

## Examen final de Introducción a la Econometría

26 de junio de 2006, a las 12:00 horas  
Duración: 2 HORAS y 15 minutos

Apellidos:	Nombre:	DNI:
Profesor/a:		Grupo:

No desgrape estas hojas

### Problemas (2.5 puntos por pregunta)

Los siguientes problemas deben ser realizados en un **ÚNICO** cuadernillo aparte

*(escriba sus datos personales, grupo y profesor en dicho cuadernillo<sup>1</sup>)*

Elija **SÓLO** dos (2) de los tres (3) problemas propuestos.

**Ejercicio 1.** Sean  $X$  e  $Y$  variables aleatorias **independientes** con función generatriz de momentos  $(1-t^2)^{-1}$  (distribución de Laplace) y con **esperanza nula**.

(a) Sea  $U = X + Y$ . Demuestre que la función generatriz de momentos de  $U$  es  $(1-t^2)^{-2}$ .

*Solución:*

$$\begin{aligned} M_U(t) &= E\left(e^{t(X+Y)}\right) = E\left(e^{tX}\right) \cdot E\left(e^{tY}\right) && \text{por independencia} \\ &= M_X(t) \cdot M_Y(t) = (1-t^2)^{-2} \end{aligned}$$

**Ejercicio 1**

(b) Sea  $V = X - Y$ . Demuestre que  $U$  y  $V$  tienen idéntica distribución.

*Solución:*

$$\begin{aligned} M_V(t) &= E\left(e^{t(X-Y)}\right) = E\left(e^{tX}\right) \cdot E\left(e^{-tY}\right) && \text{por independencia} \\ &= M_X(t) \cdot M_Y(-t) = (1-t^2)^{-2} \end{aligned}$$

Por tanto tienen idéntica distribución debido al teorema de unicidad de las funciones generatrices de momentos.

**Ejercicio 1**

(c) Calcule la función generatriz de momentos conjunta de  $U$  y  $V$

*Solución:*

$$\begin{aligned} M_{UV}(s, t) &= E\left(e^{sU+tV}\right) = E\left(e^{s(X+Y)+t(X-Y)}\right) \\ &= E\left(e^{(s+t)X+(s-t)Y}\right) = E\left(e^{(s+t)X}\right) \cdot E\left(e^{(s-t)Y}\right) && \text{por independencia} \\ &= M_X(s+t) \cdot M_Y(s-t) = (1-(s+t)^2)^{-2} \cdot (1-(s-t)^2)^{-2} \end{aligned}$$

**Ejercicio 1**

(d) Demuestre que  $U$  y  $V$  no son independientes

*Solución:* Y puesto que la función generatriz conjunta  $M_{U,V}(s, t)$  es distinta de

$$M_U(s)M_V(t) = (1-s^2)^{-2} \cdot (1-t^2)^{-2}$$

$U$  y  $V$  no son independientes.

**Ejercicio 1**

(e) No obstante, demuestre que  $U$  y  $V$  están incorreladas (**pista:** recuerde que  $E(X) = E(Y) = 0$ , y que  $E(X^2) = E(Y^2)$  puesto que ambas variables tienen idéntica distribución.)

*Solución:* Basta con demostrar que  $\text{Cov}(U, V) = 0$ .

$$\text{Cov}(U, V) = E(UV) - E(U)E(V).$$

Por una parte,

$$\begin{aligned} E(U) &= E(X + Y) = E(X) + E(Y) = 0 + 0 \\ E(V) &= E(X - Y) = E(X) - E(Y) = 0 - 0; \end{aligned}$$

<sup>1</sup> Al finalizar entregue estas paginas dentro de dicho cuadernillo

por otra

$$E(UV) = E((X + Y)(X - Y)) = E(X^2 - Y^2) = E(X^2) - E(Y^2) = 0.$$

Por tanto  $\text{Cov}(U, V) = 0$ .

**Ejercicio 1**

**Ejercicio 2.** Se cree que las terrazas y chiringuitos de verano tienen unos ingresos fijos “a”, más unos ingresos que dependen de los metros cuadrados que ocupan (además de otras circunstancias independientes de su superficie, como son su situación, la proximidad a otros locales, la calidad del servicio, etc.)

La variable  $Y$  hace referencia a los ingresos de toda la temporada de verano en **miles de euros**. La variable  $X$  son los metros cuadrados ocupados por un establecimiento.

Establecimiento	$y_i$	$x_i$	$x_i y_i$	$x_i^2$	$\hat{y}_i$	$\hat{e}_i$	$x_i \hat{e}_i$	$\hat{e}_i^2$
A	27	5	135	25	28	-1	-5	1
B	23	3	69	9	18	5	15	25
C	35	6	210	36	33	2	12	4
D	9	2	18	4	13	-4	-8	16
E	36	7	252	49	38	-2	-14	4
sumas	130	23	684	123	130	0	0	50

**Cuadro 1:**

Además se sabe que las varianzas y covarianzas muestrales son:

$$T \cdot s_x^2 = 17.2, \quad T \cdot s_y^2 = 480, \quad T \cdot s_{xy} = 86,$$

donde  $T$  es el tamaño muestral.

Suponga que plantea el siguiente modelo

$$Y_i = a + bx_i + U_i,$$

donde  $U_i$  son los otros factores que afectan a los ingresos distintos de la superficie ocupada. Se sabe que la distribución conjunta de dichos factores es:  $U \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ , donde  $\mathbf{I}$  es una matriz identidad de orden 5, y  $\sigma^2$  es la varianza de  $U_i$ , cuyo valor es desconocido.

(a) Estime por MCO los parámetros del modelo. ¿Cual es el valor esperado del ingreso fijo de un establecimiento? ¿En cuanto aumenta el ingreso esperado de un local si hace crecer su terraza en un metro cuadrado?

*Solución:*

1. Por una parte:

$$\hat{b} = \frac{\widehat{\sigma_{XY}}}{\widehat{\sigma_X^2}} = \frac{s_{xy}}{s_x^2} = \frac{86}{17.2} = 5$$

por otra, las medias muestrales son

$$\bar{x} = \frac{\sum x_i}{5} = \frac{130}{5} = 26; \quad \bar{y} = \frac{\sum y_i}{5} = \frac{23}{5} = 4.6;$$

por lo que

$$\hat{a} = \bar{y} - \hat{b} \cdot \bar{x} = 26 - 4.6 \cdot 5 = 3$$

El valor esperado del ingreso fijo es de 3000 euros. Y el incremento esperado del ingreso al hacer crecer la terraza en un metro adicional es de 5000 euros.

**Ejercicio 2**

(b) Estime por MCO la pendiente del modelo, pero dando un intervalo de confianza del 95 %.

*Solución:* El estimador MCO se distribuye Normal con esperanza igual al verdadero valor de los parámetros estimados, y varianza desconocida.

■ (**Parámetro b**) buscamos los valores  $A$  y  $B$  tales que

$$P \left( A \leq \frac{(\hat{b} - b)}{\sqrt{\frac{s^2}{\sum (x_i - \bar{x})^2}}} \leq B \right) = (1 - \alpha)$$

Donde  $\frac{(\hat{b} - b)}{\sqrt{\frac{s^2}{\sum (x_i - \bar{x})^2}}}$  se distribuye como una  $t$  de Student con  $T-2$  grados de libertad; por tanto

$$\begin{aligned} & \left[ -3.182 \leq \frac{(5 - b)}{\sqrt{\frac{16.6}{17.2}}} \leq 3.182 \right] \\ & [-3.182 \times 0.98 \leq (5 - b) \leq 3.182 \times 0.98] \\ & [5 + 3.1323 \geq b \geq 5 - 3.1323] \end{aligned}$$

es decir, el intervalo de confianza al 0.95 es

$$b \in [1.87, 8.13]$$

**Ejercicio 2**

- (c) Un empresario hostelero dice que el rendimiento de un metro adicional de establecimiento rinde una cantidad de (al menos) seis mil euros (pero nunca menos). Contraste dicha hipótesis con una significación del 10%. Bajo la hipótesis del empresario ¿cuál es el p-valor de la estimación del parámetro  $b$ ?

Solución: Las hipótesis son:

$$H_0 : b = 6$$

$$H_1 : b < 6$$

La región crítica de una sola cola es

$$RC = \left\{ \mathbf{x} \text{ tales que } \frac{\hat{b} - 6}{\sqrt{\frac{\hat{s}^2}{\sum(x_i - \bar{x})^2}}} < k \right\},$$

donde  $k$  es el valor de la tablas para una  $t$  de Student de tres grados de libertad, ya que el estadístico de la parte izquierda de la desigualdad tiene dicha distribución. Para  $\alpha = 0.1$ , tenemos que  $k = t_{3,0.1} = -1.638$  sustituyendo tenemos que

$$\frac{5 - 6}{0.98} = -1.02$$

por lo que no rechazamos  $H_0$  con un 10% de significación.

El  $p$ -valor es la probabilidad de

$$\begin{aligned} P_{H_0}(\hat{b} \leq 5) &= P_{H_0} \left( \frac{\hat{b} - b}{\sqrt{\frac{\hat{s}^2}{\sum(x_i - \bar{x})^2}}} \leq \frac{5 - b}{\sqrt{\frac{\hat{s}^2}{\sum(x_i - \bar{x})^2}}} \right) \\ &= P \left( W \leq \frac{5 - 6}{0.98} = -1.02 \right) = 0.19, \end{aligned}$$

donde  $W$  se distribuye como una  $t$  de Student con tres grados de libertad.

**Ejercicio 2**

- (d) ¿Qué establecimiento ha ingresado más respecto al nivel esperado, dada su superficie? ¿y cual menos?

Solución: El establecimiento  $B$  ha ingresado 5000 euros por encima de su ingreso esperado.

El establecimiento  $D$  ha ingresado 4000 euros por debajo de lo esperado.

**Ejercicio 2**

- (e) Dado el modelo estimado, ¿cuál es el ingreso esperado para un establecimiento con una superficie de 4 metros cuadrados?

Solución: La estimación del ingreso esperado para un establecimiento de 4 metros es:

$$\hat{Y} = 3 + 5 \times 4 = 23$$

**Ejercicio 2**

**Ejercicio 3.** Suponga que las variables aleatorias  $X$  e  $Y$  tienen distribución conjunta

$$f_{XY}(x, y) = \begin{cases} ky^2 & 0 < x < y < 1 \\ 0 & \text{en el resto de casos} \end{cases}$$

- (a) Dibuje el soporte y calcule el valor de  $k$

Solución:

$$1 = k \int_0^1 \int_0^y y^2 dx dy = k \int_0^1 y^3 dy = \frac{k}{4}$$

por tanto,  $k = 4$

**Ejercicio 3**

- (b) Calcule  $\text{Var}(Y)$

Solución:

$$f_Y(y) = 4 \int_0^y y^2 dx = 4y^3; \quad \text{para } 0 < y < 1$$

$$E(Y) = 4 \int_0^1 y^4 dy = \frac{4}{5}; \quad E(Y^2) = 4 \int_0^1 y^5 dy = \frac{2}{3}$$

$$\text{Var}(Y) = E(Y^2) - (E(Y))^2 = \frac{2}{75}$$

**Ejercicio 3**

(c) Calcule  $E_{XY}(X|y)$

Solución:

$$f_{XY}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{4y^2}{4y^3} = \frac{1}{y} \quad \text{para } 0 < x < y$$
$$E_{XY}(X|y) = \int_0^y x f_{XY}(x|y) dx = \frac{\int_0^y x dx}{y} = \frac{y}{2}$$

**Ejercicio 3**

(d) Empleando el T<sup>a</sup> de las esperanzas iteradas, calcule  $E(X)$ ; es decir, emplee la propiedad de que *la esperanza de la esperanza condicional es la esperanza incondicional*

Solución:

$$E(X) = E(E(X|Y)) = \frac{E(Y)}{2} = \frac{2}{5}$$

**Ejercicio 3**

(e) Derive la expresión general  $E(X^r Y^s)$  (donde  $r$  y  $s$  son números enteros positivos).

Solución:

$$E(X^r Y^s) = \int_0^y \int_0^1 x^r y^s f_{XY}(x,y) dx dy = 4 \int_0^1 \int_0^y x^r dx y^{s+2} dy$$
$$= \frac{4 \int_0^1 y^{s+r+3} dy}{r+1} = \frac{4}{(r+1)(s+r+4)}$$

**Ejercicio 3**

(f) Empleando la expresión obtenida en el apartado anterior calcule  $\text{Corr}(X, Y)$  (tendrá que sustituir  $r$  y  $s$  por los números enteros que necesite para cada momento calculado, por ejemplo  $E(XY) = E(X^1 Y^1)$ , es decir  $r = 1$  y  $s = 1$ ).

Solución:

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{4}{12} - \frac{2}{5} \cdot \frac{4}{5} = \frac{1}{75}$$
$$\text{Var}(Y) = E(Y^2) - (E(Y))^2 = \frac{4}{12} - \frac{4}{25} = \frac{14}{225}$$
$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \sqrt{\frac{3}{28}}$$

**Ejercicio 3**

(g) Calcule  $P(2X > Y | Y < 1/2)$

Solución:

$$P(2X > Y | Y < 1/2) = \frac{P(2X > Y, Y < 1/2)}{P(Y < 1/2)}$$

por una parte

$$P(2X > Y, Y < 1/2) = \int_0^{\frac{1}{2}} \int_{y/2}^y 4y^2 dx dy = 2 \int_0^{\frac{1}{2}} y^3 dy = \frac{1}{32}$$

por otra

$$P(Y < 1/2) = 4 \int_0^{\frac{1}{2}} y^3 dy = \frac{1}{16};$$

así pues

$$P(2X > Y | Y < 1/2) = \frac{1}{2}$$

**Ejercicio 3**

## Preguntas cortas (0.5 puntos por pregunta)

Las siguientes cuestiones deben ser respondidas en el **recuadro** que queda a continuación de cada pregunta.

**Ejercicio 4.** Se pretende contrastar si la excelencia de los alumnos se mantiene entre dos cursos consecutivos (curso 2005 y curso 2006). Para ello se dividen las calificaciones de 1200 alumnos en las siguientes categorías: “1: suspenso”, “2: aprobado”, “3: aprobado con nota”. La información muestral se resume del siguiente modo:

	(1) suspensos	(2) aprobados	(3) con nota
2005	240	190	70
2006	310	270	120

- a) Calcule la tabla de frecuencias absolutas esperadas bajo  $H_0$ . b) Para esta muestra, el estadístico Chi cuadrado arroja un valor aproximado de 2.7 ¿Cuanto vale el p-valor? (emplee el valor más próximo de los que aparecen en las tablas). c) Lleve a cabo el contraste Chi cuadrado para niveles de significación del 5 % y del 10 %.

*Solución:*  $H_0$ : idéntica distribución; N° de alumnos en 2005: 500; N° de alumnos en 2006: 700;

Probabilidades estimadas (bajo  $H_0$ ):  $P(\text{suspension}) = \frac{240+310}{1200} = 0.458$ ;  $P(\text{aprobado}) = 0.384$ ;  $P(\text{nota}) = 0.158$

Tabla de frecuencias absolutas esperadas bajo  $H_0$ :

	suspensos	aprobados	con nota
2005	$0.458 \times 500 = 229.1$	191.7	79.2
2006	320.8	268.4	110.8

El estadístico se distribuye como una Chi cuadrado con dos grados de libertad y  $\sum_{i=1}^4 \frac{(T_i - O_i)^2}{T_i} = 2.74$ ; por tanto el p-valor es aproximadamente 0.25; esto quiere que se rechaza  $H_0$  para niveles de significación superiores al 25 %. Así que al 5 % y al 10 % no se rechaza  $H_0$ . **Ejercicio 4**

**Ejercicio 5.** Calcule el valor del contraste de homogeneidad de Kolmogorov-Smirnov para dos muestras con los datos del ejercicio anterior.

*Solución:* Las funciones de distribución empíricas son:

	1	2	3
2005	0.48	0.86	1
2006	0.44	0.82	1
diferencia	0.04	0.04	0

Por lo tanto 0.04

**Ejercicio 5**

**Ejercicio 6.** Un canal de televisión realiza una encuesta a 5000 menores de 30 años, y la siguiente tabla de contingencia resume la información acerca del gusto por las películas de Van Damme:

	gusta	no gusta
chicos	348	3152
chicas	82	1418

Calcule la tabla de frecuencias absolutas teórica (o esperada) y el p-valor del estadístico Chi-cuadrado para contrastar la independencia entre el gusto por las películas y el sexo (sabiendo que para esta muestra el estadístico arroja un valor de 26.76). ¿Para qué niveles de de significación rechazaría  $H_0$ ?

*Solución:*  $H_0$ : independencia. Bajo la nula, las probabilidades estimadas son:

$P(\text{chico}) = \frac{348+3152}{5000} = 0.7$ ;  $P(\text{chica}) = 0.3$ ;  $P(\text{que guste}) = \frac{348+82}{5000} = 0.0860$ ;  $P(\text{que no guste}) = 0.914$ ;

Por tanto la tabla de frecuencias absolutas esperadas bajo  $H_0$  es

	gusta	no gusta
chicos	$0.7 \times 0.086 \times 5000 = 301$	3199
chicas	129	1371

Y el estadístico (que se distribuye como una Chi cuadrado con un grado de libertad) es  $\sum_{i=1}^4 \frac{(T_i - O_i)^2}{T_i} = 26.76$

El p-valor es aproximadamente cero; esto quiere decir que casi para cualquier nivel de significación se rechaza  $H_0$ .

**Ejercicio 6**

**Ejercicio 7.** Considere una variable aleatoria discreta  $X$ , de la que desconoce su ley de probabilidad. Usted tiene dos hipótesis posibles para la posible función de cuantía de  $X$ :

	$X$	1	2	3	4	5	6
Bajo $H_0$	$P_X(x)$	1/6	1/6	1/6	1/6	1/6	1/6
Bajo $H_1$	$P_X(x)$	2/15	1/6	1/5	1/5	1/6	2/15

Para contrastar  $H_0$ , decide realizar el siguiente contraste empleando una muestra de tamaño uno:

*Rechazar  $H_0$  si se observan los valores 4 o 5 (y no rechazar si se observa cualquier otro valor)*

Halle el nivel de significación (probabilidad de cometer el error tipo I) y la probabilidad de cometer el error tipo II.

*Solución:* La probabilidad de cometer el error tipo I es  $P_{H_0}(\text{rechazar}) = P_{H_0}(X = 4, 5)$

$$P_{H_0}(X = 4 \text{ o } 5) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

La probabilidad de cometer el error tipo II es  $P_{H_1}$  (no rechazar)

$$P_{H_1}(X = 1, 2, 3, \text{ o } 6) = \frac{2}{15} + \frac{1}{6} + \frac{1}{5} + \frac{2}{15} = \frac{19}{30}$$

### Ejercicio 7

**Ejercicio 8.** La variable aleatoria  $X$  tiene la siguiente función de densidad:  $f_x(x) = 2x/a^2$  definida en el intervalo  $[0, a]$ , con  $a > 0$ . Queremos contrastar  $H_0: a = 1$  frente a  $H_1: a > 1$ . Se dispone de una muestra aleatoria simple de tamaño 1 y de dos tests que vienen caracterizados por las siguientes regiones críticas:

**Test 1:**  $RC_1 = \{x|x > 0.987\}$

**Test 2:**  $RC_2 = \{x|x < 0.158\}$ .

Calcule el nivel de significación de ambos contrastes.

*Solución:*

$$\text{Test 1: } P_{H_0}(\text{rechazar}) = \int_{0.987}^1 2x \, dx = 0.0258$$

$$\text{Test 2: } P_{H_0}(\text{rechazar}) = \int_0^{0.158} 2x \, dx = 0.0249$$

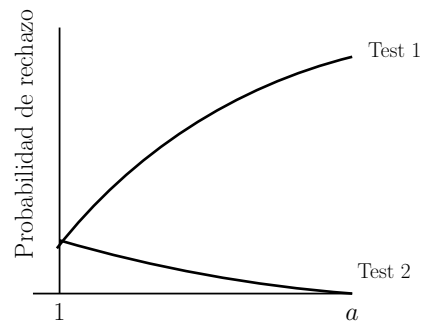
### Ejercicio 8

**Ejercicio 9.** Dibuje en un mismo gráfico las funciones potencia de los contrastes de hipótesis de la pregunta anterior.

*Solución:*

$$\text{Test 1: } W(a) = P(\text{rechazar}) = \frac{1}{a^2} \int_{0.987}^a 2x \, dx = 1 - \left(\frac{0.987}{a}\right)^2$$

$$\text{Test 2: } W(a) = P(\text{rechazar}) = \frac{1}{a^2} \int_0^{0.158} 2x \, dx = \frac{0.158^2}{a^2}$$



### Ejercicio 9

**Ejercicio 10.** Sea  $X$  una variable aleatoria que sigue una distribución  $N(\mu, \sigma^2)$ . Se desea contrastar  $H_0: \sigma^2 = 100$  frente a  $H_1: \sigma^2 < 100$  empleando el contraste uniformemente más potente. Se tiene una muestra de tamaño 31 con una cuasi-varianza  $\mathfrak{s}^2 = 75$ . Resuelva el contraste con una significación del 5% y del 1%.

*Solución:* La región crítica óptima es: el conjunto de muestras tales que  $\mathfrak{s}^2 < \lambda$ , donde  $\lambda$  satisface  $P_{H_0}(\widehat{\mathfrak{s}}^2 < \lambda) = \alpha$ , es decir  $P_{H_0}\left(\frac{30\widehat{\mathfrak{s}}^2}{100} < \frac{30\lambda}{100}\right) = \alpha$ , donde  $\frac{30\widehat{\mathfrak{s}}^2}{100}$  se distribuye  $\chi_{30}^2$ .

1. Si se dispone de las tablas de una Chi cuadrado: Para  $\alpha = 0.05$  ha de ser  $\frac{30\lambda}{100} = 18.49$ , o bien  $\lambda = 61.633$ , y por tanto para dicho  $\alpha$  no se rechaza  $H_0$ . Así pues, tampoco se rechaza para  $\alpha = 0.01$ .
2. Si NO se dispone de las tablas de una Chi cuadrado: entonces empleamos la aproximación a una distribución normal.  $\frac{30\widehat{\mathfrak{s}}^2}{100}$  se distribuye  $\chi_{30}^2$ , cuyo valor esperado es 30 y cuya varianza es 60; por tanto, lo podemos aproximar por una  $N(30, 60)$ . Así pues,  $P_{H_0}\left(\frac{30\widehat{\mathfrak{s}}^2}{100} < \frac{30\lambda}{100}\right) = \alpha$  ó bien  $P_{H_0}\left(\frac{30\widehat{\mathfrak{s}}^2/100 - 30}{\sqrt{60}} < \frac{30\lambda/100 - 30}{\sqrt{60}}\right) = P_{H_0}\left(Z < \frac{30\lambda/100 - 30}{\sqrt{60}}\right) = \alpha$ , donde aproximadamente  $Z \sim N(0, 1)$ . Para  $\alpha = 0.05$  ha de ser  $\frac{30\lambda/100 - 30}{\sqrt{60}} = -1.64$  o bien  $\lambda = 57.65$ , y por tanto para dicho  $\alpha$  no se rechaza  $H_0$ . Así pues, tampoco se rechaza para  $\alpha = 0.01$ .

### Ejercicio 10

**Ejercicio 11.** Sean  $X$  e  $Y$  los volúmenes de venta de dos bienes producidos por una empresa, con distribución normal bivalente con vector de medias  $(10 \ 15)'$  y matriz de varianzas covarianzas  $\begin{pmatrix} 9 & 2 \\ 2 & 9 \end{pmatrix}$ . ¿Calcule la probabilidad de que el volumen de ventas del primer producto supere el volumen de ventas del segundo?

*Solución:* Sea  $W = X - Y$ , entonces  $P(X > Y) = P(W > 0)$  y donde  $W \sim N(-5, 14)$

$$P(W > 0) = P\left(\frac{W+5}{\sqrt{14}} > \frac{0+5}{\sqrt{14}}\right) = P(Z > 1.3363) = 1 - P(Z \leq 1.3363) = 1 - 0.9082 = 0.0918$$

**Ejercicio 11**

**Ejercicio 12.** Con los datos del ejercicio anterior, ¿Cual es el la mejor predicción que puede hacer de las ventas de  $Y$  si sabe que las ventas del otro producto han sido  $X=12$ ?

*Solución:* Puesto que la esperanza condicional el predictor óptimo,

$$E_{Y|X}(Y|x) = E(Y) - E(X) \frac{\text{Cov}(Y, X)}{\text{Var}(X)} + \frac{\text{Cov}(X, Y)}{\text{Var}(X)} \cdot x = 15 - 10 \frac{2}{9} + \frac{2}{9} \cdot x$$

que evaluada para  $X=12$  es  $E_{Y|X}(Y|12) = 15 - 10 \frac{2}{9} + \frac{2}{9} \cdot 12 = 15.44$ ;

**Ejercicio 12**

**Ejercicio 13.** Sea la siguiente función de cuantía conjunta para  $X$  e  $Y$

$P_{XY}(x, y)$	$y=0$	$y=1$	$y=2$
$x=0$	0.33	0.12	0.05
$x=1$	0	0.25	0.25

Calcule  $E_{XY}(X|2)$ .

*Solución:*

$P_{XY}(x, y)$	$y=0$	$y=1$	$y=2$
$x=0$	0.33	0.12	0.05
$x=1$	0	0.25	0.25
Marginal de $Y$	0.33	0.37	0.3

por tanto

$P_{XY}(X 2)$	
$x=0$	0.05/0.3
$x=1$	0.25/0.3

y entonces  $E_{XY}(X|2) = 0 \times 0.05/0.3 + 1 \times 0.25/0.3 = 0.25/0.3 = 0.83$

**Ejercicio 13**

### Formulas de posible utilidad

**Transformación de variables** Sea  $f_X(x)$  en el soporte  $\mathbb{R}_X = (a, b)$ . Sea  $Y = h(X)$ , entonces  $f_Y(y) = \left| \frac{dh^{-1}(y)}{dy} \right| \cdot f_X(h^{-1}(y))$ ; en el soporte  $\mathbb{R}_Y = (h(a), h(b))$ ; donde  $h^{-1}(\cdot)$  es la *función inversa* de  $h(\cdot)$ , es decir  $h^{-1}(h(x)) = x$ .

**Función generatriz de momentos conjunta** del vector aleatorio  $\mathbf{X} = (X_1, \dots, X_n)$  evaluada en el vector  $\mathbf{t} = (t_1, \dots, t_n)$ :  $M_{\mathbf{X}}(\mathbf{t}) = M_{X_1, \dots, X_n}(t_1, \dots, t_n) = E(e^{t_1 X_1 + \dots + t_n X_n})$

### Aproximación lineal a la esperanza condicional

$$E(Y|X=x) \approx g^*(x) = E(Y) - E(X) \frac{\text{Cov}(X, Y)}{\text{Var}(X)} + \frac{\text{Cov}(X, Y)}{\text{Var}(X)} \cdot x; \quad \forall x \in \mathbb{R}_X.$$

**Varianza condicional de la normal bivalente:**  $\text{Var}_{Y|X}(Y|x) = \sigma_Y^2 - \frac{(\sigma_{XY})^2}{\sigma_X^2} = \sigma_Y^2(1 - \rho_{XY}^2)$ ,

**Modelo de regresión lineal** Sea el modelo  $Y_t = a + bX_t + U_t$ , donde  $(\mathbf{U}|\mathbf{X}) \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ ; y  $t = 1, 2, \dots, T$ . Sean  $\hat{a}$  y  $\hat{b}$  los estimadores MCO de  $a$  y  $b$ . Entonces

$$\hat{a} \sim N\left(a, \frac{\sigma^2 \sum x_t^2}{T \sum (x_t - \bar{x})^2}\right); \quad \hat{b} \sim N\left(b, \frac{\sigma^2}{\sum (x_t - \bar{x})^2}\right).$$

**Distribuciones de funciones de variables aleatorias** si  $X \sim N(\mu, \sigma^2)$ ,  $X_1 \sim N(\mu_1, \sigma_1^2)$  y  $X_2 \sim N(\mu_2, \sigma_2^2)$  con muestras de tamaños  $n, n_1$  y  $n_2$  respectivamente:

$$\frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}} \sim N(0, 1); \quad \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

$$\frac{\bar{x} - \mu}{\sqrt{s^2/n}} \sim t_{n-1}; \quad \frac{(n-1)s^2}{\sigma^2} \sim \chi_{n-1}^2; \quad \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F_{n_1-1, n_2-1}$$

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{n_1+n_2-2}, \text{ si } \sigma_1^2 = \sigma_2^2$$

donde  $s^2$  denota la *cuasivarianza* muestral  $\left(\frac{\sum (x - \bar{x})^2}{n-1}\right)$ .

**Proporciones**  $\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \overset{a}{\sim} N(0, 1)$ . Con muestras de tamaños  $n$  y  $m$ :

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{(n+m)\hat{p}_T(1-\hat{p}_T)}{n \cdot m}}} \overset{a}{\underset{H_0}{\sim}} N(0, 1); \text{ donde } \hat{p}_T = \frac{n\hat{p}_1 + m\hat{p}_2}{n+m}; H_0 : p_1 = p_2$$

**Contraste de Jarque-Bera**  $JB = n \left[ \frac{AS^2}{6} + \frac{(K-3)^2}{24} \right] \sim \chi^2_{(2)}$

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

**Kolmogorov-Smirnov para una sola muestra:**  $D_n = \sup |F_n(x) - F(x)|$ , donde  $F_n(x)$  es la función de distribución empírica (o muestral), y  $F(x)$  es la función de distribución de  $H_0$ .

Tamaño muestral	nivel de significación				
	0.2	0.1	0.05	0.02	0.01
8	0.3583	0.4096	0.4543	0.5065	0.5418
9	0.3391	0.3875	0.4300	0.4796	0.5133
10	0.3226	0.3687	0.4092	0.4566	0.4889
> 40	$1.07/\sqrt{n}$	$1.22/\sqrt{n}$	$1.36/\sqrt{n}$	$1.52/\sqrt{n}$	$1.63/\sqrt{n}$

**Kolmogorov-Smirnov para dos muestras:**  $D_n = \sup |F_1(x) - F_2(x)|$ . Cuando  $n_1$  y  $n_2$  son grandes; y donde  $F_1(x)$  y  $F_2(x)$  son funciones de distribución empíricas (o muestrales)

$$\text{Nivel crítico: } D_{\alpha, n_1, n_2} \approx k \sqrt{\frac{n_1 + n_2}{n_1 n_2}},$$

donde  $k = 1.07; 1.22; 1.52$ ; para  $\alpha$  igual a 10%, 5% y 1% respectivamente.

$\nu$	Probabilidad acumulada desde 0 hasta $x$ para $X \sim \chi^2_\nu$ con $\nu$ grados de libertad										
	60.0%	66.7%	75.0%	80.0%	87.5%	90.0%	95.0%	97.5%	99.0%	99.5%	99.9%
1	0.708	0.936	1.323	1.642	2.354	2.706	3.841	5.024	6.635	7.879	10.828
2	1.833	2.197	2.773	3.219	4.159	4.605	5.991	7.378	9.210	10.597	13.816
3	2.946	3.405	4.108	4.642	5.739	6.251	7.815	9.348	11.345	12.838	16.266
4	4.045	4.579	5.385	5.989	7.214	7.779	9.488	11.143	13.277	14.860	18.467
5	5.132	5.730	6.626	7.289	8.625	9.236	11.070	12.833	15.086	16.750	20.515

$\nu$	Probabilidad acumulada desde 0 hasta $x$ para $X \sim \chi^2_\nu$ con $\nu$ grados de libertad										
	0.1%	0.5%	1.0%	2.5%	5.0%	10.0%	12.5%	20.0%	25.0%	33.3%	50.0%
25	8.649	10.520	11.524	13.120	14.611	16.473	17.184	18.940	19.939	21.461	24.337
26	9.222	11.160	12.198	13.844	15.379	17.292	18.021	19.820	20.843	22.399	25.336
27	9.803	11.808	12.879	14.573	16.151	18.114	18.861	20.703	21.749	23.339	26.336
28	10.391	12.461	13.565	15.308	16.928	18.939	19.704	21.588	22.657	24.280	27.336
29	10.986	13.121	14.256	16.047	17.708	19.768	20.550	22.475	23.567	25.222	28.336
30	11.588	13.787	14.953	16.791	18.493	20.599	21.399	23.364	24.478	26.165	29.336
35	14.688	17.192	18.509	20.569	22.465	24.797	25.678	27.836	29.054	30.894	34.336
40	17.916	20.707	22.164	24.433	26.509	29.051	30.008	32.345	33.660	35.643	39.335

$\nu$	Probabilidad acumulada desde $-\infty$ hasta $x$ para $t$ de STUDENT con $\nu$ grados de libertad										
	60.0%	66.7%	75.0%	80.0%	87.5%	90.0%	95.0%	97.5%	99.0%	99.5%	99.9%
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
2	0.289	0.500	0.816	1.061	1.604	1.886	2.920	4.303	6.965	9.925	22.327
3	0.277	0.476	0.765	0.978	1.423	1.638	2.353	3.182	4.541	5.841	10.215
4	0.271	0.464	0.741	0.941	1.344	1.533	2.132	2.776	3.747	4.604	7.173
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
26	0.256	0.436	0.684	0.856	1.177	1.315	1.706	2.056	2.479	2.779	3.435
27	0.256	0.435	0.684	0.855	1.176	1.314	1.703	2.052	2.473	2.771	3.421
28	0.256	0.435	0.683	0.855	1.175	1.313	1.701	2.048	2.467	2.763	3.408
29	0.256	0.435	0.683	0.854	1.174	1.311	1.699	2.045	2.462	2.756	3.396
30	0.256	0.435	0.683	0.854	1.173	1.310	1.697	2.042	2.457	2.750	3.385
35	0.255	0.434	0.682	0.852	1.170	1.306	1.690	2.030	2.438	2.724	3.340

Probabilidad acumulada desde $-\infty$ hasta $z$ para $Z \sim N(0, 1)$										
$z$	x.x0	x.x1	x.x2	x.x3	x.x4	x.x5	x.x6	x.x7	x.x8	x.x9
0.0	<b>0.5000</b>	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	<b>0.5987</b>	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	<b>0.6985</b>	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	<b>0.7486</b>	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	<b>0.7995</b>	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	<b>0.8508</b>	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	<b>0.8997</b>	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	<b>0.9495</b>	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	<b>0.9750</b>	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	<b>0.9898</b>	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