# Apuntes de modelos lineales

## Clases Nos. 1 al 3

**Ernesto Ponsot Balaguer**
Dr. en Estadística, MSc. en Estadística Aplicada, Ing. de Sistemas
http://webdelprofesor.ula.ve/economia/ernesto
E-mail: ernesto@ula.ve

Departamento de Estadística. Universidad de Los Andes. Mérida, Venezuela

Abril de 2012

# Contenido

## Características del curso

- El curso comprende básicamente tres temas: el modelo lineal general, el modelo de regresión y el modelo de diseño de experimentos.
- La teoría será expuesta en clases, así como algunos ejemplos que contribuyan a su comprensión. Cuando sea pertinente, se empleará el procesador estadístico R para la solución de problemas numéricos.
- Algunos ejercicios selectos se dejarán al alumno para ser resueltos como tarea y algunas de sus soluciones deberán entregarse para ser evaluadas.
- Los textos del curso son: Graybill, Franklin A. (1976) "Theory and Application of The Linear Model". Duxbury Press, Ma, EEUU y Faraway, Julian J. (2009) "Linear models with R". Chapman & Hall/CRC, Fl, EEUU.

## Evaluación

**Tabla :** Evaluación del curso

| Concepto | Tema | Aporte |
|---|---|---|
| 1° Examen | Modelo lineal general | 25 % |
| 2° Examen | Modelo de regresión | 25 % |
| 3° Examen | Modelo de diseño | 25 % |
| Ejercicios | Todos los temas | 25 % |
| Total | | 100 % |

## Matrices

1. La matriz $\boldsymbol{A}$ con $n$ filas y $m$ columnas, se denota $\boldsymbol{A}_{n\times m}$. Sus elementos son $[a_{ij}]$ para $i = 1, 2, \cdots, n$ y $j = 1, 2, \cdots, m$.
2. $\boldsymbol{I}$ se reserva para denominar a la matriz identidad ($\boldsymbol{I}_n$ o $\boldsymbol{I}_{n\times n}$ para señalar su dimensión). Se trata de una matriz cuadrada, con unos en la diagonal y ceros en el resto, esto es, $[i_{kj}] = 1$ si $k = j$ e $[i_{kj}] = 0$ si $k \neq j$.
3. Para dos matrices cualesquiera $\boldsymbol{A}_{n\times m_1}$ y $\boldsymbol{B}_{m_2\times p}$, el producto (escalar) $\boldsymbol{AB} = \boldsymbol{C}$ está definido sí y sólo sí $m_1 = m_2 = m$. $\boldsymbol{C}_{n\times p} = [c_{ij}] : c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$, para $i = 1, \cdots, n$ y $j = 1, \cdots, p$.
4. $\boldsymbol{IA} = \boldsymbol{A}$ y $\boldsymbol{AI} = \boldsymbol{A}$, para matrices $\boldsymbol{A}$ e $\boldsymbol{I}$ de dimensiones apropiadas.
5. Como en general el producto de matrices no es conmutativo, en la operación $\boldsymbol{AB}$ se dice que $\boldsymbol{A}$ pre-multiplica a $\boldsymbol{B}$, o equivalentemente, $\boldsymbol{B}$ post-multiplica a $\boldsymbol{A}$.

## Matrices

6. Si $\boldsymbol{A}_{n\times m} = [a_{ij}]$ entonces $\boldsymbol{A}' = \boldsymbol{A}^T = [a_{ji}]$ para $i = 1, \cdots, n$, $j = 1, \cdots, m$. $\boldsymbol{A}'_{m\times n}$ se denomina la matriz traspuesta (o transpuesta) de $\boldsymbol{A}$.
7. La matriz $\boldsymbol{A}$ se dice simétrica si $\boldsymbol{A} = \boldsymbol{A}'$, esto es, si $[a_{ij}] = [a_{ji}]$, para todos los $i, j$. Obviamente, la simetría sólo puede establecerse en matrices cuadradas.
8. $\boldsymbol{A}'\boldsymbol{A}$ y $\boldsymbol{A}\boldsymbol{A}'$ son matrices simétricas.
9. $(\boldsymbol{A}\boldsymbol{B})' = \boldsymbol{B}'\boldsymbol{A}'$.
10. El determinante de la matriz $\boldsymbol{A}$ se señala como $|\boldsymbol{A}|$. Sólo se aplica a matrices cuadradas y se dice que $\boldsymbol{A}$ es una matriz singular sí y sólo sí $|\boldsymbol{A}| = 0$.

## Matrices

11. Dado un conjunto de vectores $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k\}$ con $\boldsymbol{y}_i \in \mathbb{R}^n$ para todo $i$, se dice que dicho conjunto es linealmente independiente (LI) si para $\lambda_i \in \mathbb{R}\ \forall i$, se verifica que $\sum_{i=1}^{k} \lambda_i \boldsymbol{y}_i = \mathbf{0} \Leftrightarrow \lambda_i = 0\ \forall i$. En palabras, esto es que ninguno de los $\boldsymbol{y}'$s, puede ser escrito como combinación lineal de los restantes, excepto de forma trivial. En caso contrario (al menos uno de los $\lambda'$s$\neq 0$), el conjunto se denomina linealmente dependiente (LD). Todo conjunto de vectores que incluya al vector $\mathbf{0}$ se considera LD.
12. El rango de la matriz $\boldsymbol{A}_{n\times p}$ se señala como $r(\boldsymbol{A})$ y se define como el número de columnas (o filas) linealmente independientes de $\boldsymbol{A}$. Si $n \geq p \Rightarrow r(\boldsymbol{A}) \leq p$, o bien, si $n \leq p \Rightarrow r(\boldsymbol{A}) \leq n$. Si $r(\boldsymbol{A}) = p$ se dice que $\boldsymbol{A}$ es de rango completo por columnas. De igual forma, si $r(\boldsymbol{A}) = n$ se dice que $\boldsymbol{A}$ es de rango completo por filas.

## Matrices

13. Si $\boldsymbol{A}_{n\times n}$ es tal que $r(\boldsymbol{A}) = n$ entonces $\boldsymbol{A}$ es no singular, en caso contrario ($r(\boldsymbol{A}) < n$), $\boldsymbol{A}$ es singular. Si $\boldsymbol{A}_{n\times p}$ es de rango completo por columnas, entonces $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{B}_{p\times p}$ es no singular.
14. Si $\boldsymbol{A}$ es una matriz cuadrada no singular, entonces existe (y es única) $\boldsymbol{A}^{-1}$ tal que $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$. $\boldsymbol{A}^{-1}$ se llama la matriz inversa de $\boldsymbol{A}$. Una matriz singular no tiene inversa y sobre una matriz que no sea cuadrada no está definida.
15. Si $\boldsymbol{A}$ y $\boldsymbol{B}$ son matrices equidimensionales que tienen inversa, entonces $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.
16. Una matriz simétrica $\boldsymbol{A}_{n\times n}$ es semidefinida positiva si $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} \geq 0$, para todo $\boldsymbol{y} \in \mathbb{R}^n$ y $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = 0$, para al menos un $\boldsymbol{y} \in \mathbb{R}^n$. Se verifica que:
    a) $r(\boldsymbol{A}) < n$.
    b) $a_{ii} \geq 0 \ \forall i$.

## Matrices

17. Una matriz simétrica $\boldsymbol{A}_{n\times n}$ es definida positiva si $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} > 0$, para todo $\boldsymbol{y} \in \mathbb{R}^n$. Se verifica que:
    a) $r(\boldsymbol{A}) = n$.
    b) $a_{ii} > 0 \ \forall i$.
18. Una matriz simétrica $\boldsymbol{A}$ es definida nonegativa si es definida positiva o semidefinida positiva.
19. Una matriz $\boldsymbol{P}$ se dice ortogonal, si $\boldsymbol{P}^{-1} = \boldsymbol{P}'$.
20. Dos vectores $\boldsymbol{x}$ e $\boldsymbol{y}$ son ortogonales, sí y sólo sí, $\boldsymbol{x}'\boldsymbol{y} = 0$

**Tarea 01**

Luego de repasar los conceptos de espacio vectorial, valores y vectores propios, espacio de columnas de una matriz, entre otros, resuelva los ejercicios 1.7, 1.8 y 1.9 de Graybill (p.50-51).

## Matrices

21. **Inversa generalizada** (de Moore-Penrose, o g-inversa): debido a las condiciones que impone, la inversa de una matriz no siempre existe (la matriz puede no ser cuadrada, o serlo, pero de rango incompleto y por tanto singular). Un tipo de inversa que aporta en la solución de problemas del álgebra lineal es la siguiente: sea $\boldsymbol{A}_{m \times n}$ una matriz cualquiera. La matriz $\boldsymbol{A}^{-}_{n \times m}$ (única) se denomina la inversa generalizada de $\boldsymbol{A}$ (g-inversa de $\boldsymbol{A}$), si satisface:
    - $\boldsymbol{A}\boldsymbol{A}^{-}$ es simétrica.
    - $\boldsymbol{A}^{-}\boldsymbol{A}$ es simétrica.
    - $\boldsymbol{A}\boldsymbol{A}^{-}\boldsymbol{A} = \boldsymbol{A}$.
    - $\boldsymbol{A}^{-}\boldsymbol{A}\boldsymbol{A}^{-} = \boldsymbol{A}^{-}$.
22. Se le denomina "generalizada" ya que si $\boldsymbol{A}$ es no singular, $\boldsymbol{A}^{-} = \boldsymbol{A}^{-1}$.

## Matrices

23. $(\boldsymbol{A}')^- = (\boldsymbol{A}^-)'$.
24. $(\boldsymbol{A}^-)^- = \boldsymbol{A}$.
25. $\mathrm{r}(\boldsymbol{A}^-) = \mathrm{r}(\boldsymbol{A})$.
26. Si $\boldsymbol{A} = \boldsymbol{A}'$ entonces $\boldsymbol{A}^- = (\boldsymbol{A}^-)'$.
27. Si $\boldsymbol{A} = \boldsymbol{A}'$ (simetría) y $\boldsymbol{A} = \boldsymbol{A}^2$ (idempotencia) entonces $\boldsymbol{A}^- = \boldsymbol{A}$.
28. Si $\boldsymbol{B}_{m\times r}$ y $\boldsymbol{C}_{r\times m}$ son tales que $\mathrm{r}(\boldsymbol{B}) = \mathrm{r}(\boldsymbol{C}) = r > 0$, entonces $(\boldsymbol{B}\boldsymbol{C})^- = \boldsymbol{C}^-\boldsymbol{B}^-$.
29. $(\boldsymbol{A}'\boldsymbol{A})^- = \boldsymbol{A}^-(\boldsymbol{A}')^-$.
30. $(\boldsymbol{A}\boldsymbol{A}^-)^- = \boldsymbol{A}\boldsymbol{A}^-$ y $(\boldsymbol{A}^-\boldsymbol{A})^- = \boldsymbol{A}^-\boldsymbol{A}$.
31. Una matriz $\boldsymbol{A}^c$ tal que $\boldsymbol{A}\boldsymbol{A}^c\boldsymbol{A} = \boldsymbol{A}$ se denomina **inversa condicional** de $\boldsymbol{A}$. $\boldsymbol{A}^c$ no es única para $\boldsymbol{A}$ y cuando se adopta la convención de llamar $\boldsymbol{A}^+$ a la inversa generalizada de M-P, se le denomina $\boldsymbol{A}^-$ a la inversa condicional, o simplemente inversa generalizada (sin el apelativo M-P).

## Matrices

32. La **traza** de una matriz $\boldsymbol{A}_{n\times n} = [a_{ij}]$ se define como
$$\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}.$$
33. $\mathrm{tr}(\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{BA})$.
34. $\mathrm{tr}(a\boldsymbol{A} + b\boldsymbol{B}) = a\,\mathrm{tr}(\boldsymbol{A}) + b\,\mathrm{tr}(\boldsymbol{B})$.
35. Si $\boldsymbol{A}^2 = m\boldsymbol{A}$ entonces $\mathrm{tr}(\boldsymbol{A}) = m\,\mathrm{r}(\boldsymbol{A})$.
36. Para $\boldsymbol{A}_{m\times n} = [a_{ij}]$, $\mathrm{tr}(\boldsymbol{A}'\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{A}\boldsymbol{A}') = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}^2$.
37. $\mathrm{tr}(\boldsymbol{A}') = \mathrm{tr}(\boldsymbol{A})$.
38. $\mathrm{tr}(\boldsymbol{A}^c\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^c) = \mathrm{tr}(\boldsymbol{A}^-\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^-) = \mathrm{r}(\boldsymbol{A})$.

# Matrices

39. Sean $\boldsymbol{x}$ y $\boldsymbol{A}$ un vector (columna) y una matriz de números reales, respectivamente. La función $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} x_j$ se denomina una **forma cuadrática** y a menos que se diga otra cosa, se asumirá que la matriz $\boldsymbol{A}$ es simétrica.
40. **Derivada vectorial**: sea $\boldsymbol{x} = \left[x_1, x_2, \cdots, x_n\right]'$, entonces si $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_n)$:

$$\frac{\partial f}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

## Matrices

41. $$\frac{\partial(\boldsymbol{a}'\boldsymbol{x})}{\partial \boldsymbol{x}} = \frac{\partial(\boldsymbol{x}'\boldsymbol{a})}{\partial \boldsymbol{x}} = \boldsymbol{a}$$

42. Sea $q(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$ una forma cuadrática, entonces

$$\frac{\partial q(\boldsymbol{x})}{\partial \boldsymbol{x}} = 2\boldsymbol{A}\boldsymbol{x}$$

43. Sea $\boldsymbol{W} = [W_{ij}]$ una matriz aleatoria, esto es, una matriz cuyos elementos son variables aleatorias (o funciones de variables aleatorias), entonces $\mathrm{E}[\boldsymbol{W}] = [\ \mathrm{E}[W_{ij}]\ ]$.

## Matrices

44. Sean $\boldsymbol{A}$, $\boldsymbol{B}$ dos matrices reales apropiadas y $\boldsymbol{V}$, $\boldsymbol{W}$ dos matrices aleatorias, también apropiadas, entonces:
    - $\mathrm{E}[\boldsymbol{A}] = \boldsymbol{A}$.
    - $\mathrm{E}[\boldsymbol{A}\boldsymbol{V}] = \boldsymbol{A}\mathrm{E}[\boldsymbol{V}]$.
    - $\mathrm{E}[\boldsymbol{V}\boldsymbol{B}] = \mathrm{E}[\boldsymbol{V}]\boldsymbol{B}$.
    - $\mathrm{E}[\boldsymbol{A}\boldsymbol{V}\boldsymbol{B}] = \boldsymbol{A}\mathrm{E}[\boldsymbol{V}]\boldsymbol{B}$.
    - $\mathrm{E}[\boldsymbol{V} + \boldsymbol{W}] = \mathrm{E}[\boldsymbol{V}] + \mathrm{E}[\boldsymbol{W}]$.
45. Sean $\boldsymbol{Y}_m$ y $\boldsymbol{Z}_n$ dos vectores aleatorios con vectores de medias $\boldsymbol{\mu}$ y $\boldsymbol{\psi}$ respectivamente. Entonces: $\mathrm{Cov}[\boldsymbol{Y}, \boldsymbol{Z}] = \mathrm{E}[(\boldsymbol{Y} - \boldsymbol{\mu})(\boldsymbol{Z} - \boldsymbol{\psi})'] = \boldsymbol{\Sigma}_{m\times n} = [\sigma_{ij}]$, con $\sigma_{ij} = \mathrm{Cov}(Y_i, Z_j)$ para todos los $i, j$.

## Matrices

46 En particular,
$\mathrm{V}[\boldsymbol{Y}] = \mathrm{Cov}[\boldsymbol{Y}, \boldsymbol{Y}] = \mathrm{E}[(\boldsymbol{Y} - \boldsymbol{\mu})(\boldsymbol{Y} - \boldsymbol{\mu})'] = \boldsymbol{\Sigma}_{m\times m} = [\sigma_{ij}]$,
con

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

Con $\sigma_{ii} = \sigma_i^2 = \mathrm{V}[Y_i] \;\; \forall i$, $\sigma_{ij} = \sigma_{ji} \;\; \forall i, j$ y $\boldsymbol{\Sigma}$ es una matriz definida no negativa.

## Matrices

47. Si $\boldsymbol{A}$ y $\boldsymbol{B}$ son matrices reales, $\boldsymbol{b}$ un vector de escalares, y $\boldsymbol{Y}$, $\boldsymbol{Z}$ dos vectores aleatorios, entonces:
    - $\mathrm{Cov}[\boldsymbol{A}\boldsymbol{Y} + \boldsymbol{b}] = \boldsymbol{A}\,\mathrm{Cov}[\boldsymbol{Y}]\,\boldsymbol{A}' = \boldsymbol{A}\Sigma\boldsymbol{A}'$.
    - $\mathrm{Cov}[\boldsymbol{A}\boldsymbol{Y}, \boldsymbol{B}\boldsymbol{Z}] = \boldsymbol{A}\,\mathrm{Cov}[\boldsymbol{Y}, \boldsymbol{Z}]\,\boldsymbol{B}'$.
    - $\mathrm{Cov}[\boldsymbol{A}\boldsymbol{Y} + \boldsymbol{B}\boldsymbol{Z}] =$
      $\boldsymbol{A}\,\mathrm{V}[\boldsymbol{Y}]\,\boldsymbol{A}' + \boldsymbol{B}\,\mathrm{V}[\boldsymbol{Z}]\,\boldsymbol{B}' + \boldsymbol{A}\,\mathrm{Cov}[\boldsymbol{Y}, \boldsymbol{Z}]\,\boldsymbol{B}' + \boldsymbol{B}\,\mathrm{Cov}[\boldsymbol{Z}, \boldsymbol{Y}]\,\boldsymbol{A}'$.

## El ejercicio...

Para $\boldsymbol{A} = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$, encuentre $\boldsymbol{P}_{2\times 2} : \boldsymbol{P}'\boldsymbol{A}\boldsymbol{P} = \boldsymbol{I}$.

**Solución:**

Encontrando las raíces del polinomio característico de la matriz $\boldsymbol{A}$ ($|\boldsymbol{A} - \lambda\boldsymbol{I}| = 0$) es fácil ver que $\lambda_1 = 3 + \sqrt{2}$ y $\lambda_2 = 3 - \sqrt{2}$ son los valores propios de $\boldsymbol{A}$.

Ahora bien, como $\boldsymbol{A} = \boldsymbol{A}'$, el teorema 1.2.50 de Graybill (p. 17) asegura que existe $\boldsymbol{B} : \boldsymbol{B}'\boldsymbol{A}\boldsymbol{B} = \boldsymbol{D} = \text{Diag}(\lambda_i)$. Ahora, sea $\boldsymbol{R} = \text{Diag}[(\sqrt{\lambda_i})^{-1}]$, entonces:

$$
\begin{aligned}
\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B} &= \boldsymbol{D} \\
\boldsymbol{R}'\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B} &= \boldsymbol{R}'\boldsymbol{D} \\
\boldsymbol{R}'\boldsymbol{B}'\boldsymbol{A}\boldsymbol{B}\boldsymbol{R} &= \boldsymbol{R}'\boldsymbol{D}\boldsymbol{R} \\
(\boldsymbol{B}\boldsymbol{R})'\boldsymbol{A}\boldsymbol{B}\boldsymbol{R} &= \boldsymbol{I} \\
\therefore \boldsymbol{P} &= \boldsymbol{B}\boldsymbol{R}
\end{aligned}
$$

## El ejercicio...

Por supuesto, ahora el problema es ¿quién es $\boldsymbol{B}$?. Pues bien, $\boldsymbol{B}$ es la matriz cuyas columnas forman una **base ortonormal** a partir de los vectores propios de $\boldsymbol{A}$. Recuérdese que dos vectores son ortogonales si su producto escalar es nulo, y un vector es normal si el producto consigo mismo es unitario. Entonces una matriz apropiada $\boldsymbol{Z}$ forma una base ortonormal si $\boldsymbol{Z}'\boldsymbol{Z} = \boldsymbol{I}$.
Encontrar esta base ortonormal es algebraicamente laborioso (y tedioso): deben calcularse los vectores propios y emplear el teorema de Gram-Schmidt para su ortonormalización, por lo tanto, el ejercicio continúa con R (ver programa anexo)...

## Repaso...

48. Sea $Y$ una variable aleatoria (VA) con f.d.p. dada por:

$$f(y;\mu,\sigma^2) \;=\; \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-\mu)^2}{2\sigma^2}}, \;\; \infty < y < \infty$$

Entonces decimos que $Y \sim \mathrm{N}(\mu,\sigma^2)$.

49. $Z = (Y-\mu)/\sigma \sim \mathrm{N}(0,1)$, que se denomina **normal estándar**.

50. Sean $Z_1,\cdots,Z_n$ VA's independientes tales que $Z_i \sim \mathrm{N}(0,1)\ \forall i$, entonces:

$$X^2 \;=\; \sum_{i=1}^{n} Z_i^2 \sim \chi_n^2,$$

51. Sean $Y_1,\cdots,Y_n$ VA's independientes tales que $Y_i \sim \mathrm{N}(\mu,\sigma^2)\ \forall i$, entonces:

$$X^2 \;=\; \sum_{i=1}^{n}\left(\frac{Y_i-\overline{Y}}{\sigma}\right)^2 \sim \chi_{n-1}^2,$$

## Repaso...

52. Sean $Z \sim \mathrm{N}(0,1)$ y $U \sim \chi^2_n$ dos VA's independientes, entonces $T = Z/\sqrt{U/n} \sim \mathrm{t}_n$, que se denomina **t-Student**.
53. Sean $Y_1, \cdots, Y_n$ VA's independientes tales que $Y_i \sim \mathrm{N}(\mu, \sigma^2)\ \forall i$, entonces:

$$\begin{aligned} T &= \frac{\overline{Y} - \mu}{(\widehat{\sigma}/\sqrt{n})} \sim \mathrm{t}_{n-1}, \quad \text{donde} \\ \widehat{\sigma} &= \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{n-1}. \end{aligned}$$

54. Sean $Z \sim \mathrm{N}(0,1)$ y $U \sim \chi^2_\nu$ dos VA's independientes, y $\delta$ una constante, entonces $T = (Z + \delta)/\sqrt{U/\nu} \sim \mathrm{t}_{\nu,\delta}$ denominada la **densidad t no central**.
55. Sean $U_1 \sim \chi^2_{n_1}$ y $U_2 \sim \chi^2_{n_2}$ dos VA's independientes, entonces $F = (U_1/n_1)/(U_2/n_2) \sim \mathrm{F}_{n_1,n_2}$, que se denomina **F-Snedecor**, o simplemente la distribución F.

## Repaso...

56. Sean $U_1 \sim \chi^2_{n_1,\lambda}$ y $U_2 \sim \chi^2_{n_2}$, dos VA's independientes, entonces $W = (U_1/n_1)/(U_2/n_2) \sim \mathrm{F}_{n_1,n_2,\lambda}$ denominada la **densidad F no central**.
57. Sea $\boldsymbol{Z} = \begin{bmatrix} Z_1 & Z_2 & \cdots & Z_n \end{bmatrix}'$ un vector de VA's independientes, tales que $Z_i \sim \mathrm{N}(0,1)\ \forall i$, entonces $\boldsymbol{Z} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I})$ y su f.d.p. es:

$$\mathrm{N}(\boldsymbol{z}; \boldsymbol{0}, \boldsymbol{I}) = \frac{1}{(2\pi)^{n/2}} e^{-(1/2)\boldsymbol{z}'\boldsymbol{z}}, \quad \boldsymbol{z} \in \mathbb{R}^n$$

58. Sean $\boldsymbol{Z} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I})$, $\boldsymbol{\gamma} = \begin{bmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_n \end{bmatrix}'$ un vector de constantes y $\gamma_0$ un escalar, entonces $X = \boldsymbol{\gamma}'\boldsymbol{Z} + \gamma_0 \sim \mathrm{N}(x; \gamma_0, \boldsymbol{\gamma}'\boldsymbol{\gamma})$.

## Repaso...

59 Sean:

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

donde $\boldsymbol{Y}$ es un vector aleatorio, $\boldsymbol{\mu}$ es su vector de medias y $\boldsymbol{\Sigma}$ su matriz de varianzas y covarianzas, tal que $\mathrm{r}(\boldsymbol{\Sigma}) = n$, entonces $\boldsymbol{Y} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ si su f.d.p. es:

$$\mathrm{N}(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})}, \quad \boldsymbol{y} \in \mathbb{R}^n$$

## Repaso...

60. Sean $\boldsymbol{Y} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{B}$ una matriz de constantes y $\boldsymbol{b}$ un vector también de constantes, entonces $\boldsymbol{W} = \boldsymbol{B}\boldsymbol{Y} + \boldsymbol{b} \sim \mathrm{N}(\boldsymbol{B}\boldsymbol{\mu} + \boldsymbol{b}, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}')$.
61. Sea $\boldsymbol{Y}_n \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y $\boldsymbol{Y}_m^*$ un vector aleatorio con sólo algunos de los elementos de $\boldsymbol{Y}$ $(m < n)$, entonces tomando los elementos correspondientes de $\boldsymbol{\mu}$ para formar $\boldsymbol{\mu}^*$, y los elementos de filas y columnas correspondientes de $\boldsymbol{\Sigma}$ para formar $\boldsymbol{\Sigma}^*$, se tiene que $\boldsymbol{Y}_m^* \sim \mathrm{N}(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*)$.
62. En particular, si $\boldsymbol{Y} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ entonces $Y_i \sim \mathrm{N}(\mu_i, \sigma_{ii} = \sigma_i^2)$, para todo $i$.
63. $\boldsymbol{Y}_n \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}) \Rightarrow \boldsymbol{Y}'\boldsymbol{Y} \sim \chi_n^2$.
64. $\boldsymbol{Y}_n \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{I}) \Rightarrow \boldsymbol{Y}'\boldsymbol{Y} \sim \chi_{n,\lambda}^2$. Esta distribución se llama $\chi^2$ **no central** y $\lambda = (1/2)\boldsymbol{\mu}'\boldsymbol{\mu}$ es el parámetro de no centralidad.

## Repaso...

65. $\boldsymbol{Y}_n \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \mathrm{r}(\boldsymbol{\Sigma}) = n \Rightarrow \boldsymbol{Y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} \sim \chi^2_{n,\lambda}$, con $\lambda = (1/2)\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ es el parámetro de no centralidad.
66. $\boldsymbol{Y}_n \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{I}) \Rightarrow \boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} \sim \chi^2_k \Leftrightarrow \boldsymbol{A}' = \boldsymbol{A}, \boldsymbol{A}^2 = \boldsymbol{A}, \mathrm{r}(\boldsymbol{A}) = k$.
67. $\boldsymbol{Y}_n \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{I}) \Rightarrow \boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} \sim \chi^2_{k,\lambda}$ con $\lambda = (1/2)\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$ $\Leftrightarrow \boldsymbol{A}^2 = \boldsymbol{A}, \mathrm{r}(\boldsymbol{A}) = k$.
68. $\boldsymbol{Y}_n \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con $\mathrm{r}(\boldsymbol{\Sigma}) = n$, entonces $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y} \sim \chi^2_{k,\lambda}$ con $\lambda = (1/2)\boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$ sí y sólo sí cualquiera de las siguientes condiciones se satisface:
    - $\boldsymbol{A}\boldsymbol{\Sigma}$ es idempotente de rango $k$.
    - $\boldsymbol{\Sigma}\boldsymbol{A}$ es idempotente de rango $k$.
    - $\boldsymbol{\Sigma}$ es una c-inversa de $\boldsymbol{A}$, con $\mathrm{r}(\boldsymbol{A}) = k$.
69. Si $\boldsymbol{Y}$ es un vector aleatorio con $\mathrm{E}[\boldsymbol{Y}] = \boldsymbol{\mu}$ y $\mathrm{V}[\boldsymbol{Y}] = \boldsymbol{\Sigma}$, entonces $\mathrm{E}[\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}] = \mathrm{tr}(\boldsymbol{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\boldsymbol{A}\boldsymbol{\mu}$.

## Después del repaso...

70. (Teorema 1.5.12) Si $\boldsymbol{A}_{m\times n} : \mathrm{r}(\boldsymbol{A}) = m \Rightarrow \boldsymbol{A}^{-} = \boldsymbol{A}'(\boldsymbol{A}\boldsymbol{A}')^{-1}$, $\boldsymbol{A}\boldsymbol{A}^{-} = \boldsymbol{I}$ y si $\mathrm{r}(\boldsymbol{A}) = n \Rightarrow \boldsymbol{A}^{-} = (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'$, $\boldsymbol{A}^{-}\boldsymbol{A} = \boldsymbol{I}$.
71. (Teorema 4.5.2) Dado $\boldsymbol{Y}_n \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con $\mathrm{r}(\boldsymbol{\Sigma}) = n$, si $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A} = 0$ entonces $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ es independiente de $\boldsymbol{B}\boldsymbol{Y}$.
72. (Teorema 1.2.13) $\mathrm{r}(\boldsymbol{A}\boldsymbol{B}) \leq \mathrm{r}(\boldsymbol{A})$ y $\mathrm{r}(\boldsymbol{A}\boldsymbol{B}) \leq \mathrm{r}(\boldsymbol{B})$. Dicho en otros términos, $\mathrm{r}(\boldsymbol{A}\boldsymbol{B}) \leq \text{Mín}[\mathrm{r}(\boldsymbol{A}), \mathrm{r}(\boldsymbol{B})]$.
73. (Ejercicio 1.20) Si $\boldsymbol{A}_{m\times n}$, $\boldsymbol{B}_{m\times n}$, $\boldsymbol{A}\boldsymbol{B}' = \boldsymbol{0}$ y $\boldsymbol{B}'\boldsymbol{A} = \boldsymbol{0}$, entonces: $\boldsymbol{A}^{-}\boldsymbol{B} = \boldsymbol{0}$, $\boldsymbol{B}^{-}\boldsymbol{A} = \boldsymbol{0}$, $\boldsymbol{A}\boldsymbol{B}^{-} = \boldsymbol{0}$, $\boldsymbol{B}\boldsymbol{A}^{-} = \boldsymbol{0}$, $(\boldsymbol{B}')^{-}\boldsymbol{A}^{-} = \boldsymbol{0}$ y $(\boldsymbol{A}')^{-}\boldsymbol{B}^{-} = \boldsymbol{0}$.

## Después del repaso...

74. **Espacio vectorial (EV)**: es cualquier conjunto de vectores que incluya al $\mathbf{0}$, cerrado frente a la suma y multiplicación vectoriales, y frente a la multiplicación por un escalar.
75. El **espacio de columnas** generado por la matriz $\boldsymbol{M}$, es el espacio vectorial que producen las columnas de $\boldsymbol{M}$ consideradas como vectores. Se simboliza como $\mathcal{C}(\boldsymbol{M})$. La definición del **espacio de filas** de una matriz $\mathcal{R}(\boldsymbol{M})$ es completamente análoga.
76. El **rango** de $\mathcal{C}(\boldsymbol{M})$ es el número de columnas linealmente independientes de $\boldsymbol{M}$, y por lo tanto $\mathrm{r}[\mathcal{C}(\boldsymbol{M})] = \mathrm{r}(\boldsymbol{M})$.
77. Si $\mathcal{M}, \mathcal{N}$ son dos EV's tales que $\mathcal{M} \subseteq \mathcal{N}$ entonces se dice que $\mathcal{M}$ es un **subespacio** (vectorial) de $\mathcal{N}$.
78. Sean $\mathcal{M}$ un EV y $\boldsymbol{x}$ un vector cualquiera. Se dice que $\boldsymbol{x}$ es **ortogonal** a $\mathcal{M}$ ($\boldsymbol{x} \perp \mathcal{M}$) si para todo $\boldsymbol{y} \in \mathcal{M}, \boldsymbol{x}'\boldsymbol{y} = 0$.

## Después del repaso...

79. Si $\mathcal{M} \subseteq \mathcal{N}$ son dos EV's, $\mathcal{M}_{\mathcal{N}}^{\perp} = \{\boldsymbol{x} \in \mathcal{N} : \boldsymbol{x} \perp \mathcal{M}\}$. $\mathcal{M}_{\mathcal{N}}^{\perp}$ se denomina el **complemento ortogonal de $\mathcal{M}$ respecto a $\mathcal{N}$**. Si $\mathcal{N} \equiv \mathbb{R}^n$ entonces $\mathcal{M}_{\mathcal{N}}^{\perp} = \mathcal{M}^{\perp}$ y se denomina simplemente el **complemento ortogonal de $\mathcal{M}$**.
80. Sea $\mathcal{N}$ un EV y $\mathcal{M}$ un subespacio de $\mathcal{N}$, entonces $\mathcal{M}_{\mathcal{N}}^{\perp}$ es un subespacio de $\mathcal{N}$. Si $\boldsymbol{x} \in \mathcal{N}$ entonces $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, donde $\boldsymbol{x}_1 \in \mathcal{M}$ y $\boldsymbol{x}_2 \in \mathcal{M}_{\mathcal{N}}^{\perp}$. Además, $\mathrm{r}(\mathcal{N}) = \mathrm{r}(\mathcal{M}) + \mathrm{r}(\mathcal{M}_{\mathcal{N}}^{\perp})$.
81. Si $\mathcal{M}, \mathcal{N}$ son dos EV's entonces $\mathcal{M} + \mathcal{N}$ es un EV y $\mathcal{M} + \mathcal{N} = \{\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 : \boldsymbol{x}_1 \in \mathcal{M}, \boldsymbol{x}_2 \in \mathcal{N}\}$.
82. $\mathcal{C}(A, B) = \mathcal{C}(A) + \mathcal{C}(B)$.
83. El conjunto de todos los $\boldsymbol{x}$ tales que $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ se denomina el **espacio nulo de $\boldsymbol{A}$** y se simboliza como $\mathcal{N}(A)$.
84. Si $\mathrm{r}(\boldsymbol{A}_{n \times n}) = r \Rightarrow \mathrm{r}[\mathcal{N}(\boldsymbol{A})] = n - r$.

## Después del repaso...

85. (Teorema 4.5.3) Dado $\boldsymbol{Y}_n \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con $\mathrm{r}(\boldsymbol{\Sigma}) = n$, si $\boldsymbol{A}\Sigma\boldsymbol{B} = 0$ entonces las dos formas cuadráticas $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ y $\boldsymbol{Y}'\boldsymbol{B}\boldsymbol{Y}$ son independientes.
86. (Teorema 1.5.13) Las matrices $\boldsymbol{A}\boldsymbol{A}^-$, $\boldsymbol{A}^-\boldsymbol{A}$, $\boldsymbol{I} - \boldsymbol{A}\boldsymbol{A}^-$ e $\boldsymbol{I} - \boldsymbol{A}^-\boldsymbol{A}$ son todas simétricas e idempotentes.
87. (Corolario 1.4.2) Si $\boldsymbol{B}_{p\times n}$ de rango $r$ entonces: $\boldsymbol{B}'\boldsymbol{B}$ y $\boldsymbol{B}\boldsymbol{B}'$ son matrices no negativas, $\boldsymbol{B}'\boldsymbol{B}$ es semidefinida positiva si $r < n$ y $\boldsymbol{B}'\boldsymbol{B}$ es definida positiva si $r = n$.