STA 312f2012 Quiz 10

1. For homework, you examined a log-linear model for a 3×2 contingency table of race by belief in life after death

   (a) (1 point) Write a *saturated* linear model for $\log \mu_{ij}$ using the λ notation. Race is X and Belief in afterlife is Y.

   $$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}$$

   (b) (1 point) Write a saturated linear model for $\log \mu_{ij}$ using a regression model with effect coding; this is the β notation. You do not have to say how the dummy variables are defined, but the dummy variables for race should be named r-something and the dummy variables for belief should be named b-something so we know what they are.

   $$\log \mu_{ij} = \beta_0 + \beta_1 r_1 + \beta_2 r_2 + \beta_3 b + \beta_4 r_1 b + \beta_5 r_2 b$$

   (c) (1 point) In terms of the λ notation, what is the null hypothesis when you are testing for independence?

   $$H_0 : \quad \lambda_{ij}^{XY} = 0$$

   (d) (1 point) In terms of the λ notation, how many quantities (not unique quantities) are you setting to zero under H_0? The answer is a number.

   $$6$$

   (e) (1 point) In terms of the β notation, what is the null hypothesis when you are testing for independence?

   $$H_0 : \quad \beta_4 = \beta_5 = 0$$

   (f) (1 point) In terms of the β notation, how many quantities are you setting to zero under H_0? The answer is a number.

$$2$$

2. For the table relating Alcohol use (A), Cigarette use (B) and Marijuana use (M), you used R fit the model $(AC)(AM)(CM)$.

   (a) (1 point) Under the model, what is the estimated expected number of students who do not drink alcohol, do not smoke cigarettes, and do not use marijuana? The answer is a number from your printout.

   $$279.6144$$

   (b) (1 point) What is the Likelihood Ratio test statistic G^2 for the test of model fit? The answer is a number from your printout.

   $$G^2 = 0.374$$

   (c) (2 points) You tested whether Alcohol and cigarette use were conditionally independent given Cigarette use. Give the Likelihood Ratio test statistic G^2. The answer is a number from your printout.

   $$G^2 = 91.64$$

Attach your printout for Question 2 (Homework Question 3). Make sure your name is written on the printout.

Quiz 10 Printout

```
> # Question 3, Data of Table 7.3 on p. 209.
>
> drugs = read.table(stdin())
0:     Alcohol Cigarette  Marijuana    Freq
1: 1     Yes       Yes        Yes       911
2: 2     Yes       Yes        No        538
3: 3     Yes       No         Yes        44
4: 4     Yes       No         No        456
5: 5     No        Yes        Yes         3
6: 6     No        Yes        No         43
7: 7     No        No         Yes         2
8: 8     No        No         No        279
9:
> # Take a look: See 2 sub-tables
> xtabs(Freq ~ Cigarette+Marijuana+Alcohol,data=drugs)
, , Alcohol = No

          Marijuana
Cigarette  No Yes
      No  279   2
      Yes  43   3

, , Alcohol = Yes

          Marijuana
Cigarette  No Yes
      No  456  44
      Yes 538 911
```

```
> # Order Yes and No is alphabetical, different from text.
> dtab = xtabs(Freq ~ Alcohol+Cigarette+Marijuana,data=drugs)
> drugmod = loglin(dtab,margin=list(c(1,2),c(1,3),c(2,3)),fit=T,param=T)
5 iterations: deviation 0.03408344
> drugmod
$lrt
[1] 0.3739859    2b

$pearson
[1] 0.4010998

$df
[1] 1

$margin
$margin[[1]]
[1] "Alcohol"    "Cigarette"

$margin[[2]]
[1] "Alcohol"    "Marijuana"

$margin[[3]]
[1] "Cigarette" "Marijuana"


$fit
, , Marijuana = No

      Cigarette
Alcohol        No        Yes
    No  279.614402   42.383882     2a
    Yes 455.385598  538.616118

, , Marijuana = Yes

      Cigarette
Alcohol        No        Yes
    No    1.383160    3.616919
    Yes  44.616840  910.383081
```

```
$param
$param$`(Intercept)`
[1] 4.251537

$param$Alcohol
       No       Yes
-1.503994  1.503994

$param$Cigarette
        No       Yes
-0.2822777  0.2822777

$param$Marijuana
       No       Yes
 1.196045 -1.196045

$param$Alcohol.Cigarette
       Cigarette
Alcohol        No        Yes
    No   0.5136255 -0.5136255
    Yes -0.5136255  0.5136255

$param$Alcohol.Marijuana
       Marijuana
Alcohol       No       Yes
    No   0.746502 -0.746502
    Yes -0.746502  0.746502

$param$Cigarette.Marijuana
         Marijuana
Cigarette        No        Yes
      No   0.7119739 -0.7119739
      Yes -0.7119739  0.7119739

> # Test of A*M from Table 7.5 gives G^2 = 91.64
> redAM = loglin(dtab,margin=list(c(1,2),c(2,3))) # (1,3) is missing
2 iterations: deviation 0
> G2 = redAM$lrt-drugmod$lrt; G2  # 91.64437
[1] 91.64437  2c
```