Case Study 16.1: Gender bias in admissions to Berkeley

Tou Ohone Andate - staff number 1234567

Problem

The University of California, Berkeley, was sued for gender bias against women based on the enrolment decisions given to 4526 students who applied for entry into the graduate program in 1973.

The variables of interest were:

- `Dept`: Six-level categorical variable with levels 'A', 'B', 'C', 'D', and 'E'.
- `Gender`: Two-level categorical variable with levels 'F' and 'M'.
- `Outcome`: Two-level categorical variable with levels 'In' and Out'.
- `Freq`: A count variable which corresponds to the observed combination of `Dept`, `Gender`, and `Outcome`.

Question of Interest

We want examine the association between gender and admission outcome.

### Read in and naive analysis

```
Berk.df = read.table("BerkeleyEnrolments.txt", header = TRUE)
head(Berk.df)
```

```
##   Dept Gender Outcome Freq
## 1    A      M      In  512
## 2    A      M     Out  313
## 3    A      F      In   89
## 4    A      F     Out   19
## 5    B      M      In  353
## 6    B      M     Out  207
```

```
Berk.df=transform(Berk.df,Dept=factor(Dept),Gender=factor(Gender),Outcome=factor(Outcome))
Berk.tbl = xtabs(Freq ~ Gender + Outcome, data = Berk.df)
Berk.tbl
```

```
##       Outcome
## Gender   In  Out
##      F  557 1278
##      M 1198 1493
```

```
barplot(t(Berk.tbl), main="Berkeley admissions  by gender", col=c("lightblue", "lightgreen"),
xlab="gender",legend.text=c("In", "Out"))
```

Berkeley admissions by gender
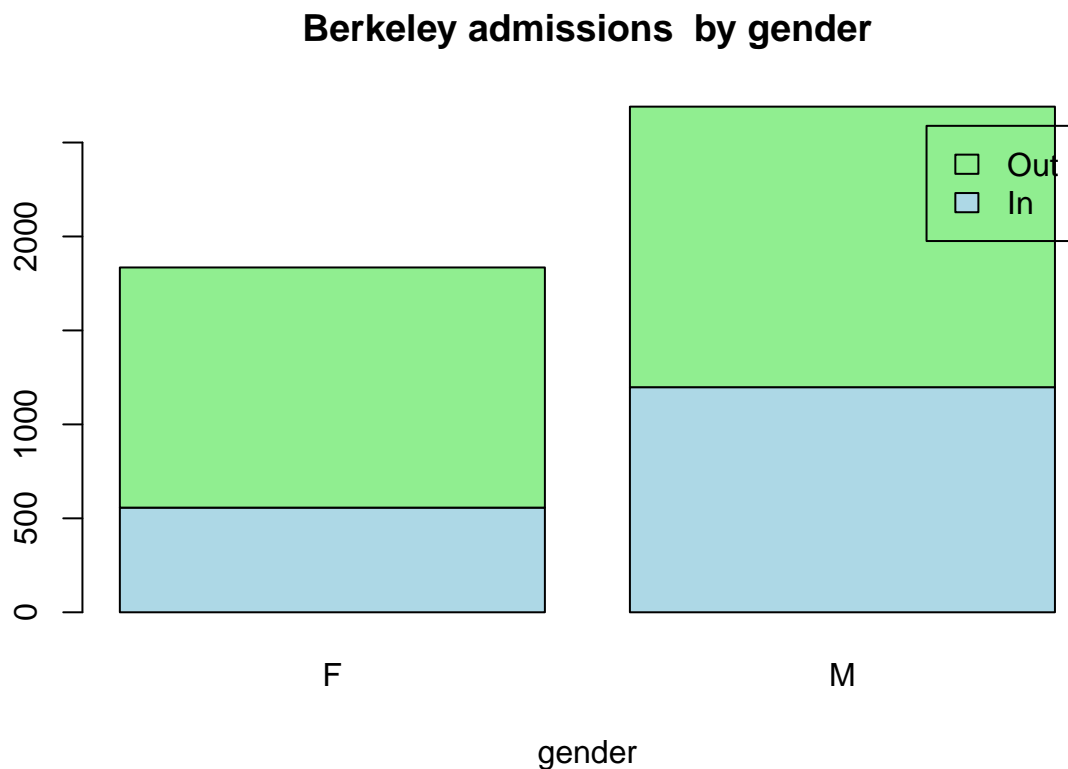

```
#Berk.gfit0 = glm(Freq ~ Gender * Outcome,family = poisson, data = Berk.df)
#The above fit is over-dispersed, so quasi fit used
Berk.quasifit0 = glm(Freq ~ Gender * Outcome,family = quasipoisson, data = Berk.df)
coef(summary(Berk.quasifit0))
```

```
##                      Estimate Std. Error    t value      Pr(>|t|)
## (Intercept)         4.5308058  0.4150710 10.9157374 7.108831e-10
## GenderM             0.7658435  0.5023801  1.5244306 1.430588e-01
## OutcomeOut          0.8304864  0.4973647  1.6697736 1.105422e-01
## GenderM:OutcomeOut -0.6103524  0.6258979 -0.9751629 3.411277e-01
```

```
exp(confint(Berk.quasifit0))[4,] #Odds ratio for females being kept out of Berkeley!!!
```

```
## Waiting for profiling to be done...
```

```
##     2.5 %    97.5 %
## 0.1512944 1.8107735
```

This appear to show an open and shut case of strong gender discrimination - the lawyers were rubbing their
hands in glee.

However, we need to be careful. In STATS 330 (which follows on from this class) we learn that an association may not be causal. Here, we have a strong association between gender and outcome, but this does **not** in itself establish a causal relationship.

In STATS 330 we would consider whether there is an indirect causal pathway. Really, we want to ask "If all things else are the same, other than gender, are the odds the same for females and males?". This means that we need to take into consideration department, and compare males and females that are applying to the same department.

Taking department in to account makes the model quite a bit more complicated. Fitting a model with 3-way interactions does not allow for meaningful interpretations, so it was decided fit the model containing all of the two-way interactions:

Model Building and Check Assumptions

```
Berk.gfit = glm(Freq ~ Dept * Gender + Dept * Outcome + Gender * Outcome,
    family = poisson, data = Berk.df)
plot(Berk.gfit, which = 1)
```
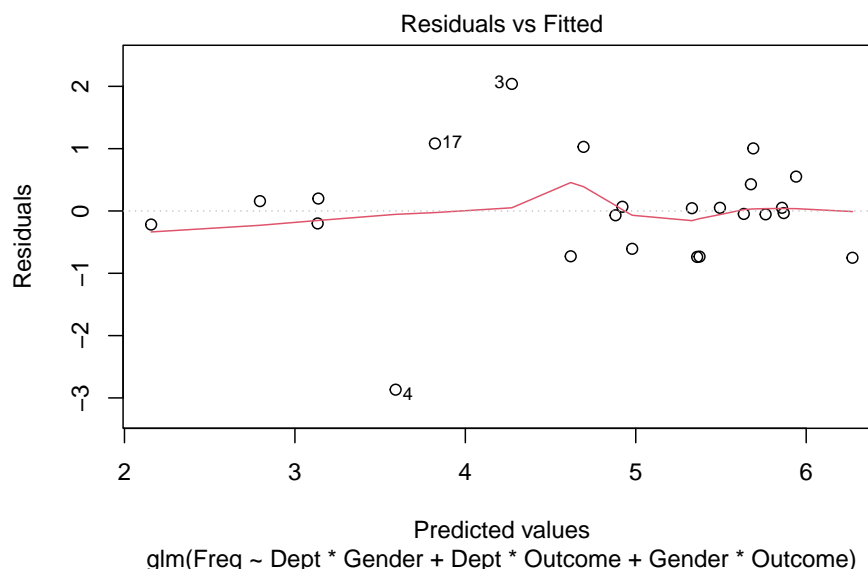


```
# The next two calls are to reduce the amount of output produced by `anova'
Berk.gfit.anova = anova(Berk.gfit, test = "Chisq")
Berk.gfit.anova[, names(Berk.gfit.anova)]
```

```
##              Df Deviance Resid. Df Resid. Dev Pr(>Chi)
## NULL                           23    2650.10
## Dept          5   159.52        18    2490.57   <2e-16 ***
## Gender        1   162.87        17    2327.70   <2e-16 ***
## Outcome       1   230.03        16    2097.67   <2e-16 ***
## Dept:Gender   5  1220.61        11     877.06   <2e-16 ***
## Dept:Outcome  5   855.32         6      21.74   <2e-16 ***
```

```
## Gender:Outcome  1     1.53          5       20.20    0.2159
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(Berk.gfit)
```

```
##
## Call:
## glm(formula = Freq ~ Dept * Gender + Dept * Outcome + Gender *
##     Outcome, family = poisson, data = Berk.df)
##
## Deviance Residuals:
##        1          2          3          4          5          6          7          8
## -0.75481    0.99471    1.96454   -3.15768   -0.03402    0.04449    0.15709   -0.22034
##        9         10         11         12         13         14         15         16
##  1.01273   -0.73839   -0.74367    0.54896    0.06760   -0.04741   -0.06911    0.05080
##       17         18         19         20         21         22         23         24
##  1.05578   -0.61236   -0.73617    0.42678   -0.20117    0.05113    0.19803   -0.05370
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)      4.27291    0.10182  41.966  < 2e-16 ***
## DeptB           -1.47804    0.22510  -6.566 5.17e-11 ***
## DeptC            1.08723    0.11610   9.365  < 2e-16 ***
## DeptD            0.60832    0.12179   4.995 5.89e-07 ***
## DeptE            0.34537    0.13179   2.621  0.00878 **
## DeptF           -1.13555    0.18196  -6.241 4.36e-10 ***
## GenderM          1.99859    0.10593  18.866  < 2e-16 ***
## OutcomeOut      -0.68192    0.09911  -6.880 5.97e-12 ***
## DeptB:GenderM    1.07482    0.22861   4.701 2.58e-06 ***
## DeptC:GenderM   -2.66513    0.12609 -21.137  < 2e-16 ***
## DeptD:GenderM   -1.95832    0.12734 -15.379  < 2e-16 ***
## DeptE:GenderM   -2.79519    0.13925 -20.073  < 2e-16 ***
## DeptF:GenderM   -2.00232    0.13571 -14.754  < 2e-16 ***
## DeptB:OutcomeOut  0.04340   0.10984   0.395  0.69277
## DeptC:OutcomeOut  1.26260   0.10663  11.841  < 2e-16 ***
## DeptD:OutcomeOut  1.29461   0.10582  12.234  < 2e-16 ***
## DeptE:OutcomeOut  1.73931   0.12611  13.792  < 2e-16 ***
## DeptF:OutcomeOut  3.30648   0.16998  19.452  < 2e-16 ***
## GenderM:OutcomeOut 0.09987  0.08085   1.235  0.21672
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 2650.095  on 23  degrees of freedom
## Residual deviance:   20.204  on  5  degrees of freedom
## AIC: 217.26
##
## Number of Fisher Scoring iterations: 4
```

```
1 - pchisq(20.204, 5)
```

```
## [1] 0.001144215
```

```
Berk.quasigfit = glm(Freq ~ Dept * Gender + Dept * Outcome + Gender * Outcome,
    family = quasipoisson, data = Berk.df)
# The next two calls are to reduce the amount of output produced by `anova'
Berk.quasigfit.anova = anova(Berk.quasigfit, test = "F")
Berk.quasigfit.anova[, names(Berk.quasigfit.anova)]
```

```
##              Df Deviance Resid. Df Resid. Dev       F    Pr(>F)
## NULL                           23    2650.10
## Dept          5   159.52        18    2490.57  8.4743 0.0174851 *
## Gender        1   162.87        17    2327.70 43.2616 0.0012192 **
## Outcome       1   230.03        16    2097.67 61.0985 0.0005495 ***
## Dept:Gender   5  1220.61        11     877.06 64.8424 0.0001519 ***
## Dept:Outcome  5   855.32         6      21.74 45.4370 0.0003612 ***
## Gender:Outcome 1    1.53         5      20.20  0.4067 0.5516975
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(Berk.quasigfit)
```

```
##
## Call:
## glm(formula = Freq ~ Dept * Gender + Dept * Outcome + Gender *
##     Outcome, family = quasipoisson, data = Berk.df)
##
## Deviance Residuals:
##        1         2         3         4         5         6         7         8
## -0.75481   0.99471   1.96454  -3.15768  -0.03402   0.04449   0.15709  -0.22034
##        9        10        11        12        13        14        15        16
##  1.01273  -0.73839  -0.74367   0.54896   0.06760  -0.04741  -0.06911   0.05080
##       17        18        19        20        21        22        23        24
##  1.05578  -0.61236  -0.73617   0.42678  -0.20117   0.05113   0.19803  -0.05370
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.27291    0.19756  21.628 3.92e-06 ***
## DeptB         -1.47804    0.43678  -3.384 0.019589 *
## DeptC          1.08723    0.22526   4.827 0.004771 **
## DeptD          0.60832    0.23631   2.574 0.049777 *
## DeptE          0.34537    0.25571   1.351 0.234719
## DeptF         -1.13555    0.35307  -3.216 0.023564 *
## GenderM        1.99859    0.20555   9.723 0.000196 ***
## OutcomeOut    -0.68192    0.19231  -3.546 0.016458 *
## DeptB:GenderM  1.07482    0.44358   2.423 0.059892 .
## DeptC:GenderM -2.66513    0.24466 -10.893 0.000113 ***
## DeptD:GenderM -1.95832    0.24707  -7.926 0.000515 ***
## DeptE:GenderM -2.79519    0.27019 -10.345 0.000145 ***
## DeptF:GenderM -2.00232    0.26333  -7.604 0.000625 ***
## DeptB:OutcomeOut 0.04340  0.21312   0.204 0.846672
## DeptC:OutcomeOut 1.26260  0.20690   6.102 0.001711 **
## DeptD:OutcomeOut 1.29461  0.20533   6.305 0.001477 **
## DeptE:OutcomeOut 1.73931  0.24470   7.108 0.000854 ***
## DeptF:OutcomeOut 3.30648  0.32982  10.025 0.000169 ***
```

```
## GenderM:OutcomeOut  0.09987      0.15687     0.637 0.552355
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasipoisson family taken to be 3.764865)
##
##      Null deviance: 2650.095  on 23  degrees of freedom
## Residual deviance:   20.204  on  5  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 4
```

Produce the Additional Output Required

```r
# Before adjusting for individual departments
chisq.test(Berk.tbl, correct = FALSE)
```

```
##
##  Pearson's Chi-squared test
##
## data:  Berk.tbl
## X-squared = 92.205, df = 1, p-value < 2.2e-16
```

```r
# After adjusting for individual departments
Berk.summ = coef(summary(Berk.quasigfit))
Berk.summ[nrow(Berk.summ), ]
```

```
##   Estimate Std. Error    t value   Pr(>|t|)
## 0.09987009 0.15686847 0.63664858 0.55235491
```

```r
c(exp(Berk.summ[nrow(Berk.summ), 1]), exp(confint(Berk.quasigfit)[nrow(Berk.summ), ]))
```

```
## Waiting for profiling to be done...
```

```
##              2.5 %    97.5 %
## 1.1050274 0.8137126 1.5058043
```

Method and Assumption Checks

Our response variable is a count broken down by three explanatory factors, gender, outcome, and department. Since we want to access the effect of gender, all else being equal, we fitted a generalized linear 3-way ANOVA model that included all possible two-way interaction terms. From looking at the data by department we stongly anticipate that some of the second-order interaction terms with department will be significant.

The residual plot shows no strong trends. The check of residual deviance had p-value 0.001 and the model was refitted as quasipoisson.

Executive Summary

We want to determine if there was gender bias at the University of Berkeley. A naive analysis of the data strongly suggests there is. However, this analysis does not compare like to like. That is, it does not compare students applying to the same department. We fitted a model that included all department in addition to gender and outcome, and found that there was no significant effect.

For completeness (despite the non-significant effect), we quantify our conclusions below:

The estimated effect is that for every female declined, 1.11 males are declined, once we adjust for the confounding effect of department. [NOTE: This is equivalent to saying that the odds of a male being declined are estimated to be 1.11 compared to females.] That is, on average about 11% more males are declined than females after adjustment for department. The confidence interval for the ratio of odds is 0.81 to 1.51.

We had evidence that the association between gender and admission outcome was significant before adjusting for individual departments ($P\text{-}value \approx 0$). However, once we adjusted for individual departments, there is insufficient evidence to say that the true ratio is different from 1 ($P\text{-}value = 0.55$).

Lecturer Comment

What is happening here is that females are tending to apply more for the departments that have lower admission rates (perhaps law and medicine). Males are tending to apply more for the departments with higher admission rates (perhaps statistics!!). The association between gender and admission outcome is genuine, but is not caused by gender, but rather the difference in departmental preferences between genders.

Plots for revealing the deparmental preferences:

Gender vs Acceptance

```
propn.gender=Berk.tbl[,1]/apply(Berk.tbl,1,sum)
barplot(propn.gender,ylim=c(0,.7), main="Berkeley admission rate by gender", col=c("lightblue", "lightg
        xlab="gender",legend.text=c("female", "male"))
```

Berkeley admission rate by gender
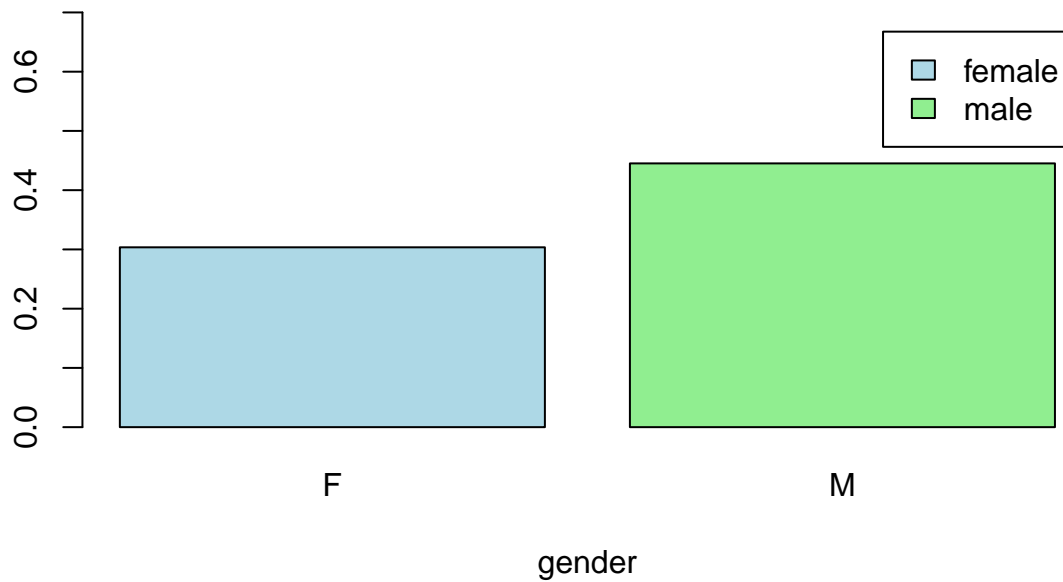
Looking at outcome by department for each gender

```
BerkByDept = xtabs(Freq ~ Gender + Outcome + Dept, data = Berk.df)

propn.gender.dept=BerkByDept[,1,]/apply(BerkByDept,c(1,3),sum)
barplot(propn.gender.dept, ylim=c(0,.85), main="Berkeley admission rate by dept and gender",
        col=c("lightblue", "lightgreen"),
        xlab="Department",legend.text=c("female", "male"),beside=T)
```

Berkeley admission rate by dept and gender
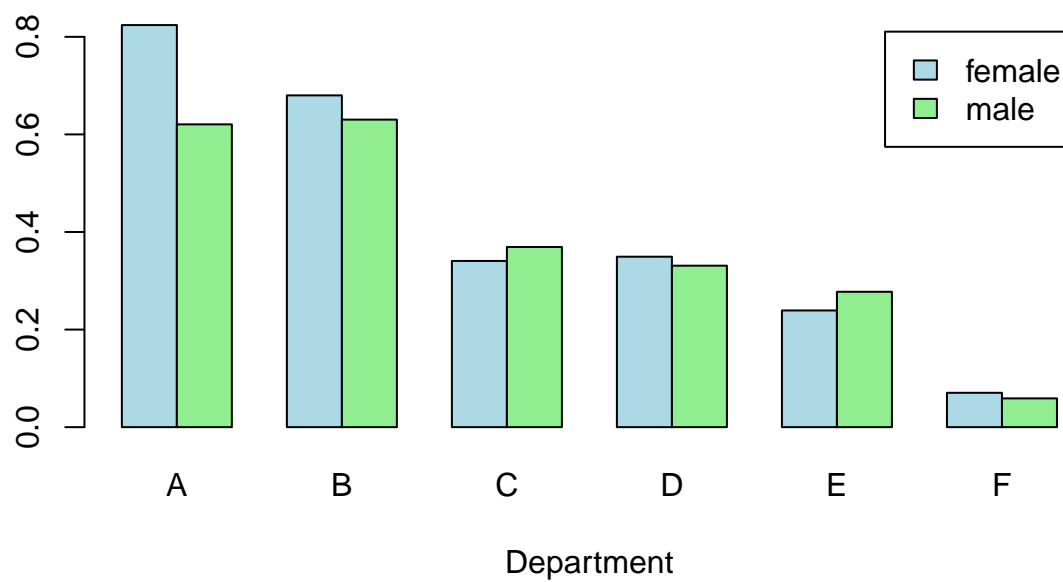
Females have higher admission rates in 4 of the 6 departments.