

Contingency analysis: R code for Chapter 9 examples

Download the R code on this page as a single file [here](#) (make sure to install the “epitools” package before running).

New methods on this page

Click on a function argument for a short description of its meaning. The variable names are plucked from the examples further below.

Odds ratio using the `epitools` package:

```
oddsratio(cancerTable, method = "wald")
```

χ^2 contingency test:

```
chisq.test(worm$fate, worm$infection, correct = FALSE)
```

Fisher’s exact test:

```
fisher.test(vampire$bitten, vampire$estrous)
```

Other new methods:

Relative risk.

G-test.

Mosaic plot for a case-control study.

Figure 9.1-1. Sex and survival on the Titanic

Mosaic plot showing association between sex and survival on the shipwrecked Titanic.

Read and inspect the data.

```
titanic <- read.csv(url("http://www.zoology.ubc.ca/~schluter/WhitlockSchluter/wp-content/data/chapter09/chap09f1.1Tit
head(titanic)
```

Contingency table of the association between sex and survival. The `addmargins` function includes the row and columns sums with the contingency table.

```
titanicTable <- table(titanic$survival, titanic$sex)
addmargins(titanicTable)
```

Mosaic plot of the association.

```
mosaicplot( t(titanicTable), col = c("firebrick", "goldenrod1"), cex.axis = 1,
  sub = "Sex", ylab = "Relative frequency", main = "")
```

Example 9.2. Aspirin and cancer

Odds ratio and **relative risk** to estimate association between aspirin treatment and cancer incidence in a 2x2 contingency table.

Read and inspect the data.

```
cancer <- read.csv(url("http://whitlockschluter.zoology.ubc.ca/wp-content/data/chapter09/chap09e2AspirinCancer.csv"))
head(cancer)
```

2 x 2 contingency table showing association between aspirin treatment and cancer incidence (Table 9.2-1).

```
cancerTable <- table(cancer$cancer, cancer$aspirinTreatment)
cancerTable
```

Mosaic plot showing the association (Figure 9.2-1).

```
mosaicplot( t(cancerTable), col = c("firebrick", "goldenrod1"), cex.axis = 1,
  sub = "Aspirin treatment", ylab = "Relative frequency", main = "")
```

Calculate odds ratio using the `epitools` package. Install the package if you haven't already done so (this needs to be done just once per computer, after you've installed R). To use, load the `epitools` package (this needs to be done just once per R session).

```
install.packages("epitools", dependencies = TRUE)
library(epitools)
```

The formulas for the odds ratio and relative risk in the book assume that the contingency table has a layout exactly like `cancerTable`, namely

```
#           treatment control
# sick      a      b
# healthy   c      d
```

This layout works for the `oddsratio` function.

```
oddsratio(cancerTable, method = "wald")
```

The resulting output of `oddsratio` includes a lot of incidental computations. To get clean output only for the odds ratio, including the 95% confidence interval, use the following instead.

```
oddsratio(cancerTable, method = "wald")$measure[-1,]
```

For step-by-step calculations of the odds ratio and confidence interval, see [here](#).

Calculate relative risk using the `epitools` package. The layout expected by the `riskratio` function is the complete opposite of the layout used in the book. To use the command with a contingency table in book style (such as `cancerTable`), we need to flip (transpose) the table and reverse the order of the rows. We can do this all at once with the following arguments to the `riskratio` function.

```
library(epitools)
riskratio(t(cancerTable), rev = "both", method = "wald")
```

Notice that the result differs slightly from the relative risk value given in the book for these same data (1.007) because rounding error here is reduced.

For clean output only for the relative risk, including the 95% confidence interval, use the following command:

```
riskratio(t(cancerTable), method = "wald", rev = "both")$measure[-1,]
```

For step-by-step calculations of the relative risk, see [here](#).

Example 9.3. Toxoplasma and driving accident

Odds ratio to estimate association between *Toxoplasma* infection and driving accident in a 2 x 2 **case-control study**.

Read and inspect the data.

```
toxoplasma <- read.csv(url("http://whitlockschluter.zoology.ubc.ca/wp-content/data/chapter09/chap09e3ToxoplasmaAndAcc
head(toxoplasma)
```

2 x 2 contingency table (Table 9.3-1).

```
toxTable <- table(toxoplasma$accident, toxoplasma$condition)
toxTable
```

Mosaic plot for a case-control study (Figure 9.3-1).

```
mosaicplot(toxTable, col=c("firebrick", "goldenrod1"), cex.axis = 1.2,
  sub = "Condition", dir = c("h", "v"), ylab = "Relative frequency")
```

Odds ratio using epitools.

```
library(epitools)
oddsratio(toxTable, method = "wald")
oddsratio(toxTable, method = "wald")$measure[-1,] # clean output
```

Example 9.4. Worm gets bird

χ^2 contingency test to test association between trematode infection status of killifish and their fate (eaten or not eaten) in the presence of predatory birds.

Read and inspect the data.

```
worm <- read.csv(url("http://www.zoology.ubc.ca/~schluter/WhitlockSchluter/wp-content/data/chapter09/chap09e4WormGets
head(worm)
```

Set the preferred order of infection categories in tables and graphs.

```
worm$infection <- factor(worm$infection,
  levels = c("uninfected", "lightly", "highly"))
```

Contingency table (Table 9.4-1). The `addmargins` function adds the row and column sums to the display of the table.

```
wormTable <- table(worm$fate, worm$infection)
addmargins(wormTable)
```

Mosaic plot (Figure 9.4-1).

```
mosaicplot( t(wormTable), col = c("firebrick", "goldenrod1"), cex.axis = 1,
  sub = "Infection status", ylab = "Relative frequency")
```

χ^2 contingency test. We include the argument `correct = FALSE` to avoid Yates' correction. This has no effect except in 2 x 2 tables, but we keep it here for demonstration purposes.

```
saveChiTest <- chisq.test(worm$fate, worm$infection, correct = FALSE)
saveChiTest
```

The **expected frequencies under null hypothesis** are included with the results, but aren't normally shown. Type `saveChiTest$expected` to extract them. Include the `addmargins` function if you want to see the row and column sums too.

```
addmargins(saveChiTest$expected)
```

G-test applied to the worm-gets-bird data (Section 9.6). R has no simple, built-in function to carry out the G-test with goodness-of-fit data. Code for a command `g.test` by Brent Large is available [here](#). Below, we source this code, which then allows you to use his function `g.test`.

```
source("http://www.stat.wisc.edu/~st571-1/gtest.R")
g.test(wormTable)
```

Example 9.5. Vampire bat feeding

Fisher's exact test of association between estrous status of cows and whether cows were bitten by vampire bats.

Read and inspect the data.

```
vampire <- read.csv(url("http://www.zoology.ubc.ca/~schluter/WhitlockSchluter/wp-content/data/chapter09/chap09e5Vampi
```

Contingency table (Table 9.5-1).

```
vampireTable <- table(vampire$bitten, vampire$estrous)
vampireTable
```

Expected frequencies under null hypothesis of independence. R complains because of the violation of assumptions. Just in case you hadn't noticed.

```
saveTest <- chisq.test(vampire$bitten, vampire$estrous, correct = FALSE)
saveTest$expected
```

Fisher's exact test. The output includes an estimate of the odds ratio, but it uses a different method than is used in our book.

```
fisher.test(vampire$bitten, vampire$estrous)
```