11

Teoría de la inferencia estadística

En el capítulo anterior se trabajó con muestras tomadas de poblaciones donde sus parámetros se conocían, o bien, se podían tomar todas las muestras posibles de tamaño k. Esto se hizo con fines didácticos porque en la práctica es muy raro encontrarse en tal situación. Lo más común es no conocer la población y el único camino que queda es tomar muestras para poder estudiarla. En este capítulo se presenta la Teoría de Estimación o Inferencia Estadística, para mostrar cómo se pueden obtener estimadores de los parámetros desconocidos de una población cualquiera. Se muestran sus requisitos y también como calcular los llamados *intervalos de confianza*. Estos son intervalos dentro de los cuales, se tiene una cierta probabilidad de encontrar al verdadero valor poblacional de la magnitud biológica en estudio. La idea general es realizar muestreos de poblaciones para luego inferir los parámetros de la misma. No importa si se conoce a priori o cuál es la forma de la función distribución de probabilidades en la población. Si se toman muestras aleatorias e independientes, lo suficientemente grandes, los estadígrafos muestrales que se calculen tendrán una función distribución gaussiana (ver Tabla 4) que permitirá efectuar estimaciones de los respectivos parámetros poblacionales y las probabilidades asociadas a las mismas.

11.1 Introducción

Las razones para efectuar una estimación en una población, en lugar de estudiarla directamente, pueden ser: que el tamaño de la población sea infinito, que el muestreo sea destructivo, que la población sea finita pero demasiado grande, y otras razones como costo o tiempo. Por esto parece más práctico tomar muestras. Hay dos maneras básicas de hacer estimaciones:

Estimaciones puntuales: se estima el parámetro desconocido con un solo valor.

Estimaciones por intervalos: se estima el parámetro desconocido con un intervalo, el cual tiene asociado una cierta probabilidad de ocurrencia.

Siempre es más aconsejable estimar con intervalos en lugar de hacerlo con un solo valor puntual, por las mismas razones que tienen los granjeros para no poner todos los huevos en una misma canasta. En efecto, suponiendo que un paciente concurre a varios bioquímicos a efectuarse el mismo análisis y estos le informan, por ejemplo: Bioq. A = 85 mg/dl; Bioq. B = 75 mg/dl; Bioq. C = 90 mg/dl. El paciente podría llegar a desconfiar de estos valores por la diversidad de resultados obtenidos: a simple vista le parecerían diferentes. Por su parte, el médico que encargó el trabajo podría mandar a hacer de nuevo el análisis. Muy diferente hubiese sido la situación si en sus informes hubieran usado intervalos como: Bioq. A: (85 ± 15)mg/dl; Bioq. B: (75 ± 4) mg/dl y el Bioq. C: (90 ± 20) mg/dl. En este caso hay coincidencia entre los tres informes, A dice entre 70 y 100, B dice entre 71 y 79 y C dice entre 70 y 110. El intervalo informado por B cae dentro de los otros dos y el de A cae dentro del intervalo informado por C. Esto es, hay una zona donde los tres informes coinciden, cosa imposible de ver si se hubieran usado estimadores puntuales. Por su parte, en la industria farmacéutica lo normal es trabajar con límites superiores e inferiores en las composiciones químicas de los medicamentos, tanto para el proceso productivo como para su posterior control de calidad. Otro tanto ocurre en la industria alimenticia o de cosméticos. La idea es siempre informar o establecer la probabilidad de que el verdadero valor caiga dentro del intervalo informado. La forma más *prudente* de informar es usando intervalos.

11.2 Estimas con intervalo

Las estimaciones por intervalo de un parámetro poblacional μ desconocido dan idea de la precisión y exactitud de la inferencia efectuada, junto con la probabilidad de que tal estimación sea cierta. Se calculan a través de los llamados: *intervalos de confianza*. Estos se construyen con la función probabilística del modelo estadístico adoptado para realizar la estimación. Se comienza con el modelo de Gauss empleado generalmente para las *grandes muestras*. En efecto, si la distribución muestral de un estadígrafo es la de Gauss de parámetros (μ_e; σ_e), porque la cantidad de muestras aleatorias e independientes tomadas es lo suficientemente grande, entonces con la Tabla 4 se pueden obtener las distribuciones de varios estadígrafos *e* y las probabilidades asociadas con la tabla de Gauss tipificada. Usando estas tablas se pueden hacer los dos tipos de estimaciones vistos.

Una *estimación puntual* se logra con: $\mu_e = \mu$ donde, por ejemplo, si *e* es la media aritmética, el valor μ_e es el promedio de las medias muestrales bajo la condición de tener una cantidad de muestras superior o por lo menos igual a 30. Este tipo de mediciones no es exacto sino que hay una cierta incertidumbre al hacerla, llamada *error típico de estimación*: σ_e ,= SE(*e*), dado por el error estándar de la distribución muestral multiplicada por un cierto factor. Todo esto se muestra en la segunda columna de la Tabla 10.1. Para el caso de la media es: SE(*e*) = $\sigma_e = \sigma / \sqrt{n}$. Entonces, se puede construir la *estimación por intervalos* con:

$$\boxed{\mu \in (\mu_e \pm Z_\alpha . \sigma_e)} \qquad \text{(a)}$$

Otra manera de informar lo mismo es calculando los límites superior e inferior del intervalo con: LS = $\mu_e + Z_\alpha . \sigma_e$. Y el inferior es LI = $\mu_e \pm Z_\alpha . \sigma_e$. Entonces, la relación (a) se escribe como

$$\boxed{\mu \in (LI \; ; \; LS)} \qquad \text{(b)}$$

Donde Z_α es un valor llamado *coeficiente de confianza.* La relación anterior (a) establece que el parámetro desconocido μ está contenido dentro de un intervalo dado por su mejor estimación puntual μ_e más o menos su error típico de estimación σ_e, multiplicado por un coeficiente relacionado con la confianza de que esto sea cierto. La relación (b) establece que el valor verdadero μ está contenido dentro de un intervalo conformado por dos límites, el inferior y el superior. Y la probabilidad de que μ se encuentre dentro de ese intervalo viene dada por el coeficiente de confianza adoptado. Al valor α se lo denomina *nivel de significación* y por lo general es elegido por el investigador. Los usos y costumbres establecidos en Biología establecen tres límites habituales: 95%, 99%, y 99,9%. En la Física o Química, donde los materiales de trabajo son más estables que los biológicos, los requerimientos son menores, llegando a lo sumo al 99%. En casi todas partes el nivel usual de trabajo es del 95%. Estos valores de *Confianza* determinan el *nivel de significación* α, y el coeficiente de confianza Z_α que se obtiene de la curva de Gauss. En la Tabla 11.1 se muestran algunos coeficientes de uso frecuente. A veces, se necesitan intervalos diferentes, de tipo abierto, donde no se necesitan dos valores extremos sino solo uno. En tal caso:

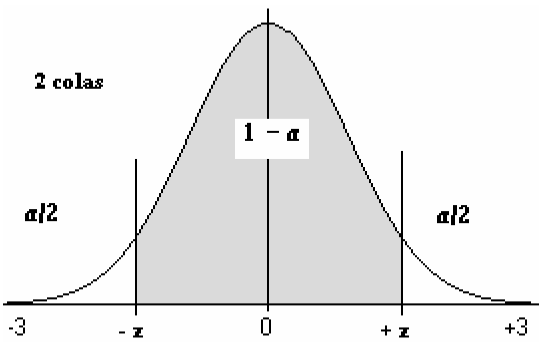
$\mu \in (-\infty\ ;\ (\mu_e + Z_\alpha \cdot \sigma_e)]$ O bien, $\mu \in [\ (\mu_e - Z_\alpha \cdot \sigma_e)\ ;\ +\infty)$

Estos intervalos abiertos por izquierda o por derecha se denominan de *una cola*, mientras que los comunes para la estimación de parámetros poblacionales son de *dos colas*. En la Figura 11.1 se muestran las dos formas usuales de una y de dos colas, con los respectivos valores de z_α en las Tablas 11.1 y 11.2.

TABLA 11.1: Niveles de confianza más usados.

Niveles de significación	Nivel de confianza	Coeficiente de confianza
α	**NC = (1-α) . 100**	Z_α
0,0005	99,95%	3,29
0,0010	99,90%	3,09
0,0013	99,87%	3,00
0,0050	99,50%	2,58
0,0100	99,00%	2,33
0,0227	97,72%	2,00
0,0250	97,50%	1,96
0,0500	95,00%	1,645
0,1587	84,13%	1,00

Figura 11.1: Probabilidad normal presentada en una y dos colas.



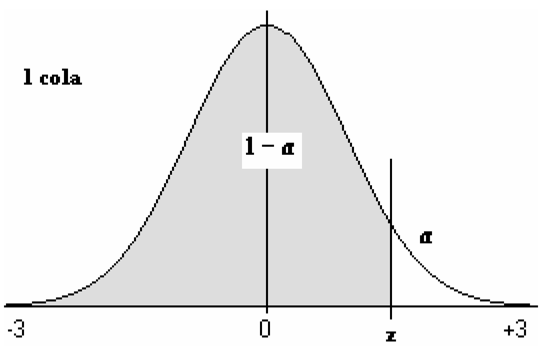


TABLA 11.2: Coeficientes de confianza más usados Z_α.

	99,90%	99,00%	95,00%
1 cola	3,09	2,33	1,645
2 colas	3,29	2,58	1,96

En clínica la notación de los intervalos de confianza para un nivel de confianza *NC* % es:

$$\text{CI } NC\,\% \ (\mu_e \pm Z_\alpha \,.\, \text{SE}(e))$$

Por ejemplo, para un 95% de confianza se expresa con:

$$\boxed{\text{CI } 95\% \ ([\mu_e - 1{,}96 \,.\, \text{SE}(e)] \ ; \ [\mu_e + 1{,}96 \,.\, \text{SE}(e)])} \qquad \text{(c)}$$

Esta es la notación recomendada para los trabajos en Medicina desde hace unos años para expresar los resultados de investigaciones en publicaciones científicas. Los tres tipos de notaciones expresados con las relaciones (a), (b) y (c) son equivalentes.

11.3 Intervalos para la media y proporciones.

Para el caso de la **media aritmética**, el intervalo se construye buscando en la Tabla 10.1 el error típico de estimación para la media y reemplazando en la ecuación (a) vista más arriba:

$$\mu \in (\overline{X} \pm Z_\alpha\, \sigma / \sqrt{n})$$

El nivel de significación adoptado hará variar el ancho del intervalo de confianza. Como puede verse en la Tabla 14.1, a medida que aumenta el Nivel de Confianza aumenta el coeficiente de confianza Z_α mientras los demás términos no se alteran. Por su parte, cuanto mayor sea el intervalo, mayor será el error $\Delta x = Z_\alpha\, \sigma / \sqrt{n}$ (recordando la expresión general del error vista antes). Esto significa que se tendrá una menor precisión. Se puede formular ahora una regla general para todos estos casos: a *mayor confianza, menor precisión*. Notar que el supuesto básico para poder usar esta relación es que el valor σ es conocido. Como generalmente esto no es así, se puede usar el valor muestral DS para estimar σ si la muestra es lo suficientemente grande, y para este caso la condición aceptable es $n \geq 30$.

Conviene desarrollar ejemplos para ilustrar esta tesitura, como sigue:

Ejemplo 1) Suponiendo que a un paciente se le extrae una muestra de sangre y al suero obtenido se lo fracciona en 50 alícuotas, luego a cada una se le determina la creatinina, y con los valores medidos se obtienen un promedio de 10 mg/dl y un desvío de 2,2 mg/dl. El verdadero valor de la creatinina en el paciente se puede estimar con un nivel de confianza del 95 % ($Z_\alpha = 1{,}96$) con:

$$\mu \in (10 \pm 1{,}96 \,.\, 2{,}2 \ / \ \sqrt{50}) \text{ mg/dl} = (10{,}0 \pm 0{,}6) \text{ mg/dl} \quad \text{O bien: CI 95\% (9,4 ; 10,6)}$$

Eso significa que se tiene una probabilidad del 95 % de encontrar la creatinina real del paciente entre 9,4 y 10,6 mg/dl. Si se quiere aumentar la confianza al 99% el nuevo intervalo tendrá una mayor indeterminación, o sea, el intervalo será más ancho: entre 9,2 y 10,8 mg/dl.

$$\mu \in (10 \pm 2{,}58 \,.\, 2{,}2 \ / \ \sqrt{50}) \text{ mg/dl} = (10{,}0 \pm 0{,}8) \text{ mg/dl} \quad \text{O bien: CI 99\% (9,2 ; 10,8)}$$

Y si todavía se aumenta un poco más al 99,9%:

$\mu \in (10 \pm 3{,}29 \,.\, 2{,}2 \;/\; \sqrt{50})$ mg/dl $= (10 \pm 1)$ mg/dl O bien: CI 99,9%(9 ; 11)

Ejemplo 2) Se toman 36 muestras de cierto antibiótico y se les mide su contenido de Sulfato de Neomicina, resultando un promedio de 92 mg con un desvío de 6 mg. Se desea saber el intervalo al 99,9% de la variable medida, tal que el contenido de Sulfato sea menor o igual al valor medio.

Se trata de un caso de intervalo de una sola cola. El 99,9% del área bajo la curva normal tipificada debe quedar a la izquierda de un valor crítico (como el dibujado en el Gráfico 14.1). O sea,

$\mu \in (-\infty \,;\, (\mu_e + z_\alpha \,.\, \sigma_e)] = (-\infty \,;\, (92 + 3{,}09 \,.\, 6 \,/\, \sqrt{36})]$ O bien: CI 99,9% (- ∞ ; 95,09)

Esto significa que hay un 99,9% de confianza de que el contenido de Sulfato de Neomicina del medicamento examinado nunca supere 95,09 mg.

Si se hubiese tratado de determinar el intervalo de estimación del verdadero contenido del Sulfato en el medicamento, entonces se debería calcular su intervalo de confianza con:

$\mu \in (92 \pm 1{,}96 \,.\, 6 \;/\; \sqrt{36})$ mg/dl $= (92 \pm 2)$mg/dl O bien: CI 95% (90 ; 94,9)
$\mu \in (92 \pm 2{,}58 \,.\, 6 \;/\; \sqrt{36})$ mg/dl $= (92{,}0 \pm 2{,}6)$mg/dl O bien: CI 99% (89,4 ; 94,6)
$\mu \in (92 \pm 3{,}29 \,.\, 6 \;/\; \sqrt{36})$ mg/dl $= (92{,}0 \pm 3{,}3)$mg/dl O bien: CI 99,9% (88,7 ; 95,3)

Ejemplo 3) Se realizaron 200 exámenes de Bioestadística. Si se toman 50 de ellos al azar y se obtiene una media de 7,5 puntos con un desvío de 1 punto: ¿cuánto vale el intervalo de estimación para la media de los 200 exámenes con una confianza del 95%?

En este problema, la población no es grande en función de la muestra, entonces conviene efectuar la corrección para estos casos vista en el capítulo anterior:

$\mu_{\bar{x}} = \mu = 7{,}5$ y $\sigma_{\bar{x}} = 0{,}123$ pues:

$\sigma^2_{\bar{x}} = (\sigma^2 \,/\, n) \; [(N_P - n) \,/\, (N_P - 1)] = [(1)^2 \,/\, 50] \; [(200 - 50) \,/\, (200 - 1)] = 0{,}0151$

Con esos datos se puede obtener: $\mu \in (7{,}5 \pm 1{,}96 \,.\, 0{,}123) = (7{,}50 \pm 0{,}24)$

O bien: CI 95% (7,26 ; 7,74)

Para el caso de las **proporciones** el intervalo queda determinado por:

$$\mu \in \left[\, p \pm Z_\alpha \,.\, \sqrt{\frac{\pi\,.\,(1-\pi)}{n}} \,\right]$$

Donde p es la proporción de éxitos obtenida en la muestra y π es la proporción de éxitos esperada o real en la población de referencia. Hay casos donde π es conocida como π = ½ en el lanzamiento de una moneda, π = 1/6 al lanzar un dado, π ≈ ½ en el sexo de un recién nacido, etc. Pero cuando no se sepa el verdadero valor de p, entonces hay que estimarlo con la proporción de éxitos en la muestra $p \approx \pi$ si y solo si, la muestra es lo suficientemente grande ($n > 40$)

Ejemplo 4) Se lanza una moneda al aire 100 veces y se obtienen 56 caras. El intervalo de confianza del 99 % para la probabilidad de sacar cara con tal moneda se calcula con:

$p = 56/100 = 0,56$ y $\pi = 0,5$ El desvío es: $\sigma_\pi = \sqrt{\dfrac{\pi.(1-\pi)}{n}} = \sqrt{\dfrac{0,5\,.\,0,5}{100}} = 0,05$ O sea,

$\mu \in (\,0,56 \pm 2,58\,.\,0,05\,) = (\,0,56 \pm 0,13\,)$ O bien: CI 99% (0,43 ; 0,69)

Expresado de otra forma: $\mu \in (56 \pm 13)$ %. Esto es, hay una probabilidad del 99 % de obtener entre 43 y 69 caras, si se arroja 100 veces esa moneda al aire.

Ejemplo 5) En un estudio por cohortes se tuvo 50 casos de enfermos en un total de 200 sujetos. Se esperaba una prevalencia del 20% en la enfermedad; hallar el intervalo de confianza:

$p = 50/200 = 0,25$ y $\pi = 0.2$. El desvío es: $\sigma_\pi = \sqrt{\dfrac{\pi.(1-\pi)}{n}} = \sqrt{\dfrac{0,2\,.\,0,8}{200}} = 0,03$ O sea,

$\mu \in (0,25 \pm 1,96\,.\,0,03)$ O bien: CI 95% (0,19 ; 0,31)

11.3.1 Intervalos "a priori" de media y proporciones.

Por intervalos "a priori" se entienden aquellos intervalos de confianza que se pueden obtener *antes* de efectuar el experimento. La idea es saber antes de medir en que intervalo se espera que caiga el resultado obtenido si todo está bien, o sea de acuerdo a las expectativas previas. Simplemente, se trata de rescribir el intervalo definido con las ecuaciones (a), (b) y (c) ya vistas:

$$\boxed{\mu_e \in (\mu \pm Z_\alpha\,.\,\sigma_e)} \qquad \text{(d)}$$

Por ejemplo para el caso de la media aritmética sería:

$$\overline{X} \in (\mu \pm Z_\alpha\, \sigma / \sqrt{n})$$

Y para las proporciones:

$$p \in [\,\pi \pm Z_\alpha\,.\,\sqrt{\frac{\pi.(1-\pi)}{n}}\,]$$

El significado de estos intervalos es estimar dentro de que valores se espera que caiga el valor muestral, antes de realizar las mediciones del experimento. Notar que esto solo puede aplicarse cuando se conocen "a priori" los valores poblacionales, por ejemplo cuando se está usando un suero control, de referencia o estándar en una medición clínica (donde μ es el valor aceptado). O bien cuando se conoce la probabilidad π de la población como en el caso de una moneda, el sexo de un recién nacido, las proporciones mendelianas, etc. Como ser:

Ejemplo 6) Suponiendo que se tiene un suero patrón cuyo valor de creatinina es μ = 12 mg/dl y con una especificación del estándar de σ = 1,22 mg/dl. Si se fracciona dicho suero en 36 alícuotas y se mide a cada una de ellas con el sistema habitual de laboratorio, se desea saber en que rango de valores debe caer la media aritmética de las mediciones a realizar, si el método clínico está bien calibrado, con un nivel de confianza del 95 %.

$$\overline{X} \in (\mu \pm Z_\alpha\, \sigma / \sqrt{n}) = [12 \pm 1{,}96 \,.\, (1{,}22 / 6)] = (12{,}0 \pm 0{,}4)$$

O bien, se espera que $\overline{X}$ caiga dentro del intervalo (11,6 ; 12,4) si el método de medición está funcionando bien, con una confianza del 95%.

Ejemplo 7) En la determinación del SIDA se usa un método cuya especificación establece que tiene una Sensibilidad del 90% y una Especificidad del 80%. Si se van a efectuar 49 determinaciones, averiguar los valores reales que se esperan obtener al 95%.

π = 0.9 y el desvío es: $\sigma = \sqrt{\dfrac{\pi\,.\,(1-\pi)}{n}} = \sqrt{\dfrac{0{,}9\,.\,0{,}1}{49}} = 0{,}043$ O sea

$p \in (0{,}9 \pm 1{,}96 \,.\, 0{,}043)$ O bien: Sensibilidad esperada es CI 95% (0,98 ; 0,82)

Mientras que la Especificidad esperada se puede calcular con:

π = 0.8 y el desvío es: $\sigma = \sqrt{\dfrac{\pi\,.\,(1-\pi)}{n}} = \sqrt{\dfrac{0{,}8\,.\,0{,}2}{49}} = 0{,}057$ O sea

$p \in (0{,}8 \pm 1{,}96 \,.\, 0{,}057)$ O bien, la Especificidad estaría en CI 95% (0,91 ; 0,69)

11.4 Intervalos para el desvío estándar y la varianza.

Para el caso del **desvío estándar** se puede construir el intervalo de confianza recurriendo a la Tabla 4 para obtener el error típico de estimación y así:

$$\mu \in (DS \pm Z_\alpha \,.\, \frac{\sigma}{\sqrt{2\,n}})$$

Y para la **varianza** es:

$$\mu \in (DS^2 \pm Z_\alpha \,.\, \sigma^2 \sqrt{\frac{2}{n}})$$

En ambos casos se usa la aproximación $\sigma \approx DS$ para resolver los problemas prácticos, siempre y cuando la muestra sea lo suficientemente grande ($n \geq 100$). Para ilustrar su aplicación se tomarán los datos de problemas anteriores:

Ejemplo 8) Usando los datos de los ejemplos 1 y 2, pero para n = 144 sería:

(1) $\sigma \in$ (2,2 $\pm$ 1,96 . 2,2 / $\sqrt{2.144}$) mg/dl = (2,20 $\pm$ 0,25) mg/dl. O bien CI 95% (2,45; 1,95)

$\sigma^2 \in$ (4,84 $\pm$ 1,96 . 4,84 $\sqrt{\frac{2}{144}}$) mg/dl = (4,84 $\pm$ 0,13) mg/dl. O bien CI 95% (4,71 ; 4,97)

(2) $\sigma \in$ (6,0 $\pm$ 1,96 . 6 / $\sqrt{2.144}$) mg/dl = (6,0 $\pm$ 0,7) mg/dl. O bien CI 95% (5,3 ; 6,7)

11.5 Propiedades de un estimador

Los conceptos básicos desarrollados en esta sección son para el caso de estimaciones paramétricas, pero se pueden extender también al caso no-paramétrico. Sea una población cualquiera de la cual se extraen *n* muestras aleatorias e independientes; los valores obtenidos de cada muestra conforman un conjunto que tendrá una cierta distribución muestral que se usará para estimar el parámetro desconocido: μ. Se desea que el estimador *e* sea bueno. El problema es que se entiende por buen estimador y la respuesta es: *La distribución muestral de e debe concentrarse lo más posible alrededor del valor verdadero del parámetro* μ *y su dispersión debe ser lo más pequeña posible.*

La técnica más importante para hallar estimadores se debe a R.A. Fisher y es el *Método de Máxima Verosimilitud (Likelihood).* Consiste en hallar una *función de verosimilitud* para cada caso y así elegir, como estimador de valor poblacional desconocido, al valor que hace máxima a esta función. Su desarrollo matemático excede los límites del presente trabajo quien desee verlo con más detalle puede recurrir a texto, más específicos como los de H. Crámer, etc. Por ahora, basta mencionar que el *estimador máximo verosímil* de un índice de tendencia central es la *media* aritmética y de un índice de dispersión es el *desvío estándar*.

Las propiedades clásicas de los estimadores son:

Estimadores insesgados: un estimador se dice insesgado cuando su valor es igual al correspondiente valor poblacional. Como la media o la mediana para la media poblacional. Caso contrario, se lo llama *estimador sesgado* como la media de una distribución muestral de varianzas.

Cuando el valor esperado del estadígrafo Ξ (*e*) = μ el estimador es insesgado.
Por ejemplo:

$$\Xi\,(\overline{X}) = \mu_{\overline{X}} = \mu$$

$$\Xi\,(\text{Med}) = \mu_{\text{Med}} = \mu$$

Sin embargo, $\Xi\,(\sigma^2) = \Xi\,[\,\sum(x_i - \mu)^2 / n] = \mu_\sigma{}^2 = (n-1) / n\,.\sigma^2$ (poblacional)

Es un estimador sesgado pues (n-1) / n siempre diferente de la unidad y eso sesga la estimación. Sin embargo, el sesgo se elimina cuando para el cálculo de la varianza se usa la fórmula

$\Xi\,(DS^2) = \Xi\,[\,\sum(x_i - \mu)^2 / (n\text{-}1)] = \mu_\sigma{}^2 = \sigma^2$ (muestral)

Estimador eficiente: si las distribuciones muestrales de dos estadígrafos tienen la misma esperanza (o media), aquel que tenga menor error de estimación (varianza) será el más eficiente de ambos. En otro caso, se dice que el *estimador es no eficiente.*

Por ejemplo, las distribuciones muestrales de la media aritmética y de la mediana tienen el mismo valor esperado, que es la media poblacional. Sin embargo, la varianza de la distribución muestral de medias es menor que la de la mediana. Por lo tanto, el estimador eficiente es la media aritmética. En efecto:

$SE(\bar{x}) = \sigma_{\bar{x}} = \sigma / \sqrt{n}$ mientras que SE (Mediana) = $(1{,}2536)\ \sigma / \sqrt{n}$ O sea $\sigma_{\bar{x}} < \sigma_{Med}$

Existen otros métodos para la obtención de los estimadores como el de los momentos desarrollados por K. Pearson, el de Ráo-Cramer o el de suficiencia para funciones de densidad conjunta de probabilidad. Sin embargo, para los alcances de esta obra basta saber que los estimadores máximos verosímiles cumplen a la vez con las propiedades de ser insesgados, eficientes y suficientes. En resumen, para efectuar estimaciones apropiadas, cuando se pueda se deben usar las estimaciones por intervalo antes que las puntuales y se deben emplear los estimadores máximos verosímiles para construirlos. Los niveles de confianza para el caso de variables biológicas se adoptan por costumbre en 95 %, 99 % y 99,9 %, y se calculan siempre los tres para poder informar en las publicaciones científicas, como se podrá ver mejor en el próximo capítulo. En Medicina, lo habitual es usar el nivel del 95% para expresar los resultados, lo que se hace extensivo a sus disciplinas derivadas como: Epidemiología, Inmunología, y otras.

11.6 Intervalos de confianza para dos muestras

Sean dos estadígrafos e_1 y e_2, extraídos de poblaciones normales o con función distribución de tipo gaussiana, el intervalo de confianza para la diferencia de sus respectivos parámetros poblacionales se obtiene con:

$$\mu_{e_1 - e_2} = (e_1 - e_2) \pm Z_\alpha . \sigma_{1,2} = (e_1 - e_2) \pm Z_\alpha . \sqrt{\sigma_1^2 + \sigma_2^2}$$

Donde el SE $(e_1 - e_2) = \sigma_{1,2}$. Los dos casos más interesantes son la media y las proporciones. Se reemplazan los valores de los estadígrafos y los de sus respectivos errores típicos (ver Tabla 4) en la fórmula anterior y se puede obtener para la media aritmética:

$$\mu_{\bar{x}_1 - \bar{x}_2} = (\bar{x}_1 - \bar{x}_2) \pm Z_\alpha . \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Cuando ambas muestras tienen la misma varianza $\sigma^2 = \sigma^2_1 = \sigma^2_2$ es:

$$\mu_{\bar{x}_1 - \bar{x}_2} = (\bar{x}_1 - \bar{x}_2) \pm Z_\alpha . \sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Y si además ambas muestras tienen el mismo tamaño muestral $n = n_1 = n_2$ es:

$$\mu\,\bar{x}_1 - \bar{x}_2 = (\bar{x}_1 - \bar{x}_2) \pm Z_\alpha \,.\, \sigma \sqrt{\frac{2}{n}}$$

Por su parte para la diferencia de proporciones es:

$$\mu\, p_1 - p_2 = (p_1 - p_2) \pm Z_\alpha \,.\, \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}$$

Cuando ambas muestras fueron extraídas de la misma población, o cuando hay razones para creer que los valores poblacionales son iguales, esto es $\pi = \pi_1 = \pi_2$ entonces la ecuación se reduce a:

$$\mu\, p_1 - p_2 = 0 = (p_1 - p_2) \pm Z_\alpha \,.\, \sqrt{\pi\ (1-\pi)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}$$

Cuando el valor poblacional π es desconocido se lo puede estimar con la aproximación:

$\pi \approx (n_1\ p_1 + n_2\ p_2) / (n_1 + n_2)$ que es el promedio ponderado de las proporciones muestrales

Para ilustrar estas aplicaciones se presentan los ejemplos siguientes:

Ejemplo 1) Se tomaron 200 muestras aleatorias de presión sistólica a niños cuyos padres son hipertensos, obteniéndose una media de 107 y un desvío de 7. Luego se tomaron 100 muestras de niños cuyos padres tienen la presión sanguínea normal, y se obtuvo una media de 98 con un desvío de 6. Obtener los límites de confianza del 95 % a la diferencia de medias.

En este caso se trata de una diferencia de medias, pero con varianzas diferentes estimadas con las muestras de la manera siguiente:

$$\mu\,\bar{x}_1 - \bar{x}_2 = (\bar{x}_1 - \bar{x}_2) = 107 - 98 = 9 \ \text{ y SE(1-2)} = \sigma_{1\text{-}2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{49}{200} + \frac{36}{100}} = 0{,}778$$

$\mu_{1\text{-}2} \in (9 \pm 1{,}96\ .\ 0{,}778) = (9{,}0 \pm 1{,}5)$. O bien CI 95% (7,5 ; 10,5)
$\mu_{1\text{-}2} \in (9 \pm 2{,}58\ .\ 0{,}778) = (9{,}0 \pm 2{,}0)$. O bien CI 99 % (7 ; 11)
$\mu_{1\text{-}2} \in (9 \pm 3{,}29\ .\ 0{,}778) = (9{,}0 \pm 2{,}6)$. O bien CI 99,9 % (6,4 ; 11,6)

Ejemplo 2) Se inoculan dos organismos aislados durante dos epidemias distintas, a dos muestras diferentes obtenidas de la misma población. A las dos semanas se enferman el 68,5 % de las 200 pruebas realizadas con el primer organismo, y el 65,3 % de las 150 pruebas efectuadas para el segundo caso. Hallar un intervalo de confianza de esta diferencia de proporciones encontradas.

En este caso se trata de una diferencia de proporciones y se asume que el valor verdadero de cada población se estima con los valores muestrales hallados: $\mu_{\pi1-\pi2} = p_1 - p_2 = (68{,}5 - 65{,}3)\% = 3{,}2\%$

$$\sigma_{\pi1-\pi2} = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}} = \sqrt{\frac{0{,}685.0{,}315}{200} + \frac{0{,}653.0{,}347}{150}} = 0{,}051 = 5{,}1\%$$

Entonces es: $\mu_{1\text{-}2} \in (\mu_{\pi1-\pi2} \pm Z_{\alpha} . \sigma_{\pi1-\pi2}) = (3{,}2 \pm 1{,}96\ 5{,}1)$ % O bien CI 95% (-6,8 ; 13,2). Y esto significa que hay un 95% de probabilidades de que el verdadero valor de la diferencia de los porcentajes esté entre -6,8% y 13,2%.

Ejemplo 3) Imaginado para el problema anterior que el supuesto básico es que no hay diferencia entre las dos poblaciones, calcular el intervalo de confianza para el valor verdadero poblacional π

Con este supuesto será $\pi = \pi_1 = \pi_2 = (n_1\ p_1 + n_2\ p_2) / (n_1 + n_2) = 67{,}13\%$.Y con esta estimación se puede calcular el desvío con:

$$\sigma \approx DS = \sqrt{\pi\ (1-\pi)\left[\frac{1}{n_1}+\frac{1}{n_2}\right]} = 4{,}35\%$$

Entonces la estimación será $\pi \in (67{,}13 \pm 1{,}96 . 4{,}35)$ O bien: CI 95% (58,60 ; 75,66) %

11.7 Intervalo para el cociente de dos proporciones

El intervalo de confianza para el cociente de dos proporciones se logra haciendo una transformación de variables. Sea el cociente $\Phi = \pi_1 / \pi_2$, entonces tomando el logaritmo neperiano de esa variable (ln Φ), se deduce que tiene una distribución de probabilidades aproximadamente normal si la muestra es lo suficientemente grande. Entonces, tomado el antilogaritmo se pueden obtener los límites del intervalo de confianza, con la ecuación:

$$\Phi\ e^{[\pm 1{,}96\ SE\ (\ln \Phi)]}$$

Donde el valor esperado es $\Xi(\Phi) = \mu_{\Phi} \cong \Phi$ se transforma en: $\Xi(\ln \Phi) \cong \ln \Phi$ y el error estándar en: SE (ln Φ). De esta forma usando los antilogaritmos es que los límites se pueden calcular tomado el signo positivo para el superior y el negativo para el inferior, como se muestra en los siguientes ejemplos clínicos:

- *Likelihood Ratio de positivos*:

Este índice clínico se calcula con: LR+ = S / (1 - E) = (vp / TE) / [1 – (vn / TS)]

O lo que es lo mismo:

LR+ = (vp / TE) / (fp / TS) que es un cociente de dos proporciones.

Donde el error estándar se calcula con:

$$SE\ [\ln (LR+)] = \sqrt{(1/vp) - (1/TE) + (1/fp) - (1/TS)}$$

Límite superior = (LR+) $e^{[+ 1{,}96\ SE\ (\ln LR+)]}$

Límite inferior = (LR+) $e^{[- 1{,}96\ SE\ (\ln LR+)]}$

- *Likelihood Ratio de negativos:*

Este índice clínico se calcula con: LR- = (1 – S) / E) = [1 - (vp / TE)] / (vn / TS)

O lo que es lo mismo:

LR- = (fn / TE) / (vp / TS) que es un cociente de dos proporciones.

Donde el error estándar se calcula con:

SE [ln (LR-)] = $\sqrt{(1/\text{fn}) - (1/\text{TE}) + (1/\text{vp}) - (1/\text{TS})}$

Límite superior = (LR-) e $^{[+ 1,96 \text{ SE } (\ln \text{LR-})]}$

Límite inferior = (LR-) e $^{[- 1,96 \text{ SE } (\ln \text{LR-})]}$

- *Riesgo Relativo*:

Este índice clínico se calcula con: RR =EER / CER

O lo que es lo mismo:

RR = [a / (a + b)] / [c / (c + d)] que es un cociente de dos proporciones.

Notar que en el caso del diagnóstico a = vp, b = fp, c = fn y d = vn. Y el error estándar será:

SE [ln (RR)] = $\sqrt{(1/a) - (1/(a+b) + (1/c) - (1/c+d)}$

Límite superior = (RR) e $^{[+ 1,96 \text{ SE } (\ln \text{RR})]}$

Límite inferior = (RR) e $^{[- 1,96 \text{ SE } (\ln \text{RR})]}$

- *Odds Ratio*:

OR = (a / c) / (b / d) = (a d) / (c b)

Es un cociente de dos proporciones pero hay una diferencia, no se trata del cociente de una frecuencia con respecto a un total marginal, sino respecto de otra frecuencia. Por eso el cálculo del error estándar es ligeramente diferente:

SE [ln (OR)] = $\sqrt{(1/a) + (1/b) + (1/c) + (1/d)}$

Límite superior = (OR) e $^{[+ 1,96 \text{ SE } (\ln \text{OR})]}$

Límite inferior = (OR) e $^{[- 1,96 \text{ SE } (\ln \text{OR})]}$

Usando el ejemplo numérico del capítulo 4 (Pág. 23), se pueden calcular: OR = 7,5 y OR = 5,6.

Los intervalos al 95% de confianza resultan:

Para OR = 7,5 será 95% CI (3,7 ; 15,4)
Para RR = 5,6 será 95% CI (2,9 ; 10,7)

Usando otro ejemplo numérico del capítulo 4 (Pág. 9) para el caso de 1000 muestras es:

Para LR+ = 3,6 será 95% CI (3,2 ; 4,1)
Para LR- = 0,13 será 95% CI (0,07 ; 0,24)

Notar que estos intervalos no son simétricos. Los índices no caen en la mitad del intervalo.

11.8 Problemas propuestos

1) Marcar la respuesta correcta a cada una de las afirmaciones siguientes, o completar la frase:

1) Una estimación de tipo puntual es más recomendable que una por intervalos. **V** **F**
2) Un intervalo de confianza se arma con el estimador y su error típico. **V** **F**
3) La fórmula para obtener un intervalo de confianza es: .... **..............**
4) El mejor estimador es el máximo verosímil. **V** **F**
5) La fórmula general para la diferencia de los estadígrafos de dos muestras es: . ...
6) Un estimador es insesgado cuando su valor esperado coincide con el valor poblacional **V** **F**
7) Los errores típicos de estimación se pueden obtener de ....
8) Un estimador es más eficiente que otro cuando su error típico es mayor. **V** **F**
9) Un estimador máximo verosímil es insesgado y eficiente a la vez. **V** **F**
10) La probabilidad asociada a un intervalo de confianza se relaciona con la significación. **V** **F**
11) El valor coeficiente de confianza se calcula con α y la tabla de Gauss. **V** **F**
12) El intervalo de confianza para la diferencia de medias es:... ....
13) El intervalo de confianza para la diferencia de proporciones es: ....
14) A mayor nivel de confianza, mayor ancho del intervalo. **V** **F**
15) La relación entre el Nivel de significación y el Nivel de confianza es: ....
16) El intervalo de confianza para el cociente de dos proporciones es simétrico. **V** **F**
17) El error estándar del OR es muy similar al del RR. **V** **F**

2) La vida útil de un medicamento A se midió con 200 muestras y dio una media de 1400 días con un desvío de 120 días. El medicamento B se determinó con 100 muestras y se obtuvo una media de 1200 días con un desvío de 80 días. Hallar el intervalo de confianza para la diferencia de vidas útiles de ambos medicamentos, al nivel del 95%, 99% y 99,9%.

3) El colesterol de una paciente medido 30 veces dio una media de 256 mg/dl y un desvío de 32 mg/dl. Encontrar el intervalo de confianza para el 95%.

4) A la muestra de suero de la paciente del problema anterior se la divide en 40 alícuotas y se le mide el colesterol con un nuevo método clínico. Su media es de 280 mg/dl y su desvío de 40 mg/dl. Encontrar el intervalo de confianza para el 99%.

5) Calcular la diferencia de medias entre el método clínico y el nuevo, usando los datos de los problemas 3 y 4 anteriores. Fijarse si el valor cero cae dentro del intervalo. Explicar los resultados obtenidos desde el punto de vista de la exactitud.

6) Calcular la diferencia de desvíos estándar encontrados en los problemas 3 y 4, expresada mediante un intervalo de confianza. Explicar los resultados desde el punto de vista de la precisión en mediciones clínicas.

7) Para hacer un test VDRL se emplea el kit A en 200 pacientes y se encuentran 120 pacientes que dieron un resultado (-). Los aciertos en (+) fueron 70 y 110 en (-). Con estos datos calcular la Sensibilidad y Especificidad de A y B. Si con los mismos pacientes se empleó el Kit B y resultaron 140 pacientes con (-), mientras el número de aciertos en (+) fue de 55 y en (-) 130. Obtener los intervalos de confianza al 95% para cada uno de los dos kits usados.

8) Se midieron 50 alícuotas de un suero con el Kit de Glucosa marca 1, y otros 50 con la marca 2. Los resultados fueron:

Marca 1: entre (80-80,9) 8 veces, (81-81,9) 22 veces, (82-82,9) 15 veces y 5 entre (83-83,9);

Marca 2: entre (80-80,9) 6 veces, (81-81,9) 30 veces, (82-82,9) 10 veces y 4 entre (83-83,9).

a) Obtener los intervalos de confianza para las medias de cada método y para los desvío estándar, a un nivel del 95% y del 99%.
b) Si el suero era uno de control con un valor de 79 mg/dl, determinar si cae dentro o fuera de los intervalos de confianza calculados en el punto anterior.
c) Si la precisión máxima admisible para una glucosa es del 8% de acuerdo con los criterios de Aspen. Comparar este valor límite contra los hallados en cada marca.

9) Se sabe que la concentración de coloides en el agua de un río es del 5,4 /ml. Se tomó una muestra del mismo a la que se le agregó un reactivo. Luego de un rato se extrajo una pequeña cantidad para analizarla en una cámara de recuento. Los resultados fueron:

Nº de coloides:	2	3	4	5	6	7	(nº/ml)
Frecuencia:	10	20	170	100	90	10	

Calcular los intervalos de confianza para el 95% y el 99%.

10) En un recuento de rojos en cámara se espera obtener 4,5 millones por cm3 de cierto paciente. Se le toma una muestra de sangre y los resultados hallados son:

Nº de rojos por millón	2	3	4	5	6
Frecuencia	10	90	110	150	40

Detectar si el valor esperado del recuento cae dentro o fuera del intervalo de confianza para el 95% de confianza. Discutir los resultados.

11) Calcular los intervalos de confianza al 95% de todos los índices clínicos de calidad

	Resultado verificado		
Test	Enfermo	Sano	Total
(+)	200	40	TP = 240
(-)	10	550	TN = 560
Total	TE = 210	TS = 590	800

Lo mismo para los índices de riesgo diagnósticos (RR y OR).