Apsley v. Boeing Co. and Spirit Aerosystems, Inc.722 F. Supp. 2d 1218, 2010 U.S. Dist. LEXIS 99515 (D. Kan. 2010)

Jian Sun
G43474152

Introduction

On June 16, 2005, Boeing terminated the Division's entire workforce of more than 10,000. The next day, Spirit rehired 8,354 employees, who had been selected by Boeing's managers.Although older employees predominated in the workforce both before and after the sale, a lower percentage of older workers than younger ones were rehired. The plaintiffs (the "Employees") sued, seeking to represent a class of about 700 former Boeing employees who were not hired by Spirit.

Combined with the data shown in the professor Gastwirth's paper, Some important statistical issues courts should consider in their assessment of statistical analyses submitted in class certification motions: implications for Dukes v. Wal-mart, we try to test if a lower percentage of older workers than younger ones were rehired. If so, there is age discrimination while rehiring.

The selected methods are Fisher's Test, Breslow-Day Test, Cochran–Mantel–Haenszel test.

The result is there is age discrimination for Boeing in the rehiring process. The rehiring probability of employees aged over 40 is lower than that of employees aged under 40.

Data and Method Explanation

The analyzed data is consisting of 5 variables,
unit, 1-19
firedunder, the number of people who is fired under 40 years old
keptunder, the number of people who is rehired under 40 years old
firedover, the number of people who is fired over 40 years old
keptover, the number of people who is fired over 40 years old

Fisher's Test, examining the significance of the association (contingency) between the two kinds of classification.
Breslow-Day Test, checking the assumption, homogeneous association of Apsley vs Boeing dataset, for Cochran–Mantel–Haenszel test.
Cochran–Mantel–Haenszel test, checking the association in stratified data, 19 units in Apsley vs Boeing case.

Data Analysis

The whole process is seperated into 3 parts.

Part 1

The first part is to check if there disparity for employee older than 40 years old in unit 1 while rehiring. The selected method is fisher's exact test.

In the fisher's exact test, we set
H0: the percentage of fired employees aged under 40 = the percentage of fired employees aged over 40;
Ha: the percentage of fired employees aged over 40 > the percentage of fired employees aged under 40.

By running code, odds ratio = 2.88, the P-value = 6.051e-05 < 0.05. So the null hypothesis is rejected. The percentage of fired employees aged over 40 > the percentage of fired employees aged under 40. There is age discrimination for employee aged over 40 in unit 1.

Part 2

The second part is proving that there is homogeneous association of Apsley vs Boeing dataset before doing Cochran–Mantel–Haenszel test. The selected method is Breslow-Day Test.

In the Breslow-Day Test, we set
H0: all 19 units have similar Odds Ratio;
Ha: at least one Odds Ratio out of 19 units is different with others.

By running code, we get the p-value = 0.2232 > 0.05, so the null hypothesis is kept. There is not sufficient evidence to reject the model of homogeneous associations.
Thus the conditional Odds Ratios of the rest 18 units, each with a different director, are similar to unit 1, the assumption of Cochran–Mantel–Haenszel test is met.

Part3

The third part is to check if 19 units are independent with each other by doing Cochran–Mantel–Haenszel test.

In the Cochran–Mantel–Haenszel test, we set
H0: The age discrimination problem on all 19 units are associated;
Ha: The age discrimination problem on all 19 units are independent, the result in unit1 will not affect the result of the rest 18 unit.

Based on the R result, the p-value is 1.346e-06 < 0.05, so the null hypothesis is rejected. And we think the 19 units, each with a different director, is independent with each otehr. Then input the selected odds ratio, 1.784, and alpha level, 0.05, into CMH.power() function. After running the code, we get the power of CMH test, 0.999. Hence, the Odds Ratio of mantal haenszel test is statistical significant.

When alpha level = 0.00015, the power.cmh=0.9. Hence, the Odds Ratio of mantal haenszel test is still statistical significant.

Thus the situation for 19 director are similar to unit 1. It is mean that after selling BCA Wichita division, there is age discrimination while rehiring employee. In Boeing, Older workers (age over 40) have lower percentage to be rehired than younger ones (age under 40). Older workers (age over 40) have higher percentage to be fired than younger ones (age under 40).

Conclusion

The above study result shows that a lower percentage of older workers than younger ones were rehired. A higher percentage of older workers than younger ones were fired.

The employees' statement is also proved by above analysis. And the statement is that Boeing, Onex, and Spirit (the "Companies") violated the Age Discrimination in Employment Act ("ADEA"), the Employee Retirement Income Security Act ("ERISA"), Title VII of the Civil Rights Act of 1964 ("Title VII"), and the Americans with Disabilities Act ("ADA").

Appendix

```
wants <- c('lawstat','BiasedUrn','Exact')
has <- wants %in% rownames(installed.packages())
if(any(!has)) install.packages(wants[!has])
library(Exact)
```

Load function 1

```
######Breslow and Day-test
breslowday.test <- function(x){
  or.hat.mh <- mantelhaen.test(x)$estimate
  K <- dim(x)[3]
  X2.HBD <- 0
  a <- tildea <- Var.a <- numeric(K)

  for (j in 1:K){
    mj <- apply(x[,,j],MARGIN = 1,sum)
    nj <- apply(x[,,j],MARGIN = 2,sum)

    coef <- c(-mj[1]*nj[1]*or.hat.mh,
              nj[2]-mj[1]+or.hat.mh*(nj[1]+mj[1]),
              1-or.hat.mh)
    sols <- Re(polyroot(coef))
    tildeaj <- sols[(0<sols) & (sols<=min(nj[1],mj[1]))]
    aj <- x[1,1,j]

    tildebj <- mj[1]-tildeaj
    tildecj <- nj[1]-tildeaj
    tildedj <- mj[2]-tildecj

    Var.aj <- (1/tildeaj+1/tildebj+1/tildecj+1/tildedj)^(-1)
    X2.HBD <- X2.HBD+as.numeric((aj-tildeaj)^2/Var.aj)

    a[j] <- aj
    tildea[j] <- tildeaj
    Var.a[j] <- Var.aj
  }

  X2.HBDT <- as.numeric(X2.HBD-(sum(a)-sum(tildea)^2)/sum(Var.aj))
  p <- 1-pchisq(X2.HBDT,df=K-1)

  res <- list(X2.HBD=X2.HBD,X2.HBDT=X2.HBDT,p=p)
  class(res) <- 'bdtest'
  return(res)
}

print.bdtest <- function(x){
  cat('Breslow and Day test (with Tarone correction):\n')
  cat('Breslow-Day X-squared =',x$X2.HBD,'\n')
  cat('Breslow-Day-Tarone X-squared =',x$X2.HBDT,'\n\n')
  cat('Test for a common OR: p-value = ',x$p,'\n\n')
}
```

Load function 2

```
##########CMH-power test###########
CMH.power <- function(data,odds,alpha,Alternative){
  s <- dim(data)[3]
  m <- n <- k <- rep(0,s)
  #m <- rep(0,s)
  #n <- rep(0,s)
  #k <- rep(0,s)
  for (i in 1:s){
    m[i] <- sum(data[1,,i])
    n[i] <- sum(data[2,,i])
    k[i] <- sum(data[,1,i])
  }
  library(BiasedUrn)
  library(lawstat)
  sim <- 10000
  pvalue.cmh <- rep(0,sim)
  pvalue.cmh.exact <- rep(0,sim)
  for (i in 1:sim){
    min <- rep(0,s)
    for (j in 1:s){
      min[j] <- rFNCHypergeo(1,m[j],n[j],k[j],odds)
    }
    maj <- k-min
    data <- NULL
    for (j in 1:s){
      data <- c(data,min[j],maj[j],m[j]-min[j],n[j]-maj[j])
    }
    sim.data <- array(data,dim=c(2,2,s))
    pvalue.cmh[i] <- mantelhaen.test(sim.data,alternative=Alternative)$p.value
    pvalue.cmh.exact[i] <- mantelhaen.test(sim.data,exact = T,alternative = Alternative)$p.value
  }
  power.cmh <- mean(pvalue.cmh<=alpha)
  power.cmh.exact <- mean(pvalue.cmh.exact<=alpha)
  return(list(power.cmh=power.cmh,power.cmh.exact=power.cmh.exact))
}
```

Data analysis process
The dataset is loaded by running the following codes.

```
WD=read.csv("/Users/LoveChina/Documents/6253/ApsleyStrati.csv",
            header = TRUE)
head(WD)
```

```
##   unit firedover firedunder keptover keptunder
## 1    1       161         13      786       183
## 2    2        22          1      617       233
## 3    3        44          0      265        34
## 4    4        15          0      106        21
## 5    5       145          7      714        63
## 6    6        26          4      214        63
```

```
summary(WD)
```

```
##       unit        firedover        firedunder        keptover
```

```
##  Min.   : 1.0   Min.   :  1.00   Min.   : 0.000   Min.   : 29.0
##  1st Qu.: 5.5   1st Qu.: 11.50   1st Qu.: 0.000   1st Qu.:102.0
##  Median :10.0   Median : 44.00   Median : 1.000   Median :288.0
##  Mean   :10.0   Mean   : 58.05   Mean   : 4.421   Mean   :376.3
##  3rd Qu.:14.5   3rd Qu.: 86.00   3rd Qu.: 7.000   3rd Qu.:623.0
##  Max.   :19.0   Max.   :192.00   Max.   :30.000   Max.   :949.0
##    keptunder
##  Min.   :  2.00
##  1st Qu.: 13.50
##  Median : 32.00
##  Mean   : 58.58
##  3rd Qu.: 63.00
##  Max.   :233.00
```

```
Wd=WD[,-1]
Wd=as.matrix(Wd)
wd=array(c(Wd[1,],Wd[2,],Wd[3,],Wd[4,],Wd[5,],Wd[6,],Wd[7,],Wd[8,],
          Wd[9,],Wd[10,],Wd[11,],Wd[12,],Wd[13,],Wd[14,],Wd[15,],Wd[16,],
          Wd[17,],Wd[18,],Wd[19,]), dim = c(2,2,19))
rownames(wd)=c("over40","under40")
colnames(wd)=c("fired","kept")
wd
```

```
## , , 1
##
##         fired kept
## over40    161  786
## under40    13  183
##
## , , 2
##
##         fired kept
## over40     22  617
## under40     1  233
##
## , , 3
##
##         fired kept
## over40     44  265
## under40     0   34
##
## , , 4
##
##         fired kept
## over40     15  106
## under40     0   21
##
## , , 5
##
##         fired kept
## over40    145  714
## under40     7   63
##
## , , 6
##
```

```
##         fired kept
## over40     26  214
## under40     4   63
##
## , , 7
##
##         fired kept
## over40     65  949
## under40     8  172
##
## , , 8
##
##         fired kept
## over40    124  629
## under40     7   52
##
## , , 9
##
##         fired kept
## over40     69  388
## under40     1   16
##
## , , 10
##
##         fired kept
## over40    103  363
## under40     7   35
##
## , , 11
##
##         fired kept
## over40     45  288
## under40     1   16
##
## , , 12
##
##         fired kept
## over40     20  257
## under40     1   32
##
## , , 13
##
##         fired kept
## over40    192  891
## under40    30  148
##
## , , 14
##
##         fired kept
## over40      3   30
## under40     0    6
##
## , , 15
##
```

```
##          fired kept
## over40       8   98
## under40      1   15
##
## , , 16
##
##          fired kept
## over40       3   29
## under40      0    2
##
## , , 17
##
##          fired kept
## over40       5   58
## under40      0    2
##
## , , 18
##
##          fired kept
## over40       1   49
## under40      0    8
##
## , , 19
##
##          fired kept
## over40      52  418
## under40      3   12
```

Doing Fisher's Exact Test.

```r
A=matrix(Wd[1,],nrow=2,byrow = FALSE)
rownames(A)=c("over40","under40")
colnames(A)=c("fired","kept")
A
```

```
##          fired kept
## over40     161  786
## under40     13  183
```

```r
B=fisher.test(A, alternative = "greater")
B
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  A
## p-value = 6.051e-05
## alternative hypothesis: true odds ratio is greater than 1
## 95 percent confidence interval:
##  1.7283    Inf
## sample estimates:
## odds ratio
##   2.881266
```

Doing Breslow-Day Test.

```r
breslowday.test(wd)
```

```
## Breslow and Day test (with Tarone correction):
## Breslow-Day X-squared = 22.19951
## Breslow-Day-Tarone X-squared = 1307157
##
## Test for a common OR: p-value =  0
```

```r
library(DescTools)
BreslowDayTest(wd)
```

```
##
##  Breslow-Day test on Homogeneity of Odds Ratios
##
## data:  wd
## X-squared = 22.2, df = 18, p-value = 0.2232
```

Doing Cochran–Mantel–Haenszel test.

```r
mantelhaen.test(wd)
```

```
##
##  Mantel-Haenszel chi-squared test with continuity correction
##
## data:  wd
## Mantel-Haenszel X-squared = 23.357, df = 1, p-value = 1.346e-06
## alternative hypothesis: true common odds ratio is not equal to 1
## 95 percent confidence interval:
##  1.409710 2.257339
## sample estimates:
## common odds ratio
##           1.78387
```

```r
CMH.power(wd,1.78,0.05,'two.sided')
```

```
## Loading required package: Hmisc

## Loading required package: lattice

## Loading required package: survival

## Loading required package: Formula

## Loading required package: ggplot2

##
## Attaching package: 'Hmisc'

## The following objects are masked from 'package:DescTools':
##
##     %nin%, Label, Mean, Quantile

## The following objects are masked from 'package:base':
##
##     format.pval, round.POSIXt, trunc.POSIXt, units

## Loading required package: Kendall

## Loading required package: mvtnorm

## Loading required package: VGAM
```

```
## Loading required package: stats4
```

```
## Loading required package: splines
```

```
## $power.cmh
## [1] 0.9994
##
## $power.cmh.exact
## [1] 0.9994
```

```r
CMH.power(wd,1.78,0.00015,'two.sided')
```

```
## $power.cmh
## [1] 0.8923
##
## $power.cmh.exact
## [1] 0.9129
```