INTRODUCCION A LA TEORIA DE LA ESTIMACION Y A LA TEORIA DE LA DECISION ESTADISTICA

JOSEP M. DOMÉNECH I MASSONS
Departamento de Psicología
Universidad de Barcelona

I. Introducción

¿Cree que la estadística es saber que durante el pasado mes 253 personas han muerto de infarto o que en España hay 35.377. 643 habitantes?

Estas informaciones numéricas, recuento de determinados acontecimientos, generalmente proporcionados por la administración de un país, reciben el nombre de *estadísticas* y no tienen nada que ver con la *estadística* que es algo muy diferente.

Analice el siguiente estudio:

El tratamiento actual A de cierta enfermedad E comporta un 50 % de curaciones. Con un nuevo tratamiento N, aplicado a un grupo experimental de 10 enfermos se observan 7 curaciones, lo que representa el 70 % de los enfermos tratados.

¿Cree que N es mejor que A?

Quién no conoce el método de razonamiento estadístico concluirá afirmativamente. Quien lo conoce dirá que es necesario aplicar una *prueba estadística* para contestar a esta pregunta.

En efecto, un análisis más detallado del anterior estudio releva los elementos siguientes:

1. Existen personas afectadas de la enfermedad E. El conjunto de todos estos enfermos —llamados en lenguaje estadístico *individuos*— recibe el nombre de *población* de los enfermos E.
2. En esta población, un 50 % de los individuos pueden ser curados mediante el tratamiento A.
3. Se ha realizado una experiencia con una parte de la población —llamada en lenguaje estadístico *muestra*— formada por 10 enfermos.
4. En esta muestra, un 70 % de los individuos tratados se curan mediante el tratamiento N.

A la vista de estos elementos, aceptar la superioridad del tratamiento N, como consecuencia del resultado observado en esta experiencia, podría ser un grave error.

Supongamos, no obstante, que se ha adoptado el tratamiento N y que una vez aplicado a toda la población de enfermos E, se han obtenido un 40 % de curaciones.

¿Cree que este resultado está en contradicción con el resultado de la experiencia realizada?

En absoluto.

El simil de la urna aclarará este aspecto fundamental.

Representemos la población de enfermos E mediante una urna con un 40 % de bolas blancas —que simbolizan los enfermos que se curan mediante el tratamiento N— y un 60 % de bolas negras —que simbolizan los enfermos de la población no sensibles al tratamiento N.

El grupo de 10 enfermos, objeto de la experiencia, equivale a una muestra extraída al azar de la urna.

¿Le parece extraño hallar en la muestra de 10 bolas, 7 de color blanco?

Aunque lo más probable es encontrar la misma proporción de bolas blancas que contiene la población (4 blancas), en una muestra tan pequeña es posible que el *azar* haga aparecer 7 bolas blancas.

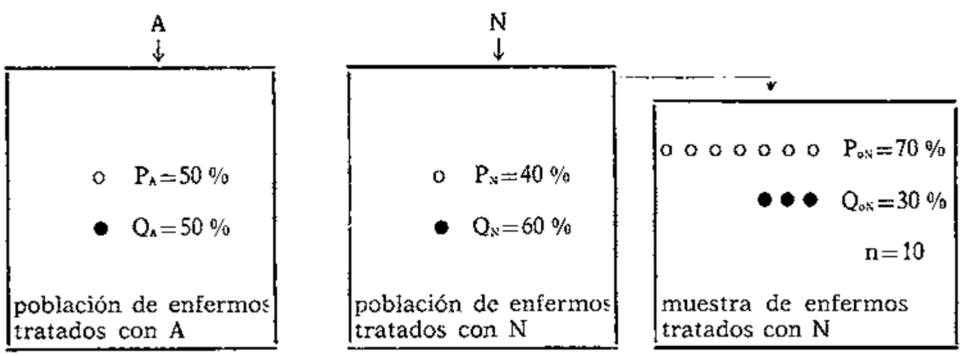
Por lo tanto no es extraño que el azar haya incluido 7 enfermos sensibles al tratamiento N en el grupo de 10 individuos objeto de la experimentación.

La figura 0 ilustra este estudio.

Ahora nos apercibimos claramente del error cometido por quienes han afirmado que el tratamiento *N* era superior al tratamiento *A*. Se han confundido la *proporción de curaciones observadas en una muestra* con la *proporción de curaciones que se obtendría aplicando el tratamiento N a toda la población* de enfermos *E*, proporción que mide la eficacia real de dicho tratamiento.

La proporción $P_{oN}=70$ % de curaciones observada en la muestra, es una *estimación* de la proporción $P_N=40$ % de curaciones que el tratamiento *N* da en la población general.

FIG. 0



El tratamiento A es mejor que el tratamiento N, ya que P_A es superior a P_N.
No debe confundirse la proporción P_N de curaciones obtenidas al aplicar el tratamiento a toda la *población* con la proporción observada P_{oN} de curaciones obtenidas al aplicar el tratamiento N a una *muestra* de individuos.

Esta estimación está afectada por el azar. Es importante recordar que la influencia del azar será tanto más importante cuanto menor sea el número de individuos de la muestra.

En lenguaje estadístico diremos que la *precisión de la estimación* aumenta con el tamaño de la muestra.

En la práctica, generalmente sólo se conoce lo que ocurre en una muestra de enfermos. La teoría estadística de la estimación permite predecir, con determinada precisión, lo que ocurre en la población mediante la información aportada por los individuos de una muestra extraída al azar de dicha población.

Por lo tanto, volviendo al ejemplo expuesto, siempre que daban darse conclusiones a partir de los resultados de una experiencia basada en el estudio de grupos de individuos, será necesario analizar si, por si solo, el azar puede explicar las diferencias encontradas.

Este análisis se hace siempre aplicando *pruebas estadísticas* que cifran, mediante probabilidades, la influencia del azar en los resultados de un experimento.

Estas pruebas, que se aplican para comparar entre sí los resultados de diversos tratamientos aplicados a grupos de individuos y que permiten decidir si las diferencias observadas en ellos pueden ser solamente explicadas por la influencia del azar o si son debidas a las diferentes eficacias de los tratamientos aplicados, están basadas en otra importante parte de la Estadística llamada *teoría estadística de la decisión.*

No es la uniformidad sino la *variabilidad* el atributo fundamental que caracteriza a la naturaleza.

No es la rigidez sino el *azar* quien está en la base de los hechos psicológicos o biológicos.

En la naturaleza, el azar hace que la variabilidad entre los individuos sea la regla. La estadística es el único instrumento matemático adecuado para analizar datos de fenómenos cuya característica fundamental es la variabilidad.

No es posible estudiar la variabilidad con una sola observación. De ahí que la *formulación estadística de un problema* no pueda ser una formulación a nievl individual sino que *debe ser una formulación a nivel de grupo.*

Los *diseños experimentales estadísticos* son experiencias formuladas a nivel de grupo. En consecuencia, una vez tratados estadísticamente los datos obtenidos, las *conclusiones de estos diseños* no pueden ser conclusiones a nivel individual sino que *son conclusiones a nivel de grupo.* Esto quiere decir, por ejemplo, que si con un diseño experimental se demuestra la superioridad de un tratamiento X respecto a otro tratamiento Y, el resultado que obtendremos al aplicar el tratamiento X a toda la población será, *en media,* superior al resultado que hubiéramos obtenido al aplicar el tratamiento Y a toda la población; sin embargo, para ciertos individuos el tratamiento Y podrá ser superior al X.

En este trabajo se desarrollan dos problemas fundamentales que estudia la estadística: los *Problemas de Estimación* y los *Problemas de Decisión.*

Con fines didácticos este estudio se particulariza a caracteres cualitativos. La *Teoría de la Estimación* se desarrolla a partir del caso particular de una proporción. La *Teoría de la Decisión* se desarrolla a partir del caso particular de la prueba de comparación de una proporción observada a una proporción teórica. Sin embargo, la generalización a otros casos es inmediata.

II. Concepto de proporción

Sea un cierto aspecto A que se observa n_A veces en una muestra de n individuos. Se define como *proporción observada* de individuos con el aspecto A presente, al cociente:

(1) $$p_o = n_A / n$$

A la proporción observada de individuos de la muestra que no presentan el aspecto A, se la designa por q_o y vale:

(2) $$q_o = 1 - p_o$$

A la verdadera proporción de individuos, de la población origen de la muestra, que presentan el aspecto A, se la designa por p y a su complemento a 1 se lo designa por q.

III. Concepto de intervalo de probabilidad $1-\alpha$

Tanto para la Teoría de la Estimación como para la Teoría de la Decisión es necesario resolver el siguiente e importante *problema de predicción* mediante la Teoría de la Probabilidad: conocer el intervalo donde fluctúan las proporciones p_o observadas en muestras de tamaño n procedentes de una población caracterizada por una proporción p.

El siguiente experimento permite resolver dicho problema.

Se tiene una población que contiene una proporción p conocida de bolas blancas y el resto $q = 1 - p$ son bolas negras.

Se extrae un conjunto muy grande de muestras con n_1 bolas cada una (muestras de tamaño n_1) y se tabulan y representan gráficamente el conjunto de las proporciones p_o de bolas blancas observadas en cada una de estas muestras.

Nuevamente se extrae un conjunto muy grande de muestras, ahora con n_2 bolas cada una (n_2 mayor que n_1) y se tabulan y representan gráficamente el conjunto de proporciones de bolas blancas p_o observadas en cada una de estas muestras.

Y así sucesivamente se extraen conjuntos muy grandes de muestras pero cada vez compuestas de un número n de bolas mayor que los anteriores y se representan gráficamente el conjunto de proporciones de bolas blancas p_o observadas en cada conjunto de muestras.

La figura 1 ilustra estas representaciones gráficas para conjuntos de muestras de tamaño $n_1=10$, $n_2=20$ y $n_3=40$, cuando en la población origen de estas muestras hay la mitad de bolas blancas y la otra mitad de bolas negras ($p=q=0.5$).

La observación de estas representaciones gráficas permite sacar las siguientes conclusiones:

a) En todas las gráficas las proporciones observadas p_o se agrupan alrededor del valor p (proporción de bolas blancas que contiene la población origen de las muestras).

b) Cuando *aumenta* el número n de bolas que hay en las muestras, las proporciones observadas p_o están *más agrupadas* alrededor del valor p. El conjunto de proporciones observadas p_o es cada vez menos disperso a medida que el tamaño n de las muestras aumenta.

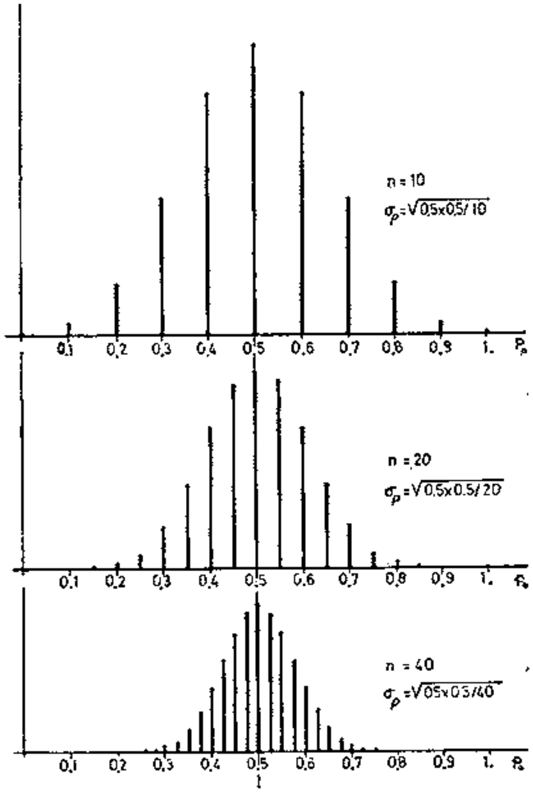


figura 1

Fluctuación de las proporciones p_0 de bolas blancas observadas en muestras de tamaño n extraídas al azar de una población que contiene una proporción de bolas blancas $p = 0.5$

Este concepto de dispersión se cuantifica mediante el índice σ_p:

(3) $$\sigma_p=\sqrt{pq/n}$$

llamado *desviación tipo.*

Por ejemplo, si las muestras proceden de una población que contiene una proporción $p=0.5$ de bolas blancas y el resto $q=1-p=0.5$ de bolas negras, cuando $n=10$ la desviación tipo de las proporciones observadas vale:

$$\sigma_p=\sqrt{0.5\times 0.5/10}=0.16$$

c) Si las muestras son grandes, se ve cómo estas gráficas son *simétricas* alrededor de un eje vertical que pasa por el valor p y tienen la misma *forma de «campana»*. La diferencia entre ellas está en que las «campanas» son más estrechas a medida que corresponden a muestras más grandes.

Este único *modelo* que siguen todas las gráficas recibe el nombre de *modelo o ley normal.* Esta ley ha sido estudiada por Laplace y Gauss, hace varios siglos.

Teorema fundamental

Se demuestra que el *desvío e*, a cada lado del punto central p, que deja dentro del 95 % de las proporciones p_o observadas en las muestras, se obtiene multiplicando el valor constante $z\alpha=1.96$ por la desviación tipo σ_p correspondiente. Es decir:

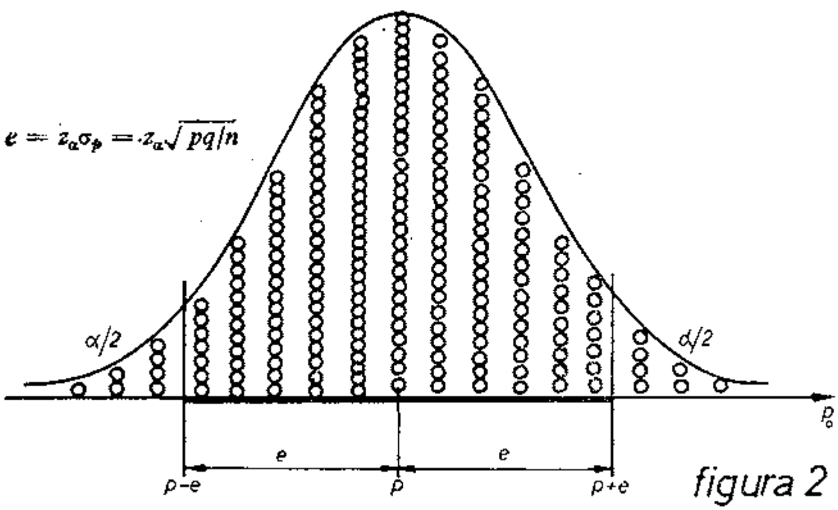


figura 2

$$e = 1.96\sqrt{pq/n}$$

El *desvío* e que deja dentro el 99 % de las p_o se obtiene multiplicando $z\alpha = 2.58$ por dicha desviación tipo.

En general estos valores constantes se simbolizan por $z\alpha$ y el desvío e viene dado por:

$$e = z\alpha\sqrt{pq/n} \tag{4}$$

En la tabla 1 se dan los valores más importantes de $z\alpha$ en función de α y en la figura 2 se ilustra el desvío e.

El valor α representa la proporción, en tantos por uno, de proporciones observadas que están fuera del intervalo alrededor de p, definido por dicho desvío e.

Condiciones de aplicación

La fórmula (4) y los valores $z\alpha$ dados en la tabla 1, son válidos exclusivamente para distribuciones que siguen el *modelo normal.*

Se comprueba experimentalmente que la distribución de las proporciones observadas sigue un modelo normal cuando se cumple que los productos np y nq son iguales o superiores a 5. Estas muestras reciben el nombre de *muestras grandes.*

Este *Teorema Fundamental* se demuestra a partir de la *Teoría de la Probabilidad* (*) y es un caso particular de un importante teorema llamado Teorema Límite Central.

DEFINICIÓN DE INTERVALO DE PROBABILIDAD

Se define como *intervalo de probabilidad* todo intervalo, simétrico alrededor de p, que contenga gran parte de las proporciones observadas en muestras de tamaño n extraídas al azar de una población caracterizada por la proporción p.

El intervalo alrededor de p, definido por el desvío e, que contiene el 95 % de las proporciones observadas, recibe el nombre de intervalo de probabilidad del 95 %. En general, recibe el nombre de *intervalo de probabilidad* $1-\alpha$, aquel intervalo que contiene el $1-\alpha$ (en tantos por uno) de las proporciones observadas p_o, es decir, aquel intervalo que sólo deja fuera una pequeña proporción α de las proporciones observadas p_o.

(*) Un estudio detallado puede verse en BONET, E.: *Fonaments d'estadística.* Teide. Barcelona, 1974.

En el *intervalo de probabilidad* $1-\alpha$, el valor α indica la probabilidad de observar, en una muestra procedente de la población p, una proporción observada p_o fuera de dicho intervalo.

Una idea importante es que un intervalo de probabilidad nunca podrá contener todas las proporciones que pueden observarse en las muestras. Para definir el intervalo de probabilidad siempre debe excluirse una proporción α, tan pequeña como queramos, de casos extremos.

La estadística analiza fenómenos cuya regla fundamental es la *variabilidad*. Si queremos dar un intervalo en el que se encuentren los valores de un cierto carácter, medido en una población de individuos, la variabilidad puede hacer que en algún individuo este carácter tome un valor excepcionalmente grande o excepcionalmente pequeño.

La estadística busca un intervalo que sea lo más pequeño posible pero que a la vez contenga la máxima información posible (el máximo número de casos posible). Para lograrlo se debe prescindir de unos pocos casos extremos que agrandan mucho el intervalo sin casi aportar información.

Por convenio se ha aceptado universalmente dar intervalos que contengan el 95 % o el 99 % de los individuos.

De esta definición y de la observación de la figura 2 se deduce fácilmente la fórmula del intervalo de probabilidad:

$$p_o \in p \pm z\alpha\sqrt{pq/n} \qquad (5)$$

C. de A.: Las muestras deben ser grandes (nq y $nq \geqslant 5$).

El intervalo de probabilidad $1-\alpha$ permite *predecir*, aceptando el riesgo α de equivocarse, el intervalo en el cual estará contenida la proporción p_o observada en una cierta muestra de tamaño n procedente de una población caracterizada por la proporción p.

Generalización

El concepto de intervalo de probabilidad se generaliza fácilmente a otras características estadísticas. Por ejemplo si se estudia un *carácter cuantitativo*, el intervalo en el cual estará contenida la media $\bar{x}$ observada en una muestra de tamaño n, procedente de una población caracterizada por una media m y una desviación tipo σ, viene dado por (6).

$$\bar{x} \in m \pm z\alpha\sqrt{\sigma^2/n} \qquad (6)$$

C. de A.: Las muestras deben ser grandes ($n \geqslant 30$).

Ejemplo 1

Dada una población que contiene un 50 % de bolas blancas y un 50 % de bolas negras, predecir el intervalo en el que se encontrará el 95 % de las proporciones observadas en muestras de 40 bolas extraídas, no exhaustivamente y al azar, de dicha población.

Aplicando la fórmula (5) se obtiene el intervalo pedido:

$$0.50 \pm 1.96\sqrt{0.50x(1-0.50)/40} = 0.50 \pm 0.16$$

$$p_o \in (0.34 \div 0.66)$$

IV. Teoría de la estimación estadística

Un *problema de estimación* consiste en hallar, con determinada precisión, el verdadero valor de un parámetro (proporción, media, variancia, etc. calculada con todos los individuos de una población), a partir solamente de la información contenida en una muestra representativa (*) de la población.

En el caso particular de estimar una proporción p, si sólo se dispone de la información contenida en una muestra de tamaño n, es decir de la proporción p_o observada en ella, se demuestra que la mejor *estimación puntual de p es la proporción p_o observada en la muestra.*

Además de esta estimación puntual, es importante conocer su *precisión*, es decir lo próximo o lo alejado que p_o puede estar de p. Este concepto de precisión se explicita con la *estimación por intervalo de p o intervalo de confianza, que es un intervalo simétrico alrededor de p_o que tiene gran probabilidad de contener el valor p.*

El intervalo de confianza se deduce a partir del intervalo de probabilidad.

Si por un momento imaginamos conocida la proporción p de la población, al ir extrayendo de ella muestras de tamaño n las proporciones p_o observadas en ellas, excepto una pequeña cantidad α de ellas, irán cayendo dentro del intervalo de probabilidad.

Salvo la pequeña cantidad α de ellas, las proporciones observadas p_o no estarán más alejadas de p que la distancia definida por el desvío e (figura 3):

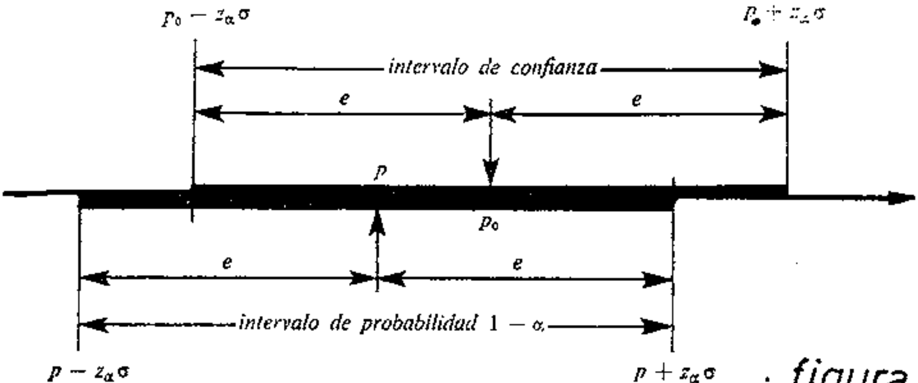
(7) $$p_o \in p \pm e = p \pm z\alpha\sqrt{pq/n}$$

Volvamos al problema real, en el que sólo se conoce p_o. Si esta proporción p_o ha caído dentro del intervalo de probabilidad no estará más alejada de p

(*) La obtención de una muestra representativa es un problema de naturaleza estadística cuya solución correcta es que la elección de los individuos de la muestra excluya cualquier sistemática. Esto se logra mediante una *elección al azar*.

que la distancia definida por el desvío e. Es decir el valor desconocido p estará en algún punto del intervalo (figura 3):

(8) $$p \in p_o \pm e = p_o \pm z\alpha \sqrt{pq/n}$$



. figura 3

La probabilidad de que el intervalo (8) no contenga el valor p coincide con la probabilidad de que p_o no esté dentro del intervalo de probabilidad. Esta probabilidad vale α.

El intervalo (8), que tiene gran probabilidad de contener p, aún no puede calcularse ya que para hallar la desviación tipo σR, debe conocerse p.

Se demuestra cuando la muestra es grande que el *parámetro* σ_p, desviación tipo teórica, puede ser estimado mediante el *estadístico* s_p, que es la desviación tipo calculada a partir de los datos observados en la muestra:

(9) $$s_p = \sqrt{p_o q_o / n}$$

Sustituyendo este valor en (8) se obtiene el *intervalo de confianza:*

(10) $$p \in p_o \pm z\alpha \sqrt{p_o q_o / n}$$

C. de A.: La muestra debe ser grande (np y $nq \geqslant 5$).

Intervalo de Probabilidad e Intervalo de Confianza

La figura 4 muestra cómo estos dos conceptos son totalmente diferentes.

El intervalo de probabilidad es la base teórica que permite deducir el intervalo de confianza. Ambos intervalos son de naturaleza radicalmente diferente. En efecto, el *intervalo de probabilidad* parte del conocimiento de la proporción p que caracteriza a la población y resuelve un *problema de predicción* al indicar dónde estarán situadas las proporciones p_o observadas en muestras de tamaño n procedentes de dicha población. El *intervalo de confianza* parte únicamente del conocimiento de la proporción p_o observada en una muestra de tamaño n y resuelve un *problema de estimación* al dar el intervalo que tiene gran probabilidad de contener la verdadera proporción p (desconocida) que caracteriza a la población origen de dicha muestra.

Dado un cierto tamaño de muestra n y un valor α, existe un único intervalo de probabilidad mientras que pueden haber infinitos intervalos de confianza, uno para cada muestra que se tenga.

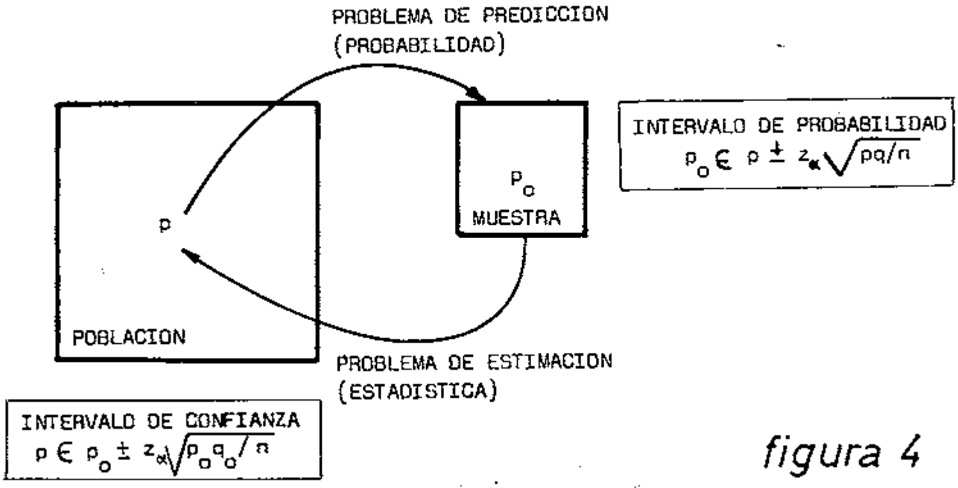


figura 4

Tamaño de la muestra

La fórmula de cálculo del intervalo de confianza indica cómo a medida que aumenta el tamaño n de la muestra, este intervalo es más estrecho, es decir la estimación de la proporción p a partir de la proporción p_o es más *precisa.*

¿Cuántos individuos deberemos elegir para estimar la proporción p? Se trata de fijar *a priori* la precisión e deseada y a partir de ella calcular el valor de n para que la proporción observada en la muestra como máximo esté a una distancia $\pm e$ de la verdadera proporción p.

Despejando el valor n en la fórmula (4) se obtiene n en función de la precisión e:

$$n = z_\alpha pq/e \tag{11}$$

El valor p es desconocido ya que precisamente se realiza el sondeo para estimarlo. En la fórmula (11) debe sustituirse p por un valor aproximado, obtenido según uno de los siguientes criterios:

a) Estudios anteriores proporcionan una idea aproximada del valor de p.

b) Realizar un sondeo previo, extrayendo una pequeña muestra, y sustituir p por la proporción p_o observada en ella.

c) Sustituir el producto pq de la fórmula (11) por el caso más desfavorable. El valor $pq=0.25$ es el mayor de todos los posibles.

Estrategia global

La determinación del tamaño de la muestra da una nueva perspectiva a los *problemas de estimación*. En efecto, es más lógico fijar la precisión e deseada que partir de un tamaño de muestra arbitrario.

El proceso metodológico óptimo para *estimar una proporción* es el siguiente:

1. Fijar la precisión e deseada y el riesgo α.
2. Determinar el tamaño n de la muestra:

$$n=z_\alpha pq/e$$

3. Realizar un sondeo para extraer una muestra al azar de tamaño n.

4. Calcular la proporción observada en dicha muestra (estimación puntual de p):

$$p_o=n_A/n$$

5. Calcular el intervalo de confianza (estimación por intervalo de p):

$$p_o \pm z\alpha\sqrt{p_o q_o/n}$$

Aunque éste es el proceso óptimo, no siempre es posible seguirlo. En efecto, si sólo se dispone de una información limitada, por ejemplo de $n=87$ individuos, y no hay posibilidad de obtener más información, la estimación se realizará con la máxima información disponible, es decir con los $n=87$ individuos, pasando directamente al punto 4.

Generalización

Cuando la característica estadística a estimar no es una proporción, cambian las tres fórmulas anteriores, pero el proceso anterior permanece invariable.

Ejemplo 2

El resultado de un primer sondeo de opinión sitúa entre un 70 % y un 83 % el porcentaje de votos SI en un referéndum.

a) ¿Cuántos individuos deberá contener el próximo sondeo, para que la proporción de votos afirmativos, dados por los individuos de la muestra, no esté más alejada de un ± 2 % de la verdadera proporción? Se acepta un riesgo $\alpha = 0.05$.

Se aplica la fórmula (11). Por estar el valor de p comprendido entre 0.70 y 0.83, el caso más desfavorable corresponde a $p = 0.70$.

$$n = 1.96^2 \times 0.70 \times 0.30 / 0.02^2 = 2.017 \text{ individuos}$$

b) Una vez realizado el sondeo al azar con 2.017 individuos, se ha observado que 1.613 votarán afirmativamente. Estimar la proporción de votos SI que dará toda la población.

La estimación puntual es:

$$p_o = 1613/2017 = 0.80$$

La estimación por intervalo es:

$$0.80 \pm 1.96 \sqrt{0.80 \times 0.20 / 2017} = 0.80 \pm 0.017$$

$$p_o \in (0.783 \div 0.817)$$

V. Teoría de la decisión estadística

Un *problema de decisión* consiste en elegir entre dos hipótesis complementarias H_0 y H_1 la más verosímil de ellas, a partir de resultados observados en muestras. Al elegir la hipótesis H_1 existe el *riesgo* α de tomar una decisión equivocada. Al elegir la hipótesis H_0 existe el *riesgo* β de tomar una decisión equivocada.

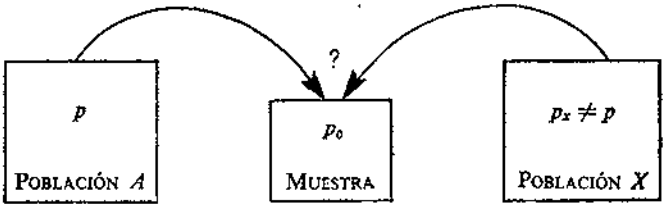
Sea una muestra que contiene n bolas, de las cuales una proporción p_o de ellas son blancas.

Se desea saber si la muestra procede o no procede de dicha población.

Hipótesis nula e hipótesis alternativa

La figura 5 esquematiza el planteo general de este problema de decisión.

Hay dos hipótesis posibles. La hipótesis principal es la *hipótesis nula* H_o, según la cual la muestra procede de la población *A*. En este caso la diferencia existente entre las proporciones p_o y p ha sido producida por la influencia del azar. Cuando la hipótesis nula es falsa significa que la muestra procede de una población *X* caracterizada por una proporción p_x desconocida, diferente de la proporción p. Esta hipótesis, complementaria a la anterior, recibe el nombre de *hipótesis alternativa.*



Hipótesis nula (H_0):
La muestra procede
de la población A.

Hipótesis alternativa (H_1):
La muestra procede de una
población X de diferente
composición que la población A.

figura 5

Prueba de decisión

La prueba es muy sencilla y está ilustrada en la figura 6. Las pruebas de decisión estadística se construyen a partir de la hipótesis nula. Si la muestra procede de la población A, la proporción p_o observada en ella estará dentro del intervalo de probabilidad, es decir dentro del intervalo en el que se encuentra la mayor parte de las proporciones observadas en muestras de tamaño n procedentes de la población p.

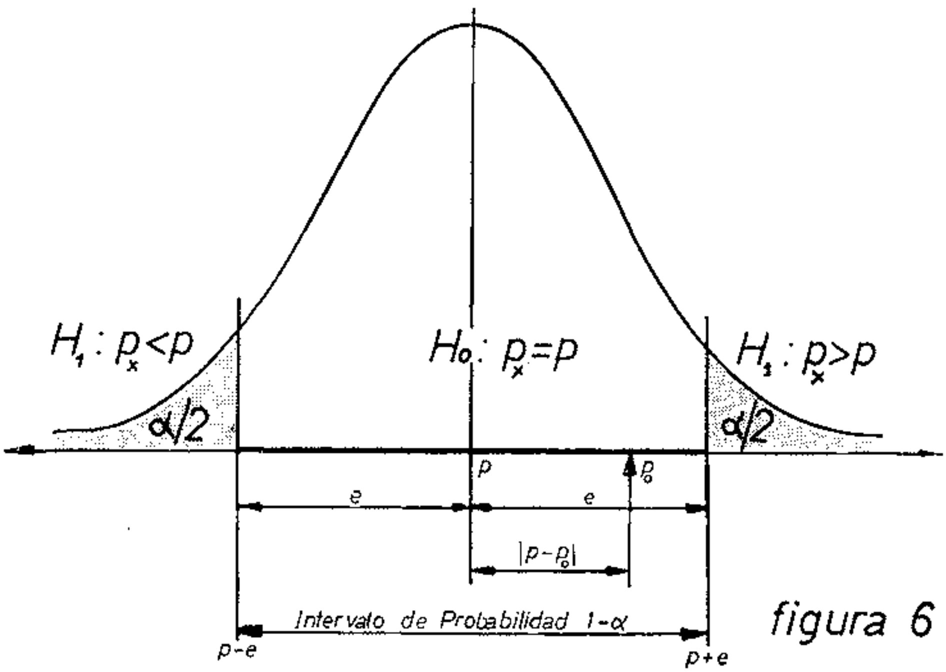


figura 6

Cuando la proporción observada p_o está fuera del intervalo de probabilidad, lo más verosímil es que la muestra proceda de una población p_x diferente de la población p. En este caso la decisión a tomar es la de quedarse con la hipótesis alternativa.

Ver si p_o está dentro o fuera del intervalo de probabilidad equivale a ver si la diferencia, en valor absoluto, entre p y p_o es menor o mayor que el desvío es:

(12)
$$|p_o - p| \begin{matrix} \leq \\ > \end{matrix} z\alpha\sqrt{pq/n} \begin{matrix} \to H_0 \\ \to H_1 \end{matrix}$$

Dividiendo ambos miembros por la raíz cuadrada:

(13)
$$|p_o - p| \,/\, \sqrt{pq/n} \begin{matrix} \leq \\ > \end{matrix} z\alpha \begin{matrix} \to H_0 \\ \to H_1 \end{matrix}$$

Si llamamos z al primer miembro, se obtiene una prueba de decisión estadística llamada *prueba de comparación de una proporción observada a una proporción teórica.* Esta prueba, por estar basada en el intervalo de probabilidad, será válida para muestras grandes.

(14)

$$z = |p - p_0| \,/\, \sqrt{pq/n}$$

$z \leqslant z\alpha$: Nada se opone en aceptar la hipótesis nula.

$z > z\alpha$: Rechazo de la hipótesis nula con un riesgo α.

C. de A.: La prueba es sólo válida para muestras grandes (nq y nq mayores o iguales a 5).

Nótese que una prueba cuando concluye aceptar H_1, afirma sólo que p_x es *diferente* de p.

Cuando la *conclusión estadística* es rechazar H_0 con riesgo α, y sólo en este caso, le corresponde la siguiente *conclusión experimental:*

Si $p_o > p$ podemos afirmar con riesgo α que la muestra procede de una población p_x mayor que p.

Si $p_o < p$ podemos afirmar con riesgo α que la muestra procede de una población p_x menor que p.

Para saber en qué cantidad p_x es mayor o menor que p, es necesario conocer el valor p_x. Esto es un problema de naturaleza diferente pero que ya ha sido estudiado: basta estimar p_x a partir de p_o. Se trata de un *problema de estimación* que se resuelve mediante el intervalo de confianza.

Los riesgos asociados a una decisión estadística

Hay sólo dos posibilidades de equivocarse al tomar la decisión:

1. Puede ser que la muestra proceda de la población *A*. En este caso p_o oscila alrededor de p; sin embargo, existe una probabilidad α de que p_o caiga fuera del intervalo de probabilidad. Cuando este caso se presenta, se toma la decisión de rechazar la hipótesis nula y se comete un error.

Este error recibe el nombre de *error tipo I*. La probabilidad de cometer un error tipo I se llama *riesgo* α o *riesgo de primera especie.*

2. Puede ser que la muestra proceda de la población *X*. En este caso p_o oscila alrededor de p_x, sin embargo, puede darse el caso de que p_o, debido a la influencia del azar, caiga dentro del intervalo de probabilidad definido alrededor de p. Cuando este caso se presente, se toma la decisión de no rechazar la hipótesis nula y se cometa un error.

Este error recibe el nombre de *error tipo II*. La probabilidad de cometer un error tipo II se llama *riesgo* β o *riesgo de segunda especie.*

El riesgo α es conocido y se fija *a priori*. En la prueba anterior la probabilidad de que una proporción observada se encuentre fuera del intervalo de probabilidad (zona de aceptación de H_1) procediendo de la población p (H_0 verdadera) viene dada por el valor α asociado al intervalo de probabilidad. Esta probabilidad es la que se define como riesgo α o riesgo de primera especie. De ahí la frase *se rechaza la hipótesis nula con riesgo* α.

Las frases *se rechaza* H_0 *al nivel de significación* α o *se rechaza* H_0 *al nivel de confianza* $1-\alpha$, son equivalentes a la anterior y expresan el mismo concepto de riesgo α.

El riesgo β es siempre desconocido. El hecho de observar una proporción p_o dentro del intervalo de probabilidad, indica un resultado que no está en contradicción con la hipótesis nula, pero en ningún caso prueba que H_0 sea verdadera.

En efecto, es posible que, debido a la influencia del azar, una proporción p_o observada es una muestra procedente de la población p_x, caiga dentro del intervalo de probabilidad de p (zona de no rechazo de H_o). Esto ocurre más a menudo cuando p_x tiene un valor próximo a p. La probabilidad de que ocurra, llamada riesgo β o riesgo de segunda especie, nunca se conoce ya que la proporción p_x es siempre desconocida.

Una frase equivalente a la de *nada se opone en aceptar la hipótesis nula* es *la diferencia encontrada no es significativa.*

Un estudio detallado del riesgo β se hace en uno de los siguientes apartados.

Se define como *potencia* de una prueba de decisión estadística el valor $1-\beta$.

Si el riesgo β indica la probabilidad de equivocarse al aceptar H_o, la *potencia* es la probabilidad complementaria al riesgo β. Es decir, *la potencia caracteriza la capacidad que tiene una prueba de decisión estadística de no equivocarse al aceptar* H_o.

Diferencia entre riesgo α, nivel de significación, nivel de confianza y grado de signiRcación.

El concepto de riesgo α ha quedado definido a nivel teórico. Sin embargo, la noción de riesgo de error es más compleja de lo que parece y se presta a muchas interpretaciones erróneas (*).

La prueba estadística descrita no está basada en calcular la probabilidad de obtener el valor p_o cuando la muestra procede de la población A (hipótesis nula). Dicha prueba estadística se basa en ver si p_o *pertenece o no al conjunto, fijado a priori, de las proporciones muy alejadas de* p. Este conjunto de proporciones alejadas de p, queda definido mediante el nivel de significación. Si se elige el nivel de significación del 5 %, el conjunto de proporciones alejadas de p, que se obtiene a partir del intervalo de probabilidad y que define la zona de rechazo de la hipótesis nula, será el *conjunto que tendrá sólo una probabilidad igual a 0,05 de que la proporción* p_o, observada en una muestra extraída al azar de la población A, *pertenezca a él.*

Se define *nivel de significación* como el valor de α a partir del cual una diferencia se considera significativa.

El nivel de significación es un valor arbitrario, pero existe el convenio, entre gran número de estadísticos, de considerar una diferencia significativa cuando el riesgo de error α es igual ó inferior a 0,05.

Si α es el valor del nivel de significación, el valor $1\text{-}\alpha$ recibe el nombre de *nivel de confianza.*

Desde un *ángulo teórico* se puede afirmar que utilizando un nivel de significación igual a 0,05, al extraer una serie de muestras de la población A (hipótesis nula verdadera) *5 veces de cada 100* nos equivocaremos en el sentido de afirmar que la muestra no procede de la población A. En este caso el riesgo de error vale exactamente $\alpha = 0{,}05$.

Desde un *ángulo práctico* se puede afirmar que un investigador utilizando un nivel de significación igual a 0,05, al realizar una serie de experiencias y analizarlas estadísticamente, generalmente se equivocará *menos de 5 veces de cada 100* en el sentido de rechazar la hipótesis nula siendo verdadera. El motivo de ello es que generalmente en algunas de las experiencias el trata-

(*) Parte de esta exposición procede del anexo titulado *La noción de riesgo en estadística* de la obra de D. SCHWARTZ: *Méthodes statistiques à l'usage des médecins et des biologistes.* Flammarion. París, 1963.

miento que se dará a los individuos de la muestra conducirá a una proporción p_x diferente de la proporción p (hipótesis alternativa verdadera).

Por lo tanto, utilizando un nivel de significación igual a 0,05, el riesgo de error vale $\alpha=0{,}05$, considerando sólo el subconjunto de experiencias en las que la hipótesis nula es verdadera. El riesgo de error es inferior a 0,05 en el conjunto de experiencias analizadas, pudiendo valer como máximo 0,05 cuando la hipótesis nula es verdadera en todas las experiencias del conjunto analizado.

En una serie de experiencias el riesgo α se interpreta como la *proporción de errores tipo I* cometidos en dicho conjunto de experiencias.

En una experiencia aislada la interpretación del riesgo α no es tan clara como en el caso de una serie de experiencias. En este caso tiene mucho más sentido calcular el llamado *grado de significación* (*).

Supongamos que se ha observado una proporción p_o. A partir de ella podemos valorar hasta qué punto la hipótesis nula puede ser verdadera. Basta calcular la probabilidad de obtener, cuando H_o es verdadera, una diferencia mayor o igual a la diferencia $|p - p_o|$ observada. Esta probabilidad recibe el nombre de *grado de significación.*

Lógicamente, cuanto más pequeño sea el grado de significación, más probable será que la hipótesis nula sea falsa.

El cálculo del grado de significación (gs) no constituye una prueba estadística tal como la hemos definido. El grado de significación conduce muchas veces a situaciones de indecisión. ¿Cuándo gs$=0{,}12$ debemos tomar la decisión de rechazar o no la hipótesis nula?

Esta indecisión está producida por no haber fijado *a priori* el nivel de significación y, por lo tanto, no se ha definido la zona de rechazo de la hipótesis nula.

Si se fija el nivel de significación en un cierto valor α, a partir del grado de significación podemos realizar la prueba de decisión estadística:

Si gs$>\alpha$: Nada se opone en aceptar H_o.

Si gs$\leqslant\alpha$: Se rechaza H_o con riesgo α.

Pruebas bilaterales y pruebas unilaterales

Se aplica una *prueba bilateral* o *prueba de dos colas* cuando se tiene la hipótesis alternativa: la población x es diferente de la población A. La prueba de decisión descrita es una prueba bilateral, ya que la hipótesis alternativa (doble) era que la proporción p_x podía ser superior o inferior a la proporción p (p_x diferente de p). La figura 7a ilustra esta prueba para un riesgo $\alpha=0{,}05$.

(*) El cálculo del grado de significación puede verse en J. M. Domenech Massons: *Métodos estadísticos para la investigación en ciencias humanas.* Herder. Barcelona, 1975.

Cuando la hipótesis alternativa (simple) es que la característica estudiada en la población x es *superior* a la de la población A, se trata de una ***prueba unilateral*** o ***prueba de una cola*** por la derecha, tal como ilustra la figura 7b.

Cuando la hipótesis alternativa (simple) es que la característica estudiada en la población x es *inferior* a la característica de la población x, se trata de una *prueba unilateral* por la izquierda, tal como ilustra la figura 7c para el mismo riesgo, $\alpha = 0{,}05$.

Las pruebas unilaterales se diferencian de las bilaterales, entre otras razones, por el hecho de que el riesgo α está situado en uno de los dos lados, en el correspondiente a la zona de aceptación de H_1, ya que el riesgo α indica

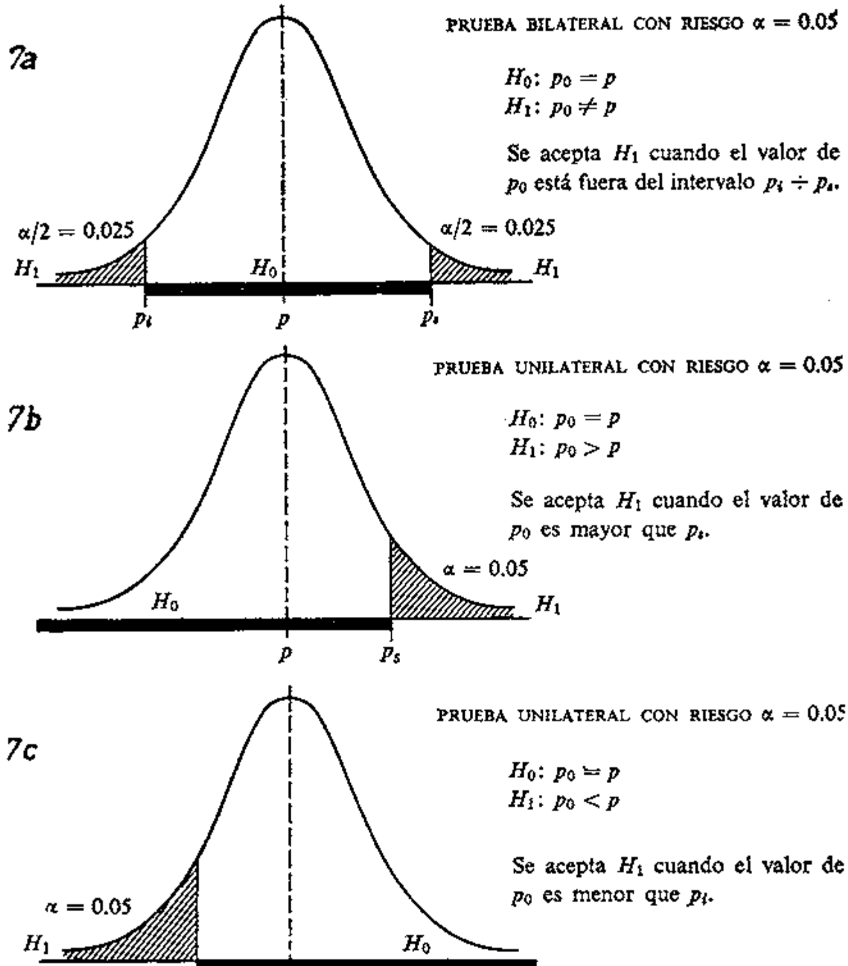


figura 7

la probabilidad de que una proporción observada procedente de p caiga en dicha zona.

De ahí se deduce que para realizar una prueba unilateral basta comparar el valor z calculado en la prueba con el valor $z_{2\alpha}$ correspondiente a dos veces el riesgo fijado. La tabla 2 da los valores con que se deben comparar z según la prueba sea bilateral o unilateral.

TABLA 2

Riesgo α de primera especie	10 %	5 %	1 %	1 ‰
Prueba bilateral $z\alpha$	1,65	1,96	2,58	3,29
Prueba unilateral $z_{2}\alpha$	1,28	1,65	2,33	3,09

Actitud explicativa y actitud pragmática

¿En qué casos se aplica una prueba bilateral y en cuáles se aplica una prueba unilateral?

Este aspecto, poco discutido en las obras de estadística, es muy importante. Para comprenderlo se necesita definir dos actitudes fundamentales que pueden adoptarse ante un problema experimental.

Se está en *actitud explicativa* cuando se tratan estadísticamente datos de problemas de *investigación fundamental.* Por ejemplo, si se elabora un nuevo método audiovisual de enseñanza x, en investigación fundamental tan interesante es demostrar que x es más eficaz que un método de referencia A, como demostrar que x es menos eficaz que A.

Para analizar los datos de este experimento con *actitud explicativa,* se aplicará una *prueba bilateral,* ya que la hipótesis alternativa planteada es doble.

Aunque el equipo de investigación *a priori* piense que x es más eficaz que A, la prueba a aplicar debe ser bilateral, ya que si finalizado el experimento se observara una diferencia significativa en sentido contrario, el investigador no dudará en concluir que A es más eficaz que x.

Se está en *actitud pragmática* cuando se tratan estadísticamente problemas de *investigación aplicada.* Continuando con el ejemplo anterior, la dirección de un centro de enseñanza que emplea el curso A, para decidir la compra del curso x estará sólo interesada en saber si el curso x es más eficaz que el A.

Probar que x es menos eficaz que A no le interesa en absoluto. Se trata de un problema de investigación aplicada.

Para analizar los datos de este experimento, en *actitud pragmática*, se aplicará una *prueba unilateral*, ya que la hipótesis alternativa planteada es simple.

En psicología experimental, las pruebas bilaterales son las más utilizadas ya que el investigador adopta una actitud explicativa ante la mayor parte de los problemas.

Determinación del número de individuos necesarios para limitar el riesgo β

El concepto de riesgo α y su interpretación gráfica han quedado claros con las anteriores explicaciones. Ahora se aborda el estudio y la interpretación gráfica del riesgo β en el caso particular de la prueba unilateral de comparación de una proporción observada a una proporción teórica. En caso de utilizar otras pruebas las fórmulas son diferentes (*), pero los conceptos aquí definidos permanecen invariables.

La prueba unilateral de comparación de una proporción observada a una proporción teórica, resuelve el problema esquematizado en la figura 8. Los valores p, p_o y n son datos conocidos. El valor p_x, correspondiente a la hipótesis alternativa, es siempre desconocido.

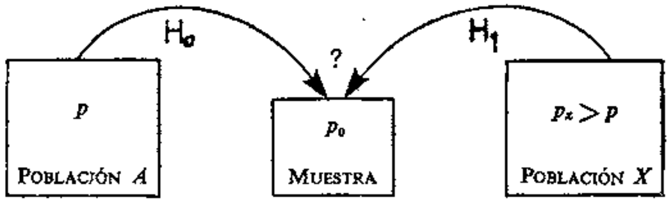


figura 8

La prueba consiste en ver si el valor p_o está dentro o fuera del intervalo de probabilidad de p. En la prueba unilateral se trata de ver si p_o está a la derecha o a la izquierda de l_s (figura 9a). Se llegan a las siguientes conclusiones:

$p_o \leqslant l_s$: Nada se opone en aceptar H_o.

$p_o > l_s$: Se rechaza H_o con riesgo α.

(*) El problema de la determinación del número de individuos necesario para las pruebas de decisión estadística más usuales, puede estudiarse en J. M. DOMENECH MASSONS: *Métodos estadísticos para la investigación en ciencias humanas*. Herder. Barcelona, 1975.

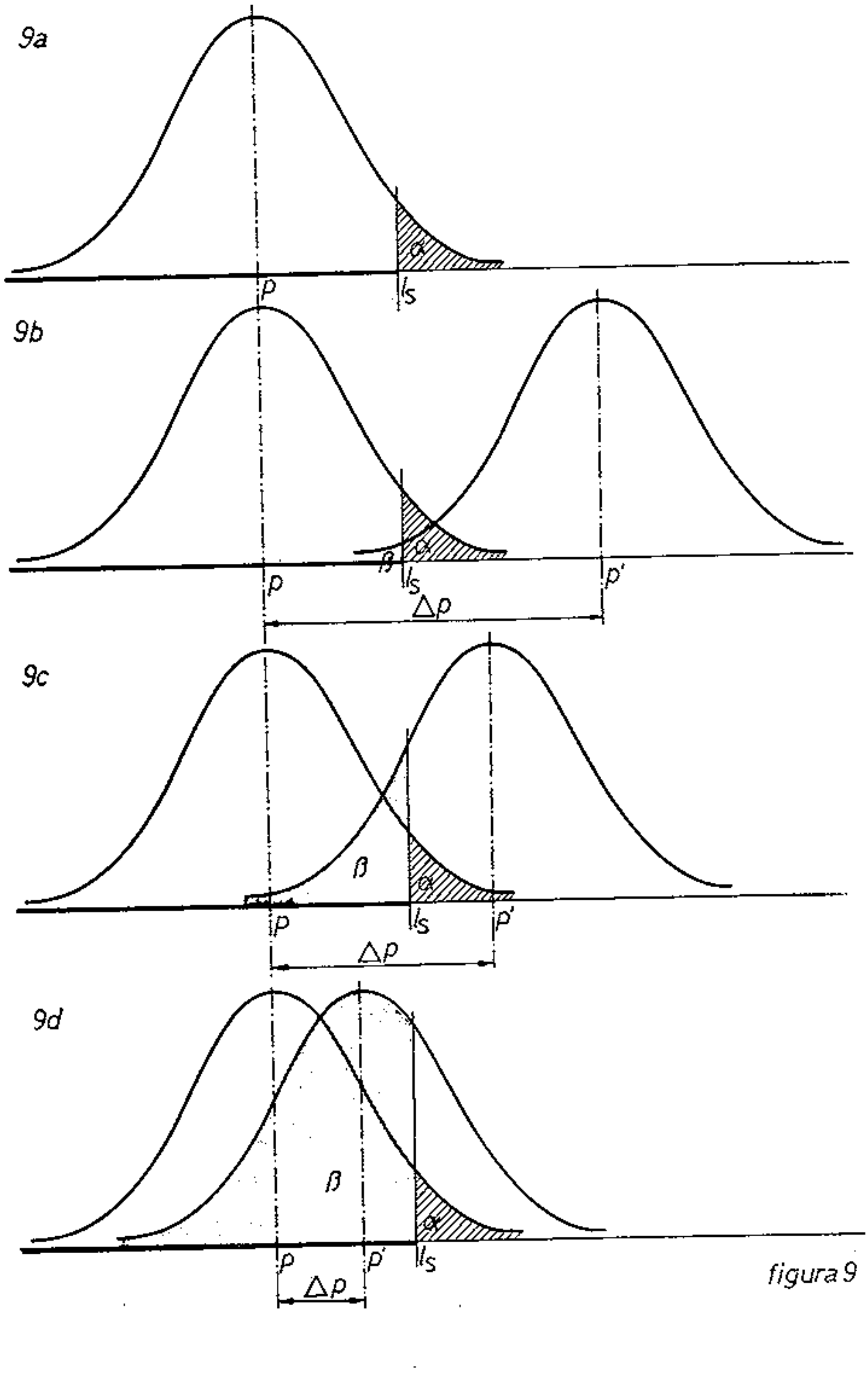


figura 9

Cuando la conclusión de la prueba es *nada se opone en aceptar* H_o, no quiere decir que la muestra proceda de la población p, ya que podría muy bien proceder de una población $p_x > p$, pero que debido a las fluctuaciones del azar p_o cayese a la izquierda de l_s, es decir, dentro de la zona de no rechazo de H_o. Por este motivo la conclusión de la prueba es indicar que el resultado observado *no está en contradicción con la hipótesis nula,* y lo de afirmar que la hipótesis nula es verdadera.

Si la población origen de la muestra fuese la población p', las proporciones observadas en muestras procedentes de p' fluctuarían alrededor de p' según una ley normal. Si se dibuja dicha distribución junto al gráfico de la figura 9a, se obtiene la figura 9b. El riesgo β será la zona sombreada ya que indica la cantidad de veces que muestras procedentes de p' dan proporciones observadas a la izquierda de l_s, es decir dentro de la zona de no rechazo de H_o.

El riesgo β, dibujada en la figura 9b, corresponde al riesgo de la prueba si la muestra procediese de la población p'. Pero el *riesgo* β *real* de la prueba es desconocido ya que la *proporción* p_x, correspondiente a la hipótesis alternativa, *es un valor que existe pero que no se conoce.*

La figura 9d ilustra el caso en que la proporción p_x coincide con un valor p' muy cercano al valor p. En el caso representado en la figura 9d el riesgo β es muy grande, supera el 50 %.

La figura 9 muestra cómo a medida que p' se aproxima a p el riesgo β de la prueba va aumentando.

Por otra parte, existe una influencia de n sobre β. La figura 10 ilustra cómo, para una misma hipótesis alternativa, al aumentar el tamaño n de la muestra disminuye el riesgo β.

El objeto de este apartado es calcular n de tal manera que los riesgos α y β sean pequeños, por ejemplo, $\alpha = \beta = 0{,}05$. ¿Cómo lograrlo si p_x es desconocido?

El conocimiento de p_x no es necesario si se analiza el problema desde un ángulo práctico. En efecto, supóngase que se trata de probar si una muestra procede de una población $p = 0{,}80$. Si en realidad procede de una población $p_x = 0{,}81$ y la prueba concluye que procede de p, este error tipo II cometido no tiene ninguna importancia ya que $p = 0{,}80$ está muy próximo de $p_x = 0{,}81$. Lo mismo ocurriría si p_x valiese 0,82, 0,83, etc. Pero llegará un momento, a partir de un cierto valor $p' = p + \Delta p$, en que el afirmar que la muestra procede de p empieza a ser un error tipo II grave y no se está dispuesto a cometerlo; es decir, a que la prueba concluya que la muestra procede de p cuando procede de una población caracterizada por la proporción p_x igual o mayor que p'.

El valor Δp se fija subjetivamente y es el *valor a partir del cual* si p_x es igual o superior a $p' = p + \Delta p$ no se está dispuesto a cometer *un error tipo II* (decir que p_o procede de p cuando en realidad procede de p_x).

Por lo tanto, para determinar el número de individuos necesario para asegurar un riesgo β pequeño, se debe evaluar subjetivamente Δp y fijar *a* *priori* unos riesgos α y β pequeños. La ecuación (15), que relaciona los datos p, p', Δp, α y β con n, se deduce de la figura 10.

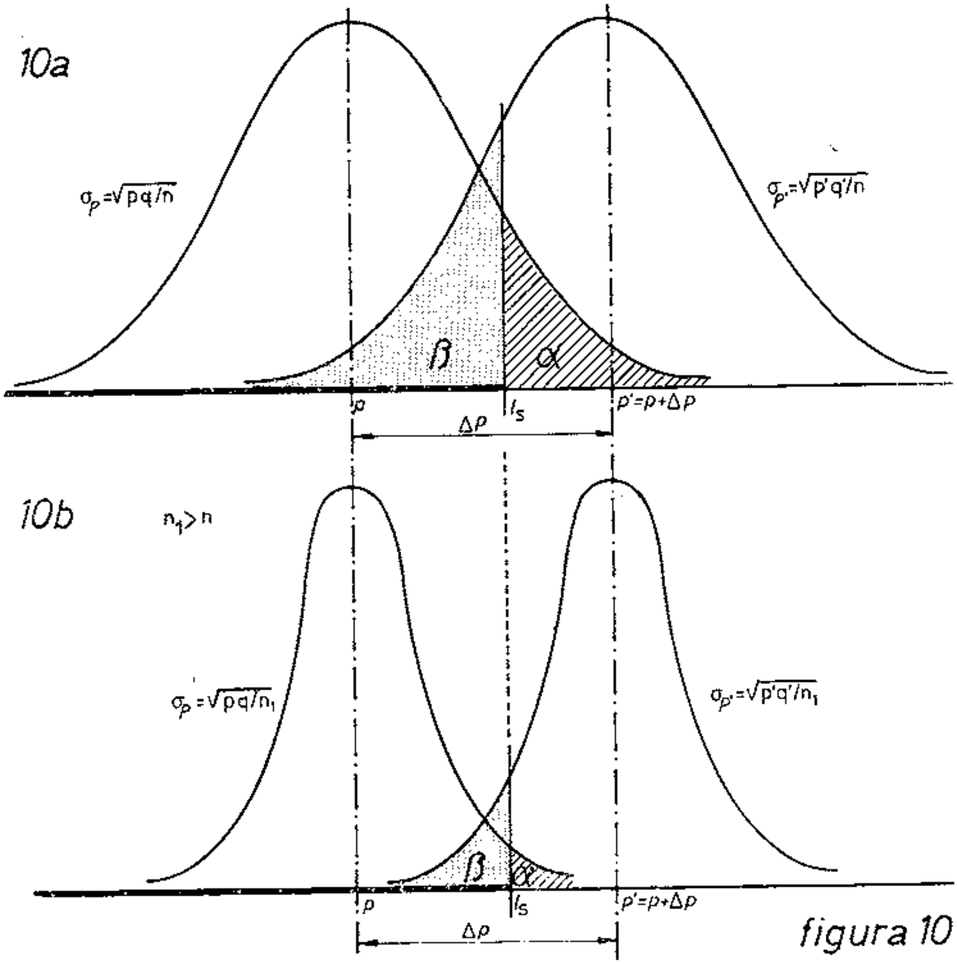


figura 10

$$\Delta p = \overline{pl_s} + \overline{l_s p'} = z_{2\alpha}\sqrt{pq/n} + z_{2\beta}\sqrt{p'q'/n} \tag{15}$$

Despejando n de esta ecuación se obtiene:

(16)

$$n = [(z_{2\alpha}\sqrt{pq} + z_{2\beta}\sqrt{p'q'})/\Delta p]^2$$

C. de A.: El valor n encontrado debe ser tal que los productos np, nq, np' y nq' sean iguales o mayores que 5.

Con un razonamiento análogo es fácil deducir (*) la fórmula (17), que da el número de individuos necesarios para una prueba bilateral de comparación de una proporción observada a una proporción teórica:

(17)

$$n=[(z\alpha\sqrt{pq}+z_2\beta\sqrt{p'q'})/\Delta p]^2$$

C. de A.: El valor encontrado n debe ser tal que los productos nq, nq, np' y nq' sean iguales o mayores que 5.

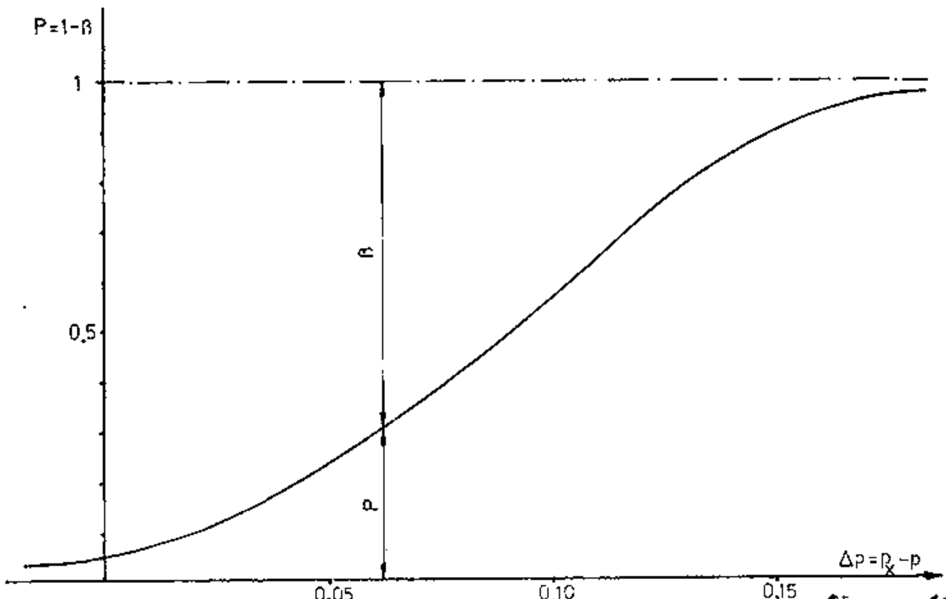


figura 11

Curva de potencia

La figura 11, llamada *curva de potencia,* da los valores de la potencia $P=1-\beta$ en función de posibles valores p_x correspondientes a la hipótesis alternativa, en una prueba unilateral de comparación de una proporción observada a una proporción teórica, en el caso particular de $p=0{,}50$, $n=100$ y un riesgo $\alpha=0{,}05$.

Esta curva es otra manera de ver el concepto ilustrado en la figura 9,

(*) J. M. Domenech Massons: *Métodos estadísticos para la investigación en ciencias humanas.* Herder. Barcelona, 1975. pp. 86 y sigs.

según el cual β aumenta ($P=1-\beta$ disminuye) a medida que Δp se hace pequeño.

No se indica la manera de dibujar esta curva ya que se trata de un sencillo problema de calcular repetidas veces superficies bajo la ley normal, según el esquema de la figura 9.

Estrategia global

La determinación previa del número de individuos completa la conclusión de la prueba cuando la decisión a tomar es la de no rechazo de la hipótesis nula.

El proceso metodológico a seguir en un proceso de decisión, mediante la prueba bilateral de comparación de una proporción observada a una proporción teórica, es el siguiente:

1. Fijar subjetivamente el valor Δp y los riesgos α y β.
2. Determinar el número n de individuos:

$$n=[(z\alpha\sqrt{pq}+z_2\beta\sqrt{p'q'})/\Delta p]^2$$

3. Realizar el experimento con la muestra de n individuos.
4. Calcular el valor z de la prueba de decisión:

$$z=|p_o-p|/\sqrt{pq/n}$$

5. Tomar la decisión dando una de las siguientes conclusiones:

$z\leqslant z\alpha$: Nada se opone en aceptar H_o. La probabilidad de equivocarse, cuando p_x está más alejado de p que la distancia Δp, es igual o inferior al valor β fijado.

$z>z\alpha$: Se acepta H_1. La probabilidad de equivocarse es igual al riesgo α fijado.

Generalización

Cuando la prueba de decisión es otra que la prueba bilateral de comparación de una proporción observada a una proporción teórica, cambian las fórmulas anteriores, pero el anterior proceso permanece invariable.

Ejemplo 3

Un grupo de investigación ha construido un nuevo cuestionario x para medir el factor neuroticismo y desea estudiar su eficacia con relación a la del cuestionario *A* de referencia, cuyo rendimiento es de 0,80 (80 de cada 100 individuos neuróticos son diagnosticados correctamente).

a) ¿Cuántos individuos neuróticos debe haber en la muestra?

Se trata de asegurar una determinada potencia a la prueba bilateral (actitud explicativa) de comparación de una proporción observada a una proporción teórica. El equipo de investigación evalúa $\Delta p = \pm 0{,}10$, ya que considera importante averiguar si el nuevo cuestionario detecta un $\pm 10\ \%$ de individuos neuróticos respecto a los del cuestionario *A*, y acepta unos riesgos $\alpha = \beta = 0{,}05$. Sustituyendo valores en (17) y teniendo en cuenta que de los dos valores $p'q'$ posibles:

$$p^s = 0{,}80 + 0{,}10 = 0{,}90 \qquad q^s = 1 - 0{,}90 = 0{,}10$$

$$p^i = 0{,}80 - 0{,}10 = 0{,}70 \qquad q^i = 1 - 0{,}70 = 0{,}30$$

se elige el correspondiente al mayor producto:

$$p'q' = \max(p^i q^i,\ p^s q^s) = 0{,}70 \times 0{,}30$$

Se obtiene:

$$n = [(1{,}96\sqrt{0{,}8 \times 0{,}2} + 1{,}65\sqrt{0{,}7 \times 0{,}3})/0{,}1]^2 = 233$$

b) En una muestra de $n = 233$ neuróticos el cuestionario x ha diagnosticado correctamente 201 de ellos. ¿x tiene eficacia diferente de *A*?

Si se acepta $\alpha = 0{,}05$, aplicando (14) para comparar $p_o = 201/233 = 0{,}86$ con $p = 0{,}80$, se obtiene:

$$z = |0{,}86 - 0{,}80| / \sqrt{0{,}8 \times 0{,}2/233} = 2{,}32 > 1{,}96$$

Conclusión estadística: Se acepta H_1 con riesgo 0,05.

Conclusión experimental: Puesto que $p_o = 0{,}86$ es superior a $p = 0{,}80$, y la diferencia encontrada es significativa, el rendimiento de x es superior al rendimiento de *A*.

c) ¿En qué cantidad el rendimiento del cuestionario x es superior al del cuestionario *A*?

Basta estimar la proporción p_x de diagnósticos correctos del cuestionario x, a partir de la proporción $p_o = 0{,}86$ observada en la muestra de $n = 233$ neuróticos. Aceptando un riesgo $\alpha = 0{,}05$ y aplicando (5) se obtiene:

$$0{,}86 \pm 1{,}96\sqrt{0{,}86 \times 0{,}14/233} = 0{,}86 \pm 0{,}84$$

$$p_x \in (0{,}82 \div 0{,}90)$$

El rendimiento de x está situado entre 0,82 y 0,90, es decir el rendimiento del cuestionario x será superior entre un 2 % y un 10 % respecto al rendimiento del cuestionario A.

d) Una vez demostrada estadísticamente la superioridad del cuestionario x, el grupo de investigación desea venderlo a un gabinete psicológico que emplea el cuestionario A. Antes de comprarlo, el director de dicho gabinete desea realizar un experimento para probar que la eficacia del cuestionario x es superior a la del cuestionario A. ¿Cuántos individuos debe haber en la muestra?

Se trata de asegurar una cierta potencia a la prueba unilateral (actitud pragmática) de comparación de una proporción observada a una proporción teórica. El director del gabinete psicológico evalúa $\Delta p = +0,07$, ya que considera importante adoptar x si como mínimo detecta un 7 % más de neuróticos que A, y acepta unos riesgos $\alpha = \beta = 0,05$. Sustituyendo valores en (16) se obtiene:

$$n = [(1,65\sqrt{0,80 \times 0,20} + 1,65\sqrt{0,87 \times 0,13})/0,07]^2 = 302$$

e) En la muestra de $n=302$ neuróticos el cuestionario x ha diagnosticado correctamente 254. ¿x es más eficaz que A?

Si se acepta $\alpha=0,05$, aplicando (14), pero teniendo en cuenta que se trata de una prueba unilateral de comparación de $p_o=254/302=0,84$ con $p=0,80$, se obtiene:

$$z = |0,84 - 0,80| / \sqrt{0,80 \times 0,20/302} = 1,74$$

El valor $z=1,74$ es superior a $z_{2\alpha}=1,65$ y además $p_o=0,84$ es superior a $p=0,80$. *Conclusión:* el cuestionario x es más eficaz que el cuestionario A.

RESUMEN

Este trabajo es una introducción al método estadístico. Aborda dos aspectos fundamentales de la estadística: la *teoría de la estimación*, que permite predecir con determinada precisión las características estadísticas de una población a partir de la información aportada por los individuos de una muestra extraída al azar de dicha población, y la *teoría de la decisión*, que, ante las diferencias observadas al aplicar diferentes tratamientos a grupos de individuos, permite saber si pueden ser explicadas solamente por la influencia del azar o si son debidas a otras causas.

Por razones didácticas, este estudio se particulariza a los caracteres cualitativos (proporciones), aunque su generalización a los caracteres cuantitativos es inmediata.

El interés de este artículo reside principalmente en el estudio detenido del significado de los riesgos α y β, en la discusión de cuándo la prueba a

aplicar debe ser unilateral y cuándo debe ser bilateral, y en el estudio detallado del número de individuos que debe haber en una experiencia, para limitar el riesgo β en un prueba de decisión estadística.

RÉSUMÉ

Le présent texte constitue une introduction à la méthode statistique. Il est axé sur deux aspects essentiels de la statistique: la *théorie de l'estimation,* qui permet de prédire avec un degré d'exactitude précis les traits statistiques d'une population donnée, à partir des informations fournies par les individus d'un échantillon tiré au hasard dans cette population; et la *théorie de la décision* qui, face aux diversités observées lors de l'application de traitements différents à des groupes d'individus, permet de savoir si les diversités ne sont explicables que par le hasard ou bien sont dûes à d'autres causes.

Pour des raisons d'ordre didactique, l'étude s'arrête sur les caractères qualitatifs (proportions), bien que la généralisation aux caractères quantitatifs en découle immédiatement.

L'intérêt de l'article se trouve surtout dans l'étude détaillée de la signification des risques α et β, dans la discussion de l'unilatéralité ou la bilatéralité de la preuve à appliquer, et dans l'examen approfondi du nombre d'individus devant participer à une expérience, dans le but de limiter le risque β dans une preuve de décision statistique.

SUMMARY

This paper is an introduction to the statistical method. Two fundamental aspects of statistics are studied: the *Estimation Theory,* which allows to foresee with a degree of precise accuracy the statistical characteristics of a given population, based on the information obtained from the individuals of a sample taken at random among this population; and the *Decision Theory,* which, in front of diversities observed when applying different treatments to groups of individuals, allows to know whether the diversities can only be explained by chance or whether they are due to other causes.

For didactic reasons, the study is centered on the qualitative characters (proportions), although its generalization to the quantitative characters follows immediately.

The interest of the article lies above all in the detailed study of the α and β risks, in the discussion of when to apply the unilateral or the bilateral test, and in the close examination of the number of individuals to be included in an experience, in order to limit the β risk in a test of statistical decision.