Examen

Salazar Vega Rodrigo

14/12/2021

{r setup} #knitr::opts\_chunk$set(echo = TRUE) library(nortest) library(factoextra) library(ggplot2) library(polycor) library(ggcorrplot) library(psych) datos <- read.table(file.choose(), header=TRUE) datos <- data.frame(datos) datos\_esc <- scale(datos)

## Resolucion de examen.

Primero cargamos los datos al entorno de trabajo y lo transformamos a un data frame para poder trabajar más comodo don dichos datos, empezaremos con la prueba de hipotesis, ya que en esta podemos manipular los datos tal cual esta, en las demás pruebas es necesario escalar los datos, por eso empezaremos con esta, aplicaremos una prueba de normalidad.

{r cars} cat("Exploracion de datos para ver si tiene una distribucion normal\n") print("Anderson-Darling normality test\n") print(ad.test(datos$laufkont)) cat("\nComo el p-value es menor a 0.5 podemos descartar que tenga una distribucion normal") print(t.test(datos$laufkont, alternative = "less", conf.level = 0.95))

## Analisis factorial.

Como se habia mencionado anteriormente para trabajar esta técnica tenemos que tener los datos escalados, por lo se realiza eso y procedemos a realizar el análisis, utilizamos una funcion la que nos muestra cual es el número de factores optimos para trabajar con nuestro data set, el resultado son los siguientes:

```{r pressure2, echo=FALSE} # Obtenemos la matriz de correlacion policorica mat\_cor <- hetcor(datos\_esc)$correlations #matriz de correlacion policorica plot(ggcorrplot(mat\_cor, type=“lower”, hc.order = T, title = “Grafica de correlaciones”))

# Verificamos que la matriz sea factoriazble cortest.bartlett(mat\_cor, n = 100)->p\_esf cat(“Test”) print(p\_esf$p) cat(“”) print(KMO(mat\_cor))

# Determinar el numero de factores plot(scree(mat\_cor)) plot(fa.parallel(mat\_cor,n.obs=200,fa=“fa”,fm=“minres”))

#Rotacion rot<-c(“varimax”)

bi\_mod<-function(tipo){ biplot.psych(fa(datos\_esc, nfactors = 4, fm = “minres”, rotate = tipo),main = paste(“Biplot con rotación”,tipo),col=c(2,3,4),pch = c(21,18))  
} sapply(rot,bi\_mod)

# Interpretacion modelo\_varimax<-fa(mat\_cor,nfactors = 4,rotate = “varimax”, fa=“minres”) fa.diagram(modelo\_varimax, main = “Gráfico AF”) print(modelo\_varimax$loadings,cut=0)

Obtenemos la matriz de correlaciones, asi como el grafico donde muestra el número de factores requeridos y la relación que tiene las variables con cada factor.  
  
## k-means  
  
Para trabajar con el k-means nos guiamos un poco con el analisis anterior, definimos 3 grupos para trabajar con este análisis y obtenemos el siguiente resultado.  
  
  
```{r pressure1, echo=FALSE}  
#se aplica el algoritmo k-means  
 grupos <- kmeans(datos\_esc, centers = 3, nstart = 25)  
   
 #Graficar los grupos  
 plot(fviz\_cluster(grupos, data = datos\_esc, main = "Clustering K medias"))  
   
 relaciones <- data.frame(grupos\_km = grupos$cluster)