

Billy



Sou aluno de doutorado pela Ecologia, trabalho com seleção sexual, mais especificamente com espécies com cuidado paternal, utilizando opiliões como modelos de estudo.

PDFs

[A protocol for data exploration to avoid common statistical problems](#)

Meus exercícios

[billy_01_f.r](#)

[billy_02_f.r](#)

[billy_03_f.r](#)

[billy_04_f.r](#)

[billy_05_f.r](#)

[billy_6_f.r](#)

[billy_07_f.r](#)

[billy_08_f.r](#)

[billy_09_p.r](#)

Trabalho Final

Problema

Tenho uma amostragem relativamente grande ($n = 600$ indivíduos) de machos de uma espécie de opilião em duas categorias comportamentais: machos que conseguem e machos que não conseguem copular. Como qualquer estimativa de tamanho, sua distribuição é normal e, devido ao tamanho amostral, diferenças encontradas entre estas duas categorias são significativas, mesmo ocorrendo num nível muito pequeno (diferenças em 0.01mm, mesmo nível de erro do instrumento utilizado para medi-los) e não sendo biologicamente relevante.

Solução A

Irei gerar uma função que simula N vezes uma regressão logística entre X elementos sorteados de cada categoria de macho e guarda o valor de p dessa regressão. Ao final, exibe uma distribuição de N valores de p (de todas essas simulações), mostrando também a probabilidade de se obter um p significativo (menor ou igual a 0.05).

Comentários PI

Olha, acho que o problema é de significância biológica x estatística. Se vc mostra que o intervalo de confiança da diferença inclui a resolução do teu aparelho já bastaria. Mesmo que passe da resolução, pode ser que a diferença não tenha nenhum significado biológico. Se você pode definir a partir de qual valor a diferença tem alguma importância biológica, o mesmo critério se aplica. Enfim, não me parece um problema estatístico, e sim da interpretação dos seus dados à luz de uma teoria.

Uma pergunta interessante que fica, no entanto, é: se a única diferença que existe está abaixo de um valor relevante definido pelo usuário, qual é a probabilidade do teste detectá-la? Esta é uma avaliação de força do teste, que poderia ser feita com uma simulação assim:

1. Defina a diferença mínima com significado biológico
2. Crie duas populações com médias que diferem neste valor
3. Sorteie amostras de um certo tamanho de cada uma das populações
4. Faça o teste e guarde o valor de p
5. Repita isto aumentando o tamanho da amostra

Com isto vc pode avaliar como a força do teste para indicar uma diferença sem significado aumenta com o tamanho da amostra. Para gerar as populações você pode usar uma distribuição teórica (e.g normal), ou usar os resíduos da análise feita com seus dados, que por definição tem o efeito da diferença entre grupos cancelada, mas preserva a variação independente desta diferença. Daí é só adicionar a a diferença sem significado à metade dos resíduos. No livro de randomização do Manly há vários exemplos deste tipo.

Solução A'

Essa função pode se tornar mais generalizada se o usuário puder escolher qual teste irá simular, randomizando comparações entre duas ou mais amostras.

Comentários PI

Se sobrar tempo resolvendo a primeira proposta é uma extensão válida e bem útil sim.

Solução B

Para resolver esse problema de “super-amostragem”, posso construir uma função que simula, a partir de uma mesma população (no caso, com distribuição normal e média e desvio-padrão iguais aos parâmetros observados na natureza), testes de comparação entre duas (regressão logística ou teste t) ou mais amostras (ANOVA). Serão realizadas N simulações sorteando X elementos da distribuição

normal e atribuindo-os a cada um dos grupos, guardando a proporção de valores de p menores ou iguais a 0.05. Além disso, essas N simulações serão realizadas também com diferentes valores de X , e uma amplitude de esforços amostrais. O resultado mostrado será um gráfico da proporção de valores de p menores ou iguais a 0.05 em função do número de elementos amostrados (X).

Comentários PI

Correto, mas o que vc deve obter aí é uma estimativa do erro tipo I, ou seja, se tudo deu certo esperaria que 5% das simulações indicassem diferença. O que vc quer é uma avaliação do aumento da força com o tamanho da amostra.

Dá para usar uma normal para isto (veja comentário acima), mas aí vc teria que ter algum palpite sobre o desvio-padrão desta normal, que podem, por exemplo, ser estimados dos desvio-padrão dos resíduos. Se vc não quiser assumir uma distribuição teórica, pode usar diretamente os resíduos.

FUNÇÃO

Página de ajuda

Ive.got.the.power
Documentation

package:nenhum

R

Poder do teste

Description:

Ive.got.the.power exibe graficamente a resposta do erro tipo I e do erro tipo II ao aumento do

esforço amostral para diversos teste de comparação de médias entre grupos. Para aqueles testes para

os quais existe solução analítica para o cálculo do poder do teste, a função Ive.got.the.power também

retorna esse valor.

Usage:

Ive.got.the.power (diff, media, sd, n.sim=3000, N=200, categorias=20, alpha=0.05, teste)

Arguments:

diff Numérico ou vetor numérico. Diferença máxima entre as médias dos grupos a serem comparados no teste.

No caso da simulação de um teste-t pareado, diff deve ser um vetor de comprimento 2, no qual o primeiro elemento representa a diferença média entre os pares e o segundo elemento, o desvio-padrão dessas diferenças. No caso de uma ANOVA, este valor é a diferença entre os grupos mais extremos.

media Numérico. Menor média a ser comparada no teste.

sd Numérico. Desvios-padrão dos grupos. Uma premissa da simulação realizada pela função Ive.got.the.power é que todos os grupos apresentam o mesmo desvio-padrão.

n.sim Numérico. Quantidade de simulações a serem realizadas para cada categoria de tamanho amostral.

N Numérico. Tamanho amostral máximo em que os testes serão simulados.

categorias Numérico. Quantidade de categorias de esforço amostral que serão simulados entre 5 unidades amostrais (esforço amostral mínimo implementado automaticamente pela função Ive.got.the.power) e o tamanho amostral máximo, indicado pelo usuário.

alpha Numérico. Nível de significância do teste a ser simulado.

teste O nome do teste a ser simulado. As opções implementadas pela função Ive.got.the.power são: "teste.t", para realizar simulações de testes-t de duas amostras; "t.pareado", para realizar simulações de testes-t pareados entre duas amostras; "anova", para realizar simulações de análises de variância entre 3 grupos (com mesmo esforço amostral em cada um); ou "logística", para realizar simulações de regressões logísticas utilizando dois grupos como variáveis independentes.

Details:

Para as simulações de erro do tipo I, é criado apenas um grupo, com $n.sim*N$ elementos, com distribuição normal de

média=media e desvio-padrão=sd. São simuladas $n.sim$ amostragens de elementos em dois grupos (ou três, no caso de um

teste de ANOVA) provindos da mesma população, para cada categoria de esforço amostral, regularmente distribuídos

entre 5 e N unidades amostrais. Para as simulações de erro do tipo II, é criado um ou dois grupos adicionais, também

com $n.sim*N$ elementos, seguindo uma distribuição normal de desvio-padrão=sd, mas de média=media+diff. Neste caso, são

simuladas $n.sim$ amostragens de elementos em cada um desses grupos para cada categoria de esforço amostral. Os gráficos

mostram os valores médios de p , a proporção das simulações que obtiveram valores iguais ou mais extremos que α e a

diferença média entre a média dos grupos, para $n.sim$ simulações em cada categoria de esforço amostral. Quando

apresentadas, a linha sólida negra representa a tendência dos valores indicados no eixo y e as linhas tracejadas

azuis representam os limites do envelope de confiança de 95%.

Adicionalmente, para o teste-t para duas amostras e

para o teste de ANOVA, também é retornado um cálculo analítico do poder do teste (equivalente aos gráficos de erro

tipo II).

Value:

São gerados seis gráficos: três ilustrando as simulações de erro tipo I e outros 3, as simulações de erro tipo II.

Quando possível, o poder do teste também é calculado analiticamente e exibido na tela, juntamente com um contador

de tempo de todo o processo.

Author(s):

Gustavo Requena Santos

billy_requena@gmail.comr

References:

Dalgaard, P. (2002) Introductory Statistics with R, Springer ISBN
0-387-95475-9

See Also:

'power.prop.test', 'power.t.test' e 'power.anova.test', do pacote (stats),
para resoluções analíticas em

situações não abordadas na função `Ive.got.the.power`

Examples:

#Teste-t para duas amostras

```
Ive.got.the.power (0.3, 5.85, 0.5, n.sim=100, N=250, categorias=30,  
alpha=0.05, teste="teste.t")
```

#Teste-t pareado

```
vetor <- c(0.3, 0.07)
```

```
Ive.got.the.power (vetor, 7.3, 1.05, n.sim=100, N=200, categorias=40,  
alpha=0.05, teste="t.pareado")
```

#ANOVA

```
Ive.got.the.power (6, 73, 8.5, n.sim=100, N=200, categorias=40, alpha=0.05,
```

```
teste="anova")
```

```
#Regressão logística
```

```
Ive.got.the.power (0.05, 5.65, 0.15, n.sim=100, N=300, categorias=40,  
alpha=0.05, teste="logistica")
```

Código da Função

```
Ive.got.the.power <- function (diff, media, sd, n.sim=3000, N=200,  
categorias=20, alpha=0.05, teste)  
{  
  if(class(diff)=="numeric")  
  {  
    if(class(media)=="numeric")  
    {  
      if(class(sd)=="numeric")  
      {  
        if(class(n.sim)=="numeric")  
        {  
          if(class(N)=="numeric")  
          {  
            if(class(categorias)=="numeric")  
            {  
              if(class(alpha)=="numeric")  
              {  
                if(alpha>= 0.001 && alpha <= 0.999)  
                {  
                  t0=proc.time()[[3]]  
                  amostras <- round(seq(5, N, by=((N-5)/(categorias-1))),0)  
  
                  teste.t <- function (a1, a2, amostras, n.sim, N, categorias, alpha,  
par="F")  
                  {  
                    dist.p <- rep(NA,categorias)  
                    sup.IC.p <- rep(NA,categorias)  
                    inf.IC.p <- rep(NA,categorias)  
                    prop.p <- rep(NA,categorias)  
                    dif.obs <- rep(NA, categorias)  
                    sup.IC.dif <- rep(NA,categorias)  
                    inf.IC.dif <- rep(NA,categorias)  
                    if(par=="F")  
                    {  
                      for(i in 1:categorias)  
                      {  
                        g1 <- matrix(sample(a1, n.sim*(amostras[i])), nrow=n.sim)  
                        g2 <- matrix(sample(a2, n.sim*(amostras[i])), nrow=n.sim)  
                        c=array(c(g1,g2),dim=c(n.sim,amostras[i],2))
```

```
      resultados <- apply( c, 1, function(x) t.test(x[,1],
x[,2])$p.value)
      resultados1 <- apply( c, 1, function(x) t.test(x[,1],
x[,2])$estimate[1])
      resultados2 <- apply( c, 1, function(x) t.test(x[,1],
x[,2])$estimate[2])
      resultados3 <- resultados1 - resultados2
      dist.p[i] <- mean(resultados)
      sup.IC.p[i] <-
sort(resultados)[ceiling(0.975*length(resultados))]
      inf.IC.p[i] <-
sort(resultados)[ceiling(0.025*length(resultados))]
      prop.p[i] <- sum(resultados<=alpha)/n.sim
      dif.obs[i] <- mean(resultados3)
      sup.IC.dif[i] <-
sort(resultados3)[ceiling(0.975*length(resultados))]
      inf.IC.dif[i] <-
sort(resultados3)[ceiling(0.025*length(resultados))]
      texto <- "TESTE-t"
    }
  }
  if(par=="T")
  {
    for(i in 1:categorias)
    {
      g1 <- matrix(sample(a1, n.sim*(amostras[i])), nrow=n.sim)
      g2 <- matrix(sample(a2, n.sim*(amostras[i])),
nrow=n.sim)
      c=array(c(g1,g2),dim=c(n.sim,amostras[i],2))
      resultados <- apply( c, 1, function(x) t.test(x[,1], x[,2],
paired=TRUE)$p.value)
      resultados1 <- apply( c, 1, function(x) t.test(x[,1],
x[,2],
paired=TRUE)$estimate[[1]][1])
      dist.p[i] <- mean(resultados)
      sup.IC.p[i] <-
sort(resultados)[ceiling(0.975*length(resultados))]
      inf.IC.p[i] <-
sort(resultados)[ceiling(0.025*length(resultados))]
      prop.p[i] <- sum(resultados<=alpha)/n.sim
      dif.obs[i] <- mean(resultados1)
      sup.IC.dif[i] <-
sort(resultados1)[ceiling(0.975*length(resultados1))]
      inf.IC.dif[i] <-
sort(resultados1)[ceiling(0.025*length(resultados1))]
      texto <- "TESTE-t PAREADO"
    }
  }
  diff.txt <- diff[1]
  plot (dist.p ~ amostras, bty="l", cex=1.1, pch=19, col="red",
```

```

xlab="Esforço
                                amostral", ylab="Valor médio de p",
ylim=c(0, sort(sup.IC.p,
                                decreasing=TRUE)[1]))
    lines(lowess(amostras, dist.p), lwd=2.3)
    lines(lowess (amostras, sup.IC.p), col="blue", lty=2,
lwd=1.7)
    lines(lowess (amostras, inf.IC.p), col="blue", lty=2,
lwd=1.7)
    abline (h=alpha, col="red")
    plot (prop.p ~ amostras, bty="l", cex=1.1, pch=19, col="red",
xlab="Esforço
                                amostral", ylab="Proporção de p
significativos")
    mtext(paste("Poder do teste de", texto), side=3, line=5,
cex=1.3, font=2)
    mtext(paste(n.sim, "simulações para cada categoria de
tamanho amostral"),
                                side=3, line=3, cex=1.15)
    mtext(paste("Média e desvio-padrão do lo grupo =", media,
"e", sd, " ;
                                Diferença entre as médias dos grupos
=", diff.txt, "; alpha =",alpha),
                                side=3, line=1, cex=1)
    plot (dif.obs ~ amostras, bty="l", cex=1.1, pch=19,
col="red", xlab="Esforço
                                amostral", ylab="Diferença encontrada
nas simulações",
                                ylim=c(sort(inf.IC.dif)[1],
sort(sup.IC.dif, decreasing=T)[1]))
    lines(lowess(amostras, dif.obs), lwd=2.3)
    lines(lowess (amostras, sup.IC.dif), col="blue", lty=2,
lwd=1.7)
    lines(lowess (amostras, inf.IC.dif), col="blue", lty=2,
lwd=1.7)
    }

anova.1 <- function(a1, a2, a3, amostras, n.sim, N, categorias, alpha)
{
  dist.p <- rep(NA,categorias)
  sup.IC.p <- rep(NA,categorias)
  inf.IC.p <- rep(NA,categorias)
  prop.p <- rep(NA,categorias)
  dif.obs.12 <- rep(NA,categorias)
  sup.IC.dif.12 <- rep(NA,categorias)
  inf.IC.dif.12 <- rep(NA,categorias)
  dif.obs.13 <- rep(NA,categorias)
  sup.IC.dif.13 <- rep(NA,categorias)
  inf.IC.dif.13 <- rep(NA,categorias)
  for(i in 1:categorias)
  {

```

```
g1 <- sample(a1, n.sim*amostras[i])
g2 <- sample(a2, n.sim*amostras[i])
g3 <- sample(a3, n.sim*amostras[i])
dados <- matrix(c(g1, g2, g3), byrow=TRUE, ncol=n.sim)
fat <- factor(rep(c("Grupo1", "Grupo2", "Grupo3"),
each=amostras[i]))
resultados <- sapply(1:ncol(dados), function(x)
anova(aov(dados[,x] ~
fat)))[5][1,1])
dist.p[i] <- mean(resultados)
sup.IC.p[i] <-
sort(resultados)[ceiling(0.975*length(resultados))]
inf.IC.p[i] <-
sort(resultados)[ceiling(0.025*length(resultados))]
prop.p[i] <- sum(resultados <= alpha)/n.sim
resultados1 <- sapply(1:ncol(dados), function(x)
summary(lm(dados[,x] ~
fat))$coefficients[2,1])
resultados2 <- sapply(1:ncol(dados), function(x)
summary(lm(dados[,x] ~
fat))$coefficients[3,1])
dif.obs.12[i] <- mean(resultados1)
sup.IC.dif.12[i] <-
sort(resultados1)[ceiling(0.975*length(resultados1))]
inf.IC.dif.12[i] <-
sort(resultados1)[ceiling(0.025*length(resultados1))]
dif.obs.13[i] <- mean(resultados2)
sup.IC.dif.13[i] <-
sort(resultados2)[ceiling(0.975*length(resultados2))]
inf.IC.dif.13[i] <-
sort(resultados2)[ceiling(0.025*length(resultados2))]
}
plot (dist.p ~ amostras, bty="l", cex=1.1, pch=19, col="red",
xlab="Esforço
amostral", ylab="Valor médio de p",
ylim=c(sort(inf.IC.p)[1],
sort(sup.IC.p, decreasing=T)[1]))
lines(lowess(amostras, dist.p), lwd=2.3)
lines(lowess (amostras, sup.IC.p), col="blue", lty=2,
lwd=1.7)
lines(lowess (amostras, inf.IC.p), col="blue", lty=2,
lwd=1.7)
abline (h=alpha, col="red")
plot (prop.p ~ amostras, bty="l", cex=1.1, pch=19, col="red",
xlab="Esforço
amostral", ylab="Proporção de p
significativos")
plot (dif.obs.12 ~ amostras, bty="l", cex=1.1, pch=19,
col="red", xlab="Esforço
amostral", ylab="Diferença entre lo e
```

```

2o grupo",
                                ylim=c(sort(inf.IC.dif.12)[1],
sort(sup.IC.dif.12, decreasing=T)[1]))
                                mtext(paste("Poder do teste de ANOVA"), side=3, line=5,
at=0, cex=1.5, font=2)
                                mtext(paste(n.sim, "simulações para cada categoria de
tamanho amostral"),
                                side=3, line=3, at=0, cex=1.15)
                                mtext(paste("Média e desvio-padrão do 1o grupo =", media,
"e", sd, " ;
                                Diferença entre as médias dos
                                grupos =", diff, "; alpha
=",alpha ), side=3, line=1, at=0, cex=1)
                                lines(lowess(amostras, dif.obs.12), lwd=2.3)
                                lines(lowess (amostras, sup.IC.dif.12), col="blue", lty=2,
lwd=1.7)
                                lines(lowess (amostras, inf.IC.dif.12), col="blue", lty=2,
lwd=1.7)
                                plot (dif.obs.13 ~ amostras, bty="l", cex=1.1, pch=19,
col="red", xlab="Esforço
                                amostral", ylab="Diferença 1o e 3o
grupo",
                                ylim=c(sort(inf.IC.dif.13)[1],
sort(sup.IC.dif.13, decreasing=T)[1]))
                                lines(lowess(amostras, dif.obs.13), lwd=2.3)
                                lines(lowess (amostras, sup.IC.dif.13), col="blue", lty=2,
lwd=1.7)
                                lines(lowess (amostras, inf.IC.dif.13), col="blue", lty=2,
lwd=1.7)
                                }

logis <- function(a1, a2, amostras, n.sim, N, categorias, alpha)
{
  dist.p <- rep(NA,categorias)
  sup.IC.p <- rep(NA,categorias)
  inf.IC.p <- rep(NA,categorias)
  prop.p <- rep(NA,categorias)
  dif.obs <- rep(NA,categorias)
  sup.IC.dif <- rep(NA,categorias)
  inf.IC.dif <- rep(NA,categorias)
  for(i in 1:categorias)
  {
    g1 <- sample(a1, n.sim*amostras[i])
    g2 <- sample(a2, n.sim*amostras[i])
    dados <- matrix(c(g1, g2), byrow=TRUE, ncol=n.sim)
    fat <- factor(rep(c(0, 1), each=amostras[i]))
    resultados <- sapply(1:ncol(dados), function(x)
summary(glm(fat ~ dados[,x],
family=binomial))$coefficients[2,4])
    dist.p[i] <- mean(resultados)
    sup.IC.p[i] <-

```

```
sort(resultados)[ceiling(0.975*length(resultados))]  
    inf.IC.p[i] <-  
sort(resultados)[ceiling(0.025*length(resultados))]  
    prop.p[i] <- sum(resultados <= alpha)/n.sim  
    resultados1 <- sapply(1:ncol(dados), function(x)  
        mean(dados[fat==0,x]) -  
mean(dados[fat==1,x]))  
    dif.obs[i] <- mean(resultados1)  
    sup.IC.dif[i] <-  
sort(resultados1)[ceiling(0.975*length(resultados1))]  
    inf.IC.dif[i] <-  
sort(resultados1)[ceiling(0.025*length(resultados1))]  
    }  
    plot (dist.p ~ amostras, bty="l", cex=1.1, pch=19, col="red",  
xlab="Esforço amostral",  
        ylab="Valor médio de p", ylim=c(0,  
sort(sup.IC.p, decreasing=TRUE)[1]))  
    lines(lowess(amostras, dist.p), lwd=2.3)  
    lines(lowess (amostras, sup.IC.p), col="blue", lty=2, lwd=1.7)  
    lines(lowess (amostras, inf.IC.p), col="blue", lty=2, lwd=1.7)  
    abline (h=alpha, col="red")  
    plot (prop.p ~ amostras, bty="l", cex=1.1, pch=19, col="red",  
xlab="Esforço amostral",  
        ylab="Proporção de p significativos")  
    mtext("Poder do teste de REGRESSÃO LOGÍSTICA", side=3, line=5,  
cex=1.3, font=2)  
    mtext(paste(n.sim, "simulações para cada categoria de  
tamanho amostral"),  
        side=3, line=3, at=0, cex=1.15)  
    mtext(paste("Média e desvio-padrão do 1o grupo =",  
media, "e", sd, " ;  
Diferença entre as médias dos grupos =",  
diff, "; alpha =",alpha ),  
        side=3, line=1, cex=1)  
    plot (dif.obs ~ amostras, bty="l", cex=1.1, pch=19, col="red",  
xlab="Esforço amostral",  
        ylab="Diferença encontrada nas simulações",  
ylim=c(sort(inf.IC.dif)[1], sort(sup.IC.dif,  
decreasing=T)[1]))  
    lines(lowess(amostras, dif.obs), lwd=2.3)  
    lines(lowess (amostras, sup.IC.dif), col="blue", lty=2, lwd=1.7)  
    lines(lowess (amostras, inf.IC.dif), col="blue", lty=2, lwd=1.7)  
    }  
  
if(teste=="teste.t")  
{  
    grupo.1 <- rnorm(n.sim*N, media, sd)  
    grupo.2 <- rnorm(n.sim*N, media+diff, sd)  
    result <- power.t.test(n=NULL, delta=diff, sd=sd,
```

```

sig.level=alpha, power=0.95,
                                type="two.sample",
alternative="two.sided")
  X11()
  par(mfrow=c(2,3), mar=c(5,5,10,2), cex.axis=1.4, cex.lab=1.3)
  teste.t(grupo.1, grupo.1, amostras, n.sim, N, categorias, alpha)
  teste.t(grupo.1, grupo.2, amostras, n.sim, N, categorias, alpha)
  par(mfrow=c(1,1))
  mtext("ERRO TIPO 1", side=3, line=8, cex=1.5, font=2, col="red")
  mtext("ERRO TIPO 2", side=3, line=-12.5, cex=1.5, font=2,
col="red")
  t1=proc.time()[[3]]
  tttotal=(t1-t0)/60
  cat("\n\t tempo de processamento: ", tttotal, "minutos\n")
  return(result)
}

else
{
  if(teste=="t.pareado")
  {
    if(length(diff)==2)
    {
      grupo.1 <- rnorm(n.sim*N, media, sd)
      diff1 <- diff[1]
      diff2 <- diff[2]
      difer <- rnorm(n.sim*N, diff1, diff2)
      grupo.2 <- grupo.1+difer
      X11()
      par(mfrow=c(2,3), mar=c(5,5,10,2), cex.axis=1.4, cex.lab=1.3)
      teste.t(grupo.1, grupo.1, amostras, n.sim, N, categorias, alpha,
par="T")
      teste.t(grupo.1, grupo.2, amostras, n.sim, N, categorias, alpha,
par="T")
      par(mfrow=c(1,1))
      mtext("ERRO TIPO 1", side=3, line=8, cex=1.5, font=2, col="red")
      mtext("ERRO TIPO 2", side=3, line=-12.5, cex=1.5, font=2,
col="red")
      t1=proc.time()[[3]]
      tttotal=(t1-t0)/60
      cat("\n\t tempo de processamento: ", tttotal, "minutos\n")
    }
    else
    {
      cat("\n\t", "Para calcular o poder do TESTE-t PAREADO, o argumento
diff deve ser um
      vetor contendo dois elementos: a média e o desvio-padrão da
diferença de cada par!\n\n")
    }
  }
  else

```

```

{
  if(teste=="anova")
  {
    grupo.1 <- rnorm(n.sim*N, media, sd)
    grupo.2 <- rnorm(n.sim*N, media+(diff/2), sd)
    grupo.3 <- rnorm(n.sim*N, media+diff, sd)
    within.var = sd^2
    medias <- c(media, media+(diff/2), media-(diff/2))
    result <- power.anova.test(groups=3, n=NULL,
between.var=var(medias),
                                within.var=within.var, sig.level=alpha,
power=0.95)
    X11()
    par(mfrow=c(2,4), mar=c(5,5,10,2), cex.axis=1.4, cex.lab=1.3)
    anova.1(grupo.1, grupo.1, grupo.1, amostras, n.sim, N, categorias,
alpha)
    anova.1(grupo.3, grupo.1, grupo.2, amostras, n.sim, N, categorias,
alpha)
    par(mfrow=c(1,1))
    mtext("ERRO TIPO 1", side=3, line=8, cex=1.5, font=2, col="red")
    mtext("ERRO TIPO 2", side=3, line=-12.5, cex=1.5, font=2,
col="red")
    t1=proc.time()[[3]]
    tttotal=(t1-t0)/60
    cat("\n\t tempo de processamento: ", tttotal, "minutos\n")
    return(result)
  }
  else
  {
    if(teste=="logistica")
    {
      grupo.1 <- rnorm(n.sim*N, media, sd)
      grupo.2 <- rnorm(n.sim*N, media+diff, sd)
      X11()
      par(mfrow=c(2,3), mar=c(5,5,10,2), cex.axis=1.4, cex.lab=1.3)
      logis(grupo.1, grupo.1, amostras, n.sim, N, categorias, alpha)
      logis(grupo.1, grupo.2, amostras, n.sim, N, categorias, alpha)
      par(mfrow=c(1,1))
      mtext("ERRO TIPO 1", side=3, line=8, cex=1.5, font=2, col="red")
      mtext("ERRO TIPO 2", side=3, line=-12.5, cex=1.5, font=2, col="red")
      t1=proc.time()[[3]]
      tttotal=(t1-t0)/60
      cat("\n\t tempo de processamento: ", tttotal, "minutos\n")
      result <- cat("\n\t NÃO EXISTE resolução analítica para avaliar
o poder
                                de uma regressão logística!\n")
      return(result)
    }
    else
    {

```

```
cat("\n\t", "ESCOLHA dentre os seguintes testes: teste-t, teste-t  
pareado,  
        regressão logística ou anova!\n\n")  
}  
}  
}  
    }  
    }  
else  
    {  
        cat("\n\t", "O argumento alpha DEVE ser um número ENTRE 0.001 e  
0.999!\n\n")  
    }  
    }  
else  
    {  
        cat("\n\t", "O argumento alpha DEVE ser um número que representa o  
NÍVEL DE SIGNIFICÂNCIA \n\t\t\t\t\t\t\t assumido a priori!\n\n")  
    }  
    }  
else  
    {  
        cat("\n\t", "O argumento categorias DEVE ser um número que representa  
o NÚMERO DE CATEGORIAS de \n\t\t\t tamanhos amostrais a serem  
simulados,  
        entre 5 amostras e", N, "amostras!\n\n")  
    }  
    }  
else  
    {  
        cat("\n\t", "O argumento N DEVE ser um número que representa o NÚMERO  
MÁXIMO DE AMOSTRAGEM!\n\n")  
    }  
    }  
else  
    {  
        cat("\n\t", "O argumento n.sim DEVE ser um número que representa o  
NÚMERO DE SIMULAÇÕES \n\t\t\t a serem realizadas para cada categoria de  
tamanho amostral!\n\n")  
    }  
    }  
else  
    {  
        cat("\n\t", "O argumento sd DEVE ser um número que representa a  
DESVIO-PADRÃO das populações amostrais!\n\n")  
    }  
}
```

```
else
{
    cat("\n\t", "0 argumento media DEVE ser um número que representa a
MÉDIA de uma
                                das populações amostrais!\n\n")
}
}
else
{
    cat("\n\t", "0 argumento diff DEVE ser um número que representa a
DIFERENÇA!\n\n")
}
}
```

ARQUIVO DA FUNÇÃO

I've got the power FUNCTION

From:

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

Permanent link:

http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2010:alunos:trabalho_final:billy.requena:start



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