

## ESCRIBIENDO HACERCA DE DISTRIBUCIONES Y ASOCIACIONES

La información distribuida en variables simples o asociaciones entre dos o tres variables provee el fundamento de selección de resultados en artículos científicos y es también incluido en artículos de interés general.

En un artículo acerca de educación elemental, se debe de describir la distribución del tamaño de la clase, posteriormente demostrar como el tamaño de clase, los gastos por estudiante y los logros de los estudiantes están relacionados entre si.

En un análisis estadístico básico o reporte acerca de datos experimentales, las descripciones de la distribución y las bi-variables o asociaciones de tres vias normalmente constituyen un análisis completo.

Para análisis de estadísticas más avanzadas como son las regresiones multivariadas se necesitan técnicas diferentes.

### ➤ DISTRIBUCIONES DE UNA SOLA VARIABLE

Las estadísticas univariantes proveen información de una variable a la vez, mostrando casos que son distribuidos a través de todos los posibles valores de tal variable.

En un artículo científico crear en una tabla con resúmenes de cada variable. Después referir a la tabla como se describen los datos.

En un artículo de interés general, indicar la medida o los valores de las variables principales en prosa, sustituyendo porcentajes por significados o por un modelo.

Si se van a comparar muestras o poblaciones, reportar la frecuencia de las distribuciones usando porcentajes para ajustar las diferencias de los tamaños de las muestras.

Las distribuciones de una sola variables se dividen en variables categóricas y variables continuas.

### EJEMPLO:

Cien exámenes aprobados en una muestra de mil estudiantes refleja una proporción muy diferente de cien exámenes aprobados entre varios millones de alumnos.

### VARIABLES CATEGORICAS

Para mostrar la composición de una variable categórica, presentar la frecuencia de cada categoría como una cuenta ( Hispanos no blancos, hispanos no negros o Niños Americanos Mexicanos) o porcentaje (el porcentaje de infantes por cada grupo racial).

Informar la categoría modal desde la media (la media aritmética). Si la variable tiene pocas categorías, reportar información numérica por cada categoría en el texto.

Para variables que contengan diversas categorías, crear un mapa, cuadro sinóptico o tabla para hacer un resumen del texto. También tener en cuenta si dos o más categorías pequeñas se pueden combinar en una sola sin ocultar algún grupo de interés para el análisis.

- En algunos casos, estas combinaciones se basan en similitud de conceptos. Por ejemplo: En instituciones públicas, privadas y escuelas parroquiales, se pueden combinar todas las escuelas parroquiales en una sola categoría independientemente de su afiliación religiosa.
- En otros casos, estas combinaciones se utilizan para evitar tabular muchas categorías raras por separado. Por ejemplo: Se pueden combinar conjuntos heterogéneos de grupos raciales poco mencionados como “otros” aunque compartan poco de su rareza. Ya sea explicar que “otros” incluye a razas distintas a las mencionadas en el gráfico o incluir una nota indicando que es lo que incluye la categoría de “otros”.

#### VARIABLES NOMINALES

- Para las variables nominales, tales como la raza/etnia o el gasto dentro del presupuesto federal, se emplean los principios tales como la frecuencia de ocurrencia o criterios teóricos para organizar números dentro del texto.
- No está de más mencionar la categoría más habitual y a continuación describir donde entran las otras categorías, utilizando una o más comparaciones cuantitativas para evaluar las diferencias de proporciones relativas de las distintas categorías.
- EJEMPLOS
- Un mal ejemplo: “La distribución de los desembolsos presupuestarios federales de los EU en el año 2000, fueron de 61% de los recursos humanos, 12% de intereses, 16% para la defensa nacional, el 6% para otras funciones y el 5% para recursos fiscales”
- Un buen ejemplo: “Más del 60% de los desembolsos presupuestarios federales de los EU en el 2000, fueron de servicios humanos, con la defensa nacional en segunda instancia (16%), y los intereses netos en el tercero (12%). El resto de los desembolsos fueron aproximadamente iguales entre los recursos fiscales y otras funciones”.
- Si una categoría es de interés especial, se puede describir antes.

#### VARIABLES ORDINALES

- Para estas variables, el orden natural de las categorías con frecuencia dicta el orden en el que se reporta de ellos: Ejemplo: grupos de jóvenes, adultos y ancianos.
- Sin embargo en ocasiones se desea destacar cierto grupo por ser mayor o de particular importancia, el cual puede ser primeramente mencionado sin importar el orden normal de la secuencia. O si se quiere hablar del auge de alguna generación se puede hacer.

- En estos casos iniciar con una descripción general. Se puede hacer en “forma de campana”, “uniformes”, “Forma de U”, o “sesgada hacia la derecha” y continuar describiendo el grupo que se desea.
- EJEMPLOS
- Mal ejemplo: “En EU, las edades de adultos para el año 2000 fueron de 8.6% de entre 20-24, 8.8% de entre 25 a 29, ..., y 1.9% de 85 años en adelante. “
- Buen ejemplo: “En el año 2000, en EU la edad adulta apenas tenía forma de campana con edades de entre 20 y 55 años, reflejando la presencia de un boom dominante de infantes nacidos a finales de los 40s y hasta los 50s. Las cifras más largas son edades de entre 35-39 y 40-44, con 10.3% y 10.1% de la población adulta. Después de los 55 años, la distribución por edades colapsa rápidamente revelando los efectos de mortalidad y nacimientos.”

#### VARIABLES CONTINUAS

- Para las variables continuas como son el ingreso, el peso, la edad, se crea una tabla para reportar los valores mínimos y máximos, las medidas de tendencia central y la desviación estándar de las variables utilizadas en el análisis. Si la distribución es inusual por una de las variables claves, trazar la distribución en un gráfico ya que es más fácil de comprender.
- En algunos casos afecta como se clasifican los datos o conduce el análisis. Para esto clasificarlos en otro análisis. Por ejemplo: Si se comparan los resultados de diferentes cuantiles de la distribución del ingreso, etiquetar cada categoría mediante la escala de ingresos pertinentes en los cuadros de estadísticas descriptivas y resultados de la regresión.
- En el texto dar las estadísticas resumidas a menos que existan muy pocos casos. Por ejemplo, en un jardín de niños que tenga 4 grupos de preescolar, enlistar el número de estudiantes de cada grupo seguido por la media general.
- Ejemplo: “La escuela primaria Kennedy abrió sus puertas con clases de preescolar de 21, 24, 27 y 28 alumnos, con una media de 25 estudiantes por clase”
- O para describir algo a nivel más general, con tan solo informar el rango y media o la mediana.
- Ejemplo: “ En un día abierto en la ciudad de Nueva York, el tamaño de las clases de preescolar, tenía un rango de entre 17 y 31 estudiantes, con un promedio de 22.1 alumnos.”
- Para proporcionar más información crear una tabla y después describir de forma general la forma de la distribución e informar los valores específicos de interés.