



**DOI:** 10.26820/reciamuc/6.(3).julio.2022.677-686

**URL:** <https://reciamuc.com/index.php/RECIAMUC/article/view/953>

**EDITORIAL:** Saberes del Conocimiento

**REVISTA:** RECIAMUC

**ISSN:** 2588-0748

**TIPO DE INVESTIGACIÓN:** Artículo de Investigación

**CÓDIGO UNESCO:** 32 Ciencias Médicas

**PAGINAS:** 677-686



## Bioestadística como Herramienta aplicada por estudiantes de la Facultad de Ciencias Médicas en el Aprendizaje dentro del Campo de la Salud

Biostatistics as a Tool applied by students of the Faculty of Medical Sciences in Learning within the Health Field

Bioestatística como Ferramenta aplicada por estudantes da Faculdade de Ciências Médicas na Aprendizagem dos Cuidados de Saúde

**Marcos Rodolfo Tobar Moran<sup>1</sup>; Sara Falconi San Lucas<sup>2</sup>; Zynnia Gabriela Reyes Sánchez<sup>3</sup>; Ana María Ramírez Hecksher<sup>4</sup>**

**RECIBIDO:** 20/06/2022 **ACEPTADO:** 10/07/2022 **PUBLICADO:** 01/08/2022

1. Magister en Sistemas de Información Gerencial; Ingeniero en Electricidad Especialización Electrónica; Universidad de Guayaquil; Universidad de Especialidades Espíritu Santo; Guayaquil, Ecuador; marcos.tobarm@ug.edu.ec;  <https://orcid.org/0000-0002-0083-5813>
2. Máster Universitario en Ingeniería de Software y Sistemas Informáticos; Magíster en Gerencia y Liderazgo Educacional; Universidad de Guayaquil; Guayaquil, Ecuador; sara.falconis@ug.edu.ec;  <https://orcid.org/0000-0002-4708-4403>
3. Master en Sistemas de Información con Mención en Inteligencia de Negocios; Ingeniera en Sistemas Computacionales; Universidad de Guayaquil; Guayaquil, Ecuador; zynnia.reyess@ug.edu.ec;  <https://orcid.org/0000-0001-6659-4399>
4. Master Universitario en Diseño y Gestión de Proyectos Tecnológicos; Licenciada en Ciencias de la Educación; Universidad de Guayaquil; Guayaquil, Ecuador; ana.ramirez@ug.edu.ec;  <https://orcid.org/0000-0002-6767-155X>

### CORRESPONDENCIA

Marcos Rodolfo Tobar Moran

marcos.tobarm@ug.edu.ec

Guayaquil, Ecuador

## RESUMEN

El profesional sanitario gracias a su experiencia y estudio, construye su ojo clínico. La estadística le ofrece conceptos que pueden facilitar este aprendizaje. La bioestadística es la estadística aplicada a las ciencias biológicas y en particular al área de ciencia de la salud, un instrumento necesario para la planificación de investigaciones. Estas investigaciones requieren del manejo teórico y práctico de estadística. Todos los datos recopilados a través de la bioestadística sirven para encontrar patrones que ayuden a luchar contra las enfermedades y a diseñar campañas de prevención. La estadística se divide en dos ramas: estadística descriptiva y estadística inferencial. La estadística descriptiva es una rama de las matemáticas que permite: ordenar, organizar y presentar la información de una determinada investigación, mediante el uso de tablas y figuras así como valores representativos de tendencia central, de forma, de dispersión. Las medidas de tendencia central tienen como objetivo describir, mediante un valor numérico, la localización de las observaciones. Son valores que representan, la posición donde se concentran los datos observados, los más usados: la media, mediana y moda. Se denominan medidas de dispersión aquellas que pretenden captar y resumir la mayor o menor variabilidad, la mayor o menor concentración, homogeneidad o parecido entre las observaciones de la variable. Se presenta a continuación las medidas de dispersión más frecuentes: el rango o recorrido, la varianza, la desviación estándar y el coeficiente de variación. La estadística inferencial se encarga de contrastar los resultados de la muestra y generalizarlos hacia toda la población, considerando un riesgo de que en este proceso se cometa algún error como por ejemplo de un "1%, o un 5% dependiendo del nivel de confianza. El modelo de regresión lineal exige que los datos de la población cumplan los supuestos de linealidad, homocedasticidad, independencia entre las observaciones y normalidad de las distribuciones condicionadas de la variable Y. Se aplicó una metodología descriptiva, con un enfoque documental, es decir, revisar fuentes disponibles en la red, con contenido oportuno y relevante para dar respuesta a lo tratado en el presente artículo.

**Palabras clave:** Bioestadística, Estadística Descriptiva, Estadística Inferencial, Media, Mediana, Moda, Inferencia, Dispersión, Error, Proyección, Comparabilidad, Representatividad.

## ABSTRACT

The health professional, thanks to his experience and study, builds his clinical eye. Statistics offers you concepts that can facilitate this learning. Biostatistics is statistics applied to biological sciences and in particular to the area of health science, a necessary instrument for research planning. These investigations require the theoretical and practical handling of statistics. All data collected through biostatistics is used to find patterns that help fight diseases and design prevention campaigns. Statistics is divided into two branches: descriptive statistics and applied statistics. Descriptive statistics is a branch of mathematics that allows: ordering, organizing and presenting the information of a certain investigation, through the use of tables and figures as well as representative values of central tendency, shape, dispersion. The measures of central tendency aim to describe, by means of a numerical value, the location of the observations. They are values that represent the position where the observed data are concentrated, the most used: the mean, median and mode. Dispersion measures are those that seek to capture and summarize the greater or lesser variability, the greater or lesser concentration, homogeneity or similarity between the observations of the variable. The most frequent measures of dispersion are presented below: the range or path, the variance, the standard deviation and the coefficient of variation. Inferential statistics is responsible for contrasting the results of the sample and generalizing them to the entire population, considering a risk that some error will be made in this process, such as "1%, or 5% depending on the level of confidence. The linear regression model requires that the population data meet the assumptions of linearity, homoscedasticity, independence between the observations and normality of the conditional distributions of the variable Y. A descriptive methodology was applied, with a documentary approach, that is, to review sources available on the network, with timely and relevant content to respond to what is discussed in this article.

**Keywords:** Biostatistics, Descriptive Statistics, Inferential Statistics, Mean, Median, Mode, Inference, Dispersion, Error, Projection, Comparability, Representativeness.

## RESUMO

O profissional de saúde, graças à sua experiência e estudo, constrói o seu olho clínico. A estatística oferece-lhe conceitos que podem facilitar esta aprendizagem. A bioestatística é a estatística aplicada às ciências biológicas e em particular à área das ciências da saúde, um instrumento necessário para o planeamento da investigação. Estas investigações requerem o tratamento teórico e prático da estatística. Todos os dados recolhidos através da bioestatística são utilizados para encontrar padrões que ajudam a combater doenças e a conceber campanhas de prevenção. A estatística está dividida em dois ramos: a estatística descritiva e a estatística aplicada. A estatística descritiva é um ramo da matemática que permite: ordenar, organizar e apresentar a informação de uma determinada investigação, através da utilização de tabelas e figuras, bem como valores representativos de tendência central, forma, dispersão. As medidas de tendência central visam descrever, por meio de um valor numérico, a localização das observações. São valores que representam a posição onde os dados observados estão concentrados, os mais utilizados: a média, a mediana e a moda. As medidas de dispersão são aquelas que procuram captar e resumir a maior ou menor variabilidade, a maior ou menor concentração, homogeneidade ou similaridade entre as observações da variável. As medidas de dispersão mais frequentes são apresentadas a seguir: o intervalo ou caminho, a variância, o desvio padrão e o coeficiente de variação. As estatísticas inferenciais são responsáveis por contrastar os resultados da amostra e generalizá-los a toda a população, considerando um risco de que algum erro seja cometido neste processo, como "1%, ou 5% dependendo do nível de confiança". O modelo de regressão linear exige que os dados da população satisfaçam os pressupostos de linearidade, homocedasticidade, independência entre as observações e normalidade das distribuições condicionais da variável Y. Foi aplicada uma metodologia descritiva, com uma abordagem documental, ou seja, rever as fontes disponíveis na rede, com conteúdo atempado e relevante para responder ao que é discutido neste artigo.

**Palavras-chave:** Bioestatística, Estatística Descritiva, Estatística Inferencial, Média, Mediana, Modo, Inferência, Dispersão, Erro, Projeção, Comparabilidade, Representatividade.

## **Introducción**

Los datos son la materia prima de la estadística. Se obtienen de las mediciones u observaciones. Al mismo tiempo, las técnicas de análisis estadísticas permiten que los datos se conviertan en información útil para la toma de decisiones. Ambos componentes son fundamentales y necesarios para que los resultados muestren lo que a simple vista no se aprecia y permitan responder la pregunta de investigación.

El profesional sanitario gracias a su experiencia y estudio, construye su ojo clínico. La estadística le ofrece conceptos que pueden facilitar este aprendizaje. Es importante considerar que los métodos estadísticos y sus resultados no pretenden ser verdades absolutas, puesto que la medicina no es exacta, solo es posible calcular las probabilidades que ocurran los eventos en salud. El concepto principal de la estadística es la variabilidad, conocer cómo medirla y modelarla comienza a ser una idea positiva, oportuna para su aplicación en salud. La estadística aporta teoremas que conectan la variabilidad e independencia con la información que se recolecta de los sujetos o unidad muestral.

La estadística permite analizar situaciones en las que los componentes aleatorios contribuyen de forma importante en la variabilidad de los datos obtenidos. En salud pública, por ejemplo, los componentes aleatorios se deben al conocimiento o a la imposibilidad de medir algunos determinantes de los estados de salud y enfermedad, así como a la variabilidad en las respuestas por los pacientes, similares entre sí, que son sometidos al mismo tratamiento.

La extensión de los conocimientos y aptitudes de carácter estadístico que necesitan adquirir los médicos son importantes, porque el conocimiento de los principios y métodos estadísticos y la competencia en su aplicación se necesitan para el ejercicio eficaz de la salud, y para la “comprensión e interpretación de los datos sanitarios; a fin

de discriminar entre opiniones arbitrarias o discrecionales, con respecto a las verdaderamente evaluadas en un contexto científico” (Roi, Pérez, Hunte, & Ruiz, 2018).

Si se desea medir la eficacia de un determinado tratamiento con respecto a otro, el desarrollo de un bebe es de acuerdo a lo esperado según su edad, los síntomas de una enfermedad en específico, la probabilidad de padecer alguna enfermedad hereditaria como la diabetes, son interrogantes que los profesionales de la salud pueden responder gracias al uso de herramientas tan poderosas como la estadística descriptiva y la estadística inferencial, las cuales serán desarrolladas a continuación.

## **Metodología**

Esta investigación está dirigida al estudio del tema “Bioestadística como Herramienta aplicada por estudiantes de la Facultad de Ciencias Médicas en el Aprendizaje dentro del Campo de la Salud”. Para realizarlo se usó una metodología descriptiva, con un enfoque documental, es decir, revisar fuentes disponibles en la red, cuyo contenido sea actual, publicados en revistas de ciencia, disponibles en Google Académico, lo más ajustadas al propósito del escrito, con contenido oportuno y relevante desde el punto de vista científico para dar respuesta a lo tratado en el presente artículo y que sirvan de inspiración para realizar otros proyectos. Las mismas pueden ser estudiadas al final, en la bibliografía

## **Resultados**

La investigación clínica se apoya en el método estadístico para la recolección, organización, análisis e interpretación de una serie de datos en la obtención de inferencias de lo que ocurre en una población determinada, a partir de una muestra representativa de la misma.

La bioestadística según Revista Electrónica de PortalesMedicos.com, (2013): “es la estadística aplicada a las ciencias biológi-

cas y en particular al área de ciencia de la salud” es un instrumento necesaria para la planificación de investigaciones. Estas investigaciones requieren del manejo teórico y práctico de estadística. Todos los datos recopilados a través de la bioestadística sirven para encontrar patrones que ayuden a luchar contra las enfermedades y a diseñar campañas de prevención.

Al ser una rama de la estadística, la bioestadística se encarga de aspectos relacionados con la recolección de datos y su correcto almacenamiento; análisis de la información a través de diversos métodos y herramientas; como la representación gráfica de los resultados obtenidos; mecanismos para la interpretación de dichos resultados; diseño y desarrollo de experimentos.

Los principales objetivos de la bioestadística se parafrasean de las ideas propuestas por Calvache, Barón, & Garret, (2006) son los siguientes:

- Decidir qué tipo de datos recoger y en qué cantidad para extraer conclusiones.
- Recoger datos suficientes, categorizarlos en atención al objeto de estudio y almacenarlos para su posterior consulta y análisis.
- Realizar análisis con herramientas computacionales y aplicar métodos estadísticos.
- Evidenciar los resultados y presentarlos de acuerdo a una visualización sencilla y rápida.
- Emitir conclusiones en base a los datos recogidos y analizados.
- Analizar el proceso una vez concluido e implementar mejoras en el método y en las herramientas para futuros estudios.

Cuando se valora la investigación, se empieza a comprender que la estadística no es una colección de fórmulas matemáticas, sino que se transforma en una herramienta invaluable que otorga sentido al manejo de

datos tanto cualitativos como cuantitativos, cuyos resultados y conclusiones no pueden ser abordados desde una perspectiva determinística como ocurre en otras ciencias, por ejemplo, la física. Las estadísticas de salud son todos aquellos datos numéricos debidamente capturados, validados, elaborados analizados e interpretados que se requieren para las acciones de salud. Se pueden agrupar para (Castro, 2019) en:

- Estadística de poblaciones: estadísticas demográficas.
- De hechos biológicos que tiene trascendencia sanitaria, como los nacimientos y defunciones: estadísticas vitales.
- De la enfermedad, que se intenta prevenir y tratar: estadísticas de morbilidad.
- De los medios tanto específicos como inespecíficos para proteger, fomentar, detectar y recuperar la salud: estadísticas de recursos.
- De las acciones e intervenciones, que desarrollan los recursos básicos para la evaluación: estadísticas de servicios.

La estadística en salud depende de los datos y su análisis, es esencial conocer la metodología que permita la correcta recolección de los datos y comprender el análisis de éstos. “Para conocer el estado de salud de la población se requiere del uso de las Estadísticas de Salud que ofrecen información numérica cuantificable, a través de diferentes indicadores” (Velázquez, González, Velázquez, & Martínez, 2019), datos de natalidad, mortalidad, morbilidad, de recursos y servicios y de calidad. Este último es una herramienta poderosa para mejorar la atención del paciente. La definición de estadística corresponde a la propuesta por Andrade, (2019):

“La estadística es la ciencia que permite recopilar, organizar y analizar la información de determinados fenómenos donde los investigadores mantienen incertidumbre, permitiéndole tomar decisiones. En el proceso

de formación se debe insistir en: la comprensión, la interpretación, la evaluación y la comunicación de los resultados sobre la base de un conocimiento estadístico básico, sin desatender, además, elementos como la responsabilidad y honradez profesionales, en tanto constituyen cualidades esenciales para su formación profesional” (p. 45).

Expresa otro autor Roi, Pérez, Hunte, & Ruiz, (2018) que el pensamiento estadístico involucra cinco componentes: “necesidad de los datos, importancia de producirlos, variabilidad, medición y modelación de estos”. Desarrollar la dinámica de la formación estadística desde el reconocimiento a la variabilidad de los datos como su epicentro, es la búsqueda e identificación de regularidades dentro de la diversidad de los datos, con la finalidad clara de utilizarlas como sustento teórico durante la realización de inferencias inductivas.

La estadística se divide en dos ramas: estadística descriptiva y estadística aplicada. Son dos partes que proveen al estudiante y/o profesionales de la salud, de herramientas científicas poderosas para el análisis de datos cuantitativos o cualitativos de la investigación.

### **Estadística descriptiva**

Para comprender el comportamiento de una variable sobre un conjunto de datos se recurre a la organización y arreglo de sus observaciones en una estructura denominada tabla de distribución de frecuencias. Es una tabla en la se disponen los diferentes valores o categorías de la variable a estudio acompañados de información sobre su número de aparición entre los individuos. Se trata de establecer a qué categoría pertenece o cuál es el valor alcanzado por cada uno de los individuos de la base de datos.

Para cualquier tipo de variable, la información básica que podemos incluir en la tabla está formada por Nolasco & Moncho, (2016) “las frecuencias absolutas ( $f_i$ ) de los valores o categorías de las variables, porcentajes

( $p_i$ ), medidas relativas de la frecuencia de aparición de cada valor o categoría entre los individuos de la base de datos”. La asignación de los individuos a las diferentes categorías o valores debe ser exclusiva y exhaustiva, es decir, cada uno de los individuos debe ser asignado a una y sólo una de las categorías o valores, mientras que todos los individuos deben ser asignados.

Cuando la variable es cuantitativa, suele ser estructurada tabulando la variable por intervalos. Se trata de definir una secuencia de intervalos de forma que un individuo será contabilizado en uno de ellos si el valor de la variable a estudio está incluido en él. Cada intervalo es definido por (Nolasco & Moncho, 2016) a través de “sus límites [ $x_i$ ,  $x_{i+1}$ ], como el conjunto de valores de la variable, digamos  $X$ , tal que  $x_i \leq X < x_{i+1}$ ”. Con esta definición se dirá que los intervalos son abiertos por la derecha y cerrados por la izquierda.

La estadística descriptiva es una rama de las matemáticas que permite ordenar, organizar y presentar la información de una determinada investigación, mediante el uso de tablas y figuras así como valores representativos de tendencia central, de forma, de dispersión, concepto que concuerda con muchos otros autores (Pantoja & Burbano, 2021).

### **Medidas de tendencia central**

Las medidas de tendencia central tienen como objetivo describir, mediante un valor numérico, la localización de las observaciones. Son valores que representan, la posición donde se concentran los datos observados. La comparación de estas medidas, calculadas para una misma variable, entre diferentes conjuntos de datos puede indicarnos las diferencias en la posición de los valores entre los conjuntos considerados. Las medidas más usuales:

Media: es la medida de tendencia central más utilizada, dada una variable genérica  $X$ , y un conjunto de  $n$  observaciones de esta

variable  $\{x_1, x_2, x_3, \dots, x_n\}$ , la media es la suma de todas las observaciones dividida por su número de observaciones o elementos:

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

Resulta un valor poco representativo cuando la existencia de datos atípicos puede incrementar o disminuir su valor de forma notable, convirtiéndola en poco representativa del centro de la estructura de los datos. Esto sucederá en mayor o menor medida cuando los datos presenten estructuras no simétricas.

**Mediana:** es aquel valor tal que la cantidad de datos inferiores a él es igual a la cantidad de datos superiores. Es una medida que busca el centro de la estructura de los datos bajo la idea de distribuir las observaciones en dos conjuntos de igual número, se conoce como el percentil 50. Para entender el concepto y proceder a su cálculo es necesario partir de que el conjunto de observaciones de la variable es ordenado (generalmente de menor a mayor).

Al igual que en el caso de la media, existe una única mediana para un conjunto determinado de datos. Una limitación es que la mediana no utiliza para su cálculo todos los valores de las observaciones de la variable, ocasionando una menor capacidad informativa. Como ventaja, la mediana no se verá afectada por observaciones extremas. Este último resultado la hace especialmente apropiada para captar la localización de un grupo de observaciones de una variable con estructura asimétrica.

**Moda:** es aquel valor que se presenta con mayor frecuencia. Para variable cuantitativa continua, su mayor variabilidad puede hacer que la inspección de los valores individuales nos lleve a que éstos se repiten con escasa frecuencia, resulta conveniente agrupar la variable en una distribución por intervalos, hablando entonces de clase o in-

tervalo modal (intervalos de igual amplitud).

**Moda =  $M_o$  = Valor de la variable con mayor frecuencia**

**Intervalo modal = Intervalo de la distribución con mayor frecuencia**

La moda puede utilizarse con datos cualitativos, representando la categoría más frecuente de la variable. En la mayoría de las ocasiones, en el caso de variables cuantitativas continuas, interesa detectar si los valores de la variable se concentran en torno a un cierto valor de la variable o éstos tienden a concentrarse alrededor de más de un valor. Estas situaciones se denominan como unimodales o multimodales.

### Medidas de dispersión

Se denominan medidas de dispersión aquellas que pretenden captar y resumir la mayor o menor variabilidad, la mayor o menor concentración, homogeneidad o parecido entre las observaciones de la variable. Se presenta a continuación las medidas de dispersión más frecuentes. Como en las medidas de tendencia central, se partirá de una variable genérica,  $X$ , y de un conjunto de  $n$  observaciones  $(x_1, x_2, x_3, x_4, \dots, x_n)$ .

**Rango o recorrido:** se define como la diferencia entre el mayor y el menor valor de la variable, se calcula:

$$\text{Rango} = R = x_{\max} - x_{\min}$$

El rango es una medida de cálculo sencillo y rápido, puesto que depende sólo del mayor y menor valor de la variable. Sin embargo, es escasamente informativa de lo que sucede con el resto de observaciones, afectándose por la existencia de observaciones extremas.

**Varianza:** es la dispersión que presentan los datos con respecto al valor promedio. Da cuenta de la variabilidad de los valores tomando como referencia al promedio. La fórmula es:

$$\text{Varianza} = s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

$$\text{Coeficiente de Variación} = CV = \frac{s}{\bar{x}} (x100)$$

Debido a que la varianza es muy compleja de interpretar, debido a que el resultado se expresa al cuadrado, por ejemplo litro cuadrado, metro cuadrado, es conveniente utilizar una medida de dispersión más comprensible como lo es la desviación estándar.

Desviación estándar: es la distancia a lo largo de un eje horizontal entre el promedio y el punto donde la curva pasa de convexa a cóncava, brinda información del porcentaje del área bajo la curva que se logra tener, es decir, el porcentaje de los valores de que se encuentran bajo la curva de distribución normal. Para Andrade, (2019) “se consideran como normales los valores ubicados entre la media y más menos dos desviaciones estándar, que equivalen al 95.4% “. La desviación estándar se obtiene simplemente calculando la raíz cuadrada de la varianza y la fórmula es:

$$\text{Desviación típica o estándar} = s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

Coeficiente de variación: la varianza y desviación típica representan medidas absolutas de la dispersión de un conjunto de observaciones de una variable. Su interpretación depende de la unidad de medida de la variable así como de su localización. Esto hace que las desviaciones típicas o varianzas de variables distintas sobre un mismo conjunto de datos o de la misma variable sobre conjuntos de datos diferentes no sean comparables, imposibilitando afirmar en qué caso hay mayor o menor variabilidad. Para resolver este problema se puede recurrir al siguiente coeficiente:

El resultado ideal para este indicador es 30% o menos, indicando poca dispersión de los datos analizados.

### Estadística inferencial

La estadística inferencial se encarga de contrastar los resultados de la muestra y generalizarlos hacia toda la población, considerando un riesgo de que en este proceso se cometa algún error como por ejemplo de un “1%, o un 5% lo que lleva a tener un nivel de confianza del 99% o 95%, dependiendo de las situaciones o del diseño y todo lo que tiene que ver con la parte metodológica de la investigación” (Pantoja & Burbano, 2021). La inferencia permite hacer estudios poblacionales con menor cantidad de recursos (usando muestras representativas, obtenidas por técnicas de muestreo apropiadas al estudio o investigación).

Cuando se habla de análisis de datos, significa la comparación de estadígrafos o porcentajes, dependiendo del tipo de escala de la variable y evaluar si existe diferencia estadísticamente significativa entre los grupos, se procede a determinar la correlación entre las variables, para luego medir la asociación entre una o más variables independientes o X y la variable dependiente, Y. Para ello el análisis estadístico se podría ordenar en tres etapas, criterio propuesto por (Castro, 2019):

1. Comparación del efecto entre grupos que están y no están expuestos a la variable independiente: se utilizan pruebas estadísticas de comparación, según la distribución de las variables, se realizarán pruebas paramétricas (si la variable distribuye normal) o no paramétricas (si la variable no distribuye normal). Según la escala de la variable que se va a comparar tendremos: para variables categóricas se usa la prueba exacta de Fischer, (muestra pequeña  $\leq 30$ ) o prue-



ba de Chi<sup>2</sup>, (muestra >30 o más de dos grupos). Para variables de escala continua, si se distribuye normal, se comparan medias. Si son dos medias se usa la prueba de T student y si son tres o más medias, la prueba de análisis de la varianza (ANOVA).

2. La determinación de la correlación entre la variable independiente y el grado de correlación: la correlación mide el grado de asociación entre dos variables numéricas. Se determina con un coeficiente de correlación, que será de Pearson (r), en el caso que la variable distribuya normal, o de Spearman (si no se distribuya normal). El valor obtenido va de -1 a 1 y el signo + o - indica correlación positiva (las variables se mueven en el mismo sentido) o negativa (las variables se mueven en sentido inverso)
3. La determinación de la asociación entre la variable independiente y el efecto, con la regresión: para determinar el grado de asociación entre una variable independiente y una dependiente, se usa la regresión. La regresión lineal simple, que es una técnica estadística que analiza la relación entre dos variables cuantitativas que distribuyen normal y donde se trata de verificar que existe una relación lineal. Supone que hay una variable fija, que está controlada por el investigador (variable independiente) y otra que no está controlada, que es la variable respuesta, o efecto o dependiente. Supone que es la respuesta al estar expuesto a la variable independiente. La variable respuesta siempre ocupa el eje de ordenadas Y y la variable independiente el eje las abscisas X. Para relacionar dos o más variables independientes a una variable dependiente, se le llama regresión bivariable, o multivariada respectivamente.

Hay dos índices estadísticos muy importantes, propone en su investigación (Gómez, Vivó, & Soria, 2001): la covariancia y el coeficiente de correlación de Pearson.

Covariancia ( $S_{xy}$ ): promedio de los productos de las desviaciones de las dos variables  $P_{xy}=(xi-x)(yi-y)$ . Puesto que las desviaciones tienen signo, la covariancia será positiva cuando la nube de puntos en la representación y vs x tenga una forma lineal ascendente y negativa en el caso contrario. La unidad de medida de la covariancia es el producto de las unidades de medida de x e y.

$$S_{xy} = \Sigma(xi-x)(yi-y)/n-1 = SP_{xy}/n-1$$

Donde:

- $S_{xy}$ : covariancia.
- $SP_{xy}$ : sumatorio de los productos cruzados de las desviaciones de las dos variables.

Coficiente de correlación de Pearson ( $r_{xy}$ ): se obtiene estandarizando la covariancia para eliminar la influencia de las unidades de medida, por lo que es una magnitud adimensional. Se define como, citando a (Gómez, Vivó, & Soria, 2001):

$$r_{xy} = S_{xy}/S_xS_y$$

$r_{xy}$  toma valores entre -1 y +1. (-1: relación lineal perfecta negativa; +1: relación lineal perfecta positiva y 0: relación nula).

El modelo de regresión lineal exige que los datos de la población cumplan los supuestos de linealidad, homocedasticidad, independencia entre las observaciones y normalidad de las distribuciones condicionadas de la variable "y". Estos 3 supuestos se pueden resumir en que los términos de error o residuales para cada una de las distribuciones condicionales deben ser variables aleatorias independientes, distribuidas según la Ley Normal (Gómez, Vivó, & Soria, 2001):

Linealidad: Las medias de "y" están sobre la recta:  $\mu_i = \alpha + \beta x_i$

Homocedasticidad: las variancias de la variable dependiente no se modifican con la variable independiente:

$$\sigma_{2y(x1)} = \sigma_{2y(x2)} \dots = \sigma^2$$

Ausencia de auto correlación: un valor  $y_i$  no influye sobre otro valor  $y_j$ .

Al igual que en el caso de la correlación, se estima el modelo de regresión mediante una recta:

$$\mu_i = \alpha + \beta x_i + \varepsilon_i \text{ Donde } \varepsilon_i: \text{ error residual.}$$

Los coeficientes A y B del ajuste de mínimos cuadrados son los mejores estimadores lineales de los parámetros a y b de la población. Así,  $\alpha = A$  (ordenada en el origen de la recta de correlación) y  $\beta = B$  (pendiente de la recta).

Un ejemplo de aplicación lo constituye el trabajo de laboratorio, muchas veces se necesitan técnicas que detecten errores sistemáticos en los resultados obtenidos con un método respecto a otro de referencia. Una posible relación entre dos métodos es una función lineal:

$$Y = \alpha + \beta x + \varepsilon$$

Donde  $\alpha$  representa el error sistemático constante,  $\beta$  el error sistemático proporcional y  $\varepsilon$  el error no sistemático (aleatorio) de los métodos.

## Conclusión

La estadística que se aplica a las ciencias biológicas se llama bioestadística, y en los últimos años ha tenido un papel fundamental en el desarrollo de las ciencias de la salud, de allí la importancia de aplicar procesos estadísticos en el desarrollo de las investigaciones, los mismos que deben ser cuidadosamente analizados para ser aplicados según las características y diseño de la investigación, considerar que por muy básicos que sean los conocimientos de estadística es necesario comprenderlos de manera adecuada para aplicarlos en una investigación, de tal manera que contribuyan con el éxito de los resultados alcanzados.

Este hecho justifica la importancia de conocer el proceso estadístico desde la obtención de los datos, el análisis, la interpretación hasta presentación de resultados y recomendaciones, ya que todos los métodos y técnicas estadísticas en la actualidad son ampliamente utilizados en todo proceso de investigación científica, la bioestadística es un instrumento de trabajo para todo profesional que aporta al desarrollo de las ciencias biológicas, ya que en todo proyecto de investigación y en todo artículo científico que publique se deberán utilizar métodos estadísticos.

Cada estudiante debe identificar el rol fundamental de la bioestadística al proporcionar técnicas que se constituyen en fundamento para dotar al profesional, de criterios en la toma de decisiones en situaciones de incertidumbre, permitiéndole, además organizar la información clínica de manera sistemática y organizada, con tablas y gráficos que facilitan su interpretación. Es oportuno destacar, que cada estudiante identifique el rol de la estadística en su trabajo diario y la importancia de esta para el avance del conocimiento y el desarrollar nuevos tratamientos que contribuyan al mejoramiento continuo de las ciencias de la salud.

La estadística descriptiva es en sí misma una herramienta fundamental para el análisis de datos de los estudios correspondientes. Los conceptos que se han presentado en este artículo son los básicos y es importante que los investigadores clínicos profundicen en el tema para lograr un mejor entendimiento de los fenómenos en el campo de la medicina. Los indicadores de tendencia central y los de dispersión son herramientas útiles y sencillas de calcular e interpretar, a partir de ellos emitir recomendaciones.

La estadística inferencial se encarga de contrastar los resultados de la muestra y generalizarlos hacia toda la población, considerando un riesgo de que en este proceso se cometa algún error dependiendo de las situaciones o del diseño y todo lo que tie-

ne que ver con la parte metodológica de la investigación. La inferencia permite hacer estudios poblacionales con menor cantidad de recursos (usando muestras representativas, obtenidas por técnicas de muestreo apropiadas al estudio o investigación), la realización de regresiones lineales permite evaluar las proyecciones realizadas al evaluar la bondad y ajuste con el cálculo del R<sup>2</sup>.

Cuando se valora la investigación, se empieza a comprender que la estadística no es una colección de fórmulas matemáticas, sino que se transforma en una herramienta invaluable que otorga sentido al manejo de datos tanto cualitativos como cuantitativos, cuyos resultados y conclusiones no pueden ser abordados desde una perspectiva determinística como ocurre en otras ciencias, por ejemplo, la física. Su potencial es un tesoro que cada especialista de salud debe descubrir.

## Bibliografía

- Andrade, H. (2019). Bioestadística Aplicada Ciencias de la Salud. Fundación Gustavo Palma Calderón, 45. Retrieved 2022, from <https://www.researchgate.net>
- Calvache, J., Barón, F., & Garret, R. (2006). La bioestadística y su aplicación a la investigación en salud. *Revista Facultad Ciencias de la Salud: Universidad del Cauca*, 8(3), 56 -59. Retrieved 2022, from <https://dialnet.unirioja.es>
- Castro, M. (2019). Bioestadística aplicada en investigación clínica: conceptos básicos. *Revista Médica Clínica Las Condes*, 30(1), 50 - 65. doi:<https://doi.org/10.1016/j.rmcl.2018.12.002>
- Gómez, S., Vivó, M., & Soria, E. (2001). Pruebas de significación en Bioestadística. *Revista de Diagnóstico Biológico*, 50(4). Retrieved 2022, from <https://scielo.isciii.es>
- Nolasco, A., & Moncho, J. (2016). Estadística básica en la Ciencias de la Salud. Universidad de Alicante. Retrieved 2022, from <https://rua.ua.es>
- Pantoja, M., & Burbano, D. (2021). Importancia de la bioestadística en odontología: estadística aplicada a la investigación prueba de Kruskal-Wallis. *Revista Universidad y Sociedad*, 13(3), 275 - 282. doi:<https://doi.org/0000-0002-5643-6511>
- Revista Electrónica de PortalesMedicos.com. (2013, Septiembre). La bioestadística aplicada a ciencias de la salud. *Enfermería. Revista Electrónica de PortalesMedicos.com*. Retrieved 2022, from <https://www.revista-portalesmedicos.com>
- Roi, A., Pérez, I., Hunte, V., & Ruiz, E. (2018, Agosto). La estadística como necesidad en la investigación en Salud. *Revista Información Científica*, 97(4). Retrieved 2022, from <http://scielo.sld.cu>
- Velázquez, V., González, R., Velázquez, G., & Martínez, P. (2019). Impacto de la Bioestadística como ciencia. *Correo Científico Médico*(23). Retrieved 2022, from <https://www.medigraphic.com/pdfs>

### CITAR ESTE ARTICULO:

Tobar Moran, M. R., San Lucas, S. F., Reyes Sánchez, Z. G., & Ramírez Hecsher, A. M. (2022). Bioestadística como Herramienta aplicada por estudiantes de la Facultad de Ciencias Médicas en el Aprendizaje dentro del Campo de la Salud. *RECIAMUC*, 6(3), 677-686. [https://doi.org/10.26820/reciamuc/6.\(3\).julio.2022.677-686](https://doi.org/10.26820/reciamuc/6.(3).julio.2022.677-686)

