Análisis Multivariado: Tarea 2

Distribuciones Multivariadas

Fecha de entrega: 17 de marzo.

Distribución Normal Multivariada

1. (1 punto) Si $\mathbf{x} \sim N_p(\mu, \Sigma)$ entonces $\mathbf{Ax}$ y $\mathbf{Bx}$ son independientes si y solo si $\mathbf{A}\Sigma\mathbf{B}^T = 0$.
2. (1 punto) Si $\mathbf{x} \sim N_p(\mu, \Sigma)$ entonces $\mathbf{x}^T\Sigma^{-1}\mathbf{x} \sim \chi^2_p(\delta)$ con parámetro de no centralidad $\delta = \mu^T\Sigma^{-1}\mu$.
3. (1 punto) Mostrar que si $\mathbf{y} \sim N_p(\mathbf{0}, \sigma^2\mathbf{I})$ y $\mathbf{A}_{p\times p}$ es una matriz simétrica entonces $\mathbf{y^T A y} \sim \sigma^2\chi^2_r$ si y solo si $\mathbf{A}$ es idempotente y tal que $ran(A) = r$. (*Hint: Para la ida obtener la función característica de $y^T\mathbf{A}y$ y compararla con la de la χ^2_r. Deducir que $\mathbf{A}$ tiene r eigenvalores iguales a 1 y así $\mathbf{A}$ es idempotente de rango r.*)

Distribuciones Elípticas y Esféricas

4. (1 punto) Si $\mathbf{y} \sim S(g)$ mostrar que al considerar la transformación a coordenadas polares

$$\begin{aligned}
y_1 &= r\sin(\theta_1) \\
y_2 &= r\cos(\theta_1)\sin(\theta_2) \\
y_3 &= r\cos(\theta_1)\cos(\theta_2)\sin(\theta_3) \\
&\vdots \\
y_{p-1} &= r\cos(\theta_1)\cos(\theta_2)\cdots\cos(\theta_{p-2})\sin(\theta_{p-1}) \\
y_p &= r\cos(\theta_1)\cos(\theta_2)\cdots\cos(\theta_{p-2})\cos(\theta_{p-1})
\end{aligned}$$

la densidad del vector aleatorio $(R, \Theta_1, \ldots, \Theta_{p-1})$ está dada por

$$r^{p-1}\cos(\theta_1)^{p-2}\cos(\theta_2)^{p-3}\cdots\cos(\theta_{p-2})g(r^2)$$

5. (1 punto) Mostrar que la densidad marginal de R es

$$\frac{2\pi^{\frac{p}{2}} r^{p-1} g(r^2)}{\Gamma\left(\frac{p}{2}\right)}$$

6. (1 punto) Mostrar que las marginales de $\Theta_1, \ldots, \Theta_{p-2}$ están dadas por

$$\frac{\Gamma\left(\frac{p-i}{2}\right)\cos(\theta_i)^{p-i-1}}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p-i-1}{2}\right)}$$

   y la densidad de Θ_{p-1} está dada por

$$\frac{1}{2\pi}$$

Distribución Dirichlet

7. (1 punto) Sean $y_1, \ldots, y_p \overset{\text{ind}}{\sim} Ga(\alpha_i, \theta)$ mostrar que

$$(x_1, \ldots, x_p) = \left(\frac{y_1}{v}, \ldots, \frac{y_p}{v}\right) \sim Dir(\alpha_1, \ldots, \alpha_p).$$

   Además mostrar que

   i. $\mathbb{E}(x_i) = \frac{\alpha_i}{\alpha_0} = \tilde{\alpha}_i$.
   ii. $\mathsf{Var}(x_i) = \frac{\tilde{\alpha}_i(1-\tilde{\alpha}_i)}{(1+\alpha_0)}$.
   iii. Para $i \neq j$, $\mathsf{Cov}(x_i, x_j) = \frac{-\alpha_i\alpha_j}{\alpha_0^2(\alpha_0+1)}$.

Distribución Wishart

8. (1 punto)[1] Sea $\mathbf{M} \sim W_p(n, \Sigma)$ y $\mathbf{A}$ una matriz tal que $\text{ran}(\mathbf{A}_{q\times p}) = q$, entonces

$$(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^{\mathbf{T}})^{-1} \sim W_q\left(n-p+q, (\mathbf{A}\boldsymbol{\Sigma}^{-1}\mathbf{A}^T)^{-1}\right).$$

9. (1 punto) Si $\mathbf{y}_{p\times 1}$ es independiente de $\mathbf{M}$ y tal que $\mathbb{P}(\mathbf{y} = 0) = 0$ entonces

$$\frac{\mathbf{y}^{\mathbf{T}}\mathbf{M}\mathbf{y}}{\mathbf{y}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{y}} \sim \chi^2_n \qquad \text{y} \qquad \frac{\mathbf{y}^{\mathbf{T}}\boldsymbol{\Sigma}^{-1}\mathbf{y}}{\mathbf{y}^{\mathbf{T}}\mathbf{M}^{-1}\mathbf{y}} \sim \chi^2_{n-p+1}.$$

[1]Ejercicio opcional

Ejercicios Prácticos

10. (1 punto) Considerar la base de datos *calificaciones.txt* que consta de 88 calificaciones de 5 exámenes de las siguientes materias: probabilidad, álgebra, análisis, ecuaciones y estadística. De estos 5 exámenes, los primeros dos fueron hechos a libro cerrado y los restantes a libro abierto. ¿Se puede asumir que los datos siguen una distribución normal multivariada? De no ser así ¿Podría elegirse un subconjunto de estas variables que sí lo sean?

11. (1 punto) Para la base de datos *wine.txt* y asumiendo normalidad en las variables *Alcohol* y *Malic acid*.

    i. Probar la hipótesis de que el vino promedio difiera de 13 grados de alcohol y 2 unidades de ácido málico.

    ii. ¿Existe una diferencia para los niveles de alcohol y ácido málico para las clases 1 y 2?