4/6/2010 36-402/608 ADA-II H. Seltman
Breakout #20 Comments

These data come from The Sleuth, chapters 18 and 19.

```
# Randomized trial of vitamin C for preventing colds
vit = matrix(c(335,302,76,105), nrow=2, dimnames=
        list(c("Placebo","Vitamin C"), c("Cold", "No Cold")))
source("http://www.stat.cmu.edu/~hseltman/files/cta.R")
cta(vit)
# $table
#           Cold No Cold   n      phat         SE      CIlo      CIhi
# Placebo    335      76 411 0.8150852 0.01914990 0.7775514 0.8526190
# Vitamin C  302     105 407 0.7420147 0.02168735 0.6995075 0.7845219
# Total      637     181 818 0.7787286 0.02902781 0.7218341 0.8356231
#
# $binDiff
#        diff      SEdiff           Z     p.value        CIlo        CIhi
# -0.07307042  0.02902781 -2.51725577  0.01155033 -0.01636372 -0.12977711
#
# $OR
#       OR     ORlo     ORih     p.value
# 1.532546 1.097770 2.139517  0.01214262
#
# $miscTests
#    p.chisq   p.Fisher
# 0.01497328 0.01444212
```

Question 1: Explain all of the numbers, including null hypotheses for the tests. Also, when is the Total CI useful? The first two lines of the "table" section gives estimates of p and the 95% CI for those estimates separately (not assuming equality). The Total line gives the pooled estimates which should be used if and only if we retain the null hypothesis of equal probabilities.

The "binDiff" section tests the difference of (independent) binomial proportions and gives the CI for the difference. We are 95% confident that the probability of a cold is 1.6 to 13.0 % lower for vitamin C than for Placebo. The best estimate of 7.3% fewer colds seems like a fairly small effect.

The OR of 1.53 represent the effect in a different way: the ratio of cold years to non-cold years you might experience with controls is 4.4:1 and with vitamin C is 2.9:1, and the ratio of these odds is 1.5. This fact that the estimated odds of getting a cold are 1.5 times

as large for control than vitamin C is often loosely and inappropriately expressed as "you are 1.5 times as likely to get a cold when not taking vitamin C".

The Z-test for OR=1 (p=0.012), the chi-square test for independence (p=0.015) and the Fisher test (p=0.014) are similar. They can disagree moderately for small samples, and it is NOT clear that any one is superior (unless the sampling scheme really does fix BOTH margins, in which case Fisher is better for small sample sizes).

```
# Retrospective Study of Lung Cancer and Smoking
# Subjects chosen to study: 86 lung cancer patients and 86 controls.
ca = matrix(c(83,3,72,14), nrow=2, dimnames=
        list(c("Smoker","Nonsmoker"), c("Cancer", "Control")))
cta(ca)
#           Cancer Control   n      phat         SE         CIlo      CIhi
# Smoker        83      72 155 0.5354839 0.04005971  0.456966849 0.6140009
# Nonsmoker      3      14  17 0.1764706 0.09245944 -0.004749916 0.3576911
# Total         86      86 172 0.5000000 0.12774500  0.249619796 0.7503802
#          diff        SEdiff             Z       p.value          CIlo
# -0.3590132827  0.1277450022 -2.8103900477  0.0003667988 -0.1615144375
#          CIhi
# -0.5565121280
#        OR      ORlo      ORih      p.value
#  5.379630  1.486341 19.470912   0.01035070

cta(t(ca))
#         Smoker Nonsmoker   n      phat         SE      CIlo      CIhi
# Cancer      83         3  86 0.9651163 0.01978573 0.9263363 1.0038963
# Control     72        14  86 0.8372093 0.03980912 0.7591834 0.9152352
# Total      155        17 172 0.9011628 0.04551218 0.8119589 0.9903667
#         diff       SEdiff            Z      p.value         CIlo         CIhi
# -0.127906977  0.045512180 -2.810390048  0.004011857 -0.040775308 -0.215038646
#        OR      ORlo      ORih      p.value
#  5.379630  1.486341 19.470912   0.01035070
```

Question 2: What do you conclude about smoking and lung cancer. What do you conclude about selection of outcome vs. explanatory variable in this setting?

Smoking is associate with lung cancer, with an estimated odds ratio of getting cancer of 5.4 (95% CI =[1.5,19.5]) comparing smokers to non-smokers. Causality is not possible in this type of study. The p-value for $H_0 : \text{OR} = 1$ is 0.010.

The OR is the same regardless of what we consider explanatory vs. outcome. The probabilities differ, and are not used in analysis of retrospective data.

```
cta(cbind(Cancer=ca[,1], Control=2*ca[,2]))
#            Cancer Control   n      phat         SE         CIlo      CIhi
# Smoker         83     144 227 0.3656388 0.03196550  0.302986386 0.4282911
# Nonsmoker       3      28  31 0.0967742 0.05310032 -0.007302425 0.2008508
# Total          86     172 258 0.3333333 0.09026301  0.156417830 0.5102488
#         diff       SEdiff            Z      p.value         CIlo         CIhi
# -2.68865e-01  9.02630e-02 -2.97868e+00  1.43804e-05 -1.47385e-01 -3.90344e-01
#        OR      ORlo      ORih       p.value
#  5.379630  1.586736 18.238959   0.006910168
```

Question 3: What are the observed pitfalls of retrospective research? Just by changing the completely arbitrary choice of how many people to study in each group, the estimation of "cancer rates difference" changes from 36% lower to 27% lower. This estimate is totally dependent on an arbitrary study design choice (in retrospective studies) so it cannot be studied with this design. Only the OR is meaningful.

This study (McCleskey vs. Zant) compares death penalty rates for black defendants in Georgia in the 1980s for 6 different (ordered) aggravation severity levels. The goal is to test whether the death penalty is applied differently depending on the race of the person killed.

```
dp = array(c(2,1,60,181, 2,1,15,21, 6,2,7,9, 9,2,3,4, 9,4,0,3, 17,4,0,0),
       dim=c(2,2,6),dimnames=list(victim=c("White","Black"),
                       DeathPen=c("Yes","No"), aggravation=1:6))
dp
# , , aggravation = 1    , , aggravation = 2
#        DeathPen               DeathPen
# victim  Yes  No        victim  Yes No
#   White   2  60          White   2 15
#   Black   1 181          Black   1 21
# , , aggravation = 3    , , aggravation = 4
#        DeathPen               DeathPen
# victim  Yes No         victim  Yes No
#   White   6  7           White   9  3
#   Black   2  9           Black   2  4
# , , aggravation = 5    , , aggravation = 6
#        DeathPen              DeathPen
# victim  Yes No         victim  Yes No
#   White   9  0           White  17  0
#   Black   4  3           Black   4  0

# Original data (collapsed over aggravation rather than incorporating it):
```

```
cta(cbind(Yes=c(sum(dp[1,1,]),sum(dp[2,1,])),
          No=c(sum(dp[1,2,]),sum(dp[2,2,]))))
#          Yes    No   n       phat         SE       CIlo       CIhi
# Group1    45    85 130 0.34615385 0.04172542 0.26437203 0.42793566
# Group2    14   218 232 0.06034483 0.01563365 0.02970288 0.09098677
# Total     59   303 362 0.16298343 0.04046480 0.08367242 0.24229443
#         diff       SEdiff            Z      p.value         CIlo          CIhi
# -2.85809e-01  4.04648e-02 -7.06315e+00  1.41467e-10 -1.98475e-01  -3.73143e-01
#           OR         ORlo         ORih      p.value
# 8.243697e+00 4.303302e+00 1.579219e+01 2.015553e-10
#      p.chisq     p.Fisher
# 4.683839e-12 5.090836e-12
```

Question 4: Ignoring aggravation level, what is the conclusion? How might this be misleading?

With a tiny p-value (<1e-11), we reject the null hypothesis that getting the death penalty is independent of the victim's race (for black defendants in Georgia in the 1980s). If whites are more often killed under aggravated circumstances (e.g., in the commission of a robbery), then this aggravation could be confounded with victim's race, and could be the actual cause of a higher death penalty rate.

```
# Per aggravation level:
apply(dp, 3, function(x)cta(x)$OR["OR"])
#     1      2      3      4      5      6
# 6.033  2.800  3.857  6.000 14.778  3.889
apply(dp, 3, function(x)cta(x)$OR["p.value"])
#     1     2     3     4     5     6
# 0.145 0.418 0.159 0.101 0.096 0.511

# Test of OR=1 in pooled tables (assuming equal ORs):
mantelhaen.test(dp)
#  Mantel-Haenszel X-squared = 9.6983, df = 1, p-value = 0.001844
# alternative hypothesis: true common odds ratio is not equal to 1
# 95 percent confidence interval:
#   1.910687 15.789312
# sample estimates:
# common odds ratio
#           5.49258

# Check assumption of common odds ratio:
woolf <- function(x) {
```

```
  x <- x + 1 / 2
  k <- dim(x)[3]
  or <- apply(x, 3, function(x) (x[1,1]*x[2,2])/(x[1,2]*x[2,1]))
  w <-  apply(x, 3, function(x) 1 / sum(1 / x))
  1 - pchisq(sum(w * (log(or) - weighted.mean(log(or), w)) ^ 2), k - 1)
}
woolf(dp)
# 0.9597382
```

Question 5: What does the woolf test p-value tell us? What does the Mantel-Haenszel p-value tell us? How do you interpret the CI? What explanations for the higher death penalty for white victims have been pretty much ruled out?

The woolf test says there is no evidence to reject the null hypothesis that the odds ratios differ across tables. Without this, the Mantel-Hansael test is invalid. The Mantel-Haenszel test reject (p=0.0018) the null hypothesis that the common odds ratio is 1, i.e, that correcting for the effects of aggravation, there is no association between victim's race and getting the death penalty. We are 95% confident that the odds of getting the death penalty are 1.9 to 15.8 times higher for a white victim than a black victim. We have pretty much ruled out chance variation and degree of aggravation as causes for the association.