UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE INGENIERÍA
DIVISIÓN DE CIENCIAS BÁSICAS
COORDINACIÓN DE CIENCIAS APLICADAS
DEPARTAMENTO DE PROBABILIDAD Y ESTADÍSTICA
SEGUNDO EXAMEN FINAL
Resoluciones

SEMESTRE 2012 - 1 — **TIPO 1**
DURACIÓN MÁXIMA 2.5 HORAS — **VIERNES, 9 DE DICIEMBRE DE 2011**

NOMBRE____
Apellido Paterno **Apellido Materno** **Nombre (s)** **Firma**

Problema 1
La siguiente muestra son las mediciones del diámetro de 36 cabezas de remache en centésimos de una pulgada.

6.62	6.66	6.7	6.72	6.75	6.76	6.78	6.82
6.62	6.66	6.7	6.72	6.76	6.76	6.78	
6.64	6.67	6.7	6.72	6.76	6.76	6.79	
6.66	6.68	6.72	6.73	6.76	6.77	6.8	
6.66	6.7	6.72	6.74	6.76	6.78	6.81	

a) Construir la tabla de distribución de frecuencias completa, comenzar en 6.62 y utilizar seis clases.
b) Calcular la media y la desviación estándar de la muestra aleatoria.
c) Construir el histograma de frecuencias relativas para la muestra aleatoria.

10 Puntos
Resolución
a) La tabla de distribución de frecuencias completa está dada por

No. Clase	Límites Aparentes		Límites Reales		Marcas de Clase	Frec. Abs.	Frec. Rel.	Frec. A. Abs.	Frec. A. Rel.
	L_i	L_s	F_i	F_s	x_i	f_i	f_i^*	F_i	F_i^*
1	6.62	6.65	6.615	6.655	6.635	3	0.083	3	0.083
2	6.66	6.69	6.655	6.695	6.675	6	0.167	9	0.250
3	6.70	6.73	6.695	6.735	6.715	10	0.278	19	0.528
4	6.74	6.77	6.735	6.775	6.755	10	0.278	29	0.806
5	6.78	6.81	6.775	6.815	6.795	6	0.167	35	0.972
6	6.82	6.85	6.815	6.855	6.835	1	0.028	36	1.000

b) La media para datos agrupados está definida por

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{6} x_i f_i = \sum_{i=1}^{6} x_i f_i^*$$

sustituyendo

$\bar{x} \approx 6.729$

La desviación estándar está definida por la raíz de la variancia, entonces

$$S_{n-1}^2 = \frac{1}{n-1}\sum_{i=1}^{6}(x_i - \bar{x})^2 f_i$$

$S_{n-1}^2 = 0.003$

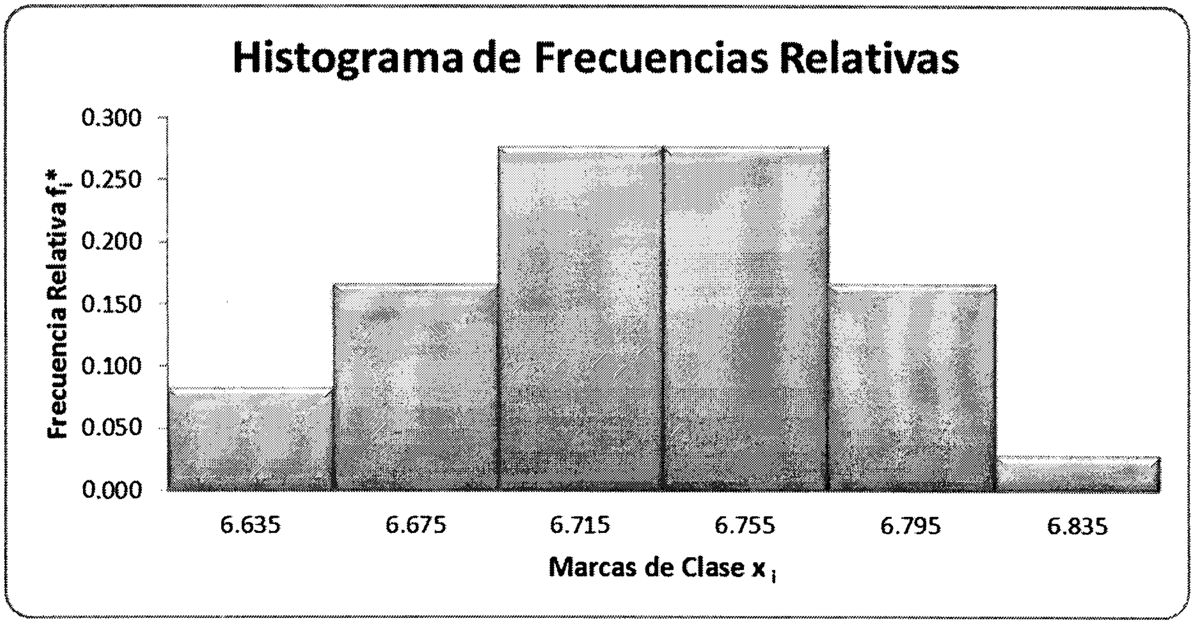
la raíz es

$$S_{n-1} = \sqrt{S_{n-1}^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{6}(x_i - \bar{x})^2 f_i}$$

al sustituir

$S_{n-1} = \sqrt{0.003} \approx 0.051$

c) El histograma de frecuencias relativas para la muestra aleatoria, con base en la tabla de a)



Problema 2

Una compañía aseguradora de automóviles clasifica a los automovilistas en tres formas: *A* (riesgo bajo), *B* (riesgo medio) y *C* (riesgo alto). En la primera se ubican el 20% de los automovilistas asegurados, el 65% en la segunda y el complemento en la tercera. También se sabe que 1, 2, 3 de cada 100 asegurados, respectivamente, llegan a tener accidentes en un periodo de un año.

a) Después de comprar una póliza con esa aseguradora, un automovilista tiene un accidente dentro del periodo del año de vigencia de la póliza, ¿cuál es la probabilidad de que el automovilista esté clasificado como de riesgo medio?

b) Si al cumplir un año un automovilista con póliza de esa aseguradora no tiene ningún accidente, ¿qué probabilidad hay de que esté clasificado como un conductor de riesgo bajo o medio?

15 Puntos

Resolución

a) Sean los eventos

A: Automovilista de riesgo bajo.
B: Automovilista de riesgo medio.
C: Automovilista de riesgo alto.
D: Accidente que puede sufrir un automovilista.

Del enunciado se tiene

$P(A) = 0.2 \qquad P(B) = 0.65 \qquad P(C) = 0.15 \qquad P(D|A) = \frac{1}{100} \qquad P(D|B) = \frac{2}{100} \qquad P(D|C) = \frac{3}{100}$

se pide calcular $P(B|D)$ que es el Teorema de Bayes, entonces

$$P(B|D) = \frac{P(B \cap D)}{P(D)} = \frac{P(B)P(D|B)}{P(A)P(D|A) + P(B)P(D|B) + P(C)P(D|C)}$$

$$= \frac{(0.65)\left(\frac{2}{100}\right)}{(0.2)\left(\frac{1}{100}\right) + (0.65)\left(\frac{2}{100}\right) + (0.15)\left(\frac{3}{100}\right)} = \frac{0.0130}{0.0195} \approx 0.6667$$

b) Se quiere calcular $P(A \cup B|\bar{D})$ que es Teorema de Bayes, sustituyendo

$$P(A \cup B|\bar{D}) = \frac{P(A \cup B \cap \bar{D})}{P(\bar{D})} = \frac{P(A \cap \bar{D} \cup B \cap \bar{D})}{P(\bar{D})} = \frac{P(A \cap \bar{D}) + P(B \cap \bar{D})}{1 - P(D)}$$

$$= \frac{P(A)P(\bar{D}|A) + P(B)P(\bar{D}|B)}{1 - [P(A)P(D|A) + P(B)P(D|B) + P(C)P(D|C)]}$$

$$= \frac{(0.2)\left(\frac{99}{100}\right) + (0.65)\left(\frac{98}{100}\right)}{1 - \left[(0.2)\left(\frac{1}{100}\right) + (0.65)\left(\frac{2}{100}\right) + (0.15)\left(\frac{3}{100}\right)\right]} = \frac{0.8350}{0.9805} \approx 0.8516$$

Problema 3

Considérese una variable aleatoria X continua, cuya función de densidad está dada por

$$f_X(x) = \begin{cases} k\cos 2x & ; \quad 0 < x < \frac{\pi}{4} \\ 0 & ; \quad en\,otro\,caso \end{cases}$$

a) Calcular el valor de k para que la función sea válida.
b) Obtener la función acumulativa.
c) Determinar el valor de la mediana.

15 Puntos

Resolución

a) Para determinar el valor de k que hace que la función sea de densidad, deberá cumplir con las propiedades

$\int_{-\infty}^{\infty} f_X(x)dx = 1$

al sustituir la función se tiene

$$\int_0^{\frac{\pi}{4}} k\cos 2x \; dx = 1$$

integrando

$$\frac{k}{2}\int_0^{\frac{\pi}{4}} \cos u \; du = 1$$

$$\int_0^{\frac{\pi}{4}} \cos u \; du = \frac{2}{k}$$

$$[sen(2x)]|_0^{\frac{\pi}{4}} = sen\left(2\left(\frac{\pi}{4}\right)\right) - sen(0) = \frac{2}{k}$$

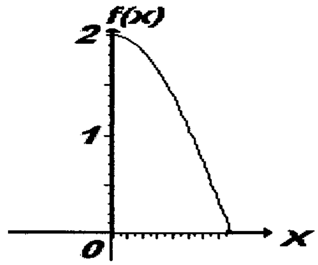
$$1 - 0 = \frac{2}{k}$$

por tanto $k = 2$

sustituyendo en la función queda dada por

$$f_X(x) = \begin{cases} 2\cos(2x) & ; \quad 0 < x < \frac{\pi}{4} \\ 0 & ; \quad en\,otro\,caso \end{cases}$$

en el intervalo $f_X(x)$ es mayor a cero con el valor obtenido, entonces es función de densidad válida.



b) Se pide determinar la función de distribución acumulada, de la definición se sabe que

$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$

sustituyendo en el caso dado

$$F_X(x) = \int_0^x 2cos(2t)dt$$

$$F_X(x) = \int_0^x 2cos(2t)dt$$

$$F_X(x) = \left(sen(2t)\right)\Big|_0^x = sen(2x) \quad en \quad 0 < x < \frac{\pi}{4}$$

por tanto la función acumulativa es

$$F_X(x) = \begin{cases} 0 & ; \quad x \leq 0 \\ sen(2x) & ; \quad 0 < x < \frac{\pi}{4} \\ 1 & ; \quad x \geq \frac{\pi}{4} \end{cases}$$

c) La mediana es el valor de la abscisa que divide en dos partes a la función de densidad, entonces

$$P(X \leq \tilde{x}) = P(X < \tilde{x}) = \frac{1}{2} = \int_{-\infty}^{x} f_X(t)dt$$

integrando e igualando, o bien, sustituyendo en la función acumulativa

$$\frac{1}{2} = \int_0^{\tilde{x}} 2cos(2t)dt = sen(2t)|_0^{\tilde{x}} = \frac{1}{2}$$

$$sen(2\tilde{x}) = \frac{1}{2}$$

$$\tilde{x} = \frac{1}{2} angsen\left(\frac{1}{2}\right)$$

entonces $\tilde{x} = \frac{\pi}{12}$ que es la medida de tendencia central a calcular.

Problema 4

La resistencia a la compresión de una serie de muestras puede modelarse mediante una distribución normal con media de 6000 kilogramos por centímetro cuadrado y una desviación estándar de 100 kilogramos por centímetro cuadrado.

a) ¿Cuál es la probabilidad de que la resistencia de una muestra sea menor que 6250 kg/cm^2?

b) ¿Cuál es la probabilidad de que la resistencia de una muestra se encuentre entre 5800 y 5900 kg/cm^2?

c) ¿Cuál es el valor de resistencia que excede el 95 % de las muestras?

15 Puntos

Resolución

Sea X la variable aleatoria que representa la resistencia a la compresión de una muestra, en kilogramos por centímetro cuadrado.

$X \sim Normal(\mu_X = 6000,\ \sigma_X = 100)$

a) Se pide calcular $P(X \leq 6250) = P(X < 6250)$, entonces

$$P(X \leq 6250) = P(X < 6250) \approx P\left(\frac{X-\mu_X}{\sigma_X} < \frac{6250-6000}{100}\right) = P\left(Z < \frac{6250-6000}{100}\right) = P(Z < 2.50)$$

de tablas de la función de distribución acumulativa normal estándar

$P(X \leq 6250) = F_Z(2.50) = 0.9938$

b) Esto es calcular $P(5800 < X < 5900)$, entonces

$$P(5800 < X < 5900) \approx P\left(\frac{5800-6000}{100} < \frac{X-\mu_X}{\sigma_X} < \frac{5900-6000}{100}\right) = P(-2.00 < Z < -1.00)$$

de tablas de la función de distribución acumulativa normal estándar

$P(5800 < X < 5900) = F_Z(-1.00) - F_Z(-2.00) = 0.1587 - 0.0228 = 0.1359$

c) El valor de resistencia que excede el 95 % de las muestras, se plantea

$P(X < x) = 0.95$

$$P\left(\frac{X-\mu_X}{\sigma_X} < \frac{x-\mu_X}{\sigma_X}\right) = 0.95$$

$$P\left(Z < z_0 = \frac{x-6000}{100}\right) = 0.95$$

de tablas de la función de distribución acumulativa normal estándar, con 0.05 y por simetría

$z_0 = 1.64$

$$\frac{x-6000}{100} = 1.64$$

despejando

$x = 1.64(100) + 6000$

$x = 1.64(100) + 6000 = 6164$

Problema 5

El tiempo de espera, en horas, que tarda un radar en detectar dos conductores sucesivos a alta velocidad es una variable aleatoria continua con función acumulativa dada por

$$F_X(x) = \begin{cases} 0 & ; \quad x \leq 0 \\ 1 - e^{-8x} & ; \quad x > 0 \end{cases}$$

a) Obtener la probabilidad de esperar menos de 12 minutos entre dos conductores sucesivos.
b) Determinar la probabilidad que el tiempo de espera sea entre 12 y 18 minutos.
c) Calcular el promedio y la variancia del tiempo de espera.

15 Puntos

Resolución

Sea X la variable aleatoria que representa el tiempo de espera entre dos automovilistas que circulan a alta velocidad.

$$X \sim Exponencial\left(\lambda = 8\ \frac{automovilistas}{hora}\right)$$

a) Se pide determinar $P(X < 12\ minutos)$, entonces

$$P\left(X < \tfrac{12}{60}\right) = P\left(X < \tfrac{1}{5}\right) = 1 - e^{-8\left(\frac{1}{5}\right)} \approx 0.7981$$

b) La probabilidad de que el tiempo de espera sea entre 12 y 18 minutos, entonces $P(12 < X < 18)$

$$P(12 < X < 18) = P\left(\frac{12}{60} < X < \frac{18}{60}\right) = F_X\left(\frac{3}{10}\right) - F_X\left(\frac{1}{5}\right) = 1 - e^{-8\left(\frac{3}{10}\right)} - \left[1 - e^{-8\left(\frac{1}{5}\right)}\right]$$

$$= 1 - e^{-8\left(\frac{3}{10}\right)} - 1 + e^{-8\left(\frac{1}{5}\right)} = e^{-8\left(\frac{1}{5}\right)} - e^{-8\left(\frac{3}{10}\right)} = e^{-\frac{8}{5}} - e^{-\frac{24}{10}} \approx 0.1111$$

c) El promedio de tiempo para detectar a dos automovilistas que circulan a alta velocidad, es la media de la distribución exponencial

$$E(X) = \frac{1}{\lambda} = \frac{1}{8} \approx 0.1250$$

La variancia de la variable aleatoria con distribución exponencial está definida por

$$Var(X) = \frac{1}{\lambda^2} = \frac{1}{64} \approx 0.0156$$

Problema 6

Se seleccionan al azar dos repuestos para una pluma de una caja que contiene tres repuestos azules, dos rojos y tres verdes. Sea X la variable aleatoria que representa el número de repuestos azules en la selección y sea Y la variable aleatoria que representa el número de repuestos rojos.

a) Determinar la distribución de probabilidad conjunta.

b) Calcular $P[(x,y) \in A]$ donde $A = \{(x,y) \mid X + Y \leq 1\}$

c) ¿Son variables aleatorias conjuntas estadísticamente independientes?

15 Puntos

Resolución

a) La distribución de probabilidad conjunta se define por

$f_{XY}(x,y) = P(X = x, Y = y) = P(X = x \cap Y = y)$

Sea X la variable aleatoria que representa el número de repuestos azules, su recorrido es

$R_X = \{0,1,2\}$

Sea Y la variable aleatoria que representa el número de repuestos rojos, su recorrido está dado por

$R_Y = \{0,1,2\}$

La función de probabilidad en forma analítica es

$$f_{XY}(x,y) = \begin{cases} \dfrac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}} & ; \quad x = 0,1,2,3 \ , \quad y = 0,1,2, \quad x + y \leq 2 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

La función de probabilidad en forma tabular, es

$$f_{XY}(0,0) = \frac{\binom{3}{0}\binom{2}{0}\binom{3}{2}}{\binom{8}{2}} = \frac{3}{28} \qquad f_{XY}(0,1) = \frac{\binom{3}{0}\binom{2}{1}\binom{3}{1}}{\binom{8}{2}} = \frac{6}{28}$$

$$f_{XY}(0,2) = \frac{\binom{3}{0}\binom{2}{2}\binom{3}{0}}{\binom{8}{2}} = \frac{1}{28} \qquad f_{XY}(1,0) = \frac{\binom{3}{1}\binom{2}{0}\binom{3}{1}}{\binom{8}{2}} = \frac{9}{28}$$

$$f_{XY}(1,1) = \frac{\binom{3}{1}\binom{2}{1}\binom{3}{0}}{\binom{8}{2}} = \frac{6}{28} \qquad f_{XY}(1,2) = 0$$

$$f_{XY}(2,0) = \frac{\binom{3}{2}\binom{2}{0}\binom{3}{0}}{\binom{8}{2}} = \frac{3}{28} \qquad f_{XY}(2,1) = 0 \qquad f_{XY}(2,2) = 0$$

entonces la forma tabular es

$f_{XY}(x,y)$		x			$f_Y(y)$
		0	1	2	
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{6}{28}$	$\frac{6}{28}$	0	$\frac{12}{28}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
$f_X(x)$		$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$	1

b) $P[(x,y) \in A]$ donde $A = \{(x,y) \mid X+Y \le 1\}$ ésto es

$$P(X+Y \le 1) = f_{XY}(0,0) + f_{XY}(0,1) + f_{XY}(1,0) = \frac{3}{28} + \frac{6}{28} + \frac{9}{28} = \frac{18}{28} = \frac{9}{14} \approx 0.643$$

c) Las variables aleatorias conjuntas son independientes, cuando se cumple la definición

$$f_{XY}(x,y) = f_X(x) f_Y(y)$$

sustituyendo se tiene

$$f_{XY}(0,0) = f_X(0) f_Y(0)$$

sustituyendo los valores de la función conjunta y de las funciones marginales, se tiene

$$\frac{3}{28} \neq \left(\frac{10}{28}\right)\left(\frac{15}{28}\right)$$

por tanto, no son estadísticamente independientes.

Problema 7

Una inversión requiere para su incremento un desembolso inicial, el cual genera un beneficio creciente durante los próximos cinco años que tiene una vida útil: Año 1, \$100; Año 2, \$150; Año 3, \$200; Año 4, \$210; Año 5, \$220.

a) Obtener la recta de regresión por el método de mínimos cuadrados.

b) ¿Existe una tendencia lineal significativa en el beneficio durante el periodo indicado? Justificar su respuesta.

c) Calcular el coeficiente de determinación e interpretar el resultado.

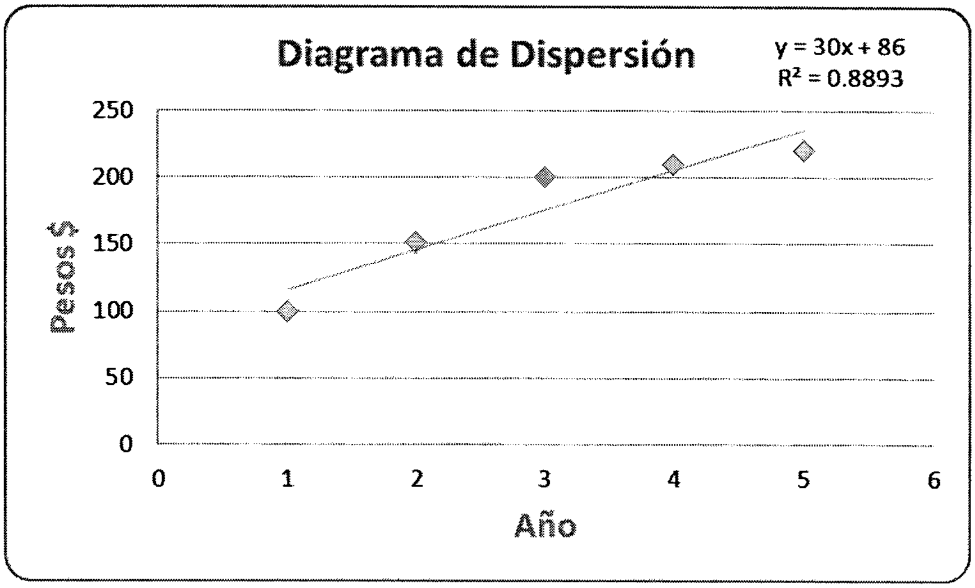
15 Puntos

Resolución

Los datos en estudio del enunciado, son

Año (x)	1	2	3	4	5
Pesos \$ (y)	100	150	200	210	220

El diagrama de dispersión es el que se muestra



a) Para obtener la recta de regresión al realizar las sumas, se tiene

	Año (x)	Pesos ($) (y)	xy	x^2	y^2
	1	100	100	1	10000
	2	150	300	4	22500
	3	200	600	9	40000
	4	210	840	16	44100
	5	220	1100	25	48400
Sumas	15	880	2940	55	165000

$$SS_{xx} = \sum_{i=1}^{5} x_i^2 - \frac{\left(\sum_{i=1}^{5} x_i\right)^2}{n}$$

sustituyendo

$$ss_{xx} = 55 - \frac{(15)^2}{5} = 10$$

$$SS_{xy} = \sum_{i=1}^{5} xy - \frac{\left(\sum_{i=1}^{5} x_i\right)\left(\sum_{i=1}^{5} y_i\right)}{n}$$

sustituyendo

$$ss_{xy} = 2940 - \frac{(15)(880)}{5} = 300$$

$$SS_{yy} = \sum_{i=1}^{5} y_i^2 - \frac{\left(\sum_{i=1}^{5} y_i\right)^2}{n}$$

sustituyendo

$$ss_{yy} = 165000 - \frac{(880)^2}{5} = 10120$$

$$\beta_1 = \frac{SS_{xy}}{SS_{xx}}$$

sustituyendo

$$\beta_1 = \frac{300}{10} = 30$$

$$\bar{X} = \frac{\sum_{i=1}^{5} x_i}{n}$$

$$\bar{x} = \frac{15}{5} = 3$$

$$\bar{Y} = \frac{\sum_{i=1}^{5} y_i}{n}$$

$$\bar{y} = \frac{880}{5} = 176$$

El estimador

$$\beta_0 = \bar{y} - \beta_1\bar{x}$$

$$\beta_0 = 176 - 30(3) = 86$$

La ecuación de la recta estimada está dada por

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1\, x$$

con los valores obtenidos

$$\hat{y} = 86 + 30\, x$$

Representa el modelo de regresión lineal de los años a partir del rendimiento obtenido.

b) Se sabe que el coeficiente de correlación está dado por el estimador

$$R = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}}$$

sustituyendo los valores obtenidos

$$r = \frac{300}{\sqrt{(10)(10120)}} \approx 0.943$$

el ajuste a una recta es de 0.943, que se puede decir que es bueno.

c) Elevando al cuadrado el coeficiente de correlación, se tiene

$$R^2 = \frac{\left(SS_{xy}\right)^2}{SS_{xx}SS_{yy}}$$

sustituyendo se tiene

$$r^2 = \frac{(300)^2}{(10)(10120)} = \frac{90000}{101200} \approx 0.889$$

Indica que los años de inversión explican el 88.9% del rendimiento de la inversión, por tanto, el modelo es bueno pues el grado de explicación es bueno.