

Material de apoyo para el curso de Estadística Multivariante

¿Qué es la estadística multivariante?

De acuerdo a Hair (2007), la estadística multivariante es la parte de la estadística y del análisis de datos que estudia, analiza, representa e interpreta los datos que resultan de observar un número de variables estadísticas sobre una muestra de individuos.

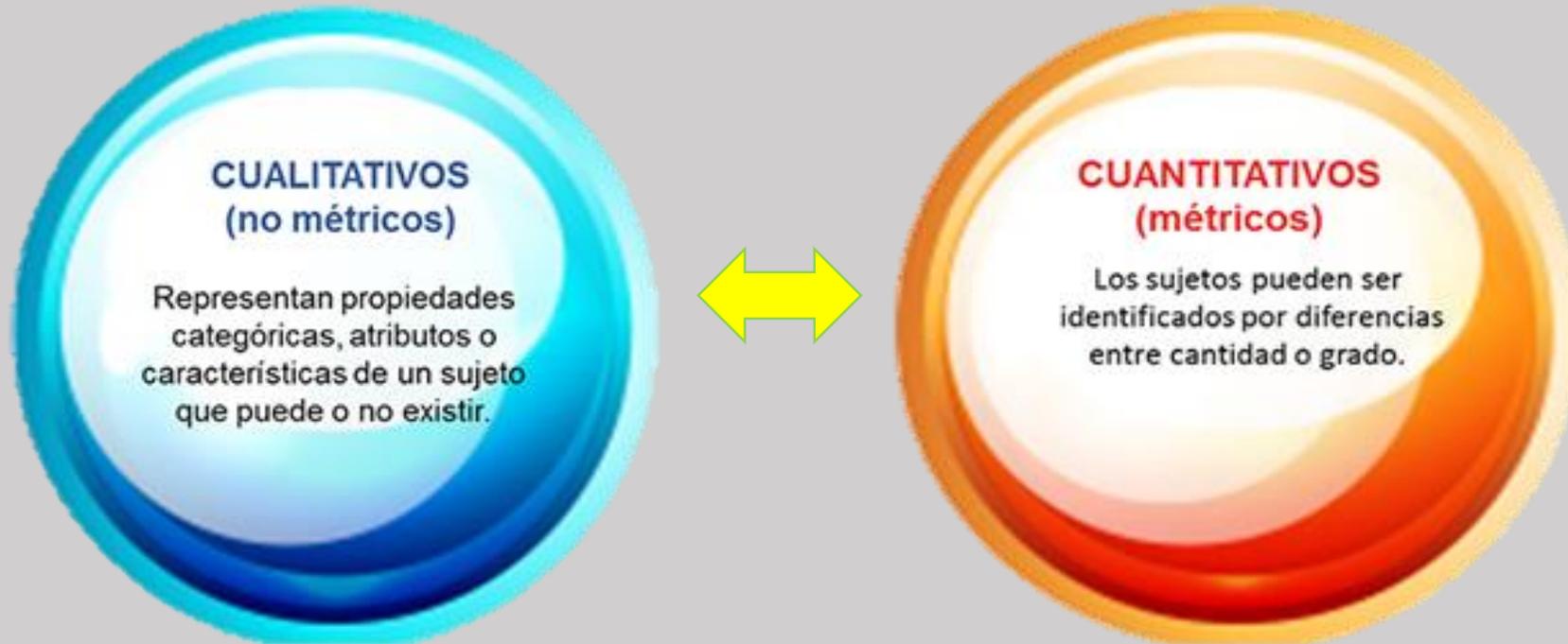
Palabra clave que según el investigador Malhotra explica que puede identificar una variación, si ésta no es cuantificable.

La medida representará el valor por el cual se selecciona el método de análisis multivariado apropiado.

Medida

Recuerda que:

Existen dos tipos básicos de datos:



En la
identificación
de datos

Es necesario identificar un **dato nominal, ordinal, de intervalo o de razón**, a fin de darle el tratamiento adecuado y poderlo interpretar correctamente.

Es indispensable para identificar si un par de variables está relacionado y de qué forma.

El coeficiente
de
correlación

Análisis de regresión

Es una técnica estadística que tiene como objetivo establecer modelos matemáticos, para representar formalmente las relaciones de dependencia existentes entre un conjunto de variables estadísticas (Malhotra, 2008; Pérez, 2014).

En este análisis están involucradas dos variables: una independiente y otra dependiente únicamente; mientras que en el análisis denominado regresión múltiple están involucradas más de dos variables independientes y una sola variable dependiente.

Análisis denominado regresión simple

ES

SIRVE

Una forma directa de buscar diferencias entre más de dos grupos de respuestas, medidas en escalas de intervalos o de razón.

Análisis de varianza

Para determinar si dos morteros de diferentes lotes son iguales o diferentes.

Tanto el **análisis de varianza**, como el **análisis de covarianza**, se utilizan para examinar las diferencias entre los valores promedio de la variable dependiente asociada con el efecto de las variables independientes controladas, después de tomar en cuenta la influencia de las variables independientes no controladas.

El análisis de varianza **ANOVA** se usa como una prueba de medias para dos o más poblaciones (Pérez, 2014).

El objetivo principal de la técnica de análisis de varianza, en su modalidad de un sólo factor, es determinar si un conjunto de muestras tienen la misma media o no.

Análisis de varianza

La ANCOVA

Es un procedimiento estadístico que permite eliminar la heterogeneidad causada en la variable de interés (dependiente), por la influencia de una o más variables cuantitativas denominadas covariables. El fundamento del ANCOVA es un ANOVA al que, a la variable dependiente, se le ha eliminado el efecto predicho por una o más covariables por regresión lineal múltiple. El incluir covariables aumenta la potencia estadística porque reduce la variabilidad.

Análisis factorial

Es una técnica estadística de reducción de datos que se utiliza para explicar qué tanto varían las **variables observadas** en función de un número menor de variables llamadas **factores**.

Se asemeja al análisis de correlación simple en el hecho de que busca encontrar relaciones entre variables. Sin embargo, este tipo de análisis va mucho más allá tratando de predecir múltiples variables dependientes, a partir de múltiples independientes demostrando su validez. El análisis de correspondencias es una técnica muy utilizada entre los investigadores, sobre todo cuando se trata de posicionar marcas, productos o servicios, formas de distribución, etcétera.

Análisis de correlación canónica

Análisis discriminantes

Es una técnica multivalente que se puede utilizar para generar reglas con las que se puede clasificar una población de forma apropiada. Existen muchas aplicaciones del análisis discriminantes en la actualidad.

Análisis de sensometría y caracterización

El análisis sensométrico o evaluación sensorial es una disciplina científica muy útil para evocar, medir, analizar e interpretar el complejo de sensaciones o reacciones que experimentan las personas, frente a las características o atributos de un determinado producto o servicio. Esta evaluación, por lo general, conlleva los cinco sentidos y puede medir la preferencia y la aceptación que dan los consumidores al producto (DosSantos, 2006).

La evaluación sensorial permite la caracterización de un producto. Al realizar una evaluación se busca que los consumidores reconozcan las cualidades y defectos del mismo. Este tipo de análisis es una técnica reproducible para identificar, cuantificar y describir las características de un producto y determinar la calidad sensorial; este es un método que permite evaluar tanto individual como integralmente, las características sensoriales y las relaciones entre ellas, obteniéndose la mayor información posible de una muestra (Sanmiguel, 2014).

Minería de datos

La minería de datos puede definirse como un proceso de descubrimiento de relaciones nuevas y significativas, patrones y tendencias al examinar grandes cantidades de datos (Hair, 2007).

Minería de datos textuales

Para ello existe la minería de datos textuales o *text mining*. Esta es el área de investigación todavía más reciente que se basa en el procesamiento de textos, y está encargada del descubrimiento de conocimientos que no existían explícitamente, pero que surgen de relacionar el contenido de varios documentos textuales.

Minería de datos textuales II

En el tema de minería de datos conociste la minería de datos textuales y algunos de los términos más comunes en su argot. Dentro de ellos cobra importancia la *lexicometría*, que se refiere a ciertos procedimientos que implican contar las ocurrencias de las unidades verbales básicas (generalmente palabras), y operar algún tipo de análisis estadístico a partir de los resultados de tales recuentos.

Ecuaciones estructurales:

Combinan la regresión múltiple y el análisis factorial, con la intención de evaluar las muy complejas interrelaciones de dependencia e incorporar los efectos del error de medida sobre los coeficientes estructurales al mismo tiempo.