Fisher's Exact Test with R[1]

```
> # Pauling's skier data again.
> ski = rbind(c(31,109),
+             c(17,122))
> rownames(ski) = c('Placebo','Ascorbic Acid')
> colnames(ski) = c('Cold','No Cold')
> addmargins(ski) # Take a look
              Cold No Cold Sum
Placebo         31     109 140
Ascorbic Acid   17     122 139
Sum             48     231 279
> prop.table(ski,margin=1) # 1=row, 2=col
                    Cold   No Cold
Placebo       0.2214286 0.7785714
Ascorbic Acid 0.1223022 0.8776978
>
> chisq.test(ski, correct=FALSE)

     Pearson's Chi-squared test

data:  ski
X-squared = 4.8114, df = 1, p-value = 0.02827

> fisher.test(ski)

     Fisher's Exact Test for Count Data

data:  ski
p-value = 0.03849
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 1.026674 4.154449
sample estimates:
odds ratio
  2.035861

> # Help file says: "Note that the conditional Maximum Likelihood
Estimate (MLE) rather than the unconditional MLE (the sample odds
ratio) is used." This means conditional on the marginal
probabilities.
```

1 See last page for copyright information.

```
>
> # Subset of Titanic data
> # Women in 1st class vs Women in crew
> ladies = Titanic[c(1,4),2,2,]; ladies
      Survived
Class   No Yes
  1st    4 140
  Crew   3  20
> X2 = chisq.test(ladies,correct=F); X2
Warning message:
In chisq.test(ladies, correct = F) :
  Chi-squared approximation may be incorrect

     Pearson's Chi-squared test

data:  ladies
X-squared = 5.2043, df = 1, p-value = 0.02253

> # Check the expected frequencies
> X2$expected
      Survived
Class         No       Yes
  1st  6.0359281 137.96407
  Crew 0.9640719  22.03593
>
> fisher.test(ladies)

     Fisher's Exact Test for Count Data

data:  ladies
p-value = 0.05547
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.03027561 1.41705937
sample estimates:
odds ratio
 0.1935113

>
> # Conclusion: Though a higher percentage of women in first class
survived than female crew, it could have been due to chance.
```

```
> # Piano data from text
>
> piano = rbind(c(4,2),
+               c(13,1),
+               c(11,2),
+               c(2,2),
+               c(9,2),
+               c(6,0)      )
> colnames(piano) = c("Steinway","Other")
> rownames(piano) = c("Boston", "Chicago", "Cleveland",
+ "Minnesota", "New York", "Philadelphia")
> piano
             Steinway Other
Boston              4     2
Chicago            13     1
Cleveland          11     2
Minnesota           2     2
New York            9     2
Philadelphia        6     0

> pianoX2 = chisq.test(piano, correct=FALSE); pianoX2
Warning message:
In chisq.test(piano, correct = FALSE) :
  Chi-squared approximation may be incorrect

     Pearson's Chi-squared test

data:  piano
X-squared = 6.5479, df = 5, p-value = 0.2565

> pianoX2$expected
             Steinway     Other
Boston       5.000000 1.0000000
Chicago     11.666667 2.3333333
Cleveland   10.833333 2.1666667
Minnesota    3.333333 0.6666667
New York     9.166667 1.8333333
Philadelphia 5.000000 1.0000000
>
> fisher.test(piano) # Good for larger than 2x2!

     Fisher's Exact Test for Count Data

data:  piano
p-value = 0.2525
alternative hypothesis: two.sided

> # In case of numerical trouble use the simulate.p.value = TRUE option.
```
