Parcial 2 Estadística Aplicada III

1

David Isaac López Romero
CU: 173993

1) a) Usamos los datos del examen y tenemos que las fórmulas son:

$Cov(X_i, Y_j) = \lambda_j a_{ji}$ e $Corr(X_i, Y_j) = \frac{\sqrt{\lambda_j} a_{ji}}{\sigma_i}$

De Excel, tenemos que

Covarianza

Variable	1er Comp	2ndo Comp	3er Comp
% Tierra Cultivable	1,6872	0,3304	0,1751
% Ganadería en PIB	1,6379	0,3304	0,3040
Industria Automovilística	0,4686	1,1566	0,2163
Industria Militar	0,3513	1,074	0,1236
Índice Democrático	0,3484	0,2478	0,8757
% Minería en PIB	0,5699	0,4956	0,1554
Industria Telecomunicaciones	0,7329	0,9914	0,1442

Correlación

Variable	1er Comp	2ndo Comp	3er Comp
% Tierra Cultivable	0,974105	0,233628	0,1751
% Ganadería en PIB	0,945353	0,233628	0,3040
Industria Automovilística	0,270546	0,817839	0,2163
Industria Militar	0,202823	0,7594327	0,1236
Índice Democrático	0,230016	0,175221	0,8757
% Minería en PIB	0,406382	0,350442	0,1554
Industria Telecomunicaciones	0,423140	0,701025	0,1442

b) Interpretación

Primer Componente: Actividades económicas primarias

Segundo Componente: Actividades económicas secundarias.

Tercer Componente: Democracia o política del país

c) Del Excel

	λ_j	$\frac{\lambda_j}{\text{Var tot}}$	% varianza explicada	% varianza acumulada explicada
Comp 1	3	0,428571	42.86%	42,86%
Comp 2	2	0,285714	28.57%	71,43%
Comp 3	1	0,142857	14.29%	85,71%

Gráfico en siguiente página.

d) Usamos los datos del examen y Excel para el país

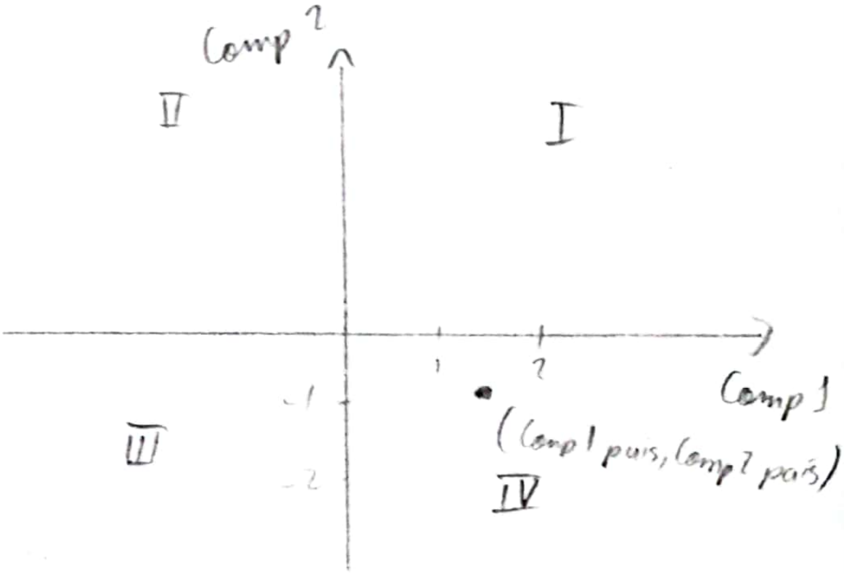
Componentes del país

Componente 1: 1.511959

Componente 2: -0.877494

Componente 3: 0,595766

El país se dedica mucho a las actividades económicas primarias y poco a las actividades económicas secundarias



Cuadrantes

I) { Mucha actividad económica primaria
Mucha actividad económica secundaria

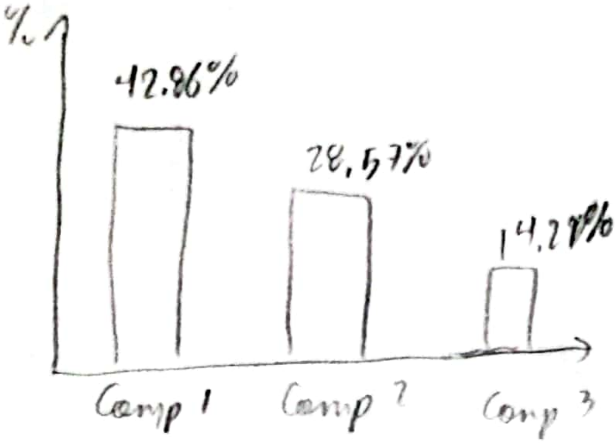
II) { Poca actividad económica primaria
Mucha actividad económica secundaria

III) Poca actividad económica primaria
Poca actividad económica secundaria

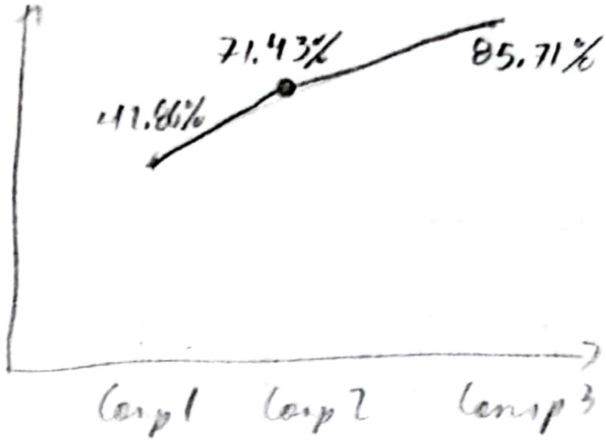
IV) { Mucha actividad económica primaria
Poca actividad económica secundaria

Gráfico inciso (c)

Varianza Explicada



Varianza Acumulada Explicada



5) $f_X(x) = \sqrt{\dfrac{\lambda}{2\pi x^3}}\; e^{-\frac{\lambda (x-\mu)^2}{2\mu^2 x}}$

a) Parámetro de interés es μ.

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}}\; e^{-\frac{\lambda (x-\mu)^2}{2\mu^2 x}} = \sqrt{\frac{\lambda}{2\pi x^3}}\; e^{-\frac{\lambda x^2}{2\mu^2 x} + \frac{2\lambda x\mu}{2\mu^2 x} - \frac{\lambda \mu^2}{2\mu^2 x}}$$

$$= \sqrt{\frac{\lambda}{2\pi x^3}}\; e^{-\frac{\lambda x}{2\mu^2}}\; e^{\frac{\lambda}{\mu}}\; e^{-\frac{\lambda}{2x}}$$

$$= \sqrt{\frac{\lambda}{2\pi x^3}}\; e^{-\frac{\lambda}{2x}}\; e^{\left\{-\frac{\lambda}{2\mu^2} x - \left(-\frac{\lambda}{\mu}\right)\right\}}$$

donde $h(x) = \sqrt{\dfrac{\lambda}{2\pi x^3}}\; e^{-\frac{\lambda}{2x}}$, $\eta(\mu) = -\dfrac{\lambda}{2\mu^2}$, $T(x) = x$

$$A(\mu) = -\frac{\lambda}{\mu}$$

b)

(4)

$A'(\mu) = \left(-\frac{\lambda}{\mu}\right)' = \frac{\lambda}{\mu^2}$

$\eta'(\mu) = \left(-\frac{\lambda}{2\mu^2}\right)' = \frac{2\lambda}{2\mu^3} = \frac{\lambda}{\mu^3}$

Así pues, como $T(x) = x$

$$\mathbb{E}(X) = \frac{A'(\mu)}{\eta'(\mu)} = \frac{\cancel{\lambda}/\mu^2}{\cancel{\lambda}/\mu^3} = \frac{\mu^3}{\mu^2} = \mu$$

c) $A''(\mu) = \left(\frac{\lambda}{\mu^2}\right)' = -\frac{2\lambda}{\mu^3}$

$\eta''(\mu) = \left(\frac{\lambda}{\mu^3}\right)' = -\frac{3\lambda}{\mu^4}$

Como $T(x) = x$

$$\mathrm{Var}(X) = \frac{A''(\mu) - \eta''(\mu)\,\mathbb{E}(X)}{[\eta'(\mu)]^2} =$$

$$= \frac{-\frac{2\lambda}{\mu^3} - \left(-\frac{3\lambda}{\mu^4}\right)\mu}{\left(\frac{\lambda}{\mu^3}\right)^2} = \frac{-\frac{2\lambda}{\mu^3} + \frac{3\lambda}{\mu^3}}{\frac{\lambda^2}{\mu^6}} = \frac{\left(\frac{\lambda}{\mu^3}\right)}{\left(\frac{\lambda^2}{\mu^6}\right)} = \frac{\cancel{\lambda}\mu^6}{\lambda^2\cancel{\mu^3}}$$

$$= \frac{\mu^3}{\lambda}$$

4) Probabilidad una persona paga o no

$$\ln\left(\frac{P_i}{1-P_i}\right) = -0{,}35 - 0.88\, Hi - 0{,}32\, Medi + 0.47\, Univi$$

$Hi = \begin{cases} 1 & \text{hombre} \\ 0 & \text{mujer} \end{cases}$, $Medi = \begin{cases} 1 & \text{ingresos menores a la media} \\ 0 & \text{e.o.c} \end{cases}$, $Univ_i = \begin{cases} 1 & \text{tiene estudios universitarios} \\ 0 & \text{e.o.c} \end{cases}$

- Mujer con ingresos superiores a la media sin universidad

$$\hat{P_i} = \frac{e^{-0{,}35}}{1+e^{-0.35}} = 0{,}4133824$$

- Hombre con ingresos superiores a la media sin universidad

$$\hat{P_i} = \frac{e^{-0{,}35-0{,}88}}{1+e^{-0{,}35-0{,}88}} = 0{,}2261814$$

- Mujer con ingresos menores a la media sin universidad

$$\hat{P_i} = \frac{e^{-0{,}35-0{,}32}}{1+e^{-0{,}35-0{,}32}} = 0{,}3384768$$

- Hombre con ingresos inferiores a la media sin universidad

$$\hat{P_i} = \frac{e^{-0{,}35-0{,}88-0{,}32}}{1+e^{-0{,}35-0{,}88-0{,}32}} = 0{,}1750863$$

- Mujer con ingresos menores a la media con universidad

$$\hat{P_i} = \frac{e^{-0{,}35-0{,}32+0{,}47}}{1+e^{-0{,}35-0{,}32+0{,}47}} = 0{,}450166$$

- Hombre con ingresos menores a la media con universidad

$$\hat{P_i} = \frac{e^{-0{,}35-0{,}88-0{,}32+0{,}47}}{1+e^{-0{,}35-0{,}88-0{,}32+0{,}47}} = 0{,}253506$$

- Mujer ingresos superiores a la media con universidad
$$\hat{p}_i = \frac{e^{-0.35+0.47}}{1+e^{-0.35+0.47}} = 0.529964$$

- Hombre ingresos superiores a la media con universidad
$$\hat{p}_i = \frac{e^{-0.35-0.88+0.47}}{1+e^{-0.35-0.88+0.47}} = 0.3186465$$

Mayor probabilidad de pago es de Mujer con ingresos superiores a la media con estudios universitarios

$\hat{p}_i = 0.529964$

Menor probabilidad de pago es del Hombre con ingresos inferiores a la media sin estudios universitarios

$\hat{p}_i = 0.1750863$

2)

matriz varianzas y covarianzas $\begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix} = \Sigma$

a)

$$|\Sigma - \lambda I| = \begin{vmatrix} \sigma_x^2-\lambda & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2-\lambda \end{vmatrix} = (\sigma_x^2-\lambda)(\sigma_y^2-\lambda) - (\rho\sigma_x\sigma_y)(\rho\sigma_x\sigma_y)$$

$$= \sigma_x^2\sigma_y^2 - \lambda\sigma_x^2 - \lambda\sigma_y^2 + \lambda^2 - \rho^2\sigma_x^2\sigma_y^2 = \lambda^2 - \lambda(\sigma_x^2+\sigma_y^2) + \sigma_x^2\sigma_y^2(1-\rho^2)$$

Cálculo raíces

$$\lambda = \frac{\sigma_x^2+\sigma_y^2 \pm \sqrt{(\sigma_x^2+\sigma_y^2)^2 - 4(1)\,\sigma_x^2\sigma_y^2(1-\rho^2)}}{2}$$

$$= \frac{\sigma_x^2+\sigma_y^2 \pm \sqrt{\sigma_x^2+\sigma_y^2 - 4\sigma_x^2\sigma_y^2(1-\rho^2)}}{2}$$

Así, $\lambda_{1,2} = \frac{\sigma_x^2+\sigma_y^2}{2} \pm \frac{\sqrt{(\sigma_x^2+\sigma_y^2)^2 - 4\sigma_x^2+\sigma_y^2(1-\rho^2)}}{2}$

Pero, como $\sum_{i\in\{x,y\}} \sigma_i^2 = \sigma_x^2+\sigma_y^2$, el porcentaje de varianza explicada del componente 1 es

$$\frac{\lambda_1}{\sigma_x^2+\sigma_y^2} = \frac{1}{\sigma_x^2+\sigma_y^2}\left\{ \frac{\sigma_x^2+\sigma_y^2}{2} + \frac{\sqrt{(\sigma_x^2+\sigma_y^2)^2 - 4\sigma_x^2+\sigma_y^2(1-\rho^2)}}{2} \right.$$

$$(\Rightarrow) \%Exp1 = \frac{1}{2} + \frac{\sqrt{(\sigma_x^2+\sigma_y^2)^2 - 4\sigma_x^2\sigma_y^2(1-\rho^2)}}{2(\sigma_x^2+\sigma_y^2)}$$

Análogamente,

$$\%Exp2 = \frac{1}{2} - \frac{\sqrt{(\sigma_x^2+\sigma_y^2)^2 + 4\sigma_x^2\sigma_y^2(1-\rho^2)}}{2(\sigma_x^2+\sigma_y^2)}$$

b) Supongamos que $|\rho| = 1$

<u>Dem:</u> $|\rho| = 1 \Leftrightarrow \rho^2 = 1 \Leftrightarrow 1-\rho^2 = 0$

$$\%Exp1 = \frac{1}{2} + \frac{\sqrt{(\sigma_x^2+\sigma_y^2)^2 - 4\sigma_x^2\sigma_y^2\cdot 0}}{2(\sigma_x^2+\sigma_y^2)} = \frac{1}{2} + \frac{\sqrt{(\cancel{\sigma_x^2+\sigma_y^2})^2}}{2(\cancel{\sigma_x^2+\sigma_y^2})}$$

$$= \frac{1}{2} + \frac{1}{2} = 1$$

$$\%Exp2 = \frac{1}{2} - \frac{\sqrt{(\sigma_x^2+\sigma_y^2)^2 - 4\sigma_x^2\sigma_y^2\cdot 0}}{2(\sigma_x^2+\sigma_y^2)} = \frac{1}{2} - \frac{\sqrt{(\cancel{\sigma_x^2+\sigma_y^2})^2}}{2(\cancel{\sigma_x^2+\sigma_y^2})}$$

$$= \frac{1}{2} - \frac{1}{2} = 0$$

c) Si $\rho = 0$

$$\% Exp 1,2 = \frac{1}{2} \pm \frac{\sqrt{(\sigma_x^2+\sigma_y^2)^2 - 4\sigma_x^2\sigma_y^2}}{2(\sigma_x^2+\sigma_y^2)} \quad \ldots (*)$$

$$\Rightarrow \sqrt{(\sigma_x^2+\sigma_y^2)^2 - 4\sigma_x^2\sigma_y^2} = \sqrt{(\sigma_x^2 - 2\sigma_x\sigma_y + \sigma_y^2)(\sigma_x^2+\sigma_y^2+2\sigma_x\sigma_y)}$$

$$= \sqrt{(\sigma_x^2+\sigma_y^2)^2(\sigma_x^2-\sigma_y^2)^2} = (\sigma_x^2+\sigma_y^2)(\sigma_x^2-\sigma_y^2)$$

Sustituyendo en $(*)$

$$\% Exp 1,2 = \frac{1}{2} \pm \frac{(\sigma_x^2+\sigma_y^2)(\sigma_x^2-\sigma_y^2)}{2(\sigma_x^2+\sigma_y^2)} = \frac{1}{2}\left\{1 \pm \sigma_x^2+\sigma_y^2\right\}$$

d) Estandarizado $\sigma_x^2 = \sigma_y^2 = 1$

$$\% Exp 1,2 = \frac{1}{2} \pm \frac{\sqrt{(1+1)^2 - 4\cdot 1\cdot 1(1-\rho^2)}}{2(1+1)} = \frac{1}{2} \pm \frac{\sqrt{4-4(1-\rho^2)}}{4}$$

$$= \frac{1}{2} \pm \frac{2\sqrt{1-1+\rho^2}}{4} = \frac{1}{2} \pm \frac{|\rho|}{2} = \frac{1}{2}(1 \pm \rho)$$

3) $\ln\left(\frac{p_i}{1-p_i}\right) = -0.5 + 1.2\, Ing_i + 0.98 M_i - 1.05\, Baja_i - 0.75\, Medi$

Ing_i = ingreso $\quad M_i = \begin{cases} 1 & \text{mujer} \\ 0 & \text{e.o.c.} \end{cases} \quad Baja_i = \begin{cases} 1 & \text{Clase baja} \\ 0 & \text{e.o.c.} \end{cases} \quad Medi = \begin{cases} 1 & \text{Clase Media} \\ 0 & \text{e.o.c.} \end{cases}$

Notemos que la base es hombre de clase alta.

a) Ingreso

$$\frac{\left(\frac{\hat{p}_i'}{1-\hat{p}_i'}\right)}{\left(\frac{\hat{p}_i}{1-\hat{p}_i}\right)} = e^{1.2} \Rightarrow (e^{1.2}-1)\times 100\% = 232.01\%$$

Momio aumenta 232.01% si aumenta en una unidad el ingreso

Hombre a mujer.

$$\frac{\left(\frac{p_i'}{1-p_i'}\right)}{\left(\frac{p_i}{1-p_i}\right)} = e^{0.98} \Rightarrow (e^{0.98}-1) \times 100\% = 166.44\%$$

Momio aumenta 166.44 % de hombre a mujer

Alta a bajo

$$\frac{\left(\frac{\hat{p}_i'}{1-\hat{p}_i'}\right)}{\left(\frac{\hat{p}_i}{1-\hat{p}_i}\right)} = e^{-1.05} \Rightarrow (e^{-1.05}-1) \times 100\% = -65.006\%$$

Momio disminuye 65% de clase alta a baja

Alta a Media

$$\frac{\left(\frac{\hat{p}_i'}{1-\hat{p}_i'}\right)}{\left(\frac{\hat{p}_i}{1-\hat{p}_i}\right)} = e^{-0.75} \Rightarrow (e^{-0.75}-1) \times 100 = -52.76\%$$

Momio disminuye siposos de alta a media en 52.76%

b) Con la tabla del examen IC = intervalo confianza)

IC (Ordenada origen) = (0.223, 1.64)

IC (Ingreso) = (1.22, 9.02)

IC (Mujer) = (1.68, 4.39)

IC (NivelBajo) = (0.2122, 0.576)

IC (Nivel Medio) = (0.173, 1.284)

Vemos que

$1 \in IC$ (Ordenada al origen)

$1 \in IC$ (Nivel Medio)

$\therefore$ Ordenada al origen y nivel medio **NO** son significativos

Asimismo

$1 \notin IC$ (Ingreso)

$1 \notin IC$ (Mujer)

$1 \notin IC$ (Nivel Bajo)

$\therefore$ Ingreso, Mujer y Nivel Bajo sí son variables significativas al 95%