

Examen de Introducción a la Econometría (LECO).

Departamento de Economía Cuantitativa. Universidad Complutense de Madrid.

8 de junio de 2011. Duración: **2 horas y media.**

Apellidos:

Nombre:

DNI:

Nombre del profesor:

Grupo:

No desgrape las hojas de este cuadernillo. El examen está compuesto por diez preguntas tipo test y diez cuestiones cortas. Responda a las preguntas tipo test en la plantilla de esta página. Las cuestiones tipo test suman tres puntos si la respuesta es correcta, restan un punto si es incorrecta y cero puntos si se deja en blanco. Debe obtener doce puntos en las preguntas tipo test para que se corrijan las cuestiones.

Pregunta 1	A	B	C	D	En blanco
Pregunta 2	A	B	C	D	En blanco
Pregunta 3	A	B	C	D	En blanco
Pregunta 4	A	B	C	D	En blanco
Pregunta 5	A	B	C	D	En blanco
Pregunta 6	A	B	C	D	En blanco
Pregunta 7	A	B	C	D	En blanco
Pregunta 8	A	B	C	D	En blanco
Pregunta 9	A	B	C	D	En blanco
Pregunta 10	A	B	C	D	En blanco

Puntos test:	Correctas:	Incorrectas:	En blanco:
Calificación test	Puntos/6=		
Calificación cuestiones			
Calificación total			

Preguntas test

Enunciado para las cuatro preguntas siguientes. Para determinar si un estudiante asiste *habitualmente a clase* (más de la mitad de los días), se han recogido firmas durante 10 clases. Se obtiene que la *mejor región crítica* para el contraste es $\{\sum_{i=1}^{10} X_i \geq k\}$, donde X_i indica si el estudiante asistió ($X_i = 1$) o no ($X_i = 0$) a clase el día i y se denota por p la probabilidad de asistencia. Considere $H_0 : p = 0,5$ frente a una alternativa $H_1 : p > 0,5$.

Pregunta 1. El valor crítico k para una significación *aproximada* del 5% es:

- A) 6.
- B) 7.
- C) 8.
- D) 9.

Pregunta 2. Suponga que la hipótesis alternativa es $H_1 : p = 0,9$. ¿Cuánto sería la potencia del contraste?

- A) 0,3758.
- B) 0,6778.
- C) 0,8791.
- D) 0,9298.

Pregunta 3. Si alguien decide que el valor crítico debería ser $k = 6$ ¿cuál sería la significación del contraste?

- A) 0,6230.
- B) 0,3770.
- C) 0,1719.
- D) 0,0107.

Pregunta 4. Si un estudiante ha asistido 9 de los 10 días muestreados ¿cuánto sería el p-valor?

- A) 0,0010.
- B) 0,0107.
- C) 0,0547.
- D) 0,3487.

Enunciado para las tres preguntas siguientes. Se dispone de datos de 100 familias sobre ingresos (x_i) y gastos corrientes (y_i) semanales. A partir de dichos datos se han calculado los siguientes estadísticos: $\sum x_i = 50,49$, $\sum y_i = 45,20$, $\sum x_i^2 = 34,66$, $\sum y_i^2 = 22,00$ y $\sum x_i y_i = 27,38$. Se desea estimar el modelo lineal simple $y_i = \alpha + \beta x_i + u_i$, donde u_i son otros factores que afectan al consumo familiar distintos de sus ingresos. [**NOTA: realice los cálculos con cuatro decimales de precisión.**]

Pregunta 5. La estimación por mínimos cuadrados de la pendiente, $\hat{\beta}$, es:

- A) 0,0531.
- B) 0,2007.
- C) 0,4972.
- D) 0,7900.

Pregunta 6. Sabiendo que la suma de los cuadrados de los residuos es $\sum \hat{u}_i^2 = 0,2963$, el valor del estadístico para contrastar $H_0 : \beta = 0,45$ frente a la alternativa $H_1 : \beta \neq 0,45$ es:

- A) 3,28.
- B) 2,73.
- C) 1,64.
- D) 0,28.

Pregunta 7. El p-valor del contraste de la cuestión previa es:

- A) 0,0016
- B) 0,0032.
- C) 0,0064.
- D) 0,0100

Enunciado para las tres preguntas siguientes. Sea la función de densidad continua, $f_{XY}(x, y) = k$, en el soporte $0 < x < 2 - y < 2$.

Pregunta 8. El soporte de la función de densidad es el triángulo de vértices:

- A) $(0,0), (2,0), (2,2)$.
- B) $(0,0), (0,2), (2,0)$.
- C) $(0,2), (2,0), (2,2)$.
- D) $(0,0), (0,2), (2,2)$.

Pregunta 9. ¿Cómo calcularía $P(X + Y \leq 1)$?

- A) $\int_0^1 \int_x^{1-x} f_{XY}(x, y) dy dx$.
- B) $\int_0^1 \int_0^{1-x} f_{XY}(x, y) dy dx$.
- C) $\int_0^1 \int_0^1 f_{XY}(x, y) dy dx$.
- D) $\int_0^2 \int_0^{1-x} f_{XY}(x, y) dy dx$.

Pregunta 10. ¿Cómo calcularía $P(Y > 1 | X = 1/2)$?

- A) $\int_0^{3/2} f(y | X = 1/2) dx$.
- B) $\int_0^{3/2} f(y | X = 1/2) dy$.
- C) $\int_{3/2}^2 f(y | X = 1/2) dy$.
- D) $\int_1^{3/2} f(y | X = 1/2) dy$.

Cuestiones cortas

Cuestión 1. (0.5 pt) Discuta la veracidad de la siguiente afirmación: “El coeficiente de correlación mide el grado de asociación entre dos variables aleatorias. Si la correlación es 0, entonces no están relacionadas”.

Enunciado para las tres cuestiones siguientes. Se dispone de las calificaciones de 10 estudiantes en un control de Introducción a la Econometría:

Estudiante	1	2	3	4	5	6	7	8	9	10
Calificación	5.9	4.5	6.8	4.6	5.5	4.3	4.6	5.6	5.9	3.2

Cuestión 2. (0.5 pt) Calcule el estadístico de contraste de Wilcoxon para la hipótesis de que las calificaciones son simétricas alrededor de una mediana de cinco puntos. Formule claramente hipótesis nula y alternativa.

Cuestión 3. (0.5 pt) Utilice la aproximación normal al estadístico de Wilcoxon para calcular el p-valor del contraste. ¿Rechazaría la hipótesis nula?

Cuestión 4. (0.5 pt) ¿Se podría haber utilizado otro estadístico para el contraste? ¿Qué sería necesario suponer?

Enunciado para las dos cuestiones siguientes. A partir de las estadísticas de la enseñanza universitaria 2009-10 recientemente publicadas por el INE, se construye la siguiente tabla de contingencia por estudios y sexo de una muestra representativa de los matriculados (se han excluido las dobles licenciaturas y los grados):

	Arquitectura e ingenierías técnicas	Diplomaturas	Licenciaturas	Arquitectura e ingenierías
Varones	130	98	231	96
Mujeres	39	217	340	43

Cuestión 5. (0.5 pt) Calcule el estadístico chi-cuadrado para contrastar la independencia entre sexo y estudios.

Cuestión 6. (0.5 pt) Obtenga el p-valor y deje clara su conclusión.

Cuestión 7. (0.5 pt) Considere una muestra aleatoria simple de tamaño n extraída de una población X , que es una variable aleatoria con distribución Poisson de parámetro λ . Obtenga el test más potente para la hipótesis nula $H_0 : \lambda = \lambda_0$ frente a la alternativa $H_1 : \lambda > \lambda_0$. [NOTA: Recuerde que la función de cuantía de una distribución Poisson es $Pr[X = x] = \frac{e^{-\lambda} \lambda^x}{x!}$, $x = 0, 1, 2, \dots$ y que la suma de variables Poisson es también Poisson.]

Enunciado para las tres cuestiones siguientes. Se desea analizar el margen de un conjunto de empresas pequeñas, cuyo comportamiento de ingresos (X) y gastos (Y) viene representado por
$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N \left[\begin{pmatrix} 6 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} \right].$$

Cuestión 8. (0.5 pt) ¿Cuánto es la probabilidad de que el margen sea superior a 2?

Cuestión 9. (0.5 pt) ¿Cuál es el valor esperado de los gastos si los ingresos son 8?

Cuestión 10. (0.5 pt) ¿Cuánto es el margen esperado para las empresas que tienen unos ingresos de 8?

Fórmulas de posible utilidad

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

Aproximación lineal a la esperanza condicional.

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

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 , y $\hat{\varepsilon}_i$ los residuos del modelo, entonces:

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

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:

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

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_1 y n_2 :

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\left(\frac{n_1+n_2}{n_1 \cdot n_2}\right) \hat{p}_T(1 - \hat{p}_T)}} \rightarrow N(0, 1),$$

donde $\hat{p}_T = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$.

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.x0	x.x1	x.x2	x.x3	x.x4	x.x5	x.x6	x.x7	x.x8	x.x9
0.0	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	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	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	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
3	1.01	1.42	1.87	2.37	2.95	3.66	4.64	6.25	7.81	9.35	11.34	16.27
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
5	2.34	3.00	3.66	4.35	5.13	6.06	7.29	9.24	11.07	12.83	15.09	20.52
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
7	3.82	4.67	5.49	6.35	7.28	8.38	9.80	12.02	14.07	16.01	18.48	24.32
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

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

p	0	1	2	3	4	5	6	7	8	9	10
0.5	0.0010	0.0107	0.0547	0.1719	0.3770	0.6230	0.8281	0.9453	0.9893	0.9990	1.0000
0.6	0.0001	0.0017	0.0123	0.0548	0.1662	0.3669	0.6177	0.8327	0.9536	0.9940	1.0000
0.7	0.0000	0.0001	0.0016	0.0106	0.0473	0.1503	0.3504	0.6172	0.8507	0.9718	1.0000
0.8	0.0000	0.0000	0.0001	0.0009	0.0064	0.0328	0.1209	0.3222	0.6242	0.8926	1.0000
0.9	0.0000	0.0000	0.0000	0.0000	0.0001	0.0016	0.0128	0.0702	0.2639	0.6513	1.0000

Cuadro 3: Función de distribución de la $B(10, p)$.

Operaciones