Fisher's Exact Test with R

```
> # A little table for testing
>
> testor = rbind(c(4,1),
+                c(20,1)  ); testor
     [,1] [,2]
[1,]    4    1
[2,]   20    1
>
> chi2 = chisq.test(testor,correct=F); chi2
Warning message:
In chisq.test(testor, correct = F) :
  Chi-squared approximation may be incorrect

	Pearson's Chi-squared test

data:  testor
X-squared = 1.3206, df = 1, p-value = 0.2505

> chi2$expected
          [,1]      [,2]
[1,]  4.615385 0.3846154
[2,] 19.384615 1.6153846
>
> # help(fisher.test)
>
> # Says "Two-sided tests are based on the probabilities of the tables, and take as
‘more extreme’ all tables with probabilities less than or equal to that of the
observed table, the p-value being the sum of such probabilities."
>
> # Also "estimate      an estimate of the odds ratio. Note that the conditional
Maximum Likelihood Estimate (MLE) rather than the unconditional MLE (the sample
odds ratio) is used."
>
>
> fisher.test(testor)
```

```
        Fisher's Exact Test for Count Data

data:  testor
p-value = 0.3538
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  0.002439905 19.594803004
sample estimates:
odds ratio
 0.2182166

> # Try some p-values by "hand"
> testor
     [,1] [,2]
[1,]    4    1
[2,]   20    1
> n = sum(testor); n
[1] 26
> a = sum(testor[1,]); a
[1] 5
> b = sum(testor[,1]); b
[1] 24
> Lo = max(0,a+b-n)
> Hi = min(a,b)

> sampletheta = testor[1,1]*testor[2,2] / (testor[1,2]*testor[2,1]); sampletheta
[1] 0.2
> x = Lo:Hi
> theta = x*(n-a-b+x)/((a-x)*(b-x))
> prob = choose(a,x)*choose(n-a,b-x)/choose(n,b)
> sum(prob)
[1] 1
> cbind(x,theta,prob)
     x theta       prob
[1,] 3   0.0 0.03076923
[2,] 4   0.2 0.32307692
[3,] 5   Inf 0.64615385
> fisher.test(testor,alternative='g')

        Fisher's Exact Test for Count Data

data:  testor
p-value = 0.9692
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 0.00496321        Inf
sample estimates:
odds ratio
 0.2182166
```

```
>
> # Try with calcpass
>
> math =read.table("http://www.utstat.toronto.edu/~brunner/312f12/code_n_data/mathcat.data")
> attach(math) # Variable names are now global
> calcpass = table(hscalc,passed); calcpass
      passed
hscalc  No Yes
   No   17   4
   Yes 141 232
>
> n = sum(calcpass); n
[1] 394
> a = sum(calcpass[1,]); a
[1] 21
> b = sum(calcpass[,1]); b
[1] 158
> Lo = max(0,a+b-n)
> Hi = min(a,b)
> sampletheta = calcpass[1,1]*calcpass[2,2] / (calcpass[1,2]*calcpass[2,1])
> sampletheta
[1] 6.992908
> x = Lo:Hi
> theta = x*(n-a-b+x)/((a-x)*(b-x))
> prob = choose(a,x)*choose(n-a,b-x)/choose(n,b)
> sum(prob)
[1] 1
```

```
> cbind(x,theta,round(prob,7))
       x       theta
 [1,]  0  0.00000000 0.0000146
 [2,]  1  0.06878981 0.0002237
 [3,]  2  0.14642375 0.0016182
 [4,]  3  0.23440860 0.0073340
 [5,]  4  0.33460657 0.0233584
 [6,]  5  0.44934641 0.0555930
 [7,]  6  0.58157895 0.1026332
 [8,]  7  0.73509934 0.1505815
 [9,]  8  0.91487179 0.1784357
[10,]  9  1.12751678 0.1725941
[11,] 10  1.38206388 0.1371548
[12,] 11  1.69115646 0.0898182
[13,] 12  2.07305936 0.0484702
[14,] 13  2.55517241 0.0214878
[15,] 14  3.18055556 0.0077747
[16,] 15  4.02097902 0.0022716
[17,] 16  5.20563380 0.0005273
[18,] 17  6.99290780 0.0000949
[19,] 18  9.98571429 0.0000128
[20,] 19 15.99280576 0.0000012
[21,] 20 34.05797101 0.0000001
[22,] 21         Inf 0.0000000
>
> fisher.test(calcpass,alternative='g')

      Fisher's Exact Test for Count Data

data:  calcpass
p-value = 0.000109
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 2.563289      Inf
sample estimates:
odds ratio
  6.959835

> sum(prob[18:22])
[1] 0.0001089755

> fisher.test(calcpass,alternative='l')

      Fisher's Exact Test for Count Data

data:  calcpass
p-value = 1
alternative hypothesis: true odds ratio is less than 1
95 percent confidence interval:
  0.00000 22.93212
sample estimates:
odds ratio
  6.959835

> sum(prob[1:18])
[1] 0.999986
```

```
> fisher.test(calcpass)

      Fisher's Exact Test for Count Data

data:  calcpass
p-value = 0.0001235
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  2.210253 29.021086
sample estimates:
odds ratio
  6.959835

> sum(prob[prob<=prob[18]])
[1] 0.0001235361


>
> # Pearson chi-squared test for comparison
> chisq.test(calcpass,correct=F)

      Pearson's Chi-squared test

data:  calcpass
X-squared = 15.4111, df = 1, p-value = 8.648e-05
```