HW7

Jared Yu

December 6, 2018

1. Assuming that $I = J = K = 3$:

1 a) In $\lambda_i^X, \lambda_j^Y, \lambda_k^Z$ there are $(I-1), (J-1), (K-1)$ or $(3-1), (3-1), (3-1)$ or 2 non-zero terms each.

1 b) In $\lambda_{ij}^{XY}, \lambda_{ik}^{XZ}, \lambda_{jk}^{YZ}$ there are $(I-1)(J-1), (I-1)(K-1), (J-1)(K-1)$ or $(2)(2), (2)(2), (2)(2)$ or 4 non-zero terms for each.

1 c) In λ_{ijk}^{XYZ} there are $(I-1)(J-1)(K-1)$ free parameters, therefore there are $(3-1)(3-1)(3-1) = (2)(2)(2) = 8$ non-zero terms.

1 d) Thus in the saturated model with $log(\mu_{ijk}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}$ there are $2 + 2 + 2 + 4 + 4 + 4 + 8 = 26$ non-zero terms.

2.

```
HD <- read.csv('HD.csv')
HD
```

```
##      weight heart Freq
## 1 NotOver    Yes   55
## 2    Over    Yes   30
## 3 NotOver     No   65
## 4    Over     No   50
```

2 a) The estimates for the λ parameters are below:

```
# declare weight and heart as factors
HD$weight=as.factor(HD$weight)
HD$heart=as.factor(HD$heart)
# fit a log-linear model
fit=glm(Freq ~weight + heart,data=HD, family = "poisson")
fit
```

```
##
## Call:  glm(formula = Freq ~ weight + heart, family = "poisson", data = HD)
##
## Coefficients:
## (Intercept)    weightOver       heartYes
##      4.2341       -0.4055        -0.3023
##
## Degrees of Freedom: 3 Total (i.e. Null);  1 Residual
## Null Deviance:        13.94
## Residual Deviance: 1.371     AIC: 30.23
```

2 b) Below are the estimated frequencies.

```
fit$fitted.values
```

```
##  1  2  3  4
## 51 34 69 46
```

Below are the matching calculations for the estimated frequencies.

```
n1p=55+30
n2p=65+50
np1=55+65
np2=30+50
n=sum(HD$Freq)
c(n1p*np1/n, n1p*np2/n, n2p*np1/n, n2p*np2/n)
```

```
## [1] 51 34 69 46
```

2 c) G^2 :

```
G=1.371
```

AIC :

```
AIC=30.23
```

$p - value$:

```
p_value=1-pchisq(G,1)
```

2 d) $H_0 : log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y$ for all i, j

$H_1 : log(\mu_{ij}) \neq \lambda + \lambda_i^X + \lambda_j^Y$ for some i, j

The p-value is 0.2416401, which is larger than $\alpha = 0.05$, therefore we fail to reject the null hypothesis. The conclusion is that we cannot reject the null hypothesis that $log(\mu_{ij}) = \lambda + \lambda_i^X + \lambda_j^Y$ for all i, j at the significance level $\alpha = 0.05$.

3.

```
bp <- read.csv('bepatient.csv')
bp
```

```
##    Cured Gender Group Freq
## 1    No      F   One   16
## 2   Yes      F   One   32
## 3    No      M   One    8
## 4   Yes      M   One   59
## 5    No      F Three   56
## 6   Yes      F Three   80
## 7    No      M Three   17
## 8   Yes      M Three  107
## 9    No      F   Two   33
## 10  Yes      F   Two   24
## 11   No      M   Two   12
## 12  Yes      M   Two   55
```

```
bp$Cured=as.factor(bp$Cured)
bp$Gender=as.factor(bp$Gender)
bp$Group=as.factor(bp$Group)
fit2=glm(Freq ~Cured+Gender+Group + Cured*Group + Gender*Group, data=bp,family = "poisson")
fit2
```

```
##
## Call:  glm(formula = Freq ~ Cured + Gender + Group + Cured * Group +
##      Gender * Group, family = "poisson", data = bp)
##
## Coefficients:
##         (Intercept)                CuredYes                 GenderM
##              2.3043                  1.3328                  0.3335
##           GroupThree                GroupTwo   CuredYes:GroupThree
##              1.3381                  0.7251                 -0.3922
##    CuredYes:GroupTwo   GenderM:GroupThree       GenderM:GroupTwo
##             -0.7700                 -0.4259                 -0.1719
##
## Degrees of Freedom: 11 Total (i.e. Null);  3 Residual
## Null Deviance:       236.7
## Residual Deviance: 54.9  AIC: 136.5
```

3 a)

G^2 :

```
G2=54.9
G2
```

```
## [1] 54.9
```

$p - value :$

```
p_value2=1-pchisq(G2,1)
p_value2
```

```
## [1] 1.267875e-13
```

$X = Cured, Y = Gender, Z = Group$

$H_0 : log(\mu_{ijk}) = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$ for all i, j, k

$H_1 : log(\mu_{ijk}) \neq \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$ for some i, j, k

Here the conditional independence of Cured and Gender given Group has a $p - value$ that is less than 0.05. Therefore we reject the null hypothesis in favor of the alternative hypothesis at $\alpha = 0.05$. So in the end we cannot conclude that the model of conditional independence is reasonable for this data.

3 b)

```
library(MASS)
fit3=glm(Freq ~Cured+Gender+Group+Cured*Gender+Cured*Group+Gender*Group+Cured*Gender*Grou
p, data=bp,family = "poisson")
fit4=stepAIC(fit3,direction = "backward", k=2)
```

```
## Start:  AIC=87.61
## Freq ~ Cured + Gender + Group + Cured * Gender + Cured * Group +
##      Gender * Group + Cured * Gender * Group
##
##                            Df Deviance    AIC
## - Cured:Gender:Group  2  0.79416 84.409
## <none>                        0.00000 87.615
##
## Step:  AIC=84.41
## Freq ~ Cured + Gender + Group + Cured:Gender + Cured:Group +
##      Gender:Group
##
##                       Df Deviance    AIC
## <none>                     0.794   84.409
## - Gender:Group   2     5.041   84.655
## - Cured:Group    2     8.146   87.761
## - Cured:Gender   1    54.898 136.513
```

Below is an output of the parameter estimates, standard errors, z-values and p-values:

```
summary(fit4)
```

```
##
## Call:
## glm(formula = Freq ~ Cured + Gender + Group + Cured:Gender +
##     Cured:Group + Gender:Group, family = "poisson", data = bp)
##
## Deviance Residuals:
##        1         2         3         4         5         6         7
## -0.24972   0.18222   0.37711  -0.13201  -0.09302   0.07838   0.17191
##        8         9        10        11        12
## -0.06743   0.30456  -0.34280  -0.47428   0.23422
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)          2.83438    0.21848  12.973  < 2e-16 ***
## CuredYes             0.59897    0.25411   2.357 0.018418 *
## GenderM             -0.89129    0.26779  -3.328 0.000874 ***
## GroupThree           1.20338    0.24439   4.924 8.48e-07 ***
## GroupTwo             0.60864    0.26599   2.288 0.022123 *
## CuredYes:GenderM     1.55262    0.22332   6.952 3.59e-12 ***
## CuredYes:GroupThree -0.26348    0.28176  -0.935 0.349721
## CuredYes:GroupTwo   -0.79477    0.31343  -2.536 0.011222 *
## GenderM:GroupThree  -0.35524    0.23783  -1.494 0.135267
## GenderM:GroupTwo     0.06704    0.27905   0.240 0.810142
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 236.74514  on 11  degrees of freedom
## Residual deviance:   0.79416  on  2  degrees of freedom
## AIC: 84.409
##
## Number of Fisher Scoring iterations: 4
```

3 c)

$$log\theta_{XY(1)} = log\frac{\mu_{111}\mu_{221}}{\mu_{121}\mu_{211}} = (\lambda + \lambda_1^X + \lambda_1^Y + \lambda_1^Z + \lambda_{11}^{XZ} + \lambda_{11}^{YZ}) + (\lambda + \lambda_2^X + \lambda_2^Y + \lambda_1^Z + \lambda_{21}^{XZ} + \lambda_{21}^{YZ})$$
$$-(\lambda + \lambda_1^X + \lambda_2^Y + \lambda_1^Z + \lambda_{11}^{XZ} + \lambda_{21}^{YZ}) - (\lambda + \lambda_2^X + \lambda_1^Y + \lambda_1^Z + \lambda_{21}^{XZ} + \lambda_{11}^{YZ}) = 0$$
$$log\theta_{XY(2)} = log\frac{\mu_{112}\mu_{222}}{\mu_{122}\mu_{212}} = (\lambda + \lambda_1^X + \lambda_1^Y + \lambda_2^Z + \lambda_{12}^{XZ} + \lambda_{12}^{YZ}) + (\lambda + \lambda_2^X + \lambda_2^Y + \lambda_2^Z + \lambda_{22}^{XZ} + \lambda_{22}^{YZ})$$
$$-(\lambda + \lambda_1^X + \lambda_2^Y + \lambda_2^Z + \lambda_{12}^{XZ} + \lambda_{22}^{YZ}) - (\lambda + \lambda_2^X + \lambda_1^Y + \lambda_2^Z + \lambda_{22}^{XZ} + \lambda_{12}^{YZ}) = 0$$
$$log\theta_{XY(3)} = log\frac{\mu_{113}\mu_{223}}{\mu_{123}\mu_{213}} = (\lambda + \lambda_1^X + \lambda_1^Y + \lambda_3^Z + \lambda_{13}^{XZ} + \lambda_{13}^{YZ}) + (\lambda + \lambda_2^X + \lambda_2^Y + \lambda_3^Z + \lambda_{23}^{XZ} + \lambda_{23}^{YZ})$$
$$-(\lambda + \lambda_1^X + \lambda_2^Y + \lambda_3^Z + \lambda_{13}^{XZ} + \lambda_{23}^{YZ}) - (\lambda + \lambda_2^X + \lambda_1^Y + \lambda_3^Z + \lambda_{23}^{XZ} + \lambda_{13}^{YZ}) = 0$$