

1. (25 pts) Sea U_1 y U_2 dos variables aleatorias independientes $U(0, 1)$. Suponga $\mathbf{X} = (X_1, X_2, X_3, X_4)^\top$ donde

$$X_1 = U_1, \quad X_2 = U_2, \quad X_3 = U_1 + U_2, \quad X_4 = U_1 - U_2.$$

Calcule la matrix de correlación \mathbf{R} de \mathbf{X} . ¿Cuántas componentes (PC) son de interés? Muestre que

$$\mathbf{t}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{t}_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \\ 1 \end{pmatrix},$$

son vectores propios de \mathbf{R} asociados a λ 's no triviales. Interprete las dos primeras PC obtenidas.

2. (25 pts) Considere un modelo de análisis factorial

$$\mathbf{x}_i = \boldsymbol{\mu} + \boldsymbol{\Gamma} \mathbf{z}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, n,$$

donde $\boldsymbol{\mu} \in \mathbb{R}^p$, $\mathbf{z}_i \sim N_q(\mathbf{0}, \mathbf{I})$, $\boldsymbol{\epsilon}_i \sim N_p(\mathbf{0}, \boldsymbol{\Psi})$, y $\boldsymbol{\Psi} = \text{diag}(\psi_1, \dots, \psi_p)$. Sea $\hat{\boldsymbol{\Sigma}} = \hat{\boldsymbol{\Gamma}} \hat{\boldsymbol{\Gamma}}^\top + \hat{\boldsymbol{\Psi}}$, con $\hat{\boldsymbol{\Gamma}}$ y $\hat{\boldsymbol{\Psi}}$ los MLE de $\boldsymbol{\Gamma}$ y $\boldsymbol{\Psi}$, respectivamente. Muestre que $\text{tr}(\mathbf{S} \hat{\boldsymbol{\Sigma}}^{-1}) = p$.

Recuerde que: por el Teorema de Sherman-Morrison-Woodbury, tenemos

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{C}^{-1} + \mathbf{DA}^{-1} \mathbf{B})^{-1} \mathbf{DA}^{-1}$$

3. (25 pts) Sea $\mathbf{Y} = \mathbf{XB} + \mathbf{E}$, donde \mathbf{X} es $n \times k$ de rango k , y las filas de \mathbf{E} son IID $N_p(\mathbf{0}, \boldsymbol{\Sigma})$. Sea $\mathbf{XB} = \mathbf{X}_1 \mathbf{B}_1 + \mathbf{X}_2 \mathbf{B}_2$, donde \mathbf{X}_2 es $n \times k_2$. Derive el test de razón de verosimilitudes para probar $H_0 : \mathbf{B}_2 = \mathbf{0}$

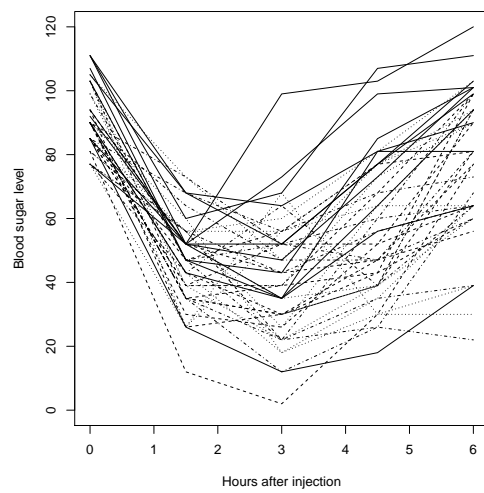
4. (25 pts) Dos mezclas de insulina, una estándar (A) y una conteniendo 5% menos protamina (B), fueron probadas en conejos. Dos grupos de 11 conejas recibieron inyecciones de insulina. Para cada tratamiento se registró el nivel de azúcar en la sangre al momento de la administración de la insulina, así como en 4 tiempos a posterior igualmente espaciados. Proponga un modelo GMANOVA y escriba un **muy breve** reporte de sus resultados.

El conjunto de datos, llamado `bloodsugar.rda`, se encuentran disponibles en la página:

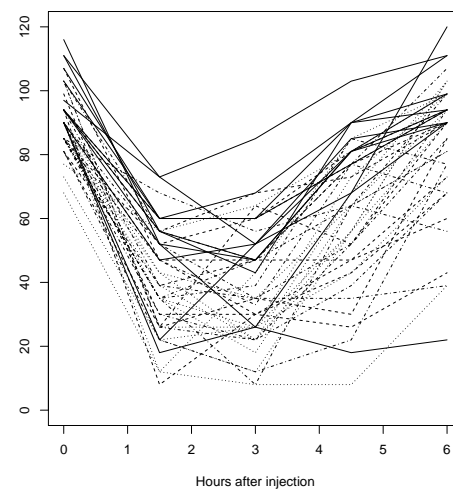
<http://fosorios.mat.utfsm.cl/teaching.html#MAT269>

A continuación se presenta el gráfico de los perfiles del nivel de azúcar en la sangre para cada una de los conejos en el estudio

Observación: La **Pregunta 4** puede ser entregada hasta el día **Jueves 13** a las 18:00 hrs.



(a) Tratamiento A



(b) Tratamiento B

Figura 1: Diagrama del nivel de azucar en la sangre versus tiempo (en horas) para los dos tratamientos en un grupo de conejos.