Contraste de hipótesis

1.- CONTRASTE DE HIPÓTESIS.

En la unidad anterior extraíamos muestras de una población para obtener información acerca de un parámetro que desconocíamos. Una vez conseguida la muestra se obtenía información sobre el parámetro desconocido a través de una estimación puntual o mediante intervalos de confianza.

A continuación veremos una tercera forma de conseguir información del parámetro poblacional, denominada *contraste de hipótesis.*

*Llamamos **contraste de hipótesis** la procedimiento estadístico mediante el cual tratamos de cuantificar las diferencias o discrepancias entre una hipótesis estadística y una realidad de la que poseemos una información muestral, estableciendo una regla de decisión para juzgar si las discrepancias son excesivamente grandes y, por tanto, para rechazar la hipótesis.*

Una situación en la podemos utilizar el contraste hipótesis es la de decidir, de entre dos poblaciones en las que estudiamos una característica común, aquella con la que obtengamos mejores resultados, basándonos en datos observados experimentalmente.

Ej. Según una información dada por cierta Universidad, sabemos que la proporción de aprobados en las pruebas de acceso a la Universidad es del 95%. Si queremos conocer la veracidad de esta información, consideramos la hipótesis: *La proporción de aprobados en las pruebas de acceso a la Universidad es igual al 95%*, y la contrastamos con la información obtenida a partir de una muestra. Si ambas informaciones coinciden dentro de un margen de error considerado admisible, mantendremos dicha hipótesis como cierta; en caso contrario, la rechazaremos y buscaremos nuevas hipótesis capaces de explicar los datos observados.

Hemos considerado una hipótesis que en principio admitimos como válida y que deseamos contrastar, a la que denominaremos **hipótesis nula, H_0,** y una hipótesis contraria a esta, que llamaremos **hipótesis alternativa, H_1,** y que admitiremos si el contraste de hipótesis da como resultado el rechazo de H_0.

- ***Hipótesis nula, H_0,*** *es la hipótesis que deseamos contrastar, considerada en principio como verdadera y que aceptaremos o rechazaremos como consecuencia de este contraste.*

- ***Hipótesis alternativa, H_1,*** *es cualquier otra hipótesis que nos sitúe frente a H_0 y que aceptaremos si, como consecuencia del contraste, rechazamos H_0.*

En el ejemplo anterior H_0 será: *la proporción de aprobados en las pruebas de acceso a la Universidad es del 95%*. Por el contrario H_1 podría ser: *la proporción de aprobados en las pruebas de acceso a la Universidad es distinta del 95%.*

La decisión de aceptar o rechazar la hipótesis nula dependerá de:

–La discrepancia entre la hipótesis estadística H_0 y una realidad de la que poseemos una información muestral.

–Que la discrepancia entre las dos hipótesis sea menor que un valor que consideraremos aceptable, en cuyo caso mantendremos la hipótesis nula y, en caso contrario, la rechazaremos.

El conjunto de valores que puede tomar la discrepancia da lugar a una variable aleatoria llamada **estadístico de contraste**. Los valores de este estadístico, que nos llevarán a aceptar la hipótesis nula, forman la **región de aceptación;** los que nos conduzcan a rechazar H_0 constituyen la **región de rechazo**.

- **Estadístico de contraste** es una variable aleatoria cuyo valor para una muestra determinada nos permitirá tomar la decisión sobre la aceptación o rechazo de la hipótesis.
- **Región de aceptación** es el conjunto de valores del estadístico de contraste que nos lleva a aceptar la hipótesis nula.
- **Región de rechazo** es el conjunto de valores del estadístico de contraste que nos lleva a rechazar la hipótesis nula.
- **Valores críticos** son los valores frontera entre la región de aceptación y la región de rechazo.

Cualquier contraste de hipótesis lleva aparejado el rechazo a aceptación de la hipótesis nula. Al tomar cualquiera de estas dos decisiones, existirá un riesgo de equivocarnos; es decir, si aceptamos la hipótesis nula cuando esta es cierta, no cometemos ningún error; del mismo modo tampoco cometemos error si la rechazamos siendo falsa; pero, ¿qué ocurre cuando aceptamos la hipótesis nula siendo falsa, o, por el contrario, la rechazamos cuando es cierta? Esta equivocación da lugar a dos tipos de errores:

- **Error de tipo I** es el que se produce cuando rechazamos la hipótesis nula siendo cierta. La probabilidad de cometer este error se llama **nivel de significación** del contraste y se designa por α. Este valor es conocido de antemano, pues es determinado previamente por la persona que realice el contraste.
- **Error de tipo II** es el que se produce cuando aceptamos la hipótesis nula siendo falsa.

	Si H_0 es cierta	**Si H_0 es falsa**
Aceptamos H_0	No hay error	Error tipo II
Rechazamos H_0	Error tipo I	No hay error

Ej. Si suponemos como hipótesis nula H_0: $p = 10\%$ y como hipótesis alternativa H_1: $p \neq 10\%$, indica, en los siguientes casos, si se comete o no error y, en caso de cometerlo, de qué tipo es:

a) $p = 10\%$ y rechazamos H_0 **b)** $p = 10\%$ y aceptamos H_0.

c) $p = 20\%$ y rechazamos H_0 **d)** $p = 20\%$ y aceptamos H_0.

2.- FASES DE UN CONTRASTE DE HIPÓTESIS.

Una dama afirma que el sabor de una taza de té con leche es distinto cuando se vierte antes la leche que el té.

Para contrastar esta información se preparan diez tazas de té; en cinco de ellas se vierte antes la leche y en las cinco restantes, antes el té. A continuación, la dama prueba en orden aleatorio las diez tazas y acierta en ocho de las diez. ¿Es este hecho una evidencia significativa a favor de la hipótesis?

Para poder contestar estudiaremos cada una de las fases en que se divide cualquier problema de contraste de hipótesis.

Fase 1. Definir la hipótesis nula y la hipótesis alternativa.

Definiremos la hipótesis nula y la hipótesis alternativa siguiendo las indicaciones del apartado anterior. Dependiendo del sentido de la hipótesis alternativa, podemos hablar de **contraste bilateral** o de **contraste unilateral.**

- **Contraste bilateral:** H_1: $p \neq p_o$
- **Contraste unilateral:** H_1: $p < p_o$; o bien, H_1: $p > p_o$

Ej. En el ejemplo anterior tomaremos como hipótesis nula la más conservadora, H_0 ={*el sabor de la taza de té es independiente del orden en que se viertan la leche y el té*} y como hipótesis alternativa, la nueva información, H_1 = {*el sabor de una taza de té es distinto si se vierte primero la leche y luego el té o si se hace al contrario*}. Estas hipótesis se verifican si al elegir una muestra, la proporción de aciertos es igual a 0.5 o mayor que 0.5, respectivamente.

Por tanto las hipótesis nula y alternativa serán:

H_0: $p = 0.5$ y H_1: $p > 0.5$

Siendo p la proporción de aciertos. En este caso se trataría de un contraste unilateral.

Fase 2. Determinación del estadístico de contraste.

Todos los estadísticos Z que vamos a utilizar dependerán del parámetro sobre el que se elabore la hipótesis nula.

- Si la hipótesis es sobre la media poblacional: $\mathbf{Z} = \dfrac{\overline{\mathbf{X}} - \mu}{\sigma / \sqrt{\mathbf{n}}}$

- Si la hipótesis es sobre la proporción poblacional: $\mathbf{Z} = \dfrac{\hat{\mathbf{P}} - \mathbf{p}}{\sqrt{\dfrac{\mathbf{p} \cdot (1 - \mathbf{p})}{\mathbf{n}}}}$

- Si la hipótesis es sobre la diferencia de medias: $\mathbf{Z} = \dfrac{(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{\mathbf{n}_1} + \dfrac{\sigma_2^2}{\mathbf{n}_2}}}$

Todos estos estadísticos se aproximan a una distribución normal N(0, 1)

Ej. En nuestro ejemplo, la hipótesis se realiza sobre la proporción poblacional, siendo, por tanto, el estadístico de contraste:

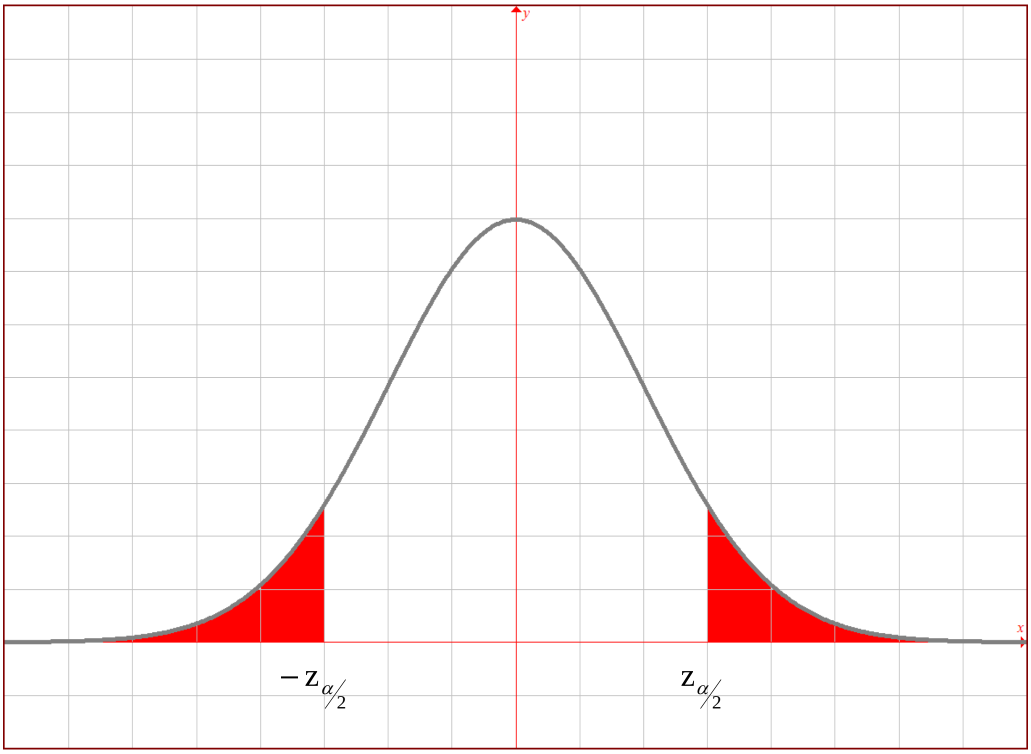
$$Z=\frac{\hat{P}-p}{\sqrt{\frac{p\cdot(1-p)}{n}}}=\frac{\hat{P}-0.5}{\sqrt{\frac{0.5\cdot(1-0.5)}{10}}}=\frac{\hat{P}-0.5}{0.16}$$

Fase 3. Determinación de las regiones de aceptación y de rechazo.

Determinaremos las regiones de aceptación y de rechazo a partir del *nivel de significación,* α, es decir, de la probabilidad de rechazar la hipótesis nula cuando esta es cierta. Tendremos que prefijar este nivel de antemano y, normalmente, tomará valores muy pequeños, como 0.05; 0.01…

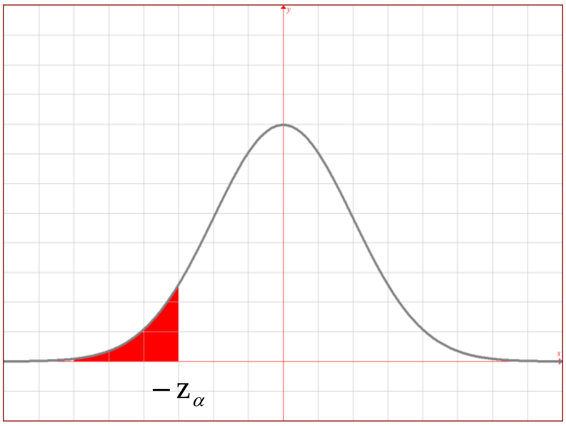
Una vez fijado el nivel de significación, y puesto que todas las distribuciones que siguen los estadísticos de contraste se aproximan a una N(0, 1), podremos hallar las regiones de aceptación y de rechazo según las siguientes indicaciones:

- Si se trata de un *contraste bilateral*, la región de rechazo está formada por las dos zonas que quedan a ambos lados de los *valores críticos*: $-z_{\alpha/2}$ y $z_{\alpha/2}$ tal y como se muestra en la figura. El área de estas dos zonas juntas es α, y los valores críticos se calculan de igual forma a como se hacía en el tema anterior.

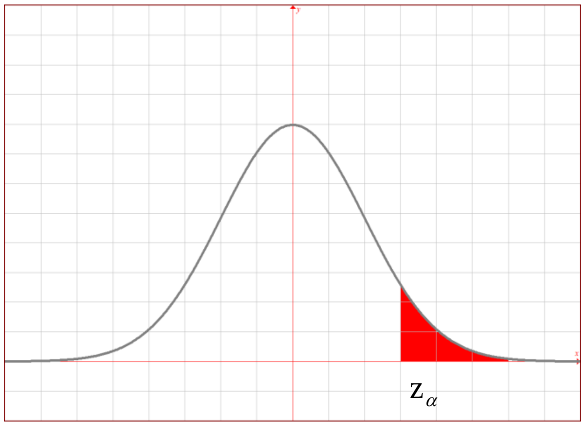


Contraste bilateral

- Si se trata de un *contraste unilateral*, la región de rechazo será una única zona, cuyo área será igual a α, y el valor crítico, z_{α}, se calculará a partir de las tablas de la normal N(0, 1).



Contraste unilateral H_1: $p < p_o$



Contraste unilateral H_1: $p > p_o$

Fase 4. Valor del estadístico y obtención de conclusiones.

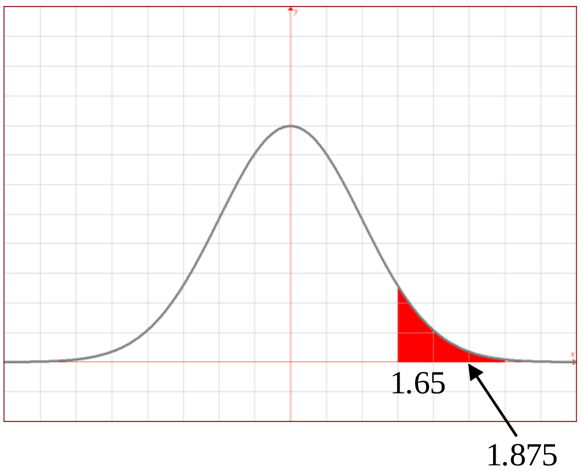
A partir de la muestra observada podemos obtener un valor concreto de la medida de la discrepancia o estadístico de contraste. Una vez obtenido este valor, y puesto que conocemos la región de aceptación y de rechazo, determinaremos si este valor es considerado aceptable o no y, por tanto, si aceptamos o rechazamos la hipótesis nula.

Ej. Continuando con el ejemplo anterior, podemos tomar como un buen nivel de significación α = 0.05, y como se trata de un contraste unilateral, tenemos que $z_{\alpha} = 1.65$ es el valor crítico que separa la región de aceptación de la de rechazo.

Obtuvimos que el estadístico era: $Z = \dfrac{\hat{P} - 0.5}{0.16}$; como $\hat{P} = \dfrac{8}{10} = 0.8$, se tiene que el valor concreto del estadístico es:

$$\frac{0.8 - 0.5}{0.16} = 1.875$$

Como 1.875 > 1.65, este valor cae en la región de rechazo y, por tanto, rechazamos la hipótesis de que el sabor de una taza de té es independiente del orden en que se mezcla el té y la leche, con un nivel de significación del 5%.



Ej. Si en el ejemplo anterior consideramos un nivel de significación del 0.01, ¿aceptaríamos o rechazaríamos la hipótesis?

Ej. Calcula el valor que separa la región de aceptación de la de rechazo para un nivel de significación $\alpha = 0.05$, en caso de contraste bilateral.

3.- CONTRASTE PARA LA MEDIA DE UNA POBLACIÓN.

Sea una población normal $N(\mu, \sigma)$ con desviación típica σ conocida y en la que deseamos contrastar la hipótesis $\mu = \mu_0$. Frente a esta hipótesis tenemos varias opciones de hipótesis alternativas: $\mu \neq \mu_0$, $\mu > \mu_0$ o $\mu < \mu_0$. La primera de estas opciones da lugar a un contraste bilateral, mientras que los otros dos casos dan lugar a un contraste unilateral. La elección de uno u otro tipo de hipótesis alternativa dependerá del contexto del problema.

En cualquiera de estas situaciones, si tomamos una muestra de tamaño n, el estadístico de contraste es $Z = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$, que sigue una normal $N(0, 1)$, pues $\overline{X}$ sigue una distribución $N\left(\mu_0, \dfrac{\sigma}{\sqrt{n}}\right)$.

De este modo, si $\overline{x}$ es la media de la muestra, aceptaremos la hipótesis nula si el valor $Z = \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$ cae en la región de aceptación. Esta región dependerá del tipo de contraste (bilateral o unilateral) y del valor del nivel de significación α a partir del cual podemos calcular $z_{\alpha/2}$ o z_α .

Por tanto, aceptamos H_0: $\mu = \mu_0$ con un nivel de significación α si:

H_1: $\mu \neq \mu_0$	H_1: $\mu > \mu_0$	H_1: $\mu < \mu_0$
$-z_{\alpha/2} < \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} < z_{\alpha/2}$	$\dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} < z_\alpha$	$-z_\alpha < \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$

Si σ es desconocida, podemos tomar como valor de σ la desviación típica muestral cuando el tamaño de la muestra es grande ($n \geq 30$).

Ej. La velocidad media de la última contrarreloj en la vuelta ciclista a España se cree que fue de 40 km/h, con una desviación típica de 5 km/h. Para contrastar esta información se decidió estudiar la velocidad media de 15 corredores elegidos aleatoriamente, resultando que las velocidades medias de cada uno de ellos eran, en km/h: 43 43.2 45 43.1 39.5 45 43 42.3 39.6 40 39 45 39 44.1 42.6

¿Se puede aceptar la hipótesis de que la velocidad media de los ciclistas fue de 40 km/h con un nivel de significación del 5%?

<u>Fase 1</u>: H_0: $\mu = 40$ y H_1: $\mu \neq 40$. Se trata de un contraste bilateral.

Fase 2: El estadístico de contraste es:

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{\overline{X} - 40}{5/\sqrt{15}}$$

Fase 3: Como el nivel de significación es del 5% se tiene que α = 0.05.

Al ser contraste bilateral, el valor crítico es $z_{\alpha/2} = 1.96$, siendo, por tanto, la región de aceptación (–1.96, 1.96).

Fase 4: Como el valor de la media muestral es $\bar{x} = 42.23$, el valor de este estadístico para esta muestra es:

$$\frac{42.43 - 40}{5/\sqrt{15}} = 1.71$$

Ya que $1.71 \in [-1.96, 1.96]$, el valor del estadístico se encuentra dentro del nivel de aceptación; por consiguiente, admitimos la hipótesis de que la velocidad media ha sido de 40 km/h con un nivel de significación del 5%.

**En la aceptación o rechazo de H_0 un parámetro importante es el nivel de significación. Así, si en el ejemplo anterior α = 0.1, tendríamos que $z_{\alpha/2} = 1.65$ y la región de aceptación sería (–1.65, 1.65).*

Como $1.71 \notin [-1.65, 1.65]$, se rechazaría la hipótesis nula. Aceptaríamos que $\mu \neq 40$.

Nota: También se puede aceptar H_0: μ = μ_0 con un nivel de significación α si:

$$\bar{x} \in \left[\mu_0 - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \mu_0 + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right]$$

En caso contrario se acepta H_1.

En nuestro caso:

$$\left[\mu_0 - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \mu_0 + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right] = \left[40 - 1.96 \cdot \frac{5}{\sqrt{15}}, 40 + 1.96 \frac{5}{\sqrt{15}} \right] = [37.47,\ 42.53]$$

Como 42.43 $\in$ [37.47, 42.53] se acepta la hipótesis nula.

*Si hubiésemos deseado contrastar la hipótesis H_0: $\mu \geq \mu_0$, la hipótesis alternativa sería H_1: $\mu < \mu_0$. En este caso se resuelve igual que H_0: $\mu = \mu_0$ y H_1: $\mu < \mu_0$

De igual modo, si H_0: $\mu \leq \mu_0$, entonces H_1: $\mu > \mu_0$, siendo su resolución idéntica a la realizada para el caso: H_0: $\mu = \mu_0$ y H_1: $\mu > \mu_0$

Ej. Según los datos aportados por el partido en la oposición de un Ayuntamiento de una ciudad, el salario medio de los habitantes de esa ciudad es de 1300 €, con una desviación típica de 400 €.

El partido que gobierna no comparte estos resultados y afirma que el sueldo medio es mayor.

A fin de ratificar esta información, el Ayuntamiento decide realizar un contraste de hipótesis con un nivel de significación del 5%. La empresa encargada de este contraste selecciona 100 personas aleatoriamente y obtiene que su sueldo medio es de 1450 €.

Podemos considerar:

$$H_0:\ \mu_0 \le 1300$$

$$H_1:\ \mu_0 > 1300$$

Se trata de un contraste unilateral del que conocemos α = 0.05 $\Rightarrow z_\alpha = 1.65$. El intervalo de aceptación será pues: (–∞, 1.65).

Como $\bar{x} = 1450$, tenemos:

$$\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{1450-1300}{400/\sqrt{100}} = 3.75 > 1.65$$

Luego rechazamos la hipótesis nula con un nivel de significación del 5%; es decir, se rechazará la información dada por la oposición y admitiremos la nueva información H_1 como cierta.

Observamos que también habríamos podido considerar:

$$H_0:\ \mu_0 = 1300 \quad \text{y} \quad H_1:\ \mu_0 > 1300$$

sin que ello influyera en la resolución del problema.

Ej. La nota media de los estudiantes que finalizaron el módulo de Educación Infantil en una determinada Comunidad Autónoma en el curso 01-02 es de 5.5 con una desviación típica de 2.

a) Si elegimos al azar 30 estudiantes que han finalizado los estudios de dicho módulo en el curso 02-03 y obtenemos que su nota media es de 6, ¿podemos mantener los datos censados en el año anterior con un nivel de significación del 0.05?

b) Si al elegir la muestra de 30 estudiantes obtenemos una media de 7, ¿qué podemos deducir con el mismo nivel de significación?

4.- CONTRASTE PARA LA PROPORCIÓN.

Partimos de una distribución binomial B(n, p) y suponemos que deseamos contrastar la hipótesis de que la proporción de individuos con un determinado atributo es p_0. En este caso, H_0: $p = p_0$ y H_1 puede ser $p \neq p_0$, $p > p_0$ o $p < p_0$ dependiendo del contexto del problema.

Si tomamos una muestra de tamaño *n*, el estadístico de contraste es $Z = \dfrac{\hat{P}-p}{\sqrt{\dfrac{p\cdot(1-p)}{n}}}$,

que se aproxima a una N(0, 1), pues la distribución muestral de proporciones $\hat{P}$, se aproxima a una $N\left(p, \sqrt{\dfrac{p\cdot(1-p)}{n}}\right)$.

Aceptaremos H_0 si el valor del estadístico de contraste para la muestra observada cae dentro de la región de aceptación.

Al igual que en el caso anterior, la región de aceptación dependerá del tipo de contraste (bilateral o unilateral) y del valor del nivel de significación α a partir del cual podemos calcular $z_{\alpha/2}$ o z_α.

Por tanto, aceptaremos la hipótesis H_0: $p = p_0$ con un nivel de significación α si:

H_1: $p \neq p_0$	**H_1: $p > p_0$**	**H_1: $p < p_0$**
$-z_{\alpha/2} < \dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0\cdot(1-p_0)}{n}}} < z_{\alpha/2}$	$\dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0\cdot(1-p_0)}{n}}} < z_\alpha$	$-z_\alpha < \dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0\cdot(1-p_0)}{n}}}$

Como vimos en el apartado anterior, la hipótesis nula no interviene en la forma de resolver el problema; por tanto si tenemos otra hipótesis nula diferente de $p = p_0$, resolveremos el problema fijándonos en el sentido de la hipótesis alternativa.

Ej. El fabricante de una crema para aliviar el dolor de las quemaduras afirma que esta tiene efectividad del 90% durante un periodo de cinco horas. Seleccionamos una muestra de 40 personas que sufren quemaduras y les aplicamos dicha crema, de forma que 32 de ellas resultan aliviadas. Si deseamos contrastar la afirmación del fabricante a partir de los resultados obtenidos en esta muestra, con un nivel de significación del 5%, tenemos que:

H_0: p = 0.9 $\quad$ H_1: p ≠ 0.9 $\qquad$ α = 0.05 $\quad$ $z_{\alpha/2} = 1.96$ $\quad$ $\hat{p} = \dfrac{32}{40} = 0.8$

Aceptamos la hipótesis nula si:

$-z_{\alpha/2} < \dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0\cdot(1-p_0)}{n}}} < z_{\alpha/2}$ $\rightarrow$ como: $-1.96 \not< \dfrac{0.8-0.9}{\sqrt{\dfrac{0.9\cdot 0.1}{40}}} < -2.1 < 1.96$, rechazamos la hipótesis del fabricante con un 5% de nivel de significación.

Nota: También se puede aceptar H_0: p = p_0 con un nivel de significación α si:

$$\hat{p} \in \left[p_0 - z_{\alpha/2}\cdot\sqrt{\frac{p_0\cdot(1-p_0)}{n}},\ p_0 + z_{\alpha/2}\cdot\sqrt{\frac{p_0\cdot(1-p_0)}{n}} \right]$$

En nuestro ejemplo la región de aceptación es:

$$\left[0.9-1.96\cdot\sqrt{\frac{0.9\cdot 0.1}{40}},\ 0.9+1.96\cdot\sqrt{\frac{0.9\cdot 0.1}{40}}\right] = [0.807,\ 0.993]$$

Como $0.8 \notin [0.807,\ 0.993]$ rechazamos la hipótesis nula.

Ej. En un producto de comida para perros se especifica que el contenido mínimo de proteínas es del 40%. Con el fin de comprobar esta especificación analizamos 40 latas para analizar su contenido en proteínas y obtenemos un contenido medio de 39%, con una desviación típica del 3%.

¿Es correcta la especificación del producto con un nivel de significación del 0.05?

5.- EJERCICIOS PROPUESTOS.

1.- Contrasta la hipótesis nula H_0: $\mu = 5$ frente a la hipótesis alternativa H_1: $\mu > 5$, sabiendo que $\sigma = 2$, $n = 16$ y $\alpha = 0.01$.

2.- En cada uno de los siguientes casos, indica si rechazaremos o aceptaremos la hipótesis nula con un nivel de significación del 0.01:

a) $n = 10$	$\hat{p} = 0.6$	H_0: $p = 0.5$	H_1: $p \neq 0.5$
b) $n = 100$	$\hat{p} = 0.6$	H_0: $p = 0.5$	H_1: $p \neq 0.5$
c) $n = 1000$	$\hat{p} = 0.8$	H_0: $p = 0.5$	H_1: $p \neq 0.5$
d) $n = 10$	$\hat{p} = 0.6$	H_0: $p = 0.5$	H_1: $p > 0.5$
e) $n = 100$	$\hat{p} = 0.6$	H_0: $p = 0.5$	H_1: $p > 0.5$
f) $n = 1000$	$\hat{p} = 0.8$	H_0: $p = 0.5$	H_1: $p > 0.5$

3.- La duración media de 100 tubos fluorescentes producidos por una fábrica A resultó ser de 1500 horas, con una desviación típica de 115. Sin embargo, la dirección de dicha fábrica afirma que la duración media es de 1550 horas.

Contrasta la información dada por la dirección de la empresa con un nivel de significación del 0.05:

a) Si la hipótesis alternativa es que la duración media es distinta de 1550 horas.

b) Si la hipótesis alternativa es que la duración media es menor de 1550 horas.

4.- Según una encuesta, la proporción de población que lee el periódico diariamente en cierta Comunidad Autónoma es del 40%. Como esta encuesta fue realizada delante de los puestos de prensa, nosotros creemos que esta información es incorrecta y que este tanto por ciento es distinto, por lo que decidimos contrastar la información dada por la empresa. Para ello seleccionamos una muestra de 100 personas aleatoriamente y les preguntamos si leen o no el periódico diariamente.

¿Cuál es el número mínimo de personas que tienen que responder afirmativamente a esta pregunta para aceptar la afirmación dada por la encuesta con un nivel de significación del 5%?

5.- Una cadena de hoteles afirma que el grado medio de satisfacción de sus clientes en una escala de uno a cien es de 83, con una desviación típica de 7. A fin de contrastar esta información, preguntamos a 50 clientes del hotel, elegidos aleatoriamente, y obtenemos un grado de satisfacción medio de 79.

¿Podemos admitir como cierta la información dada por la cadena de hoteles sobre el grado de satisfacción de sus clientes con un nivel de significación del 0.06?

6.- Un laboratorio farmacéutico afirma que un calmante de su fabricación quita el dolor en menos de 14 minutos. Para comprobar esto se han tomado 38 personas y se ha medido el tiempo transcurrido entre la toma del medicamento y el momento en que desaparece el dolor de cabeza, observando una media de 19 minutos y una desviación típica de 7 minutos. A un nivel de significación de 0,05, ¿qué podemos decir de la afirmación hecha por el laboratorio?. (Suponemos que la población se distribuye normalmente).

7.- Una academia afirma que al menos el 80% de sus alumnos aprueban. Con el fin de contrastarlo se elige aleatoriamente una muestra de 50 alumnos de los cuales aprueban 38. ¿Se puede admitir la propaganda de la academia a un nivel de significación del 5%?

8.- En las últimas elecciones, celebradas hace un año, el 52 por ciento de los votantes de una ciudad estaban a favor del alcalde. Una encuesta, realizada recientemente, indica que, de 350 ciudadanos elegidos al azar, 198 están a favor del alcalde:

a) ¿Se puede afirmar con un nivel de confianza del 90 %, que el alcalde gana popularidad?.

b) ¿Se obtiene la misma respuesta en el apartado anterior si el nivel de confianza es igual a 0,99?

9.- Una encuesta realizada a 64 empleados de una fábrica, concluyó que el tiempo medio de duración de un empleo en la misma era de 6,5 años, con una desviación típica de 4. ¿Sirve esta información para aceptar, con un nivel de significación del 5%, que el tiempo medio de empleo en esa fábrica es menor o igual que 6?

Justifica adecuadamente la respuesta.

10.- Un equipo de psicólogos ha comprobado que en cierta población infantil, el tiempo (en minutos) empleado en realizar determinada actividad manual sigue un modelo normal de probabilidad. Un grupo de 36 niños, seleccionados aleatoriamente, en dicha población, realizaron esa actividad manual en un tiempo medio de 6,5 minutos con una desviación típica muestral de 1,5 minutos. A partir de esa información:

a) ¿Qué error máximo cometeremos, con una confianza del 95%, si estimamos en 6,5 minutos el tiempo medio empleado en realizar la actividad manual en dicha población infantil?

b) Para un nivel de significación del 1% ($\alpha = 0{,}01$), ¿podríamos rechazar la hipótesis de que el tiempo medio en la población es de 7 minutos?

(*Junio 96*)

11.- A una muestra de 169 deportistas seleccionados aleatoriamente se les preguntó cuánto tiempo dedicaban diariamente a su entrenamiento. Como resumen de la información recogida, se obtuvo un tiempo medio de 4,3 horas y una desviación típica de 1,5 horas. Para un nivel de significación del 1 % ($\alpha = 0{,}01$), ¿podríamos rechazar la hipótesis de que el tiempo medio al día que dedica un deportista de dicha población a su entrenamiento es de 4 horas? Justificar la respuesta.

(*Septiembre 95*)

12.- Se ha comprobado que el tiempo de espera (en minutos) hasta ser atendido, en un cierto servicio de urgencias, sigue un modelo normal de probabilidad. A partir de una muestras de 100 personas que fueron atendidas en dicho servicio, se ha calculado un tiempo medio de espera de14,25 minutos y una varianza de 6,25 minutos^2.

a) ¿Podríamos afirmar con un nivel de significación del 5% ($\alpha = 0{,}05$) que el tiempo medio de espera, en este servicio de urgencias, no es 15 minutos?

b) ¿Qué podríamos concluir, si el nivel de significación hubiera sido del 0,1%?.

c) ¿Existe contradicción en ambas situaciones?

Justificar las respuestas. (Junio 97)

13.- Un equipo de educadores ha comprobado que en cierta población infantil, el tiempo de reacción (en centésimas de segundo) ante determinado estímulo educativo sigue un modelo normal de probabilidad. A partir de una muestra de 100 niños de dicha población se ha obtenido una cuasi-desviación típica de 12. Teniendo en cuenta la información proporcionada se ha contrastado la hipótesis de que la media poblacional es 5 frente a la hipótesis de que es distinta de 5, resultando que para un nivel de significación de un 5% se rechaza que la media poblacional sea 5, mientras que para un nivel de significación de un 1% se acepta que dicha media sea 5.

a) ¿Cuál entre los siguientes sería el valor experimental con el que se ha realizado el contraste de hipótesis : 1.85 , 2.5 , 2.75 ?

b) ¿Cuál sería el tiempo medio de reacción obtenido con los 100 niños de la muestra? Justificar las respuestas.

(*Septiembre 98*)

14.- A partir de los datos recogidos sobre una muestra aleatoria de 121 pequeñas y medianas empresas de una región se ha calculado, para el año 2001, un beneficio medio de 89 millones de euros con una cuasivarianza de 30.25 millones de euros^2.

Contesta justificando las respuestas:

a) ¿Podríamos rechazar (con un nivel de significación del 0.001) la afirmación de que los beneficios medios en la pequeña y mediana empresa de dicha región son de 90 millones de euros?

b) ¿Qué ocurriría para el nivel de significación 0.05?

(*Junio 2001*)

15.- Si el gerente de una empresa selecciona aleatoriamente entre sus trabajadores una muestra de 169 y anota el número de horas de trabajo que cada uno de ellos ha perdido por causa de accidentes laborales en el año 2001. A partir de la información obtenida determina, en esos 169 trabajadores, un número medio de horas perdidas por accidentes laborales en el 2001 de 36.5 horas. Sabiendo que:

$$\sum_{1}^{169}(x_i - 36.5)^2 = 15970.5$$

donde x_i representa el número de horas perdidas por el i-ésimo trabajador, i=1,......,169.

a) ¿Podríamos rechazar, con un nivel de significación del 1 %, la hipótesis de que el número medio de horas perdidas por causa de accidentes laborales en esa empresa durante el año 2001 fue de 35 horas?.

b) ¿Y para un nivel de significación del 5 %?.

Justifica las respuestas.

(*Septiembre 2002*)