

# ANOVA\_\_METOPROLOL

*Zongyan Wang*

*Monday, December 07, 2015*

## Load package

```
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 3.1.3
```

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.1.3
```

```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:data.table':
##
##   between, last
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(agricolae)
```

```
## Warning: package 'agricolae' was built under R version 3.1.3
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

## data loading and clean

```
phy_drugs = fread("F:/Academic/Stat 992/group project/toy_data/wi_cardi2.tab", sep="\t", header=FALSE)
hospitals = unique(phy_drugs[, 20])
setkey(phy_drugs, V1)
#phy_drugs$zip = DT[as.character(phy_drugs$V1), mult = "first"]$`Zip Code`
if(T){
  phy_drugs = phy_drugs[!V9=="METOPROLOL SUCCINATE"]
}
```

```

}

#First aggregate ratios over physicians
grouped_by_physician = phy_drugs %>%
  group_by(V1) %>%
  dplyr::select(drug_name = V8, generic_name = V9, total_claim_cnt = V11)

#Calculates ratios based on number of brand name vs total claims
phy_bg_ratios = dplyr::summarise(grouped_by_physician,
                                bg_ratio = (sum(as.vector(total_claim_cnt[as.vector(drug_name) != as.vector(g
)

phy_bg_ratios$hospital <- phy_drugs[.(phy_bg_ratios$V1), mult = "first"]$V20

```

## ANOVA

```

data <- phy_bg_ratios
data <- cbind(data, trt=factor(data$hospital, labels=1:length(unique(data$hospital))))
model <- aov(bg_ratio ~ trt, data = data)
summary(model)

```

```

##              Df Sum Sq Mean Sq F value Pr(>F)
## trt          60 0.5210 0.008684   2.705 7.5e-08 ***
## Residuals    216 0.6935 0.003211
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The p-value is  $7.5 \times 10^{-8}$ , it shows there are significant different among hospital.

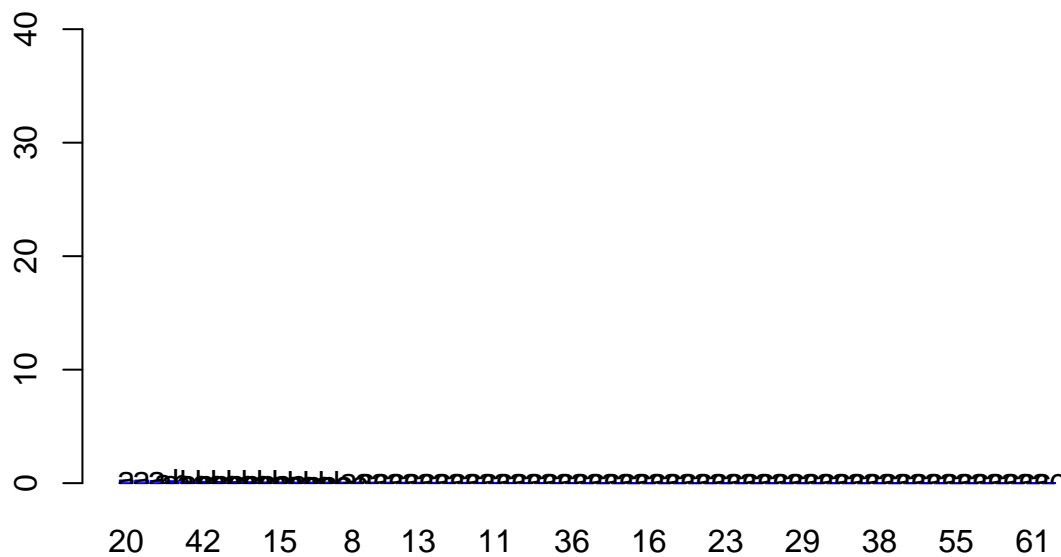
## Doing LSD and HSD test

```

lsd.out <- LSD.test(model, "trt", p.adj = "bonferroni")
hsd.out <- HSD.test(model, "trt", group = T)

bar.group(lsd.out$groups, ylim=c(0,45), density=4, border="blue")
bar.group(hsd.out$groups, ylim=c(0,45), density=4, border="blue")

```



```
#lsd.out$groups
#hsd.out$groups
```

You can match the group number with hospital with the data we use

```
id.group <- data %>%
  dplyr::select(hospital, trt) %>%
  unique()

hsd.out$groups.id <- merge(id.group, hsd.out$groups, by = "trt")
hsd.out$groups.id
```

```
##      trt
## 1:  10
## 2:  11
## 3:  12
## 4:  13
## 5:  14
## 6:  15
## 7:  16
## 8:  17
## 9:  18
##10:  19
##11:  20
```

## 12: 21  
 ## 13: 22  
 ## 14: 23  
 ## 15: 24  
 ## 16: 25  
 ## 17: 26  
 ## 18: 27  
 ## 19: 28  
 ## 20: 29  
 ## 21: 30  
 ## 22: 31  
 ## 23: 32  
 ## 24: 33  
 ## 25: 34  
 ## 26: 35  
 ## 27: 36  
 ## 28: 37  
 ## 29: 38  
 ## 30: 39  
 ## 31: 40  
 ## 32: 41  
 ## 33: 42  
 ## 34: 43  
 ## 35: 44  
 ## 36: 45  
 ## 37: 46  
 ## 38: 47  
 ## 39: 48  
 ## 40: 49  
 ## 41: 50  
 ## 42: 51  
 ## 43: 52  
 ## 44: 53  
 ## 45: 54  
 ## 46: 55  
 ## 47: 56  
 ## 48: 57  
 ## 49: 58  
 ## 50: 59  
 ## 51: 60  
 ## 52: 61  
 ## trt

##	hospital
## 1:	BAYCARE AURORA, LLC
## 2:	BELLIN MEMORIAL HOSPITAL, INC
## 3:	BELOIT HEALTH SYSTEM INC
## 4:	CARDIOLOGY ASSOCIATES OF WAUKESHA, S.C.
## 5:	COLUMBIA ST MARY'S HOSPITAL OZAUKEE, INC.
## 6:	COLUMBIA ST. MARY'S HOSPITAL MILWAUKEE, INC
## 7:	CULLINANE CARDIOVASCULAR CONSULTANTS, SC
## 8:	DEAN HEALTH SYSTEMS, INC
## 9:	EAU CLAIRE HEART INSTITUTE, SC
## 10:	EAU CLAIRE MEDICAL CLINIC SC
## 11:	EDDY D. CO, MD, SC

```

## 12: FARZAD KAMRANI,MD.SC.
## 13: FROEDTERT AND THE MEDICAL COLLEGE OF WISCONSIN COMMUNITY PHYSICIANS IN
## 14: GUNDERSEN CLINIC LTD
## 15: GUNDERSEN LUTHERAN MEDICAL CENTER, INC.
## 16: HOLY FAMILY MEMORIAL INC
## 17: HUDSON PHYSICIANS SC
## 18: JOAN T. HARNEY GNADT, MD, SC
## 19: LAKE COUNTRY CARDIOVASCULAR ASSOCIATES LTD
## 20: LAKESHORE MEDICAL CLINIC LTD
## 21: LOUIE COULIS MD SC
## 22: MARSHFIELD CLINIC INC
## 23: MAYO CLINIC HEALTH SYSTEM-RED CEDAR, INC
## 24: MAYO CLINIC HEALTH SYSTEM - EAU CLAIRE CLINIC, INC
## 25: MAYO CLINIC HEALTH SYSTEM - FRANCISCAN MEDICAL CENTER, INC.
## 26: MERCY HEALTH SYSTEM CORPORATION
## 27: MERITER HOSPITAL, INC.
## 28: MERITER MEDICAL GROUP INC
## 29: MILE BLUFF MEDICAL CENTER INC
## 30: MILWAUKEE CARDIOVASCULAR CNTR S.C.
## 31: NETWORK HEALTH SYSTEM, INC.
## 32: OAKLEAF CLINICS SC
## 33: PRACTICE MANAGEMENT GROUP INC
## 34: PROHEALTH CARE MEDICAL ASSOCIATES INC
## 35: SIXTEENTH STREET COMMUNITY HEALTH CENTERS INC
## 36: ST VINCENT HOSPITAL-HOSPITAL SISTERS-THIRD ORDER OF ST FRANCIS
## 37: ST. JOSEPH'S ELECTROCARDIOGRAPHIC ASSOCIATES
## 38: ST. NICHOLAS HOSPITAL-SISTERS OF THE THIRD ORDER OF ST. FRANCIS
## 39: SUHAS K SHELGIKAR MD SC
## 40: THE MEDICAL COLLEGE OF WISCONSIN INC
## 41: THE MONROE CLINIC, INC.
## 42: THEDACARE, INCORPORATED
## 43: UNITED HOSPITAL SYSTEM INC
## 44: UNIVERSITY OF WISCONSIN MEDICAL FOUNDATION INC
## 45: WATERTOWN HEART INSTITUTE SC
## 46: WATERTOWN REGIONAL MEDICAL CENTER, INC
## 47: WAUKESHA HEART INSTITUTE
## 48: WESTFIELDS HOSPITAL, INC.
## 49: WHEATON FRANCISCAN MEDICAL GROUP
## 50: WISCONSIN HEALTH FUND
## 51: WISCONSIN HEART CENTER S C
## 52: ZENITH HEALTHCARE SC
## hospital
##
## means M
## 1: 0.000000000 ac
## 2: 0.023809524 ac
## 3: 0.052272727 abc
## 4: 0.032323040 ac
## 5: 0.000000000 ac
## 6: 0.058406175 abc
## 7: 0.000000000 ac
## 8: 0.013143483 ac
## 9: 0.000000000 ac
## 10: 0.000000000 ac
## 11: 0.236641221 a

```

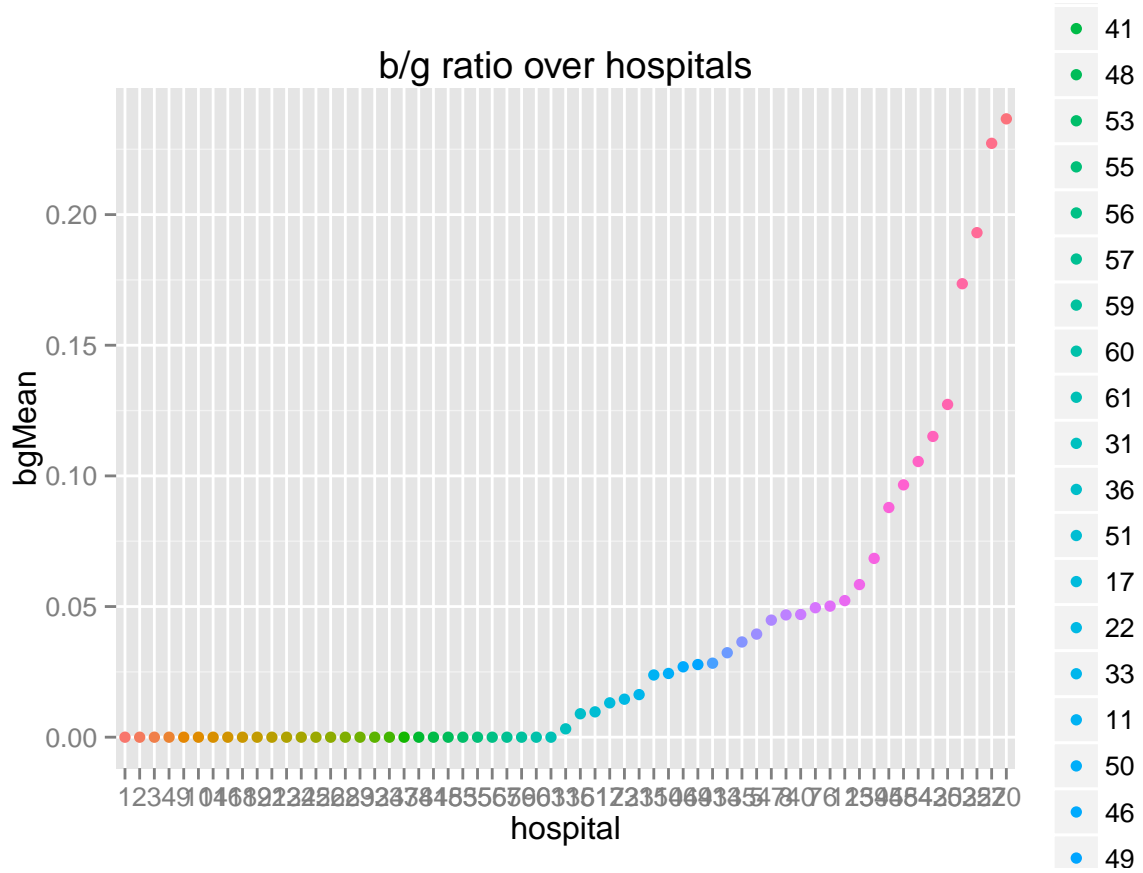
```
## 12: 0.000000000 ac
## 13: 0.014541387 ac
## 14: 0.000000000 ac
## 15: 0.000000000 ac
## 16: 0.000000000 ac
## 17: 0.000000000 ac
## 18: 0.227272727 a
## 19: 0.000000000 ac
## 20: 0.000000000 ac
## 21: 0.127340824 abc
## 22: 0.003196347 ac
## 23: 0.000000000 ac
## 24: 0.016300177 ac
## 25: 0.000000000 ac
## 26: 0.193082340 a
## 27: 0.008970984 ac
## 28: 0.000000000 ac
## 29: 0.000000000 ac
## 30: 0.068421053 abc
## 31: 0.046974190 abc
## 32: 0.000000000 ac
## 33: 0.115107914 abc
## 34: 0.028322440 ac
## 35: 0.087893864 abc
## 36: 0.036438385 ac
## 37: 0.026936027 ac
## 38: 0.044802328 ac
## 39: 0.000000000 ac
## 40: 0.027799108 ac
## 41: 0.024390244 ac
## 42: 0.009696970 ac
## 43: 0.173538512 ab
## 44: 0.000000000 ac
## 45: 0.105504587 abc
## 46: 0.000000000 ac
## 47: 0.000000000 ac
## 48: 0.000000000 ac
## 49: 0.096568840 abc
## 50: 0.000000000 ac
## 51: 0.000000000 ac
## 52: 0.000000000 ac
##          means  M
```

There are three hospitals stand out, which are EDDY D. CO, MD, SC; JOAN T. HARNEY GNADT, MD, SC; MERCY HEALTH SYSTEM CORPORATION.

## plot

```
bgmean <- data %>%
  group_by(trt) %>%
  summarize(bgMean = mean(bg_ratio)) %>%
  mutate(hospital = reorder(x = trt, X = bgMean, min))
```

```
ggplot(bgmean, aes(x = hospital, y = bgMean)) +  
  geom_point(aes(color = hospital)) +  
  ggtitle("b/g ratio over hospitals")
```



```
# geom_text(aes(x = hospital, y = bgMean, label = hospital), size = 2.5)

ggsave(filename = "wi_cardio_METO.png", plot = last_plot())
```

```
## Saving 6.5 x 4.5 in image
```