Publicado en: Morales Vallejo, Pedro (2008) ***Estadística aplicada a las Ciencias Sociales.*** **Madrid: Universidad Pontificia Comillas (edit@pub.upcomillas.es)**

Estadística inferencial: el *error típico* de la media

©Pedro Morales Vallejo,
Universidad Pontificia Comillas, Madrid
Facultad de Ciencias Humanas y Sociales
(última revisión 21 de Septiembre de 2007)

Índice

1. Introducción: estadística descriptiva y estadística inferencial: *estadísticos* y *parámetros, poblaciones y muestras*

Recordamos algunos conceptos básicos:

Una *población* es un conjunto de elementos (sujetos, objetos) cuyos límites los define el investigador; por ejemplo los alumnos de una universidad, o los de una sola facultad o los de todo el país…

Una *muestra* es un número concreto de elementos extraídos de una población.

Una *muestra aleatoria* es aquella en la que todos los sujetos (u objetos) han tenido la misma probabilidad de ser escogidos; las muestras aleatorias son las que mejor representan las características de la población[1].

La estadística *descriptiva* tiene por objeto describir las *muestras:* por ejemplo, la media aritmética (una medida de tendencia central) y la desviación típica (una medida de dispersión) son *estadísticos* o medidas propias de la estadística descriptiva: nos describen cómo es una muestra.

La estadística *inferencial* nos permite hacer inferencias, sacar conclusiones con respecto a una población: a partir de los datos descriptivos de una muestra, deducimos los datos o *medidas* de la población, que en este caso se denominan *parámetros*.

Normalmente el investigador trabaja con muestras, grupos concretos a los cuales tiene acceso o que ha buscado y que puede medir en alguna característica. Las poblaciones son en general inasequibles; se trabaja con pequeñas muestras y se generalizan las conclusiones a las poblaciones a las que pertenecen las muestras. Lo que vamos a ver ahora tiene que ver sobre todo (no exclusivamente) con la generalización a la población de los datos que encontramos en muestras concretas.

2. Las distribuciones muestrales y el error típico

Dos conceptos previos importantes son los de *distribución muestral* y *error típico*. En definitiva nos vamos a encontrar con una aplicación de lo que ya sabemos de la *distribución normal* y de las *puntuaciones típicas:* en la distribución normal conocemos las probabilidades de obtener una puntuación superior o inferior a cualquier puntuación típica. Ahora se trata básicamente de una aplicación de esta relación.

Qué es una *distribución muestral* lo podemos ver con facilidad con un caso concreto:

1º Imaginemos una *población* de sujetos; por ejemplo los alumnos de una universidad. Los *límites de la población* (qué sujetos, u objetos, pertenecen a una población) lo determina el que investiga. De la misma manera que ponemos como ejemplo de población a los alumnos de una universidad, podríamos decidir que la población que vamos a estudiar son los alumnos de una sola facultad, o los alumnos de todas las universidades del país.

2º De esta población podemos extraer una *muestra aleatoria* de, por ejemplo, 30 sujetos.

Muestra *aleatoria* quiere decir que *todos los sujetos de la población han tenido en principio la misma oportunidad de ser elegidos*. Las muestras aleatorias son en principio las que mejor representan

[1] Los diversos tipos de muestreo, aleatorio otros, y cómo llevarlos a cabo, pueden verse en muchos textos (como Hernández Sampieri, Fernández Collado y Baptista Lucio, 2000; Salkind, 1998) y en monografías específicas (como Rodríguez Osuna, 1993). Una breve exposición de los tipos de muestras puede verse en Internet, en StatPac Inc (2003)

las características de la población. Hay varios métodos para elegir muestras aleatorias pero no los tratamos aquí.

3º De esta muestra podemos calcular la *media.* Seguimos extrayendo muestras aleatorias y calculando sus medias.

4º Al disponer de un *número grande* de medias tendríamos una *distribución* de estas medias; esa distribución es una *distribución muestral:* no se trata de una distribución de puntuaciones individuales sino de *medias* de *muestras*.

Un punto importante es que aunque las muestras no tengan una distribución normal, las medias de estas muestras sí tienden a seguir la *distribución normal*.

5º La *desviación típica* de estas *distribuciones muestrales* se denomina *error típico* y se puede *estimar* a partir de los datos de una muestra. Por lo tanto ***un error típico es la desviación típica de una distribución muestral,*** y se interpreta como cualquier desviación típica.

Dos distribuciones muestrales, con sus errores típicos, nos van a interesar de manera especial:

1) la *distribución muestral de las medias;*
2) la *distribución muestral de las diferencias entre medias de la misma población.*

Estas distribuciones muestrales son modelos teóricos que a partir de los datos de una muestra nos van a permitir *inferir* conclusiones acerca de la población a la que pertenece la muestra. Conociendo el error típico de estas distribuciones podemos estimar entre qué limites se encuentra la media de la población o si dos muestras proceden de poblaciones distintas con media distinta. Ahora nos centramos en el *error típico de la media*.

Conviene caer en la cuenta desde el principio de *la utilidad del error típico de la media*. Es fácil obtener la *media* de una muestra en cualquier variable de interés, pero con frecuencia lo que nos interesa no es la media como dato descriptivo de una muestra, sino conocer o tener una idea de *por dónde anda* la media en la población representada por esta muestra. La media de la población no la vamos a conocer, pero sí podremos estimar *entre qué valores se encuentra.*

La media de una muestra podemos interpretarla como una *estimación* (solamente una estimación sujeta a error) de la media de la población. Esta estimación será *más precisa:*

1º Si la muestra es aleatoria porque en ese caso representa mejor las características de la población

2º Si la muestra es grande (si la muestra comprendiera a toda la población tendríamos el dato exacto, no una estimación).

El error típico, como es la desviación típica de todas las posibles muestras de esa población, nos va a permitir localizar entre qué límites se encuentra la media de la población.

Este planteamiento es semejante al que nos encontramos en los sondeos de opinión, como son las *encuestas pre-electorales*. Si el 48% de los sujetos entrevistados dice que va a votar a un determinado candidato, esto no quiere decir que el 48% exacto de la población le vaya a votar. Sin embargo los datos obtenidos de una muestra nos van a permitir estimar un tanto por ciento mínimo probable y un tanto por ciento máximo probable de votantes a ese candidato: entre esos dos tantos por ciento se va a encontrar el tanto por ciento definitivo cuando todos hayan votado. De los datos de una muestra extrapolamos a la población, por eso se trata de estadística *inferencial*.

De manera análoga podemos pensar en *distribuciones muestrales* de otros estadísticos como proporciones, medianas, coeficientes de correlación, etc., y también en *distribuciones muestrales de las diferencias* entre proporciones, medianas, coeficientes de correlación, etc., con aplicaciones semejantes a las que vamos a ver con respecto a la media que son las de utilidad más inmediata y frecuente.

3. El *error típico* de la media

Según el *teorema del límite central,* si de cualquier población se extraen muestras *aleatorias* del mismo tamaño N, al aumentar el número de muestras sus medias se distribuyen *normalmente,* con media μ y una desviación típica, o *error típico* $\sigma_{\overline{X}} = \sigma/\sqrt{N}$

Esta *distribución muestral de las medias* es independiente de la distribución de la población: aunque la distribución en la población no sea normal, las medias de las muestras aleatorias extraídas de esa población sí tienden a tener una distribución normal.

El *error típico* de la media *(desviación típica de la distribución muestral de las medias)* podemos expresarlo de dos maneras:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}} \qquad [1]$$

En la fórmula [1] la desviación típica del numerador se supone calculada dividiendo por N-1 la *suma de cuadrados* (o la suma de las puntuaciones diferenciales, $X-\overline{X}$, elevadas previamente al cuadrado).

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N-1}} \qquad [2]$$

En la fórmula [2] la desviación típica se ha calculado dividiendo por N, como es normal hacerlo cuando se calcula la desviación típica como dato *descriptivo* de la muestra. Ambas fórmulas son equivalentes y dan el mismo resultado; la única diferencia está en cuándo se ha restado 1 a N.

En principio suponemos que la desviación típica de la muestra la hemos calculado dividiendo por N, como dato descriptivo de la dispersión en la muestra, por eso al calcular el error típico de la media utilizaremos la fórmula [2].

La desviación típica del numerador en ambas fórmulas es la calculada en la muestra, pero debería ser la desviación típica calculada con todos los sujetos de la población. Como desconocemos la desviación típica de la población, utilizamos la de la muestra como una *estimación* de la desviación típica de la población.

Observando la fórmula del *error típico de la media* podemos ver que:

1º Es claro que el *error típico* de la media será *menor* que la *desviación típica* de cualquier muestra: el cociente siempre será menor que el numerador. Esto quiere decir que las medias de las muestras son *más estables* y tienden a *oscilar menos* que las puntuaciones individuales; dicho de otra manera, las medias de muestras de la misma población se parecen entre sí más que los sujetos (u objetos) de una muestra entre sí.

2º Observando las fórmulas vemos también que el error típico de la media será más pequeño en la medida en que N sea grande: si aumentamos el denominador, disminuirá el cociente.

Es natural que al aumentar el número de sujetos (N) el *error* sea menor: la media de la muestra se aproximará más a la media de la población. Si N es muy grande, el error tiende a cero; y si N no comprende a una muestra sino a toda la población, el error sería cero: en este caso la media de la población coincide con la media de la muestra y no hay *error muestral* (o variación esperable de muestra a muestra).

3º Por otra parte si la desviación típica de la muestra es grande, el error típico estimado de la media será también mayor: si aumentamos el numerador, el cociente será mayor.

También esto es lógico: una desviación típica grande en una muestra quiere decir que las diferencias entre los sujetos son mayores, y consecuentemente las medias de las diferentes muestras también diferirán más entre sí.

4. Utilidad del *error típico* de la media

Vamos a exponer dos *usos* del error típico de la media. Aquí el más importante es el primero, establecer los límites probables *(intervalos de confianza)* entre los que se encuentra la media de la población, un planteamiento típico y frecuente en estadística inferencial. Veremos también lo mismo aplicado a una *proporción,* que es la media cuando se trata de datos dicotómicos (1 ó 0).

En segundo lugar el *error típico* de la media nos permite comprobar si una muestra con una determinada media puede considerarse como perteneciente a una población cuya media conocemos, es también de interés y es simplemente una aplicación del anterior. Igualmente podemos aplicarlo si la media es una proporción (una proporción es la media cuando los datos son *unos* y *ceros*). Es conveniente exponerlo aquí brevemente, pero lo volveremos a encontrar al tratar del contraste de medias, pues allí veremos un procedimiento más sencillo. Son procedimientos equivalentes.

Podemos añadir un tercer uso del error típico de la media, que es determinar el número de sujetos que necesitamos en la muestra para extrapolar los resultados a la población. Cuando a partir de los datos de una muestra nos interesa extrapolar los resultados a la población (por ejemplo cuántos van a votar a un partido político en unas elecciones), lo hacemos con *un margen de error* (en cuyo cálculo tenemos en cuenta el *error típico* y nuestro *nivel de confianza):* si queremos un margen de error pequeño, necesitaremos más sujetos… por eso en las fórmulas para determinar el número de sujetos de la muestra entrará el error típico. Este punto lo veremos de manera más sucinta, porque suele verse con más detalle en otro contexto más práctico, al tratar de las muestras, tipos de muestras, número de sujetos necesario según distintas finalidades, etc.

No sobra por último repetir una observación ya hecha: estamos tratando del error típico de la *media* (o desviación típica de una hipotética distribución de medias), pero de manera análoga podríamos tratar de *otros* errores típicos y con las mismas aplicaciones: de los coeficientes de correlación, de las proporciones, y de cualquier otro estadístico.

4.1. Establecer entre qué limites *(intervalos de confianza)* se encuentra la media (μ) de la población *(establecer parámetros poblacionales)*

La media de una muestra ($\overline{X}$) es una *estimación* de la media de la población (μ); pero decir que es una *estimación* quiere decir que está sujeta a error. La media exacta de la población no la conocemos; pero sí podemos estimar entre qué *límites extremos* se encuentra, y esto a partir de la media de una muestra y del error típico de la media.

El *error típico de la media* no es otra cosa que una estimación de la *desviación típica de las medias* (de muestras de la misma población), y se interpreta de la misma manera; así por ejemplo según la distribución normal, el 95% de las medias se encontrará entre -1.96σ y + 1.96σ; aquí σ es propiamente $\sigma_{\overline{X}}$, el error típico de la media.

Si tenemos estos datos de una muestra: N = 30, $\overline{X}$ = 62.8 y σ = 7.9, tendremos que (fórmula [2]):

$$\sigma_{\overline{X}} = \frac{7.9}{\sqrt{30-1}} = 1.47$$

El *error típico de la media* (o desviación típica de las medias posibles) es en este caso igual a 1.47, y según las probabilidades de la distribución normal podremos afirmar que:

Hay un 68% de probabilidades de que la media de la población se encuentre entre la media de la muestra *más menos* un error típico:

entre (62.8 - 1.47) y (62.8 + 1.47) = entre 61.33 y 64.27.

Hay un 95% de probabilidades de que la media de la población se encuentre entre la media de la muestra *más menos* 1.96 errores típicos:

entre [62.8 - (1.96 x 1.47)] y [62.8 + (1.96 x 1.47)] = entre 59.92 y 65.68.

Si deseamos *mayor seguridad* al establecer los *límites probables* entre los que se encuentra la media de la población, podemos tomar como límite 2.57 errores típicos, porque sabemos que entre la media *más menos* 2.57 desviaciones típicas se encuentra el 99% de los casos. En este caso:

El *límite inferior* de la media de la población sería [62.8 - (2.57 x 1.47)] = 59.02
El *límite superior* de la media de la población sería [62.8 + (2.57 x 1.47)] = 66.58

A estos límites, o *valores extremos, superior e inferior, de la media en la población* se les denomina *intervalos de confianza,* porque eso es precisamente lo que expresan: entre qué límites podemos situar la *media de la población* con un determinado *grado de confianza* o de seguridad (o de probabilidades de no equivocarnos). Los intervalos de confianza suelen denominarse en las encuestas de opinión *márgenes de error*.

Estos *intervalos de confianza* podemos establecerlos con diversos *niveles de seguridad,* que vendrán dados por el valor de z que escojamos, por lo que podemos expresarlos así:

$$\textit{intervalos de confianza de la media} = \overline{X} \pm (z)\frac{\sigma}{\sqrt{N-1}} \qquad [3]$$

La cantidad que sumamos y restamos a la media de la muestra podríamos denominarla *margen de error* al estimar los límites probables de la media en la población y que podemos expresar de esta manera:

$$\begin{matrix}\text{margen de error al estimar}\\ \text{la media de la población}\end{matrix} = \begin{bmatrix}\text{nivel de confianza}\\ \text{expresado por un valor de z}\end{bmatrix} x \begin{bmatrix}\text{error típico}\\ \text{de la media}\end{bmatrix} \qquad [4]$$

Como ya hemos indicado estos límites o márgenes de error serán más ajustados cuando el número de sujetos sea mayor. Es útil visualizar el efecto del tamaño de la muestra en los intervalos de confianza (tabla 1). Queremos saber, por ejemplo, entre qué límites se encuentra la media de la población, estimada a partir de una muestra pequeña (N = 10) y de una muestra grande (N = 500), y con un niveles de confianza de .05 (que corresponde a z = 1.96). En ambos casos suponemos en las muestras una media = 8 y una desviación típica = 2.

N	*Error típico de la media*	*Nivel de confianza*	*Límite mínimo de la media en la población*	*Límite máximo de la media en la población*
10	$\frac{2}{\sqrt{9}} = .67$	.05 (z = 1.96)	8-(1.96)(.67) = 6.68	8+ (1.96)(.67)= 9.31
500	$\frac{2}{\sqrt{499}} = 09$	.05 (z = 1.96)	8-(1.96)(.09) = 7.82	8+ (1.96)(.09) = 8.18

Tabla 1

Lo vemos con más claridad con una representación gráfica:

	6 7 8 9 10
Límites extremos de la media de la población estimados a partir de N = 10	6.68 9.31
Límites extremos de la media de la población estimados a partir de N = 500	7.82 8.18

Con más sujetos los límites son más ajustados, hay más precisión; con 10 sujetos situamos la media de la población entre 6.68 y 9.13 (una diferencia de 2.45 puntos), y con 500 sujetos entre 7.82 y 8.18 (una diferencia entre ambos límites de sólo .36).

También con un nivel de confianza más estricto (.01, que corresponde a z = 2.57, en vez de .05) tenemos una menor probabilidad de salir falsos profetas, más seguridad, pero los límites son más amplios (una mayor seguridad pero menor precisión). Si en el ejemplo anterior utilizamos .01 en vez de .05 con 500 sujetos veremos la diferencia (tabla 2).

N	*Error típico de la media*	*Nivel de confianza*	*Límite mínimo de la media en la población*	*Límite máximo de la media en la población*
500	$\frac{2}{\sqrt{499}} = 09$	.05 (z = 1.96)	8-(1.96)(.09) = 7.82	8+ (1.96)(.09) = 8.18
500	$\frac{2}{\sqrt{499}} = 09$	.01 (z = 2.57)	8- (2.57)(.09) = 7.77	8+ (2.57)(.09) = 8.23

Tabla 2

Con una menor probabilidad de error (.01 en vez de .05) los límites extremos de la media en la población son 7.77 y 8.23, una diferencia de .46 en vez de .36

Tanto $\overline{X}$ como σ son los valores calculados en una muestra. Naturalmente el valor exacto de la media de la población (μ) no lo conocemos: puede estar en cualquier punto entre los valores extremos indicados. También puede estar fuera de los límites indicados, pero esto va siendo más improbable cuando establecemos unos intervalos de confianza más estrictos.

Es normal operar con un *nivel de confianza* del 95% (o, lo que es lo mismo, con una probabilidad de error, al situar los límites extremos de la media, de un 5%); en este caso z en la fórmula [3] será igual a 1.96; como se desprende de esta fórmula, a mayor valor de z *(mayor seguridad)* los límites serán más extremos.

Cuando calculamos la media de una muestra en una variable de interés ¿Es útil calcular además entre qué límites se encuentra la media de la población?

Con frecuencia nos bastará conocer la media de una muestra concreta como dato informativo, pero con frecuencia *extrapolamos informalmente* de la muestra a la población. Siempre es útil *relativizar* este tipo de información, y con mayor razón si *de hecho* (como es frecuente) estamos utilizando la media de una muestra como estimación de la media de la población[2].

4.2. Establecer los intervalos de confianza de una proporción

El *error típico de una proporción* es un caso particular del error típico de la media pero dado el uso frecuente de proporciones y porcentajes es útil verlo por separado y con ejemplos ilustrativos.

[2] Una de las recomendaciones de la *American Psychological Association* es calcular siempre los intervalos de confianza (Wilkinson, Leland and Task Force on Statistical Inference APA Board of Scientific Affairs 1999; American Psychological Association, 2001).

Cuando los datos son *dicotómicos* (1 ó 0) la media *p* es la proporción de sujetos que responden *sí* o que escogen la respuesta codificada con un 1. Si de 200 sujetos 120 responden *sí* (ó 1) a una pregunta y 80 responden *no* (0), la media *p* es igual a 120/200 = .60: el 60% de los sujetos (o *una media* del 60%) han respondido *sí*.

El error típico de una proporción es el mismo que el error típico de cualquier media, solo que en este caso la *media* es *p*, la *varianza* es *pq [proporción de unos por proporción de ceros]* y la *desviación típica* es $\sqrt{pq}$.

La fórmula del *error típico de una proporción* (σ_p) será por lo tanto:

$$\sigma_p = \frac{\sigma}{\sqrt{N}} = \frac{\sqrt{pq}}{\sqrt{N}} \quad \text{por lo tanto} \quad \sigma_p = \sqrt{\frac{pq}{N}} \qquad [5]$$

En el ejemplo anterior tenemos que N = 200, p =120/200 = .60 y q =.40 por lo tanto el error típico de la proporción será:

$$\sigma_p = \sqrt{\frac{(.60)(.40)}{200}} = .0346$$

De manera análoga a lo que hemos visto en los intervalos de confianza de la media en variables continuas (fórmulas 3 y 4), los intervalos de confianza de una proporción p serán:

$$\textit{Intervalos de confianza de una proporción} = p \pm [z]\left[\sqrt{\frac{pq}{N}}\right] \qquad [6]$$

Ahora podemos hacernos esta pregunta: en esa muestra de 200 sujetos han respondido *sí* 120 sujetos (una *media* de .60 o el 60%), pero ¿cuántos responderán *sí* en la población representada por esa muestra? Ya podemos intuir la importancia de esta pregunta si pensamos en los *sondeos pre-electorales;* lo que interesa realmente no es conocer cuántos sujetos *de esa muestra* van a votar a un candidato, sino cuántos le votarán el día de las elecciones.

La proporción de votantes que dirán *sí* a ese candidato (o la *media* de votantes) en la población no la sabemos (habría que preguntar a todos y eso se hará el día de las elecciones), pero sí podemos *estimar* entre qué límites máximo y mínimo se encuentra esa proporción con un determinado *nivel de confianza* (o seguridad de acertar en la predicción); es decir, podemos establecer *los márgenes de error.*

Para responder a esta pregunta calculamos los *intervalos de confianza* de la media (p = .60) con un nivel de confianza de .05 (un 5% de probabilidades de equivocarnos) que equivale a z = 1.96.

La proporción de los que dirán *sí* a juzgar por los datos de esa muestra estará entre .60 *menos* 1.96 *errores típicos* y .60 *más* 1.96 errores típicos:

Límite mínimo: .60 – (1.96)(.0346) = .60 – .0678 = .5322 (el 53%)
Límite máximo: .60 + (1.96)(.0346) = .60 + .0678 = .6678 (el 67%)

El *margen de error* en nuestra predicción es .0678 (casi un 7% redondeando). En la muestra encuestada ha respondido *sí* el 60%, pero en la población representada por esa muestra esperamos que responda *sí* entre un 53% y un 67%.

El ejemplo de los sondeos pre-electorales pone de relieve la importancia de calcular los *intervalos de confianza de una proporción* (y es lo que se hace y comunica cuando se publican estas encuestas), pero estos intervalos de confianza son informativos casi en cualquier situación. Cuando se hacen sondeos de opinión en grupos diversos (alumnos, padres de alumnos, grupos profesionales, etc.) prácticamente se tienen muestras (no responde *toda* la población) pero los resultados suelen interpretarse como si todos hubieran respondido; lo realmente informativo es aportar los intervalos de confianza, o entre qué límites se encuentran con toda probabilidad las respuestas si *todos* hubieran respondido.

Cuando distintos grupos responden a la misma pregunta (*sí* o *no* en este caso, pero puede tratarse también de respuestas con valores continuos) es útil especificar el error típico de la proporción en cada muestra y los intervalos de confianza entre los que se encuentra la proporción de *síes* (o *unos)* en las poblaciones representadas por esas muestras (ejemplo en la tabla 3, con un nivel de confianza de .05 ó z = 1.96).

Muestras de distinto tamaño	*Proporción (ó %) en la muestra*	*Error típico*	*Intervalos de confianza en las poblaciones representadas por esas muestras* 45 50 55 **60** 65 70 75 80
A, N = 300	.60	.0283	54 **60** 65
B, N = 80	.60	.0548	49 **60** 71

Tabla 3

En la tabla 3 podemos observar que en las muestras A y B responde afirmativamente la misma proporción de sujetos (un 60%), pero al extrapolar los resultados a las poblaciones representadas por esas muestras el margen de error es mucho menor en la muestra A porque se trata de más sujetos.

Al hablar de extrapolar a la población los resultados de una muestra (en este caso y en cualquier otro) hay que hacer una observación importante. Estamos suponiendo que esa muestra es representativa de la población, que no está *sesgada*, y es esto lo se intenta conseguir con las muestras aleatorias. Cuando éste no es el caso (responden los sujetos disponibles, los que quieren, etc.) siempre podemos pensar en la población que *pueda* estar representada por esa muestra y ser cautelosos al generalizar los resultados. En cualquier caso siempre es más seguro informar sobre los intervalos de confianza sin limitarnos a una proporción o porcentaje aparentemente exacto.

4.3. Comparar la media de una muestra con la media de una población

Se trata ahora de verificar si podemos considerar que una muestra, cuya media conocemos, pertenece a una población cuya media también conocemos. Si tenemos la media de una muestra ($\overline{X}$) y la media de una población (μ), podemos preguntarnos ¿Es posible afirmar que nuestra muestra, cuya media conocemos, pertenece a (es una muestra aleatoria de) una población con media μ? Si la respuesta es *no,* podremos afirmar que la muestra pertenece a una *población distinta,* con una media distinta.

Al hablar de diferencias *estadísticamente significativas* estamos hablando de diferencias *no aleatorias,* no explicadas por el *error muestral,* no *esperables por azar*. Esto lo afirmaremos con una determinada probabilidad de error; es el *nivel de significación* o *nivel de confianza.*

Es más frecuente comparar las medias de dos muestras (para comprobar si proceden de o pertenecen a poblaciones distintas), pero también tiene su interés el comparar la media de una muestra con la media de una población cuando ésta es conocida por otras investigaciones o estudios, o es la conclusión lógica de una determinada teoría, o simplemente la *media de la población* es una hipótesis de trabajo; siempre podemos pensar en *medias hipotéticas*.

Lo veremos con un ejemplo. Un profesor pone a sus alumnos una serie de problemas y obtiene estos resultados: N = 40, $\overline{X}$= 12.6 y σ = 4.25. El profesor piensa que un resultado óptimo y posible hubiera sido obtener una media de 15, y se pregunta ¿puede considerarse esta muestra de 40 alumnos como una muestra aleatoria de una población cuya media fuera μ = 15?

Este tipo de planteamientos puede tener su interés cuando la media de la población es una hipótesis plausible o hay datos de otros estudios, etc. Vamos a suponer que el nivel de confianza que nos ponemos es de α = .01 (que corresponde a z = 2.57; probabilidad de equivocarnos: 1% o menos; sólo el 1% de los casos *cae* más allá de ±2.57).

Podemos solucionar el problema de dos maneras.

1º Nuestra muestra pertenece a una población cuya media en principio desconocemos. Lo que sí podemos hacer es estimar el *límite máximo* de la media de la población a la que pertenece nuestra muestra, tal como hemos visto antes, y con un riesgo máximo de error del 1%, tal como hemos fijado previamente.

1. Calculamos el *error típico de la media,* $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{N-1}} = \dfrac{4.25}{\sqrt{40-1}} = 0.68$

2. ¿Cuáles serán los límites superior e inferior de la media de la población, con una probabilidad de error del 1%?

   El *límite superior* será $\overline{X}$ +(2.57)($\sigma_{\overline{X}}$) = 12.6 + (2.57)(.68) = 14.35
   El *límite inferior* será $\overline{X}$ - (2.57)($\sigma_{\overline{X}}$) = 12.6 – (2.57)(.68) = 10.85

   Podemos considerar que nuestra muestra, con una media de 12.6, pertenece a una población cuya media estará entre 10.85 y 14.34, y esto podemos afirmarlo con una probabilidad de error del 1%.

3. Nuestra conclusión es clara: nuestra muestra con media de 12.6 no pertenece a una población hipotética cuya media fuera 15 porque el límite máximo de la población de nuestra media es 14.35 y no llega a 15, luego nuestra muestra pertenece a *otra población* con otra media, cuyo límite inferior no es 15.

Podemos visualizar el resultado con un sencillo gráfico:

Límites probables extremos de la media de la población a la que pertenece esta muestra:		*Media de la población de referencia, término de comparación*
↓	↓	↓
10.85	14.35	15

Salta a la vista que la media de la población de referencia es mayor que el límite superior de la media de la población representada por esa muestra.

2º De hecho el procedimiento utilizado habitualmente para comprobar si la media de una muestra difiere significativamente de la media de una población suele ser otro que nos permite llegar a las mismas conclusiones. Nos basta calcular una *puntuación típica* (z), que nos dirá en cuántos *errores típicos* se aparta nuestra media de la media de la población. El procedimiento y la fórmula apropiada están puestos y explicados como un caso más del contraste de medias.

4.4. Calcular el tamaño N de la muestra para extrapolar los resultados a la población

No es éste el lugar apropiado para tratar con cierta extensión sobre el tamaño necesario de la muestra, pero sí es útil, tratando del error típico de la media o de una proporción, ver y entender en este contexto la relación entre la magnitud de los intervalos de confianza de la media y el número necesario de sujetos en la muestra para extrapolar los resultados a la población con un determinado margen de error.

De manera análoga a lo que hemos visto en [4] y en [6] el margen de error cuando la proporción encontrada en una muestra la extrapolamos a la población es:

$$\text{e } \textit{(margen de error)} = z\left[\sqrt{\frac{pq}{N}}\right] \qquad [7]$$

Si en [7] despejamos N (el tamaño de la muestra) tendremos:

$$N = \frac{z^2 pq}{e^2} \qquad [8]$$

En [8] conocemos todos los valores que nos interesan para calcular N

z Este valor corresponde al *nivel de confianza* y lo establecemos nosotros; habitualmente utilizaremos un nivel de confianza del .05 y z = 1.96 (ó z = 2.57 si nuestro nivel de confianza es de .01)

pq Es *la varianza de la población,* no la varianza de la muestra. Esta varianza no la conocemos, pero como a mayor varianza en la población hará falta una muestra mayor, nos situamos en la situación en que la varianza es la máxima posible; en este caso p = q = .50, y pq = .25, que es un valor constante.

e Es *el margen de error que estamos dispuestos a aceptar* y también lo establece el investigador. Si por ejemplo estamos dispuestos a aceptar un margen de error del 5%, esto quiere decir que si en la muestra encuestada en esta caso responde sí el 35%, en la población esperamos que responda sí entre el 30% y el 40%. Éste 5% lo expresaremos en forma de *proporción* (o *tanto por uno):* .05

Vemos de nuevo que si queremos un margen de error pequeño (*e,* el denominador en 8) necesitaremos una muestra mayor.

Podemos ver la aplicación de esta fórmula [8] con un ejemplo. Vamos a hacer una encuesta para extrapolar los resultados a una población mayor (muy grande, de tamaño indefinido).

El *margen de error* que estamos dispuestos a aceptar es del 5% (e = .05), de manera que si nos responden sí el 50% de la muestra ya sabemos que en la población el sí estará entre el 45% y el 55%

El *nivel de confianza* es del .05, que corresponde a z = 1.96

Necesitaremos una muestra de este tamaño: $N = \frac{(1.96^2)(.25)}{.05^2} = 384$ sujetos

Si el margen de error máximo que nos interesa es del 3% (e = .03), la muestra necesaria sería de 1067 sujetos.

Hacemos algunas observaciones ya que el exponer y justificar brevemente estas fórmulas tiene un valor complementario para entender mejor el concepto y utilidad del error típico, pero no tratamos aquí de manera expresa sobre el tamaño de la muestra, tipos de muestreos y cómo hacerlos, etc.[3]

a) Estas fórmulas para calcular el tamaño de la muestra son válidas aun cuando las preguntas no sean dicotómicas (estamos utilizando el error típico de una proporción, cuya varianza máxima es pq = .25).
b) Son válidas cuando se hace un muestreo aleatorio simple; hay variantes cuando se utilizan otros tipos de muestreo (como el estratificado).
c) Suponemos que la población a la que se extrapolan los resultados es grande, de tamaño indefinido y que podemos no conocer con exactitud. Con poblaciones menores y cuyo tamaño conocemos hay fórmulas más ajustadas; más o menos a partir de poblaciones en torno a los 30.000 sujetos el tamaño necesario de la muestra no varía mucho; al aumentar el tamaño de la población no aumenta proporcionalmente el tamaño necesario de la muestra.

5. Referencias bibliográficas

AMERICAN PSYCHOLOGICAL ASSOCIATION (2001). *Publication manual of the American Psychological Association* (5th Edit). Washington D.C.: Author

HERNÁNDEZ SAMPIERI, ROBERTO; FERNÁNDEZ COLLADO, CARLOS Y BAPTISTA LUCIO, PILAR (2000). *Metodología de la Investigación*. Segunda Edición. México: McGraw-Hill

MORALES VALLEJO, PEDRO. *Tamaño necesario de la muestra: ¿Cuántos sujetos necesitamos?* http://www.upco.es/personal/peter/investigacion/Tama%F1oMuestra.pdf (última revisión, 23, Nov., 2006)

RODRÍGUEZ OSUNA, JACINTO (1993). *Métodos de muestreo. Casos prácticos*. Cuadernos metodológicos. Madrid: Centro de Investigaciones Sociológicas (CIS).

SALKIND, NEIL J. (1998). *Métodos de Investigación*, 3ª edición, México: Prentice-Hall

STATPAC INC (2003) *Questionnaires & Survey Design* http://www.statpac.com/surveys/index.htm#toc (en *Sampling Methods)*

WILKINSON, LELAND AND TASK FORCE ON STATISTICAL INFERENCE *APA BOARD OF SCIENTIFIC AFFAIRS* (1999*)* Statistical Methods in Psychology Journals: Guidelines and Explanations *American Psychologist* August 1999, Vol. 54, No. 8, 594–604 (http://www.apa.org/journals/amp/amp548594.html).

[3] Puede verse más información en la bibliografía mencionada y en otras muchas publicaciones; sobre el *tamaño de la muestra* necesario también con otras finalidades (construir una escala de actitudes, hacer un análisis factorial, etc.) puede verse Morales (2006)

Anexo. Los intervalos de la media y de las proporciones en Internet

Varios programas disponibles en Internet nos dan los intervalos de confianza de una media o proporción para un determinado nivel de confianza, lo mismo que el tamaño de la muestra necesario pra determinados márgenes de error.

Algunas direcciones relacionadas con *encuestas de opinion*:

CREATIVE RESEARCH SYSTEMS. The Survey System Sample Size Calculador http://www.surveysystem.com/sscalc.htm

CUSTOMINSIGHT.COM. Survey Random Sample Calculator (Home: http://www.custominsight.com/index.asp) http://www.custominsight.com/articles/random-sample-calculator.asp

DIMENSION RESEARCH, INC. Confident Intervals for Means Calculador http://www.dimensionresearch.com/resources/calculators/conf_means.html (*home:* http://www.dimensionresearch.com/index.html**)**

Además, entre otras direcciones.

LOWRY, RICHARD, VASSARSTATS: WEB SITE FOR STATISTICAL COMPUTATION,Vassar College, Poughkeepsie, NY, USA; http://faculty.vassar.edu/lowry/VassarStats.html (*menú* en *proportions:* the confidence interval of a proportion; en *t test & procedures:* .95 and .99 Confidence Intervals for the Estimated Mean of a Population).