} e^{-t/2}$$

Esta distribucion es Gama $\Gamma(\frac{1}{2}, \frac{1}{2})$

⊗

Ejercicio

5) Sea $f : [0, 1] \rightarrow \mathbb{R}$ una funcion continua. Probar que

$$\lim_{d \rightarrow \infty} \int_0^1 \int_0^1 \dots \int_0^1 f\left(\frac{x_1 + \dots + x_d}{d}\right) dx_1 dx_2 \dots dx_d = f\left(\frac{1}{2}\right)$$

Solucion ejercicio 5:

Empecemos reescribiendo la expresion a calcular. Tenemos:

$$\lim_{d \rightarrow \infty} \int_{[0,1]^d} f\left(\frac{x_1 + \dots + x_d}{d}\right) dx$$

Defino $\Omega = ([0, 1], \mathcal{B}([0, 1]), \mathcal{L}([0, 1]))$ espacio de probabilidad, y $(X_i)_{1 \leq i \leq d}$ independientes tal que $X_i(w) : [0, 1] \rightarrow [0, 1]$ es una funcion biyectiva. Notar que

$$E[X_i] = \frac{1}{2}, \forall 1 \leq i \leq d,$$

y que, por LGN, $\exists D \subset \Omega : \mu(D) = 1, \overline{X_d(D)} \rightarrow \frac{1}{2}$. Ademas, como $|[0, 1]| = |[0, 1]^d| = |[0, 1]^N|$, sabemos que existe $g : [0, 1] \rightarrow [0, 1]^d$ biyectiva. Entonces:

$$\lim_{d \rightarrow \infty} \int_{[0,1]^d} f\left(\frac{x_1 + \dots + x_d}{d}\right) dx = \lim_{d \rightarrow \infty} \int_{[0,1]} f(\overline{X_d}) 1_D$$

Intercambio integral y limite:

$$\int_{[0,1]} \lim_{d \rightarrow \infty} f(\overline{X_d}) 1_D = \int_{[0,1]} f\left(\frac{1}{2}\right) = f\left(\frac{1}{2}\right)$$

Justificacion mas extensa:

Queremos calcular $\lim_{d \rightarrow \infty} \int_{[0,1]^d} f\left(\frac{x_1 + \dots + x_d}{d}\right) dx$, y sabemos que $[0, 1] \sim [0, 1]^d$. Es decir, hay una funcion biyectiva $g : [0, 1] \rightarrow [0, 1]^d, g(t) = (t_1, \dots, t_d)$. Entonces esto equivale a lo siguiente:

$$\lim_{d \rightarrow \infty} \int_{[0,1]} f\left(\frac{\sum_{i=1}^d g(t)_i}{d}\right) dt$$

Podemos pensar a cada $g(t)_i$ como la imagen de una variable aleatoria X_i evaluada en t . Es decir, $X_i(t) = g(t)_i, \forall t \in [0, 1]$. Notar que $X_i \sim \mathcal{U}([0, 1]) \forall 1 \leq i \leq d$, y son independientes entre si.

⊗

Ejercicio

6) Sean X e Y variables aleatorias independientes con distribucion exponencial de parametro λ . Hallar la distribucion de $Z = X + Y$ y $W = \min\{X, Y\}$

Solucion ejercicio 6:

Buscamos primero la distribucion de Z :

$$\varphi_{X+Y}(t) = \varphi_X(t) \varphi_Y(t) = \frac{\lambda^2}{(\lambda - it)^2} = \frac{1}{(1 - \frac{it}{\lambda})^2} = \varphi_Z(t),$$

donde $Z \sim \Gamma(2, \lambda)$.

Ahora busquemos $W = \min\{X, Y\}$:

$$P(W \leq t) = 1 - P(W > t) = 1 - (1 - P(X \leq t))^2 = 1 - e^{-2\lambda t}$$

Por lo tanto, $W \sim \mathcal{E}(2\lambda)$.



Ejercicio

7) Para las variables X e Y del ejercicio anterior, hallar $X|X = Y$.

Solucion ejercicio 7:

Busco distribucion conjunta de $Z = X - Y$ y X . Por Teorema de Cambio de Variable:

$$\begin{aligned} f_{X,Z}(x, z) &= \lambda e^{-\lambda x} 1_{\{x>0\}} \lambda e^{-\lambda(x-z)} 1_{\{x-z>0\}} \\ f_{X,Z}(x, z) &= \lambda^2 e^{-2\lambda x} 1_{\{x>0\}} e^{\lambda z} 1_{\{x>z\}} \end{aligned}$$

Ahora calculo la densidad de Z :

$$\begin{aligned} f_Z(z) &= e^{\lambda z} 1_{\{z>0\}} \int_z^{+\infty} \lambda^2 e^{-2\lambda x} dx + e^{\lambda z} 1_{\{z \leq 0\}} \int_0^{+\infty} \lambda^2 e^{-2\lambda x} dx \\ f_Z(z) &= 1_{\{z>0\}} e^{\lambda z} \frac{\lambda}{2} e^{-2\lambda z} + 1_{\{z \leq 0\}} e^{\lambda z} \frac{\lambda}{2} \end{aligned}$$

Finalmente, calculo $X|Z = 0$:

$$f_{X|Z=0}(x) = 2\lambda e^{-2\lambda x} 1_{\{x>0\}}$$

Concluyo que $X|X = Y \sim \mathcal{E}(2\lambda)$.



Ejercicio

8) Una funcion suave $\phi : \mathbb{R} \rightarrow \mathbb{R}$ se dice convexa si $\phi''(x) \geq 0, \forall x \in \mathbb{R}$.

a) Probar que si ϕ es convexa, entonces

$$\phi(y) \geq \phi(x) + \phi'(x)(y - x), \forall x, y \in \mathbb{R}.$$

b) Probar que

$$\phi\left(\frac{x+y}{2}\right) \leq \frac{1}{2}\phi(x) + \frac{1}{2}\phi(y), \forall x, y \in \mathbb{R}.$$

Antes de pasar a la resolucion del problema, demuestro la siguiente prop que voy a usar para el item b:

Claim 1.0.1

Sea $\phi : \mathbb{R} \rightarrow \mathbb{R}$ una funcion suave, y $x, y \in \mathbb{R} : y > x$. Si $\frac{\phi(\frac{x+y}{2}) - \phi(x)}{\frac{y}{2} - \frac{x}{2}} > \frac{\phi(y) - \phi(x)}{y - x}$, entonces hay un intervalo $I \subset [x, y]$ donde ϕ' decrece.

Demostracion: Se puede deducir que bajo estas hipotesis

$$\frac{\phi(y) - \phi(\frac{y+x}{2})}{y - \frac{y+x}{2}} < \frac{\phi(\frac{x+y}{2}) - \phi(x)}{\frac{y}{2} - \frac{x}{2}}.$$

Ademas, por el Teorema del Valor Medio, $\exists c \in [x, \frac{x+y}{2}] / \phi'(c) = \frac{\phi(\frac{x+y}{2}) - \phi(x)}{\frac{y}{2} - \frac{x}{2}}$, y $\exists c' \in [\frac{x+y}{2}, y] / \phi'(c') = \frac{\phi(y) - \phi(\frac{x+y}{2})}{y - \frac{x+y}{2}}$.
Entonces,

$$\phi'(c) > \phi'(c'), c < c',$$

lo que demuestra la prop. ☺

Solucion ejercicio 8:

a)

Consideremos el desarrollo de Taylor de orden 2 de $\phi(y)$ centrado en x :

$$\phi(y) = \phi(x) + \phi'(x)(y - x) + \frac{\phi''(\eta)}{2}(y - x)^2,$$

donde η es el error. Sabiendo que $\phi''(\eta) > 0$, entonces:

$$\phi(x) + \phi'(x)(y - x) + \frac{\phi''(\eta)}{2}(y - x)^2 \geq \phi(x) + \phi'(x)(y - x)$$

Por lo tanto:

$$\phi(y) \geq \phi(x) + \phi'(x)(y - x)$$

b)

Lo encaramos por absurdo: suponemos que $\phi(\frac{x+y}{2}) > \frac{1}{2}\phi(x) + \frac{1}{2}\phi(y)$, y usamos el resultado del item a.

$$\begin{aligned} \phi(\frac{x+y}{2}) &> \frac{1}{2}\phi(x) + \frac{1}{2}\phi(y) = \phi(x) + \frac{\phi(y) - \phi(x)}{y - x} \frac{y - x}{2} \\ \frac{\phi(\frac{x+y}{2}) - \phi(x)}{\frac{y-x}{2}} &> \frac{\phi(y) - \phi(x)}{y - x} \end{aligned}$$

Por la prop enunciada, sabemos que $\exists c, c' : c < c', \phi'(c) > \phi'(c')$, pero ϕ' es monotona creciente dado que ϕ convexa. Por lo tanto, llegamos a un absurdo. ☺