

## Tema 2

# Contrastes de hipótesis

- 2.1 Introducción.
- 2.2 Contrastes paramétricos: medias, varianzas y proporciones.
- 2.3 Test de la  $\chi^2$
- 2.4 Otros contrastes no paramétricos.

### 2.1. Introducción

Hay muchos problemas (como fueron los ejemplos 1.3. o 1.4.) en los que más que estimar el valor de un parámetro, debemos decidir si un enunciado referente a un parámetro es cierto o falso, o sea, debemos probar una hipótesis sobre un (o más) parámetro(s), o bien, comprobar si una teoría sobre la población es verosímil dados los datos muestrales.



**Contraste de hipótesis:** es un método numérico para comprobar una teoría o hipótesis sobre una población.

En todo contraste de hipótesis nos encontramos con una hipótesis nula ( $H_0$ ) y una hipótesis alternativa ( $H_1$  o  $H_A$ ). Cuando la meta de un experimento sea establecer una afirmación, ésta se convertirá en la hipótesis alternativa y su negación será la hipótesis nula. La hipótesis nula se supone cierta hasta que los datos indiquen lo contrario, por tanto la que se ha de demostrar que es cierta es la hipótesis alternativa,  $H_1$ . Podríamos plantear un símil con los juicios. En principio, se parte de la hipótesis nula de que un acusado es inocente hasta que se demuestre su culpabilidad. El fiscal es el que debe demostrar la culpabilidad y no será hallado culpable a menos que la hipótesis nula de su inocencia sea claramente rechazada. Con lo cual, si es hallado no culpable, no implica que el acusado haya demostrado su inocencia, sino que sólo implica que no se ha demostrado su culpabilidad.

**Ejemplo 2.1.:** En un entorno interactivo, definimos el *tiempo de respuesta* como el tiempo transcurrido entre que un usuario pulsa la tecla *INTRO* y el comienzo de la respuesta. Desea reducirse el tiempo de respuesta en un sistema concreto, así que supongamos que para reducirlo, se realizan diversas reconfiguraciones. Actualmente, sin esta reconfiguración, el tiempo de respuesta medio es de  $\mu_0 = 3$  segundos. Se realiza una prueba con la nueva actualización y durante

una semana recogemos aleatoriamente 50 tiempos de respuesta y obtenemos una media  $\bar{x} = 2.7$  segundos y una desviación típica  $s = 0.68$ . Denotemos por  $\mu$  el tiempo de respuesta medio con la actualización, así que  $\bar{x}$  es un valor estimado de  $\mu$ . ¿Existe diferencia significativa entre la actualización y sin ella en cuanto al tiempo de respuesta? Podríamos plantear el siguiente contraste:


$$\begin{cases} H_0 : \mu \geq \mu_0 = 3 \text{ (o simplemente } \mu = \mu_0 = 3, \text{ no hay diferencia entre ambos)} \\ H_1 : \mu < \mu_0 = 3 \end{cases}$$


Planteamos dicho contraste ya que la hipótesis que queremos demostrar es que la actualización es mejor que la configuración existente. Implícitamente se está suponiendo que la configuración existente es buena y que los gastos de actualización son lo suficiente importantes como para necesitar una justificación.

🎵 Nota: Si hubiésemos partido de que la configuración existente es mala y los gastos del cambio no fueran importantes, entonces habría que demostrar que la existente es superior y plantearíamos:




$$\begin{cases} H_0 : \mu \leq \mu_0 = 3 \\ H_1 : \mu > \mu_0 = 3 \end{cases}$$

Aunque  $H_0$  sea verdadera y no exista diferencia entre los procesos, para una muestra es difícilísimo que obtengamos exactamente  $\bar{x} = \mu_0$ , y por tanto, podemos cometer errores que serán:

 **Error de tipo I:** se produce cuando  $H_0$  es cierta pero se rechaza. La probabilidad de cometerlo se designa por  $\alpha$ .

 **Error de tipo II:** se produce cuando  $H_0$  es falsa pero se "acepta". La probabilidad de incurrir en él se designa por  $\beta$ .


	"Aceptar" $H_0$	Rechazar $H_0$
$H_0$ cierta	Decisión correcta	Error I
$H_0$ falsa	Error II	Decisión correcta

Se llama   **nivel de significación**  $\alpha$  (recuerda el tema 1 ) de un contraste de hipótesis a la probabilidad de cometer un error de tipo I.

**Ejemplo 1.3.:** Recordemos este ejemplo presentado en el tema anterior: a la hora de diseñar un sistema de servicios, resulta adecuado contar con diversas hipótesis de trabajo respecto al número de unidades que se deben servir en un cierto momento. Una estimación demasiado baja conducirá a un servicio inadecuado, mientras que una estimación demasiado alta conduciría a un desperdicio de recursos. Supongamos que a una compañía de distribución de recursos informáticos le interesa el número de usuarios interactivos durante una hora dada y desea comprobarse si el número medio de usuarios se desvía de 110<sup>1</sup>. Se considera una muestra de 25 horas observadas

<sup>1</sup>nótese que en realidad la variable es discreta pero que las observaciones se aproximan a una normal

y se obtiene una media de 112.2 y una desviación típica de 8.4.

 **Ejemplo 2.2.:** Se ha probado la duración de 9 baterías de cierta marca, para ordenadores portátiles. Los tiempos, que consideraremos normal, obtenidos son (en horas): 11.7, 12.2, 10.9, 11.4, 11.3, 12, 11.1, 10.7, 11.6. Según el fabricante de las baterías, su duración es de 12.1 horas. Se quiere saber con el 99% de certeza si las baterías de esta marca tienen una duración  $\mu$  significativamente inferior a  $\mu_0 = 12.1$ .

### Pasos a seguir para resolver los problemas de pruebas de hipótesis:

1) Formular una hipótesis nula y alternativa apropiada:

a) es un contraste bilateral mientras que b) y c) son contrastes unilaterales.

$$\begin{array}{lll}
 \text{a)} & H_0 : \theta = \theta_0 & H_0 : \theta \geq \theta_0 \ (\theta = \theta_0) & H_0 : \theta \leq \theta_0 \ (\theta = \theta_0) \\
 & H_1 : \theta \neq \theta_0 & H_1 : \theta < \theta_0 & H_1 : \theta > \theta_0 \\
 \text{b)} & & & \text{c)}
 \end{array}$$

### Ejemplo 1.3.:

$$\begin{cases} H_0 : \mu = 110 \\ H_1 : \mu \neq 110 \end{cases}$$

### Ejemplo 2.2.:

$$\begin{cases} H_0 : \mu \\ H_1 : \mu \end{cases}$$

2) Especificar la probabilidad de error de tipo I e identificar los datos con los que contamos.

### Ejemplo 2.1.:

$\mu_0 = 3$ ,  $s = 0.68$ ,  $\bar{x} = 2.7$ ,  $N = 50$ , tomaremos  $\alpha = 0.05$ .

### Ejemplo 1.3.:

$\mu_0 = 110$ ,  $s = 8.4$ ,  $\bar{x} = 112.2$ ,  $N = 25$ , tomaremos  $\alpha = 0.05$ .

### Ejemplo 2.2.:

3) Elegir un estadístico de contraste adecuado, así como su distribución (véase la página 52):

**Ejemplo 2.1.:**

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{50}} \sim N(0, 1)$$

**Ejemplo 1.3.:**

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{25}} \sim t_{25-1}$$

**Ejemplo 2.2.:**

4) Cálculo del valor observado del estadístico de contraste según los datos observados:

**Ejemplo 2.1.:**

$$z = \frac{2,7 - 3}{0,68/\sqrt{50}} = -3,12$$

**Ejemplo 1.3.:**

$$t = \frac{112,2 - 110}{8,4/\sqrt{25}} = 1,31$$

**Ejemplo 2.2.:**

5) Determinación de la región crítica y el/los valor/es crítico/s:



**Región crítica:** son los valores del estadístico de contraste que nos conducen a rechazar la hipótesis nula.



**Región de aceptación:** son los valores del estadístico de contraste que nos llevan a "aceptar" = *no rechazar* la hipótesis nula.



**Valor crítico:** valor/es que separan la región crítica de la de aceptación.

**Ejemplo 2.1.:**

región crítica:  $(-\infty, -z_\alpha)$

Como  $\alpha = 0.05$  y el estadístico sigue una  $N(0,1)$ ,  $z_\alpha = z_{0,05} = 1.64$  y la región crítica quedaría:  $(-\infty, -1.64)$ .

**Ejemplo 1.3.:**

región crítica:  $(-\infty, -t_{\alpha/2}) \cup (t_{\alpha/2}, +\infty)$

Como  $\alpha = 0.05$  y el estadístico sigue una  $t$  de Student con  $25 - 1 = 24$  grados de libertad,  $t_{\alpha/2} = t_{0,025}$  es: 2.064 y la región crítica queda:  $(-\infty, -2.064) \cup (2.064, +\infty)$



**Ejemplo 2.2.:**

6) Decisión: rechazar  $H_0$  si el valor observado pertenece a la región crítica, sino no rechazar  $H_0$ .

**Ejemplo 2.1.:**

$-3.12 \in (-\infty, -1.64) \rightarrow$  Rechazamos  $H_0$ . La actualización es mejor que la configuración existente.

**Ejemplo 1.3.:**

1.31 no pertenece a  $(-\infty, -2.064) \cup (2.064, +\infty) \rightarrow$  No rechazamos  $H_0$ , no tenemos pruebas para afirmar que el número de usuarios es distinto de 110, la diferencia ( $\bar{x} = 112.2$ ) puede atribuirse al azar.



**Ejemplo 2.2.:**

## 2.2. Contrastes paramétricos: medias, varianzas y proporciones.



### A) Contraste de hipótesis para la media $\mu$ , con $N$ grande ( $N \geq 30$ ):

$$Z \approx \frac{\bar{X} - \mu_0}{S/\sqrt{N}} \sim N(0, 1) \quad \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : 3 \text{ casos posibles} \rightarrow \end{cases}$$

$H_1$	Región crítica
$\mu < \mu_0$	$(-\infty, -z_\alpha)$
$\mu \neq \mu_0$	$(-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, \infty)$
$\mu > \mu_0$	$(z_\alpha, \infty)$

Ejemplo 2.1.: ya visto



### B) Contraste de hipótesis para la media $\mu$ en una población Normal con $\sigma^2$ desconocida:

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{N}} \sim t_{N-1} \quad \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : 3 \text{ casos posibles} \rightarrow \end{cases}$$

$H_1$	Región crítica
$\mu < \mu_0$	$(-\infty, -t_\alpha)$
$\mu \neq \mu_0$	$(-\infty, -t_{\alpha/2}) \cup (t_{\alpha/2}, \infty)$
$\mu > \mu_0$	$(t_\alpha, \infty)$

[🎵 Nota: la distribución  $t$  se denomina  $t$  de Student por el seudónimo empleado por Gosset para publicar sus trabajos. Gosset trabajaba para la Guinness Brewers (la cervecería) en Irlanda, y debido a que su patrón desaprobaba la publicación de investigaciones de un empleado, tuvo que publicar sus resultados bajo el seudónimo Student.]



Ejemplo 1.3.: ya visto

Ejemplo 2.2.: ya visto




**P-valor:** Además de calcular la región crítica puede calcularse el  $p$ -valor (asociado a nuestros datos). El  $p$ -valor es el menor valor de  $\alpha$  que nos conduciría a rechazar  $H_0$ . Un  $p$ -valor se determina como la probabilidad de que el estadístico de contraste pertenezca a la región crítica cuando el valor observado se considera valor crítico. Valores pequeños del  $p$ -valor (por ejemplo menor que 0.05) nos llevan a rechazar  $H_0$ . Si  $\alpha$  es menor que el  $p$ -valor, no

rechazamos  $H_0$ . En cambio, si  $\alpha$  es mayor que el  $p$ -valor, rechazamos  $H_0$ .


 **Observación:** Existe una relación entre los intervalos de confianza y los contrastes de hipótesis. Los intervalos de confianza (bilaterales) vistos en el tema anterior  (exceptuando el  $L$ ) nos dan la región de aceptación de contrastes bilaterales al  $100 \cdot (1 - \alpha) \%$  y por tanto,  $H_0$  no será rechazada al nivel  $\alpha$  si  $\theta_0$  pertenece al intervalo. O sea, intervalo de confianza  $(1 - \alpha) =$  conjunto de hipótesis aceptables a nivel  $\alpha$ . Por ejemplo, para el caso de la media de una población Normal, se acepta al nivel  $\alpha$  la hipótesis  $\mu = \mu_0$  cuando el intervalo de confianza  $1 - \alpha$  construido para  $\mu$  incluye a  $\mu_0$  y viceversa. Ejemplo: consideremos un proceso de fabricación que, en condiciones normales, produce elementos cuya vida se distribuye normalmente con media 5.100 horas. Se introducen ciertos cambios en el proceso que pueden afectar a la media pero no a la variabilidad. Para contrastar si estos cambios han producido efectos, tomamos una muestra de 4 elementos cuyas vidas resultan ser (en horas): 5010, 4750, 4826, 4953. ¿Hay evidencia de un efecto sobre la media?

$\bar{x} = 4884.75$ ,  $s = 118.258$ , intervalo de confianza para  $\mu$  al 95%:  $(\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{N}}) = (4884.75 \pm 3.182 \frac{118.258}{\sqrt{4}}) = (4696.6, 5072.9)$ , como  $5100 \notin (4696.6, 5072.9) \rightarrow$  sí que habrá afectado a la media

 **C) Contraste para la diferencia de medias  $\mu_1 - \mu_2$ , con  $\sigma_1^2$  y  $\sigma_2^2$  desconocidas, para muestras aleatorias independientes y tamaños muestrales grandes** ( $N_1 =$  tamaño muestral de la muestra de la población 1,  $N_2 =$  tamaño muestral de la muestra de la población 2):

$$Z \approx \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{s_1^2/N_1 + s_2^2/N_2}} \sim N(0, 1) \quad \begin{cases} H_0 : \mu_1 - \mu_2 = \Delta_0 \\ H_1 : 3 \text{ casos posibles} \rightarrow \end{cases}$$

$H_1$	Región crítica
$\mu_1 - \mu_2 < \Delta_0$	$(-\infty, -z_\alpha)$
$\mu_1 - \mu_2 \neq \Delta_0$	$(-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, \infty)$
$\mu_1 - \mu_2 > \Delta_0$	$(z_\alpha, \infty)$

 **Ejemplo 2.3.:** En el diseño de un sistema operativo, la planificación del uso de la CPU por los distintos procesos usa un algoritmo basado en colas. Sin embargo, en la práctica es posible que la disciplina de cola no sea tan simple como una FIFO (primero en entrar, primero en salir), ni que las distribuciones de llegada y servicio sean exponenciales o Erlang. Dados dos algoritmos de planificación, podría ser más conveniente compararlos en la práctica. Para hacer esta comparación, supongamos que se registran 40 tiempos (tiempos de espera y servicio) de 40 segmentos de programas seleccionados aleatoriamente para el método A y otros 40 para el método B. Se afirma que el tiempo medio para el método A es más de 0.4 segundos más rápido que el método B. Según los tiempos observados, se tiene una media de 3.64 y una desviación típica de 0.53 para el método A, mientras que para el método B tenemos una media de 4.25 y

una desviación típica de 0.44. ¿Confirma este estudio el argumento a nivel de significación  $\alpha = 0.05$ ?



**D) Contraste para la diferencia de medias  $\mu_1 - \mu_2$  de poblaciones normales independientes, con varianzas poblacionales desconocidas pero iguales ( $\sigma_1^2 = \sigma_2^2$ )** ( $N_1 =$  tamaño muestral de la muestra de la población 1,  $N_2 =$  tamaño muestral de la muestra de la población 2):

$$T = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{\frac{(N_1-1)s_1^2 + (N_2-1)s_2^2}{N_1+N_2-2}}} \sqrt{\frac{N_1 \cdot N_2}{N_1 + N_2}} \sim t_{N_1+N_2-2} \quad \begin{cases} H_0 : \mu_1 - \mu_2 = \Delta_0 \\ H_1 : 3 \text{ casos posibles} \rightarrow \end{cases}$$

$H_1$	Región crítica
$\mu_1 - \mu_2 < \Delta_0$	$(-\infty, -t_\alpha)$
$\mu_1 - \mu_2 \neq \Delta_0$	$(-\infty, -t_{\alpha/2}) \cup (t_{\alpha/2}, \infty)$
$\mu_1 - \mu_2 > \Delta_0$	$(t_\alpha, \infty)$

**Ejemplo 2.4.:** Muchos paquetes de software utilizan interfaces con el usuario controladas por menús a fin de mejorar la "amabilidad con el usuario". Una característica que puede incorporarse en la interfaz es una presentación de menús apilados: cada vez que se selecciona una opción de un menú, aparece un submenú superpuesto parcialmente al menú original, creando así una serie de menús "apilados". En el *Special Interest Group on Computer Human Interaction Bulletin* (julio 1993) se recoge un estudio para determinar los efectos de la presencia o ausencia de una estructura de menús apilados sobre el tiempo de búsqueda. Se colocó a 22 sujetos aleatoriamente en uno de dos grupos, y a cada uno se le pidió buscar una opción en particular en un paquete de software controlado por menús. En el grupo experimental ( $N_1 = 11$ ) se utilizó el formato de menús apilados, en el grupo de control ( $N_2 = 11$ ) se exhibió únicamente el menú actual. Se midieron los tiempos de búsqueda, obteniéndose para el grupo experimental una media de 11.02 segundos y una desviación de 0.13 y para el grupo de control una media de 11.07 con una desviación típica de 0.11. ¿Hay alguna diferencia entre los tiempos medios de búsqueda? Usa  $\alpha = 0.05$ , supón normalidad e igualdad de varianzas, para resolver el contraste adecuado.




**E) Contraste para la diferencia de medias  $\mu_1 - \mu_2$  de poblaciones normales**





**independientes, con varianzas poblacionales  $\sigma_1^2, \sigma_2^2$  desconocidas y desiguales** ( $N_1 =$  tamaño muestral de la muestra de la población 1,  $N_2 =$  tamaño muestral de la muestra de la población 2):

$$T = \frac{\bar{X}_1 - \bar{X}_2 - \Delta_0}{\sqrt{s_1^2/N_1 + s_2^2/N_2}} \sim t_{g.l.} \quad g.l. = \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\frac{(s_1^2/N_1)^2}{N_1-1} + \frac{(s_2^2/N_2)^2}{N_2-1}} \quad \begin{cases} H_0 : \mu_1 - \mu_2 = \Delta_0 \\ H_1 : 3 \text{ casos posibles} \rightarrow \end{cases}$$


$H_1$	Región crítica
$\mu_1 - \mu_2 < \Delta_0$	$(-\infty, -t_\alpha)$
$\mu_1 - \mu_2 \neq \Delta_0$	$(-\infty, -t_{\alpha/2}) \cup (t_{\alpha/2}, \infty)$
$\mu_1 - \mu_2 > \Delta_0$	$(t_\alpha, \infty)$

 **Ejemplo 2.5.:** Se están probando los componentes electrónicos de dos proveedores, midiéndose el tiempo hasta el fallo (que consideraremos normal). Para el primero de ellos, se obtuvo para  $N_1 = 10$ , una media de 2545 horas y una desviación típica de 437 horas. Para el segundo proveedor, se obtuvo para  $N_2 = 15$ , una media de 2800 horas y una desviación típica de 231 horas. ¿Presentan los datos suficiente evidencia para asegurar que hay diferencia significativa entre los tiempos medios hasta el fallo de los componentes de los dos proveedores? Usa  $\alpha = 0.05$  para resolver el contraste adecuado, ten en cuenta que consideraremos varianzas poblacionales distintas.

  **F) Contraste para la diferencia de medias  $\mu_1 - \mu_2$  para muestras apareadas, cuya diferencia es normal:**  $\bar{D}$  y  $S_D$  son la media y desviación típica de las diferencias

$$T = \frac{\bar{D} - \Delta_0}{S_D/\sqrt{N}} \sim t_{N-1} \quad \begin{cases} H_0 : \mu_D = \Delta_0 \\ H_1 : 3 \text{ casos posibles} \rightarrow \end{cases}$$

$H_1$	Región crítica
$\mu_D < \Delta_0$	$(-\infty, -t_\alpha)$
$\mu_D \neq \Delta_0$	$(-\infty, -t_{\alpha/2}) \cup (t_{\alpha/2}, \infty)$
$\mu_D > \Delta_0$	$(t_\alpha, \infty)$

 **Ejemplo 2.6.:** En un sistema de reconocimiento automático diversos algoritmos se encargan de procesar la imagen de entrada, extraer las características de interés y tomar una decisión (identificación). El éxito de cualquier procedimiento de reconocimiento depende de las condiciones externas, como pueden ser la luz, oclusiones del objetivo o la textura del fondo. Consideremos que se quiere comparar dos sistemas de reconocimiento, evaluándolos bajo 10 conjuntos

de condiciones distintas. Para cada escenario, cada sistema tratará 50 escenas que contienen objetos diseñados para reconocerlos y se guardará el número de identificaciones correctas. Aunque los datos son discretos, se asumirá que la distribución de la diferencia es aproximadamente normal. Los datos obtenidos fueron:

Escenario	SISTEMA 1	SISTEMA 2	DIFERENCIA (Sistema 1 - Sistema 2)
1	42	39	3
2	47	38	9
3	15	16	-1
4	42	36	6
5	38	36	2
6	25	24	1
7	46	39	7
8	35	30	5
9	12	12	0
10	40	35	5

↓  
 $\bar{d} = 3.7$   
 $s_d = 3.234$

Se quiere probar que el sistema 1 (¡el que hemos hecho nosotros!) es mejor que el sistema 2. Usa  $\alpha = 0.05$ .



### G) Contraste para $\sigma^2$ en una población normal:

$$\chi_0^2 = \frac{(N-1)S^2}{\sigma_0^2} \sim \chi_{N-1}^2 \quad \begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : 3 \text{ casos posibles} \rightarrow \end{cases}$$

$H_1$	Región crítica
$\sigma^2 < \sigma_0^2$	$(0, \chi_{1-\alpha}^2)$
$\sigma^2 \neq \sigma_0^2$	$(0, \chi_{1-\alpha/2}^2) \cup (\chi_{\alpha/2}^2, \infty)$
$\sigma^2 > \sigma_0^2$	$(\chi_{\alpha}^2, \infty)$

**Ejemplo 2.7.:** La variable aleatoria  $X$  cuenta el número de bits en mensajes codificados procedentes de cierta fuente. Asumiendo que  $X$  es aproximadamente normal, desea comprobarse la hipótesis nula  $H_0: \sigma^2 = 160.000$  al nivel  $\alpha = 0.05$ . Construye la región crítica de dos colas apropiada y toma una decisión en base a los siguientes datos: 4532, 4606, 3511, 4201, 3392, 4639, 4021, 4722, 3470, 3100, 4212, 4165.



### H) Contraste para el cociente $\sigma_1^2/\sigma_2^2$ de varianzas de dos poblaciones normales independientes:

$$F = \frac{S_1^2}{S_2^2} \sim F_{(N_1-1, N_2-1)} \quad \begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 \\ H_1 : 3 \text{ casos posibles} \rightarrow \end{cases}$$

$H_1$	Región crítica
$\sigma_1^2 < \sigma_2^2$	$(0, F_{1-\alpha}) = (0, \frac{1}{F_{\alpha}^{(N_2-1, N_1-1)}})$
$\sigma_1^2 \neq \sigma_2^2$	$(0, F_{1-\alpha/2}) \cup (F_{\alpha/2}, \infty)$
$\sigma_1^2 > \sigma_2^2$	$(F_{\alpha}, \infty)$

**Ejemplo 2.8.:** Los diseñadores de un sistema operativo han propuesto dos algoritmos de colas bastante complicados para la planificación del uso de la CPU. En lugar de compararlos usando una aproximación analítica mediante un modelo matemático, los diseñadores deciden compararlos colocando máquinas test en dos escenarios considerados uniformes en términos de la proporción y duración de trabajos que llegan. Los siguientes son tiempos de espera (en segundos) para trabajos seleccionados aleatoriamente destinados a las dos máquinas test:

CPU A: 4.62 2.51 0.12 0.67 2.3 6.12 3.01 5.55 3 2.42 2.61 0.52

CPU B: 2.5 2.01 3.21 1.89 1.76 2.06 2.53 2.2 1.9 2.45 2.1 1.6

¿Puede concluirse al nivel  $\alpha = 0.01$  que la varianza resultante de la disciplina de cola de la CPU-A es mayor que para la CPU-B?



### I) Contraste para una proporción $p$ (de una Binomial) cuando $N$ es grande y la proporción no es cercana a cero ni a uno:

$\hat{p} = X/N$  ( $X$  = número de éxitos en las  $N$  pruebas),  $q_0 = 1 - p_0$

$$Z \approx \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / N}} \sim N(0, 1) \quad \begin{cases} H_0 : p = p_0 \\ H_1 : 3 \text{ casos posibles} \rightarrow \end{cases}$$

$H_1$	Región crítica
$p < p_0$	$(-\infty, -z_{\alpha})$
$p \neq p_0$	$(-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, \infty)$
$p > p_0$	$(z_{\alpha}, \infty)$

**Ejemplo 2.9.:** Para comprobar si un circuito está libre de fallos, de una secuencia de 100 entradas observamos 37 unos a la salida (63 ceros). Si el circuito está libre de fallos, se

esperan 50 unos. A un nivel de significación de  $\alpha = 0.05$ , ¿podemos rechazar la hipótesis de que el circuito está libre de fallos?



### J) Contraste para la diferencia de dos proporciones, con $N_1$ y $N_2$ grandes

( $N_1 =$  tamaño muestral de la muestra de la población 1,  $N_2 =$  tamaño muestral de la muestra de la población 2):

$\hat{p}_1 = X_1/N_1$  ( $X_1 =$  número de éxitos en las  $N_1$  pruebas),  $\hat{p}_2 = X_2/N_2$  ( $X_2 =$  número de éxitos en las  $N_2$  pruebas),  $\hat{p} = (X_1 + X_2)/(N_1 + N_2)$

$$Z \approx \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/N_1 + 1/N_2)}} \sim N(0, 1) \quad \begin{cases} H_0 : p_1 = p_2 \\ H_1 : 3 \text{ casos posibles} \rightarrow \end{cases}$$

$H_1$	Región crítica
$p_1 < p_2$	$(-\infty, -z_\alpha)$
$p_1 \neq p_2$	$(-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, \infty)$
$p_1 > p_2$	$(z_\alpha, \infty)$



**Ejemplo 2.10.:** Se realiza una encuesta en dos zonas distintas de un país para conocer el grado de implantación de Internet en los hogares. En la zona norte se visitaron 200 domicilios seleccionados al azar y el 38 % se había dado de alta para navegar por Internet. Este número descendía al 33 % en la zona sur, donde se habían visitado 240 hogares. ¿Estaría justificado afirmar que el porcentaje de gente conectada desde sus domicilios a la red en la zona norte es superior a la zona sur? (usa  $\alpha = 0.05$ )



Al probar hipótesis, el analista selecciona directamente la probabilidad del error tipo I. Sin embargo, la probabilidad  $\beta$  del error tipo II depende de la elección del tamaño de la muestra. Cuando se hacen cálculos del tamaño de la muestra o del error tipo II, puede resultar útil usar curvas de operación característica, que representan  $\beta(\theta) = P(\text{"no rechazar } H_0" | \theta)$ . Estas curvas permiten determinar  $\beta$  cuando fijamos los restantes parámetros y recíprocamente  $N$  cuando los restantes vienen dados. En Montgomery puedes encontrarlas. En la prácticas con el Statgraphics, se trabajará este punto.

## 2.3. Test de la $\chi^2$

En el contexto de la inferencia paramétrica clásica existen tres hipótesis cruciales a la hora de poder llevar a cabo las técnicas y los análisis disponibles. Estas tres hipótesis son: la independencia de los datos, la homogeneidad (es decir, el hecho de que todos provengan de la misma distribución de probabilidad) y el ajuste a la distribución especificada.

Teniendo en cuenta lo anterior, la inferencia no paramétrica presenta dos objetivos bien diferenciados: construir procedimientos estadísticos que permitan concluir sobre el grado de cumplimiento de las tres hipótesis estructurales antes mencionadas y elaborar procedimientos alternativos (métodos de distribución libre) que puedan usarse cuando dichas hipótesis fallan para el modelo paramétrico postulado.

En este apartado, veremos el test chi-cuadrado,  $\chi^2$ , que puede adoptar dos formas que nos permitirán contrastar la bondad de ajuste y la independencia u homogeneidad en tablas de contingencia, como veremos a continuación.



### K) Prueba de la bondad de ajuste con la $\chi^2$ :

El objetivo de este contraste es aclarar si es cierta la hipótesis nula  $H_0$  de que una variable sigue una distribución teórica determinada. Por ello, se tratará de ver si las frecuencias de las observaciones se ajustan bien con la distribución.

Para esta prueba, las observaciones de la muestra aleatoria de tamaño  $N$  de la población cuya distribución de probabilidad es desconocida se ordenan en un histograma de frecuencia, con  $k$  intervalos de clase. Denotaremos por  $o_i$  la frecuencia observada en el intervalo de clase  $i$ . Calcularemos la frecuencia esperada,  $e_i$ , para el intervalo  $i$ -ésimo, a partir de la distribución de probabilidad hipotética.

El estadístico que usaremos es:


$$\chi_0^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}, \quad (2.1)$$

que sigue aproximadamente una distribución  $\chi^2$  con  $k - 1$  grados de libertad, siempre que la distribución especificada sea la correcta. Sin embargo, es usual aplicar el test aun en casos en los que la distribución de la variable no está totalmente especificada, sino que depende de algún parámetro que, en consecuencia, deberá ser estimado (por ejemplo, el caso en que se suponga que la variable en concreto sigue una distribución de Poisson y falta por especificar su parámetro  $\lambda$ ). En estos casos la distribución aproximada del test ha de ser corregida para incorporar esta información pasando a ser una  $\chi^2$  con  $k - r - 1$  grados de libertad, siendo  $r$  el número de parámetros estimados por máxima verosimilitud.

La hipótesis nula de que la distribución de la población es la distribución hipotética se rechazará si el valor calculado del estadístico anterior  $\chi_0^2$  es mayor que  $\chi_\alpha^2$ , o sea, la región crítica (a nivel  $\alpha$ ) es:  $(\chi_\alpha^2, \infty)$ .


Una limitación bastante recomendable en la práctica, es la de no llevar a cabo el contraste cuando la frecuencia esperada de alguna clase sea menor que 5, para evitar problemas de mala aproximación de la distribución usada a la verdadera distribución. Entonces, en los casos en los que esta condición falle, podríamos agrupar varios valores adyacentes hasta que se cumpla la restricción.

Veamos primero un ejemplo (ejemplo 2.11.) en el que la distribución viene completamente determinada, y posteriormente otro (ejemplo 2.12) en el que la distribución depende de uno o más parámetros desconocidos.

 **Ejemplo 2.11.:** Una forma de medir el rendimiento de un ordenador es a través de los tiempos de ejecución de las diversas instrucciones que ejecuta el procesador. Puesto que existen una serie de instrucciones usuales, que se pueden considerar afines, una manera apropiada sería considerar una media ponderada (pesada):  $\sum_{i=1}^r p_i t_i$ , donde  $p_i$  sería la probabilidad de llamar la instrucción  $I_i$ , mientras que  $t_i$  sería el tiempo de ejecución de  $I_i$ . Un conjunto específico de instrucciones  $I_1, I_2, \dots, I_r$  junto con sus probabilidades de ocurrencia constituyen un *grupo de instrucciones* (*instruction mix*). La clasificación por tiempos de ejecución dependerá de lo cuán correcta o apropiada sea la división en grupos que se escoja. Una de las divisiones en grupos habitualmente usada es la *clasificación de Gibson* que aparece en la tabla siguiente:


Tipo de instrucción	Probabilidad del tipo de instrucción $p_i$	Frecuencia esperada $e_i$	Frecuencia observada $o_i$
Transfer to and from main memory	0.31		72
Indexing	0.18		30
Branching	0.17		32
Floating-point arithmetic	0.12		14
Fixed-point arithmetic	0.07		22
Shifting	0.04		10
Miscellaneous	0.11		20

En esta tabla aparecen también las frecuencias observadas correspondientes a una muestra de 200 instrucciones aleatoriamente seleccionadas de programas típicos en uso. Consideraremos el problema de determinar si la clasificación de Gibson es apropiada para este entorno particular (usar  $\alpha = 0.05$ ).

 **Ejemplo 2.12.:** Se propone (y veremos si es apropiada) hipotéticamente que el número

de defectos en tarjetas de circuitos impresos sigue una distribución de Poisson. Se obtiene una muestra aleatoria de  $N = 60$  tarjetas de circuitos impresos, y se observa el número de defectos, que se recogen en la tabla siguiente:

Número de defectos	Frecuencia observada
0	32
1	15
2	9
3	4
$\geq 4$	0

El contraste  $\chi^2$ , se usa no sólo para variables discretas o cualitativas, sino incluso con variables continuas. En este caso, dicha variable ha de ser agrupada en intervalos. Esto plantea el problema de la subjetividad en la elección de los mismos, así como la dependencia del resultado del test de los intervalos elegidos. El siguiente sería un ejemplo que muestra el procedimiento para variables continuas, aunque en este curso para variables continuas nosotros optaremos por usar otros tests, que veremos en el apartado siguiente  2.4.


**Ejemplo 2.13.:** Un ingeniero está probando una fuente de alimentación usada en un portátil. Utilizando  $\alpha = 0.05$ , él quiere determinar si una distribución normal describe adecuadamente el voltaje de salida. La solución de este ejemplo puedes encontrarlo en las fotocopias de problemas resueltos.



#### L) Pruebas con tablas de contingencia:

Una tabla de contingencia puede surgir en dos contextos diferentes:

a) Una muestra y dos variables cada una de ellas con  $r$  y  $c$  valores. En este caso podría interesarnos **contrastar la hipótesis de independencia de las dos variables.**

[ Nota: de manera intuitiva (existe una definición formal), dos variables serán independientes (fíjate que la propia palabra lo expresa) cuando el conocimiento sobre el valor de una de ellas (fijamos el valor de una de ellas), no altera la distribución de valores de la otra, o sea, no nos aportaría información acerca de esta variable].

**Ejemplo 2.14.:** Se efectúa un estudio sobre los fallos de unos componentes electrónicos. Existen cuatro tipos de fallos posibles y dos posiciones de montaje para el dispositivo. Los datos obtenidos fueron:

Posición de montaje	Tipo de fallo			
	A	B	C	D
1	12	36	8	5
2	5	16	6	12

Se quiere probar la hipótesis (con  $\alpha = 0.05$ ) de que el tipo de fallo es independiente de la posición de montaje.

b)  $r$  muestras independientes y una variable observada con  $c$  categorías. En este caso, podría interesarnos contrastar la hipótesis de que todas las distribuciones de donde se seleccionan las  $r$  muestras son semejantes, es decir, que **las  $r$  distribuciones son homogéneas**. Por ejemplo, cuando existen sólo dos categorías, tales como éxito y fracaso, defectuoso y no defectuoso, etc., entonces la prueba de homogeneidad es en realidad una prueba sobre la igualdad de  $r$  parámetros binomiales.

**Ejemplo 2.15.:** Consideremos 4 proveedores  $A, B, C$  y  $D$ ; 200 componentes de cada proveedor son aleatoriamente seleccionadas para examinarlas obteniéndose los siguientes resultados:

Calidad	Proveedor			
	A	B	C	D
Defectuoso	6	4	4	16
No defectuoso	194	196	196	184

¿La proporción de defectuosos es la misma para los 4 proveedores, o sea, son homogéneas las poblaciones? (considera  $\alpha = 0.01$ ).

Para ambos casos, el cálculo del estadístico de contraste es el mismo, aunque la forma de plantear  $H_0$  y de enunciar las conclusiones sean distintas.

Dada la tabla de contingencia siguiente:

$X \setminus Y$	$y_1$	...	$y_j$	...	$y_c$	Total
$x_1$	$o_{11}$	...	$o_{1j}$	...	$o_{1c}$	$T_{1.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$x_i$	$o_{i1}$	...	$o_{ij}$	...	$o_{ic}$	$T_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$x_r$	$o_{r1}$	...	$o_{rj}$	...	$o_{rc}$	$T_{r.}$
Total	$T_{.1}$	...	$T_{.j}$	...	$T_{.c}$	$T$

$T_{i.}$  es el total de observaciones de la fila  $i$ -ésima,  $T_{.j}$  es el total de observaciones de la columna  $j$ -ésima y  $T$  es el total de observaciones.



El estadístico que usaremos es:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}}, \quad (2.2)$$

siendo  $e_{ij} = T_i \cdot T_j / T$

Bajo  $H_0$ , sigue aproximadamente una distribución  $\chi^2$  con  $(r - 1) \cdot (c - 1)$  grados de libertad. La región crítica (a nivel  $\alpha$ ) es:  $(\chi_\alpha^2, \infty)$ . Para que la aproximación sea correcta, todas las  $e_{ij}$  deben ser al menos 5.

 **Ejemplo 2.14.:** Primero calculamos  $\chi^2 = 9.36$ .

$$\chi_{0,05}^2 = (\text{con } (2 - 1) \cdot (4 - 1) \text{ grados de libertad}) =$$



 **Ejemplo 2.15.:**

$$\chi^2 = 13.71$$

$$\chi_{0,01}^2 = (\text{con } (2 - 1) \cdot (4 - 1) \text{ grados de libertad}) =$$


## 2.4. Otros contrastes no paramétricos.


En este apartado veremos algunos procedimientos diseñados para el contraste de ajuste a distribuciones continuas, aunque los cálculos los realizaremos en prácticas con el Statgraphics

 Existen muchos otros contrastes no paramétricos con otros objetivos.  En Peña y Cao podéis encontrar una introducción a la inferencia no paramétrica.


El primero de los contrastes considerados es el de Kolmogorov-Smirnov, que se basa en la diferencia máxima entre la función de distribución empírica y teórica. Cuando no se especifican los parámetros, sino que éstos han de estimarse, se debe corregir la distribución del estadístico.

Debido a la gran importancia de la distribución Normal, existen diversos contrastes específicos para estudiar la bondad de ajuste a esta distribución. Algunos de ellos son:

- Contraste de Shapiro-Wilks: que se basa en el ajuste de la muestra a una recta al dibujarla en papel probabilístico normal. En prácticas , se verá cómo obtener este gráfico (*normal probability plot*).
- Contraste de asimetría (recordad que para una Normal, el coeficiente de asimetría vale cero).
- Contraste de curtosis o apuntamiento

 Fíjate que estos tres últimos contrastes son contrastes específicos de normalidad.]

A continuación se muestra un ejemplo donde se contrasta si unos datos provienen de una distribución normal (sin especificar a priori sus parámetros).

 **Ejemplo 2.16.:** Las cuotas de disco ocupado (en Mbytes) para distintos usuarios de una estación de trabajo son: 35, 45, 47, 50, 31, 30, 25, 33, 35, 40, 45, 47, 49, 42, 40, 50, 46, 55, 42, 46. Usemos los contrastes anteriores para estudiar su posible normalidad.

/\*\*\*\* Contraste de Kolmogorov-Smirnov \*\*\*\*\*/

EDF Statistic	Value	Modified Form	P-Value
Kolmogorov-Smirnov D	0,164318	0,764441	>=0.10*

/\*\*\*\* Contrastes de normalidad \*\*\*\*\*/


Tests for Normality for Disco

Computed Chi-Square goodness-of-fit statistic = 9,9 P-Value = 0,44931

Shapiro-Wilks W statistic = 0,9564 P-Value = 0,482919

Z score for skewness = 0,711306 P-Value = 0,476892

Z score for kurtosis = -0,366703 P-Value = 0,713837

Tras aplicar contrastes de normalidad como los anteriores, es posible que no se pueda aceptar que la distribución poblacional sea normal. En ese caso, o bien se utiliza otro modelo paramétrico que se ajuste a los datos o bien se trata de aplicar alguna transformación sobre la variable para tratar de conseguir que los nuevos datos se ajusten a una normal, como son las transformaciones de Box-Cox .

# Problemas del tema 2

En reprografía, podrás encontrar más problemas resueltos.

1. Una compañía sostiene que sus focos son superiores a los de su principal competidor. Se realizó un estudio: una muestra de 40 focos tuvo un ciclo medio de vida de 647 horas con una desviación típica de 27 horas, mientras que una muestra de 40 focos de la compañía rival dio un ciclo medio de vida de 638 horas con una desviación típica de 31 horas, ¿confirma este estudio el argumento a nivel de significación de 0.05?

*(Sol. :  $z = 1.38$  no pertenece a  $(1.64, \infty)$ , así que no rechazamos  $H_0$ , la diferencia entre ambas medias no es significativa).*

2. Al estudiar una muestra de 1000 personas clasificadas en fumadores y no fumadores, con o sin problemas en las vías respiratorias, se obtuvo:

	$E = \text{"Enfermos"}$	$E^c = \text{"No Enfermos"}$
$F = \text{"Fumadores"}$	183	289
$F^c = \text{"No Fumadores"}$	56	472

Determina si las variables son independientes ( $\alpha = 0.05$ ).

*(Sol. :  $\chi^2 = 108.7$ , rechazamos  $H_0$ , no son independientes)*

3. El proceso de alisado que se usa para pulir ciertos discos de silicio a fin de que su grosor sea el adecuado, sólo es aceptable si  $\sigma$ , la desviación típica de la población del grosor de cubos cortados de los discos (supuesta normal), es como mucho de 0.5 milésimas de milímetro. Utiliza el nivel de significación de 0.05 para probar la hipótesis nula de  $\sigma = 0.5$  frente a la alternativa  $\sigma > 0.5$ , si el grosor de 15 cubos cortados de esos discos tiene una desviación típica de 0.64 milésimas de milímetro.

*(Sol. :  $22.94$  no pertenece a  $(23.68, \infty)$ , no hay evidencias suficientes para rechazar  $H_0$  y a pesar que  $s = 0.64$  no podemos concluir que el proceso de alisado es insatisfactorio.)*

4. Un cierto algoritmo genera números pseudoaleatorios. "Pseudoaleatorio" significa que los dígitos generados deberían estar distribuidos (aproximadamente) uniformemente entre 0 y 9, de forma que la probabilidad de obtener un dígito dado fuera 0.1 Tras observar aleatoriamente 100 dígitos, se obtuvo:

Dígito:	0	1	2	3	4	5	6	7	8	9
Frecuencia	12	13	9	12	12	10	9	12	8	3

Realiza un contraste  $\chi^2$  para probar la "aleatoriedad" usando  $\alpha = 0.05$ .

(Sol. :  $\chi^2 = 8$  no pertenece a  $(16.92, \infty)$ , por ello, no rechazamos  $H_0$ , no hay razones para dudar de la aleatoriedad.)

5. Un artículo publicado en la revista *Materials Engineering* (1989, Vol. II, No. 4, págs. 275-281) describe los resultados de pruebas de resistencia a la adhesión de 22 especímenes de aleación U-700. La carga para la que cada espécimen falla es la siguiente (en MPa):

19.8 18.5 17.6 16.7 15.8 15.4 14.1 13.6 11.9 11.4 11.4 8.8 7.5 15.4  
15.4 19.5 14.9 12.7 11.9 11.4 10.1 7.9

Suponiendo que la carga donde se presenta el fallo es Normal y usando  $\alpha = 0.05$ , construye la prueba de hipótesis adecuada para responder a la siguiente pregunta: ¿sugieren los datos que la carga promedio de fallo es mayor que 10 MPa?

(Sol. :  $t=4.90 > 1.721$ , se rechaza  $H_0$ , la carga de fallo promedio es mayor que 10 MPa.)

6. Debido a interferencias, una proporción de señales de satélites son demasiado incomprensibles para interpretarse. Se prueban 3 filtros, A, B y C para ver si poseen distintos grados de efectividad. Se eligen 50 mensajes aleatoriamente para cada filtro, recogiendo el número de interpretaciones con éxito en la siguiente tabla:

Filtro	A	B	C
Éxitos	42	36	20

¿Puede rechazarse la hipótesis de homogeneidad al nivel de significación de  $\alpha = 0.1$ ?

(Sol. :  $\chi^2 = 22.842$ , rechazamos  $H_0$ , no son homogéneas)

7. En una muestra de 100 pistones de cerámica fabricados para un motor diesel experimental, 18 se rompieron. Prueba la hipótesis  $p = 0.2$  contra la alternativa  $p < 0.2$  al nivel 0.05.

(Sol. :  $-0.5$  no pertenece  $(-\infty, -1.64)$ , no podemos rechazar  $H_0$ .)

8. A continuación se presenta un resumen de la información obtenida de una muestra de 200 partes maquinadas.

	Profundidad de barrenado	
Condición de la arista	A = "Mayor de la necesaria"	N = "menor de la necesaria"
B = "Burda"	15	10
M = "Moderada"	25	20
S = "Suave"	60	70

Determina si las variables son independientes ( $\alpha = 0.05$ ).

(Sol. :  $\chi^2 = 2.32$ , no rechazamos  $H_0$ , no hay razones para pensar que no sean independientes)

9. Se pueden utilizar dos pruebas analíticas diferentes para determinar el nivel de impureza en aleaciones de acero. Tomamos 8 especímenes y los probamos con ambos procedimientos. Los resultados aparecen en la tabla siguiente:

	Prueba 1	Prueba 2
Especímen 1	1.3	1.5
Especímen 2	1.4	1.8
Especímen 3	1.6	1.7
Especímen 4	1.5	1.4
Especímen 5	1.8	2.1
Especímen 6	1.9	2.2
Especímen 7	1.5	1.8
Especímen 8	1.4	1.7

¿Hay evidencia suficiente para concluir que ambas pruebas producen el mismo nivel de impureza medio usando  $\alpha = 0.01$ ?

(Sol. :  $t = -4.025 \in (-\infty, -3.499) \cup (3.499, \infty)$ , rechazamos  $H_0$ , los resultados de ambas pruebas son diferentes.)

10. Dos proveedores, A y B, suministran a una compañía un determinado tipo de componentes electrónicos. Se seleccionan aleatoriamente 200 componentes de A y B, resultando que 3 de los componentes de A no funcionan adecuadamente, mientras que 6 de B son insatisfactorios. ¿Puede concluirse al nivel  $\alpha = 0.05$  que la proporción de componentes defectuosas de B es mayor que la de A?

(Sol. :  $z = -1.01$ , no pertenece a  $(-\infty, -1.64)$ , no rechazamos  $H_0$ , no podemos concluir al nivel  $\alpha = 0.05$  que la proporción de componentes defectuosas de B es mayor que la de A.)

11. En base a la siguiente tabla de frecuencias de las observaciones de la variable aleatoria  $X$ :

Valores	0	1	2	3	4
Frecuencia	4	21	10	13	2

¿La distribución Binomial con  $n = 6$  y  $p = 0.25$  es un modelo apropiado? Realiza el procedimiento de bondad de ajuste adecuado con  $\alpha = 0.05$ .

(Sol. :  $\chi^2 = 9.86$ , se rechaza  $H_0$ , no se ajustaría a una Binomial(6,0.25))



## Autoevaluación del tema 2

Las soluciones de algunos ejercicios están en reprografía, para los restantes, las soluciones se encuentran en estas mismas hojas.

**Ejercicio 1.:** Una ingeniera quiere mejorar la resistencia a los golpes de un envase que se produce en una planta y diseña un nuevo envase. Con el proceso existente, la resistencia del plástico del envase se distribuye normalmente con  $\mu_0 = 1250$ . La ingeniera produce en el laboratorio un lote de envases con el nuevo proceso y escoge una muestra aleatoria de tamaño 25. La media de esta muestra es  $\bar{x} = 1312$  y  $s = 150$ . Denotemos por  $\mu$  la resistencia media del envase producido por el nuevo proceso, así que  $\bar{x}$  es un valor estimado de  $\mu$ . ¿Existe diferencia significativa entre ambos envases?

Podríamos plantear el siguiente contraste:

$$\begin{cases} H_0 : \mu \leq \mu_0 = 1250 \text{ (o simplemente } \mu = \mu_0 = 1250, \text{ no hay diferencia entre ambos)} \\ H_1 : \mu > \mu_0 = 1250 \end{cases}$$

Planteamos dicho contraste ya que la hipótesis que queremos demostrar es que el nuevo es mejor que el existente. Implícitamente se está suponiendo que el proceso existente es bueno y que los gastos de cambio de proceso son lo suficiente importantes como para necesitar una justificación.

● Nota: Si hubiésemos partido de que el proceso existente es malo y los gastos del cambio no fueran importantes, entonces habría que demostrar que el existente es superior y plantearíamos:

$$\begin{cases} H_0 : \mu \geq \mu_0 = 1250 \\ H_1 : \mu < \mu_0 = 1250 \end{cases}$$

Paso 2)

$\mu_0 = 1250$ ,  $s = 150$ ,  $\bar{x} = 1312$ ,  $N = 25$ , tomaremos  $\alpha = 0.05$ .

Paso 3)

$$T = \frac{\bar{X} - \mu_0}{s/\sqrt{25}} \sim t_{25-1}$$

Paso 4)

$$t = \frac{1312 - 1250}{150/\sqrt{25}} = 2,06667$$

Paso 5)

región crítica:  $(t_\alpha, +\infty)$

Como  $\alpha = 0.05$  y el estadístico sigue una  $t$  de Student con  $25 - 1 = 24$  grados de libertad, el valor crítico es: 1.711 y la región crítica queda:  $(1.711, +\infty)$

Paso 6)

$2.06667 \in (1.711, +\infty) \rightarrow$  Rechazamos  $H_0$ . El proceso nuevo es mejor que el existente.

**Ejercicio 2:** Se quiere verificar si la cantidad de agua envasada en botellas de 0.5 l. es efectivamente dicha cantidad. Se toma una muestra de  $N = 60$  botellas de las que se obtienen  $\bar{x} = 0.51$  l. y  $s = 0.1$ .

$$\begin{cases} H_0 : \mu = 0.5 \\ H_1 : \mu \neq 0.5 \end{cases}$$

Paso 2)

$\mu_0 = 0.5$ ,  $s = 0.1$ ,  $\bar{x} = 0.51$ ,  $N = 60$ , tomaremos  $\alpha = 0.05$ .

Paso 3)

$$Z = \frac{\bar{X} - \mu_0}{s/\sqrt{60}} \sim N(0, 1)$$

Paso 4)

$$z = \frac{0,51 - 0,5}{0,1/\sqrt{60}} = 0,7746$$

Paso 5)

región crítica:  $(-\infty, -z_{\alpha/2}) \cup (z_{\alpha/2}, +\infty)$

Como  $\alpha = 0.05$  y el estadístico sigue una  $N(0,1)$ ,  $z_{\alpha/2} = z_{0,05/2} = z_{0,025} = 1.96$  y la región crítica quedaría:  $(-\infty, -1.96) \cup (1.96, +\infty)$

Paso 6)

0.77 no pertenece a  $(-\infty, -1.96) \cup (1.96, +\infty) \rightarrow$  No rechazamos  $H_0$ , no tenemos pruebas para afirmar que la cantidad es distinta de 0.5l., la diferencia ( $\bar{x} = 0.51$ ) puede atribuirse al azar.



**Ejercicio 3.:** Desea determinarse si la cantidad de ingresos está relacionada con la última marca comprada de cierto producto. Por ello, se analizó una muestra de 800 clientes, cuya información aparece en la tabla siguiente:

Ingresos	Marca		
	M1 = "Marca 1"	M2 = "Marca 2"	M3 = "Marca 3"
B = "Bajos"	50	125	75
M = "Medios"	125	65	190
A = "Altos"	75	45	50

Plantearemos el siguiente contraste:

$$\begin{cases} H_0 : \text{La marca comprada es independiente de los ingresos} \\ H_1 : \text{La marca comprada no es independiente de los ingresos} \end{cases}$$

Usaremos  $\alpha = 0.05$  y calcularemos el valor del estadístico:

$$\begin{aligned} \chi^2 = & \frac{(50 - 250 \cdot 250/800)^2}{250 \cdot 250/800} + \frac{(125 - 250 \cdot 235/800)^2}{250 \cdot 235/800} + \frac{(75 - 250 \cdot 315/800)^2}{250 \cdot 315/800} \\ & + \frac{(125 - 380 \cdot 250/800)^2}{380 \cdot 250/800} + \frac{(65 - 380 \cdot 235/800)^2}{380 \cdot 235/800} + \frac{(190 - 380 \cdot 315/800)^2}{380 \cdot 315/800} \\ & + \frac{(75 - 170 \cdot 250/800)^2}{170 \cdot 250/800} + \frac{(45 - 170 \cdot 235/800)^2}{170 \cdot 235/800} + \frac{(50 - 170 \cdot 315/800)^2}{170 \cdot 315/800} = 96,39 \end{aligned}$$

Para determinar la región crítica necesitamos conocer  $\chi_{0,05}^2$ , con  $(3 - 1) \cdot (3 - 1) = 4$  grados de libertad, que vale 9.49. Por tanto, la región crítica vendría dada por:  $(9.49, \infty)$  y como 96.39 pertenece a la región crítica, rechazaremos  $H_0$ , o sea, ambas variables no serían independientes, son dependientes.

**Ejercicio 4.:** Pueden utilizarse diversas técnicas para minimizar el ruido de fondo en un entorno de producción. Para comparar los niveles de ruido de fondo en dos localizaciones distintas A y B, se tomaron 15 medidas aleatorias (independientes) en cada lugar de la planta, que asumiremos normales. Los datos dieron para la posición A  $\bar{x} = 82.4$  y  $s_X^2 = 2.3$ , y para la posición B  $\bar{y} = 78.3$  y  $s_Y^2 = 4.1$ . Si las condiciones en los dos sitios difieren, especialmente con respecto al tipo de maquinaria en operación en cada lugar, podría ser arriesgado asumir varianzas iguales, por ello asumiremos varianzas distintas. Se desea mostrar que el nivel de ruido medio en la posición A es 3 decibelios mayor que el nivel de ruido medio en la posición B. Usa  $\alpha = 0.05$  con el contraste adecuado.

**Ejercicio 5.:** Se hace un estudio sobre el número de defectos de fabricación de una marca de ordenadores. En una muestra de 100 ordenadores, se observan los resultados siguientes:

Número de defectos por ordenador	0	1	2	3	4	5	6
Número de ordenadores	36	40	19	2	0	2	1

Decide al 95 % si se puede ajustar el número de defectos con una distribución Poisson.

**Ejercicio 6.:** Los tiempos entre averías consecutivas de una máquina en una cadena productiva se anotaron durante un período de 6 meses. Los datos (en horas) son los siguientes: 1, 10, 20, 30, 40, 52, 63, 70, 80, 90, 100, 102, 130, 140, 190, 210, 266, 310, 530, 590, 640, 1340. A continuación aparecen distintas salidas del Statgraphics, para realizar contrastes de bondad de ajuste a la distribución normal y exponencial, por ese orden. En base a estas salidas, contrasta las hipótesis anteriores.

/\*\*\*\* Normal \*\*\*\*/

EDF Statistic	Value	Modified Form	P-Value
Kolmogorov-Smirnov D	0,249665	1,21378	<0.01*

Tests for Normality for Averaia

Computed Chi-Square goodness-of-fit statistic = 39,4545 P-Value = 0,000021132

Shapiro-Wilks W statistic = 0,683604 P-Value = 0,00000361872

Z score for skewness = 2,85285 P-Value = 0,00433299

Z score for kurtosis = 3,46451 P-Value = 0,000531285

/\*\*\*\* Exponencial \*\*\*\*/

EDF Statistic	Value	Modified Form	P-Value
Kolmogorov-Smirnov D	0,184077	0,884909	>=0.10*