Exámenes pasados EAII

1. **(32 Puntos)** Las siguientes afirmaciones se refieren al Modelo de Regresión Lineal Simple, diga si cada una de ellas es CIERTA o FALSA:

   (a) Considere el modelo $Y_i - \beta_0 = (\beta_1 X_i) u_i$ con $i = 1, \dots, n$, donde u_i es un error aleatorio que satisface los supuestos clásicos. Entonces β_0 y β_1 pueden **estimarse por MCO**. Falsa

   (b) La siguiente igualdad **se cumple cuando** el modelo de regresión lineal simple contiene la ordenada al origen y se estiman los parámetros por Máxima Verosimilitud. Falsa

   $$\sum_{i=1}^n (Y_i - \hat{Y}_i) = \sqrt{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}.$$

   (c) Si en el modelo aparece la ordenada al origen y se estiman los parámetros por MCO, entonces **se cumple que** Falsa

   $$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n (Y_i - \bar{Y})^2 + \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$$

   (d) El Teorema de Gauss – Markov garantiza la **eficiencia** de los estimadores Mínimo – Cuadráticos. Falsa

1a) $Y_i - \beta_0 = (\beta_1 x_i) u_i, \quad i = 1, \dots, n$ Falsa

$\mathbb{E}(u_i) = 0, \quad Var(u_i) = \sigma^2, \quad Cov(u_i, u_j) = 0 \quad \forall i \neq j.$

Si X es fija

$\mathbb{E}(Y_i) = \mathbb{E}(\beta_0 + (\beta_1 x_i) u_i) = \beta_0 + \beta_1 x_i \overset{=0}{\cancel{\mathbb{E}(u_i)}} = \beta_0$

$Var(Y_i) = Var(\beta_0 + (\beta_1 x_i) u_i) = \beta_1^2 x_i^2 Var(u_i) = \beta_1^2 x_i^2 \sigma^2$

1b) P.D. $\sum (Y_i - \hat{Y}_i) = \sqrt{\sum (Y_i - \hat{Y}_i)^2}$ Falsa

Argumento ①

$e_i = Y_i - \hat{Y}_i$ y $\sum (Y_i - \hat{Y}_i) = 0$

Asimismo $SCR = \sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2 \neq 0$ en general

$\therefore \sqrt{\sum (Y_i - \hat{Y}_i)^2} \neq 0 = \sum (Y_i - \hat{Y}_i)$

1c) Si $\beta_0 \neq 0$ ($\Rightarrow b_0 \neq 0$) y se usa MCO Falsa

$\Rightarrow \sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \bar{Y})^2 + \sum (\hat{Y}_i - \bar{Y})^2$

Sabemos que

$SCT = SCR + SCE$

$\Leftrightarrow \sum (Y_i - \bar{Y})^2 = \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2$

$\Leftrightarrow \sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \bar{Y})^2 - \sum (\hat{Y}_i - \bar{Y})^2$

1d) Teo Gauss Markov indica que b_0 y b_1 son MELI

= Mejores Estimadores lineales insesgados

(No habla de eficiencia)

Intento fallido ② ⓐ

¿ $\beta_1 = \alpha_1$?

$$\begin{cases} \dfrac{\sum (Y_i - \bar{Y})}{S_Y} = (\bar{Y} - \beta_1 \bar{X}) + \beta_1 \dfrac{\sum (X_i - \bar{X})}{S_X} \\ \dfrac{\sum (X_i - \bar{X})}{S_X} = (\bar{X} - \alpha_1 \bar{Y}) + \alpha_1 \dfrac{\sum (Y_i - \bar{Y})}{S_Y} \end{cases}$$

$$\Leftrightarrow \begin{cases} \dfrac{\sum (Y_i - \bar{Y})}{S_Y} = \bar{Y} - \beta_1 \left(\bar{X} - \dfrac{\sum (X_i - \bar{X})}{S_X} \right) \\ \dfrac{\sum (X_i - \bar{X})}{S_X} = \bar{X} - \alpha_1 \left(\bar{Y} - \dfrac{\sum (Y_i - \bar{Y})}{S_Y} \right) \end{cases}$$

Intentos Fallidos

1b) ¿Si el modelo de regresión lineal simple tiene a β_0 y se estima por Máxima Verosimilitud

$\Rightarrow \sum_{i=1}^n (Y_i - \hat{Y}_i) = \sqrt{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}$?

Verifiquemos $SCT = SCR + SCE$

$\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2 = SCR$

$\Rightarrow SCR = \sum (Y_i - \hat{Y}_i)^2$

$= \sum (Y_i - \bar{Y})^2 - \sum (\hat{Y}_i - \bar{Y})^2$

$= \sum (Y_i^2 - 2Y_i \bar{Y} + \bar{Y}^2) - \sum (\hat{Y}_i^2 - 2\hat{Y}_i \bar{Y} + \bar{Y}^2)$

$= \sum Y_i^2 - 2Y_i \bar{Y} + \cancel{\bar{Y}^2} - \hat{Y}_i^2 + 2\hat{Y}_i \bar{Y} - \cancel{\bar{Y}^2}$

$= \sum Y_i^2 - 2Y_i \bar{Y} - \hat{Y}_i^2 + 2\hat{Y}_i \bar{Y}$

$= \sum Y_i^2 - \hat{Y}_i^2 - 2\bar{Y}(Y_i - \hat{Y}_i)$

$= \sum (Y_i - \hat{Y}_i)(Y_i + \hat{Y}_i) - 2\bar{Y}(Y_i - \hat{Y}_i)$

$= \sum e_i Y_i + \overset{=0}{\cancel{e_i \hat{Y}_i}} - 2\bar{Y}(Y_i - \hat{Y}_i)$

$= \sum e_i Y_i - 2\bar{Y}(Y_i - \hat{Y}_i)$

$= \sum e_i Y_i - 2\bar{Y}(\sum Y_i - \sum \hat{Y}_i)$

$= \sum e_i Y_i - 2\bar{Y}(\sum Y_i - \overset{=0}{\cancel{\sum \hat{Y}_i}})$

$= \sum e_i Y_i = \sum (Y_i - \hat{Y}_i) Y_i$

$= \sum (Y_i - b_0 - b_1 x_i) Y_i = \sum Y_i^2 - b_0 \sum Y_i - b_1 \sum x_i Y_i$

$= \sum Y_i^2 - (\bar{Y} - b_1 \bar{x}) Y_i - b_1 x_i Y_i$

$= \sum Y_i^2 - \bar{Y} Y_i - b_1 \bar{x} Y_i - b_1 x_i Y_i$

$= \sum Y_i^2 - \bar{Y} Y_i - b_1 (\bar{x} Y_i - x_i Y_i)$

1b otra vez

Sea $y_i = Y_i - \hat{Y}_i$ entonces

$\sum y_i - \hat{Y}_i = \sqrt{\sum (Y_i - \hat{Y}_i)^2}$

$\Leftrightarrow (\sum y_i)^2 = \sum (y_i^2)$

$\Leftrightarrow \cancel{\sum y_i^2} + \sum_{\substack{i=1 \\ i \neq j}}^n y_i y_j = \cancel{\sum y_i^2}$

$\Leftrightarrow \sum_{\substack{i=1 \\ i \neq j}}^n y_i y_i = 0 \Leftrightarrow \sum_{i \neq j} (Y_i - \hat{Y}_i)(Y_j - \hat{Y}_j) = 0$

$\Leftrightarrow \sum_{i \neq j} Y_i Y_j - Y_i \hat{Y}_j - \hat{Y}_i Y_j + \hat{Y}_i \hat{Y}_j = 0$

2. **(33 Puntos)** Elija la opción que corresponda, para los enunciados que siguen.

   (a) En caso de que se postulen los dos modelos de regresión,

   $$\tilde{Y}_i = \beta_0 + \beta_1 \tilde{X}_i + \varepsilon_i \quad \text{y} \quad \tilde{X}_i = \alpha_0 + \alpha_1 \tilde{Y}_i + \delta_i,$$

   donde ε y δ son errores aleatorios que satisfacen los supuestos clásicos, con $\tilde{X}_i = (X_i - \bar{X})/S_X$ y $\tilde{Y}_i = (Y_i - \bar{Y})/S_Y$, entonces los estimadores de MCO de β_1 y de α_1 **son idénticos**.

   OPCIONES: ~~i)~~ Cierto sólo si $\alpha_0 = \beta_0$ ii) Cierto
   ~~iii)~~ Cierto sólo si $S_X = S_Y$ iv) Falso

   (b) Si $\bar{r}^2$ denota al coeficiente de determinación ajustado por grados de libertad y r es el coeficiente de correlación, entonces se cumple que $\bar{r}^2 < |r|$.

   OPCIONES: i) Cierto sólo si $r = 0$ ii) Cierto
   iii) Cierto sólo si $r \neq 0$ iv) Falso

   (c) La covarianza entre los estimadores de MCO, b_0 y b_1, en un modelo de regresión lineal simple será ____________ si todos los valores de las variables X y Y son positivos.

   OPCIONES: i) cero, ii) negativa, iii) positiva, iv) ninguna de las anteriores.

i) No Funciona

Si $\alpha_0 = \beta_0$

$\Rightarrow \tilde{Y}_i = \beta_0 + \beta_1 \tilde{X}_i + \varepsilon_i$

$\Rightarrow \tilde{X}_i = \alpha_0 + \alpha_1 \tilde{Y}_i + \varepsilon_i$

$\Rightarrow \tilde{Y}_i = \tilde{X}_i - \alpha_1 \cancel{\tilde{Y}_i} + \beta_1 \tilde{X}_i$

$\Leftrightarrow (1 + \alpha_1) \tilde{Y}_i = (1 + \beta_1) \tilde{X}_i$

$\Leftrightarrow \left(1 + \dfrac{S_{XY}}{S_{YY}}\right) \dfrac{\sum (Y_i - \bar{Y})}{S_{YY}} = \left(1 + \dfrac{S_{XY}}{S_{XX}}\right) \dfrac{\sum (X_i - \bar{X})}{S_{XX}}$

$\Leftrightarrow \left(\dfrac{S_{YY} + S_{XY}}{S_{YY}^2}\right) \sum (Y_i - \bar{Y}) = \left(\dfrac{S_{XX} + S_{XY}}{S_{XY}^2}\right) \sum (X_i - \bar{X})$

Falso en general

2a) La asociación lineal para $\tilde{X}$ o $\tilde{Y}$ se da como sigue

¿ $\alpha_1 = \beta_1$?

$\hat{\beta}_1 = \dfrac{n \sum \tilde{x}_i \tilde{Y}_i - \sum \tilde{x}_i \sum \tilde{Y}_i}{n \sum \tilde{x}_i^2 - (\sum \tilde{x}_i)^2} = \dfrac{n \sum \frac{(x_i - \bar{x})}{S_x} \frac{(Y_i - \bar{Y})}{S_Y} - \sum \left(\frac{x_i - \bar{x}}{S_x}\right) \sum \left(\frac{Y_i - \bar{Y}}{S_Y}\right)}{n \sum \left(\frac{x_i - \bar{x}}{S_x}\right)^2 - \left(\sum \frac{x_i - \bar{x}}{S_x}\right)^2}$

$\hat{\alpha}_1 = \dfrac{n \sum \tilde{x}_i \tilde{Y}_i - \sum \tilde{x}_i \sum \tilde{Y}_i}{n \sum \tilde{Y}_i^2 - (\sum \tilde{Y}_i)^2} = \dfrac{n \sum \frac{(x_i - \bar{x})}{S_x} \frac{(Y_i - \bar{Y})}{S_Y} - \sum \left(\frac{x_i - \bar{x}}{S_x}\right) \sum \left(\frac{Y_i - \bar{Y}}{S_Y}\right)}{n \sum \left(\frac{Y_i - \bar{Y}}{S_Y}\right)^2 - \sum \left(\frac{Y_i - \bar{Y}}{S_Y}\right)^2}$

$\hat{\alpha}_1 = \hat{\beta}_1 \Leftrightarrow n \sum \left(\frac{x_i - \bar{x}}{S_x}\right)^2 - \left[\sum \left(\frac{x_i - \bar{x}}{S_x}\right)\right]^2 = n \sum \left(\frac{Y_i - \bar{Y}}{S_Y}\right)^2 - \left[\sum \left(\frac{Y_i - \bar{Y}}{S_Y}\right)\right]^2$

$\Leftrightarrow n \dfrac{\sum (x_i - \bar{x})^2}{S_x^2} - \dfrac{1}{S_x^2} \left(\sum (x_i - \bar{x})\right)^2 = n \dfrac{\sum (Y_i - \bar{Y})^2}{S_Y^2} - \dfrac{\left[\sum (Y_i - \bar{Y})\right]^2}{S_Y^2}$

Notación Viejo

$\Leftrightarrow n \dfrac{\cancel{S_x^2}}{\cancel{S_x^2}} - \dfrac{\left[\sum (x_i - \bar{x})\right]^2}{S_x^2} = n \dfrac{\cancel{S_Y^2}}{\cancel{S_Y^2}} - \dfrac{\left[\sum (Y_i - \bar{Y})\right]^2}{S_Y^2}$

$\Leftrightarrow \dfrac{\left[\sum (x_i - \bar{x})\right]^2}{S_x^2} = \dfrac{\left[\sum (Y_i - \bar{Y})\right]^2}{S_Y^2}$

(ii) $\Leftrightarrow \left[\sum (x_i - \bar{x})^2\right] = \left[\sum (Y_i - \bar{Y})\right]^2$ Falso en general

2 b) ¿ $\bar{r}^2$ = coef det ajustado con g.l. $\Rightarrow \bar{r}^2 < |r|$?

$\bar{r}^2 = 1 - \dfrac{\left(\frac{SCR}{n-2}\right)}{\left(\frac{SCT}{n-1}\right)}$

$-1 \leq r^2 \leq 1$

$\Leftrightarrow |r^2| \leq 1$

$\Leftrightarrow r^2 \leq 1$

$\Leftrightarrow r \leq 1$

$|r| = \sqrt{1 - \dfrac{\left(\frac{SCR}{n-2}\right)}{\left(\frac{SCT}{n-1}\right)}}$

Como $|r| < 1 \Rightarrow r^2 < r < 1$

Verdadero si $r \neq 0$

2c) X, Y positivos

i) 0 ii) negativa iii) positiva iv) ninguna

$Cov(b_0, b_1) = Cov(\bar{Y} - b_1 \bar{x}, b_1) = Cov(\bar{Y}, b_1) + Cov(-b_1 \bar{x}, b_1)$

$= -\bar{x}\, Cov(b_1) = -\bar{x}\, Var(b_1) = -\bar{x} \dfrac{\sigma^2}{\sum (x_i - \bar{x})^2}$

// $Cov(\bar{Y}, b_1) = 0$

$= \mathbb{E}(\bar{Y} b_1) - \mathbb{E}(\bar{Y}) \mathbb{E}(b_1) = \mathbb{E}\left(\frac{1}{n} \sum Y_i\, b_1\right) - \mathbb{E}\left(\frac{1}{n} \sum Y_i\right) \mathbb{E}(b_1)$

$= \frac{1}{n} \left\{ \sum \mathbb{E}\left(Y_i \frac{\sum (x_i - \bar{x})(Y_i - \bar{Y})}{\sum (x_i - \bar{x})^2}\right) - \sum \underbrace{\mathbb{E}(Y_i)}_{=Y_i} \frac{\sum (x_i - \bar{x})(Y_i - \bar{Y})}{\sum (x_i - \bar{x})^2} \right\}$

$= \frac{1}{n} \left\{ \mathbb{E}\left[\sum_i \sum_j \frac{Y_i \cancel{(x_j - \bar{x})(Y_j - \bar{Y})}}{\cancel{\sum (x_i - \bar{x})^2}}\right] - \sum_i \sum_j \frac{Y_i \cancel{(x_j - \bar{x})(Y_j - \bar{Y})}}{\cancel{\sum (x_i - \bar{x})^2}} \right\}$

$= 0$

1. **(24 Puntos)** Diga si cada una de las siguientes afirmaciones es CIERTA o FALSA.

(a) El Teorema de Gauss – Markov garantiza la consistencia de los estimadores de MCO. Falso $\Rightarrow$ MELI

(b) El método de MCO se puede usar para estimar los parámetros β_0 y β_1 en el modelo

$$\log(Y_i) = (\beta_0 + \beta_1 X_i + \varepsilon_i)^{-1} \text{ con } i = 1, \dots, n,$$

donde ε_i es un error aleatorio que satisface los supuestos clásicos. Cierto

(c) En un intervalo de 90% de confianza para la pendiente de la recta de regresión, siempre se encontrará el valor estimado de la pendiente. Cierta

(d) En el modelo de regresión lineal simple con distribución Normal, un estimador lineal de la pendiente puede ser de MCO, aunque no sea insesgado. Falsa

(Si se cumple Normalidad $\Rightarrow$ $b_{0_{MC}} = b_{0_{MV}}$

$b_{1_{MC}} = b_{1_{MV}}$

2. **(36 Puntos)** Elija la opción que corresponda para llenar el espacio en blanco de los siguientes enunciados:

(a) Si todos los valores de las variables X y Y son positivos, la covarianza entre los estimadores de MCO, b_0 y b_1, en un modelo de regresión lineal simple, será ___________.

OPCIONES: 1. cero, 2. negativa, 3. positiva, 4. ninguna de las anteriores.

(b) En un estudio se consideran las variables X y Y, y dos posibles modelos estimados por MCO: (1) la regresión de Y sobre X (con pendiente $b_1 = 5$) y (2) la regresión de X sobre Y (con pendiente $b_1' = 10$). Entonces las varianzas muestrales guardan la relación ___________.

OPCIONES: 1. $S_X^2 = 2S_Y^2$, 2, $S_X^2 = 50S_Y^2$, 3. $S_Y^2 = 2S_X^2$, 4. $S_Y^2 = 50S_X^2$, 5. ninguna de las anteriores.

(c) Los estimadores de MCO para el modelo de regresión lineal simple, sin el supuesto de Normalidad, son lineales debido a que ___________.

OPCIONES: 1. minimizan la Suma de Cuadrados Residual, 2. son insesgados, 3. maximizan la Función de Verosimilitud, 4. ninguna de las anteriores.

2b) X en Y $b_1 = 5$

Y en X $b_1' = 10$

$b_1 = \frac{S_{xy}}{S_{xx}} = 5 \qquad b_1' = \frac{S_{xy}}{S_{yy}} = 10$

$\Leftrightarrow S_{xy} = 5 S_{xx} \quad , \quad S_{xy} = 10 S_{yy}$

$\therefore 5 S_{xx} = 10 S_{yy} \Leftrightarrow S_{xx} = 2 S_{yy}$

3. **(35 Puntos)** Considere las variables X y Y, así como el modelo

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

en donde ε es un error aleatorio que cumple con los supuestos clásicos.

a) Parta de la siguiente definición del coeficiente de correlación muestral

$$r_{XY} = \frac{\hat{\text{Cov}}(X,Y)}{S_X S_Y}$$

para **demostrar** que si b_1 es el estimador mínimo-cuadrático de β_1, entonces

$$r_{XY} = b_1 S_X / S_Y.$$

b) Considere un tamaño de muestra de $n = 30$ y que se conocen los datos siguientes

$$\bar{X} = 150, \; S_X = 100$$
$$\bar{Y} = 2000, \; S_Y = 1000, \; r_{XY} = 0.8$$

- **Estime** los parámetros β_0 y β_1 de la regresión de Y sobre X.
- **Estime** los parámetros α_0 y α_1 en el modelo $X = \alpha_0 + \alpha_1 Y + \delta$ donde δ es un error aleatorio que cumple con los supuestos clásicos
- ¿Explica más X a Y que Y a X?. **Justifique** su respuesta.

3) $Y = \beta_0 + \beta_1 X + \varepsilon$, ε error aleatorio clásico

a) $r_{xy} = \frac{Cov(x,y)}{S_x S_y}$ y $b_{1_{MCO}}$ de β_1

P.D. $r_{xy} = \frac{b_1 S_x}{S_y}$

Dem.

$$r_{xy} = \frac{Cov(x,y)}{S_x S_y} = \frac{S_{xy}}{S_x S_y} = \frac{\left(\frac{\sum x_i y_i - \sum x_i \sum \frac{y_i}{n}}{n-1}\right)}{\sqrt{\left(\frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}\right)}\sqrt{\left(\frac{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}{n-1}\right)}}$$

$$= \frac{\left(\frac{\sum x_i y_i - \sum x_i \sum \frac{y_i}{n}}{\cancel{n-1}}\right)}{\frac{\sqrt{\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right)\left(\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right)}}{\cancel{n-1}}}$$

$$= \frac{\sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i}{\sqrt{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2}\sqrt{\sum y_i^2 - \frac{1}{n}(\sum y_i)^2}} \cdot \frac{\sqrt{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2}}{\sqrt{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2}}$$

$$= \frac{\sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i}{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2}\left(\frac{\sqrt{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2}}{\sqrt{\sum y_i^2 - \frac{1}{n}(\sum y_i)^2}}\right)$$

$$= b_1 \frac{S_x}{S_y}$$

3b) $n = 30 \quad \bar{x} = 150 \quad \bar{y} = 2000 \quad r_{xy} = 0.8$

$s_x = 100 \quad s_y = 1000$

i) β_0 y β_1 de X en Y

$b_1 = \frac{S_{xy}}{S_x^2} = \frac{80{,}000}{(100)^2} = 8 \; , \; b_0 = \bar{y} - b_1\bar{x} = 2000 - 8(150) = 800$

// $r_{xy} = \frac{S_{xy}}{S_x S_y} \Leftrightarrow 0.8(100)(1000) = S_{xy} \Leftrightarrow S_{xy} = 80000$

$\therefore b_0 = 800, \; b_1 = 8 \Rightarrow \hat{y} = 800 + 8X$

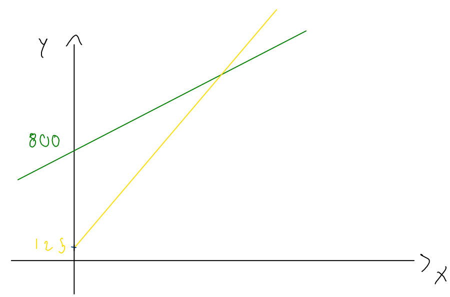
ii) α_0 y α_1 de Y en X

$\alpha_1 = \frac{S_{xy}}{S_y^2} = \frac{80{,}000}{1000^2} = 0{,}08, \; \alpha_0 = \bar{x} - \alpha_1\bar{y} = 150 - 0.08(2000)$

$= -10$

$\alpha_0 = -10, \; \alpha_1 = 0.08$

$\Rightarrow \hat{X} = \alpha_0 + \alpha_1 Y \Leftrightarrow Y = -\frac{\alpha_0}{\alpha_1} + \frac{\hat{X}}{\alpha_1} \Leftrightarrow Y = -125 + 12.5X$



X explica mejor a Y porque la varianza muestral de X es menor a la varianza muestral de Y

X en Y $\rightarrow$ $b_1 = \frac{S_{xy}}{S_{xx}}$ (b_1 tiene menor varianza que α_1)

Y en X $\Rightarrow$ $\alpha_1 = \frac{S_{xy}}{S_{yy}}$

3. (40 Puntos) Se desea estudiar la posible relación entre el Ingreso personal disponible (I) y el Gasto de las personas en un cierto producto (Y), expresadas en miles de pesos mensuales. Estas dos variables se observaron para ocho individuos y se obtuvieron los siguientes resultados:

$$\sum_{i=1}^{8} Y_i = 15.50,\ \sum_{i=1}^{8} Y_i^2 = 34.07,\ \sum_{i=1}^{8} Y_i I_i = 55.30$$

$$\sum_{i=1}^{8} I_i = 25.12 \ \text{y} \ \sum_{i=1}^{8} I_i^2 = 89.87$$

El objetivo del estudio consiste en explicar el gasto personal en función del ingreso disponible, usando para ello un modelo de regresión lineal simple.

USE TRES DECIMALES EN SUS CÁLCULOS NUMÉRICOS

(a) **Postule** el modelo correspondiente, **estime** los parámetros involucrados e **interprete** los resultados en el contexto del estudio.

(b) **Indique** el porcentaje de variabilidad del gasto que se explica con el modelo.

(c) **Calcule** el valor estimado de la desviación estándar del error.

(d) **Pruebe** la hipótesis nula de que el ingreso no afecta al gasto, al nivel de significancia del 5% (use para ello el punto porcentual, dado por el valor 2.447, de la distribución t correspondiente).

Relación entre I e Y $n=8$

$\sum Y_i = 15.50,\ \sum Y_i^2 = 34.07,\ \sum Y_i I_i = 55.30$

$\sum I_i = 25.12,\ \sum I_i^2 = 89.87$

$$b_1 = \frac{S_{IY}}{S_I^2} = \frac{n\sum I_i Y_i - \sum I_i \sum Y_i}{n\sum I_i^2 - (\sum I_i)^2} = \frac{8(55.30) - (25.12)(15.50)}{8(89.87) - (25.12)^2}$$

$= 0.6031$ // $\bar{Y} = \frac{31}{16} = 1.937$ //

$$b_0 = \bar{Y} - b_1 \bar{I} = \frac{1}{8}(15.50) - 0.603\left(\frac{25.12}{8}\right) = \frac{4}{625} = 6.7\times10^{-3}$$

$= 0.0437$

$\therefore\ \hat{Y} = b_0 + b_1 X$, $b_0 = 0.6031$, $b_1 = 0.0437$

b) Variabilidad se calcula con R^2 y $\bar{R}^2$ (ajustada)

$$R^2 = \frac{SCE}{SCT} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = \frac{\overbrace{\sum(\hat{Y}_i - \bar{Y})^2}^{SCE}}{\sum Y_i^2 - n\bar{Y}^2}$$

$$= b_1^2 \frac{\sum x_i^2}{\sum y_i^2} = b_1^2 \frac{\sum(X_i - \bar{X})^2}{\sum(Y_i - \bar{Y})^2} = b_1^2 \frac{\sum x_i^2 - n\bar{x}^2}{\sum y_i^2 - n\bar{y}^2}$$

$$= 0.6031^2 \frac{89.87 - 8(3.14)^2}{34.07 - 8(1.9375)^2} = 0.990047$$

El modelo explica el 99% de la variabilidad

c) $$S^2 = \frac{\overset{(SCE)}{SCR_{\text{egresión}}}}{g.l.} = \frac{\overset{(SCE)}{SCR_{\text{egresión}}}}{n-2} = \frac{3.4485}{6} = 0.66642$$

$\therefore S = 0.08163789$

$\Rightarrow$ 8.1% de error estándar

// $SCT = SCE + SCR$

$R^2 = \frac{SCE}{SCT} \Leftrightarrow SCE = R^2 \underbrace{SCT}_{=4.03875} = 3.998552321$

d) $H_0: \beta_1 = 0$ v.s. $H_A: \beta_1 \neq 0$

significancia $\alpha = 5\%$ punto de t-student de 2.447

$$t = \frac{b_1 - 0}{\hat{e}e(b_1)} = \frac{b_1}{\sqrt{\hat{Var}(\hat{b}_1)}} = \frac{b_1}{\left(\frac{S}{\sqrt{\sum(x_i - \bar{x})^2}}\right)} = \frac{b_1}{\left(\frac{S}{\sqrt{\sum x_i^2 - n\bar{x}^2}}\right)}$$

$$= \frac{b_1\left(\sqrt{\sum x_i^2 - n\bar{x}^2}\right)}{S} = 2.4494 > 2.447$$

$\therefore$ Se rechaza H_0 con $\alpha = 0.05$ de significancia

① $\ln(y_i) = [\beta_0 + \beta_1 X_i + u_i]^{-1}$ $\forall i = 1, \dots, n$.
t.q. u_i cumple supuestos clásicos

$E(u_i) = 0$
$var(u_i)$ cte, $\neq 0$, finita
$cov(u_i, u_j) = 0 \ \forall i \neq j$

¿se puede usar MCO?
sí es solo una transformación:

$$y_i^{*} = [\ln(y_i)]^{-1} = \beta_0 + \beta_1 X_i + u_i$$

② En el modelo de regresión lineal simple con distribución Normal para los errores, un estimador lineal de la pendiente puede ser de MCO, aunque no necesariamente sea insegado.

Falso

③ en un intervalo $100(1-\alpha)\%$ de confianza para la pendiente de la recta de regresión t.q. $\alpha \leq 1$, siempre se encontrará el valor estimado de la pendiente.

cierto siempre:
$\beta_1 \in [b_1 \pm \_\_\_]$ con $100(1-\alpha)\%$ de confianza
$\Rightarrow$ $b_1 \in$ intervalo siempre

④ $n \geq 3$, $\sum_{i=1}^{n} [y_i - \hat{y}_i]^2 = \Sigma(y_i - \bar{y})^2 + \Sigma(\hat{y}_i - \bar{y})^2$

$$SCR = SCT + SCE$$

Falso

$$SCR = SCT - SCE$$

$n-2$ gl, $n-1$ gl, 1 gl

Si $n=2 \Rightarrow$ no hay SCR

⑤ El Teorema de Gauss - Markov establece que los estimadores de MCO son:

- ~~Invariantes e insesgados~~ MV
- ~~Consistentes~~ y eficientes MV
- ~~Consistentes~~ y lineales MV
- Eficientes y lineales

Esta respuesta es correcta, porque el Teorema de Gauss-Markov establece que los estimadores de MCO son MELI's, o sea mejores (en términos de eficiencia), lineales e insesgados.

- ~~Suficientes~~ e ~~invariantes~~ → MV

MV $\neq$ varianza mínima

⑥ $y_i^{\circ} = \beta_0 + \beta_1 X_i^{\circ} + u_i$
$X_i^{\circ} = \alpha_0 + \alpha_1 y_i^{\circ} + v_i$

$\Rightarrow \beta_1 = \alpha_1$.

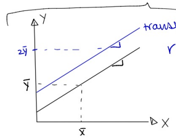
$$\underbrace{y_i^{\circ} = \frac{y_i - \bar{y}_i}{s_y} \quad X_i^{\circ} = \frac{X_i - \bar{X}_i}{s_x}}$$

estandarizadas $\Rightarrow \bar{y}_i^{\circ} = 0$
$\Rightarrow s^2_{y^{\circ}} = 1$

verdadero:

$$b_1 = r_{y^{\circ},x^{\circ}} \frac{s_{y^{\circ}}}{s_{x^{\circ}}} = r_{y^{\circ},x^{\circ}} = r_{yx}$$

$$a_1 = r_{x^{\circ},y^{\circ}} \frac{s_{x^{\circ}}}{s_{y^{\circ}}} = \underbrace{r_{x^{\circ},y^{\circ}} = r_{xy}} = r_{yx}$$



transformación lineal

$r = r$

LO VIMOS EN LA TAREA
cambia ordenada al origen
no cambia la pendiente
ni asociación

⑦ En un análisis de regresión, se consideran las variables X y Y, y dos posibles modelos estimados por MCO:

(1) la regresión de Y sobre X (con pendiente $b_1 = 5$)

y

(2) la regresión de X sobre Y (con pendiente $b_1' = 10$).

Entonces, las varianzas muestrales mantienen la siguiente relación

- $S_X^2 = 50 S_Y^2$
- $S_Y^2 = 50 S_X^2$
- Ninguna de las otras opciones
- $S_X^2 = 2 S_Y^2$

$$r = b_1 \frac{s_x}{s_y} = b_1' \frac{s_y}{s_x}$$

$$b_1 s_x^2 = b_1' s_y^2$$

$$s_x^2 = \left(\frac{b_1'}{b_1}\right) s_y^2 = 2 s_y^2$$

⑧ Si $\bar{r}^2$
$r \neq 0$, $-1 < r < 1$
$\Rightarrow \bar{r}^2 < |r|$ cierto

$$\bar{r}^2 = 1 - \frac{\frac{SCR}{n-2}}{\frac{SCT}{n-1}} \qquad r^2 = 1 - \frac{SCR}{SCT}$$

$$\bar{r}^2 = 1 - \left(\frac{n-1}{n-2}\right)\frac{SCR}{SCT}$$

$$= 1 - \left(\frac{n-1}{n-2}\right)(1 - r^2) < 1 - (1 - r^2) = r^2 < |r|$$

Pregunta 8 8 / 8 pts

Si no se cumple el supuesto de media cero del error:

- Se produce no-normalidad de los errores.
- Se generan observaciones aberrantes.
- Ninguna de las otras opciones.

Esta respuesta es correcta, porque el supuesto de media cero del error no está ligado a ningún otro supuesto del modelo.

- Se viola el supuesto de varianza constante.

Pregunta 2 6 / 6 pts

La gráfica de los residuos *vs.* los valores observados estimados de la variable dependiente permite detectar la no-linealidad de efectos de las variables explicativas.

- Verdadero
- Falso

Esta respuesta es correcta porque la gráfica de residuos *vs.* valores observados de Y siempre va a mostrar alguna asociación lineal, que no es relevante para detectar no-linealidad de efectos de las Xs.

Pregunta 4

La linealidad de los estimadores de MCO no puede verificarse gráficas de residuos *vs.* cada una de las variables explicativas.

- Verdadero

Esta respuesta es correcta, ya que la linealidad de los estimadore MCO no es verificable con los datos, sino que es resultado del m de estimación utilizado.

- Falso

Pregunta 7 8 / 8 pts

El problema de autocorrelación en los errores del modelo se debe a que:

- Existe alguna variable omitida del modelo.

Esta respuesta es correcta, ya que la omisión de variables ocasiona la presencia de estructuras del tipo de autocorrelación.

- Existe un problema de multicolinealidad en el modelo.
- El modelo incluye variables que no explican a la variable endógena.
- Existe estructura de asociación entre las variables exógenas.