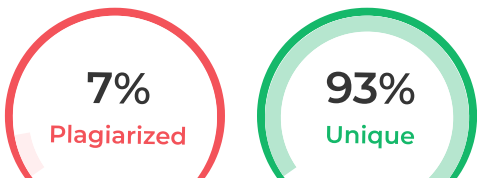


Plagiarism Scan Report



Characters:6702	Words:949
Sentences:46	Speak Time:8 Min

Excluded URL	None
--------------	------

Content Checked for Plagiarism

Marco teórico ### ¿Por qué utilizar técnicas de estadística multivariante?

Las técnicas de estadística multivariante son ideales para analizar conjuntos de datos que abarcan múltiples variables simultáneamente. A diferencia de los modelos univariados, que se centran en una sola variable a la vez, estas técnicas exploran las complejas interrelaciones entre varias variables. Esto permite una comprensión más profunda y completa de los datos. El autor [anderson], explica lo útiles que son para identificar interdependencias entre variables que podrían pasarse por alto en análisis univariados. En la investigación moderna, donde las relaciones entre variables son la norma, esta capacidad para descubrir conexiones ocultas es fundamental. Junto a anderson el autor [hair] resalta los beneficios de estas técnicas, como la capacidad para capturar la complejidad inherente de conjuntos de datos multidimensionales que permite comprender mejor las relaciones subyacentes. No obstante, a pesar de sus ventajas, las técnicas multivariadas presentan limitaciones importantes. La interpretación de los resultados puede ser complicada, especialmente en la gestión de grandes conjuntos de datos con numerosas variables. En situaciones donde la interpretación precisa es crucial, como en estudios médicos, esta complejidad puede representar un desafío significativo. La asunción de linealidad también constituye una limitación relevante. Aunque algunos métodos permiten extensiones no lineales, si las relaciones de las variables no son lineales los modelos podrían dar resultados no confiables al momento de predecir. En nuestro contexto, donde buscamos predecir los tiempos de reparación de máquinas con fallos, la diversidad de técnicas disponibles nos proporciona la capacidad de modelar nuestra variable dependiente mediante múltiples técnicas buscando la que mejor se acople a la naturaleza de nuestros datos. Las técnicas que utilizaremos para modelar nuestra variable dependiente son las siguientes: ##### Modelo de regresión múltiple El modelo de regresión lineal múltiple es una extensión del modelo de regresión lineal simple que permite explorar la relación entre una variable dependiente y múltiples variables independientes @montgomery2012introduction. En este modelo, la relación entre las variables se modela mediante un plano o un hiperplano en un espacio de varias dimensiones. Los coeficientes de regresión representan el efecto de cada variable independiente en la variable dependiente @fox2015applied. Los modelos de regresión lineal múltiple son ampliamente utilizados en diversas áreas, como la economía, la sociología, la psicología y la

epidemiología, entre otros @draper2014applied. Se utilizan para predecir o explicar el valor de una variable dependiente en función de múltiples variables independientes. la fórmula general para un modelo de regresión lineal es la siguiente: \myequations{Modelo de regresión lineal}

$$\begin{equation} \large Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \end{equation}$$

- Y es la variable dependiente que estamos tratando de predecir. - β_0 es el intercepto o término constante. - $\beta_1, \beta_2, \dots, \beta_p$ son los coeficientes asociados con las variables predictoras X_1, X_2, \dots, X_p , respectivamente. - ε es el término de error, que representa la variabilidad no explicada por el modelo. Los modelos de regresión lineal se basan en múltiples supuestos, como los siguientes:

- Linealidad**: Se asume que la relación entre las variables predictoras y la variable dependiente es lineal. Esto implica que los cambios en las variables predictoras están asociados con cambios proporcionales en la variable de interés.
- Independencia de errores**: Se asume que los errores (residuos) de la regresión no están correlacionados entre sí.
- Homocedasticidad**: La varianza de los errores debe ser constante en todos los niveles de las variables predictoras. Esto significa que la dispersión de los errores es uniforme en toda la gama de valores de las variables predictoras.
- Normalidad de errores**: Se asume que los errores de la regresión se distribuyen normalmente. Esto significa que los residuos siguen una distribución normal con una media de cero.
- Independencia de variables predictoras**: Es crucial que las variables predictoras en un modelo de regresión sean independientes unas de otras. Una elevada correlación entre estas variables puede entorpecer la interpretación precisa de los coeficientes, afectando la validez del modelo.

Modelos Lineales Mixtos (LMM) Los Modelos Lineales Mixtos (LMM) son una clase de modelos estadísticos utilizados para el análisis de datos en los que las observaciones no son independientes entre sí, como es común en datos longitudinales o de panel @pinheiro2000mixed. Los LMM combinan elementos de los Modelos Lineales Generalizados (GLM) con la capacidad de manejar la estructura de dependencia entre las observaciones. En un LMM, se asume que los datos tienen una estructura jerárquica, donde las observaciones están agrupadas en unidades o clústeres, como en nuestro caso las diferentes disciplinas encargadas de las reparaciones. La inclusión de efectos aleatorios en el modelo permite capturar la variabilidad entre estos grupos, lo que mejora la precisión de las estimaciones y la capacidad de generalización del modelo @gelman2006data. Los LMM se pueden utilizar para modelar tanto respuestas continuas como categóricas, y pueden incluir tanto efectos fijos como efectos aleatorios. Los **efectos fijos** representan las relaciones promedio entre las variables independientes y la variable dependiente, mientras que los **efectos aleatorios** capturan la variabilidad entre los grupos o clústeres @gelman2006data. Algunas aplicaciones comunes de los LMM incluyen estudios longitudinales, para analizar datos recogidos a lo largo del tiempo, estudios multinivel, para analizar datos estructurados en múltiples niveles jerárquicos, y estudios espaciales, para analizar datos recogidos en ubicaciones geográficas diferentes

@gelman2006data. Los LMM se pueden ajustar utilizando diferentes métodos de estimación, como la máxima verosimilitud restringida (REML) o la estimación de la máxima verosimilitud (MLE). Para trabajar con este modelo en R podemos utilizar paquetes como lme4 creado por @bates2015lme4 siendo este paquete útil para ajustar y analizar modelos lineales mixtos de manera eficiente y efectiva. La fórmula general de un modelo LMM es la siguiente:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + \epsilon_{ij}$$

Sources

2% Plagiarized

Aug 29, 2023 — - y es la variable dependiente que estamos tratando de predecir. - x es la variable independiente que utilizamos para hacer la predicción ...

<https://rpubs.com/FaridSayago/1075386>

2% Plagiarized

La variable ϵ se conoce como el error aleatorio del modelo y se ... error, que representa la variabilidad no explicada por el modelo de regresión.

<https://git.cepal.org/cursos/asd/-/blob/main/07-Regresi%C3%B3n.Rmd>

2% Plagiarized

El enfoque principal del tidyverse es promover un flujo de trabajo coherente y fácil de entender, centrado en la manipulación y visualización de datos. readxl: ...

<https://rpubs.com/Jlastra/1147049>

