

# Two way ANOVA and ANCOVA

Hossein Vatani

April 13, 2016

## بنام خدا

## تحلیل پراکنش و هموردایی دو طرفه

در این آموزش ما در مورد برازش دوطرفه تحلیل پراکنش (fitting of two-way analysis of variance) و همچنین تحلیل هموردایی نمونه (analysis of covariance) صحبت خواهیم کرد. چون در این بخش از رابطه نمونه پراکنش استفاده شده است، موارد آموزشی مطرح شده در این بخش، برای نمونه پسرفتهای چندگانه نیز قابل استفاده است. بهرحال، در این بخش می توانیم متغیرهای تشریحی را کیفی و یا دسته ای (categorical) در نظر بگیریم.

## الف-تحلیل پراکنش دوطرفه

ANOVA دوطرفه برای مقایسه میانگین جامعه ای که از دوطریق متفاوت رده بندی (Calsified) شده است بکار می رود و یا میانگین متغیر پاسخ در یک تجربه حاصل از متغیری دو عاملی (two factors). جهت برازش مدل پراکنش دوعاملی، می توان از همان دستور برازش نمونه پسرفت استفاده کرد. همچنین با کمک رابطه دوم، می توان اثر مشترک دو عامل را نیز محاسبه کرد

```
Fit=lm(Response ~ FactorA + FactorB)
Fit=lm(Response ~ FactorA + FactorB + FactorA*FactorB)
anova(Fit)
```

که در اینجا هر دو متغیر FactorA,FactorB متغیرهای قیاسی (Categorical) هستند و متغیر پاسخ متغیری کمی (Quantitative).

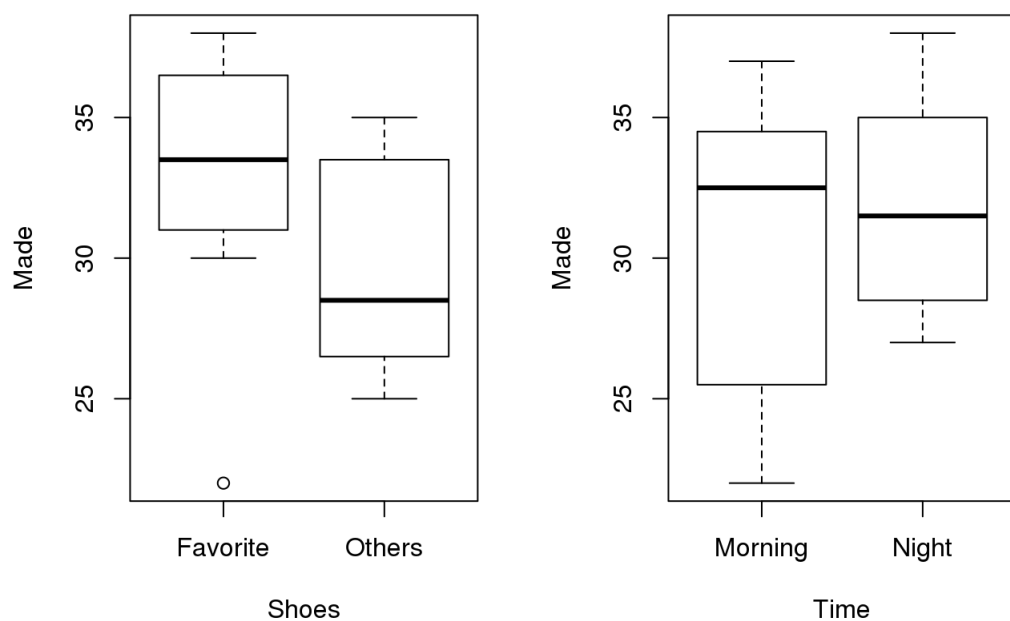
مثال-(فایل موجود نیست)

```
Time=c("Morning","Morning","Night","Night","Morning","Morning","Night","Night","Morning","Morning","Night","Night","Morning","Morning","Night","Night")
Shoes=c("Others","Others","Others","Others","Favorite","Favorite","Favorite","Favorite","Others","Others","Others","Others","Favorite","Favorite","Favorite","Favorite")
Made=c(25,26,27,27,32,22,30,34,35,34,33,30,33,37,36,38)
BaskBall=data.frame(Made,Shoes,Time)
BaskBall
```

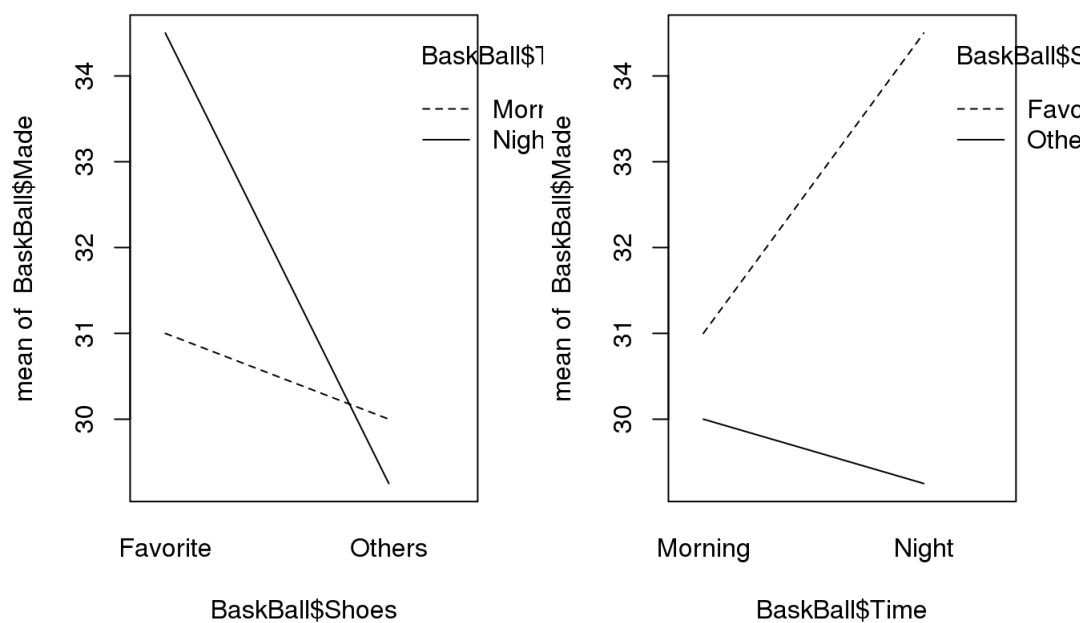
```
##      Made   Shoes   Time
## 1     25    Others Morning
## 2     26    Others Morning
## 3     27    Others   Night
## 4     27    Others   Night
## 5     32 Favorite Morning
## 6     22 Favorite Morning
## 7     30 Favorite   Night
## 8     34 Favorite   Night
## 9     35    Others Morning
## 10    34    Others Morning
## 11    33    Others   Night
## 12    30    Others   Night
## 13    33 Favorite Morning
## 14    37 Favorite Morning
## 15    36 Favorite   Night
## 16    38 Favorite   Night
```

ابتدا نمایش تصویری داده های فوق بهراه نمایش اثر مشترک آنها بر روی هم بررسی می کنیم.

```
par(mfrow=c(1,2))
plot(Made~Shoes+Time,data = BaskBall)
```



```
interaction.plot(BaskBall$Shoes,BaskBall$Time,BaskBall$Made)
interaction.plot(BaskBall$Time,BaskBall$Shoes,BaskBall$Made)
```



درمورد نمودارهای اثر مشترک، به این نکته توجه شود که اگر خطهای نمایش داده شده با اغماض بایکدیگر موازی باشند آنگاه می توان گفت که اثر مشترک نامحسوس است. بهر حال باید ابتدا نمونه را بسازیم تا بتوان با کمک آن بهتر تصمیم گرفت.

```
Fit=lm(Made~Shoes+Time+Shoes*Time,data = BaskBall)
anova(Fit)
```

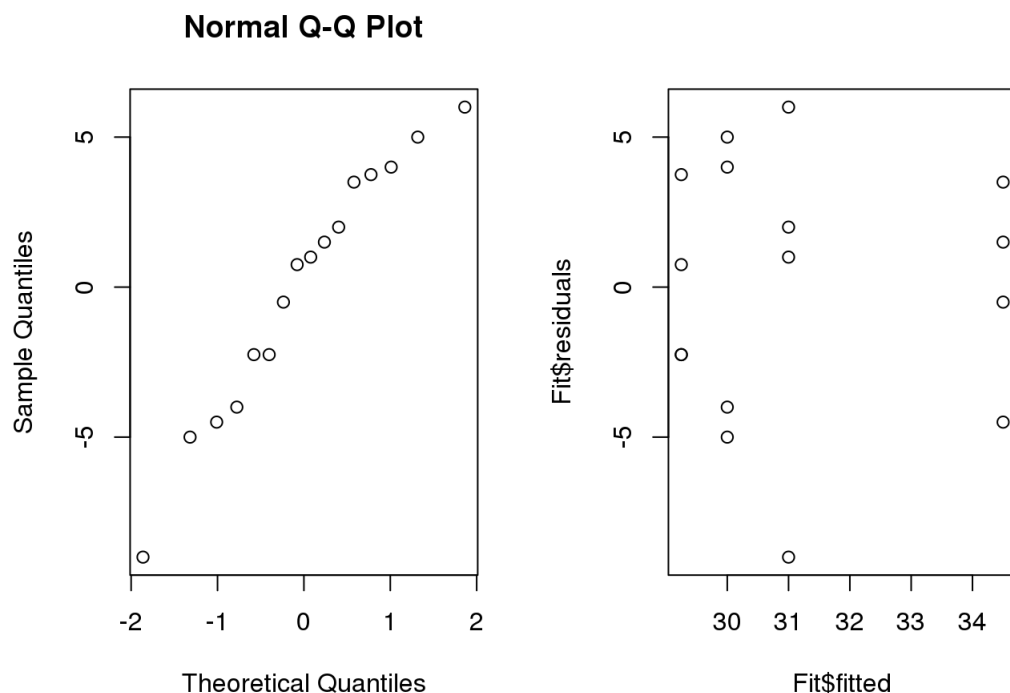
```
## Analysis of Variance Table
##
## Response: Made
##           Df Sum Sq Mean Sq F value Pr(>F)
## Shoes      1  39.062   39.062   1.7773 0.2072
## Time       1   7.563    7.563   0.3441 0.5684
## Shoes:Time  1  18.062   18.062   0.8218 0.3825
## Residuals 12 263.750   21.979
```

خوب، خدایوشکر نتیجه آزمایش به ما می گوید که هیچکدام برهم هیچ اثری ندارند:-). جهت اطلاعات بیشتر نمودار کلی آنرا نیز رسم میکنم تا مطلب کاملاً روشن شود. همینطور با توجه به اینکه دیگر نیازی به نمودارهای باقیمانده ها نیست ولی آنها را نیز رسم می کنیم.

```
plot(BaskBall)
```



```
par(mfrow=c(1,2))
qqnorm(Fit$residuals)
plot(Fit$fitted,Fit$residuals)
```



نمودارهای باقیمانده مبین صحیح بودن نمونه ایجاد شده می باشد.

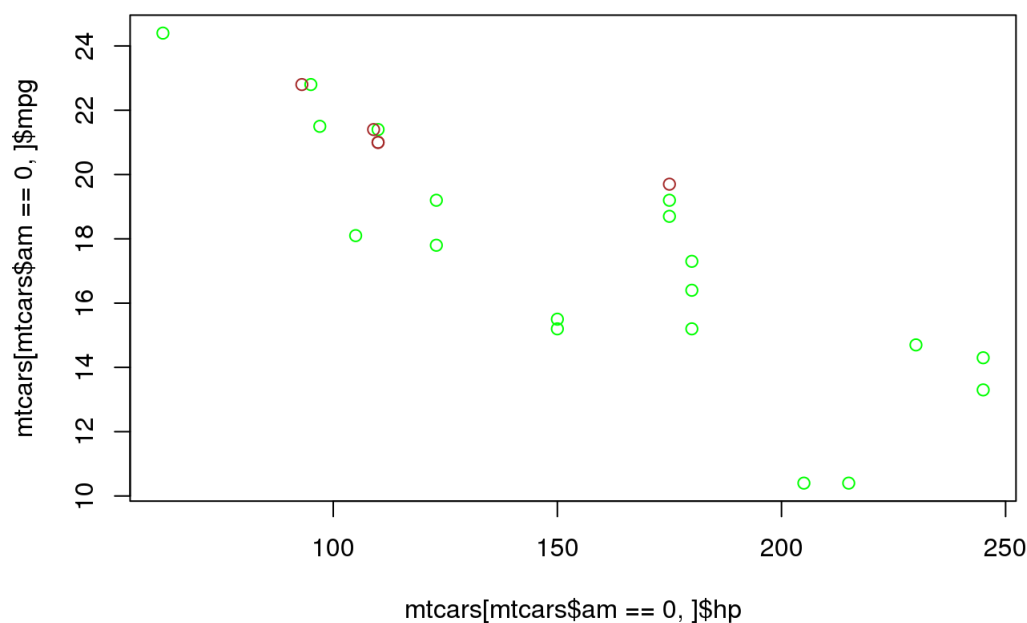
## ب-تحلیل هموردایی (Covariance)

تحلیل هموردایی، ترکیبی از خصوصیات هر دو تحلیل پراکنش (ANOVA) و پسرفت (Regression) می باشد. این تحلیل، چند متغیر کمی که به آنها هموردا گفته می شود را به تحلیل پراکنش (ANOVA) اضافه می کند. اینان مربوط به اثر مشترک متغیرها می باشند. این متغیرهای اضافی جهت کاهش اثر جملات خطا در پراکنش (Varaince) و همچنین ارایه ابزاری دقیقتر جهت بررسی رفتار متغیرها می باشد. این نوع تحلیل بیشتر در تحلیل اثر عاملها (Factors) و همچنین تحلیل اثر مشترک آنها کاربرد دارد. زمانی که کنترل هموردایی (covariate) مورد نظر باشد، تحلیل ANCOVA مورد استفاده قرار می گیرد. این تحلیل هم اثر عامل (Factors) و هم اثر متقابل آنها (interaction effects of the factors) را نشان می دهد.

نکته: برای این مورد همان دستور aov استفاده می شود و فقط در نمونه (Model) ما جمله حاصل ضرب عاملهای مورد نظر را نیز اضافه می کنیم.

مثال- (فایل در دسترس نیست)- می خواهیم بدانیم در داده های خودروی موجود در نرم افزار) آیا ارتباطی بین عائل های hp و am وجود دارد یاخیر. ابتدا نمودار داده را رسم می کنیم تا بطور شهودی بتوانیم بررسی کنیم که آیا ایندو عامل تاثیری برهم دارند یاخیر

```
mtcars=as.data.frame(mtcars)
mtcars$am=as.factor(mtcars$am)#عنوان فاکتور معرفی کنیم.
plot(mtcars[mtcars$am==0,]$hp,mtcars[mtcars$am==0,]$mpg,col="green")
points(mtcars[mtcars$am==1,]$hp,mtcars[mtcars$am==1,]$mpg,col="brown")
```



باتوجه به آنچه که حاصل شده است، بطور شهودی می توان گفت که ایندو فاقد اثر برهم هستند.آما بررسی روابط

```
Fit1=lm(mpg~hp*am,data=mtcars) # می توان اینگونه هم نوشت lm(mpg~hp+am+hp*am,data=mtcars)
anova(Fit1)
```

```
## Analysis of Variance Table
##
## Response: mpg
##          Df Sum Sq Mean Sq F value    Pr(>F)
## hp         1  678.37   678.37  77.3912 1.505e-09 ***
## am         1  202.24   202.24  23.0717 4.749e-05 ***
## hp:am      1    0.01     0.01  0.0006  0.9806
## Residuals 28  245.43     8.77
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

مقدار P-Value در نتیجه آزمایش بوضوح می گوید که ایندو متغیر نسبت به یکدیگر تاثیری بر عامل مورد نظر ندارند.

می توان جهت اطمینان بیشتر مراحل زیر را نیز انجام داد.

```
Fit2=lm(mpg~hp+am,data=mtcars)
anova(Fit1,Fit2)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ hp * am
## Model 2: mpg ~ hp + am
##   Res.Df  RSS Df Sum of Sq    F Pr(>F)
## 1      28 245.43
## 2      29 245.44 -1 -0.0052515 6e-04 0.9806
```

مقدار P-Value حاصله بوضوح می گوید که دو نمونه ایجاد شده نسبت بهم اختلاف معنی داری ندارند.