

Análisis de datos multivariados

Mtro. Carlos Estrada Zamora



Análisis multivariado

Término general para los procedimientos estadísticos que analizan simultáneamente múltiples mediciones de cada individuo u objeto en estudio.

Algunas técnicas de análisis multivariado que analizaremos:

Análisis de regresión múltiple

Análisis discriminante simple

Análisis de clústers

Análisis factorial

Análisis conjunto



Análisis de regresión múltiple

Procedimiento para predecir el nivel o magnitud de una variable dependiente (métrica) con base en los niveles de múltiples variables independientes.

Aplicaciones del análisis de regresión múltiple

Estimación de los efectos de diversas variables de la mezcla de marketing sobre ventas o participación en el mercado.

Estimación de la relación entre varios factores demográficos o psicográficos y la frecuencia con que son visitadas ciertas empresas prestadoras de servicios.

Determinación de la influencia relativa de elementos particulares de satisfacción en la satisfacción general.

Cuantificación de la relación entre diversas variables de clasificación, como edad e ingreso, y la actitud general hacia un producto o servicio.

Determinación de qué variables son predictivas de ventas de un producto o servicio particular.

El análisis de regresión múltiple puede servir para:

Predecir el nivel de la variable dependiente con base en niveles dados de las variables independientes.

Comprender la relación entre las variables independientes y la variable dependiente.

Medidas del análisis de regresión múltiple

Coefficiente de determinación (R^2)

Medida del porcentaje de la variación en la variable dependiente explicado por variaciones en las variables independientes.

R^2 ajustada

Corrige el coeficiente de determinación con base en la relación entre el número de variables predictoras y el tamaño de muestra total, produciendo una estimación más racional del ajuste del modelo cuando se incluyen diversas variables independientes. Siempre será menor a R^2 y negativa en algunos casos de muestras pequeñas.

Coefficiente de regresión

Estimación del efecto de variables independientes individuales en la variable dependiente.

Variables ficticias

En el análisis de regresión, forma de representar variables independientes nominalmente escaladas de dos grupos, o dicotómicas, codificando un grupo como 0 y el otro como 1.

Posibles problemas de uso e interpretación

Colinealidad. Correlación de variables independientes entre sí, que puede sesgar estimaciones de coeficientes de regresión.

Causalidad. Inferencia de que un cambio en una variable es responsable de (o causó) un cambio observado en otra variable.

Escalamiento de coeficientes. Método para comparar directamente las magnitudes de los coeficientes de regresión de variables independientes escalándolas en las mismas unidades o estandarizando los datos.

Tamaño de la muestra.

ANÁLISIS DISCRIMINANTE MÚLTIPLE

Procedimiento para predecir la membresía a un grupo de una variable dependiente (nominal o categórica) con base en dos o más variables independientes.



En el caso del análisis de regresión múltiple, la variable dependiente debe de ser de escala métrica; en el **análisis discriminante múltiple**, la variable dependiente es de naturaleza nominal o categórica (uso de números para indicar membresía a un grupo (p. ej. 1=hombre, 2=mujer)).

Metas del análisis discriminante múltiple

Determinar si hay diferencias estadísticamente significativas entre los perfiles promedio de puntaje discriminante de dos (o más) grupos (usuarios o no usuarios).

Establecer un modelo para clasificar individuos u objetos en grupos con base en sus valores en las variables independientes.

Determinar qué tanto de la diferencia en los perfiles promedio de puntajes de los grupos es explicado por cada variable independiente.



Puntaje discriminante

Puntuación que sirve de base para predecir a qué grupo pertenece un objeto o individuo particular; también llamado puntaje de Z.

Coefficiente discriminante

Estimación de la potencia discriminante de una variable independiente particular; también llamado ponderación discriminante.

Aplicaciones del análisis discriminante múltiple

¿En qué se diferencian los consumidores que adquieren varias marcas de aquellos que no adquieren esas marcas?

¿Cómo seleccionar a probables compradores de un nuevo producto a partir de nuestra base de datos de clientes existentes a fin de realizar la más efectiva campaña de prelanzamiento de marketing?

¿En qué se diferencian los consumidores que frecuentan un restaurante de comida rápida en características demográficas y de estilo de vida de aquellos que frecuentan otro restaurante de comida rápida?

¿En qué se diferencian entre si los consumidores que han elegido seguro de indemnización, cobertura HMO o cobertura PPO respecto al uso de servicios de salud, percepciones y actitudes?

Contenido obtenido de:

INVESTIGACIÓN DE MERCADOS

McDaniel y Gates (2016)

Cengage Learning

Décima edición