Name Jerry

Student Number ______

STA 312f2012 Quiz 11

1. (2 points) For the Florida Death Penalty data, you were asked whether the model $(PR, VR)(DP)$ fits adequately.

   (a) Give the value of the Likelihood Ratio (not Pearson) test statistic. The answer is a number from your printout.

   $G^2 = 8.13$

   (b) What is the critical value at $\alpha = 0.05$? The answer is a number.

   7.815

   (c) By this criterion, does the model fit adequately? Answer Yes or No.

   No

2. (3 points) Planning another study of race and the death penalty, assume there are approximately equal numbers of Black and White prisoners, which is the case with the Florida data. In a 2×2 table of race by death penalty, suppose the probability of a White prisoner getting the death penalty is 10%, while the probability of a Black prisoner getting the death penalty is 15%. What sample size is required for the power of the Pearson chi-square test to reach 0.80? The answer is a number that should be on your printout.

   1374

3. (3 points) For the study of gender (G), being on the honour roll (H) and traffic accidents (A), you were asked for the smallest sample size that could reject the conditional independence model (AG)(GH) with probability 0.90 at $\alpha = 0.05$, given a set of cell probabilities. It does not matter whether you did it for a Pearson test or a likelihood ratio test; either will do. What is the required sample size? The answer is an integer.

   n = 3613 for Likelihood ratio,
   3982 for Pearson

Please attach your printout or printouts. Make sure your name is written on each one.

STA 312f2012 Formulas

$$Z_1 = \frac{\sqrt{n}(p-\pi_0)}{\sqrt{\pi_0(1-\pi_0)}} \quad Z_2 = \frac{\sqrt{n}(p-\pi_0)}{\sqrt{p(1-p)}} \quad p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

```
> qnorm(0.975)
[1] 1.959964
> qnorm(0.995)
[1] 2.575829
```

$$Z_2^2 \dot{\sim} \chi^2(1,\lambda), \text{ with } \lambda = n\,\frac{(\pi-\pi_0)^2}{\pi(1-\pi)}$$

$$\theta = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}$$

$$P(n_1,\ldots,n_c) = \binom{n}{n_1 \cdots n_c}\pi_1^{n_1}\cdots\pi_c^{n_c}$$

$$\ell(\boldsymbol{\pi}) = \prod_{i=1}^n \pi_1^{y_{i,1}}\pi_2^{y_{i,2}}\cdots\pi_c^{y_{i,c}} = \pi_1^{n_1}\pi_2^{n_2}\cdots\pi_c^{n_c}$$

$$G^2 = -2\log\left(\frac{\max_{\beta\in\mathcal{B}_0}\ell(\beta)}{\max_{\beta\in\mathcal{B}}\ell(\beta)}\right) = -2\log\left(\frac{\ell_0}{\ell_1}\right)$$

$$\widehat{\mu}_j = n\,\widehat{\pi}_j \quad , \quad \widehat{\mu}_{ij} = \frac{n_{i+}n_{+j}}{n}$$

$$X^2 = \sum_{j=1}^c \frac{(n_j-\widehat{\mu}_j)^2}{\widehat{\mu}_j} = n\sum_{j=1}^c \frac{(p_j-\widehat{\pi}_j)^2}{\widehat{\pi}_j}$$

$$G^2 = 2\sum_{j=1}^c n_j\log\left(\frac{n_j}{\widehat{\mu}_j}\right) = 2n\sum_{j=1}^c p_j\log\left(\frac{p_j}{\widehat{\pi}_j}\right)$$

$$\lambda = n\sum_{j=1}^c \frac{[\pi_j-\pi_j(M)]^2}{\pi_j(M)}$$

$$\lambda = 2n\sum_{j=1}^c \pi_j\log\left(\frac{\pi_j}{\pi_j(M)}\right)$$

$$X^2 = n\sum_{i=1}^I\sum_{j=1}^J \frac{(p_{ij}-p_{i+}p_{+j})^2}{p_{i+}p_{+j}}$$

$$G^2 = 2n\sum_{i=1}^I\sum_{j=1}^J p_{ij}\log\left(\frac{p_{ij}}{p_{i+}p_{+j}}\right)$$

$$\lambda = n\sum_{i=1}^I\sum_{j=1}^J \frac{(\pi_{ij}-\pi_{i+}\pi_{+j})^2}{\pi_{i+}\pi_{+j}}$$

$$\lambda = 2n\sum_{i=1}^I\sum_{j=1}^J \pi_{ij}\log\left(\frac{\pi_{ij}}{\pi_{i+}\pi_{+j}}\right)$$

$$\log\mu_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}$$

```
> df = 1:8
> CriticalValue = qchisq(0.95,df)
> round(rbind(df,CriticalValue),3)
               [,1]  [,2]  [,3]  [,4]  [,5]   [,6]   [,7]   [,8]
df            1.000 2.000 3.000 4.000  5.00  6.000  7.000  8.000
CriticalValue 3.841 5.991 7.815 9.488 11.07 12.592 14.067 15.507

> # Table of non-centrality values needed for specified power
   Power
df       0.5       0.7       0.8      0.9
  1 3.841022  6.172007  7.848883 10.50739
  2 4.956736  7.701774  9.634685 12.65394
  3 5.760482  8.792389 10.902570 14.17149
  4 6.419476  9.682473 11.935286 15.40503
  5 6.991270 10.452523 12.827607 16.46946
  6 7.503313 11.140677 13.624286 17.41883
  7 7.971192 11.768443 14.350527 18.28355
  8 8.404641 12.349332 15.022138 19.08270
```

Quiz 11 Printout

```
> # DeathPen (Q1)
>
> deathpen =
read.table("http://fisher.utstat.utoronto.ca/~brunner/312f12/code_n_data/de
athpen2.data",header=T)
> deathpen[1:4,]
  Prace Vrace Death
1 White White   Yes
2 White White   Yes
3 White White   Yes
4 White White   Yes
> attach(deathpen)
> deathrace = table(Prace, Vrace, Death)
> deathrace
, , Death = No

       Vrace
Prace    Black White
  Black     97    52
  White      9   132

, , Death = Yes

       Vrace
Prace    Black White
  Black      6    11
  White      0    19

> Q1 = loglin(deathrace, margin=list(c(1,2),3)); Q1
2 iterations: deviation 0
$lrt
[1] 8.131611

$pearson
[1] 6.977343

$df
[1] 3

$margin
$margin[[1]]
[1] "Prace" "Vrace"
```

```
$margin[[2]]
[1] "Death"


>
> # Q2 Power for another DP study.
>
> dummy2 = rbind(c(10,90),
+                c(15,85))
> eff2 = loglin(dummy2,margin=list(1,2))$pearson/200
2 iterations: deviation 0
> 7.848883/eff2
[1] 1373.555
>
>
> # Q3 AHG test (AG)(GH). Want power of 0.90
>
> Male = rbind(c(0.005777638, 0.144224),
+              c(0.024222362, 0.325776) )
>
> Female =  rbind(c(0.004222372,  0.04577688),
+                 c(0.065777628,  0.38422312) )
>
> truth = numeric(8); dim(truth) = c(2,2,2)
> truth[,,1]=Male; truth[,,2]=Female
> truth
, , 1

            [,1]     [,2]
[1,] 0.005777638 0.144224
[2,] 0.024222362 0.325776

, , 2

            [,1]       [,2]
[1,] 0.004222372 0.04577688
[2,] 0.065777628 0.38422312

> truth=truth*1000
>
> dummy = loglin(truth,margin=list(c(1,3),c(2,3)))
2 iterations: deviation 5.684342e-14
```

```
> dummy
$lrt
[1] 3.501477

$pearson
[1] 3.177402

$df
[1] 2

$margin
$margin[[1]]
[1] 1 3

$margin[[2]]
[1] 2 3


> efflrt = dummy$lrt/1000; effpearson = dummy$pearson/1000
> nlrt = 12.65394/efflrt; nlrt
[1] 3613.886
> npearson = 12.65394/effpearson; npearson
[1] 3982.48
```