Lesson 5 HW 1 Chancey

Ellen Chancey

September 22, 2018

3.11 parts a and b

Set Up

```
raceparty <- matrix(c(192,75,8,459,586,471), byrow = TRUE, nrow = 2)
dimnames(raceparty) <- list(race=c("Black", "White"),
                            party=c("Democrat", "Independent", "Republican"))
raceparty

##        party
## race     Democrat Independent Republican
##    Black      192          75          8
##    White      459         586        471
```

G Test

```
library("DescTools")

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

rpg <- GTest(raceparty)
rpg

##
##  Log likelihood ratio (G-test) test of independence without
##  correction
##
## data:  raceparty
## G = 197.39, X-squared df = 2, p-value < 2.2e-16
```

The G^2 value of this contingency table is very high. This provides evidence to reject the null hypothesis, $H_0: \pi_{ij} = \pi_{i+}\pi_{+j}$. The p-value indicates that this method will rarely reject the H_0 when its true.

Standardized Residuals

```
rppf<- procfreq(raceparty)
rppf$adjusted.residuals

##         party
## race      Democrat Independent Republican
##    Black  12.54209   -3.598555   -9.70631
##    White -12.54209    3.598555    9.70631
```

Additional Problem 1

Investigators studied the relationship between stage of breast cancer at diagnosis (local or advance) and a woman's living arrangement. Of 144 women living alone, 41.0% had an advanced case; of 209 living with a spouse, 52.2% were advanced; and of 89 living with others, 59.6% were advanced. The investigators reported their P-value for the relationship as 0.02 (Moritz and Satariano, J. Clin. Epidemiol. 45:443-454, 1993). Reconstruct the analysis performed to obtain this P-value.

Set Up

```r
cancer <- matrix(c(
            round(144*.41),
            round(209*.522),
            round(89*.596),
            round(144*.59),
            round(209*.478),
            round(89*.404)),
            byrow = TRUE,
            nrow = 2)
dimnames(cancer) <- list(stage=c("Advanced", "Local"),
                         living=c("Alone", "Spouse", "Others"))
cancer <- t(cancer)
cancer

##          stage
## living   Advanced Local
##    Alone       59    85
##    Spouse     109   100
##    Others      53    36
```

Method One

```r
library(MASS)
chisq.test(cancer)

##
##  Pearson's Chi-squared test
##
## data:  cancer
## X-squared = 8.3292, df = 2, p-value = 0.01554
```

Method Two

```r
cancerpf <- procfreq(cancer)
cancerpf$chi.square

## [1] 8.329195
## attr(,"P-value")
## [1] 0.01553596
```

I assume that researchers round up, given that I obtain a p value of 0.015.

Additional Problem 2

Table 3.13 (p 106) presents a contingency table for a sample of patients by their diagnosis and whether they were prescribed medication. Obtain standardized Pearson residuals for the independence model. Interpret these results.

Set Up

```
drug <- matrix(c(105,8,12,2,18,19,47,52,0,13), byrow = TRUE, nrow = 5)
dimnames(drug) <- list(diagnosis=c("Schizophrenia", "Affective", "Neurosis",
"Personality", "Special"),
                              drug=c("Yes", "No"))
drug
```

```
##                 drug
## diagnosis        Yes No
##    Schizophrenia 105  8
##    Affective      12  2
##    Neurosis       18 19
##    Personality    47 52
##    Special         0 13
```

```
chisq.test(drug, correct = TRUE)
```

```
## Warning in chisq.test(drug, correct = TRUE): Chi-squared approximation may
## be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  drug
## X-squared = 84.188, df = 4, p-value < 2.2e-16
```

The X^2 value is 84.188 with a very small p value, well below 0.05. This result provides evidence to reject the null hypothesis - $H_0: \pi_{ij} = \pi_{i+}\pi_{+j}$. The p-value indicates that this method will rarely reject the H_0 when its true.

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] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] MASS_7.3-47       DescTools_0.99.25
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.13     lattice_0.20-35  mvtnorm_1.0-6    digest_0.6.12
##  [5] rprojroot_1.2    grid_3.4.1       backports_1.1.0  magrittr_1.5
##  [9] evaluate_0.10.1  stringi_1.1.5    Matrix_1.2-10    boot_1.3-19
## [13] rmarkdown_1.6    tools_3.4.1      stringr_1.2.0    foreign_0.8-69
## [17] yaml_2.1.14      compiler_3.4.1   manipulate_1.0.1 htmltools_0.3.6
## [21] knitr_1.16       expm_0.999-2
```