



Mi Universidad

BIOESTADISTICA

Nombre del Alumno: Gabriel de Jesús Martínez Zea.

Nombre del tema: Planificación Estadística, Diseño muestral Calculo de tamaño muestral, Definición de variable y síntesis.

Nombre de la Materia: Investigación Epidemiológica Avanzada

Nombre del profesor: DR. Agenor Abarca Espinoza

Nombre de la Licenciatura: Medicina Humana

Comitán de Domínguez, Chiapas. 23 de abril del 2024

PLANIFICACION ESTADISTICA

La planificación estadística es el proceso de definir y organizar la recopilación, análisis e interpretación de datos para tomar decisiones informadas, este debe estar enfocado en cumplir cada uno de los objetivos específicos y el objetivo.

Los pasos de la planificación estadística suelen incluir:

1.- Diseño Muestral	<ul style="list-style-type: none"> Define y diseña el método de selección de una muestra
2.- Calculo del tamaño muestral	<ul style="list-style-type: none"> Define el numero de elementos que conforma la muestra
3.- Definir variables	<ul style="list-style-type: none"> El tipo de variables, la escala de medición y la unidad de medida si es que aplica
4.- Plan de tabulación de la base de datos	<ul style="list-style-type: none"> Describen y analizan de la información recolectada Definen las formas de presentar y analizar la información recolectada

DISEÑO MUESTRAL

MUESTRA: se trata de un subconjunto de individuos o elementos de una población definida que cumple con ciertas propiedades comunes.

- Representativa de la población de estudio: promueven que los resultados tengan validez interna y externa.
- Selección aleatoria: los sujetos blancos deben ser escogidos al azar.
- Tamaño mínimo adecuado: se trata de el numero de individuos necesarios aleatoriamente que debe contener la muestra.

La selección de la muestra aleatoria es indispensable para a obtención de resultados válidos.

Diseño Muestral	
1.- Muestreo probabilístico	<ul style="list-style-type: none"> Aleatorio simple Estratificado Sistemático Por conglomerasa
2.- Muestreo no probabilístico	<ul style="list-style-type: none"> No probabilístico o sin asignación aleatoria De asignación aleatoria

MUESTREO PROBABILISTICO:

Aleatorio simple: asume que la población es homogénea y que todos los elementos de la población tienen la misma probabilidad de ser elegidos en la muestra.

- VENTAJAS:
 - Es relativamente técnico sencillo.
- DESVENTAJAS:
 - La muestra puede ser desequilibrada si no se toma en cuenta subgrupos y su peso ponderado o proporcional respecto a la población de estudio.

Muestreo estratificado: asume que la población es heterogénea y que existen subgrupos, es por aquello que se agrupa la población en unidades homogéneas llamadas “Estratos”. Este diseño tiene por objetivo evitar que algún grupo este menos representado que otro.

- VENTAJAS:
 - previene que la muestra quede desequilibrada respecto a la representación de cada subgrupo o estrato.
 - Disminuye la variabilidad dentro de los estratos
 - Permite resultados más precisos.
- DESVENTAJAS:
 - Complica un poco más el diseño muestral
 - Existen muchos estratos, puede reducir el número muestral para cada estrato
 - Necesita mayor tamaño muestral.

Muestreo sistemático: utilizado cuando los elementos de la población están ordenados

- VENTAJAS:
 - Obtiene buenas propiedades de representatividad
 - Puede garantizar una selección perfectamente equitativa de la población
- DESVENTAJAS:
 - Se puede obtener una muestra que contenga un sesgo en la selección.

Muestreo por Conglomerado: utilizado cuando el muestreo aleatorio simple es demasiado caro, por la magnitud de población y tampoco se tiene un listado de los individuos que la componen.

- VENTAJAS:
 - útil para estudios epidemiológicos que desean abarcar poblaciones de diferentes regiones
 - útil para grandes números de muestras
 - útil para estudios ecológicos

- DESVENTAJAS:
 - exige un conocimiento previo de las zonas de estudio.

No probabilístico o sin asignación aleatoria: es un método de muestreo que no se basa en probabilidades, en un muestreo por conveniencia o de cuota.

- VENTAJAS:
 - es elegido

- DESVENTAJAS:
 - existe un sesgo de selección.

De asignación aleatoria: este método usa los ensayos clínicos controlados , aleatorizados por doble ciego multicéntricos.

- VENTAJAS:
 - conveniencia para el acceso a las unidades muestrales
 - simple
 - económico
 - rápido

- DESVENTAJAS:
 - Falta de representatividad y validez externa.
 - Riesgo de incurrir en sesgos debido al criterio de muestreo empleado.

CALCULO DE TAMAÑO MUESTRAL

Una de las exigencias que debe tener una muestra es que tenga el tamaño mínimo adecuado para que represente cuantitativamente a la población de estudio o blanca.

Este tipo de muestra:

- Estima un parámetro: información cálculo de tamaño muestral para la estimación de un parámetro.
 - Variabilidad del parámetro a estimar
 - Precisión de la estimación amplitud del intervalo de confianza
 - Nivel de confianza: 95%

- Variabilidad del parámetro: estima la probabilidad del evento que se desea estimar, a mayor variabilidad, tiene un atributo, se necesita de medidas y de manera consecutiva una muestra más grande.

- Precisión de la estimación del parámetro: es la amplitud del intervalo de confianza, la amplitud del IC dependerá del objetivo de estudio, debido a que necesita gran precisión o solo una aproximación, se denota con una "d o l".
 - Puede ir desde 0.1 – 0.010

- Comparar proporciones: El objetivo es obtener el tamaño muestral necesario que permite detectar estadísticamente una magnitud de diferencia entre dos grupos , en el supuesto que realmente exista , es decir conseguir la máxima exactitud al medir , significa procurar que exista validez en la medición y prevenir el error aleatorio.

Validez	Se mide lo que se desea son realizar sesgo
Sesgo	Es el error sistemático prevenible , que se comete al efectuar las mediciones y que puede ocurrir por un defecto: <ul style="list-style-type: none"> - En el observador - En el instrumento de medición - En el observador

El sesgo se previene y se controla atravez de la metodología de investigación del estudio que finalmente es el modo de diseñar, ejecutar y analizar el estudio.

- Error aleatorio: también llamado accidental o error de precisión , es debido a pequeñas causas que son imposibles de controlar por el investigador.
 - Error tipo 1 o ALFA
 - Error tipo 2 o BETA

1. Error aleatorio A: Definido como el error que comete el investigador al rechazar la hipótesis nula (H_0) siendo esta verdadera.
Concluye que si existe asociación entre una variable independiente (X) y una variable dependiente (Y) cuando realmente no existe.}
2. Error tipo 1 o A (Valor P): al fijar el error A se fija en el valor p , debido que el valor de p es el valor de la probabilidad de cometer error A, es decir cometer error aleatorio.
 - El valor P es equivalente al valor de significancia A con la diferencia que el valor de p , se calcula a partir de una muestra A posteriori.
 - Es necesario considerar que el valor p , mide la fuerza de la evidencia estadística de un estudio realizado en una muestra.
 - Cuando la menor sea el valor p , mayor fuerza tienen las evidencias del estudio y menor probabilidad de rechazar la hipótesis nula, siendo esta cierta.
 - Un valor p , derivado de un estudio que no ha tenido un cálculo de tamaño muestral para probar una diferencia, que no tiene mayor validez, porque no hay certeza de donde proviene el valor <0.05 .
3. Error tipo 2 o B: es aceptar H_0 cuando no es cierta.
 - Concluye que no existe asociación entre una variable independiente y una variable dependiente cuando realmente si existe
 - Generalmente se usa el error B de 20%, lo que da una potencia de estudio de 80%

TIPO DE HIPOTESIS

- BILATERAL O DOS COLAS: Postula que existe diferencia o que no existe diferencia.
- UNILATERAL O A UNA COLA: postula que existe una diferencia mayor o menor que.

Calculo de tamaño muestral incorrecto:

- EXCESO DE TAMAÑO MUESTRAL:
 - Encarece el estudio
 - No es ético someter a mas pacientes de lo necesario a la intervención o la vulneración de los datos.
- INSUFICIENTE TAMAÑO MUESTRAL:
 - El parámetro estimado será poco preciso.

- Si se desea probar una hipótesis , una muestra con escaso tamaño , será incapaz de detectar la diferencia entre los grupos de estudio y se llegara a una conclusión errónea que no existe diferencia estadística significativa.
- **CALCULO DE TAMAÑO MUESTRAL:** se realiza para que el investigador sepa el numero mínimo , necesario de unidades muestrales que debe estudiar , con el fin que los resultados estadísticos sean creíbles y permitan estimar un parámetro extrapolable a la población de estudio .
 - En caso de probar una inferencia , permite tomar la decisión de aprobar o rechazar la hipótesis.

DEFINICION DE VARIABLES

- **VARIABLES:**
 - INDEPENDIENTE
 - DEPENDIENTE
 - CO-VARIABLES
- **VARIABLES CUANTITATIVAS:**
 - Los que miden cantidad, que pueden ser discretas o continuas.
 - **DISCRETA:** Atributo que no puede tomar valores decimales.
 - **CONTINUA:** Aquel atributo que puede asumir un numero infinito de valores dentro de un determinado rango. Tienen una unidad de medida y tiene una distribución en el plano cartesiano del eje x , también llamadas intervalares.
- **VARIABLES CUALITATIVAS:**
 - Las que miden una cualidad (nominal o categóricas y la ordinal)
 - **NOMINAL O CATEGORICAS:** son aquellas que se clasifican en categorías
 - **DICOTOMICAS:** En que solo hay dos categorías y son autoexcluyentes. Vivo / muerto , enfermo / no enfermo , sexo masculino / femenino.
 - **POLICOTOMICAS:** Hay mas de dos categorías y no necesariamente son autoexcluyentes, causas de insuficiencia cardiaca , hipertensión arterial , arritmia , enfermedad valvular , hipertrofia ventricular.
 - **ORDINAL:** tiene un sentido de orden , que esta implicito , sin que sea una magnitud o cantidad , por lo que no tienen unidad de medida: 1, 2 , 3 ,4,5.

- LA DETERMINACION DE LA ASOCIACION ENTRE LA VARIABLE INDEPENDIENTE Y EL EFECTO CON LA REGRESION.
 - La regresión supone que hay una variable fija, que esta controlada por el investigador (variable independiente) o predictora (que predice el efecto). También considerada como la exposición o posible causa. Y otra que no está controlada que es la variable respuesta o efecto o dependiente.
- ESTADIGRAFOS DE CENTRALIZACION O DE TENDENCIA CENTRAL:
 - MODA: Es el dato de mayor frecuencia de aparición. Apropiaada para describir datos medidas en escala categórica o nominal.
 - MEDIANA (P50): Es el punto que divide a la muestra en dos partes iguales. Es apropiada para describir datos medidos en escala: Ordinal, discreta o continua , es un estadígrafo de posición y de centralización.
 - MEDIA O PROMEDIO: Es el punto donde se ubica el centro de masas de la muestra Describe bien solo si la muestra es homogénea y/o simétrica.
- COMPARACION DEL EFECTO ENTRE GRUPOS QUE ESTAN Y NO ESTAN EXPUESTAS A LA VARIABLE INDEPENDIENTE:
 - Pruebas estadísticas de comparación, distribución de la variable, variable distribuye normal: pruebas paramétricas, variable no distribuye normal: pruebas no paramétricas.
 - VARIABLES CATEGORICAS: prueba exacta de Fischer
 - VARIABES DE ESCALA CONTINUA: se distribuye normal, se comparan medias.
- MEDIDAS DE ASOCIACION:
 - RIESGO RELATIVO: se puede estimar solo en los estudios prospectivos
 - ODDS RATIO: solo se calcula en estudios con respuesta dicotómica, sale de una regresión logística.
 - HAZARD RATIO (HR): se puede calcular en estudios prospectivos y es el resultado de la regresión de riesgos proporcionales de Cox.
- ESTADIGRAFOS DE ORDEN, CENTRALIZACION Y VARIABILIDAD:
 - Remarca la importancia de utilizar los estadígrafos correctos de acuerdo con el tipo de variable que se desea describir.
 - Estadígrafos de orden
 - Estadígrafos de tendencia central
 - Estadígrafos de variabilidad

- PLAN DE TABULACION DE LA BASE DE DATOS
 - Una base de datos es una estructura en la cual se almacenan con un orden definido, un grupo de descripciones sobre determinados sujetos o unidades de análisis, ya medidas (variables).

SINTESIS

Permite cumplir con los objetivos del estudio y responder la pregunta de investigación.

- Los resultados dependerán la validez interna del estudio y eso es, una rigurosa prevención y control de sesgos de información, selección y confusión.
- Hay que prevenir el error en quien mide (observador) el error con que se mide (instrumento) y el error a quien se mide (observado).

ANALISIS DE DATOS

La estadística analítica tiene por objetivo comprobar hipótesis o establecer relaciones de causalidad en un determinado fenómeno, también llamado estadística inferencial.

- Es un área de la estadística que se dedica al análisis a la elaboración de los datos con métodos basados en probabilidades para la toma de decisiones en salud.
- Comparación de estadígrafos o porcentajes
- Asociación entre una o mas variables independientes
- Correlación entre las variables

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.
-
- REGRESION LINEAL SIMPLE: analiza la relación entre dos variables cuantitativas que distribuyen normal.
- Una asociación lineal , significa que a medida que la variable independiente está presente , también esta presente en la misma proporción la variable dependiente.

La estadística es una herramienta que no es posible soslayar en la investigación científica , porque es lo que permite probar matemáticamente una hipótesis y extraer conclusiones validas.

- Junto con las medidas de asociación son aquellas medidas que reporta la magnitud de asociación entre la variable explicadora X y la variable dependiente Y, la cual dice la fuerza con que se asocian amabas variables.

- **PLAN DE TABULACION DE LA BASE DE DATOS:** descripción y análisis de la información recolectada.
Protección de la confidencialidad de la información: solicita su consentimiento informado para registrarlos, describirlos y analizarlos
Los análisis estadísticos deben realizarse sobre bases de datos codificadas, donde no exista información que se pueda identificar al sujeto de estudio.
- **LA DETERMINACION DE LA CORRELACIONM ENTRE LA VARIABLE INDEENEDIENTE Y EL EFECTO CON EL COEFICIENTE DE CORRELACION.**
 - La correlación es el grado de variación conjunta de dos variables numéricas o continuas.
- **DESCRIPCION DE DATOS:**
La estadística descriptiva tiene el objetivo de describir cuantitativamente un conjunto de datos, que para ello se utilizan diferentes recursos estadísticos , tales como los estadígrafos de orden , centralización y variabilidad , que son números resúmenes de los datos recolectados , que pueden venir de estudios poblacionales o muestrales.

CONCLUSION

En conclusión, la planificación estadística desempeña un papel fundamental en la obtención de datos confiables y representativos para la toma de decisiones informadas. A lo largo de este trabajo, hemos explorado la importancia del diseño muestral, La planificación estadística que es crucial en cualquier estudio o investigación, ya que nos permite obtener datos representativos y confiables para sacar conclusiones válidas, los muestreos probabilísticos, el cálculo del tamaño muestral, los tipos de hipótesis, la definición de variables y el análisis de datos en la obtención de resultados significativos y generalizables.