

## Decisión estadística. Contraste de hipótesis

**Nota.-** Cuando tratábamos la estimación de parámetros, intentábamos obtener un valor o un intervalo de valores que constituyesen la mejor estimación del parámetro desconocido a partir de la información muestral.

**Nota.-** Este nuevo ámbito denominado teoría de la decisión estadística trata de utilizar los datos obtenidos a partir de una muestra para tomar decisiones sobre la población.

◆◆ Por ejemplo, se puede decidir, a partir de los datos del muestreo, si un fármaco nuevo es realmente efectivo para el tratamiento de una enfermedad, si un método de aprendizaje es mejor que otro, si unos nuevos productos manufacturados son mejores que los fabricados con anterioridad, si una determinada moneda está o no cargada, etc.

**Nota.-** Estas decisiones se llaman **decisiones estadísticas**. Hay que tener en cuenta que estas decisiones están tomadas sobre una base probabilística, es decir, el acierto de la decisión se mide objetivamente en términos de probabilidad.

**Nota.-** En las situaciones en las que deseamos tomar decisiones es conveniente realizar determinados supuestos o conjeturas acerca de las poblaciones que se estudian. Estos supuestos que pueden ser o no ciertos se denominan **hipótesis estadísticas**.

**Nota.-** En muchos casos se formulan las hipótesis estadísticas con el único propósito de rechazarlas o invalidarlas. Por ejemplo, se dice de un parámetro de la población que no posee variación, o que un procedimiento no es mejor que otro, que una moneda no está cargada. Estas hipótesis se denominan **hipótesis nulas** y se denotan por  $H_0$ .

Cualquier hipótesis que difiera de una hipótesis dada se llama **hipótesis alternativa** y se denotan por  $H_1$ .

◆◆ Por ejemplo, si la hipótesis nula  $H_0$  es que una moneda no está cargada, es decir,  $p=0,5$ ; las posibles hipótesis alternativas  $H_1$ , serían  $p \neq 0,5$ ,  $p > 0,5$  o  $p < 0,5$ .

**Def.- Hipótesis nula** es la denominación que recibe la hipótesis a contrastar, que se considera provisionalmente como verdadera y que es revisada tras haber obtenido información de los datos muestrales. Se designa por  $H_0$ .

**Def.-** Hipótesis alternativa es la denominación que recibe el conjunto de situaciones restantes posibles, o admitidos como posibles, en una situación experimental dada. Se designa por  $H_1$ .

**Nota.-** En la toma de decisiones estadísticas, toda hipótesis nula  $H_0$  ha de ir acompañada de una alternativa  $H_1$  que es la que aspira a desplazar a la nula. Para una hipótesis nula concreta puede haber diversas hipótesis alternativas.

## Pruebas de Hipótesis. Errores

**Def.-** Se llaman **pruebas de hipótesis** a los procedimientos que permiten decidir si una hipótesis se acepta o se rechaza o determinar si las muestras observadas difieren significativamente de los resultados esperados.

**Nota.-** Estos procedimientos, también reciben los nombres de **ensayo de hipótesis**, **ensayos (test) de significación** o **reglas de decisión**.

**Nota.-** La metodología usual del contraste de la hipótesis nula frente a una alternativa, en base a la información suministrada por la muestra, puede conducir a **dos tipos de errores**, debidos a la aleatoriedad del muestreo.

**Nota.-** Hay ocasiones en las que es cierta la hipótesis nula, pero, como a la vista de la información muestral es muy poco probable, nuestra decisión puede ser rechazarla. En estos casos cometemos un error. El error consistente en rechazar la hipótesis nula cuando es verdadera se denomina **error de tipo I**.

**Nota.-** Puede ocurrir que, basándonos en la información muestral, decidamos aceptar  $H_0$  siendo falsa, que sería lo mismo que rechazar  $H_1$  siendo verdadera. En estas situaciones se comete un error denominado **error de tipo II**.

**Nota.-** Cuando se comete uno u otro tipo de error en una prueba de hipótesis, éstos conducen a decisiones inadecuadas.

**Nota.-** Siendo  $H_0$  y  $H_1$  la hipótesis nula y la hipótesis nula alternativa, respectivamente, podemos resumir las decisiones correctas y erróneas que se pueden producir en el contraste, mediante el siguiente cuadro.

	$H_0$ es verdadera	$H_0$ es falsa
Se acepta $H_0$	acierto	ERROR DE TIPO II
Se rechaza $H_0$	ERROR DE TIPO I	acierto

## Etapas de las Pruebas de Hipótesis

**Nota.-** Los procedimientos seguidos en las pruebas de hipótesis correspondientes a las situaciones de decisión estadística se encuentran totalmente estandarizados y los pasos o etapas a seguir se describen a continuación.

- *Etapas 1 : Enunciar la hipótesis nula ( $H_0$ ) y la alternativa o de investigación ( $H_1$ ).*

**Nota.-** Ambas hipótesis han de ser excluyentes entre sí. La hipótesis nula ( $H_0$ ) es la hipótesis que se desea contrastar. Se suele formular con el propósito de ser rechazada. Cuando se consigue rechazarla, se acepta la hipótesis alternativa ( $H_1$ ).

**Nota.-** En las pruebas de hipótesis para la media la hipótesis nula  $H_0$  y la alternativa  $H_1$ , podrán ser:

+ *En contrastes bilaterales*, que contrastan la hipótesis nula  $H_0: \mu = \mu_0$ , contra la hipótesis alternativa,  $H_1: \mu \neq \mu_0$ .

+ *En contrastes unilaterales por la derecha*, que contrastan la hipótesis nula  $H_0: \mu \leq \mu_0$ , contra la hipótesis alternativa  $H_1: \mu > \mu_0$ .

+ *En contrastes unilaterales por la izquierda*, que contrastan la hipótesis nula  $H_0: \mu \geq \mu_0$ , contra la hipótesis alternativa  $H_1: \mu < \mu_0$ .

**Nota.-** En las pruebas de hipótesis para la proporción la hipótesis nula  $H_0$  y la alternativa  $H_1$ , podrán ser:

+ *En contrastes bilaterales*, que contrastan la hipótesis nula  $H_0: p = p_0$ , contra una hipótesis alternativa,  $H_1: p \neq p_0$ .

+ *En contrastes unilaterales por la derecha*, que contrastan la hipótesis nula  $H_0: p \leq p_0$ , contra  $H_1: p > p_0$ .

+ *En contrastes unilaterales por la izquierda*, que contrastan la hipótesis nula  $H_0: p \geq p_0$ , contra  $H_1: p < p_0$ .

**Nota.-** En las pruebas de hipótesis para la diferencia de medias la hipótesis nula  $H_0$  y la alternativa  $H_1$ , podrán ser:

+ *En contrastes bilaterales*, que contrastan la hipótesis nula  $H_0: \mu_1 = \mu_2$ , contra una hipótesis alternativa,  $H_1: \mu_1 \neq \mu_2$ .

+ *En contrastes unilaterales por la derecha*, que contrastan la hipótesis nula  $H_0: \mu_1 \leq \mu_2$ , contra  $H_1: \mu_1 > \mu_2$ .

+ *En contrastes unilaterales por la izquierda*, que contrastan la hipótesis nula  $H_0: \mu_1 \geq \mu_2$ , contra  $H_1: \mu_1 < \mu_2$ .

- *Etapa 2 : Determinar la región crítica, prefijado el nivel de significación ( $\alpha$ ).*

**Nota.-** El nivel de significación “ $\alpha$ ” es un valor probabilístico muy pequeño que se establece con anterioridad a la obtención del resultado.

Se rechazará  $H_0$  en favor de  $H_1$  cuando la prueba estadística dé lugar a un valor cuya probabilidad de que ocurra bajo  $H_0$  es igual o menor que el nivel de significación “ $\alpha$ ”.

**Nota.-** Dado el nivel de significación “ $\alpha$ ”, sabemos que el nivel de confianza es “ $1 - \alpha$ ”. Evidentemente la suma de ambos es 1.

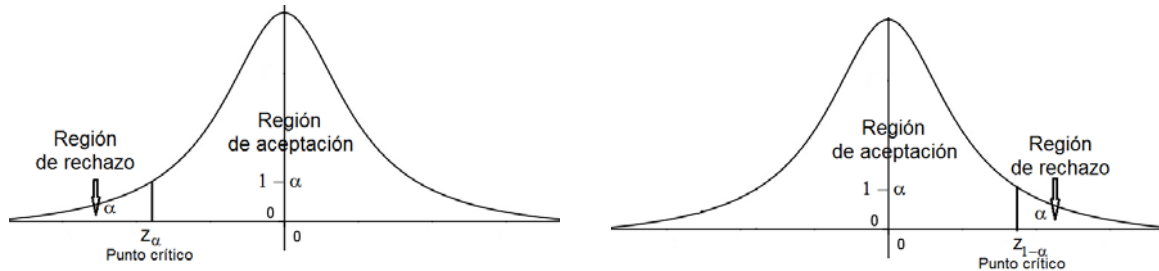
**Nota.-** La región crítica está limitada por uno o dos **valores críticos**, que se determinan de tal modo que la probabilidad de que el valor del estadístico de prueba que se obtenga de la muestra elegida esté en la región crítica no sea mayor que el nivel de significación.

En un contraste unilateral los valores críticos serán  $z_\alpha$  y  $z_{1-\alpha}$ , según sea la dirección de la hipótesis alternativa  $H_1$ .

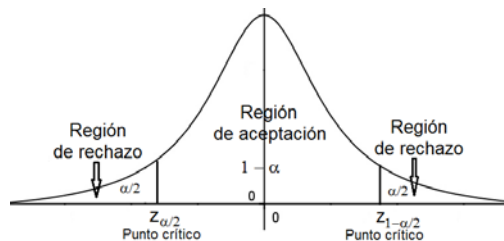
Como las distribuciones sabemos que son simétricas tenemos que  $z_\alpha = -z_{1-\alpha}$ .

Recordamos que para calcular  $z_{1-\alpha}$ , utilizábamos la fórmula  $p(Z < z_{1-\alpha}) = 1 - \alpha$ , lo mirábamos en la tabla de la  $N(0, 1)$  y obteníamos el correspondiente valor de  $z_{1-\alpha}$ .

Vamos a indicarlo gráficamente:



En un contraste bilateral (nos lo dice la hipótesis alternativa  $H_1$ ) los valores críticos serán  $z_{\alpha/2}$  y  $z_{1-\alpha/2}$ , y al ser simétricos sabemos que  $z_{\alpha/2} = -z_{1-\alpha/2}$ . Veámoslo gráficamente:



- *Etapa 3: Determinar el estadístico apropiado para la prueba.*

Nota.- Los estadísticos que van a ser utilizados en las pruebas de hipótesis que se desarrollan y se plantean en esta unidad son la media, las proporciones y la diferencia de medias. Todos ellos serán estudiados suponiendo que en la mayoría de las situaciones seguirán una distribución normal.

- Recordamos que si en la *población*  $X \rightarrow N(\mu_0; \sigma)$ , en las *muestras*  $\bar{X} \rightarrow N(\mu_0; \sigma/\sqrt{n})$ . En este caso el **estadístico de prueba** es  $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ , que sigue una normal tipificada  $N(0, 1)$ .

- Recordamos que en la *población*  $X \rightarrow B(n; p_0) \approx N(n \cdot p_0; \sqrt{n \cdot p_0 \cdot q_0})$ , en las *muestras*

$\hat{p} \rightarrow N(p_0; \sqrt{\frac{p_0 \cdot q_0}{n}})$ . En este caso el **estadístico de prueba** es  $Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot q_0}{n}}}$ , que sigue una

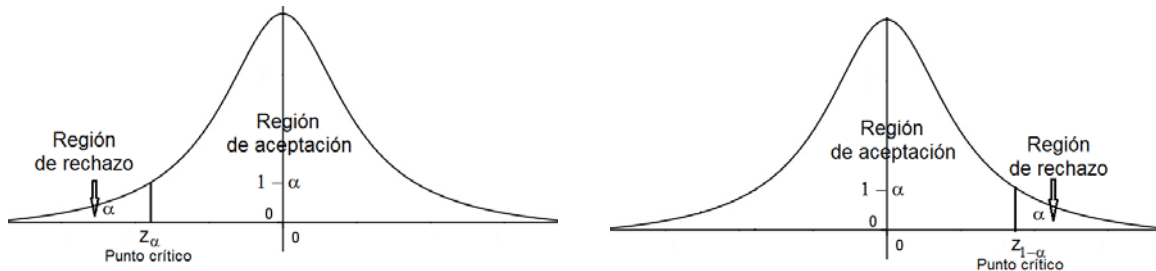
normal tipificada  $N(0, 1)$ .

- Recordamos que si en las *población*  $X_1 \rightarrow N(\mu_1; \sigma)$ ,  $X_2 \rightarrow N(\mu_2; \sigma)$  en las *muestras* resulta que  $\bar{X}_1 \rightarrow N(\mu_1; \sigma_1/\sqrt{n_1})$ ,  $\bar{X}_2 \rightarrow N(\mu_2; \sigma_2/\sqrt{n_2})$ . En este caso el **estadístico de prueba**

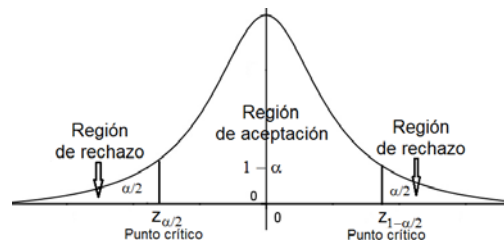
es  $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ , que sigue una normal tipificada  $N(0,1)$ .

**Nota.-** Recordamos las regiones de rechazo, que ya hemos puesto, que dependían del nivel de significación “ $\alpha$ ”, y por supuesto de la hipótesis alternativa  $H_1$ .

En contrastes unilaterales, tendríamos:



En un contraste bilateral, tendríamos:



- *Etap 4 : Calcular el valor observado del estadístico de prueba.*

- Si el **estadístico de prueba** es la variable  $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ , el **valor observado del estadístico de prueba** será número  $z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ , que resulta de sustituir la variable  $\bar{X}$ , por la

media aritmética,  $\bar{x}$ , de los datos de la muestra obtenida.

- Si el **estadístico de prueba** es la variable  $Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot (1-p_0)}{n}}}$ , el **valor observado del estadístico de prueba** será número  $z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot (1-p_0)}{n}}}$ , donde  $\hat{p} = x/n$ , es la proporción de éxitos “ $x$ ” que se han observado en una muestra de tamaño “ $n$ ”.

- Si el **estadístico de prueba** es la variable  $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ , el **valor observado**

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

del estadístico de prueba será número  $z_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ , que resulta de sustituir

la variable  $\bar{X}_1$  y  $\bar{X}_2$ , por la media aritmética,  $\bar{x}_1$  y  $\bar{x}_2$ , respectivamente de los datos de la muestra obtenida.

- *Etapas 5: Tomar la decisión e interpretarla.*

**Nota.-** Una vez calculados los **valores críticos** ( $z_\alpha$ ,  $z_{1-\alpha}$  ó  $z_{\alpha/2}$  y  $z_{1-\alpha/2}$ ) y los **valores ob-**

**servados de los estadísticos de prueba** ( $z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ ,  $z_0 = \frac{\hat{p} - p_0}{\sqrt{p_0 \cdot (1-p_0)/n}}$  ó

$z_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ ), se tomará la decisión pertinente a la hipótesis nula.

**Nota.-** *La aceptación o rechazo de una hipótesis la tomaremos en función de los valores críticos. Sólo tendremos que ver si los valores observados de los estadísticos de prueba están en la región de aceptación o de rechazo, limitados por los puntos críticos.*

## Veamos unos cuantos ejemplos:

*Para el contraste de medias*

♦♦ Una encuesta a 64 profesionales de una institución reveló que el tiempo medio de empleo en dicho campo era de 5 años, con una desviación típica de 4. Considerando un nivel de significación del 0,05, ¿sirven estos datos de soporte de que el tiempo medio de empleo de los profesionales de esta institución está por debajo de los 6 años? Suponemos que la población de profesionales se distribuye normalmente.

sol

A continuación se describen las etapas de la prueba de hipótesis para la media:

Etapas 1: Las hipótesis son:  $H_0 : \mu = 6$  y  $H_1 : \mu < 6$

Los datos de la muestra nos hacen intuir que el tiempo medio de empleo es menor de 6 años, lo que marca la dirección de la prueba.

Etapas 2: La prueba es unilateral por la izquierda, y para el nivel de significación  $\alpha = 0,05$ , tenemos  $1 - \alpha = 0,95$ .

De  $p(Z \leq z_{1-\alpha}) = 0,95$ , mirando en las tablas de la  $N(0,1)$  obtenemos  $z_{1-\alpha} = 1,645$ , con lo cual el **valor crítico** es  $z_\alpha = -z_{1-\alpha} = -1,645$  que separa las zonas de aceptación y rechazo.

Etapas 3 y 4: En este caso el **estadístico de prueba** es  $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ , que sigue una normal

tipificada  $N(0,1)$ , y el **valor observado del estadístico de prueba** será el número

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{5 - 6}{4/\sqrt{64}} = -2, \text{ puesto que } \bar{x} = 5, \mu_0 = 6, \sigma = 4 \text{ y } n = 64.$$

Etapas 5: Como puede observarse, el **valor observado del estadístico de prueba**  $z_0 = -2$  es  $< z_{\alpha} = -1,645$ , que se encuentra en la zona de rechazo. Por tanto, tomamos la decisión de rechazar la hipótesis nula, es decir, aceptamos la hipótesis alternativa, por la cual el tiempo medio de los empleados está por debajo de los 6 años.

♦ ♦ La vida media de una muestra de 100 tubos fluorescentes producidos por una empresa es de 1570 horas, con una desviación típica de 120 horas. Si  $\mu$  es la vida media de todo lo producido en esa empresa, contrasta la hipótesis de que  $\mu = 1600$  horas contra la hipótesis alternativa  $\mu \neq 1600$  horas, utilizando un nivel de significación de 0,05.

sol

Etapas 1: Las hipótesis son:  $H_0: \mu = 1600$  horas y  $H_1: \mu \neq 1600$  horas, es decir, suponemos que la vida media de los tubos es de 1600 horas.

Etapas 2: La prueba es bilateral y para un nivel de significación  $\alpha = 0,05$ , con lo cual  $\alpha/2 = 0,05/2 = 0,025$ . De  $p(Z \leq z_{1-\alpha/2}) = 1 - \alpha/2 = 1 - 0,025 = 0,975$ , mirando en las tablas de la  $N(0,1)$  obtenemos  $z_{1-\alpha} = 1,96$ , con lo cual le corresponde por **valores críticos**  $z_{1-\alpha} = 1,96$  y  $z_{\alpha} = -z_{1-\alpha} = -1,96$  que separan las zonas de aceptación y de rechazo.

Etapas 3 y 4: En este caso el **estadístico de prueba** es  $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ , que sigue una normal

tipificada  $N(0,1)$ , y el **valor observado del estadístico de prueba** será el número

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{1570 - 1600}{120/\sqrt{100}} = -2,5, \text{ puesto que } \bar{x} = 1570, \mu_0 = 1600, \sigma = 120 \text{ y } n = 100.$$

Etapas 5: Como el **valor observado del estadístico de prueba**  $z_0 = -2,5$  es menor que el valor crítico  $z_{\alpha/2} = -1,96$ , vemos que se encuentra en la zona de rechazo. Por tanto, tomamos la decisión de rechazar la hipótesis nula  $\mu = 1600$  horas, es decir, aceptamos la hipótesis alternativa  $\mu \neq 1600$  horas.

♦ ♦ El salario medio correspondiente a una muestra de 600 personas de cierta población es de 1127 €. Se sabe que la desviación típica de los salarios en la población es 241 €. ¿Se puede afirmar, con un nivel de significación de 0,01, que el salario medio en dicha población es de 1145 €? ?

sol

Etapa 1: Las hipótesis son:  $H_0: \mu = 1145 \text{ €}$  y  $H_1: \mu \neq 1145 \text{ €}$ .

Etapa 2: La prueba es bilateral y para un nivel de significación  $\alpha = 0'01$ , con lo cual  $\alpha/2 = 0,051/2 = 0'005$ . De  $p(Z \leq z_{1-\alpha/2}) = 1 - \alpha/2 = 1 - 0'005 = 0'995$ , mirando en las tablas de la  $N(0,1)$  obtenemos  $z_{1-\alpha} = 2'58$ , con lo cual le corresponde por **valores críticos**  $z_{1-\alpha} = 2'58$  y  $z_{\alpha} = -z_{1-\alpha} = -2'58$ , que separan las zonas de aceptación y de rechazo.

Etapas 3 y 4: En este caso el **estadístico de prueba** es  $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ , que sigue una normal tipificada  $N(0,1)$ , y el **valor observado del estadístico de prueba** será el número

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{1127 - 1145}{241/\sqrt{600}} = -1'83, \text{ puesto que } \bar{x} = 1127, \mu_0 = 1145, \sigma = 241 \text{ y } n = 600.$$

Etapa 5: Como el **valor observado del estadístico de prueba**  $z_0 = -1'83$  es mayor que el valor crítico  $z_{\alpha/2} = -2'58$ , vemos que se encuentra en la zona de aceptación. Por tanto, tomamos la decisión de rechazar la hipótesis nula  $\mu = 1145 \text{ €}$ .

*Para el contraste de proporciones*

♦ ♦ Un experto en temas electorales, basado en los resultados de anteriores comicios, sostiene que, si se celebran elecciones generales en la actualidad, tan sólo acudiría a votar el 48% del electorado. Sin embargo, en un sondeo electoral realizado recientemente con una muestra de 1500 personas, 800 manifiestan su intención de votar. Plantea la prueba de hipótesis más adecuada, para un nivel de significación del 0,05 y comenta el resultado.

sol

Etapa 1: Las hipótesis son:  $H_0: p = 0,48$  y  $H_1: p > 0,48$

Para plantear la hipótesis nula nos basamos en la información previa. Sostenemos en ella que la población sigue manteniendo las mismas pautas de comportamiento electoral.

En la hipótesis alternativa queda reflejado el hecho de que la muestra actual nos dice que la situación ha cambiado, con una proporción mayor de votantes  $\hat{p} = 800/1500 = 0,53 > 0'48$ , lo que marca la dirección de la prueba.

Etapa 2: El nivel de significación es  $\alpha = 0'05$ , luego tenemos  $1 - \alpha = 0,95$ .

De  $p(Z \leq z_{1-\alpha}) = 1 - \alpha = 1 - 0'05 = 0'95$ , mirando en las tablas de la  $N(0,1)$  obtenemos  $z_{1-\alpha} = 1,645$ , con lo cual el **valor crítico** es  $z_{1-\alpha} = 1,645$  que separa las zonas de aceptación y rechazo.



Etapa 3 y 4: En este caso el **estadístico de prueba** es  $Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot (1-p_0)}{n}}}$ , que sigue una

normal tipificada  $N(0,1)$ , y el **valor observado del estadístico de prueba** será el número

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot (1-p_0)}{n}}} = \frac{0'53 - 0'48}{\sqrt{\frac{0'48 \cdot 0'52}{1500}}} = 3'85, \text{ puesto que } \hat{p} = 800/1500 = 0,53, p_0 = 0'48, q_0 = 0'52$$

y  $n = 1500$ .

Etapa 5: Como el **valor observado del estadístico de prueba**  $z_0 = 3'85$  es mayor que el **valor crítico**  $z_{\alpha/2} = 1,645$ , vemos que se encuentra en la zona de rechazo. Por tanto, tomamos la decisión de rechazar la aceptar hipótesis nula  $H_0: p = 0'48$ , y aceptamos la hipótesis alternativa  $H_1: p > 0'48$ . Con lo cual con una probabilidad de equivocarnos del 5% afirmamos que la tendencia actual de voto se ha incrementado respecto a comicios anteriores.

♦ ♦ Existe la hipótesis de que en la Comunidad Autónoma de Castilla y León realizan estudios de nivel medio el mismo número de varones que de mujeres. Mediante un sistema aleatorio tomamos una muestra de 1000 expedientes escolares, de los cuales 532 son varones y 468 son mujeres. ¿Es éste un resultado poco probable o, por el contrario, se ajusta a la gran mayoría del 99% de los resultados?

sol

Etapa 1: Las hipótesis son:  $H_0: p = 0,5$  y  $H_1: p > 0,5$

En el planteamiento se ha tomado como referencia para la hipótesis alternativa la proporción de varones en la muestra, que es  $\hat{p} = 800/1500 = 0,532 > 0,5$ .

Etapa 2: El nivel de significación es  $\alpha = 0'01$ , luego tenemos  $1 - \alpha = 0,99$ .

De  $p(Z \leq z_{1-\alpha}) = 1 - \alpha = 0'99$ , mirando en las tablas de la  $N(0,1)$  obtenemos  $z_{1-\alpha} = 2'33$ , con lo cual el **valor crítico** es  $z_{1-\alpha} = 2'33$  que separa las zonas de aceptación y rechazo.

Etapa 3 y 4: En este caso el **estadístico de prueba** es  $Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot (1-p_0)}{n}}}$ , que sigue una

normal tipificada  $N(0,1)$ , y el **valor observado del estadístico de prueba** será el número

$$z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot (1-p_0)}{n}}} = \frac{0'53 - 0'48}{\sqrt{\frac{0'48 \cdot 0'52}{1500}}} = 2,025, \text{ puesto que } \hat{p} = 532/1000 = 0'532, p_0 = 0'5, q_0 = 0'5$$

y  $n = 1000$ .

**Etapa 5:** Como el **valor observado del estadístico de prueba**  $z_0 = 2'025$  es menor que el **valor crítico**  $z_{\alpha/2} = 2,33$ , vemos que se encuentra en la zona de aceptación. Por tanto, tomamos la decisión de aceptar hipótesis nula  $H_0: p = 0'5$ , es decir no podemos rechazar la hipótesis de que estudien el mismo número de varones que de mujeres.

*Para el contraste de diferencia de medias*

◆ ◆ Estudiamos dos muestras de ciudadanos de la Comunidad Autónoma de Andalucía (A y B), de 80 miembros cada una para conocer el sentimiento nacionalista. Sobre una escala de 1 a 10, la primera alcanzó una media de 7,2 con una desviación típica de 3,1 mientras que la segunda alcanzó una media de 8,1 con una desviación típica de 4,2. Nuestra hipótesis de investigación es que el colectivo B tiene un sentimiento nacionalista mayor que el colectivo A. Comprueba la hipótesis para un nivel de significación del 0,01.

sol

**Etapa 1:** Las hipótesis son:  $H_0: \mu_1 = \mu_2$  y  $H_1: \mu_1 < \mu_2$

**Etapa 2:** La prueba es unilateral por la izquierda, y para el nivel de significación  $\alpha = 0,01$ , tenemos  $1 - \alpha = 0,99$ .

De  $p(Z \leq z_{1-\alpha}) = 0'99$ , mirando en las tablas de la  $N(0,1)$  obtenemos  $z_{1-\alpha} = 1,645$ , con lo cual el **valor crítico** es  $z_\alpha = -z_{1-\alpha} = -2,33$  que separa las zonas de aceptación y rechazo.

**Etapa 3 y 4:** En este caso el **estadístico de prueba** es  $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ , que sigue

una normal tipificada  $N(0,1)$ , y el **valor observado del estadístico de prueba** será el

número  $z_0 = \frac{(7'2 - 8'1) - (0)}{\sqrt{\frac{3'1^2}{80} + \frac{4'2^2}{80}}} = -1'55$ , puesto que  $\bar{x}_1 - \bar{x}_2 = 7,2 - 8,1 = -0,9$ ,  $\mu_1 - \mu_0 = 0$ ,  $\sigma_1 = 3'1$ ,

$\sigma_2 = 4'2$  y  $n_1 = n_2 = 80$ .

**Etapa 5:** Como el **valor observado del estadístico de prueba**  $z_0 = -1'55$  es mayor que el **valor crítico**  $z_{\alpha/2} = -2'33$ , vemos que se encuentra en la zona de aceptación, luego se acepta la hipótesis nula y se rechaza la hipótesis alternativa.

No existen diferencias significativas en cuanto al sentimiento nacionalista en los dos colectivos.