Lesson 5 HW 2 Ellen Chancey

Ellen Chancey

September 23, 2018

Question 3.15

For this table X^2 = 11.5 with P = 0.24 where using scores (3,10,20,35) for income and (1,3,4,5) for job satisfaction. M^2 = 7.04 with P = 0.008. Explain why the results are so different.

The results are quite different because the Pearson Likelihood ratio is detecting any pattern in the table, and because it aims to detect any departure from independence, it is not sensitive to a particular kind of departure, in this case from an ordinal variable. The linear trend alternative however aims to detect only departure from independence based on a directional association, so it is more sensitive to this trend.

How to obtain these values in R:

Set Up

```
job <- matrix(c(2,4,13,3,2,6,22,4,0,1,15,8,0,3,13,8), byrow = TRUE, nrow = 4)
dimnames(job) <- list(income=c("<5", "5-15", "15-25", ">25"),
                      satisfaction=c("very dis", "little satis", "mod satis",
"very satis"))
job
```

```
##        satisfaction
## income  very dis little satis mod satis very satis
##   <5            2            4        13          3
##   5-15          2            6        22          4
##   15-25         0            1        15          8
##   >25           0            3        13          8
```

χ^2 Pearson Likelihood Ratio

```
chisq.test(job)
```

```
## Warning in chisq.test(job): Chi-squared approximation may be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  job
## X-squared = 11.524, df = 9, p-value = 0.2415
```

M^2 Linear Trend Alternative

```
library("vcdExtra")
```

```
## Warning: package 'vcdExtra' was built under R version 3.4.4
```

```
## Loading required package: vcd
```

```
## Loading required package: grid
```

```
## Loading required package: gnm
```

```
## Warning: package 'gnm' was built under R version 3.4.4
```

```
CMHtest(job, rscores = c(3,10,20,35), cscores = c(1,3,4,5))
```

```
## Cochran-Mantel-Haenszel Statistics for income by satisfaction
##
##                   AltHypothesis   Chisq Df      Prob
## cor         Nonzero correlation  7.0449  1 0.0079493
## rmeans   Row mean scores differ 10.0388  3 0.0182395
## cmeans   Col mean scores differ  7.4797  3 0.0580826
## general     General association 11.4134  9 0.2484298
```

Question 3.17b

Use table 2.13. Obtain a 95% CI for gamma. Interpret the assocation.

```
sex <- matrix(c(300,4,4,17,78,15,3,14,107,16,46,54,234,32,35,336), byrow =
TRUE, nrow = 4)
dimnames(sex) <- list(premarital=c("always", "almost always","sometimes",
"not"),
                      homosexual=c("always", "almost always","sometimes",
"not"))
sex
```

```
##                homosexual
## premarital      always almost always sometimes not
##   always           300             4         4  17
##   almost always     78            15         3  14
##   sometimes        107            16        46  54
##   not              234            32        35 336
```

```
library("DescTools")
```

```
## Warning: package 'DescTools' was built under R version 3.4.4
```

```
GoodmanKruskalGamma(sex, conf.level = 0.95)
```

```
##     gamma    lwr.ci    upr.ci
## 0.6573527 0.6053989 0.7093064
```

```
# -1 to +1 association, 0 indicates no association
```

The 95% CI for gamma is (0.6573527,0.7093064). This indicates that there is a very strong positive association between opinion on premarital sex and homosexual sex, as these values are close to 1.

Additional Problem

A study of hospitalized patients found that, among the subjects that were hospitalized for lung cancer the smoking histories defined by number of cigarettes per day (CPD) is in the following table.Using these data, assess whether CPD is associated with lung cancer using the Cochrane-Armitage test. Interpret your results.

Set Up

```
smoking <- matrix(c(7,55,489,475,293,38,61,129,570,431,154,12), byrow = TRUE,
nrow = 2)
dimnames(smoking) <- list(cancer=c("Lung", "Other"),
                          CPD=c("0", "<5", "5-14", "15-24", "25-49",
"50+"))
smoking
```

```
##        CPD
## cancer   0  <5 5-14 15-24 25-49 50+
##   Lung   7  55  489   475   293  38
##   Other 61 129  570   431   154  12
```

Cochran-Armitage Test

```
library("DescTools")
DescTools::CochranArmitageTest(smoking)
```

```
##
##  Cochran-Armitage test for trend
##
## data:  smoking
## Z = 11.368, dim = 6, p-value < 2.2e-16
## alternative hypothesis: two.sided
```

The z statistic value of 11.368 indicates a moderate positive linear association between lung cancer and cigarettes per day with a significant p value.

Session Info

```
sessionInfo()
```

```
## R version 3.4.1 (2017-06-30)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 16299)
##
## Matrix products: default
##
## locale:
```

```
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] grid      stats     graphics  grDevices utils     datasets  methods
## [8] base
##
## other attached packages:
## [1] DescTools_0.99.25 vcdExtra_0.7-1    gnm_1.1-0         vcd_1.4-3
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.13     ca_0.70          knitr_1.16       magrittr_1.5
##  [5] MASS_7.3-47      colorspace_1.3-2 lattice_0.20-35  stringr_1.2.0
##  [9] tools_3.4.1      nnet_7.3-12      htmltools_0.3.6  qvcalc_0.9-1
## [13] yaml_2.1.14      rprojroot_1.2    digest_0.6.12    lmtest_0.9-35
## [17] manipulate_1.0.1 Matrix_1.2-10    relimp_1.0-5     evaluate_0.10.1
## [21] rmarkdown_1.6    stringi_1.1.5    compiler_3.4.1   backports_1.1.0
## [25] boot_1.3-19      expm_0.999-2     mvtnorm_1.0-6    foreign_0.8-69
## [29] zoo_1.8-0
```