Practice problems and their solutions – do NOT need to submit

1. The table below cross-classifies a sample of Americans according to their gender and opinion about an afterlife. Does an association exist between gender and belief in an afterlife? Is one gender more likely than the other to believe in an afterlife, or is belief in an afterlife independent of gender?

	Belief in Afterlife	
	Yes	No or Undecided
Females	509	116
Males	398	104

Let X =gender and Y =belief in an afterlife.

(a) The output below shows the results of fitting the independence log-linear model. Report and interpret the results of goodness-of-fit test.

(b) Report $\{\hat{\lambda}_j^Y\}$. Interpret $\hat{\lambda}_1^Y - \hat{\lambda}_2^Y$.

First note that the model of independence is

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

$$\log \mu_{ij} = \lambda + \lambda_i^{gender} + \lambda_j^{belief}$$

Here are three different outputs for the same problem, and the corresponding solutions:

- 1) Based on R, glm() function; see R code

```
Call:
glm(formula = Freq ~ gender + belief, family = poisson(), data = belief)

Deviance Residuals:
      1        2        3        4
 0.2672  -0.2995  -0.5482   0.6006

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  6.22058    0.04261 146.004  < 2e-16 ***
gendermale  -0.21915    0.05993  -3.657 0.000256 ***
beliefno    -1.41651    0.07515 -18.848  < 2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 463.85468  on 3  degrees of freedom
Residual deviance:   0.82242  on 1  degrees of freedom
AIC: 35.795

Number of Fisher Scoring iterations: 3
```

solution

- ”Residual deviance” gives us the deviance value, $G^2 = 0.82242$, $df = 1$ which indicates good fit. The p-value is $1 - pchisq(0.8224, df = 1) = 0.3645$. The model of independence fits well.

  We can round off these values and report to 2 or 3 decimal places, e.g., $G^2 = 0.82$, $df = 1, p - value = 0.36$.

  Some extra info: for example, the log of expected counts is:

  $$\log \hat{\mu}_{11=female,yes} = 6.22 + 0 + 0,$$

  and thus the expected count is $exp(6.22058) = 502.9949$
  You can get the rest of the values by running

```
>m.ind1$fitted
        1         2         3         4
502.99468 404.00532 122.00532  97.99468
```

- Report $\{\hat{\lambda}_j^Y\}$. More specifically $\hat{\lambda}_1^Y = \hat{\lambda}_{yes}^Y = 0$ and $\hat{\lambda}_2^Y = \hat{\lambda}_{no}^Y = -1.41651$. Then $\hat{\lambda}_1^Y - \hat{\lambda}_2^Y = 0 - (-1.41651) = 1.41651$.

  Interpret: Given gender, estimated odds of belief in afterlife equal $exp(1.41651) \approx 4.1$.

  Notice that since this model of independence, odds-ratios are assumed to be equal to 1, and the odds are the same for each gender.

2) Based on R, loglin() function; see R code

```
$lrt
[1] 0.8224158

$pearson
[1] 0.8245764

$df
[1] 1

$margin
$margin[[1]]
[1] "gender"

$margin[[2]]
[1] "belief"


$fit
        belief
gender         yes        no
  female 502.99468 122.00532
  male   404.00532  97.99468

$param
$param$`(Intercept)`
[1] 5.402746

$param$gender
    female       male
 0.1095758 -0.1095758

$param$belief
       yes         no
 0.7082575 -0.7082575
```

solution

- "\$lrt" gives us the deviance value, $G^2 = 0.82242$, and "\$pearson" gives the $X^2 = 0.82458$, $df = 1$, which indicates good fit. The p-value is $1 - pchisq(0.8224, df = 1) = 0.3645$. The model of independence fits well.

  We can round off these values and report to 2 or 3 decimal places, e.g., $G^2 = X^2 = 0.82$, $df = 1$, $p - value = 0.36$.

  Some extra info: for example, the log of expected counts is:

  $$\log \hat{\mu}_{11=female,yes} = 5.403 + 0.1096 + 0.7083,$$

  and thus the expected count is $exp(6.2209) = 502.9947$

  You can get all the fitted (expected) values from the output labeled "\$fit"; see above.

- Report $\{\hat{\lambda}_j^Y\}$. More specifically, $\hat{\lambda}_1^Y = \hat{\lambda}_{yes}^Y = 0.7082575$ and $\hat{\lambda}_2^Y = \hat{\lambda}_{no}^Y = -0.7082575$. Then $\hat{\lambda}_1^Y - \hat{\lambda}_2^Y = 0.7082575 - (-0.7082575) = 1.41651$.

  Interpret: Given gender, estimated odds of belief in afterlife equal $exp(1.41651) \approx 4.1$.

  Notice that since this model of independence, odds-ratios are assumed to be equal to 1, and the odds are the same for each gender.

2. 3) Based on SAS, PROC GENMOD; see SAS code

```
Criteria For Assessing Goodness Of Fit

Criterion                 DF           Value        Value/DF

Deviance                   1          0.8224          0.8224
Scaled Deviance            1          0.8224          0.8224
Pearson Chi-Square         1          0.8246          0.8246
Scaled Pearson X2          1          0.8246          0.8246
Log Likelihood                     5461.9458

                            Algorithm converged.

Analysis Of Parameter Estimates

                                            Standard     Wald 95% Confidence       Chi-
 Parameter             DF     Estimate       Error           Limits            Square    Pr > ChiSq

Intercept              1       4.5849       0.0752       4.4376      4.7322    3720.50        <.0001
gender       female    1       0.2192       0.0599       0.1017      0.3366      13.37        0.0003
gender       male      0       0.0000       0.0000       0.0000      0.0000        .            .
belief       yes       1       1.4165       0.0752       1.2692      1.5638     355.26        <.0001
belief       no        0       0.0000       0.0000       0.0000      0.0000        .            .
Scale                  0       1.0000       0.0000       1.0000      1.0000
```

solution

- ”Deviance” gives us the deviance value, $G^2 = 0.82242$, and ”Pearson Chi-Squre” gives the $X^2 = 0.8246$, $df = 1$, which indicates good fit. The p-value is $1 - pchisq(0.8224, df = 1) = 0.3645$. The model of independence fits well.

  We can round off these values and report to 2 or 3 decimal places, e.g., $G^2 = X^2 = 0.82$, $df = 1$, $p - value = 0.36$.

  Some extra info: for example, the log of expected counts is:

  $$\log \hat{\mu}_{11=female,yes} = 4.5849 + 0.2192 + 1.4165,$$

  and thus the expected count is $exp(6.2209) = 502.9947$

  You can get all the fitted (expected) values look at the portion of the output labeled ”Observation Statistics” that you get by invoking OPTION ”obstats”: e.g.,

  ```
  model count=gender belief /link=log dist=poisson obstats residuals
  ```

  Values corresponding to the column label ”Pred” are the expected (predicted) values, e.g., fourth value in each row out of possible 15 statistics that the SAS gives you. For example, for the first cell with gender=female and belief=yes, we observe 509 counts and the expected count is 502.99467.

```
                                        The GENMOD Procedure

                                       Observation Statistics

Observation       count    gender    belief          Pred         Xbeta           Std       HessWgt         Lower         Upper
                              Resraw        Reschi        Resdev      StResdev      StReschi        Reslik

          1         509    female    yes        502.99467     6.2205796     0.0426054     502.99467     462.69793     546.80088
                            6.005331     0.2677659     0.2672357     0.9062646     0.9080626     0.9079064
          2         116    female    no         122.00546     4.8040658      0.072513     122.00546     105.84152     140.63793
                           -6.005459     -0.543697     -0.548251     -0.915689     -0.908082     -0.910816
          3         398    male      yes        404.00534     6.0014281     0.0469814     404.00534     368.46516     442.97353
                           -6.005336     -0.298775     -0.299519     -0.910327     -0.908063     -0.908309
          4         104    male      no         97.994789     4.5849143     0.0751676     97.994789     84.570753     113.54964
                           6.0052111     0.6066341     0.6005912     0.8989992     0.9080444     0.9040186
```

- Report $\{\hat{\lambda}_j^Y\}$. More specifically, $\hat{\lambda}_1^Y = \hat{\lambda}_{yes}^Y = 1.4165$ and $\hat{\lambda}_2^Y = \hat{\lambda}_{no}^Y = 0$. Then $\hat{\lambda}_1^Y - \hat{\lambda}_2^Y = 1.4165 - 0 = 1.41651$.

  Interpret: Given gender, estimated odds of belief in afterlife equal $exp(1.41651) \approx 4.1$.

  Notice that since this model of independence, odds-ratios are assumed to be equal to 1, and the odds are the same for each gender.