Dpto. de Economía Cuantitativa
Universidad Complutense de Madrid

Examen de Introducción a la Econometría

21 de junio de 2005, 12:00h.
Duración: 2 HORAS y 30 MINUTOS

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

No desgrape las hojas de este cuadernillo

Problemas largos (3 puntos por problema)

Los siguientes problemas deben ser realizados en un **ÚNICO** cuadernillo a parte

(*escriba sus datos personales, grupo y profesor en dicho cuadernillo*[1])

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

Ejercicio 1. (Consta de 4 apartados)
La variable aleatoria X mide los rendimientos semanales de un determinado Fondo de Inversión. La información que presentamos a continuación resume las rentabilidades de los últimos 100 días. Los datos de las rentabilidades han sido previamente estandarizadas, por lo que tienen media cero y varianza uno. Sus realizaciones han sido ordenadas en 7 intervalos (pérdidas mayores al 1.5 %, pérdidas de entre el uno y el uno y medio, etc.).

Rentabilidades estandarizadas	< -1.5	$[-1.5; -1)$	$[-1; -0,5)$	$[-0,5; 0,5]$	$(0,5; 1]$	$(1; 1.5]$	> 1.5
Nº de días	9	6	12	41	19	8	5

(a) Queremos contrastar (al 10 % de significación) si los rendimientos (estandarizados) siguen una distribución N(0,1). Lleve a cabo el contraste por los tres métodos vistos en el curso (Jarque-Bera, Kolmogorov-Smirnov, y Chi cuadrado) y discuta si cambian las conclusiones de un contraste a otro
(**Dato**: el coeficiente de curtosis de los datos estandarizados es 3,23 y el de asimetrıa -0,24).

Solución:

1. **Jarque-Bera:** $RC = \{\mathbf{x} \,|\, JB > 4.605\}$

$$100 \cdot \left(\frac{-0.24^2}{6} + \frac{0.23^2}{24} \right) = 1.18 < 4.605 = \chi^2_{2,\ 90\,\%}$$

Por lo que no rechazamos H_0 de normalidad.

2. **Kolmogorov-Smirnov:** Primero calculo la probabildad de cada intervalo bajo H_0 empleando las tablas de la $N(0\,,\,1)$:
   - $a = P(x > 1.5) = 1 - .9332 = 0.0668$
   - $b = P(1 < x \leq 1.5) = (1 - .8413) - a = 0.0919$
   - $c = P(0.5 < x \leq 1) = (1 - .6915) - a - b = 0.1498$
   - $b = P(-0.5 < x \leq 0.5) = 2 \cdot P(0 > x \leq 0.5) = 2 * (.6915 - .5) = 0.383$

   Y puesto que la función de densidad de una Normal es simétrica, no es necesario calcular más probabilidades:

[1] Al finalizar entregue estas paginas dentro de dicho cuadernillo

Tabla probabilidades teóricas bajo H_0:

Intervalos	< -1.5	$[-1.5; -1)$	$[-1; -0,5)$	$[-0,5; 0,5]$	$(0,5; 1]$	$(1; 1.5]$	> 1.5
prob.$_{H_0}$	0.0668	0.0919	0.1498	0.383	0.1498	0.0919	0.0668

Acumulando las probabilidades teóricas obtenemos la función de distribución teórica bajo H_0, $F_Z(z)$; y para calcular la función de distribución empírica, $F_n(x)$, debemos dividir el número de días de la tabla del enunciado por 100, y acumular

z	-1.5	-1	$-0,5$	$0,5$	1	1.5	∞
$F_Z(z)$	0.0668	0.1587	0.3085	0.6915	0.8413	0.9332	1
$F_n(x)$	0.0900	0.1500	0.2700	0.6800	0.8700	0.9500	1
D_n = Max diff			0.0385				

Y puesto que $0.0385 < \frac{1.22}{\sqrt{100}} = 0.122$ no podemos rechazar H_0 de distribución $N(0\,,\,1)$ al 10 % de significación.

3. **Chi cuadrado:** Sólo queda calcular las frecuencias relativas (dividiendo el número de días de la tabla del enunciado por 100) para poder calcular este contraste; que dá:

$$\frac{(0.0668 \;-\; 0.0900)^2}{0.0900} + \frac{(0.0919 \;-\; 0.0600)^2}{0.0600} + \cdots + \frac{(0.0668 \;-\; 0.0500)^2}{0.0500} = 0.048$$

Puesto que este valor es menor que lo que aparece en las tablas de la Chi cuadrado con *seis grados de libertad* (10.64) no podemos rechazar H_0 con un nivel de significación del 10 %.

Fin **ejercicio** 1

(b) Suponga que, a partir de los datos originales sin estandarizar, comprueba que la media muestral es 1,15 y su cuasi-varianza muestral es 0,04; y teniendo en cuenta la conclusión a la que ha llegado en el apartado anterior, quiere contrastar que la media poblacional de la variable X es igual a 1 frente a que es mayor que 1. Defina correctamente la hipótesis nula y la alternativa de este contraste y la región crítica más apropiada. Resuelva el contraste para un nivel de significación del 5 % y del 10 %.

Solución: Puesto que en el apartado anterior no hemos rechazado la hipótesis de distribución normal, podemos emplear los estadísticos basados en distribución normal.

En este caso podemos emplear el estadístico: $\frac{\bar{x}-\mu}{\sqrt{\mathfrak{s}^2/n}} \sim t_{n-1}$

1. $H_0 : \mu = 1; \qquad H_1 : \mu > 1$
2. $RC = \left\{ \mathbf{x} \mid \frac{\bar{x}-1}{\sqrt{\mathfrak{s}^2/100}} > t_{99,\,\alpha} \right\}$
3. Por una parte, $t_{99,\,5\,\%} = 1.645$, y $t_{99,\,10\,\%} = 1.282$; por otra $\frac{1.15-1}{\sqrt{0.04/100}} = 7.5$ que supera ambos niveles críticos. Por tanto se rechaza H_0 tanto al 10 % como al 5 % de nivel de significción.

Fin **ejercicio** 1

(c) Halle el p-valor del contraste anterior. ¿Qué significado tiene el p-valor? ¿Qué implicaciones tiene a la hora de rechazar o no la hipótesis nula?

Solución:

- El valor 7.5 es tan elevado, que no aparece en las tablas por lo que podemos afirmar que el p-valor es aproximadamente cero.
- Es la probabilidad, calculada bajo H_0, de obtener un valor numérico que aporte más evidencia en contra de H_0, que la aportada por el estadístico que hemos obtenido (en nuestro caso que 7.5).
- Cuanto menor es el p-valor, menor ha de ser el nivel de significación (menor que el p-valor) para poder aceptar la hipótesis nula.

Fin **ejercicio** 1

(d) Explique (sin necesidad de resolver el contraste) cómo cambiaría el planteamiento del contraste del apartado (b) si la conclusión del apartado (a) hubiese sido la contraria a la que usted ha llegado.

Solución: Si hubiesemos rechazado la hipótesis de normalidad, no habriamos podido emplear el estadístico $\frac{\bar{x}-\mu}{\sqrt{s^2/n}}$ Fin **ejercicio** 1

Ejercicio 2. (Consta de 5 apartados)
Las demandas de dos bienes, X e Y, se pueden aproximar por una normal bivariante con vector de medias $\begin{pmatrix} 5 & 10 \end{pmatrix}'$ y matriz de varianzas covarianzas $\begin{pmatrix} 4 & 1 \\ 1 & 9 \end{pmatrix}$.

(a) Calcule el coeficiente de correlación lineal entre las demandas de los dos bienes. Según su resultado, ¿los bienes son complementarios o sustitutivos?

Solución:

$$\operatorname{Corr}(X,Y) = \rho_{XY} = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}} = \frac{1}{2\cdot 3} = \frac{1}{6}$$

Puesto que la relación lineal es positiva, son bienes complementarios. Fin **ejercicio** 2

(b) ¿Cuál es la probabilidad de que la demanda del bien X supere la del bien Y?

Solución: Puesto que $P(X > Y) = P(X - Y > 0)$; vamos a definir la variable W=X-Y que por ser combinación lineal de normales tiene distribución normal, por tanto:

$$W \sim N(5-10\,,\, 4+9-2\cdot 1) = N(-5\,,\, 11)$$

Así pues $P(X > Y) = P(W > 0)$ y calculando

$$P(W>0) = P\left(\frac{W-(-5)}{\sqrt{11}} > \frac{0-(-5)}{\sqrt{11}}\right) = P(Z > 1.51) =$$
$$= 1 - P(Z \leq 1.51) = 0.0655$$

Fin **ejercicio** 2

(c) Calcule la esperanza de la demanda del bien Y. ¿Cuál es la esperanza de la demanda del bien Y si del bien X se han demandado 8 unidades?

Solución:

1. $\operatorname{E}(Y) = 10$
2.

$$\operatorname{E}_{Y|X}(Y|\,8) = \operatorname{E}(Y) - \operatorname{E}(X)\frac{\operatorname{Cov}(Y,X)}{\operatorname{Var}(X)} + \frac{\operatorname{Cov}(X,Y)}{\operatorname{Var}(X)}\cdot 8 =$$
$$= 10 - 5\frac{1}{4} + \frac{1}{4}\cdot 8 = 10.750$$

Fin **ejercicio** 2

(d) Suponga que los consumidores son precio aceptantes, y que los precios (en euros) de estos dos bienes son $p_X = 3$ y $p_Y = 1$. Si la renta es de 30 euros ¿Cuál es la probabilidad de que tenga que pedir prestado (gasto de más de 30 euros) al consumir una cesta de bienes X e Y?

Solución: La probabilidad de pedir prestado es la probabilidad de que el valor de la cesta de bienes supere la renta de 30 euros; es decir $P(3\cdot X + Y > 30)$.

Si definimos la nueva variable $W = 3\cdot X + Y \sim N(3\cdot 5 + 10\,,\, 9\cdot 4 + 9 + 2\cdot 3\cdot 1)$, es decir $W \sim N(25\,,\, 51)$, el calculo se reduce a

$$P(W > 30) = P\left(\frac{W-25}{\sqrt{51}} > \frac{30-25}{\sqrt{51}}\right) = P(Z > 0.7) = 1 - P(Z \leq 0.7) = 0.242$$

Fin **ejercicio** 2

(e) ¿Cómo cambiaría la respuesta del apartado anterior si sabemos que del bien X se han demandado 8 unidades?

Solución: Primero calculamos cómo es la distribución conjunta de W y X. Puesto que sabemos que ha de ser conjuntamente normal, sólo necesitamos calcular $\text{Cov}\,(W, X)$

$$\text{Cov}\,(W,X) = \text{Cov}\,(3\cdot X + Y, X) = 3\cdot\text{Cov}\,(X,X) + \text{Cov}\,(Y,X) = 3\cdot 4 + 1 = 13$$

Por tanto

$$\begin{bmatrix} W \\ X \end{bmatrix} \sim N\left(\begin{bmatrix} 25 \\ 5 \end{bmatrix}, \begin{bmatrix} 51 & 13 \\ 13 & 4 \end{bmatrix}\right)$$

Y por tanto, W condicionada a X tiene distribución normal

$$\left[W|X=8\right] \sim N\left(\text{E}_{W|X}\,(W\,|\,8)\;,\;\text{Var}_{W|X}\,(W\,|\,8)\right)$$

Calculando $\text{E}_{W|X}\,(W\,|\,8)$ y $\text{Var}_{W|X}\,(W\,|\,8)$ ya podemos responder a la pregunta

- Por una parte

$$\text{E}_{W|X}\,(W\,|\,8) = \text{E}\,(W) - \text{E}\,(X)\,\frac{\text{Cov}\,(W,X)}{\text{Var}\,(X)} + \frac{\text{Cov}\,(X,W)}{\text{Var}\,(X)}\cdot 8 =$$

$$= 25 - 5\frac{13}{4} + \frac{13}{4}\cdot 8 = 34.750$$

- Por otra parte

$$\text{Var}_{W|X}\,(W\,|\,8) = \text{Var}\,(W)\left(1-\rho^2_{WX}\right) \simeq 51\cdot(1-0.91^2) = 8.77$$

donde $\rho_{WX} = \frac{\text{Cov}(W,X)}{\sqrt{\text{Var}(W)\text{Var}(X)}} = \frac{13}{\sqrt{51\cdot 4}} \simeq 0.91$

Por tanto $\left[W|X=8\right] \sim N\,(34.75\,,\,8.77)$

Realizando los cálculos de manera similar al apartado anterior se calcula la probabilidad solicitada.

$$P\,(W > 30\,|\,X=8) = P\left(\frac{W-34.75}{\sqrt{8.77}} > \frac{30-34.75}{\sqrt{8.77}}\right) =$$

$$= P\,(Z > -1.6) = P\,(Z \leq 1.6) = 0.9452$$

Fin **ejercicio** 2

Ejercicio 3. (Consta de 9 apartados)

Sea (X,Y) una variable aleatoria bidimensional continua con función de densidad conjunta

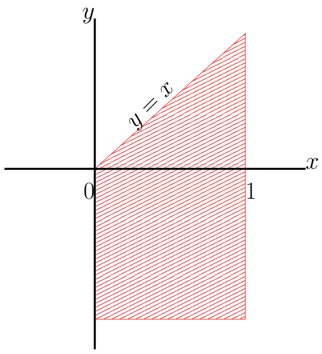
$$f_{XY}\,(x,y) \text{ definida sobre el soporte } -1 \leq y \leq x; \quad 0 \leq x \leq 1$$

Deje indicado cómo calcularía las siguientes cuestiones:

(En este ejercico no se le pide realizar calculos, tan sólo plantear y/o expresar adecuadamente lo requerido en cada apartado; por ello *preste especial atención en indicar cláramente los **límites de integración** y **dominio** de las funciones de densidad en cada caso; ¡no hacerlo significará no haber respondido a la pregunta!*):

(a) Dibuje el soporte de $f_{XY}\,(x,y)$.

Solución:



Fin **ejercicio** 3

(b) ¿Cómo verificaría que $f_{XY}(x, y)$ es función de densidad?

1. Primero integrando respecto de y y después respecto de x
2. Primero integrando respecto de x y después respecto de y

Solución: Comprobando si se verifican las siguientes igualdades

1.
$$\int_0^1 \int_{-1}^x f_{XY}(x, y)\, dy\ dx = 1$$

2.
$$\int_0^1 \int_y^1 f_{XY}(x, y)\, dx\ dy + \int_{-1}^0 \int_0^1 f_{XY}(x, y)\, dx\ dy = 1$$

Fin **ejercicio** 3

(c) Las funciones de densidad marginales de X y Y, especificando explicitamente sus soportes (dominio de cada función de densidad).

Solución:

$$f_X(x) = \int_{-1}^x f_{XY}(x, y)\, dy; \qquad x \in [0,\ 1]$$

$$f_Y(y) = \begin{cases} \int_y^1 f_{XY}(x, y)\, dx; & y \in (0,\ 1] \\ \\ \int_0^1 f_{XY}(x, y)\, dx; & y \in [-1,\ 0] \end{cases}$$

Fin **ejercicio** 3

(d) ¿Cómo verificaría si X e Y son independientes o no?

Solución: Debería comprobarse si se verifica

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y) \ \text{ para } \ -1 \le y \le x; \quad 0 \le x \le 1$$

en caso afirmativo, X e Y son independientes.

Fin **ejercicio** 3

(e) La Varianza de X

Solución: Puesto que $\text{Var}(X) = \text{E}\left(X^2\right) - \left[\text{E}(X)\right]^2$, entonces

$$\text{Var}(X) = \int_0^1 \int_{-1}^x x^2 f_{XY}(x, y)\, dy\ dx - \left[\int_0^1 \int_{-1}^x x f_{XY}(x, y)\, dy\ dx\right]^2$$

o lo que es lo mismo

$$\text{Var}(X) = \int_0^1 x^2 f_X(x)\, dx - \left[\int_0^1 x f_X(x)\, dx\right]^2$$

Fin **ejercicio** 3

(f) La esperanza de Y condicionanda a $X = x$.
Solución: Para ello es necesario calcular la función de densidad condicionada, que en este caso es

$$f_{Y|X}(y \mid x_0) = \frac{f_{XY}(x_0, y)}{f_X(x_0)}; \qquad -1 \le y \le x_0$$

por lo tanto

$$\mathrm{E}_{Y|X}(Y \mid x) = \int_{-1}^{x} y \frac{f_{XY}(x,y)}{f_X(x)} dy; \qquad x \in [-1,\ 0]; \quad -1 < y < x.$$

Fin **ejercicio** 3

(g) Cual sería la expresión exacta de la anterior esperanza condicional en el caso de ser una función lineal de x.
Solución:

$$\mathrm{E}_{Y|X}(Y \mid x) = \mathrm{E}(Y) - \mathrm{E}(X) \frac{\mathrm{Cov}(Y,X)}{\mathrm{Var}(X)} + \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)} \cdot x; \qquad \forall x \in [0,\ 1].$$

Fin **ejercicio** 3

(h) $P(Y > d \mid X = 0.5)$, donde $-1 \le d \le 1$.
Solución:

$$P(Y > d \mid X = 0.5) = \int_{d}^{0.5} f_{Y|X}(y \mid 0.5)\, dy = \int_{d}^{0.5} \frac{f_{XY}(0.5, y)}{f_X(0.5)} dy \qquad \text{si } d < 0.5$$

$$P(Y > d \mid X = 0.5) = 0 \qquad \text{si } d \ge 0.5$$

Fin **ejercicio** 3

(i) $P(Y > d \mid X > .5)$, donde $-1 \le d \le 1$.
Solución: En este caso no podemos emplear la función de densidad condicionada $f_{Y|X}(y \mid x)$; aquí debemos emplear directamente la definición de probabilidad condicionada.

$$P(Y > d \mid X > 0.5) = \frac{\int_{0.5}^{1} \int_{d}^{x} f_{XY}(x,y)\, dy\, dx}{\int_{0.5}^{1} f_X(x)\, dx} \qquad \text{si } d < 0.5$$

$$P(Y > d \mid X > 0.5) = \frac{\int_{d}^{1} \int_{d}^{x} f_{XY}(x,y)\, dy\, dx}{\int_{0.5}^{1} f_X(x)\, dx} \qquad \text{si } d \ge 0.5$$

Fin **ejercicio** 3

Sólo se le pide que **deje indicado** cómo calcular cada apartado, no que lo calcule (por lo tanto, preste especial atención a dejar cláramente indicados los límites de integración y/o los soportes allí donde es necesario).

Preguntas cortas (0.5 puntos por pregunta)

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

▼ comienzo de un grupo de preguntas ▼

El siguiente texto es válido para las tres siguientes preguntas:

Sea Y la creación de empleo trimestral de las 17 Comunidades Autónomas (CCAA) españolas (excluidas Ceuta y Melilla) y X un indicador que mide el crecimiento en el nivel de actividad de la comunidad a lo largo del último año. Queremos estimar un modelo de regresión lineal simple que relacione Y en funcion de X. La media muestral de Y es 0.04, la de X es 0.03; sus desviaciones típicas son 0.2 y 0.15, respectivamente, y la covarianza es 0.01.

Ejercicio 4. Presente el modelo de regresión lineal y estime por MCO sus parámetros.

Solución: El modelo es $Y_n = a + x_n \cdot b + U_n$

Las estimaciones MCO de a y b son:

$$\widehat{b} = \frac{s_{xy}}{s_x^2} = \frac{0.01}{0.15^2} = 0.44; \qquad \widehat{a} = \overline{y} - \overline{x} \cdot \frac{s_{xy}}{s_x^2} = 0.04 - 0.03 \cdot 0.44 = 0.027;$$

Las estimación MCO de cada y_n es: $\widehat{y_n} = 0.027 + 0.44 \cdot x_n$

Fin **ejercicio** 4

Ejercicio 5. La siguiente tabla muetra los residuos MCO de la estimación anterior correspondientes a cada una de las 17 comunidades autonomas (CCAA)

CCAA Nº	residuo	CCAA Nº	residuo	CCAA Nº	residuo
1	0.79	7	0.24	13	-0.68
2	1.00	8	0.67	14	1.10
3	0.84	9	0.45	15	0.61
4	-0.51	10	-0.81	16	-0.60
5	-0.33	11	1.54	17	-2.40
6	-1.33	12	-0.60		

¿Qué interpretación tienen los residuos en el marco de este modelo? En función de dicha interpretación, comente los aspectos más destacables de esta tabla de residuos con relación a cómo ha sido la creación de empleo en las diferentes CCAA en el trimestre.

Solución: El residuo n-ésimo es: $\widehat{e_n} = y_n - \widehat{y_n} = y_n - \operatorname{E}_{Y_n|X_N}(Y_n \mid x_n)$ es decir, es la diferencia entre la creación de empleo observada en la comunidad n-ésima y la creación de empleo *esperada* dado el dato de crecimiento de la actividad de dicha comunidad.

Por tanto, residuos positivos indican mayor creación de empleo de lo esperado (dado el crecimiento de la actividad), y residuos negativos indican menor creación de empleo de lo esperado (dado el crecimiento de la actividad).

Por tanto, la comunidad que parece haber tenido un comportamiento mejor (dada la información disponible) es la comunidad número 11, y la que peor lo ha tenido es la número 17.

Fin **ejercicio** 5

Ejercicio 6. Suponga que desea llevar a cabo el contraste de significación individual de la pendiente del modelo de regresión. Indique qué hipótesis alternativa es la más apropiada para este contraste. ¿Qué estadístico de contraste y qué región crítica son más apropiados?.

Solución: No cabe esperar que crecimientos positivos del nivel de actividad perjudiquen la creación de empleo; por lo tanto el contraste debe ser de una sóla cola. Así pues:

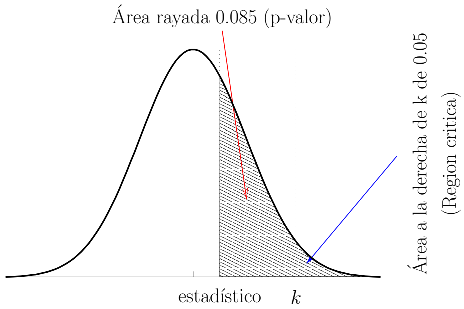
$$H_0 : b = 0; \quad H_1 : b > 0; \quad \text{estadístico: } \frac{(\widehat{b} - b)}{\sqrt{\frac{\widehat{s^2}}{\sum(x_i - \overline{x})^2}}} \sim t_{15}; \quad RC = \left\{ \mathbf{x} \;\middle|\; \frac{\widehat{b}}{\sqrt{\frac{\widehat{s^2}}{\sum(x_i - \overline{x})^2}}} > k \right\}$$

Fin **ejercicio** 6

■ fin del grupo de preguntas ■

Ejercicio 7. ¿Es cierta la siguiente afirmación? *"El p-valor de un contraste de normalidad es 0,085. Esto quiere decir que con un nivel de significación del 0,05 rechazaría que la distribución es normal"*. Razone su respuesta.

Solución: **La afirmación es falsa.** Un p-valor igual a 0.085 significa que la probabilidad, calculada bajo H_0, de obtener un valor numérico que aporte más evidencia en contra de H_0, que la aportada por el estadístico que hemos obtenido es del 8.5 %; es decir, el area que queda a la derecha del estadístico es de 0.085 (mayor por tanto al área de la región critica, de solo 0.05). Por tanto, el estadístico empleado para contrastar normalidad queda fuera de la región crítica (a la izquierda de k en el dibujo).



Fin **ejercicio** 7

Ejercicio 8. Para puestos de trabajo similares, sean $X \sim N\left(\mu_X\,,\, \sigma_X^2\right)$ los salarios de los hombres e $Y \sim N\left(\mu_Y\,,\, \sigma_Y^2\right)$ los salarios de las mujeres. Se quiere contrastar si los salarios de los hombres son mayores que los de las mujeres (si hay discriminación salarial).

1. Proponga la hipótesis nula y alternativa que considere más apropiada para llevar a cabo este contraste.
2. Proponga el estadístico de contraste y la región crítica para llevar a cabo este contraste, suponiendo independencia entre X e Y y conocidas las varianzas $(\left(\sigma_X^2, \sigma_Y^2\right) = (4,9))$.

(NO resuelva el contraste, limitese a contestar lo que se le pide)

Solución: Lamentablemente, y dada la experiencia, no cabe esperar que los salarios de la mujeres seán superiores; por tanto el contraste debe ser de una sóla cola. Así pues:

1. $H_0 : \mu_X - \mu_Y = 0; \quad H_1; \mu_X - \mu_Y > 0;$
2. estadístico: $\frac{(\overline{x}-\overline{y})-(\mu_X-\mu_Y)}{\sqrt{\frac{\sigma_X^2}{n_x}+\frac{\sigma_Y^2}{n_y}}} \sim N(0,1); \quad RC = \left\{ \mathbf{x} \;\middle|\; \frac{(\overline{x}-\overline{y})}{\sqrt{\frac{4}{n_x}+\frac{9}{n_y}}} > z_{1-\alpha} \right\}$

Fin **ejercicio** 8

Ejercicio 9. La proporción de alumnos aprobados en junio fue de 150 de los 350 presentados por la mañana y de 75 de los 180 presentados por la tarde. Queremos hacer un contraste no paramétrico chi-cuadrado de independencia entre el turno y aprobar el examen. ¿Cuáles son las frecuencias absolutas teóricas para llevar a cabo este contraste?

(NO resuelva el contraste, limitese a contestar lo que se le pide)

Solución:

	Aprobado	suspenso
Mañana	148.58	201.41
Tarde	76.41	103.58

Fin **ejercicio** 9

Ejercicio 10. La funcion generatriz de momentos de una chi-cuadrado de r grados de libertad es: $M(t) = (1-2t)^{-r/2}$. Use su función generatriz de momentos para calcular la varianza de una chi-cuadrado de r grados de libertad.

Solución: $\mathrm{Var}\,(X) = \mathrm{E}\left(X^2\right) - \left[\mathrm{E}\,(X)\right]^2$

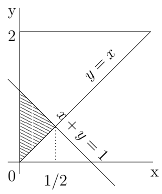
$M'(t) = \frac{-r}{2}(1-2t)^{\frac{-r}{2}-1}(-2) = r(1-2t)^{\frac{-r}{2}-1}$ y $\mathrm{E}\,(X) = M'(0) = r$

$M''(t) = r(\frac{-r}{2} - 1(1-2t)^{\frac{-r}{2}-2}(-2) = (r^2+2r)(1-2t)^{\frac{-r}{2}-2}$ y $\mathrm{E}\left(X^2\right) = M''(0) = r^2 + 2r$.

Por tanto, $\mathrm{Var}\,(X) = r^2 + 2r - r^2 = 2r$

Fin **ejercicio** 10

Ejercicio 11. Sea la función de densidad continua, $f_{XY}(x,y) = \begin{cases} k & \text{si } 0 < x < y < 2 \\ 0 & \text{resto} \end{cases}$. Indique cómo calcularía la siguiente probabilidad: $P\left(X + Y \leq 1\right)$.



Solución: $P\left(X + Y \leq 1\right) = \int_0^{1/2} \int_x^{1-x} k\; dy\; dx$

Fin **ejercicio** 11

Formulas de posible utilidad

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

Función generatriz de momentos conjunta del vector aleatorio $\boldsymbol{X} = (X_1, \ldots, X_n)$ evaluada en el vector $\boldsymbol{t} = (t_1, \ldots, t_n)$: $M_{\boldsymbol{X}}(\boldsymbol{t}) = M_{X_1,\ldots,X_n}(t_1, \ldots, t_n) = \operatorname{E}\left(e^{t_1 X_1 + \cdots + t_n X_n}\right)$

Aproximación lineal a la esperanza condicional

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

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

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

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

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

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

$$\frac{\bar{x} - \mu}{\sqrt{\mathfrak{s}^2/n}} \sim t_{n-1}; \qquad \frac{(n-1)\mathfrak{s}^2}{\sigma^2} \sim \chi^2_{n-1}; \qquad \frac{\mathfrak{s}_1^2/\sigma_1^2}{\mathfrak{s}_2^2/\sigma_2^2} \sim F_{n_1-1, n_2-1}$$

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

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

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

$$\frac{(\widehat{p_1} - \widehat{p_2}) - (p_1 - p_2)}{\sqrt{\frac{(n+m)\widehat{p_T}(1 - \widehat{p_T})}{n \cdot m}}} \underset{H_0}{\overset{a}{\sim}} N\left(0\,,\, 1\right); \quad \text{donde } \widehat{p_T} = \frac{n\widehat{p_1} + m\widehat{p_2}}{n + m}; \ H_0:\ p_1 = p_2$$

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: $D_n = \sup\left|F_n(x) - F(x)\right|$, 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: $D_n = \sup\left|F_1(x) - F_2(x)\right|$. 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 α igual a 10 %, 5 % y 1 % respectivamente.

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	**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

Puntos porcentuales de la **fun. distribución** t de STUDENT con ν grados de libertad

ν	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
6	0.265	0.453	0.718	0.906	1.273	1.440	1.943	2.447	3.143	3.707	5.208
7	0.263	0.449	0.711	0.896	1.254	1.415	1.895	2.365	2.998	3.499	4.785
8	0.262	0.447	0.706	0.889	1.240	1.397	1.860	2.306	2.896	3.355	4.501
9	0.261	0.445	0.703	0.883	1.230	1.383	1.833	2.262	2.821	3.250	4.297
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
11	0.260	0.443	0.697	0.876	1.214	1.363	1.796	2.201	2.718	3.106	4.025
12	0.259	0.442	0.695	0.873	1.209	1.356	1.782	2.179	2.681	3.055	3.930
13	0.259	0.441	0.694	0.870	1.204	1.350	1.771	2.160	2.650	3.012	3.852
14	0.258	0.440	0.692	0.868	1.200	1.345	1.761	2.145	2.624	2.977	3.787
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
16	0.258	0.439	0.690	0.865	1.194	1.337	1.746	2.120	2.583	2.921	3.686
17	0.257	0.438	0.689	0.863	1.191	1.333	1.740	2.110	2.567	2.898	3.646
18	0.257	0.438	0.688	0.862	1.189	1.330	1.734	2.101	2.552	2.878	3.610
19	0.257	0.438	0.688	0.861	1.187	1.328	1.729	2.093	2.539	2.861	3.579
20	0.257	0.437	0.687	0.860	1.185	1.325	1.725	2.086	2.528	2.845	3.552
21	0.257	0.437	0.686	0.859	1.183	1.323	1.721	2.080	2.518	2.831	3.527
22	0.256	0.437	0.686	0.858	1.182	1.321	1.717	2.074	2.508	2.819	3.505
23	0.256	0.436	0.685	0.858	1.180	1.319	1.714	2.069	2.500	2.807	3.485
24	0.256	0.436	0.685	0.857	1.179	1.318	1.711	2.064	2.492	2.797	3.467
∞	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090

Puntos porcentuales de la **fun. distribución** CHI-CUADRADO con ν grados de libertad

ν	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
6	6.211	6.867	7.841	8.558	9.992	10.645	12.592	14.449	16.812	18.548	22.458
7	7.283	7.992	9.037	9.803	11.326	12.017	14.067	16.013	18.475	20.278	24.322
8	8.351	9.107	10.219	11.030	12.636	13.362	15.507	17.535	20.090	21.955	26.125
9	9.414	10.215	11.389	12.242	13.926	14.684	16.919	19.023	21.666	23.589	27.877
10	10.473	11.317	12.549	13.442	15.198	15.987	18.307	20.483	23.209	25.188	29.588
11	11.530	12.414	13.701	14.631	16.457	17.275	19.675	21.920	24.725	26.757	31.264
12	12.584	13.506	14.845	15.812	17.703	18.549	21.026	23.337	26.217	28.300	32.910
13	13.636	14.595	15.984	16.985	18.939	19.812	22.362	24.736	27.688	29.819	34.528
14	14.685	15.680	17.117	18.151	20.166	21.064	23.685	26.119	29.141	31.319	36.123
15	15.733	16.761	18.245	19.311	21.384	22.307	24.996	27.488	30.578	32.801	37.697
16	16.780	17.840	19.369	20.465	22.595	23.542	26.296	28.845	32.000	34.267	39.252
17	17.824	18.917	20.489	21.615	23.799	24.769	27.587	30.191	33.409	35.718	40.790
18	18.868	19.991	21.605	22.760	24.997	25.989	28.869	31.526	34.805	37.156	42.312
19	19.910	21.063	22.718	23.900	26.189	27.204	30.144	32.852	36.191	38.582	43.820
20	20.951	22.133	23.828	25.038	27.376	28.412	31.410	34.170	37.566	39.997	45.315
21	21.991	23.201	24.935	26.171	28.559	29.615	32.671	35.479	38.932	41.401	46.797
22	23.031	24.268	26.039	27.301	29.737	30.813	33.924	36.781	40.289	42.796	48.268
23	24.069	25.333	27.141	28.429	30.911	32.007	35.172	38.076	41.638	44.181	49.728
24	25.106	26.397	28.241	29.553	32.081	33.196	36.415	39.364	42.980	45.559	51.179