

Trabalho Final

normal.r

```
normal<-function(x, transf="todos", teste="SW", colu)
{
  library(plyr) #biblioteca para a utilização da função arrange.

  #####
  #transformações
  #####
  lg<-log(x) #logaritmo
  sq<-sqrt(x) #raiz quadrada
  cub<-x^1/3 #raiz cubica
  inv<-1/x #inverso
  media<-mean(x)/x #percentual da media

  #####3
  #Criando lista com as transformações
  #####
  trans<-list(lg,sq, cub,inv, media)
  names(trans)<-c("log","Rquadrada","Rcubica","inverso","media")

  #####
  #Testes de normalidade
  #####
  if(teste=="SW") #testes de Shapiro-Wilk para cada uma das transformações
  realizadas
  {
    plg<-shapiro.test(lg)
    psq<-shapiro.test(sq)
    pcub<-shapiro.test(cub)
    pinv<-shapiro.test(inv)
    pmedia<-shapiro.test(media)
  }

  if(teste=="KS") #testes de Kolmogorov-Smirnov para cada uma das
  transformações realizadas
  {
    plg<- ks.test(unique(lg), "pnorm", mean(lg),sd(lg))
    psq<-ks.test(unique(sq),"pnorm", mean(sq), sd(sq))
    pcub<-ks.test(unique(cub),"pnorm", mean(cub), sd(cub))
    pinv <-ks.test(unique(inv),"pnorm", mean(inv), sd(inv))
    pmedia <-ks.test(unique(media),"pnorm", mean(media), sd(media))
  }
```

```
}

if(transf=="todos")
{
  escolha<-c(plg$p.value, psq$p.value, pcub$p.value, pinv$p.value,
  pmedia$p.value) #string com os valores de p encontrado no SWtest
  tipos<-c("lg", "sq", "cub", "inv", "media") #string com os nomes das
  transformações realizadas
  te<-data.frame(tipos) #criando data frame com os nomes das
  transformações
  te[,2]<-data.frame(escolha) #incluindo o p encontrado
  te<-arrange(te, desc(escolha)) #organizando em ordem decrescente os
  valores de p encontrados
  melhor<-te[1,1] #obtendo o nome do teste cujo p apresenta maior chance
  da hipótese nula do SWtest ser verdadeiro

#####
#Histogramas, curvas de densidade e curva normal dos dados transformados
#####
x11() #janela grafica
par(mfrow=c(3,2)) #janela com 3 linhas e 2 colunas

for (i in names(trans)) #loop para construção de gráficos
{
  hist(trans[[i]], probability = T, main=i, xlab = i)
  lines(density(trans[[i]]), col="blue", lwd=2) #densidade      estimada
  da amostra transformada
  curve(dnorm(x, mean=mean(trans[[i]]), sd=sd(trans[[i]])), add=T,
  col="red") #curva padrão para media e DP da amostra transformada
  legend("topright", c("Densidade estimada","Curva padrão"),
  fill=c("blue","red"), horiz=F, cex=0.7)
}
par(mfrow=c(1,1))

#####
#Bloco output melhor transformação
#####
#se a transformação pela raiz quadrada for melhor
if (melhor=="sq")
{
  cat( "\n A matrix com os dados transformados está armazenada na
variável de nome sq")
  cat("\n A transformação pela Raiz Quadrada foi a que apresentou melhor
resultado no teste de normalização com p de ")
  sq<-sq
  return(psq$p.value)
}

#se a transformação pela raiz cubica for melhor
```

```
if (melhor=="cub")
{
  cat( "\n A matrix com os dados transformados está armazenada na
variável de nome cub")
  cat("\n A transformação pela Raiz Cúbica foi a que apresentou melhor
resultado no teste de normalização com p de ")
  cub<-cub
  return(pcub$p.value)
}

#se a transformação logaritmica for melhor
if (melhor=="lg")
{
  cat( "\n A matrix com os dados transformados está armazenada na
variável de nome lg")
  cat("\n A transformação logarítimica foi a que apresentou melhor
resultado no teste de normalização com p de ")
  lg<-lg
  return(plg$p.value)
}

#se a inversão for melhor
if (melhor=="inv")
{
  cat( "\n A matrix com os dados transformados está armazenada na
variável de nome inv")
  cat("\n A transformação pela Inversão foi a que apresentou melhor
resultado no teste de normalização com p de ")
  inv<-inv
  return(inv$p.value)
}

#se a normalização pela media for melhor
if (melhor=="media")
{
  cat( "\n A matrix com os dados transformados está armazenada na
variável de nome media")
  cat("\n A transformação pela Média foi a que apresentou melhor
resultado no teste de normalização com p de ")
  media<-media
  return(psq$p.value)
}
#####
}
#####
#Escolha da transformações
#####

#função para criação de histogramas da transformação escolhida
his<-function(w,z)
{
```

```
x11()
par(mfrow=c(1,1))
hist(w, probability = T, xlab=z, main = z)
lines(density(w), col="blue", lwd=2) #densidade      estimada da amostra
transformada
curve(dnorm(x, mean=mean(w), sd=sd(w)), add=T, col="red")
legend("topright", c("Densidade estimada","Curva padrão"),
fill=c("blue","red"), horiz=F, cex=0.7)
}

#Criação da variável, histograma e demonstração do p encontrado para a
transformação escolhida
if(transf=="sq")
{
  cat("\n A transformação pela Raiz Quadrada foi escolhida, com a matriz
transformada armazenada na variável sq e o teste de normalização com p
de ")
  his(sq, "raiz quadrada")
  sq<-sq
  return(psq$p.value)
}

if(transf=="inv")
{
  cat("\n A transformação pelo Inverso foi escolhida, com a matriz
transformada armazenada na variável inv e o teste de normalização com p
de ")
  his(inv, "inverso")
  inv<-inv
  return(pinv$p.value)
}

if(transf=="media")
{
  cat("\n A transformação pela Razão da Média foi escolhida, com a
matriz transformada armazenada na variável media e o teste de
normalização com p de ")
  his(media, "media")
  media<-media
  return(pmedia$p.value)
}

if(transf=="lg")
{
  cat("\n A transformação logarítmica foi escolhida, com a matriz
transformada armazenada na variável lg e o teste de normalização com p
de ")
  his(lg, "log")
  lg<-lg
  return(plg$p.value)
}

if(transf=="cub")
```

```
{  
  
  cat("\n A transformação pela Raiz Cúbica foi escolhida, com a matriz  
transformada armazenada na variável cub e o teste de normalização com p  
de ")  
  his(cub, "raiz cúbica")  
  cub<-cub  
  return(pcub$p.value)  
}  
}
```

[help_normal.txt](#)

Help

normal{stats} R Documentation

Descrição:

Função para normalizar dados de uma matriz, com opção de testar quais das transformações possíveis apresentam maior chance de produzir uma matriz normal. Será apresentado um histograma com a matriz transformada e linhas da densidade estimada e a curva normal para a uma amostra hipotética com a média e desvio padrão da amostra transformada.

Uso:

```
normal(x, transf="todos", teste="SW")
```

Argumentos:

x: uma matriz de dados numéricos

transf: Transformação a ser realizada. "sq" - Raiz quadrada, "cub" - Raiz cúbica, "inv" - inverso do número, "lg" - Transformação logarítmica, "media" - Razão pela média do total, "todos" - Todas as transformações anteriores serão realizadas e a que apresentar a maior chance da transformação apresentar uma distribuição normal será escolhida.

teste: Escolha de qual teste de normalidade será utilizado. "SW" - Teste de Shapiro-Wilk. "KS" - Teste de kolmogorov-Smirnov.

Detalhes:

A função deverá ser alimentada com uma matriz de uma linha, onde estarão os valores a ser transformados e analisados.

Valores

Será retornada uma variável contendo a matriz transformada e o(s) gráfico(s) de distribuição e o p do teste de normalizade.

Autor

Raymond Sehiji Takiguchi

Referência.

Ausente

Exemplos

#Obtém uma matriz com a transformação que tem a maior chance de apresentar distribuição normal utilizando-se do teste de Shapiro-Wilk

```
x<-matrix(iris[,2])
```

```
normal(x)
```

#Obtém uma matriz com tranformação logaritmica e teste de Kolgoromov-Smirnov

```
x<-matrix(iris[,2])
```

```
normal(x, transf="lg", teste="KS")
```

From:

<http://ecor.ib.usp.br/> - **ecoR**

Permanent link:

http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2016:alunos:trabalho_final:raymond.takiguchi:trabalho_final 

Last update: **2020/08/12 06:04**