

Examen de Introducción a la Econometría (LECO).

Departamento de Economía Cuantitativa. Universidad Complutense de Madrid.

16 de junio de 2008. Duración: **2 horas y 30 minutos**.

Apellidos:

Nombre:

DNI:

Nombre del profesor:

Grupo:

No desgrape las hojas de este cuadernillo. El examen está compuesto por tres problemas y 10 cuestiones cortas. Elija **SÓLO DOS DE LOS TRES** problemas propuestos y respóndalos en hojas aparte. Las cuestiones cortas deben ser respondidas en el espacio debajo del enunciado en este mismo cuadernillo. Junto a cada pregunta y apartado se indica la puntuación.

Problemas

Problema 1. (2.5 pt.) Se dispone de datos sobre la velocidad en km/h y la distancia recorrida por un automóvil hasta que se detiene (en metros) para cinco vehículos.

i	Vel. (x_i)	Dist. (y_i)
1	18	5
2	24	16
3	11	7
4	19	7
5	40	26
Suma	112	61

Con los datos de la tabla se han calculado los siguientes estadísticos: $nS_x^2 = 473,2$, $nS_y^2 = 310,8$ y $nS_{xy} = 357,6$. Se pide (realice los cálculos con dos decimales de precisión):

- (0.25 pt) Formule un modelo teórico de relación lineal entre distancia de frenado y velocidad del coche.
- (0.25 pt) Estime el modelo por mínimos cuadrados.
- (0.25 pt) Interprete los parámetros estimados.
- (0.25 pt) Calcule los residuos. ¿Qué vehículo frena peor?
- (0.75 pt) Contraste la hipótesis nula de que la velocidad influye en la distancia de frenado. Exprese claramente la hipótesis nula, la alternativa y la región crítica del contraste.
- (0.75 pt) Obtenga un intervalo de confianza al 95% para el efecto de la velocidad sobre la distancia de frenado. ¿Qué interpretación tiene el intervalo calculado?

Problema 2. (2.5 pt) En la centralita de una empresa hay un empleado que atiende las llamadas entrantes y que últimamente se siente saturado de trabajo. La empresa supone que las llamadas que entran por minuto siguen una distribución de Poisson. Se ha recogido, aleatoriamente, el número de llamadas que han entrado en la centralita por minuto en cien minutos diferentes. Los resultados aparecen en la siguiente tabla:

Llamadas	0	1	2	3	4	5	6	7	8
Veces	1	2	11	28	20	23	12	3	1

- (1 pt) Contraste si **los datos se ajustan** a los supuestos de la empresa (las llamadas son una $\mathcal{P}(1)$). Formule claramente la hipótesis nula, la alternativa y halle el p-valor.
- (1.5 pt) Realice un **contraste paramétrico** para la hipótesis nula de que el parámetro que caracteriza la entrada de llamadas es $\lambda = 2$ frente a la alternativa de que es mayor. Para ello:
 - (0.5 pt) Encuentre la **mejor región crítica**.
 - (0.5 pt) Halle el valor crítico a un 5% de significación.
 - (0.5 pt) Aplique el contraste obtenido a los datos de la tabla.

Nota: La función de cuantía de una distribución $\mathcal{P}(\lambda)$ es $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \dots$

Problema 3. (2.5 pt) Sea una población X que se distribuye $N(\mu, 2)$. Se quiere llevar a cabo el contraste $H_0 : \mu = 0$ frente a $H_1 : \mu = 1$ con una significación $\alpha = 0,05$. Para ello se dispone de una muestra aleatoria simple de tamaño $n = 10$. Se consideran dos regiones críticas: $RC_1 = \{\bar{x} > k_1\}$ y $RC_2 = \{\frac{x_1+x_n}{2} > k_2\}$, donde \bar{x} es la media muestral y x_1 y x_n son, respectivamente, el primer y el último elemento de la muestra.

- (0.5 pt) Halle los valores de k_1 y k_2 para que ambas regiones críticas posean el mismo nivel de significación.
- (0.5 pt) Dado que ambos contrastes tienen el mismo nivel de significación, ¿esto implica que ambos son igualmente adecuados para realizar el contraste?
- (1 pt) Calcule la potencia de ambos contrastes ¿Qué se puede decir acerca de la idoneidad de las regiones críticas?
- (0.5 pt) ¿Cómo cambiaría su respuesta a los apartados anteriores si aumentara el tamaño muestral a $n = 20$?

Cuestiones cortas

Debe obtener más de 2 puntos en estas preguntas para que se califiquen los problemas.

Enunciado para las tres cuestiones siguientes. Sea el par de variables aleatorias (X, Y) , cuya distribución de probabilidad conjunta viene dada por la siguiente tabla:

	$X = 1$	$X = 2$	$X = 3$
$Y = 0$	$1/9$	$1/3$	$2/9$
$Y = 1$	$1/18$	$1/6$	$1/9$

Cuestión 1. (0.5 pt) Halle el coeficiente de correlación lineal ρ_{XY} .

Cuestión 2. (0.5 pt) Calcule $E(Y/X = 2)$.

Cuestión 3. (0.5 pt) Formule la función generatriz de momentos conjunta $M_{XY}(t, s)$.

Cuestión 4. (0.5 pt) Demuestre que la suma de los datos de una muestra tamaño n en desviaciones a su media es cero; esto es, que $\sum_{i=1}^n (x_i - \bar{x}) = 0$.

Cuestión 5. (0.5 pt) “El p-valor de un contraste de normalidad es 0,085. Esto quiere decir que para un nivel de significación del 0,05 no se puede aceptar la hipótesis nula de que la distribución es normal” ¿Es cierta la afirmación? Razone la respuesta.

Cuestión 6. (0.5 pt) De 2.000 encuestados, 1.420 son favorables a la retirada de los parquímetros y 580 no lo son. El Ayuntamiento manifiesta que “la mayoría de la población está a favor de los parquímetros”. Contraste la veracidad de la afirmación del Ayuntamiento con un nivel de significación del $\alpha = 0,05$ **utilizando un contraste paramétrico**. Indique claramente hipótesis nula y alternativa.

Cuestión 7. (0.5 pt) Ilustre gráficamente y obtenga numéricamente el p-valor del contraste de la cuestión anterior.

Cuestión 8. (0.5 pt) Discuta la veracidad o falsedad de la siguiente afirmación: “Si se rechaza H_0 con un nivel de significación $\alpha = 0,05$ también se rechaza H_0 con un nivel de significación $\alpha = 0,10$ ”.

Cuestión 9. (0.5 pt) Suponga la siguiente función de densidad marginal de una variable aleatoria X : $f_1(x) = \frac{1}{x^2}, x \geq 1$. Compruebe si es función de densidad y calcule $E(X)$.

Cuestión 10. (0.5 pt) Sea $(X, Y)'$ una normal bivalente con vector de esperanzas $\mu = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ y matriz de varianzas-covarianzas $\Sigma = \begin{pmatrix} 1 & \sigma_{XY} \\ \sigma_{XY} & 1 \end{pmatrix}$. ¿Cuánto es la $P(X - Y > 0)$ si $\sigma_{XY} = 0$? ¿Y si $\sigma_{XY} = 0,5$?

Fórmulas de posible utilidad

Transformación de variables. Sea $X \sim f_X(x)$ y se define $Y = h(X)$. Entoces $f_Y(y) = f_X(h^{-1}(y)) \left| \frac{dh^{-1}(y)}{dy} \right|$ donde $h^{-1}(\cdot)$ es la *función inversa* de $h(\cdot)$.

Aproximación lineal a la esperanza condicional.

$$E^*(Y/X = x) = E(Y) - E(X) \frac{Cov(X, Y)}{V(X)} + E(X) \frac{Cov(X, Y)}{V(X)} \cdot x$$

Varianza condicional de la normal bivalente. $V(Y/X = x) = V(Y)(1 - \rho_{XY}^2)$.

Modelo de regresión lineal. Sea $E(Y_i/X_i = x_i) = a + bx_i$ (o también $Y_i = a + bx_i + \varepsilon_i$, $\varepsilon_i \stackrel{iid}{\sim} N(0, \sigma_\varepsilon^2)$). Si \hat{a} y \hat{b} son los estimadores por el método de los momentos (o de mínimos cuadrados) de a y b , entonces:

$$\frac{\hat{a} - a}{\sqrt{\frac{\hat{\sigma}_\varepsilon^2 \sum x_i^2}{T \sum (x_i - \bar{x})^2}}} \sim t_{n-2}; \quad \frac{\hat{b} - b}{\sqrt{\frac{\hat{\sigma}_\varepsilon^2}{\sum (x_i - \bar{x})^2}}} \sim t_{n-2}; \quad \hat{\sigma}_\varepsilon^2 = \frac{1}{n-2} \sum_{i=1}^n \hat{\varepsilon}_i^2.$$

Distribuciones de funciones de variables aleatorias. Si $X \sim N(\mu, \sigma^2)$, $X_1 \sim N(\mu_1, \sigma_1^2)$ y $X_2 \sim N(\mu_2, \sigma_2^2)$ e independientes y se dispone de muestras de tamaños n , n_1 y n_2 respectivamente:

$$\begin{aligned} \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}} &\sim N(0, 1); & \frac{\bar{x} - \mu}{\sqrt{s^2/n}} &\sim t_{n-1}; & \frac{(n-1)s^2}{\sigma^2} &\sim \chi_{n-1}^2; & \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} &\sim F_{n_1-1, n_2-1} \\ \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} &\sim N(0, 1); & \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}} &\sim t_{n+m-2} \end{aligned}$$

donde s^2 denota la *cuasivarianza* muestral.

Proporciones. $\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \rightarrow N(0, 1)$. Con dos poblaciones y muestras de tamaños n y m :

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\left(\frac{n+m}{n \cdot m}\right) \hat{p}_T (1 - \hat{p}_T)}} \rightarrow N(0, 1),$$

donde $\hat{p}_T = \frac{n\hat{p}_1 + m\hat{p}_2}{n+m}$.

Contraste de Jarque-Bera. $JB = n \left[\frac{AS^2}{6} + \frac{(K-3)^2}{24} \right] \rightarrow \chi_2^2$.

Contraste Chi cuadrado. $\sum_{i=1}^k \frac{(T_i - O_i)^2}{T_i} \sim \chi^2$ donde T_i y O_i son, respectivamente las i -ésimas frecuencias absolutas esperadas y observadas.

Contrastes de Kolmogorov-Smirnov. Para una muestra $D_n = \sup |F_n^*(x) - F(x)|$. Para dos muestras $D_{n,m} = \sup |F_n^*(x) - G_m^*(x)|$. $F_n^*(x)$ y $G_m^*(x)$ son funciones de distribución empíricas (o muestrales) y $F(x)$ es una función de distribución teórica.

Contraste de Wilcoxon. El estadístico $T = T^+ - T^-$, bajo H_0 cumple $E(T) = 0$ y $V(T) = \frac{n(n+1)(2n+1)}{6}$.

Contraste de Mann-Whitney. $U = \min(U_1, U_2)$, donde $U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$ y $U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$. Bajo H_0 se cumple $E(U) = \frac{n_1 n_2}{2}$ y $V(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$.

Aproximación a los valores críticos en los contrastes de Kolmogorov-Smirnov. Para el contraste de una muestra, el valor crítico c^* con un nivel de significación α se aproxima mediante $c_\alpha^* = k_\alpha \sqrt{1/n}$, donde k_α es 1.07, 1.22, 1.36, 1.52 y 1.63 para niveles de significación del 20 %, 10 %, 5 %, 2 % y 1 %, respectivamente.

Para el contraste de dos muestras, el valor crítico aproximado se calcula:

$$c_\alpha^* = k_\alpha \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}},$$

donde k_α es 1.07, 1.22 y 1.52 para niveles de significación α del 10 %, 5 % y 1 %, respectivamente.

Tablas estadísticas

	x.xx	x.x1	x.x2	x.x3	x.x4	x.x5	x.x6	x.x7	x.x8	x.x9
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993

Cuadro 1: Función de distribución de la $N(0, 1)$

r	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.975	0.99	0.999
2	0.45	0.71	1.02	1.39	1.83	2.41	3.22	4.61	5.99	7.38	9.21	13.82
4	1.65	2.19	2.75	3.36	4.04	4.88	5.99	7.78	9.49	11.14	13.28	18.47
6	3.07	3.83	4.57	5.35	6.21	7.23	8.56	10.64	12.59	14.45	16.81	22.46
8	4.59	5.53	6.42	7.34	8.35	9.52	11.03	13.36	15.51	17.53	20.09	26.12
10	6.18	7.27	8.30	9.34	10.47	11.78	13.44	15.99	18.31	20.48	23.21	29.59
12	7.81	9.03	10.18	11.34	12.58	14.01	15.81	18.55	21.03	23.34	26.22	32.91
14	9.47	10.82	12.08	13.34	14.69	16.22	18.15	21.06	23.68	26.12	29.14	36.12
16	11.15	12.62	13.98	15.34	16.78	18.42	20.47	23.54	26.30	28.85	32.00	39.25
18	12.86	14.44	15.89	17.34	18.87	20.60	22.76	25.99	28.87	31.53	34.81	42.31
20	14.58	16.27	17.81	19.34	20.95	22.77	25.04	28.41	31.41	34.17	37.57	45.31

Cuadro 2: Función de distribución de la χ_r^2 .

r	0.6	0.7	0.8	0.9	0.95	0.975	0.99	0.999
1	0.3249	0.7265	1.3764	3.0777	6.3138	12.7062	31.8205	318.3088
2	0.2887	0.6172	1.0607	1.8856	2.9200	4.3027	6.9646	22.3271
3	0.2767	0.5844	0.9785	1.6377	2.3534	3.1824	4.5407	10.2145
4	0.2707	0.5686	0.9410	1.5332	2.1318	2.7764	3.7469	7.1732
5	0.2672	0.5594	0.9195	1.4759	2.0150	2.5706	3.3649	5.8934
6	0.2648	0.5534	0.9057	1.4398	1.9432	2.4469	3.1427	5.2076
7	0.2632	0.5491	0.8960	1.4149	1.8946	2.3646	2.9980	4.7853
8	0.2619	0.5459	0.8889	1.3968	1.8595	2.3060	2.8965	4.5008
9	0.2610	0.5435	0.8834	1.3830	1.8331	2.2622	2.8214	4.2968
10	0.2602	0.5415	0.8791	1.3722	1.8125	2.2281	2.7638	4.1437
11	0.2596	0.5399	0.8755	1.3634	1.7959	2.2010	2.7181	4.0247
12	0.2590	0.5386	0.8726	1.3562	1.7823	2.1788	2.6810	3.9296
13	0.2586	0.5375	0.8702	1.3502	1.7709	2.1604	2.6503	3.8520
14	0.2582	0.5366	0.8681	1.3450	1.7613	2.1448	2.6245	3.7874
15	0.2579	0.5357	0.8662	1.3406	1.7531	2.1314	2.6025	3.7328
16	0.2576	0.5350	0.8647	1.3368	1.7459	2.1199	2.5835	3.6862
17	0.2573	0.5344	0.8633	1.3334	1.7396	2.1098	2.5669	3.6458
18	0.2571	0.5338	0.8620	1.3304	1.7341	2.1009	2.5524	3.6105
19	0.2569	0.5333	0.8610	1.3277	1.7291	2.0930	2.5395	3.5794
20	0.2567	0.5329	0.8600	1.3253	1.7247	2.0860	2.5280	3.5518

Cuadro 3: Función de distribución de la t_r .

λ	x												
	0	1	2	3	4	5	6	7	8	9	10	11	12
1	0.3679	0.7358	0.9197	0.9810	0.9963	0.9994	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0.1353	0.4060	0.6767	0.8571	0.9473	0.9834	0.9955	0.9989	0.9998	1.0000	1.0000	1.0000	1.0000
3	0.0498	0.1991	0.4232	0.6472	0.8153	0.9161	0.9665	0.9881	0.9962	0.9989	0.9997	0.9999	1.0000
4	0.0183	0.0916	0.2381	0.4335	0.6288	0.7851	0.8893	0.9489	0.9786	0.9919	0.9972	0.9991	0.9997
5	0.0067	0.0404	0.1247	0.2650	0.4405	0.6160	0.7622	0.8666	0.9319	0.9682	0.9863	0.9945	0.9980

Cuadro 4: Función de distribución de la $\mathcal{P}(\lambda)$.