



Bioestadística

Programa de Biología

Kevin Pérez - Ing de Sistemas - Estadístico - (E) MSc. Ciencia de Datos
Departamento de Matemáticas y Estadística - Universidad de Córdoba

Genralidades del curso

Objetivo general.

Fomentar la aplicación de las técnicas estadísticas en la solución de problemas que se presentan en el trabajo cotidiano de un Biólogo, mediante la formación integral de profesionales con sólidas bases tanto en su saber específico como en las otras disciplinas.

- Dar a conocer a los estudiantes los conceptos de estadísticos fundamentales en su formación básica.
- Orientar a los estudiantes en los procesos que involucran la aplicación de métodos estadísticos.
- Inducir a los estudiantes en la utilización de herramientas estadísticas apoyándose en software especializados.
- Motivar y fomentar en los estudiantes el espíritu investigativo y creativo mediante el uso de las técnicas y principios estadísticos.

Competencias

Al terminar el curso, el estudiante estará en capacidad de:

- Organizar de forma adecuada un conjunto de datos, teniendo en cuenta sus características.
- Reducir un conjunto de datos mediante el cálculo de medidas de tendencia central y dispersión.
- Identificará de manera adecuada el tipo de distribución que modela un determinado problema.
- Utilizará las diferentes técnicas de estimación para hacer inferencias sobre uno o mas parámetros de una o mas poblaciones.
- Aplicará las herramientas estadísticas en la toma de decisiones.

Contenido programatico

Unidad de aprendizaje N° 1. Estadística descriptiva

- Conceptos Básicos.
- Organización de datos.
- Medidas de tendencia central.
- Medidas de dispersión.
- Medidas de posición.

Contenido programatico

Unidad de aprendizaje N° 2. Teoría de probabilidad

- Elementos de probabilidad.
- Variables aleatorias
- Distribuciones discretas de probabilidad
- Distribuciones continuas de probabilidad

Contenido programatico

Unidad de aprendizaje N° 3. Distribuciones muestrales

- Distribución muestral de un parámetro.
- Distribución muestral de dos parámetros

Contenido programatico

Unidad de aprendizaje N° 4. Estimación de parámetros

- Estimación puntual.
- Estimación por intervalos.

Contenido programatico

Unidad de aprendizaje N° 5. Pruebas de Hipótesis

- Pruebas de hipótesis sobre un parámetro.
- Pruebas de hipótesis sobre varios parámetros.

Conceptos generales

- Orígenes de la estadística
- ¿Qué es Estadística?
- Conceptos Generales. (Población, Muestra, Muestreo, Estadígrafo, Parámetros, variable, dato, clasificación de la estadística, Etapas de una investigación estadística.)
- ¿Qué es variable?.
- Clasificación de variables

Orígenes de la Estadística

- Los orígenes históricos de la Estadística (descriptiva) hay que buscarlos en los procesos de recogida de datos, censos y registros sistemáticos, asumiendo un papel asimilable a una aritmética estatal para asistir al gobernante, que necesitaba conocer la riqueza y el número de sus súbditos con fines tributarios y políticos.
- Los primeros registros de riqueza y población que se conocen se deben a los egipcios. Ramsés II en el 1400 a.C. realizó el primer censo conocido de las tierras de Egipto, no siendo éste, se supone, ni el primero ni el último que se hiciera en las tierras bañadas por el Nilo.
- Posteriormente, desde el siglo III a.C., en las civilizaciones china y romana se llevan a cabo censos e inventarios de posesiones, que pueden considerarse precedentes institucionalizados de la recogida de datos demográficos y económicos de los Estados Modernos.

Orígenes de la Estadística

- Hay que realizar una mención especial del período helénico, en el que las escuelas matemáticas se suceden. Centros como el de Quios, donde estudió Hipócrates (Hipócrates de Quios) el matemático, considerado como el inventor del método matemático y escuelas como las de Cirene, Megara y al final Atenas, donde se reúnen los matemáticos, unos alrededor de Protágoras y otros en torno a Sócrates.
- En la Edad Media se vuelve a la utilización de la Aritmética para la recogida de datos, existiendo menos interés por la elucubración matemática abstracta. Es en este período de tiempo cuando Carlomagno ordenó en su *“Capitulare de villis”* la creación de un registro de todos sus dominios y bienes privados.
- En el siglo XVII se producen avances sustanciales, y así, en las universidades alemanas se imparten enseñanzas de “Aritmética Política”, término con el que se designa la descripción numérica de hechos de interés para la Administración Pública. Destacados autores de Aritmética Política fueron los ingleses Graunt (1620-1674) y Petty (1623-1687).

Orígenes de la Estadística

Con métodos de estimación en los que cabía la conjetura, la experimentación y la deducción, Graunt llega a estimar tasas de mortalidad para la población londinense, analizando además la verosimilitud de la información de que disponía. Por su parte, Petty, cuyas aportaciones estadísticas fueron menos relevantes.

Conceptos Generales

- ¿Qué es Estadística?

Conceptos Generales

La **Estadística** se ocupa del manejo de la información que pueda ser cuantificada. Implica esto la descripción de conjuntos de datos y la inferencia a partir de la información recolectada de un fenómeno de interés. La función principal de la estadística abarca: Resumir, Simplificar, Comparar, Relacionar, Proyectar. Entre las tareas que debe enfrentar un estudio estadístico están:

- Delimitar con precisión la población de referencia o el conjunto de datos en estudio, las unidades que deben ser observadas, las características o variables que serán medidas u observadas.
- Estrategias de Observación: Censo, Muestreo, Diseño de Experimental
- Recolección y Registro de la información

Conceptos Generales

- Depuración de la información.
- Construcción de Tablas.
- Análisis Estadístico:
 - Producción de resúmenes gráficos y numéricos.
 - Interpretación de resultados.

Cuando los datos comprenden toda la población de referencia, hablamos de un Censo y cuando solo comprometen una parte de ella, hablamos de una muestra. En ambos casos es pertinente un análisis Descriptivo. En el segundo caso un análisis Inferencial.

Conceptos Generales

A grandes rasgos podemos decir que una **Población** es el conjunto de toda posible información, o de los objetos, que permite estudiar un fenómeno de interés.

- Una **muestra** es un subconjunto de información representativa de una población.
- Las **Variables** resultan ser aquellas características de interés que desean ser medidas sobre los objetos o individuos seleccionados. En la mayoría de los casos lo que se pretende es estimar, a partir de la información recolectada de una muestra, características desconocidas de los objetos en dicha población de interés.
- Las características desconocidas de una población serán llamadas **Parámetros**. Las características calculadas a partir de una muestra son llamadas **Estadísticas**. Una Inferencia es una generalización obtenida a partir de una muestra aleatoria.

Conceptos Generales

La Estadística puede dividirse en dos grandes ramas: Estadística Descriptiva y Estadística Inferencial.

- ***Estadística descriptiva:*** Es el conjunto de métodos usados para la organización y presentación (descripción) de la información recolectada. La información recolectada puede ser catalogada de dos maneras: Datos Cualitativos y Cuantitativos.
- ***Estadística inferencial:*** Comprende los métodos y procedimientos para deducir propiedades (hacer inferencias) de una población, a partir de una pequeña parte de la misma (muestra).

Variables y atributos

Una primera clasificación del tipo de datos en función de que las observaciones resultantes del experimento sean de tipo cualitativo o cuantitativo, en el primero de los casos se tiene un atributo y en el segundo una variable. Para hacer referencia genéricamente a una variable o a un atributo se utilizará el término carácter.

- **Ejemplo 1:** Como ejemplos de **atributos** pueden considerarse el color del pelo de un colectivo de personas, su raza o el idioma que hablan y como **variables** su estatura, peso o edad.
- Para poder operar con un atributo es necesario asignar a cada una de sus clases un valor numérico, con lo que se transforma en una variable, esta asignación se hará de forma que los resultados que se obtengan al final del estudio sean fácilmente interpretables.

Variables discretas y continuas

Dentro del conjunto de las variables se distingue entre **discretas** y **continuas**. Se dice que una variable es discreta cuando entre dos valores consecutivos no toma valores intermedios y que es continua cuando puede tomar cualquier valor dentro de un intervalo.

- **Ejemplo 2:** La **estatura** de un grupo de personas sería una variable **continua**, mientras que el **número de cabellos** que tienen en la cabeza sería una variable **discreta**.
- En la práctica todas las variables son discretas debido a la limitación de los aparatos de medida, y así, en el ejemplo de las estaturas, quizás se podría detectar una diferencia de una cienmilésima de metro, o a lo más, de una millonésima, pero dados dos individuos que se diferencien en una millonésima no puede detectarse otro que tenga una estatura intermedia. De todas formas, en general se trata a las variables “*teóricamente*” continuas como tales, por razones que se pondrán de manifiesto más adelante.

Tipos de variables y escalas de medida

Se denomina escalamiento al desarrollo de reglas sistemáticas y de unidades significativas de medida para identificar o cuantificar las observaciones empíricas. La clasificación más común distingue cuatro conjuntos de reglas básicas que producen cuatro escalas de medida; éstas son:

- La escala de medida más simple implica una relación de identidad entre el sistema de números y el sistema empírico objeto de medida. La escala resultante se denomina **nominal**, porque los números empleados se consideran como “etiquetas” las cuales se asignan a los objetos con el propósito de clasificarlos, pero no poseen el significado numérico usual, aparte de la relación de igualdad; por tanto, tienen una naturaleza no métrica. *El género, la raza, la profesión, el credo religioso, son variables observadas en este tipo de escala.*

Tipos de variables y escalas de medida

- Una escala más compleja, implica además de la relación de igualdad como el caso anterior, una relación de orden que se preserva tanto en el sistema numérico como en el sistema empírico (medidas sobre los objetos). Éste tipo de escalas se denomina **ordinal** porque los números que se asignan a los atributos deben respetar (conservar) el orden de la característica que se mide. El tipo de datos que resulta tiene naturaleza no métrica. La valoración de la opinión en *“de acuerdo”*, *“indiferente”* o *“en desacuerdo”*, constituye un ejemplo de una variable típica de esta escala.

Tipos de variables y escalas de medida

- El siguiente nivel de escalamiento implica, además de una relación de orden como la escala anterior, una relación de igualdad de diferencias entre pares de objetos respecto a una característica determinada. La escala resultante se denomina de **intervalo** porque las diferencias entre los números se corresponden con las diferencias entre la propiedad medida sobre los objetos, y por tanto tiene naturaleza métrica. La medición de la *temperatura, la altura física, constituyen ejemplos de esta escala de medida*. Una característica adicional de esta escala es la necesidad de precisar un origen o punto “cero” respecto al cual la medida tiene sentido, esto no necesariamente significa ausencia del atributo. En el ejemplo de la temperatura, el cero en la escala Celsius, es la temperatura de congelación del agua al nivel del mar; nótese que este cero no corresponde con el de la escala Fahrenheit.

Tipos de variables y escalas de medida

- El nivel más complejo de escalamiento implica, además de una relación de igualdad de diferencias como en la escala anterior, un punto de origen fijo o natural, el cero absoluto. El resultado es la escala de **razón**, que tiene también naturaleza métrica. Ejemplos de este tipo de escala son el peso, la talla o la edad de los individuos.

Resumen

- Las variables pueden ser **cuantitativas**, cuando se realiza una medición y el resultado es un número, o pueden ser **cualitativas**, cuando solamente registran una cualidad o atributo del objeto o persona en estudio. *La edad, el peso y la estatura son ejemplos de variables cuantitativas en una población de personas, mientras que el sexo y el estado civil son variables cualitativas.*
- De acuerdo al número total de sus posibles valores, una variable cuantitativa puede ser clasificada como discreta cuando sólo puede tomar un número discreto (es decir, finito o numerable) de valores, o continua cuando puede tomar cualquier valor dentro de un intervalo (a, b) de la recta real.

Resumen

De acuerdo con la posible relación que pudieran guardar los valores de una variable, se cuenta por lo menos con cuatro escalas de medición. Las variables cualitativas pueden ser clasificadas de acuerdo a dos escalas: escala nominal o escala ordinal. Mientras que las variables **cuantitativas** pueden clasificarse por: **escala de intervalo** o **escala de razón**.

Ejemplo de formula

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

