



Nombre del alumno: Karina Montserrat Méndez Lara.

Nombre del profesor: Agenor Abarca Espinosa.

Nombre del trabajo: Resumen y conclusión del tema asignado.

Materia: Investigación epidemiológica.

PASIÓN POR EDUCAR

Grado: 4

Grupo: "C"

Comitán de Domínguez Chiapas a 22 de abril de 2024.

BIOESTADÍSTICA

Haciendo hincapié la palabra "estadística", derivada del latín "status", que significa posición o modo de pararse. En el lenguaje cotidiano, se utiliza el término estadístico de dos maneras: 1) para referirse a conjuntos de datos específicos y 2) como una disciplina matemática. La estadística matemática es una rama de la ciencia matemática que se aplica en diversos campos prácticos, como la administración, economía, biología y medicina, donde se la conoce como bioestadística. Se presentan varias definiciones de estadística, destacando su papel en la recolección, elaboración, análisis e interpretación de datos. La estadística se compone de dos elementos principales: datos y análisis. Los datos son la materia prima de la estadística, obtenidos a través de mediciones u observaciones, mientras que las técnicas de análisis convierten estos datos en información útil, ambos componentes son fundamentales y se necesitan mutuamente para obtener resultados precisos. Se enfatiza la importancia de comprender y predecir la variabilidad en diversos contextos, como la medicina, donde la estadística puede ayudar a modelar y medir la variabilidad. Se destaca que la estadística proporciona teoremas que relacionan la variabilidad e independencia con la información recolectada de muestras. La bioestadística se define como un método objetivo y racional para probar hipótesis científicas, en ella se describen las estadísticas de salud, que abarcan datos numéricos relacionados con poblaciones, hechos biológicos relevantes para la salud, enfermedades, recursos y acciones de intervención. Es importante conocer los conceptos básicos de bioestadística aplicados y combinados en investigación clínica: **Parámetro:** Es un número que resume los elementos de una población, se denota con letras griegas como μ para la media y σ para la varianza. Es importante diferenciarlo de los parámetros del estudio, que se obtienen de una muestra y se llaman estadísticos. **Estadísticos:** Son números que resumen una muestra y se utilizan para estimar parámetros con como \bar{X} para la media muestral y S para la varianza muestral. **Observar:** Significa medir; y las observaciones son las mediciones realizadas en las unidades muestrales. **Unidad de análisis o unidad muestral:** Es el objeto de interés que se observa o mide, como personas, objetos, etc. **Atributo:** Es la característica de la unidad de análisis que se observa, como la edad, el peso, etc. **Variable:** Es cualquier característica que toma dos o más valores en una población. Los atributos medidos se llaman variables porque cambian entre unidades de análisis. **Variable aleatoria:** Son atributos medidos en un conjunto de individuos de una muestra aleatoria cuyos resultados no pueden anticiparse. Estos son solo datos que nos proporcionan una base sólida para comprender los principios fundamentales de la bioestadística en la investigación clínica.

El proceso de planificación estadística en un proyecto de investigación que implica varios pasos esenciales. Primero, se debe definir el método de selección de la muestra y en segundo lugar calcular su tamaño; luego, se determinan las variables a estudiar, incluyendo su tipo y escala de medición; finalmente, se establece un plan para tabular y analizar la información recolectada. Es crucial que el diseño muestral asegure la representatividad, selección aleatoria y tamaño adecuado de la muestra para obtener resultados válidos y generalizables. Muestreo probabilístico: el método de **aleatorio simple**, supone su homogeneidad en la población y que todos los elementos tienen la misma probabilidad de ser seleccionados; asimismo se señalan ventajas como su simplicidad, pero se advierte sobre la posibilidad de desequilibrio si no se consideran los subgrupos y su ponderación respecto a la población de estudio. El **Muestreo Estratificado:** Agrupa la población en

estratos homogéneos para evitar la subrepresentación de ciertos grupos, calcula el peso ponderado de cada estrato para determinar el tamaño de la muestra; sus ventajas son; prevenir desequilibrios en la muestra, reduce la variabilidad dentro de los estratos y permitir resultados más precisos, como desventajas; complica el diseño muestral, puede reducir el tamaño para cada estrato, no se recomienda más de 8 o 10 estratos y requiere un tamaño muestral más grande. **Muestreo Sistemático:** selecciona elementos de la población a intervalos regulares después de elegir uno al azar; ventajas: buena representatividad y rapidez, puede garantizar una selección equitativa de la población; desventajas: riesgo de sesgo de selección si hay una periodicidad oculta en el orden de la muestra. **Muestreo por Conglomerado:** este es utilizado cuando el muestreo aleatorio simple es costoso y no hay una lista de individuos de la población; por lo que se toma una muestra aleatoria de conglomerados y se seleccionan individuos al azar de cada uno; dentro de sus ventajas será su utilidad para estudios epidemiológicos en diferentes regiones, grandes muestras, estudios ecológicos y en diferentes localizaciones; como desventajas, requiere conocimiento previo de las zonas de estudio, información sobre las áreas a muestrear y uso de estratificación y ponderación de grupos. Por otra parte el muestreo no probabilístico se divide en: **No probabilístico o sin asignación aleatoria:** se basa en la conveniencia o en criterios de cuotas en lugar de probabilidades; ventaja: conveniencia en el acceso, simplicidad, economía y rapidez; desventajas: falta de representatividad y validez externa, imposibilidad de generalizar los resultados, riesgo de sesgos y posible sesgo sistemático. **De asignación aleatoria:** la muestra no se selecciona aleatoriamente, la asignación de la intervención se realiza de manera aleatoria por lo que permite que los grupos sean comparables, su ventaja es la conveniencia en el acceso, simplicidad, economía y rapidez y sus desventajas; la falta de representatividad y validez externa, dificultad para hacer generalizaciones estadísticas, riesgo de sesgos y posible sesgo sistemático.

Para definir el número de elementos que conforman una muestra se basa en dos elementos básicos: **1. Estimación de un parámetro:** busca obtener una estimación representativa de un parámetro de la población de estudio; se consideran tres aspectos: 1) variabilidad del parámetro o probabilidad del evento a estimar; 2) precisión de la estimación, medida por la amplitud del intervalo de confianza, y 3) nivel de confianza, habitualmente fijado en 95%, en donde se ofrecen fórmulas para calcular el tamaño de muestra requerido para estimar una media o una proporción. **2. Comparación de proporciones:** detecta estadísticamente diferencias entre dos grupos, en caso de que existan, requieren supuestos previos, incluyendo la variabilidad del parámetro a estimar, la precisión de la estimación y el nivel de confianza y enfatiza la importancia de evitar errores en el número y calidad de las mediciones, incluyendo sesgos y errores aleatorios, así como la importancia de controlar el sesgo a través de una metodología de investigación adecuada. Supuestos del cálculo: se deben establecer supuestos a prioridad, incluyendo el error tipo 1 (alfa) y el error tipo 2 (beta). El error tipo alfa es la probabilidad de rechazar la hipótesis nula cuando es verdadera. Se fija previamente y combinado se establece en un 5%, aunque puede variar según el estudio. El valor p es la probabilidad de cometer error tipo alfa y se fija a priori junto con el nivel de confianza, compuesto en un 95%. El error tipo beta es la probabilidad de aceptar la hipótesis nula cuando es falsa, además se establece previamente y su valor suele ser del 20%, lo que proporciona una potencia del estudio del 80%. **Tipos de hipótesis:** Las hipótesis pueden ser bilaterales (o de dos colas), lo que significa que postulan la existencia de una diferencia, ya sea mayor o menor, entre los grupos de estudio; también pueden ser unilaterales (o de una cola), donde se postula específicamente una

diferencia mayor o menor que cierto valor. **Probabilidad del efecto estándar (P1) y la probabilidad del efecto propuesto (P2):** Estos términos se refieren a la probabilidad de observar ciertos resultados en el contexto de la hipótesis y se utilizan en el cálculo del valor p. Las **implicaciones que conlleva un cálculo de tamaño muestral incorrecto:** exceso de tamaño muestral puede resultar costoso y poco ético, un tamaño muestral insuficiente puede conducir a conclusiones erróneas debido a la falta de precisión en la estimación o a la incapacidad para detectar diferencias significativas entre los grupos de estudio. La **Importancia del cálculo de tamaño muestral** es el tamaño de muestra para garantizar que los resultados del estudio sean creíbles y extrapolables a la población de estudio y tomar decisiones basadas en evidencia estadística sólida y evitar conclusiones erróneas debido a muestras inadecuadas. Paralelamente se menciona que existen varios métodos para calcular el tamaño de muestra, adaptados a diferentes diseños de estudio, como estudios de correlación, pruebas diagnósticas, estudios de no inferioridad y de control. Según la escala de medición hay dos tipos de variables; las cualitativas y cuantitativas. Las **cualitativas** incluyen las nominales y las ordinales. Los nominales se clasifican en dicotómicas (dos categorías) y policotómicas (más de dos categorías), no tienen un orden establecido y no se pueden cuantificar. Los ordinales tienen un sentido de orden implícito pero no son cantidades, aunque pueden tener puntajes asignados, estos no representan cantidades reales. Las **cuantitativas** reflejan cantidades y se dividen en discretas y continuas; las discretas son valores enteros y las continuas pueden asumir infinitos valores dentro de un rango y se representan en un plano cartesiano.

El plan de tabulación de una base de datos implica almacenar sistemáticamente descripciones de sujetos o unidades de análisis junto con las variables medidas. La estructura se compone de columnas y filas, donde cada columna representa una variable y cada fila corresponde a una unidad muestral. La estadística descriptiva se emplea para describir cuantitativamente conjuntos de datos, utilizando diversos recursos estadísticos, estos incluyen estadígrafos de orden, centralización y variabilidad, los cuales son números resúmenes de los datos recolectados, ya sea de estudios poblacionales o muestras. Los estadígrafos de centralización son importantes para comprender cómo se agrupan los datos en torno a ciertos valores representativos; **Moda:** valor que aparece con mayor frecuencia en los datos, utilizada para describir variables medidas en escalas categóricas o nominales, como el sexo y en escalas ordinales. **Mediana (P50):** es el valor que divide a la muestra en dos partes iguales, es útil para datos medidos en escalas ordinales, discretas o continuas y funciona como un estadístico de posición y de centralización. **Media:** representa el centro de masa de la muestra; es un estadístico central utilizado para variables medidas en escalas continuas con distribución normal y muestras homogéneas y simétricas. El análisis de datos tiene como objetivo principal comprobar hipótesis o establecer relaciones de causalidad en un fenómeno determinado, se basa en el análisis y la elaboración de los datos utilizando métodos probabilísticos para la toma de decisiones; este proceso se puede dividir en tres etapas principales: 1) Comparación entre grupos: se compara el efecto entre grupos que están expuestos a la variable independiente y aquellos que no lo están. Se utilizan pruebas estadísticas para realizar estas comparaciones. 2) Determinación de correlación: analiza la relación entre la variable independiente y el efecto. 3) Asociación mediante regresión: determina el grado de asociación entre una variable independiente y una dependiente. Las medidas de asociación son herramientas estadísticas que nos permiten entender la relación entre una variable explicativa (X) y una variable dependiente. 1) **Riesgo Relativo (RR):** se utiliza en estudios prospectivos para estimar el riesgo de un evento en un grupo expuesto

en comparación con un grupo no expuesto, además es útil en estudios de cohortes. 2) **Odds Ratio (OR)**: representa la relación entre la probabilidad de que ocurra un evento en el grupo expuesto en comparación con el grupo no expuesto. 3) **Hazard Ratio (HR)**: esta medida se utiliza en estudios prospectivos y es el resultado de un análisis de regresión de riesgos proporcionales de Cox.

A modo de cierre, la estadística abarca tanto conjuntos de datos específicos como una disciplina matemática, su aplicación en diversos campos prácticos, como la administración, economía, biología y medicina donde se conoce como bioestadística, destaca su papel crucial en la recolección, elaboración, análisis e interpretación de datos. Es fundamental comprender la variabilidad en diferentes contextos, como la medicina, donde la estadística ayuda a modelar y medir dicha variabilidad. La bioestadística, como método objetivo y racional para probar hipótesis científicas, se fundamenta en conceptos como parámetros, estadísticos, observaciones y variables, esenciales para la investigación clínica. El proceso de planificación estadística en un proyecto de investigación que involucra pasos esenciales, desde la definición del método de selección de muestra hasta el establecimiento de un plan para tabular y analizar la información recolectada. Los diferentes métodos de muestreo, ya sean probabilísticos o no probabilísticos, tienen ventajas y desventajas que deben considerarse en función del diseño del estudio y la representatividad de la muestra. El cálculo del tamaño de muestra es crucial para garantizar resultados válidos y generalizables, evitando muestras inadecuadas que puedan conducir a conclusiones erróneas. La estadística descriptiva y el análisis de datos, mediante estadígrafos y herramientas como el riesgo relativo, el Odds ratio y el Hazard Ratio, permiten comprender y establecer relaciones entre variables en un fenómeno determinado. En resumen, la estadística y la bioestadística son fundamentales para la investigación científica, proporcionando herramientas y métodos para comprender, analizar e interpretar datos de manera objetiva y rigurosa.