

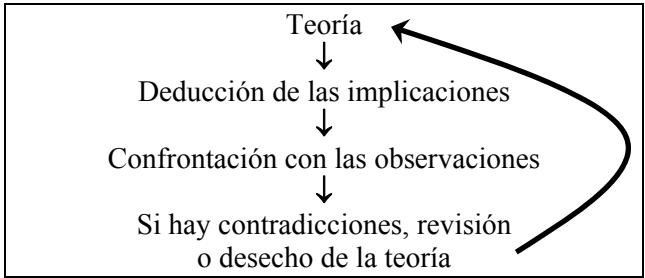
CAPÍTULO 2-2 LA INDUCCIÓN ESTADÍSTICA

2-2.1 LA INDUCCIÓN ESTADÍSTICA EN EL MÉTODO CIENTÍFICO: MODELOS TEÓRICOS Y MODELOS ALEATORIOS⁸⁸

Estudiaremos después la lógica fundamental de las pruebas de hipótesis, la cual es el punto final de la inducción estadística. Mientras, antes, es importante situar la inducción estadística en el marco del método científico. Nuestro punto inicial será un esquema simplista del método científico hipotético-deductivo.⁸⁹

⁸⁸ Malinvaud (1969), caps. 1 y 2 y Blalock (1972), caps. 2 y 8.

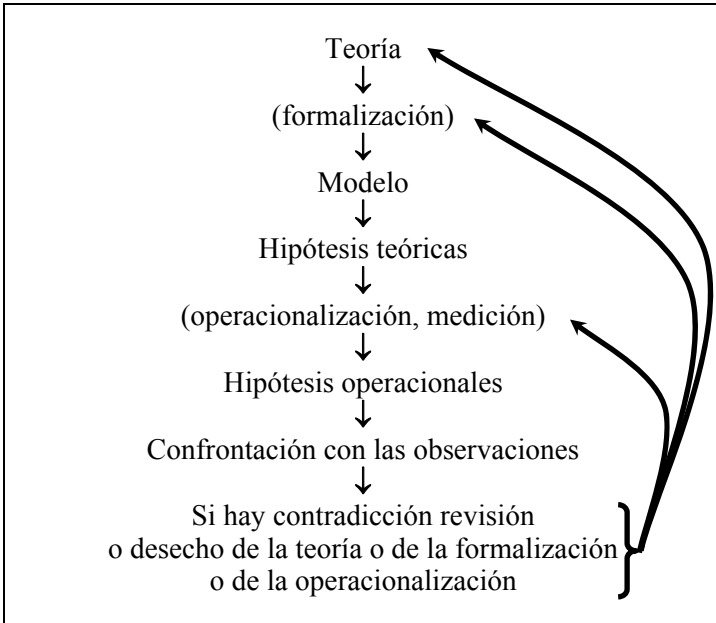
⁸⁹ Debemos reconocer que el esquema es truncado. Para tener acceso a una presentación más detallada del proceso de desarrollo del conocimiento, vea Robert (1993). Lo que se presenta ahora es más bien lo que Kuhn, en *La estructura de las revoluciones científicas* (1983), llama la ciencia “normal”.



Tanto en ciencias sociales como en ciencias físicas, se formaliza, a menudo, las teorías como modelos: “Un modelo consiste en la representación formal de ideas o de conocimientos con relación a un fenómeno” (Malinvaud, 1969, p 45).

Cuando, como pasa muchas veces, se selecciona una representación formal del tipo matemático, tenemos un modelo matemático.

Una teoría es una construcción intelectual de carácter hipotético, es decir, una hipótesis global emitida con relación a un fenómeno. De igual manera, un modelo, y también las partes del modelo, son una representación formal del fenómeno. No obstante, tanto en el modelo como en la teoría, no se definen los conceptos en términos operacionales, esto es, no se pueden confrontar las hipótesis teóricas directamente con las observaciones. Para esto es necesario traducir las hipótesis teóricas en hipótesis operacionales, lo que implica definir mediciones (vea el capítulo 1-1). En estas condiciones, el esquema complejo es el siguiente:



En este esquema es fácil observar que una contradicción entre las observaciones y las predicciones de la teoría puede venir no solamente de la teoría misma o también de su formalización o de su operacionalización. En la realidad el proceso de la investigación no es tan fluido como lo enseña el primer esquema...

La estadística interviene al momento de la confrontación con las observaciones visto que, en la mayoría de los casos, se efectúan éstas sobre un muestreo cuando el modelo abarca una población. Ahora bien, la mayoría de los modelos matemáticos son deterministas, o sea que las variables que representan los conceptos se vinculan entre sí gracias a relaciones funcionales (funciones matemáticas) que no tienen ningún elemento aleatorio.

Sin embargo, la relación entre una muestra y la población de donde proviene, es esencialmente aleatoria puesto que cada muestra no es más que una de muchas muestras posibles.

Para confrontarlos con las observaciones, se necesita complementar los modelos deterministas para tomar en cuenta este elemento aleatorio. Al combinar un modelo determinista y un modelo de la relación aleatoria de la muestra con la población, obtenemos un modelo aleatorio (probability model).⁹⁰ Para traducir esta distinción, algunos autores designan como “modelo estructural” (structural model) el modelo teórico determinista y como “modelo muestral” (sampling model) el modelo de relación entre la muestra y la población.⁹¹

Hasta el momento mencionamos una sola fuente de lo aleatorio. De hecho, existen tres “puertas” por las cuales lo aleatorio se introduce en los modelos:⁹²

1. Primero, existe el carácter aleatorio que ya se mencionó de la relación entre la muestra y la población de donde proviene.

⁹⁰ “[...] para todo conjunto de valores que se dan a las variables exógenas, un modelo aleatorio define la ley de probabilidad correspondiente a las variables endógenas” (Malinvaud, 1969, p. 59).

⁹¹ Más precisamente, Upton y Fingleton (1985, p. 264) nombran “structural model” a la especificación de la relación funcional entre la variable dependiente y las variables independientes; nombran “sampling model” a la hipótesis con relación a la distribución de probabilidad de la variable dependiente (o, lo que equivale, del término de error).

⁹² Malinvaud escribe: “Sabemos que se justifica el empleo del cálculo de las probabilidades para el análisis de los datos estadísticos con una u otra de las dos consideraciones siguientes: o bien se asimila el fenómeno estudiado con un proceso que tenga una determinación aleatoria de algunas magnitudes; se consideran, entonces, estas magnitudes como aleatorias en el universo (NDLR: es decir, en la población) como en la muestra observada. O bien, la selección de los elementos observados resulta de un sorteo aleatorio; entonces, la composición de la muestra es aleatoria, y por lo tanto, los datos obtenidos lo son también aunque se refieran a variables no aleatorias.” (Malinvaud, 1969, p. 62). Malinvaud prosigue diciendo que, en el contexto de la econometría, la primera consideración se adapta mejor.

2. Segundo, las variables operacionales son mediciones imperfectas de los conceptos, así que podemos considerar que el error de medición es aleatorio (es decir, determinado al azar). Por lo tanto, se puede representar con un modelo aleatorio, la influencia de los errores de medición que intervienen en el momento de la traducción de las hipótesis teóricas en hipótesis operacionales (los modelos de la “teoría de los errores” en ciencias físicas fueron de por sí unos de los primeros modelos aleatorios).
3. Finalmente, percibimos algunos fenómenos como aleatorios de por sí y no se pueden representar adecuadamente con modelos teóricos no aleatorios. El azar en estos modelos constituye un concepto que encierra unas veces una indeterminación fundamental (como en física de las partículas), otras, una multitud de factores que no se pueden observar (como pasa en la mayoría de los casos en ciencias sociales),⁹³ y cuyas manifestaciones aparecen como consecuencias de la aplicación de las leyes de probabilidad.

De todas maneras, un modelo aleatorio tiene un carácter hipotético puesto que se apoya en hipótesis sobre la estructura aleatoria, es decir sobre las leyes de probabilidad que rigen el azar. Durante la confrontación con las observaciones, estas hipótesis no se cuestionan (por lo menos no todas). Se consideran, por decirlo así, como el peaje que se exige para cruzar el puente de lo conocido a lo desconocido, dado que la inducción estadística va “más allá” de los datos observados. Sin embargo, aunque la inducción se basa en hipótesis, existe una ganancia epistemológica cuando las hipótesis que fun-

⁹³ En particular, pensamos en los modelos de utilidad aleatoria (random utility) subyacentes a los modelos de selecciones discretas (discrete choice) logit, probit, etc. Vamos a encontrar este tipo de modelos en el apartado 4-3.

damentan la inducción son menos restrictivas que los resultados obtenidos por medio de la inducción.

En resumen, lo mismo si la simplicidad del esquema que presentamos con anterioridad aparenta lo contrario, debemos reconocer que la confrontación de la teoría, de los modelos y de las hipótesis con las observaciones es, pocas veces, total. Cada ejercicio de confrontación se basa, de hecho, en un modelo más general que no se cuestiona. Esto es ciertamente el caso de la inducción estadística y de los tests de hipótesis, temas que trataremos más adelante. En efecto, los tests de hipótesis se aplican, casi siempre, a unas formas particulares de un modelo teórico general que no se cuestiona y se basan en un modelo aleatorio que no se cuestiona tampoco.⁹⁴

2-2.2 ALGUNOS CONCEPTOS CLAVE DE LA TEORÍA DE LAS PROBABILIDADES*

Antes de estudiar propiamente la inducción estadística, es necesario recordar, aunque de manera resumida, las definiciones de algunos conceptos claves de la teoría de las probabilidades.

⁹⁴ Una “confrontación total” sería propia de una revolución científica a la Khun, Sin embargo, es dudoso que la inducción estadística tenga un papel predominante en el proceso de cambio de paradigma de una revolución científica. No obstante, es cierto que existen tests “de nivel superior”, por así nombrarlos, que se aplican a ciertos aspectos del modelo aleatorio. Pero, estos mismos tests se basan en modelos aleatorios más generales los cuales, a este punto, no se cuestionan. Podemos imaginar un test del modelo aleatorio del test del modelo aleatorio... Antes bien, poco importa la “altura” del nivel al cual llegamos, siempre existirá un nivel superior donde el modelo de muestreo no se cuestiona.

* Referencias: Wonnacott y Wonnacott (1992, caps. 3 y 4, apartados 4.1-4.2)

2-2.2.1 Conceptos fundamentales

Azar (de la palabra árabe *az-zahr*, “el dado”). El Diccionario Enciclopédico Planeta lo define como “suceso que se presenta fortuitamente sin venir motivado por intención o plan alguno”.

*Evento aleatorio.*⁹⁵ Evento cuya realización o no depende del azar. Por ejemplo, en el caso de las muestras que se pueden sortear de una población, cada posibilidad es un evento aleatorio. Cuando se sortea una muestra, solo uno de estos eventos se realiza mientras que los demás no se realizan.

Variable aleatoria. Es una variable cuyo valor es el resultado de eventos aleatorios.⁹⁶ Puesto que el resultado del sorteo de una muestra es un evento aleatorio, todas las mediciones que se pueden efectuar sobre una muestra son variables aleatorias. Esto se aplica tanto en los datos brutos como en las estadísticas calculadas a partir de estos datos.

Distinguimos las variables aleatorias discretas que no pueden más que tomar ciertos valores (números enteros en la mayoría de los casos), y las variables aleatorias continuas, cuyo valor puede ser cualquier número real en un intervalo dado (abierto o cerrado). Las variables aleatorias continuas forman un conjunto de posibilidades infinitas, cuando las va-

⁹⁵ “Alea” significa dado de jugar en latín. Recordemos el famoso “Alea Jacta est” (se jugaron los dados), significa la suerte está echada”, de Julio Cesar al momento de cruzar el Rubicon.

⁹⁶ En la mayoría de los manuales de estadística, la distinción entre la variable aleatoria y sus valores posibles o observados se manifiesta a través de una simbolización donde X es la variable aleatoria y x sus valores posibles o observados. En nuestro contexto, escribiremos “variable aleatoria” textualmente cuando será necesario; de otra manera, usaremos una x para designar los dos.

riables aleatorias discretas pueden formar un conjunto de posibilidades finito cuando su campo de variación es finito.⁹⁷

Probabilidad de un evento aleatorio. Todos tenemos una noción intuitiva de lo que es una probabilidad, sin embargo, no es dar de este concepto una definición rigurosa. Se puede tratar la noción de probabilidad de tres maneras.

Podemos concebir la probabilidad en el contexto de una serie de “experimentos” o de “pruebas”, donde el resultado de cada intento es un “éxito” (el evento acontece) o de una “falla” (el evento no acontece); ésta es la definición “frecuentista” de la probabilidad en términos de frecuencias relativas de un evento aleatorio. Durante una serie de experimentos de este tipo (sortear cara o cruz o echar los dados), se define la probabilidad de un evento aleatorio (como obtener “sol” o un “seis”) como la proporción de los experimentos cuando este evento se realiza en promedio.

Se puede definir la probabilidad de un evento como la suerte que pensamos tenga un evento de acontecer en una escala de 0 a 100% (definición subjetiva o bayesiana).

Finalmente, podemos considerar el concepto de probabilidad como primero y no definible para luego enunciar un sistema de axiomas al cual se debe conformar cualquier medición de probabilidad.

2-2.2.2 Distribuciones de probabilidad

Función de distribución de probabilidad o distribución de probabilidad. Es una correspondencia que asocia una probabilidad a cada evento de un conjunto exhaustivo de eventos que sean mutuamente exclusivos (posibilidades). Por ejem-

⁹⁷ Una variable aleatoria cuyo campo de variación es el conjunto de los enteros naturales es una variable discreta; sin embargo, el conjunto de sus valores posibles es infinito.

plo, cuando jugamos a cara o sol una vez con una moneda que no es trucada, la función de probabilidad es

$$\text{Prob}(\textit{cara}) = \text{Prob}(\textit{sol}) = 0.5$$

La distribución de probabilidad se parece a una distribución de frecuencia relativa, pero se distingue en el hecho de que la distribución de frecuencia específica frecuencias observadas cuando la distribución de probabilidad asigna a cada evento la frecuencia relativa que tendría en promedio en un contexto de una serie infinita de experimentos (vea más arriba, la definición “frecuentista” de la probabilidad).

Función de distribución acumulada de una variable aleatoria. La función de distribución acumulada de una variable aleatoria es una función $F(x)$ (una correspondencia) la cual, por cada valor posible de x de la variable aleatoria, da la probabilidad de que la variable aleatoria tome un valor inferior o igual a x .⁹⁸

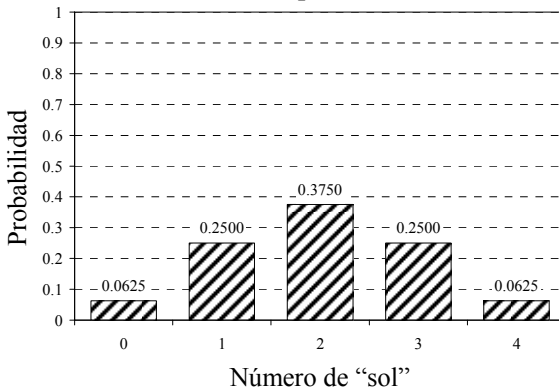
Por ejemplo, si sorteamos cara o sol cuatro veces, el número de veces que obtenemos cara es una variable aleatoria discreta cuya función de probabilidad y función de distribución acumulada se representan en la tabla que sigue (esta distribución se llama distribución binomial). Las figuras que acompañan la tabla ilustran las nociones de distribución de probabilidad y de función de distribución acumulada.

⁹⁸ En caso que los valores de la variable aleatoria no sean numéricos – como “cara” o “sol” – es necesario definir con anterioridad el orden en el cual se acomodan estos valores para que la relación “inferior o igual” tenga sentido.

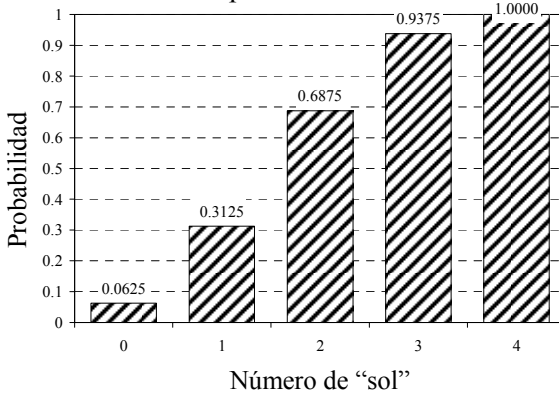
Función de distribución acumulada de una variable aleatoria

Número de "sol"	Probabilidad	Probabilidad acumulada	
x_i	$f(x_i)$	$F(x_{i-1})$	$F(x_i)$
0	1/16		1/16
1	4/16	+ 1/16	= 5/16
2	6/16	+ 5/16	= 11/16
3	4/16	+ 11/16	= 15/16
4	1/16	+ 15/16	= 16/16

Función de probabilidad



Función de probabilidad acumulada



2-2.2.3 Distribución de muestreo*

El concepto de distribución de muestreo (sampling distribution) es primordial en inducción estadística. Es la forma operacional que toma el modelo de muestreo (el modelo de la relación entre una muestra y la población; vea 2-2.1).

En efecto, una distribución de muestreo es una distribución de probabilidades asociada a una estadística. Recordemos que una estadística es una característica de una muestra, mientras que un parámetro es una característica de una población.

Ahora bien, vimos que una muestra no es más que una de las muestras del mismo tamaño que se podría obtener de la población estudiada. Por consiguiente, según la muestra obtenida, la estadística podría tomar valores diferentes. Y, puesto que la muestra se obtiene al azar, el valor de la estadística es aleatorio y la estadística misma es una variable aleatoria. La distribución de muestreo de la estadística es su distribución de probabilidad en la población de las muestras de un tamaño específico que se pueden obtener al azar en una población estudiada.

En general, la distribución de muestreo de una estadística depende de los parámetros de la población estudiada. Es esta dependencia la que permite, a partir del valor observado de una estadística, formular enunciados probabilistas con relación con los parámetros. Explicitaremos este proceso cuando tratemos el tema de los tests de hipótesis.

Por ejemplo, imaginemos que queremos saber si una moneda que usamos para jugar a cara o sol es trucada. Si la moneda no es trucada, debería “en promedio” caer con la misma frecuencia sobre cara o sol. Sin embargo, para conocer el verdadero promedio, se tendría que lanzar la moneda un número infinito de veces porque, independientemente del nú-

* Referencias: Wonnacott y Wonnacott (1992, p. 224-226).

mero de lanzes, nunca estaremos seguros del resultado de los lanzes suplementarios que podríamos efectuar. La población estudiada es, por lo tanto, infinita. La única manera que queda para decidir si debemos considerar si la moneda es trucada o no es efectuar un cierto número de lanzes, calcular la proporción de cara y sol y aceptar la moneda como honesta si esta proporción (frecuencia relativa) es lo suficientemente cercana a 50%. La distribución de muestreo de esta proporción es la distribución de probabilidad de esta estadística. Esta distribución depende del número de lanzes y del verdadero valor de la probabilidad.

Diagrama 1

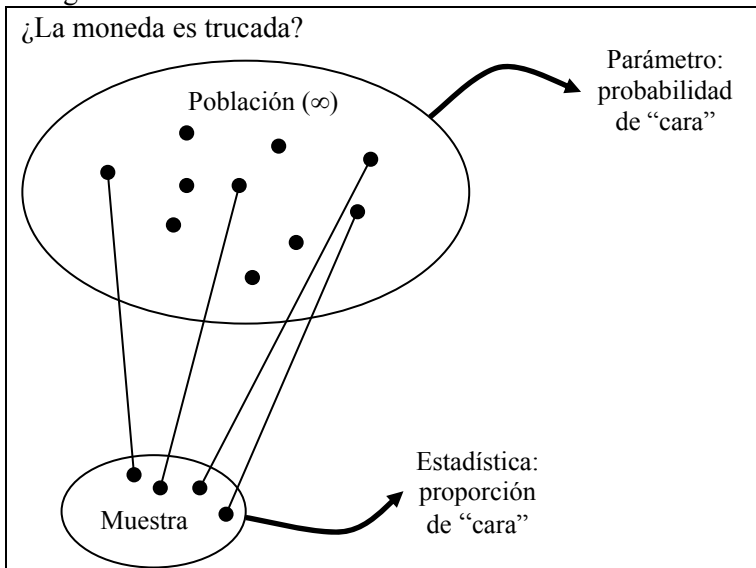
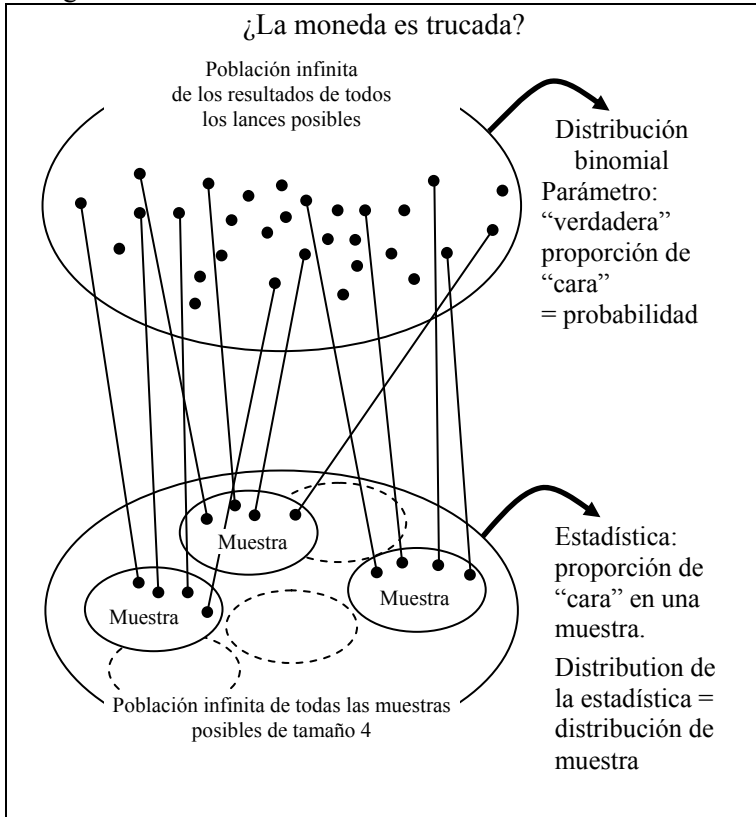


Diagrama 1b



2-2.2.4 Variables aleatorias continuas: función de densidad de probabilidad y esperanza matemática

Función de densidad de probabilidad de una variable aleatoria continua*

Con el fin de entender correctamente la lógica de los tests estadísticos, es de suma importancia captar la diferencia entre una función de densidad de probabilidad y una función de distribución de probabilidad (aunque la expresión más usual, “distribución de probabilidad” se aplique indiferentemente para los dos).

Son variables aleatorias continuas. Ahora bien, con una variable aleatoria continua el número de valores posibles es infinito. En consecuencia, la probabilidad de que la variable aleatoria pueda tomar un valor específico es normalmente infinitamente pequeña; en otras palabras, no podemos asociar una probabilidad a cada valor posible de la variable aleatoria, así que el concepto de función de probabilidad como lo definimos anteriormente no se aplica.

Es la razón por la cual, cuando tratamos con variables continuas, es necesario recurrir a la noción de función de densidad de probabilidad. Se define la función de densidad de probabilidad a partir de la función de probabilidad acumulada puesto que, aunque la probabilidad que la variable aleatoria pueda tomar un valor x específico es infinitamente pequeña, existe una probabilidad positiva⁹⁹ que su valor no rebese este valor x : tenemos por lo tanto una función

$$F(x) = \text{Prob (variable aleatoria} \leq x)$$

* Referencias: Wonnacott y Wonnacott (1992, p. 138-140) proponen otra presentación de la función de densidad de probabilidad.

⁹⁹ Es decir, no infinitamente pequeña.

Lo cual no es otra cosa que la distribución acumulada en función de los valores posibles de x .

Por ejemplo, se puede considerar el tiempo de vida de un foco eléctrico incandescente como una variable aleatoria continua. ¿Cuál es la probabilidad de que el foco dure exactamente 112 horas, 23 minutos, 14 segundos y tres centésimas? Es fácil entender que esta probabilidad es infinitamente pequeña. Sin embargo, la probabilidad que este foco dure 112 horas, 23 minutos, 14 segundos y tres centésimas o menos es, de seguro, positiva. Esta última probabilidad es la probabilidad acumulada

$$F(x) = \text{Prob}(\text{tiempo de vida del foco} \leq x),$$

donde $x = 112 \text{ hrs. } 23 \text{ mn } 14.03 \text{ s.}$

En resumen, con una variable aleatoria continua, el concepto de función de probabilidad tal como se definió más arriba no se aplica, pero la función de distribución acumulada existe por lo general. Y es a partir de la función de distribución acumulada $F(x)$ que se define la función de densidad $f(x)$; es una función que, por cada valor posible de la variable aleatoria, da la tasa (velocidad, densidad) a la cual aumenta la probabilidad acumulada en este punto de la función. Técnicamente, la función de densidad de probabilidad es una derivada (una pendiente) de la función de distribución acumulada de una variable continua:¹⁰⁰

$$f(x) = \frac{d}{dx} F(x)$$

Con el fin de entender correctamente los tests estadísticos, es de suma importancia captar la diferencia entre una función

¹⁰⁰ Con relación en la función de distribución acumulada, la densidad juega el mismo papel que la velocidad con relación a la distancia recorrida: en un gráfico de la distancia recorrida en función del tiempo transcurrido, la pendiente de la curva da la velocidad en ese instante. La derivada es la velocidad instantánea, que es diferente de la velocidad promedio en un intervalo dado la cual corresponde en una gráfica de la distancia recorrida en la pendiente promedio en ese intervalo.

de densidad de probabilidad y una función de distribución de probabilidad, porque los tests son, de hecho, razonamientos sobre las probabilidades y se calculan estas probabilidades con la ayuda de funciones de densidad de probabilidad. Ahora bien, la ordenada de una función de densidad (su altura) no es una probabilidad (cuando la ordenada de una función de probabilidad sí es una probabilidad). Por lo contrario, la superficie debajo de una curva de una función de densidad es una probabilidad; técnicamente esto es cierto ya que si $f(x)$ es la derivada de $F(x)$ entonces se obtiene $F(x)$ con la integral de $f(x)$.

De esta manera,

$$\text{Prob}(a \leq \text{variable aleatoria} \leq b) = F(b) - F(a)$$

$$\text{Prob}(a \leq \text{var. aleatoria} \leq b) = \left(\begin{array}{l} \text{superficie debajo} \\ \text{de } f(x) \text{ entre } a \text{ y } b \end{array} \right)$$

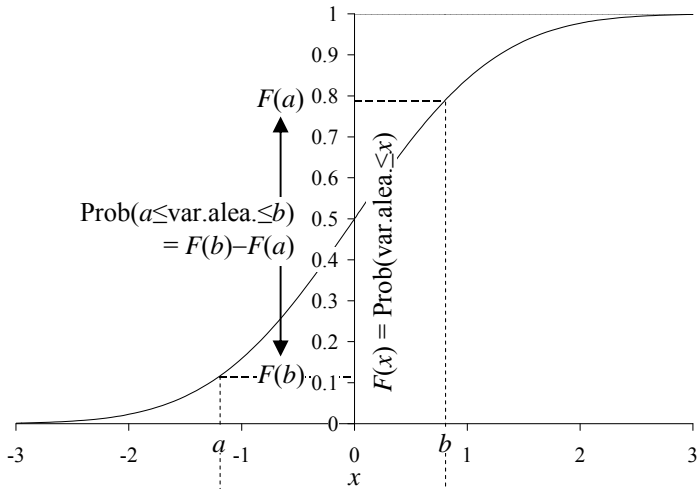
$$\text{Prob}(a \leq \text{variable aleatoria} \leq b) = \int_a^b f(x) dx$$

Naturalmente,

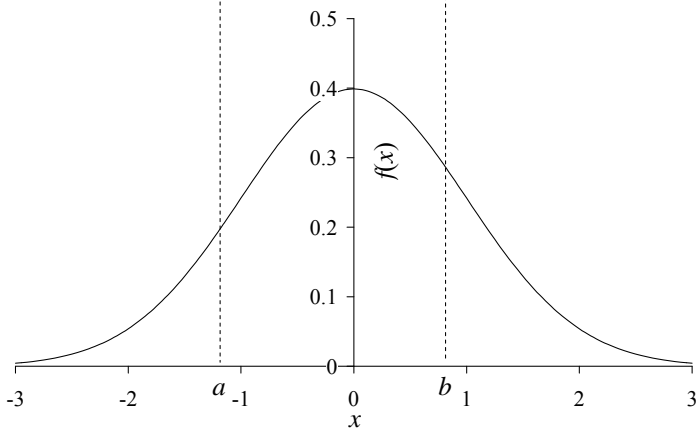
$$\int_{-\infty}^{+\infty} f(x) dx = 1 = F(+\infty)$$

Se ilustra en la figura 1 la relación entre la función de probabilidad acumulada y la función de densidad de probabilidad de una variable aleatoria continua.

Figura 1
 Función de probabilidad acumulada y función de densidad
 Función de distribución acumulada



Función de densidad de probabilidad



Esperanza matemática*

El promedio de una variable aleatoria continua en una población infinita no puede calcularse por medio de la famosa fórmula:

$$\mu_x = \frac{1}{n} \sum_{i=1}^n x_i$$

(donde x_i son los valores posibles de la variable aleatoria) simplemente porque el número de valores posibles n es infinito. El concepto de esperanza matemática es una generalización del promedio. Para una variable continua, la esperanza matemática es¹⁰¹

$$E(x) = \int_{-\infty}^{+\infty} f(x) x dx$$

Así, cuando hablamos del promedio de una variable aleatoria continua en una población, nos referimos a

$$\mu_x = E(x)$$

Y cuando hablamos de la varianza de una variable aleatoria continua en una población, nos referimos a:

$$\sigma_x^2 = E\{[x - E(x)]^2\} = \int_{-\infty}^{+\infty} f(x) [x - E(x)]^2 dx$$

En el marco de este curso, sólo es necesario acordarse que las fórmulas de cálculo del promedio y de la varianza se pueden generalizar en el caso de una variable aleatoria continua. Por lo que resta, la intuición que se tenga del concepto del promedio y de la varianza a partir de las fórmulas de la estadística descriptiva es suficiente.

* Referencias: Wonnacott y Wonnacott (1992, pp. 154-155, 184-185).

¹⁰¹ Si queremos comparar las dos fórmulas, se puede decir que \int juega el papel de \sum y $f(x)$ el papel de $(1/n)$.

Ley normal*

La ley normal es un ejemplo de la distribución de probabilidad de una variable aleatoria continua. Es una distribución cuya función de densidad tiene la forma de una campana simétrica, como lo enseña la figura 2. Esta distribución es una buena aproximación de varias distribuciones de probabilidad observadas de manera empírica. Es, también, la distribución asintótica hacia la cual tienden muchas otras distribuciones (en el tema de distribución asintótica, vea 3.2).

Una de las características más importantes de las distribuciones normales es no tener más que dos parámetros: el promedio y la desviación estándar. Esto significa que, en caso de saber que una variable tiene una distribución normal y de conocer su promedio y su desviación estándar, se conoce entonces perfectamente su función de densidad de probabilidad.

Además, si la variable x tiene una distribución normal con un promedio μ_x y una desviación estándar σ_x , entonces la variable “estandarizada”

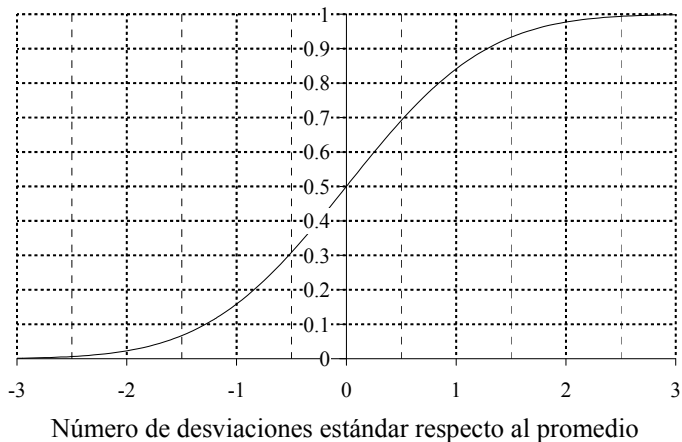
$$z = \frac{x - \mu_x}{\sigma_x}$$

posee una distribución normal de promedio 0 y desviación estándar 1.¹⁰²

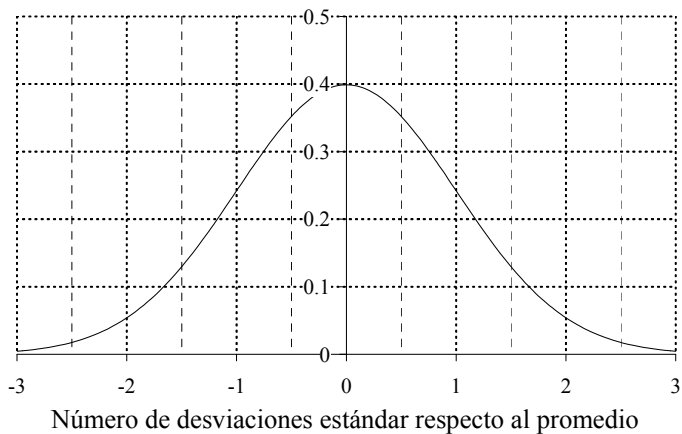
* Referencias: Wonnacott y Wonnacott (1992, pp. 142-148).

¹⁰² Para diferentes valores de la variable normal estándar, las tablas estadísticas dan el valor de la densidad de probabilidad (ordenada de la función de densidad normal) y de la probabilidad acumulada (ordenada de la función de distribución acumulada normal). Estas tablas nos informan también de lo inverso, a saber, de el valor de la variable normal estándar que corresponde a diferentes probabilidades acumuladas. Se encuentra lo equivalente en el software Excel con las funciones *NORMSDIST* y *NORMSIMV*, las funciones que corresponden para las variables normales no estándares son *NORMLIST* y *NORMINV*.

Figura 2
Distribución normal
Función de distribución normal cumulativa
 $F(x)$



Función de densidad normal
 $f(x)$



2-2.3 MUESTREO, ESTIMACIÓN Y TESTS DE HIPÓTESIS

En grandes rasgos, el proceso de la inducción consiste en usar las estadísticas relativas a una muestra con el fin de aprender algo con relación a los parámetros de la población.

Este proceso conlleva tres tipos de preguntas:

- ¿Cuáles propiedades conviene que la muestra tenga? ¿Bajo cuáles condiciones es razonable considerar esta muestra como “representativa”?
- ¿Cuáles son las estadísticas de la muestra que se pueden usar para estimar el valor de los parámetros de la población? ¿Cuáles son las propiedades de estos estimadores?
- ¿Cómo se puede evaluar la fiabilidad de las estimaciones obtenidas? De manera más general, ¿qué se puede afirmar de la población y con qué grado de confianza?

Con relación a este tercer tipo de preguntas, los principios epistemológicos subyacentes a la inducción conducen a los tests de hipótesis.

Se puede, por lo tanto, dividir la inducción estadística en tres partes:

- La muestra.
- La estimación.
- Los tests de hipótesis.

2-2.3.1 Muestrario *

Un plan de muestreo debe contestar las preguntas siguientes:

1. ¿Cómo seleccionar la muestra de manera a respetar las condiciones que requiere el modelo de muestreo (*sampling model*) que corresponde al modelo de relación entre la muestra y la población? O, a la inversa, ¿cuál

* Referencias: Wonnacott y Wonnacott (1992, cap. 23).

modelo de muestreo refleja correctamente el modo de seleccionar la muestra?

2. ¿Cuál tamaño de muestra es necesario para obtener la precisión y el nivel de confianza deseados?

Antes de definir el plan de muestreo, hay que escoger la unidad de observación. Por ejemplo, en una encuesta a los consumidores, la unidad de observación puede ser la persona o el hogar. El conjunto de las unidades de observación constituye la población de la que se quiere sacar una muestra. Y la población también tiene que ser bien definida. Por ejemplo, en una encuesta a los hogares, por supuesto, hay que circunscribir el universo muestrario por límites geográficos u otros. Pero hay también que circunscribir el universo como concepto: por ejemplo, en una encuesta a los hogares, ¿se quiere tomar en cuenta los hogares de una persona sola? ¿Hogares “colectivos” (prisiones, conventos, cuarteles militares)?

Selección

Mencionemos los tres tipos más importantes de muestras.

La primera distinción que debemos operar en los métodos de selección es entre aquellas que conducen a una muestra aleatoria y las demás. Una muestra aleatoria es una muestra que se constituye por medio de un método de selección que permite conocer, para cada una de las muestras posibles, cuál es la probabilidad de que se seleccione. En la mayoría de los casos, esto equivale a conocer, para cada individuo, la probabilidad que se seleccione sabiendo que las probabilidades de los individuos son independientes entre sí (la probabilidad que se seleccione a un individuo no se ve afectada en caso de que se seleccione cualquier otro individuo).

Muestra aleatoria simple (sorteada): cada individuo de la población tiene igual probabilidad de ser seleccionado. Este método de muestreo exige, por lo general, que se haga un inventario previo de la población.

Un inventario incompleto puede crear un sesgo; por ejemplo, una selección aleatoria en el directorio telefónico descarta a priori los individuos que no tengan teléfonos o cuyo teléfono es confidencial. Se cometerá un error si los descartados son diferentes a los demás. Hay otras posibilidades de sesgo en la recolección de datos. Por ejemplo, los que la ciencia política suele llamar los “electores discretos” (que no contestan o “no saben”) quizá tienen en promedio una opinión diferente de los que se expresan más espontáneamente.

Se hace la distinción entre el muestreo simple con o sin reemplazo: en el primer caso, se sortean sucesivamente los individuos miembros de la muestra y después de cada sorteo, el individuo seleccionado vuelve a ser elegible durante el sorteo que sigue (así, un individuo puede ser seleccionado más de una vez en la misma muestra); en el segundo caso, se quita el individuo que ha sido seleccionado de donde se sorteará los demás miembros de la muestra.

El muestreo sistemático es un método que se acerca del muestrario aleatorio simple.¹⁰³ Consiste en seleccionar un individuo a cada n vez (lo que supone que los casos ya sean ordenados: por ejemplo, el orden alfabético en el directorio telefónico o el orden de los números de dirección en las calles. Para sacar una muestra sistemática, se divide el tamaño de la población entre el tamaño de muestra que se quiere: eso es el intervalo muestral n . Después, basta sacar al azar el primer individuo y los demás siguen.

Muestra aleatoria estratificada: cuando la población se divide en varias subpoblaciones (“estratos”) con parámetros que pueden ser diferentes, queremos que la muestra sea representativa no sólo de la población en general sino también de la subpoblación.

Esta representatividad no implica forzosamente que cada estrato de la muestra sea proporcional a la subpoblación que

¹⁰³ May (1993), p. 70.

representa. De hecho, si buscamos la misma precisión para todas las estratos de la población, es necesario que las subpoblaciones menos numerosas tengan una mayor representatividad. Eso porque la precisión en la estimación de los parámetros no es proporcional al tamaño de la muestra. Retomaremos este tema (vea capítulo 2-3, al final del apartado 2-3.5).

Mientras que la muestra aleatoria simple exige un inventario de la población, la muestra aleatoria estratificada necesita un inventario por estrato. Eso no es siempre disponible. A menudo se trata de aproximar la muestra aleatoria estratificada gracias al muestreo por cuotas.¹⁰⁴ Este método consiste en clasificar a los individuos al momento que son seleccionados, hasta que se cumpla el número esperado (cuota) en cada estrato. En una encuesta por cuestionario basada en el muestreo por cuotas, la primera parte del cuestionario sirve para clasificar a los individuos; no se completa el cuestionario con los individuos supernumerarios.

Muestra en racimos (cluster): este método consiste en dividir la población en grupos (racimos) para luego sortear un cierto número de éstos; los miembros de estos racimos seleccionados constituyen la muestra. Recurrimos a menudo a este método cuando no se tiene un inventario previo de la población.

Por ejemplo, para efectuar una encuesta en los hogares de una zona habitacional informal (donde los datos del censo no son fiables), se puede subdividir la zona en cuadras de casas, luego seleccionar una cierta proporción éstas y proceder a la entrevista en todos los hogares en el interior de las cuadras seleccionadas. La inducción estadística es más difícil con una muestra en racimos, porque las distribuciones muestrales de las estadísticas son más complejas.

¹⁰⁴ May (1993), p. 71.

Por supuesto, hay métodos de muestreos no aleatorios que, sin embargo, son adecuados en contextos no estadísticos. Así, encuestas de tipo cualitativo se basan a veces en el método “bola de nieve” o en el método de saturación. Pero esos métodos muestrarios no son pertinentes aquí.

Tamaño

Por lo general, más grande es la muestra, más probabilidad tiene de ser representativa y más alto es el grado de precisión de la estimación para un mismo grado de confianza. Sin embargo, el grado de precisión no es directamente proporcional al tamaño de la muestra (examinaremos este fenómeno con más precisión al momento de estudiar un test de hipótesis sobre el promedio; vea 2-3.5). Según los análisis que queremos efectuar, existen reglas que permiten determinar el tamaño de la muestra requerida para alcanzar la precisión y la confianza deseadas. Pero los costos de la recolección crecen con el tamaño.

2-2.3.2 *Estimación*

- ¿Cuáles son las estadísticas de la muestra que se pueden usar para estimar el valor de los parámetros de la población? ¿Cuáles son las propiedades de estos estimadores?

En este momento, hacemos una distinción entre un estimador que es una fórmula, un método de cálculo, y una estimación o valor estimado como el resultado de la aplicación de esta fórmula. Un estimador es una variable aleatoria puesto que el mismo estimador que se aplica a datos de muestras diferentes arroja, por lo general, valores estimados diferentes.

Métodos*

Existen tres enfoques:

1. Analógico.
2. Menores cuadrados.
3. Máxima verosimilitud (Theil, 1971, p. 89, Freund, 1962, p. 223).

1. Estimación según el enfoque analógico. El principio de estimación analógica (conocido también como método de los momentos) es simple: para estimar un parámetro, se aplica a la muestra la misma fórmula matemática que a la población.

Ejemplo:

Para estimar el valor promedio μ_x de una variable x en una población, se calcula el promedio de la misma variable en la muestra.

$$m_x = \frac{1}{n} \sum_i x_i$$

Este procedimiento es totalmente mecánico. Sin embargo, en general un estimador es la expresión matemática de un principio de selección del “mejor” valor como estimación del parámetro. Diferentes principios de selección conducen a diferentes estimadores, sabiendo que los principios que más se emplean son el principio de mínimos cuadrados y el principio de máxima verosimilitud.

Se podría comparar la estimación con el hecho de sintonizar una radio: se prueba diferentes frecuencias hasta optimizar la recepción de la señal para que, al final, el valor seleccionado en el receptor sea un valor estimado del parámetro que se busca, o sea la frecuencia de emisión. La frecuencia seleccionada dependerá del criterio de selección usado (supongamos que se usa un solo criterio a la vez para

* Referencia: Wonnacott y Wonnacott (1992, cap. 18).

lograr la comparación): fuerza de la señal, ausencia de ruidos y de distorsión, ausencia de parásitos...¹⁰⁵

2. *Principio de los menores cuadrados.* El principio de menores cuadrados puede aplicarse sin necesidad de modelo aleatorio. Consiste en “sintonizar” los valores estimados de los parámetros del modelo de tal manera que, cuando se aplica este modelo a la muestra, sus errores de predicción sean tan pequeños como se pueda. La expresión “menores cuadrados” se refiere a la medición de errores como la suma de los cuadrados de los errores de predicción, sabiendo que se mide cada error con la diferencia entre un valor observado y el valor predicho correspondiente. Esta medición de errores es, por lo tanto, el cuadrado de la distancia euclidiana generalizada entre la serie de las observaciones y la serie de las predicciones.

3. *Principio del máximo de verosimilitud.* La aplicación del principio del máximo de verosimilitud se refiere directamente al modelo aleatorio seleccionado para representar la relación aleatoria entre la muestra y la población o para representar el carácter aleatorio del fenómeno estudiado. Por consiguiente y contrario al principio de los menores cuadrados, el principio del máximo de verosimilitud no puede aplicarse sin modelo aleatorio. El principio del máximo de verosimilitud consiste en “sintonizar” los valores estimados de los parámetros del modelo de manera que, suponiendo que estos valores fueran los buenos, la muestra sea la más “verosímil” posible. Se mide la verosimilitud con la función de verosimilitud, la cual es

¹⁰⁵ Hay que tener cuidado con seguir demasiado lejos esta analogía puesto que cuando sintonizar un radio cuya frecuencia no se conoce se hace por lo general a tuestas, la aplicación de uno u otro de los principios de estimación conduce, con frecuencia, a una fórmula que permite calcular directamente el valor estimado correspondiente.

la función de densidad de probabilidad de la muestra tomando los valores de los parámetros.

Cuando maximizamos la función de verosimilitud, los papeles de los valores observados de la muestra y de los parámetros en la función de densidad de probabilidad son invertidos: en lugar de considerar los valores observados como variables aleatorias cuya función de densidad de probabilidad depende del valor de los parámetros, son, al contrario, los valores observados que se consideran como fijos y se hace variar los valores estimados de los parámetros de tal manera que la verosimilitud alcance su máximo. Los valores seleccionados como valores estimados de los parámetros son, por lo tanto, los valores cuando la densidad de probabilidad de la muestra es la más grande (el modo de la distribución) y, por consiguiente, los intervalos alrededor de este punto tienen la probabilidad más alta.

En ciertas condiciones, los principios de los menores cuadrados y del máximo de verosimilitud conducen al mismo estimador. En algunos casos (como la estimación del promedio), este estimador es al mismo tiempo el estimador del proceso analógico.

Propiedades deseables*

Los estimadores son variables aleatorias de modo que sus propiedades son las propiedades de su distribución de muestreo.

1. Ausencia de sesgo. Entre las propiedades deseables de un estimador, la ausencia de sesgo es de suma importancia. Un estimador no sesgado es un estimador cuyo valor será en promedio igual al valor del parámetro estimado. La expresión

* Referencias: Wonnacott y Wonnacott (1992, pp. 262-266, 275-276); Freund (1962, p. 215-220).

“en promedio” implica, en este momento, examinar la distribución de muestreo del estimador.

Por ejemplo, si queremos estimar la varianza de una variable en la población por medio de los datos de la muestra, y si aplicamos la fórmula del método analógico

$$\frac{1}{n} \sum_i (x_i - m_x)^2$$

podemos demostrar que obtenemos un estimador sesgado: en caso de repetir el cálculo con muy grande número de muestras (un infinidad), el resultado sería, en promedio, diferente de la verdadera varianza. Es la razón por la cual usamos de preferencia un estimador corregido con tal de eliminar el sesgo; la fórmula de este estimador no sesgado es

$$s_x^2 = \frac{1}{n-1} \sum_i (x_i - m_x)^2$$

De la misma manera, el estimador de la covarianza mediante la fórmula del método analógico

$$\frac{1}{n} \sum_i (x_i - m_x)(y_i - m_y)$$

da un estimador sesgado de la covarianza entre x e y en la población, mientras que

$$s_{xy} = \frac{1}{n-1} \sum_i (x_i - m_x)(y_i - m_y)$$

es un estimador no sesgado.

2. *Eficacia relativa: los estimadores “best unbiased”*. En el universo de las muestras posibles, los resultados que cualquier estimador no sesgado arroja y apuntan en promedio hacia el objetivo que constituye el valor del parámetro que se pretende estimar. ¿Cómo escoger, en estas condiciones, entre dos estimadores no sesgados? Es obvio que se escogerá el estimador que apunta más al centro del objetivo y se evitará el estimador que arroja resultados más dispersos.

Es justamente la varianza que mide la dispersión de una variable aleatoria alrededor de su promedio. Llamamos “varianza de muestreo” la varianza de un estimador en la población de las muestras posibles (es la varianza de la distribución de muestreo); la raíz cuadrada de la varianza de muestreo constituye el “error de muestreo” (sampling error). Se dice que un estimador no sesgado es más eficaz que otro cuando su varianza de muestreo es inferior a la varianza del otro.

Llamamos “best unbiased” un estimador no sesgado cuya eficacia relativa es superior a la eficacia de cualquier otro estimador no sesgado. Esta misma apelación se usa de manera más restrictiva para una clase dada de estimadores. Por ejemplo, en la clase de estimadores cuyo valor es una función lineal de los datos, el estimador que detenta la mejor eficacia relativa es conocido como el “Best Linear Unbiased Estimate” o “BLUE”.

3. *Convergencia.* Otra propiedad deseable de un estimador es que su precisión sea superior en cuanto la muestra sea de tamaño más grande o, dicho de otra manera, que su varianza de muestreo sea más pequeña cuando la muestra es más grande. Se dice que un estimador es convergente si su varianza de muestreo tiende hacia cero cuando el tamaño de la muestra tiende hacia el infinito (la distribución de muestreo tiende a concentrarse en un solo punto).

4. *Suficiencia.* Finalmente, un estimador es suficiente cuando incorpora toda la información contenida en la muestra con relación al parámetro que se pretende estimar; en cuanto se calculó el valor del estimador (a partir de los datos de la muestra), no se podrá aprender algo más sobre el valor del parámetro aunque se examinen nuevamente los datos de la muestra.

Técnicamente, esta propiedad se traduce de la manera siguiente: si un estimador es suficiente entonces la probabilidad de la muestra (su verosimilitud) dada el valor estimado es independiente del valor del parámetro.

2-2.3.3 *La lógica fundamental de las pruebas de hipótesis**

Volvamos a examinar el esquema del método científico estudiado en el apartado 2-2.1. La lógica fundamental de este proceso es el siguiente:

- Si una teoría (o un modelo o una hipótesis) es verdadera, entonces sus implicaciones son también verdaderas.
- Por lo tanto, si las observaciones contradicen las implicaciones de una teoría, esta teoría no es verdadera; es falsa.

Con este razonamiento, se pretende aclarar algo de suma importancia: ¡si las observaciones no contradicen las implicaciones de una teoría, este hecho no nos da el derecho de concluir que esta teoría es verdadera! Con más precisión, para poder concluir que esta teoría es verdadera, es necesario que no exista otra teoría posible que sea compatible con las observaciones. En la práctica, esta condición es tan exigente que nunca se cumple.

En resumen, al momento de confrontar las implicaciones de una teoría con las observaciones, se rechaza la teoría cuando las observaciones contradicen las implicaciones; en el caso contrario, cuando las observaciones no contradicen las implicaciones, no se puede todavía confirmar la teoría, sólo queda en la categoría de “no rechazada”.¹⁰⁶

* Referencias: Blalock (1972), cap. 8, “The fallacy of affirming the consequent”.

¹⁰⁶ Personalmente prefiero la expresión “no rechazada” en lugar de la palabra “aceptable” que usa Wonnacott y Wonnacott (1992) por el riesgo de pasar de la categoría “aceptable” a la categoría “aceptada”, que no es lo mismo. Se reconocerá, aquí un parentesco con el falsificacionismo poppe-

Es esta misma lógica de rechazo/no rechazo que prevalece en los tests de hipótesis. Sin embargo, existe una diferencia capital: en los tests de hipótesis, la relación entre las hipótesis y la muestra observada es aleatoria, lo que implica que el razonamiento ya no puede ser determinista sino más bien probabilista.

En una lógica determinista, una observación es compatible con la hipótesis o no lo es, es decir, no existe un punto intermedio. En la lógica probabilista, una observación es más o menos compatible con la hipótesis: cuanto más improbable una observación mientras que se supone la hipótesis verdadera, menos compatible con la hipótesis.

Para ilustrar esto, enseñaremos un ejemplo un tanto caricatural:

Consideremos la hipótesis de que el dromedario no es parte de la fauna salvaje del continente australiano. Supongamos que un viajero, con gran estupefacción, encuentre un dromedario sin amo en el desierto australiano. Esta observación contradice su hipótesis. Sin embargo, este dromedario pudiera haberse escapado de un circo o de un zoológico o bien pudiera ser un espejismo. La observación de un dromedario no es imposible, más bien es improbable: la observación de un solo dromedario o hasta de algunos no podría considerarse como incompatible con la hipótesis. Ahora bien, supongamos que el mismo viajero ubique dromedarios en diferentes momentos. Si la hipótesis fuese verdadera, estas repetidas observaciones serían extraordinariamente improbables. Al paso de tiempo, el observador acabará por concluir que estas observaciones no son compatibles con su hipótesis.¹⁰⁷

riano; para Popper, una hipótesis que no es posible rechazar lógica o empíricamente no es “científica”.

¹⁰⁷ Los dromedarios salvajes son parte de la fauna de los desiertos australianos desde que fueron abandonados por caravaneros afganos, quienes los

En este ejemplo, nuestro viajero se contentará, de seguro, con un enfoque intuitivo. En caso de los tests de hipótesis, está claro que se formaliza mucho más el proceso; en particular,

- Se debe cuantificar las probabilidades de las cuales trata el razonamiento. (Si bien se sabe que el dromedario no es parte de la fauna australiana, ¿cuál es la probabilidad exacta de, no obstante, encontrar un dromedario? ¿De encontrar dos? ¿Tres?);
- Se debe tomar la decisión de rechazar la hipótesis bajo criterios precisos, definido de antemano (¿cuál es la probabilidad arriba de la cual decidiremos que las observaciones son incompatibles con la hipótesis?)

Es la primera de estas exigencias la que, por mucho, causa grandes dificultades tanto conceptuales como prácticas. En cuanto a la segunda, veremos que es, ni más ni menos, una exigencia de transparencia.

importaron para asegurar las comunicaciones trans-continetales antes la construcción del ferrocarril.

