Noncentrality Parameter

- If H_0 is true, i.e., if $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}\boldsymbol{\beta} \in \mathcal{C}(\boldsymbol{X}_0)$, then the noncentrality parameter θ is 0 because

end lecture 4
01-30-25

$$\begin{aligned}(\boldsymbol{P_X} - \boldsymbol{P_{X_0}})\boldsymbol{X}\boldsymbol{\beta} &= \boldsymbol{P_X}\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{P_{X_0}}\boldsymbol{X}\boldsymbol{\beta} \\ &= \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{0}.\end{aligned}$$

  Hence,

$$\boldsymbol{y}^\top(\boldsymbol{I} - \boldsymbol{P_X})\boldsymbol{y}/\sigma^2 \sim \chi^2_{\mathrm{rank}(\boldsymbol{X}) - \mathrm{rank}(\boldsymbol{X}_0)},$$

  a central χ^2 distr.

- If H_0 is false and $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}\boldsymbol{\beta} \notin \mathcal{C}(\boldsymbol{X}_0)$, then $(\boldsymbol{P_X} - \boldsymbol{P_{X_0}})\boldsymbol{X}\boldsymbol{\beta} \neq 0$ and $\theta > 0$. Hence,

$$\boldsymbol{y}^\top(\boldsymbol{I} - \boldsymbol{P_X})\boldsymbol{y}/\sigma^2 \sim \chi^2_{\mathrm{rank}(\boldsymbol{X}) - \mathrm{rank}(\boldsymbol{X}_0)}(\theta),$$

In general, the noncentrality parameter quantifies how far the mean of $\boldsymbol{y}$ is from $\mathcal{C}(\boldsymbol{X}_0)$ because

$E(y)$

we could divide $(P_X - P_{X_0})/\sigma^2$ but the result would be the same

$$\underbrace{\boldsymbol{\beta}^\top \boldsymbol{X}^\top}_{\mu^T} \underbrace{(\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_{\boldsymbol{X}_0})}_{A} \underbrace{\boldsymbol{X}\boldsymbol{\beta}}_{\mu}$$

$$= \boldsymbol{\beta}^\top \boldsymbol{X}^\top (\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_{\boldsymbol{X}_0})^\top (\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_{\boldsymbol{X}_0}) \boldsymbol{X}\boldsymbol{\beta}$$

$a^T a = ||a||^2$

bc $(P_X - P_{X_0})$ is symmetric & idempotent

$$= \;||\, (\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_{\boldsymbol{X}_0}) \boldsymbol{X}\boldsymbol{\beta} \,||^2 = \;||\, \boldsymbol{P}_{\boldsymbol{X}} \boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{P}_{\boldsymbol{X}_0} \boldsymbol{X}\boldsymbol{\beta} \,||^2$$

$$= \;||\, \boldsymbol{X}\boldsymbol{\beta} - \underline{\boldsymbol{P}_{\boldsymbol{X}_0} \boldsymbol{X}\boldsymbol{\beta}} \,||^2 = \;||\, \mathrm{E}(\boldsymbol{y}) - \boldsymbol{P}_{\boldsymbol{X}_0} \mathrm{E}(\boldsymbol{y}) \,||^2 .$$

if $E(y)$ indeed lives in $\mathcal{C}(X_0)$ then $P_{X_0} E(y) = E(y)$

Note that

$$
\begin{aligned}
\boldsymbol{y}^\top(\boldsymbol{P}_{\boldsymbol{X}} - \boldsymbol{P}_{\boldsymbol{X}_0})\boldsymbol{y} &= \boldsymbol{y}^\top\left[(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{X}_0}) - (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{X}})\right]\boldsymbol{y} \\
&= \boldsymbol{y}^\top \boxed{(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{X}_0})}\boldsymbol{y} - \boldsymbol{y}^\top(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{X}})\boldsymbol{y}
\end{aligned}
$$

reduced model   full model

$$
= \underbrace{SSE_{\text{REDUCED}}} - SSE_{\text{FULL}}.
$$

$SSE_{Red} > SSE_{Full}$

Also rank($\boldsymbol{X}$)−rank($\boldsymbol{X}_0$)

$$
\begin{aligned}
&= [n - \text{rank}(\boldsymbol{X}_0)] - [n - \text{rank}(\boldsymbol{X})] \\
&= DFE_{\text{REDUCED}} - DFE_{\text{FULL}},
\end{aligned}
$$

where $DFE =$ Degrees of Freedom for Error.

Thus, the F statistic has the familiar form

$$\frac{(SSE_{\text{REDUCED}} - SSE_{\text{FULL}})/(DFE_{\text{REDUCED}} - DFE_{\text{FULL}})}{SSE_{\text{FULL}}/DFE_{\text{FULL}}}.$$

Equivalence of F-Tests

It turns out that this reduced vs. full model F-test is equivalent to the F-test for testing

$$H_0 : \boldsymbol{C}\boldsymbol{\beta} = \boldsymbol{d} \qquad \text{vs.} \qquad H_A : \boldsymbol{C}\boldsymbol{\beta} \neq \boldsymbol{d}$$

with an appropriately chosen $\boldsymbol{C}$ and $\boldsymbol{d}$.

The equivalence of these tests is proved in STAT 6110.

Example: F-Test for Lack of Linear Fit

Suppose a balanced, completely randomized design is used to assign 1, 2, or 3 units of fertilizer to a total of 9 plots of land.

treatments

The yield harvested from each plot is recorded as the response.

$y_1, y_2, y_3, \dots, y_9$

1	2	1
3	1	2
2	3	3

3 repl. per trt

Let y_{ij} denote the yield from the jth plot that received i units of fertilizer ($i, j = 1, 2, 3$).

constant

Suppose all yields are independent and $y_{ij} \sim \mathcal{N}(\mu_i, \sigma^2)$ for all $i, j = 1, 2, 3$.

We are allowing treatment specific means

μ_1, μ_2, μ_3

$$\text{If } \boldsymbol{y} = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{32} \\ y_{33} \end{bmatrix}, \text{ then } \mathrm{E}(\boldsymbol{y}) = \begin{bmatrix} \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_3 \\ \mu_3 \\ \mu_3 \end{bmatrix}.$$

Suppose we wish to determine whether there is a linear relationship between the amount of fertilizer applied to a plot and the expected value of a plot's yield.

In other words, suppose we wish to know if there exists real numbers β_1 and β_2 such that

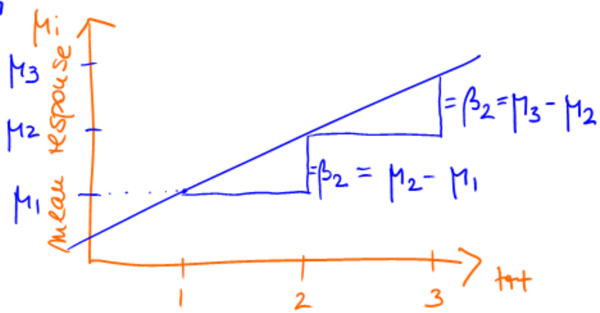
$$\mu_i = \beta_1 + \beta_2(i) \text{ for all } i = 1, 2, 3.$$

under H_0
(linear model)

$\mu_3 - \mu_2 = \mu_2 - \mu_1$

$(\mu_3 - \mu_2) - (\mu_2 - \mu_1) = 0$

$\mu_3 - 2\mu_2 + \mu_1 = 0$



Consider testing

$H_0 : \mathrm{E}(\boldsymbol{y}) \in \mathcal{C}(\boldsymbol{X}_0)$ vs. $\mathrm{H_A} : \mathrm{E}(\boldsymbol{y}) \in \mathcal{C}(\boldsymbol{X}) \setminus \mathcal{C}(\boldsymbol{X}_0)$, where

if linear relationship indeed holds true

$$\boldsymbol{X}_0 = \begin{array}{c} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \\ 1 & 2 \\ 1 & 2 \\ 1 & 3 \\ 1 & 3 \\ 1 & 3 \end{bmatrix} \\ \beta_1 \quad \beta_2 \end{array} \quad \text{and} \quad \underline{\boldsymbol{X}} = \begin{array}{c} \mu_1 \quad \mu_2 \quad \mu_3 \\ \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \end{array}.$$

original model

Note $H_0 : \mathrm{E}(\boldsymbol{y}) \in \mathcal{C}(\boldsymbol{X}_0) \iff \exists\, \boldsymbol{\beta} \in \mathbb{R}^2 \ni \mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}_0\boldsymbol{\beta} \iff$

$$\exists \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \in \mathbb{R}^2 \ni \begin{bmatrix} \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_3 \\ \mu_3 \\ \mu_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \\ 1 & 2 \\ 1 & 2 \\ 1 & 3 \\ 1 & 3 \\ 1 & 3 \end{bmatrix}}_{X_0} \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}}_{\beta} = \begin{array}{r} \mu_1 = \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array} \begin{bmatrix} \beta_1 + \beta_2(1) \\ \beta_1 + \beta_2(1) \\ \beta_1 + \beta_2(1) \\ \beta_1 + \beta_2(2) \\ \beta_1 + \beta_2(2) \\ \beta_1 + \beta_2(2) \\ \beta_1 + \beta_2(3) \\ \beta_1 + \beta_2(3) \\ \beta_1 + \beta_2(3) \end{bmatrix} \begin{array}{l} \\ \\ \\ \\ \mu_2 \\ \\ \\ \mu_3 \\ \\ \end{array}$$

$\iff \mu_i = \beta_1 + \beta_2(i)$ for all $i = 1, 2, 3$.

Note $\mathrm{E}(\boldsymbol{y}) \in \mathcal{C}(\boldsymbol{X}) \iff \exists\, \boldsymbol{\beta} \in \mathbb{R}^3 \ni \mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}\boldsymbol{\beta} \iff$

$$\exists \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \in \mathbb{R}^3 \ni \begin{bmatrix} \mu_1 \\ \mu_1 \\ \mu_1 \\ \mu_2 \\ \mu_2 \\ \mu_2 \\ \mu_3 \\ \mu_3 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{array}{c} \left[\begin{array}{c} \beta_1 \\ \beta_1 \\ \beta_1 \\ \hline \beta_2 \\ \beta_2 \\ \beta_2 \\ \hline \beta_3 \\ \beta_3 \\ \beta_3 \end{array}\right] \end{array} \begin{array}{c} \\ \mu_1 \\ \\ \\ \mu_2 \\ \\ \\ \mu_3 \\ \\ \end{array}$$

β has one more element

This condition clearly holds with $\boxed{\beta_i = \mu_i}$ for all $i = 1, 2, 3$.

The alternative hypothesis

$$H_A : \mathrm{E}(\boldsymbol{y}) \in \mathcal{C}(\boldsymbol{X}) \setminus \mathcal{C}(\boldsymbol{X}_0)$$

is equivalent to

H_A : There do not exist $\beta_1, \beta_2 \in \mathbb{R}$ such that

$$\mu_i = \beta_1 + \beta_2(i) \qquad \forall\ i = 1, 2, 3.$$

Because the lack of fit test is a reduced vs. full model F test, we can also obtain this test by testing

$$H_0 : \boldsymbol{C\beta} = \boldsymbol{d} \quad \text{vs.} \quad H_A : \boldsymbol{C\beta} \neq \boldsymbol{d}$$

for appropriate $\boldsymbol{C}$ and $\boldsymbol{d}$.

going back to slide 26

we got $1\mu_3 - 2\mu_2 + 1\mu_1 = \boxed{0}$

$C^T = (1 \;\; -2 \;\; 1)$

$$\boldsymbol{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} \qquad \boldsymbol{C} = ? \qquad \boldsymbol{d} = ?$$

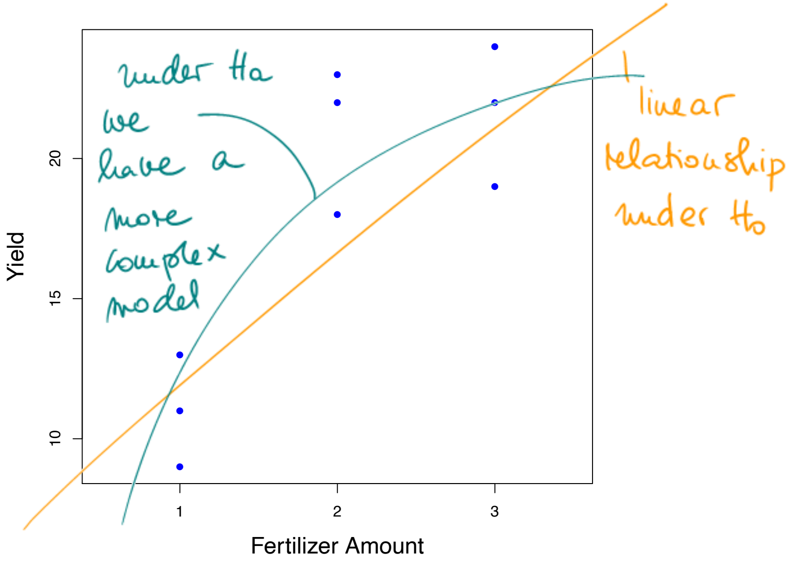
or $C^T = (-1 \;\; +2 \;\; -1)$

R Code and Output

```
> x=rep(1:3,each=3)
> x
[1] 1 1 1 2 2 2 3 3 3
>
> y=c(11,13,9,18,22,23,19,24,22)
```

y_{ij}

```
>
> plot(x,y,pch=16,col=4,xlim=c(.5,3.5),
+       xlab="Fertilizer Amount",
+       ylab="Yield",axes=F,cex.lab=1.5)
> axis(1,labels=1:3,at=1:3)
> axis(2)
> box()
```


under Ha
we
have a
more
complex
model
linear
relationship
under Ho
Yield
20
15
10
1
2
3
Fertilizer Amount

```
> X0=model.matrix(~x)
> X0
  (Intercept) x
1           1 1
2           1 1
3           1 1
4           1 2
5           1 2
6           1 2
7           1 3
8           1 3
9           1 3
```

model matrix, assuming an intercept and slope

$= X_0$

eliminates the intercept in our model

```
> X=model.matrix(~0+factor(x))
> X
  factor(x)1 factor(x)2 factor(x)3
1          1          0          0
2          1          0          0
3          1          0          0
4          0          1          0
5          0          1          0
6          0          1          0
7          0          0          1
8          0          0          1
9          0          0          1
```

μ_1

μ_2

μ_3

$= X$

```
> proj=function(x){
+   x%*%ginv(t(x)%*%x)%*%t(x)
+ }
>
> library(MASS)
> PX0=proj(X0)
> PX=proj(X)
```

$P_X = X(X^T X)^- X^T$

to get generalized inverse

$= P_{X_0}$

$= P_X$

Option 1

$y^T(P_X - P_{X_0})y/1$ ← $q = 1$

```
> Fstat=(t(y)%*%(PX-PX0)%*%y/1)/
+         (t(y)%*%(diag(rep(1,9))-PX)%*%y/(9-3))
> Fstat
         [,1]
[1,] 7.538462
>
> pvalue=1-pf(Fstat,1,6)
> pvalue
           [,1]
[1,] 0.03348515
```

$y^T(I - P_X)y^T/(9-3)$

n $\qquad$ rank(X)

Option 2

```
> reduced=lm(y~x)
> full=lm(y~0+factor(x))
>
> rvsf=function(reduced,full)
+ {
+   sser=deviance(reduced)
+   ssef=deviance(full)
+   dfer=reduced$df
+   dfef=full$df
+   dfn=dfer-dfef
+   Fstat=(sser-ssef)/dfn/
+         (ssef/dfef)
+   pvalue=1-pf(Fstat,dfn,dfef)
+   list(Fstat=Fstat,dfn=dfn,dfd=dfef,
+        pvalue=pvalue)
+ }
```

F as seen on slide 22

```
> rvsf(reduced,full)
$Fstat
[1] 7.538462

$dfn
[1] 1

$dfd
[1] 6

$pvalue
[1] 0.03348515
```

Option 3

```
> anova(reduced,full)
Analysis of Variance Table

Model 1: y ~ x
Model 2: y ~ 0 + factor(x)
  Res.Df    RSS Df Sum of Sq      F  Pr(>F)
1      7 78.222
2      6 34.667  1    43.556 7.5385 0.03349 *
---
Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1
```

```
> test=function(lmout,C,d=0){
+   b=coef(lmout)
+   V=vcov(lmout)
+   dfn=nrow(C)
+   dfd=lmout$df
+   Cb.d=C%*%b-d
+   Fstat=drop(
       t(Cb.d)%*%solve(C%*%V%*%t(C))%*%Cb.d/dfn)
+   pvalue=1-pf(Fstat,dfn,dfd)
+   list(Fstat=Fstat,pvalue=pvalue)
+ }
> test(full,matrix(c(1,-2,1),nrow=1))
$Fstat
[1] 7.538462
$pvalue
[1] 0.03348515
```

Option 4 is F in the form of $(C\beta - d)$

$\{C(X^TX)^- C^T\}$

$$F = \frac{(C\hat{\beta}-d)^T\{\qquad\}^{-1}(C\hat{\beta}-d)}{\hat{\sigma}^2}$$

SAS Code and Output

See the annotated code in Canvas!

```
data d;
  input x y;
  cards;
1 11
1 13
1 9
2 18
2 22
2 23
3 19
3 24
3 22
;
run;
```

```
proc glm;
  class x;
  model y=x;
  contrast 'Lack of Linear Fit' x 1 -2 1;
run;
```


The SAS System

The GLM Procedure

Dependent Variable: y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	214.2222222	107.1111111	18.54	0.0027
Error	6	34.6666667	5.7777778		
Corrected Total	8	248.8888889			

R-Square	Coeff Var	Root MSE	y Mean
0.860714	13.43684	2.403701	17.88889

Source	DF	Type I SS	Mean Square	F Value	Pr > F
x	2	214.2222222	107.1111111	18.54	0.0027

Source	DF	Type III SS	Mean Square	F Value	Pr > F
x	2	214.2222222	107.1111111	18.54	0.0027

Conclusion: We have some (moderately strong) evidence in favor of H_a suggesting that the model allowing for a treatment specific mean is better at modeling $E(y)$.

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
Lack of Linear Fit	1	43.55555556	43.55555556	7.54	0.0335

// end lecture 5
01-31-25