

```

## Kappa de Cohen
library(vcd)

# Dados
# Tabela 1011, p. 376 em Agresti (1990), Categorical Data Analysis

tab1011 <- matrix(c(50, 45, 8, 18, 8,
                   28, 174, 84, 154, 55,
                   11, 78, 110, 223, 96,
                   14, 150, 185, 714, 447,
                   3, 42, 72, 320, 411), ncol = 5, byrow = TRUE)
rownames(tab1011) <- paste("Pai", 1:5, sep = ".")
colnames(tab1011) <- paste("Filho", 1:5, sep = ".")
tab1011
margin.table(tab1011, margin = 1)
margin.table(tab1011, margin = 2)
n <- sum(tab1011)
cat("\n n =", n, "\n")

k1011 <- Kappa(tab1011)
k1011
confint(k1011)

## Teste de McNemar
# Dados
# Tabela 10.1, p. 350 em Agresti (1990), Categorical Data Analysis
tab101 <- matrix(c(794, 150, 86, 570), byrow = TRUE, ncol = 2)
rownames(tab101) <- c("1. Aprovação", "1. Desaprovação")
colnames(tab101) <- c("2. Aprovação", "2. Desaprovação")
tab101
n <- sum(tab101)
cat("\n n =", n, "\n")

mcnemar.test(tab101, correct = FALSE)

# Teste exato (dist. binomial)
# H1:  $\pi_{12} > \pi_{21}$  (aprovação 1o. > 2o.)
ns <- tab101[1, 2] + tab101[2, 1]
valorp <- pbinom(tab101[1, 2] - 1, ns, 0.5, lower.tail = FALSE)
cat("\n Teste exato para H1:  $\pi_{12} > \pi_{21}$  \n n12 =", tab101[1, 2],
    "\n n* =", ns, "(p =", valorp, ") \n")

# Proporções amostrais
(prop101 <- prop.table(tab101))

# Diferença de aprovação (2o. - 1o.)
(estd <- prop101[2, 1] - prop101[1, 2])

# Erro padrão da estimativa
(marg101lin <- margin.table(prop101, margin = 1))
(marg101col <- margin.table(prop101, margin = 2))
epest <- sqrt( (marg101lin[1] * (1 - marg101lin[1]) +
               marg101col[1] * (1 - marg101col[1]) - 2 * det(prop101)) / n)
cat("\n e.p.(d) =", epest, "\n")

```

```

# IC para a diferença de aprovação (2o. - 1o.)
conf <- 0.95
icdif <- estd + c(-1, 1) * qnorm((1 + conf) / 2) * epest
cat("\n IC de", conf * 100, "% para a diferença (2o. - 1o.): \n",
icdif, "\n")

## Modelos de simetria
# Dados
# Tabela 10.2, p. 357 em Agresti (1990), Categorical Data Analysis

resid80 <- (c("NE", "MW", "S", "W"))
resid85 <- (c("NE", "MW", "S", "W"))

resid80 <- factor(resid80, levels = resid80)
resid85 <- resid80

tabl02 <- expand.grid(res80 = resid80, res85 = resid85)
tabl02$n <- c(11607, 100, 366, 124, 87, 13677, 515, 302, 172, 225,
             17819, 270, 63, 176, 286, 10192)

## Modelo de simetria
# Variável auxiliar com um valor para cada diferente probabilidade
tabl02$simaux <- factor(paste(
  pmin(as.numeric(tabl02$res80), as.numeric(tabl02$res85)),
  pmax(as.numeric(tabl02$res80), as.numeric(tabl02$res85)), sep =
  ", "))

msim <- glm(n ~ simaux, family = poisson, data = tabl02)
msim
cat("\n Modelo de simetria\n G2 = ", msim$deviance,
    "(p =", pchisq(msim$deviance, msim$df.residual, lower.tail =
FALSE),
    ", g.l. = ", msim$df.residual, ")")

## Modelo de quase simetria
mqsim <- glm(n ~ res80 + simaux, family = poisson, data = tabl02)
mqsim
cat("\n Modelo de quase simetria\n G2 = ", mqsim$deviance,
    "(p =", pchisq(mqsim$deviance, mqsim$df.residual, lower.tail =
FALSE),
    ", g.l. = ", mqsim$df.residual, ")")

# Frequências esperadas estimadas (calcule a estatística X2)
muqsim <- matrix(mqsim$fitted.values, ncol = 4, byrow = TRUE,
dimnames =
  list(levels(resid80), levels(resid85)))
muqsim

#  $P(\text{res85} = \text{MW}, \text{res80} = \text{W}) / P(\text{res85} = \text{W}, \text{res80} = \text{MW})$ 
muqsim["W", "MW"] / muqsim["MW", "W"]

## Modelo de quase independência
# Variáveis auxiliares
tabl02$D1 <- as.numeric(tabl02$simaux == "1,1")
tabl02$D2 <- as.numeric(tabl02$simaux == "2,2")
tabl02$D3 <- as.numeric(tabl02$simaux == "3,3")

```

```

tab102$D4 <- as.numeric(tab102$simaux == "4,4")

mqind <- glm(n ~ res80 + res85 + D1 + D2 + D3 + D4, family = poisson,
            data = tab102)

mqind
cat("\n Modelo de quase independência\n G2 = ", mqind$deviance,
    "(p =", pchisq(mqind$deviance, mqind$df.residual, lower.tail =
FALSE),
    ", g.l. = ", mqind$df.residual,)")

# Frequências esperadas estimadas (calcule a estatística X2)
muqind <- matrix(mqind$fitted.values, ncol = 4, byrow = TRUE,
                dimnames =
                list(levels(resid80), levels(resid85)))

# Exemplos de razões de chances (explicação?)
muqind["MW", "S"] * muqind["NE", "W"] / (muqind["NE", "S"] *
muqind["MW", "W"])
muqind["NE", "MW"] * muqind["W", "S"] / (muqind["W", "MW"] *
muqind["NE", "S"])

# Um modelo com delta1 = delta2 = delta3 = delta4 faz um bom ajuste?

## Teste de homogeneidade marginal
g2hm <- msim$deviance - mqsim$deviance
gl <- msim$df.residual - mqsim$df.residual
cat("\n Teste de homogeneidade marginal \n G2 = ", g2hm,
    "(p =", pchisq(g2hm, gl, lower.tail = FALSE), ", g.l. = ", gl,)")

```