PUCV - Instituto de Estadística

Distribución gaussiana multivariada

Ejercicio 1: Sea $(X, Y)' \sim \mathcal{N}(0, \Sigma)$, con

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Nos vamos a especificar los valores que puede tomar ρ en las preguntas siguientes.

1- Mostrar que si $\rho = 1$, Σ no es definida positiva. Se recuerde que Σ es definita positiva si y solo si $x'\Sigma x > 0, \forall x \neq 0$.

2- Mostrar que en este caso tenemos $\mathrm{Var}(X - Y) = 0$. Deducir que $X - Y = 0$ casi-seguro.

- Nota: cuando $\rho = 1$ estamos en una situación donde la correlación entre X y Y es perfecta. La información de X es repetida con Y. Podemos entonces resumir $(X, Y)'$ en X por ejemplo. En regresión, si los regresores son X y Y, el modelo es, mal puesto (problema de colinealidad).

3- Mostrar que los valores propios de Σ son $\lambda_1 = 1 + \rho$ y $\lambda_2 = 1 - \rho$.

4- Mostrar que los vectores propios de Σ son de la forma $V_1 = (v, v)'$ y $V_2 = (v, -v)'$, $v \in \mathbb{R}$. ¿Los V_1 y V_2 son ortogonales?

5- Vamos a considerar V_1 y V_2 con norma 1. O sea

$$V_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{y} \quad V_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Nota que dado que V_1 y V_2 son de norma 1 y ortogonales, forman una base ortonormal de $\mathbb{R}^2$.
Sea la matriz $P = (V_1 : V_2)$ (simetrica y ortogonal) de dimensiones 2×2 y $D = diag(1 + \rho, 1 - \rho)$. Mostrar que: $P\Sigma P' = D$.

6- Sea $U = P(X, Y)'$. Especificar la distribución de U. ¿Puede ser que $\rho > 1$?

7- ¿Los componentes de U son independientes?

- Nota: Hemos resumido la variabilidad de $(X,Y)'$ según dos ejes ortonormales. Nos vamos a ver que esos ejes se llaman componentes principales en ACP (análisis por componentes principales).

8- Dado que P es ortogonal, mostrar que $\Sigma = P'DP$

Ejercicio 2: Sea la matriz

$$\Gamma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & a \end{pmatrix},$$

y $X \sim \mathcal{N}(0, \Gamma)$. Nos vamos anotar X_1, X_2 y X_3 los componentes de X.

1- Mostrar que Γ es definida positiva si y solo si $a > 5$. Se recuerde que Γ es definita positiva si y solo si $x'\Gamma x > 0, \forall x \neq 0$.

2- Suponemos que $a = 5$. Mostrar que $X_1 - 2X_2 + X_3 = 0$ casi seguro. Recordamos por eso que si una variable aleatoria Z tiene una varianza cero, entonces Z es constante casi seguro.

3- Calcular el determinante de Γ cuando $a = 5$. ¿La matriz Γ es invertible? ¿Cual tipo de problemas podemos tener en este caso?

- Nota: en regresión lineal se considera la matriz inversa de los regresores. Si tenemos colinealidad casi-perfecta, o sea que existe una combinación lineal de los regresores igual a un constante, entonces el modelo es mal puesto...

Ejercicio 3: Sean las variables aleatorias reales $X_1, \ldots, X_n$, iid de distribución normal esperanza 0 y varianza 1. Sean dos variables:

$$M_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{y} \quad V_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - M_n)^2.$$

Sea una matriz A cuadrada $n \times n$, ortogonal (o sea $AA' = A'A = I_n$), que tiene sus componentes de la primera fila todos iguales a $1/\sqrt{n}$. Sea el vector $Y = AX$, donde $X = (X_1, \ldots, X_n)'$.

1- ¿Cual es la ley de Y?

2- Expresar M_n y V_n en funcción de los componentes de Y.

3- ¿Las variables M_n y V_n son independientes?

4- Dar la ley de M_n.

En las 3 próxima preguntas, el objetivo es determinar la ley de $n \times V_n$.

5- Mostrar que si $Z \sim \mathcal{N}(0,1)$, entonces $R = Z^2 \sim \chi_1^2$. Recordamos que la densidad una variable que sigue la ley χ_1^2 es dada por

$$f(r) = \frac{1}{\sqrt{2\pi}} r^{-\frac{1}{2}} \exp(-\frac{1}{2}r) 1_{(0,+\infty)}(r).$$

En particular la ley χ_1^2 corresponde a la ley $\gamma(\frac{1}{2}, \frac{1}{2})$.

6- Mostrar que si $U \sim \gamma(\theta_1, \lambda)$ y $W \sim \gamma(\theta_2, \lambda)$, y que U y W son independientes, entonces $(U + W) \sim \gamma(\theta_1 + \theta_2, \lambda)$. Con razonamiento por recurrencia se puede generalizar a la suma de más de dos variables. Recordamos que la funcción caracteristica de una variable que sigue la ley $\gamma(\theta, \lambda)$ es dada por

$$\left(\frac{\lambda}{\lambda - it} \right)^{\theta}$$

7- Dado el resultado anterior, deducir la ley de $n \times V_n$.