

```

# Stratified two by two example

df <- read.csv("~/teach/data/berkelyAdmissions", skip=1, header=TRUE)
berkAdmits <- aperm(array(data.matrix(df[2:5]), dim=c(6,2,2),
  dimnames=list(program=LETTERS[1:6], admit=c("Y", "N"), sex=c("M", "F"))),
  3:1)
print(percents(berkAdmits, denom=c(1,3)))

or2 <- function(m) { (m[1,1]*m[2,2])/(m[2,1]*m[1,2]) }
apply(berkAdmits, 3, or2)

collapseTab <- apply(berkAdmits, 1:2, sum)
print(percents(collapseTab, 1))

mantelhaen.test(berkAdmits)

# Analysis by logistic regression

df1 <- df[1:3]
df2 <- df[c(1,4:5)]
names(df1) <- names(df2) <- c("program", "yes", "no")
df <- rbind(df1, df2)
df$sex <- factor(rep(1:2, c(6,6)), labels=c("M", "F"))
attach(df)
logitFit <- glm(cbind(yes, no) ~ program + sex, family=binomial)
summary(logitFit)
print(exp(cbind(logitFit$coef, confint.default(logitFit))))
logitFit <- glm(cbind(yes, no) ~ program * sex, family=binomial)
print(anova(logitFit, test="Chisq"))

# 2x2x2 log-linear model - diagnostic testing

diagTest <- array(scan(), dim=c(2,2,2),
  dimnames=list(d.dimer=c("+", "-"), ultrasound=c("+", "-"), venography=c("+", "-")))
61 24 15 6 4 15 35 150

print(diagTest)
count.df <- data.frame(counts=as.vector(diagTest),
  d.dimer=factor(c(1,2,1,2,1,2,1,2), labels=c("+", "-")),
  ultrasound=factor(c(1,1,2,2,1,1,2,2), labels=c("+", "-")),
  venography=factor(c(1,1,1,1,2,2,2,2), labels=c("+", "-")))
print(count.df)

attach(count.df)
model1 <- glm( counts ~ d.dimer + ultrasound + venography, family=poisson)
model2 <- glm( counts ~ d.dimer + ultrasound * venography, family=poisson)
model3 <- glm( counts ~ d.dimer + ultrasound + venography +
  venography:(d.dimer + ultrasound), family=poisson)
model4 <- glm( counts ~ d.dimer * ultrasound + venography +
  venography:(d.dimer + ultrasound), family=poisson)
model5 <- glm( counts ~ d.dimer * ultrasound * venography , family=poisson)

print(anova(model1, model2, model3, model4, model5, test="Chisq"))

```