

Diplomado en Análisis de Información Geoespacial

Conceptos básicos de estadística



Autor:
M. en G. Alberto Porras Velázquez

Introducción

La estadística se refiere a un conjunto de **métodos** para manejar la obtención, presentación y análisis de observaciones numéricas. Sus objetivos son **describir** el conjunto de datos obtenidos y **tomar decisiones** o **realizar generalizaciones** acerca de las características de todas las posibles observaciones bajo consideración.

Los métodos estadísticos son ampliamente utilizados en diferentes áreas del quehacer humano: el gobierno, los negocios, la educación, la psicología, la sociología, la antropología, las ciencias del comportamiento, la agricultura, la medicina, la biología y la física, entre otras.

Existen dos grandes ramas en esta disciplina, una es la estadística descriptiva y se refiere a la obtención, organización, presentación y descripción de la información numérica. Por otro lado, la estadística inferencial es una técnica mediante la cual se obtienen generalizaciones o se toman decisiones con base en información parcial o incompleta, obtenida mediante técnicas descriptivas.

En la ilustración 1 se muestra el proceso de aplicación de los métodos estadísticos en su dos grandes ramas. Con la estadística descriptiva se obtienen los estadísticos muestrales utilizados por la estadística inferencial para hacer inferencias y estimaciones de los parámetros de una población.

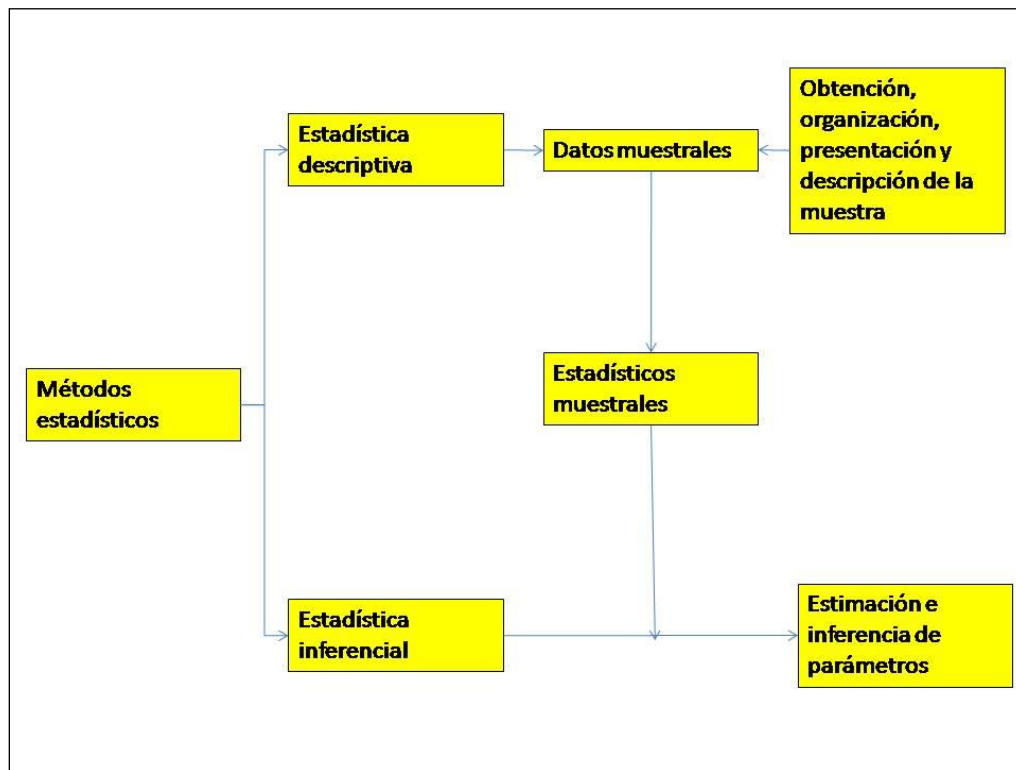


Ilustración 1. La estadística descriptiva y la estadística inferencial. Porras, A. (2014).

Definiciones básicas

- **Población (universo)**

Conjunto de todas las posibles unidades de observación que son objeto del problema a considerar. Es el objeto real de interés del cual la muestra escogida constituye un subconjunto particular. Por ejemplo: los niños de diez años en México.

Una población es finita si el proceso de conteo de las unidades que la conforman puede completarse o si incluye un número limitado de medidas u observaciones. Ejemplo: todas las personas que viven en el hemisferio norte o los estudiantes de secundaria en México.

En cambio, una población infinita incluye un gran conjunto de medidas que no puede obtenerse por conteo. Ejemplo: la población de dígitos extraída con reemplazo de una urna o la colección de todos los números positivos.

- **Muestra**

Parte o subconjunto de una población. Subconjunto de medidas u observaciones tomadas a partir de una población dada. Se utiliza una muestra por razones prácticas, económicas o de tiempo que no permiten considerar a toda la población. Ejemplo: un centenar de niños de diez años.

- **Parámetros**

Se refieren a las características medibles de una población y a los valores verdaderos que las describen. Ejemplo: estatura promedio de la población de 15 años en México. No es posible conocer este dato si la población es infinita. Una población tiene sólo un valor para un parámetro en particular (valor verdadero), pero puede tener diferentes valores para el estadístico muestral correspondiente.

- **Estadístico muestral**

Característica medible de una muestra. Coeficiente intelectual, estatura, peso, etc.

- **Inferencia**

La propuesta de una característica en la población (parámetro) estimada según la evaluación de la característica en la muestra (estadístico). Ejemplo: el número de piezas defectuosas obtenidas en un proceso de fabricación.

Nomenclatura

- **Parámetros**

Se trata de los valores calculados de las características medibles que describen a la población, los cuales son representados con letras griegas. Por ejemplo, la media y la desviación estándar se denotan por μ y σ .

- **Estadísticos muestrales**

Los valores calculados para una muestra son designados por las letras del alfabeto latino. La media se representa con m o \bar{X} y la desviación estándar, mediante la letra s .

Abusos de la estadística

En la vida cotidiana es común encontrar ejemplos del uso indebido de la estadística, ya sea por una interpretación incorrecta de la información o por presentarla de manera parcial.

Una conocida frase, que algunas veces se atribuye a Benjamin Disraeli y otras a Mark Twain, dice “Existen tres tipos de mentiras: las mentiras, las malditas mentiras y la estadística”.

El uso incorrecto de la estadística puede deberse a la realización de un muestreo erróneo y la presentación engañosa de los datos; contar con datos insuficientes, o no representativos llevará a conclusiones falsas.

Características de un estimador

Como se mencionó, los estadísticos muestrales pueden fungir como estimadores de los parámetros de una población. Algunas características de los estimadores son:

- **Exactitud**

Es el grado de aproximación de los resultados de la observación (de la medida, del cálculo o de la estimación) con respecto al valor verdadero (o que se considera como tal) o a un valor de referencia.

- **Precisión**

Desde el punto de vista estadístico, resulta del agrupamiento de los datos obtenidos por la repetición de una medida. Sólo se ve afectada por factores aleatorios y no guarda relación con el valor aceptado como real. Una varianza pequeña nos indica que hay mayor precisión.

- **Sesgo**

Se dice que los datos presentan un sesgo cuando todas las mediciones están cargadas hacia un mismo lado del valor real. En este caso, el valor esperado del estadístico muestral difiere del parámetro correspondiente.

En la ilustración 2 se muestran gráficamente las características de un estimador. Supongamos que te encuentras frente a un tablero de tiro al blanco y que el punto central representa el valor del parámetro real de una población; mientras que cada tiro (representado en la ilustración mediante una estrella de color rojo) equivale al valor de un estimador de dicho parámetro, tomado a partir de una muestra.

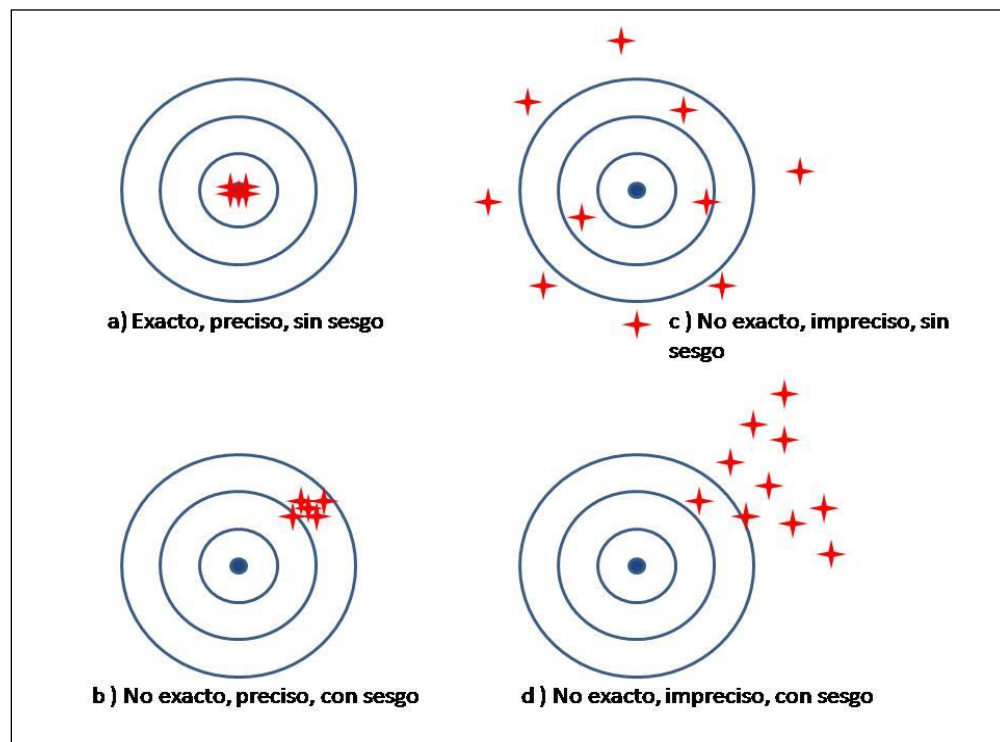


Ilustración 2. Características de un estimador. Porras, A. (2014).

En el caso “a” se tiene que los tiros realizados (estimadores) son exactos porque tienden a dar en el blanco, es decir que son precisos, están agrupados y varían poco; no hay sesgo porque los estimados son similares al valor real.

En el caso “b”, a diferencia del primero, los estimadores no dieron en el blanco (no son exactos), sin embargo varían poco (son precisos) y tienen un sesgo porque los tiros están cargados hacia un lado del valor real.

En lo que respecta a la opción “c”, los estimadores no son exactos porque no dan en el blanco, son imprecisos en tanto que los tiros están dispersos en el espacio y no hay sesgo, es decir, los estimadores no se agrupan en una dirección preferencial.

Finalmente, en el caso “d”, los estimadores no son exactos porque no dan en el blanco, son imprecisos porque están dispersos, y hay un sesgo debido a que sí hay una dirección en la cual tienden a agruparse los estimadores.

Algunos ejemplos de abusos en el uso de la estadística

Aquí se muestran algunas afirmaciones que denotan el uso inadecuado de la estadística:

- **Datos insuficientes.** Un vendedor de pólizas, basado en cuatro visitas, afirma que un cuarto de sus reuniones son exitosas.
- **Datos muestrales no representativos.** Supongamos que los gastos mensuales de una muestra de estudiantes son: 2000, 1900, 2300, 1800, 1700, 2100, 2200 y 10, 000. Con todos los estudiantes, el gasto promedio total asciende a 3000, pero sin el último de ellos, decae a 2000.
- **Falsas conclusiones.** Según las estadísticas, un estadounidense que viaja en avión está más seguro que otro que permanece en casa, pues en un año hubo solamente 395 accidentes aéreos fatales, mientras que ocurrieron 29,000 accidentes fatales en el hogar.

Referencias:

- Canavos, G. (1994). *Probabilidad y Estadística*. México: McGraw-Hill.
- Huff, D. y Geis, I. (1993). *How to lie with Statistics*. Nueva York: W. W. Norton.
- Levin, J. y Levin, W. (1997). *Fundamentos de Estadística en la Investigación Social*. México: Oxford University Press.