
Segundo Semestre 2009

Curso : Probabilidad y Estadística
Sigla : EYP1113
Pauta : I3
Profesores : Ricardo Aravena (Sec 01) y Ricardo Olea (Sec 02)
Ayudante : José Quinlan y Constanza Quezada.

Problema 1

Un empresa constructora solicita a un proveedor de vigas de acero un lote de vigas cuya resistencia sea superior a 38 kg/cm^2 . Antes de enviar el lote, como parte del protocolo de control de calidad el proveedor toma una muestra 24 vigas y las somete a ensayos de resistencia cuyos resultados fueron los siguientes:

$$\bar{X} = 40 \text{ kg/cm}^2, \quad S_X = 5 \text{ kg/cm}^2$$

Asumiendo una distribución normal para las resistencias observadas.

- (a) ¿Existe evidencia estadística que permita afirmar que las vigas de acero que se enviarán a la empresa constructora presentan una resistencia media superior a los 38 kg/cm^2 ? Plantee hipótesis, indique test a utilizar y fundamente su conclusión. Use un nivel de significancia $\alpha = 5 \%$.
- (b) Un proveedor (competidor) dice que no es factible evaluar estas vigas, dado que sobrepasan con holgura la varianza máxima permitida que es de $18 \left(\text{kg/cm}^2\right)^2$. ¿Es valida la aseveración? Use un nivel de significancia $\alpha = 1 \%$.

Solución

- (a) Sean las hipótesis:

$$H_0 : \mu = 38 \quad \text{vs.} \quad H_0 : \mu > 38 \quad \text{[1.0 Ptos.]}$$

El estadístico de prueba bajo la hipótesis nula (considerando una estimación de la varianza) es:

$$T_c = \frac{\bar{X} - \mu_0}{S_X / \sqrt{n}} = \frac{40 - 38}{5 / \sqrt{24}} = 1,96 \sim t(n-1) \quad \text{[0.5 Ptos.]}$$

Se rechaza H_0 a un nivel α si $T_c > t_{1-\alpha}(n-1)$ [0.5 Ptos.]. Para este caso tenemos que

$$T_c = 1,96 > t_{1-0,05}(23) = 1,714 \quad \text{[0.5 Ptos.]}$$

Por lo tanto se rechaza H_0 , es decir, las vigas presentan una resistencia superior a los 38 kg/cm^2 solicitados. [0.5 Ptos.]

- (b) Sean las hipótesis

$$H_0 : \sigma^2 = 18 \quad \text{vs.} \quad H_0 : \sigma^2 > 18 \quad \text{[1.0 Ptos.]}$$

El estadístico de prueba bajo H_0 es

$$X^2 = \frac{(n-1) S_X^2}{\sigma^2} \sim \chi^2(n-1) \quad \text{[0.5 Ptos.]}$$

Se rechaza H_0 a un nivel α si $X^2 > \chi_{1-\alpha}^2(n-1)$ [0.5 Ptos.]. Para este caso tenemos que

$$X^2 = \frac{23 \cdot 5^2}{18} = 31,9 \not> \chi_{1-0,01}^2(23) = 41,64 \quad \text{[0.5 Ptos.]}$$

Por lo tanto, no existe suficiente evidencia para rechazar H_0 a un nivel α , es decir, NO es válida la aseveración de la competencia. [0.5 Ptos.]

+ 1 Punto Base

Problema 2

Con el fin de realizar simulaciones referente a los retrasos de los vuelos del aeropuerto de Pudahuel. Se toma una muestra 64 vuelos que sufrieron retrasos durante el mes de octubre y obtiene la siguiente tabla.

	Tiempo de Retraso (min.)			
	< 20	[20 – 30)	[30 – 45)	≥ 45
N° de Vuelos	26	16	13	9

Se proponen dos modelos de probabilidad:

- Distribución exponencial con una media de 30 minutos,
- Distribución normal con media 30 y desviación estándar de 20 minutos.

Decida, a través de una prueba estadística, su decisión. Sea explícito: hipótesis, test y criterios en los cuales basa su decisión.

Solución

Test de Bondad de Ajuste χ^2 para una distribución Exponencial y Normal.

Sean las siguientes hipótesis:

$$H_0 : \text{Datos} \sim \text{Exponencial}(\lambda = 1/30) \quad \text{vs.} \quad H_a : \text{Datos} \not\sim \text{Exponencial}(\lambda = 1/30) \quad [0.5 \text{ Ptos.}] \quad (1)$$

$$H_0 : \text{Datos} \sim \text{Normal}(\mu = 30, \sigma = 20) \quad \text{vs.} \quad H_a : \text{Datos} \not\sim \text{Normal}(\mu = 30, \sigma = 20) \quad [0.5 \text{ Ptos.}] \quad (2)$$

0.92 5.53 1.20 1.83

La siguiente tabla muestra los valores observados y esperados según las distribuciones propuestas:

			Exponencial			Normal		
Intervalo	i	O_i	p_i [0.5 Ptos.]	E_i	$(O_i - E_i)^2/E_i$	p_i [0.5 Ptos.]	E_i	$(O_i - E_i)^2/E_i$
< 20	1	26	0,49	31,36	0,92	0,31	19,84	1,91
[20 – 30)	2	16	0,15	9,60	5,53	0,19	12,16	1,21
[30 – 45)	3	13	0,14	8,96	1,20	0,27	17,28	1,06
≥ 45	4	9	0,22	14,08	1,83	0,23	14,72	2,22
Total		64	1,00	64,00	9,48	1,00	64,00	6,40

[1.0 Ptos.]

[1.0 Ptos.]

Nota: TODOS los resultados fueron redondeados al segundo decimal.

Como en ambos casos no se estimaron parámetros el estadístico de prueba $X^2 = \sum_{i=1}^4 \frac{(O_i - E_i)^2}{E_i}$ se compara con un valor $\chi^2(\nu)$, con $\nu = (k - 1) - 0 = 3$. [0.2 Ptos.]

Se rechaza H_0 si $X^2 > \chi^2_{1-\alpha}(3)$. Para un nivel $\alpha = 5\%$ se tiene que:

- Caso 1: $X^2 = 9,48 > \chi^2_{0,95}(3) = 7,81$. [0.7 Ptos.]
- Caso 2: $X^2 = 6,40 > \chi^2_{0,95}(3) = 7,81$. [0.7 Ptos.]

Por lo tanto, a un nivel del 5% de significancia la distribución Normal propuesta no es rechazada, mientras que la distribución exponencial propuesta si lo es. [0.4 Ptos.]

+ 1 Punto Base

Problema 3

Suponga que postula un modelo de regresión lineal simple sin intercepto, es decir,

$$Y = \beta X.$$

Para una muestra aleatoria $(x_1, y_1), \dots, (x_n, y_n)$, determine el estimador de mínimos cuadrados ordinario de β .

Solución

Interesa determinar la recta de regresión bajo la restricción que pase por el origen y que mejor se ajuste al comportamiento de las observaciones.

La técnica de mínimos cuadrados toma la suma de las distancias al cuadrado entre las observaciones y_i y lo planteado por el modelo $y'_i = \beta \cdot x_i$.

$$\sum_{i=1}^n (y_i - \beta x_i)^2 \quad [1.0 \text{ Ptos.}]$$

Esta suma para un conjunto dado de observaciones $(x_1, y_1), \dots, (x_n, y_n)$, es solo una función de β

$$S(\beta) = \sum_{i=1}^n (y_i - \beta x_i)^2 \quad [1.0 \text{ Ptos.}]$$

Lo que busca este método en este caso, es determinar el valor de β tal que mín $S(\beta)$. [1.0 Ptos.]

Derivando e igualando a cero se puede obtener el valor de β estimado como una función de los datos para la cual la suma de cuadrados es mínima.

$$\frac{\partial}{\partial \beta} S(\beta) = 2 \sum_{i=1}^n (y_i - \beta x_i) \cdot (-x_i) = 0 \Rightarrow \sum_{i=1}^n y_i x_i = \beta \sum_{i=1}^n x_i^2 \quad [1.0 \text{ Ptos.}]$$

Por lo tanto, el estimador de mínimos cuadrados de β está dado por:

$$\hat{\beta} = \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2} \quad [1.0 \text{ Ptos.}]$$

Como

$$\frac{\partial^2}{\partial \beta^2} S(\beta) = 2 \sum_{i=1}^n x_i^2 > 0$$

se asegura que $\hat{\beta}$ minimiza $S(\beta)$. [1.0 Ptos.]

+ 1 Punto Base

Problema 4

Se postula que el consumo per cápita de energía Y (en toneladas de carbón equivalente) es función lineal del ingreso per cápita de ingreso X (en M\$US). Para ilustrar la relación anterior, se toman 8 países en vías de desarrollo, obteniéndose los siguientes resultados

$$\sum_{i=1}^8 x_i = 38, \quad \sum_{i=1}^8 y_i = 17, \quad \sum_{i=1}^8 x_i^2 = 250, \quad \sum_{i=1}^8 y_i^2 = 45, \quad \sum_{i=1}^8 x_i y_i = 100$$

- (a) Obtenga el modelo de regresión estimado, la varianza condicional $s_{Y|x}^2$ y el coeficiente de determinación (interpretelo).
- (b) ¿Cuál es el coeficiente de correlación? Interprete y bosqueje un grafico que ilustre la relación.
- (c) Asumiendo el modelo normal, determine la probabilidad que el consumo per cápita sea inferior a 2 (en ton de carbón equivalente), dado que consideramos aquellos países cuyo ingreso per cápita alcanza los 5 (en M\$US)

Solución

- (a) Del formulario se tiene que

$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2} = \frac{100 - 8 \cdot \left(\frac{38}{8}\right) \cdot \left(\frac{17}{8}\right)}{250 - 8 \cdot \left(\frac{38}{8}\right)^2} \approx 0,28 \quad \text{[0.3 Ptos.]}$$
$$\hat{\alpha} = \bar{y} - \hat{\beta} \cdot \bar{x} = \left(\frac{17}{8}\right) - 0,28 \cdot \left(\frac{38}{8}\right) \approx 0,8 \quad \text{[0.2 Ptos.]}$$

Por lo tanto el modelo de regresión estimado esta dado por:

$$y' = 0,8 + 0,28 \cdot x$$

La varianza condicional $s_{Y|x}^2$ esta dada por

$$s_{Y|x}^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - y'_i)^2 = \frac{1}{n-2} \left[\sum_{i=1}^n (y_i - \bar{y}_i)^2 - \hat{\beta}^2 \sum_{i=1}^n (x_i - \bar{x})^2 \right], \quad \text{[0.1 Ptos.]}$$

donde

$$\sum_{i=1}^n (y_i - \bar{y}_i)^2 = \sum_{i=1}^n y_i^2 - n \bar{y}^2 = 45 - 8 \cdot \left(\frac{17}{8}\right)^2 = 8,88 \quad \text{[0.1 Ptos.]}$$
$$\sum_{i=1}^n (x_i - \bar{x}_i)^2 = \sum_{i=1}^n x_i^2 - n \bar{x}^2 = 250 - 8 \cdot \left(\frac{38}{8}\right)^2 = 69,5 \quad \text{[0.1 Ptos.]}$$

Reemplazado se tiene que

$$s_{Y|x}^2 = \frac{1}{6} [11,88 - 0,28^2 \cdot 69,5] = 0,57 \quad \text{[0.2 Ptos.]}$$

El coeficiente de determinación esta dado por

$$r^2 = 1 - \frac{s_{Y|x}^2}{s_Y^2},$$

donde

$$s_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{8,88}{7} = 1,27 \quad [0.4 \text{ Ptos.}]$$

Luego

$$r^2 = 1 - \frac{0,57}{1,27} = 0,55 \quad [0.3 \text{ Ptos.}]$$

Es decir, el 55 % de la variabilidad de y es explicada por x . [0.3 Ptos.]

(b) **Alternativa 1**

El coeficiente de correlación esta dado por:

$$\hat{\rho} = \hat{\beta} \cdot \frac{s_X}{s_Y} = \hat{\beta} \cdot \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}} = 0,28 \cdot \frac{\sqrt{69,5}}{\sqrt{8,88}} = 0,78 \quad [2.0 \text{ Ptos.}]$$

Alternativa 2

$$\hat{\rho}^2 = 1 - \frac{n-2}{n-1} \frac{s_{Y|x}^2}{s_Y^2} = 1 - \frac{6}{7} \cdot \frac{0,57}{1,27} = 0,615 \Rightarrow \hat{\rho} = 0,78 \quad [2.0 \text{ Ptos.}]$$

(c) Asumiendo el modelo normal tenemos que

$$Y | X = x \sim \text{Normal}(\hat{\alpha} + \hat{\beta} x, s_{Y|x}) \quad [0.5 \text{ Ptos.}]$$

Si $x = 5$, tenemos que

$$Y | X = 5 \sim \text{Normal}(2,2, 0,75) \quad [0.5 \text{ Ptos.}]$$

Por lo tanto

$$P(Y < 2 | X = 5) = \Phi\left(\frac{2-2,2}{0,75}\right) = \Phi(-0,27) = 1 - \Phi(0,27) = 1 - 0,6064 = 0,3936 \quad [1.0 \text{ Ptos.}]$$

+ 1 Punto Base

Formulario:

- Si $X \sim \text{Normal}(\mu, \sigma^2)$, entonces $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\}$ para $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ y $\sigma > 0$.
 $E(X) = \mu$ y $\text{Var}(X) = \sigma^2$.
- Si $X \sim \text{Exponencial}(\lambda)$, entonces $f_X(x) = \lambda e^{-\lambda x}$, para $x \in \mathbb{R}^+$ y $\lambda > 0$.
 $E(X) = 1/\lambda$ y $\text{Var}(X) = 1/\lambda^2$.
- Sean X_1, \dots, X_n una muestra aleatoria con distribución $\text{Normal}(\mu, \sigma^2)$, entonces

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \text{Normal}(0, 1), \quad \frac{\bar{X}_n - \mu}{s/\sqrt{n}} \sim t(n-1), \quad \frac{(n-1)s^2}{\sigma^2} \sim \chi_{(n-1)}^2,$$

$$\text{con } s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

- Para el modelo de regresión lineal simple $Y = \alpha + \beta x$ se tiene que

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x}, \quad \hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad r^2 = 1 - \frac{s_{Y|x}^2}{s_Y^2}$$

$$s_{Y|x}^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - y'_i)^2$$

Tablas de Percentiles p

Distribución Normal Estándar											Distribución t-student $t_p(\nu)$				
Z_p	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	ν	$t_{0,90}$	$t_{0,95}$	$t_{0,975}$	$t_{0,99}$
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359	1	3,078	6,314	12,706	31,821
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753	2	1,886	2,920	4,303	6,965
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141	3	1,638	2,353	3,182	4,541
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517	4	1,533	2,132	2,776	3,747
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879	5	1,476	2,015	2,571	3,365
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224	6	1,440	1,943	2,447	3,143
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549	7	1,415	1,895	2,365	2,998
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852	8	1,397	1,860	2,306	2,896
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133	9	1,383	1,833	2,262	2,821
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389	10	1,372	1,812	2,228	2,764
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621	11	1,363	1,796	2,201	2,718
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830	12	1,356	1,782	2,179	2,681
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015	13	1,350	1,771	2,160	2,650
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177	14	1,345	1,761	2,145	2,624
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319	15	1,341	1,753	2,131	2,602
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441	16	1,337	1,746	2,120	2,583
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545	17	1,333	1,740	2,110	2,567
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633	18	1,330	1,734	2,101	2,552
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706	19	1,328	1,729	2,093	2,539
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767	20	1,325	1,725	2,086	2,528
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817	21	1,323	1,721	2,080	2,518
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857	22	1,321	1,717	2,074	2,508
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890	23	1,319	1,714	2,069	2,500
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916	24	1,318	1,711	2,064	2,492
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936	25	1,316	1,708	2,060	2,485
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952	26	1,315	1,706	2,056	2,479
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964	27	1,314	1,703	2,052	2,473
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974	28	1,313	1,701	2,048	2,467
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981	29	1,311	1,699	2,045	2,462
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986	30	1,310	1,697	2,042	2,457
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990	∞	1,282	1,645	1,960	2,326

Distribución Chi-Cuadrado $\chi_p^2(\nu)$								
ν	$\chi_{0,025}^2$	$\chi_{0,05}^2$	$\chi_{0,10}^2$	$\chi_{0,90}^2$	$\chi_{0,95}^2$	$\chi_{0,975}^2$	$\chi_{0,99}^2$	$\chi_{0,995}^2$
1	0,00	0,00	0,02	2,71	3,84	5,02	6,63	7,88
2	0,05	0,10	0,21	4,61	5,99	7,38	9,21	10,60
3	0,22	0,35	0,58	6,25	7,81	9,35	11,34	12,84
4	0,48	0,71	1,06	7,78	9,49	11,14	13,28	14,86
5	0,83	1,15	1,61	9,24	11,07	12,83	15,09	16,75
6	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	1,69	2,17	2,83	12,02	14,07	16,01	18,48	20,28
8	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,95
9	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	3,82	4,57	5,58	17,28	19,68	21,92	24,72	26,76
12	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
17	7,56	8,67	10,09	24,77	27,59	30,19	33,41	35,72
18	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
19	8,91	10,12	11,65	27,20	30,14	32,85	36,19	38,58
20	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
21	10,28	11,59	13,24	29,62	32,67	35,48	38,93	41,40
22	10,98	12,34	14,04	30,81	33,92	36,78	40,29	42,80
23	11,69	13,09	14,85	32,01	35,17	38,08	41,64	44,18
24	12,40	13,85	15,66	33,20	36,42	39,36	42,98	45,56
25	13,12	14,61	16,47	34,38	37,65	40,65	44,31	46,93