*Stat 622/422 (Dr. Baron) Advanced Biostatistics*

**Epidemiology: risks and contingency tables**

**# Enter the contingency table of counts data. The data from Rosner, page 373, Example 10.2.**

> counts = c(683,2573,1498,8747)

> Table = matrix(counts,2,2)

> Table

[,1] [,2]

[1,] 683 1498

[2,] 2573 8747

**# Calculate the absolute, relative, and attributable risks and the odds ratio**

> R = 683/(683+1498)

> R

[1] 0.3131591

> RR = (683/(683+1498)) / (2573/(2573+8747))

> RR

[1] 1.377754

> AR = (683/(683+1498)) - (2573/(2573+8747))

> AR

[1] 0.08586228

> OR = (683/1498) / (2573/8747)

> OR

[1] 1.549988

**# We estimate the probability of developing a disease at 0.210, for people exposed to the risk factor. # This probability is 37.8% higher than for people who are not exposed. The attributable risk of**

**# 0.086 represents the additional probability of developing the disease that can be attributed solely**

**# to the risk factor. The odds ratio of developing the disease is 55.0% higher for people exposed to**

**# the risk factor.**

**# Chi-square test of association**

> TestResults = chisq.test(Table, correct=0)

> TestResults

Pearson's Chi-squared test

X-squared = 73.668, df = 1, p-value < 2.2e-16

**# The association between the disease and the factor is statistically significant. Next, we can see the**

**# observed and expected counts that were used in the calculation of the chi-square statistic.**

> TestResults$observed

[,1] [,2]

[1,] 683 1498

[2,] 2573 8747

> TestResults$expected

[,1] [,2]

[1,] 525.9859 1655.014

[2,] 2730.0141 8589.986

**# The chi-square statistic measured how far the observed counts are from the expected counts. The**

**# difference appeared statistically significant.**