ANOVA

Peter Hoff
Duke STA 610

Review of ANOVA

Analysis of Variance (ANOVA)

For two-level data, ANOVA provides an additive *decomposition of variance:*

$$\text{total variation} = \text{across-group variation} + \text{within-group variation}$$

A typical ANOVA includes

- an estimate of within-group variance;
- an estimate of between-group variance (variance of subpopulation means);
- estimates and tests of contrasts of subpopulation means.

ANOVA decomposition

ANOVA estimation

ANOVA inference

One-factor balanced design

Example (wheat yield):

- $m = 10$ regions of land available,
- $n = 5$ plots of land seeded within each region.
- $y_{i,j}$=the yield of plot i in region j.

One-factor balanced design

Example (wheat yield):

- $m = 10$ regions of land available,
- $n = 5$ plots of land seeded within each region.
- $y_{i,j}$=the yield of plot i in region j.

One-factor balanced design

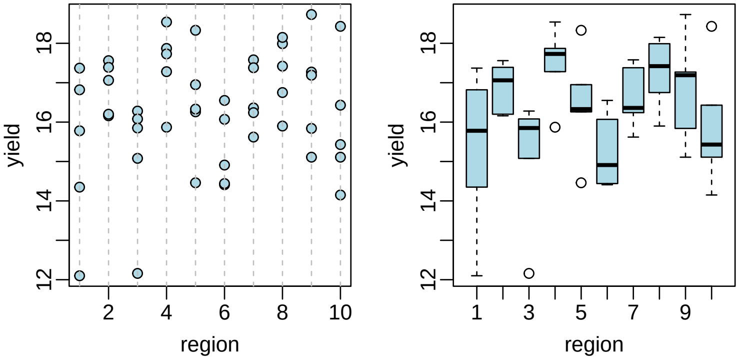
Example (wheat yield):

- $m = 10$ regions of land available,
- $n = 5$ plots of land seeded within each region.
- $y_{i,j}$=the yield of plot i in region j.

One-factor balanced design

Example (wheat yield):

- $m = 10$ regions of land available,
- $n = 5$ plots of land seeded within each region.
- $y_{i,j}$=the yield of plot i in region j.



ANOVA decomposition

Every observation can be written as equal to

- the grand mean, plus
- the difference between its group mean and the grand mean, plus
- the difference between the observation and the group mean.

$$
\begin{aligned}
y_{i,j} &= \bar{y}_{\cdot\cdot} + (\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot}) + (y_{i,j} - \bar{y}_{\cdot j}) \\
&\equiv \hat{\mu} + \hat{a}_j + \hat{\epsilon}_{i,j}.
\end{aligned}
$$

ANOVA decomposition

Total		Across		Within
$y_{11} - \bar{y}_{..}$	$=$	$(\bar{y}_{.1} - \bar{y}_{..})$	$+$	$(y_{11} - \bar{y}_{.1})$
$y_{21} - \bar{y}_{..}$	$=$	$(\bar{y}_{.1} - \bar{y}_{..})$	$+$	$(y_{21} - \bar{y}_{.1})$
.	$=$	.	$+$	.
.	$=$	.	$+$	.
.	$=$	.	$+$	.
$y_{n1} - \bar{y}_{..}$	$=$	$(\bar{y}_{.1} - \bar{y}_{..})$	$+$	$(y_{n1} - \bar{y}_{.1})$
$y_{12} - \bar{y}_{..}$	$=$	$(\bar{y}_{.2} - \bar{y}_{..})$	$+$	$(y_{12} - \bar{y}_{.2})$
.	$=$	.	$+$	.
.	$=$	.	$+$	.
.	$=$	.	$+$	.
$y_{n2} - \bar{y}_{..}$	$=$	$(\bar{y}_{.2} - \bar{y}_{..})$	$+$	$(y_{n2} - \bar{y}_{.2})$
$\vdots$		$\vdots$		$\vdots$
$y_{1m} - \bar{y}_{..}$	$=$	$(\bar{y}_{.m} - \bar{y}_{..})$	$+$	$(y_{1m} - \bar{y}_{.m})$
.	$=$	.	$+$	.
.	$=$	.	$+$	.
.	$=$	.	$+$	.
$y_{nm} - \bar{y}_{..}$	$=$	$(\bar{y}_{.m} - \bar{y}_{..})$	$+$	$(y_{nm} - \bar{y}_{.m})$
$\mathbf{r}_T$	$=$	$\mathbf{r}_A$	$+$	$\mathbf{r}_W$
$mn - 1$	$=$	$m - 1$	$+$	$m(n-1)$

Degrees of freedom

Residual vectors: Each vector $\mathbf{r}_T, \mathbf{r}_A, \mathbf{r}_W$ in the preceding table is of length $N = m \times n$, but lives in a lower-dimensional space:

Total: $\mathbf{r}_T$ lives in an $mn - 1$ dimensional subspace;
Across groups: $\mathbf{r}_A$ lives in an $m - 1$ dimensional subspace;
Within groups: $\mathbf{r}_W$ lives in an $mn - m$ dimensional subspace.

These subspace dimensions are known as *degrees of freedom*.

Exercise: Explain the above results.

Degrees of freedom

Residual vectors: Each vector $\mathbf{r}_T, \mathbf{r}_A, \mathbf{r}_W$ in the preceding table is of length $N = m \times n$, but lives in a lower-dimensional space:

Total: $\mathbf{r}_T$ lives in an $mn - 1$ dimensional subspace;
Across groups: $\mathbf{r}_A$ lives in an $m - 1$ dimensional subspace;
Within groups: $\mathbf{r}_W$ lives in an $mn - m$ dimensional subspace.

These subspace dimensions are known as *degrees of freedom*.

Exercise: Explain the above results.

ANOVA decomposition

SST$= ||\mathbf{r}_T||^2$ = Total sum of squares variation = variation of $y_{i,j}$'s around $\bar{y}_{\cdot\cdot}$;

SSA$= ||\mathbf{r}_A||^2$ = Across group variation = variation of $\bar{y}_j$'s around $\bar{y}_{\cdot\cdot}$;

SSW$= ||\mathbf{r}_W||^2$ = Within group variation = variation of $y_{i,j}$'s around $\bar{y}_j$'s.

Exercise: Show that

- $\mathbf{r}_T = \mathbf{r}_A + \mathbf{r}_W$;
- $\mathbf{r}_A \cdot \mathbf{r}_W = 0$.

Sum of squares decomposition: You can show that

SST	=	SSA	+	SSW
total variation	=	between group variation	+	within group variation

ANOVA decomposition

SST$= ||\mathbf{r}_T||^2$ = Total sum of squares variation = variation of $y_{i,j}$'s around $\bar{y}_{\cdot\cdot}$;

SSA$= ||\mathbf{r}_A||^2$ = Across group variation = variation of $\bar{y}_j$'s around $\bar{y}_{\cdot\cdot}$;

SSW$= ||\mathbf{r}_W||^2$ = Within group variation = variation of $y_{i,j}$'s around $\bar{y}_j$'s.

Exercise: Show that

- $\mathbf{r}_T = \mathbf{r}_A + \mathbf{r}_W$;
- $\mathbf{r}_A \cdot \mathbf{r}_W = 0$.

Sum of squares decomposition: You can show that

$$\begin{array}{ccccc} \text{SST} & = & \text{SSA} & + & \text{SSW} \\ \text{total variation} & = & \text{between group variation} & + & \text{within group variation} \end{array}$$

ANOVA decomposition

SST$= ||\mathbf{r}_T||^2 =$ Total sum of squares variation $=$ variation of $y_{i,j}$'s around $\bar{y}_{\cdot\cdot}$;

SSA$= ||\mathbf{r}_A||^2 =$ Across group variation $=$ variation of $\bar{y}_j$'s around $\bar{y}_{\cdot\cdot}$;

SSW$= ||\mathbf{r}_W||^2 =$ Within group variation $=$ variation of $y_{i,j}$'s around $\bar{y}_j$'s.

Exercise: Show that

- $\mathbf{r}_T = \mathbf{r}_A + \mathbf{r}_W$;
- $\mathbf{r}_A \cdot \mathbf{r}_W = 0$.

Sum of squares decomposition: You can show that

$$\begin{array}{ccccc} \text{SST} & = & \text{SSA} & + & \text{SSW} \\ \text{total variation} & = & \text{between group variation} & + & \text{within group variation} \end{array}$$

ANOVA decomposition

SST= $||\mathbf{r}_T||^2$ = Total sum of squares variation = variation of $y_{i,j}$'s around $\bar{y}_{\cdot\cdot}$;

SSA= $||\mathbf{r}_A||^2$ = Across group variation = variation of $\bar{y}_j$'s around $\bar{y}_{\cdot\cdot}$;

SSW= $||\mathbf{r}_W||^2$ = Within group variation = variation of $y_{i,j}$'s around $\bar{y}_j$'s.

Exercise: Show that

- $\mathbf{r}_T = \mathbf{r}_A + \mathbf{r}_W$;
- $\mathbf{r}_A \cdot \mathbf{r}_W = 0$.

Sum of squares decomposition: You can show that

SST	=	SSA	+	SSW
total variation	=	between group variation	+	within group variation

ANOVA decomposition

SST$= ||\mathbf{r}_T||^2$ = Total sum of squares variation = variation of $y_{i,j}$'s around $\bar{y}_{\cdot\cdot}$;

SSA$= ||\mathbf{r}_A||^2$ = Across group variation = variation of $\bar{y}_j$'s around $\bar{y}_{\cdot\cdot}$;

SSW$= ||\mathbf{r}_W||^2$ = Within group variation = variation of $y_{i,j}$'s around $\bar{y}_j$'s.

Exercise: Show that

- $\mathbf{r}_T = \mathbf{r}_A + \mathbf{r}_W$;
- $\mathbf{r}_A \cdot \mathbf{r}_W = 0$.

Sum of squares decomposition: You can show that

SST	=	SSA	+	SSW
total variation	=	between group variation	+	within group variation

ANOVA for wheat yield

```
y

##  [1] 17.37 15.78 14.35 12.10 16.82 16.16 16.20 17.56 17.39 17.06 16.28 16.08
## [13] 15.08 12.16 15.85 17.87 17.73 15.87 17.28 18.54 18.33 16.26 16.95 16.33
## [25] 14.46 14.41 14.44 14.91 16.55 16.07 16.36 16.24 17.58 17.38 15.62 15.90
## [37] 17.42 17.99 16.75 18.15 17.27 15.84 17.19 15.11 18.73 18.43 15.11 14.15
## [49] 16.43 15.43

g

##  [1]  1  1  1  1  1  2  2  2  2  2  3  3  3  3  3  4  4  4  4  4  5  5  5  5  5
## [26]  6  6  6  6  6  7  7  7  7  7  8  8  8  8  8  9  9  9  9  9 10 10 10 10 10

ybGrand<-mean(y)
ybGroup<-tapply(y,g,mean) ; a<-ybGroup-ybGrand

ybGrand

## [1] 16.3064

mean(ybGroup)

## [1] 16.3064

a

##       1       2       3       4       5       6       7       8       9      10
## -1.0224  0.5676 -1.2164  1.1516  0.1596 -1.0304  0.3296  0.9356  0.5216 -0.3964

mean(a)

## [1] -1.776628e-16
```

ANOVA for wheat yield

```
SST<-sum( (y-ybGrand)^2 )
SST

## [1] 104.8566

ybGroup[ g ]

##      1      1      1      1      1      2      2      2      2      2      3
## 15.284 15.284 15.284 15.284 15.284 16.874 16.874 16.874 16.874 16.874 15.090
##      3      3      3      3      4      4      4      4      4      5      5
## 15.090 15.090 15.090 15.090 17.458 17.458 17.458 17.458 17.458 16.466 16.466
##      5      5      5      6      6      6      6      6      7      7      7
## 16.466 16.466 16.466 15.276 15.276 15.276 15.276 15.276 16.636 16.636 16.636
##      7      7      8      8      8      8      8      9      9      9      9
## 16.636 16.636 17.242 17.242 17.242 17.242 17.242 16.828 16.828 16.828 16.828
##      9     10     10     10     10     10
## 16.828 15.910 15.910 15.910 15.910 15.910

SSA<-sum( (ybGroup[g]-ybGrand)^2 )
SSA

## [1] 33.36831

n*sum( (ybGroup-ybGrand)^2 )

## [1] 33.36831

n*sum(a^2)

## [1] 33.36831
```

ANOVA for wheat yield

```
rA<-ybGroup[g]-ybGrand
rW<-y-ybGroup[g]
sum(rA*rW)

## [1] 1.029211e-14
```

```
SSW<-sum( (y-ybGroup[g])^2 )
SSW

## [1] 71.48824

SSW+SSA

## [1] 104.8566

SST

## [1] 104.8566
```

ANOVA table

The ANOVA decomposition is usually summarized with an ANOVA table:

source	deg of freedom	SS	MS	F-ratio
across	$m-1$	SSA	MSA=SSA$/(m-1)$	MSA/MSW
within	$m(n-1)$	SSW	MSW=SSW$/m(n-1)$	
total	$mn-1$	SST		

```
anova( lm(y~as.factor(g)) )

## Analysis of Variance Table
##
## Response: y
##              Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(g)  9 33.368  3.7076  2.0745 0.0555 .
## Residuals    40 71.488  1.7872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA table

The ANOVA decomposition is usually summarized with an ANOVA table:

source	deg of freedom	SS	MS	F-ratio
across	$m-1$	SSA	MSA=SSA/$(m-1)$	MSA/MSW
within	$m(n-1)$	SSW	MSW=SSW/$m(n-1)$	
total	$mn-1$	SST		

```
anova( lm(y~as.factor(g)) )

## Analysis of Variance Table
##
## Response: y
##              Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(g)  9 33.368  3.7076  2.0745 0.0555 .
## Residuals    40 71.488  1.7872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA table

```
anova( lm(y~as.factor(g)) )

## Analysis of Variance Table
##
## Response: y
##              Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(g)  9 33.368  3.7076  2.0745 0.0555 .
## Residuals    40 71.488  1.7872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

SSA

## [1] 33.36831

SSA/(m-1)

## [1] 3.70759

SSW

## [1] 71.48824

SSW/(m*(n-1))

## [1] 1.787206

(SSA/(m-1)) / (SSW/(m*(n-1)))

## [1] 2.074518
```

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$
\begin{aligned}
\text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\
&= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m)
\end{aligned}
$$

- within group variability

$$
\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2
$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

ANOVA decomposition as a description

The ANOVA decomposition and sums of squares provide

Descriptions of center:

- overall mean: $\bar{y}_{\cdot\cdot}$
- group means: $\bar{y}_1, \ldots, \bar{y}_m$
- group effects: $\bar{y}_1 - \bar{y}_{\cdot\cdot}, \ldots, \bar{y}_1 - \bar{y}_{\cdot\cdot}$

Descriptions of variability:

- across group variability

$$\begin{aligned} \text{SSA} &= \sum_j \sum_i (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 \\ &= n \sum_j (\bar{y}_j - \bar{y}_{\cdot\cdot})^2 = n \times (m-1) \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \end{aligned}$$

- within group variability

$$\text{SSW} = \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2 = \sum_j (n-1) s_j^2$$

Checking calculations

SSA:

```
SSA

## [1] 33.36831

n*(m-1)*var(ybGroup)

## [1] 33.36831
```

SSW:

```
SSW

## [1] 71.48824

tapply(y,g,var)

##       1       2       3       4       5       6       7       8       9      10
## 4.49173 0.43388 2.88970 0.99197 1.94843 0.95908 0.67748 0.86467 1.96792 2.64720

sum( ( n-1)* tapply(y,g,var)  )

## [1] 71.48824
```

Checking calculations

SSA:

```
SSA
```

```
## [1] 33.36831
```

```
n*(m-1)*var(ybGroup)
```

```
## [1] 33.36831
```

SSW:

```
SSW
```

```
## [1] 71.48824
```

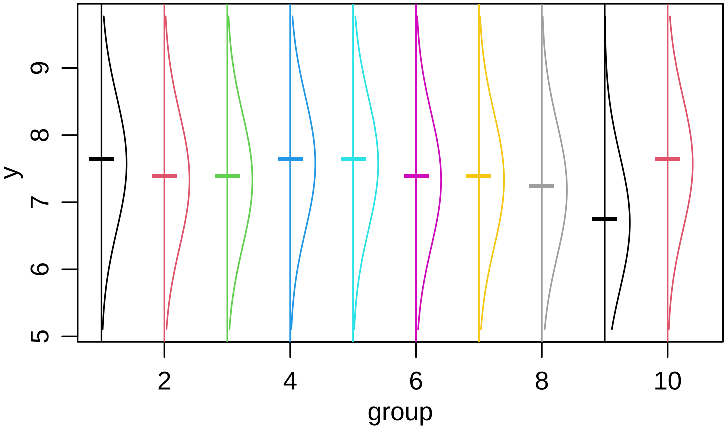
```
tapply(y,g,var)
```

```
##       1       2       3       4       5       6       7       8       9      10
## 4.49173 0.43388 2.88970 0.99197 1.94843 0.95908 0.67748 0.86467 1.96792 2.64720
```

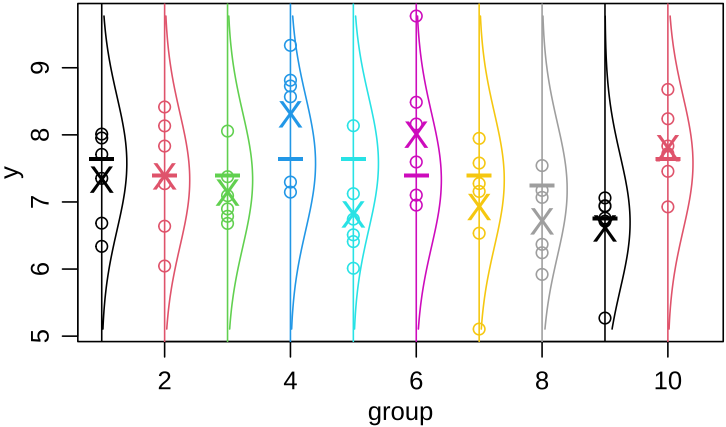
```
sum( ( n-1)* tapply(y,g,var)  )
```

```
## [1] 71.48824
```

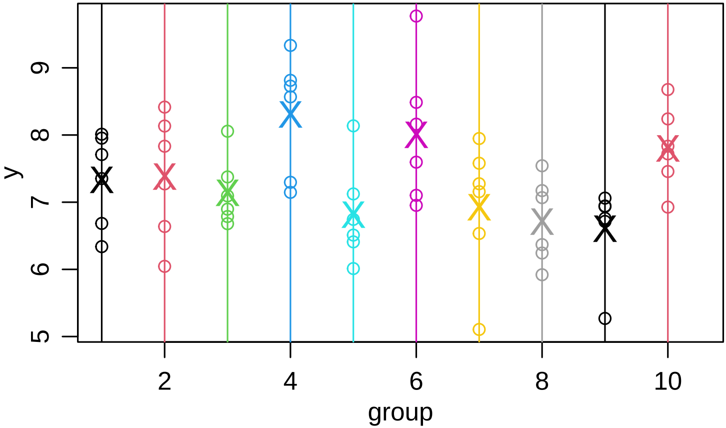
One-way means model



One-way means model



One-way means model



One-way means model

$$y_{i,j} = \mu + a_j + \epsilon_{i,j} \quad \text{(treatment effects model) , or}$$
$$y_{i,j} = \theta_j + \epsilon_{i,j} \qquad \text{(treatment means model),}$$

where $\theta_j = \mu + a_j$.

- μ is expected yield across all regions;
- θ_j is expected yield from region j;
- a_j is the deviation of region-specific expected yield from μ;

$$\theta_j = \mu + a_j \;\; \Leftrightarrow a_j = \theta_j - \mu$$

- $\epsilon_{i,j}$ is the deviation of an observed yield from its region-specific expectation.

One-way means model

$$y_{i,j} = \mu + a_j + \epsilon_{i,j} \quad \text{(treatment effects model) , or}$$
$$y_{i,j} = \theta_j + \epsilon_{i,j} \qquad \text{(treatment means model),}$$

where $\theta_j = \mu + a_j$.

- μ is expected yield across all regions;
- θ_j is expected yield from region j;
- a_j is the deviation of region-specific expected yield from μ;

$$\theta_j = \mu + a_j \ \Leftrightarrow a_j = \theta_j - \mu$$

- $\epsilon_{i,j}$ is the deviation of an observed yield from its region-specific expectation.

One-way means model

$$y_{i,j} = \mu + a_j + \epsilon_{i,j} \quad \text{(treatment effects model) , or}$$
$$y_{i,j} = \theta_j + \epsilon_{i,j} \qquad \text{(treatment means model),}$$

where $\theta_j = \mu + a_j$.

- μ is expected yield across all regions;
- θ_j is expected yield from region j;
- a_j is the deviation of region-specific expected yield from μ;

$$\theta_j = \mu + a_j \;\Leftrightarrow a_j = \theta_j - \mu$$

- $\epsilon_{i,j}$ is the deviation of an observed yield from its region-specific expectation.

One-way means model

$$
\begin{aligned}
y_{i,j} &= \mu + a_j + \epsilon_{i,j} \quad \text{(treatment effects model) , or} \\
y_{i,j} &= \theta_j + \epsilon_{i,j} \qquad \text{(treatment means model),}
\end{aligned}
$$

where $\theta_j = \mu + a_j$.

- μ is expected yield across all regions;
- θ_j is expected yield from region j;
- a_j is the deviation of region-specific expected yield from μ;

$$\theta_j = \mu + a_j \;\Leftrightarrow a_j = \theta_j - \mu$$

- $\epsilon_{i,j}$ is the deviation of an observed yield from its region-specific expectation.

One-way means model

$$y_{i,j} = \mu + a_j + \epsilon_{i,j} \quad \text{(treatment effects model) , or}$$
$$y_{i,j} = \theta_j + \epsilon_{i,j} \qquad \text{(treatment means model),}$$

where $\theta_j = \mu + a_j$.

- μ is expected yield across all regions;
- θ_j is expected yield from region j;
- a_j is the deviation of region-specific expected yield from μ;

$$\theta_j = \mu + a_j \;\Leftrightarrow a_j = \theta_j - \mu$$

- $\epsilon_{i,j}$ is the deviation of an observed yield from its region-specific expectation.

One-way means model

$$y_{i,j} = \mu + a_j + \epsilon_{i,j} \quad \text{(treatment effects model) , or}$$
$$y_{i,j} = \theta_j + \epsilon_{i,j} \qquad \text{(treatment means model),}$$

where $\theta_j = \mu + a_j$.

- μ is expected yield across all regions;
- θ_j is expected yield from region j;
- a_j is the deviation of region-specific expected yield from μ;

$$\theta_j = \mu + a_j \;\Leftrightarrow a_j = \theta_j - \mu$$

- $\epsilon_{i,j}$ is the deviation of an observed yield from its region-specific expectation.

One-way means model

$$y_{i,j} = \mu + a_j + \epsilon_{i,j} \quad \text{(treatment effects model) , or}$$
$$y_{i,j} = \theta_j + \epsilon_{i,j} \qquad \text{(treatment means model),}$$

where $\theta_j = \mu + a_j$.

- μ is expected yield across all regions;
- θ_j is expected yield from region j;
- a_j is the deviation of region-specific expected yield from μ;

$$\theta_j = \mu + a_j \;\Leftrightarrow a_j = \theta_j - \mu$$

- $\epsilon_{i,j}$ is the deviation of an observed yield from its region-specific expectation.

Identifiability

The standard "ANOVA" model parameterizes things so that

- $\sum_j a_j = 0$ (sum-to-zero side conditions),
- $\{\epsilon_{i,j}\} \sim$ i.i.d. from some mean-zero distribution.

In this case,

$$\begin{aligned} \mathrm{E}[y_{i,j}|\mu, a_1, \ldots, a_m] &= \mathrm{E}[\mu + a_j + \epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\ &= \mathrm{E}[\mu|\mu, a_1, \ldots, a_m] + \mathrm{E}[a_j|\mu, a_1, \ldots, a_m] + \mathrm{E}[\epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\ &= \mu + a_j \\ &= \theta_j \end{aligned}$$

If we assume $\{\epsilon_{i,j}\} \sim$ i.i.d. $N(0, \sigma^2)$, then the model is

$$y_{i,j} \sim N(\mu + a_j, \sigma^2) \text{ or equivalently,}$$
$$y_{i,j} \sim N(\theta_j, \sigma^2).$$

Identifiability

The standard "ANOVA" model parameterizes things so that

- $\sum_j a_j = 0$ (sum-to-zero side conditions),
- $\{\epsilon_{i,j}\} \sim$ i.i.d. from some mean-zero distribution.

In this case,

$$
\begin{aligned}
\mathrm{E}[y_{i,j}|\mu, a_1, \ldots, a_m] &= \mathrm{E}[\mu + a_j + \epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\
&= \mathrm{E}[\mu|\mu, a_1, \ldots, a_m] + \mathrm{E}[a_j|\mu, a_1, \ldots, a_m] + \mathrm{E}[\epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\
&= \mu + a_j \\
&= \theta_j
\end{aligned}
$$

If we assume $\{\epsilon_{i,j}\} \sim$ i.i.d. $N(0, \sigma^2)$, then the model is

$$
\begin{aligned}
y_{i,j} &\sim N(\mu + a_j, \sigma^2) \quad \text{or equivalently,} \\
y_{i,j} &\sim N(\theta_j, \sigma^2).
\end{aligned}
$$

Identifiability

The standard "ANOVA" model parameterizes things so that

- $\sum_j a_j = 0$ (sum-to-zero side conditions),
- $\{\epsilon_{i,j}\} \sim$ i.i.d. from some mean-zero distribution.

In this case,

$$\begin{aligned}
\mathrm{E}[y_{i,j}|\mu, a_1, \ldots, a_m] &= \mathrm{E}[\mu + a_j + \epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\
&= \mathrm{E}[\mu|\mu, a_1, \ldots, a_m] + \mathrm{E}[a_j|\mu, a_1, \ldots, a_m] + \mathrm{E}[\epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\
&= \mu + a_j \\
&= \theta_j
\end{aligned}$$

If we assume $\{\epsilon_{i,j}\} \sim$ i.i.d. $N(0, \sigma^2)$, then the model is

$$\begin{aligned}
y_{i,j} &\sim N(\mu + a_j, \sigma^2) \text{ or equivalently,} \\
y_{i,j} &\sim N(\theta_j, \sigma^2).
\end{aligned}$$

Identifiability

The standard "ANOVA" model parameterizes things so that

- $\sum_j a_j = 0$ (sum-to-zero side conditions),
- $\{\epsilon_{i,j}\} \sim$ i.i.d. from some mean-zero distribution.

In this case,

$$\begin{aligned}
E[y_{i,j}|\mu, a_1, \ldots, a_m] &= E[\mu + a_j + \epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\
&= E[\mu|\mu, a_1, \ldots, a_m] + E[a_j|\mu, a_1, \ldots, a_m] + E[\epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\
&= \mu + a_j \\
&= \theta_j
\end{aligned}$$

If we assume $\{\epsilon_{i,j}\} \sim$ i.i.d. $N(0, \sigma^2)$, then the model is

$$\begin{aligned}
y_{i,j} &\sim N(\mu + a_j, \sigma^2) \text{ or equivalently,} \\
y_{i,j} &\sim N(\theta_j, \sigma^2).
\end{aligned}$$

Identifiability

The standard "ANOVA" model parameterizes things so that

- $\sum_j a_j = 0$ (sum-to-zero side conditions),
- $\{\epsilon_{i,j}\} \sim$ i.i.d. from some mean-zero distribution.

In this case,

$$
\begin{aligned}
E[y_{i,j}|\mu, a_1, \ldots, a_m] &= E[\mu + a_j + \epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\
&= E[\mu|\mu, a_1, \ldots, a_m] + E[a_j|\mu, a_1, \ldots, a_m] + E[\epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\
&= \mu + a_j \\
&= \theta_j
\end{aligned}
$$

If we assume $\{\epsilon_{i,j}\} \sim$ i.i.d. $N(0, \sigma^2)$, then the model is

$$
\begin{aligned}
y_{i,j} &\sim N(\mu + a_j, \sigma^2) \text{ or equivalently,} \\
y_{i,j} &\sim N(\theta_j, \sigma^2).
\end{aligned}
$$

Identifiability

The standard "ANOVA" model parameterizes things so that

- $\sum_j a_j = 0$ (sum-to-zero side conditions),
- $\{\epsilon_{i,j}\} \sim$ i.i.d. from some mean-zero distribution.

In this case,

$$
\begin{aligned}
\mathrm{E}[y_{i,j}|\mu, a_1, \ldots, a_m] &= \mathrm{E}[\mu + a_j + \epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\
&= \mathrm{E}[\mu|\mu, a_1, \ldots, a_m] + \mathrm{E}[a_j|\mu, a_1, \ldots, a_m] + \mathrm{E}[\epsilon_{i,j}|\mu, a_1, \ldots, a_m] \\
&= \mu + a_j \\
&= \theta_j
\end{aligned}
$$

If we assume $\{\epsilon_{i,j}\} \sim$ i.i.d. $N(0, \sigma^2)$, then the model is

$$
\begin{aligned}
y_{i,j} &\sim N(\mu + a_j, \sigma^2) \text{ or equivalently,} \\
y_{i,j} &\sim N(\theta_j, \sigma^2).
\end{aligned}
$$

Parameter estimates

Parameters to estimate include

- $\{\theta_1, \ldots, \theta_m, \sigma^2\}$, or equivalently
- $\{\mu, a_1, \ldots, a_m, \sigma^2\}$

If $\hat{\theta}_j$ is an estimate of θ_j, we say that

- $\hat{y}_{i,j} = \hat{\theta}_j$ is the *fitted value* of $y_{i,j}$;
- $\hat{\epsilon}_{i,j} = y_{i,j} - \hat{y}_{i,j} = y_{i,j} - \hat{\theta}_j$ is the *residual* for $y_{i,j}$.

Parameter estimates

Parameters to estimate include

- $\{\theta_1, \ldots, \theta_m, \sigma^2\}$, or equivalently
- $\{\mu, a_1, \ldots, a_m, \sigma^2\}$

If $\hat{\theta}_j$ is an estimate of θ_j, we say that

- $\hat{y}_{i,j} = \hat{\theta}_j$ is the *fitted value* of $y_{i,j}$;
- $\hat{\epsilon}_{i,j} = y_{i,j} - \hat{y}_{i,j} = y_{i,j} - \hat{\theta}_j$ is the *residual* for $y_{i,j}$.

OLS estimation

The OLS estimates of $\theta_1, \ldots, \theta_m$ are the values that minimize SSR:

$$\begin{aligned} \text{SSR}(\hat{\theta}_1, \ldots, \hat{\theta}_m) &= \sum_{j=1}^{m} \sum_{i=1}^{n} (y_{i,j} - \hat{\theta}_j)^2 \\ &= \sum_{i=1}^{n} (y_{i,1} - \hat{\theta}_1)^2 + \cdots + \sum_{i=1}^{n} (y_{i,m} - \hat{\theta}_m)^2 \end{aligned}$$

Exercise: Show that $\hat{\theta}_j = \bar{y}_j$ is the OLSE/MLE for θ_j.

Note: For $\hat{\theta}_j = \bar{y}_j$, $SSR = SSW$.

OLS estimation

The OLS estimates of $\theta_1, \ldots, \theta_m$ are the values that minimize SSR:

$$\begin{aligned} \text{SSR}(\hat\theta_1, \ldots, \hat\theta_m) &= \sum_{j=1}^{m} \sum_{i=1}^{n} (y_{i,j} - \hat\theta_j)^2 \\ &= \sum_{i=1}^{n} (y_{i,1} - \hat\theta_1)^2 + \cdots + \sum_{i=1}^{n} (y_{i,m} - \hat\theta_m)^2 \end{aligned}$$

Exercise: Show that $\hat\theta_j = \bar y_j$ is the OLSE/MLE for θ_j.

Note: For $\hat\theta_j = \bar y_j$, $SSR = SSW$.

OLS estimation

The OLS estimates of $\theta_1, \dots, \theta_m$ are the values that minimize SSR:

$$\begin{aligned} \text{SSR}(\hat\theta_1, \dots, \hat\theta_m) &= \sum_{j=1}^{m} \sum_{i=1}^{n} (y_{i,j} - \hat\theta_j)^2 \\ &= \sum_{i=1}^{n} (y_{i,1} - \hat\theta_1)^2 + \cdots + \sum_{i=1}^{n} (y_{i,m} - \hat\theta_m)^2 \end{aligned}$$

Exercise: Show that $\hat\theta_j = \bar y_j$ is the OLSE/MLE for θ_j.

Note: For $\hat\theta_j = \bar y_j$, $SSR = SSW$.

OLS estimation

For the "treatment effects" parametrization, we have that

- $\theta_j = \mu + a_j$
- $\sum_j a_j = 0$,

which together imply that

$$\mu = \sum \theta_j / m.$$

So our OLS estimates of $\{\mu, a_1, \ldots, a_m\}$ are

$$\hat{\mu} = \sum \hat{\theta}_j / m$$
$$\hat{a}_j = \hat{\theta}_j - \hat{\mu}.$$

OLS estimation

For the "treatment effects" parametrization, we have that

- $\theta_j = \mu + a_j$
- $\sum_j a_j = 0$,

which together imply that

$$\mu = \sum \theta_j / m.$$

So our OLS estimates of $\{\mu, a_1, \ldots, a_m\}$ are

$$\hat{\mu} = \sum \hat{\theta}_j / m$$
$$\hat{a}_j = \hat{\theta}_j - \hat{\mu}.$$

Unbiased variance estimation

Recall we assumed that within each group j,

$$y_{i,j} = \theta_j + \epsilon_{i,j}$$

where the $\epsilon_{i,j}$'s are independent with mean 0 and variance σ^2.

This implies

- $\mathrm{E}[y_{i,j}|\theta_j, \sigma^2] = \theta_j$;
- $\mathrm{Var}[y_{i,j}|\theta_j, \sigma^2] = \sigma^2$;
- $y_{1,j}, \ldots, y_{n,j}$ are uncorrelated with each other.

This further implies that the sample variance is an unbiased estimator of σ^2:

$$s_j^2 = \sum_i (y_{i,j} - \bar{y}_j)^2/(n-1)$$

$$\mathrm{E}[s_j^2|\theta_j, \sigma^2] = \sigma^2$$

Unbiased variance estimation

Recall we assumed that within each group j,

$$y_{i,j} = \theta_j + \epsilon_{i,j}$$

where the $\epsilon_{i,j}$'s are independent with mean 0 and variance σ^2.

This implies

- $\mathrm{E}[y_{i,j}|\theta_j, \sigma^2] = \theta_j$;
- $\mathrm{Var}[y_{i,j}|\theta_j, \sigma^2] = \sigma^2$;
- $y_{1,j}, \ldots, y_{n,j}$ are uncorrelated with each other.

This further implies that the sample variance is an unbiased estimator of σ^2:

$$s_j^2 = \sum_i (y_{i,j} - \bar{y}_j)^2/(n-1)$$

$$\mathrm{E}[s_j^2|\theta_j, \sigma^2] = \sigma^2$$

Unbiased variance estimation

Recall we assumed that within each group j,

$$y_{i,j} = \theta_j + \epsilon_{i,j}$$

where the $\epsilon_{i,j}$'s are independent with mean 0 and variance σ^2.

This implies

- $\mathrm{E}[y_{i,j}|\theta_j, \sigma^2] = \theta_j$;
- $\mathrm{Var}[y_{i,j}|\theta_j, \sigma^2] = \sigma^2$;
- $y_{1,j}, \ldots, y_{n,j}$ are uncorrelated with each other.

This further implies that the sample variance is an unbiased estimator of σ^2:

$$s_j^2 = \sum_i (y_{i,j} - \bar{y}_j)^2/(n-1)$$

$$\mathrm{E}[s_j^2|\theta_j, \sigma^2] = \sigma^2$$

Pooled sample variance

We pool all the sample variances to obtain an unbiased estimate of σ^2:

$$\hat{\sigma}^2 = (s_1^2 + \cdots + s_m^2)/m$$
$$\mathrm{E}[\sigma^2|\theta,\sigma^2] = (\mathrm{E}[s_1^2|\theta,\sigma^2] + \cdots + \mathrm{E}[s_m^2|\theta,\sigma^2])/m$$
$$= (\sigma^2 + \cdots \sigma_m^2)/m = \sigma^2.$$

Variance estimate via SSW

$$SSW = \sum_{i=1}^{n}(y_{i,1} - \hat{\theta}_1)^2 + \cdots + \sum_{i=1}^{n}(y_{i,m} - \hat{\theta}_m)^2$$
$$= \sum_{i=1}^{n}(y_{i,1} - \bar{y}_1)^2 + \cdots + \sum_{i=1}^{n}(y_{i,m} - \bar{y}_m)^2$$
$$= (n-1)s_1^2 + \cdots + (n-1)s_m^2$$

so

$$\hat{\sigma}^2 = SSW/[m(n-1)].$$

The estimate $\hat{\sigma}^2$ is sometimes called the MSW, MSR or MSE.

Pooled sample variance

We pool all the sample variances to obtain an unbiased estimate of σ^2:

$$\begin{aligned} \hat{\sigma}^2 &= (s_1^2 + \cdots + s_m^2)/m \\ \mathrm{E}[\sigma^2|\theta, \sigma^2] &= (\mathrm{E}[s_1^2|\theta, \sigma^2] + \cdots + \mathrm{E}[s_m^2|\theta, \sigma^2])/m \\ &= (\sigma^2 + \cdots \sigma_m^2)/m = \sigma^2. \end{aligned}$$

Variance estimate via SSW

$$\begin{aligned} SSW &= \sum_{i=1}^{n}(y_{i,1} - \hat{\theta}_1)^2 + \cdots + \sum_{i=1}^{n}(y_{i,m} - \hat{\theta}_m)^2 \\ &= \sum_{i=1}^{n}(y_{i,1} - \bar{y}_1)^2 + \cdots + \sum_{i=1}^{n}(y_{i,m} - \bar{y}_m)^2 \\ &= (n-1)s_1^2 + \cdots + (n-1)s_m^2 \end{aligned}$$

so

$$\hat{\sigma}^2 = SSW/[m(n-1)].$$

The estimate $\hat{\sigma}^2$ is sometimes called the MSW, MSR or MSE.

Wheat yield data

```
## pooled sample variance
s2groups<-tapply(y,g,var)

s2groups

##       1       2       3       4       5       6       7       8       9      10
## 4.49173 0.43388 2.88970 0.99197 1.94843 0.95908 0.67748 0.86467 1.96792 2.64720

mean(s2groups)

## [1] 1.787206

## SSW and MSW
SSW<-sum( (y-ybGroup[g])^2 ) # was SSW

SSW

## [1] 71.48824

MSW<-SSW/(m*(n-1))

MSW

## [1] 1.787206
```

MSW, MSA and the F-statistic

```
anova(lm( y ~ as.factor(g) ))

## Analysis of Variance Table
##
## Response: y
##              Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(g)  9 33.368  3.7076  2.0745 0.0555 .
## Residuals    40 71.488  1.7872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

MSW<-SSW/(m*(n-1))
MSW

## [1] 1.787206

MSA<-SSA/(m-1)
MSA

## [1] 3.70759

MSA/MSW

## [1] 2.074518
```

Testing for across-group heterogeneity

Model:

$$y_{i,j} = \mu + a_j + \epsilon_{i,j} \quad \{\epsilon_{i,j}\} \sim \text{iid } N(0, \sigma^2)$$

Hypotheses: Consider deciding between the following hypotheses:

$$H_0 : a_j = 0 \text{ for all } j$$
$$H_1 : a_j \neq 0 \text{ for some } j$$

H_0 implies all group means are the same, H_1 implies the opposite.

Statistical inference: How to evaluate H_1 versus H_0 using the observed data?

Testing for across-group heterogeneity

Model:

$$y_{i,j} = \mu + a_j + \epsilon_{i,j} \quad \{\epsilon_{i,j}\} \sim \text{iid } N(0, \sigma^2)$$

Hypotheses: Consider deciding between the following hypotheses:

$$H_0 : a_j = 0 \text{ for all } j$$
$$H_1 : a_j \neq 0 \text{ for some } j$$

H_0 implies all group means are the same, H_1 implies the opposite.

Statistical inference: How to evaluate H_1 versus H_0 using the observed data?

Testing for across-group heterogeneity

Model:

$$y_{i,j} = \mu + a_j + \epsilon_{i,j} \quad \{\epsilon_{i,j}\} \sim \text{iid } N(0, \sigma^2)$$

Hypotheses: Consider deciding between the following hypotheses:

$$H_0 : a_j = 0 \text{ for all } j$$
$$H_1 : a_j \neq 0 \text{ for some } j$$

H_0 implies all group means are the same, H_1 implies the opposite.

Statistical inference: How to evaluate H_1 versus H_0 using the observed data?

MSA as a measure of across-group heterogeneity

$$
\begin{aligned}
SSA &= \sum_{i=1}^{n} \sum_{j=1}^{m} (\bar{y}_j - \bar{y})^2 \\
&= n \times \sum_{j=1}^{m} (\bar{y}_j - \bar{y})^2 \\
MSA &= SSA/(m-1) \\
&= n \times \sum_{j=1}^{m} (\bar{y}_j - \bar{y}_{\cdot\cdot})^2/(m-1) \\
&= n \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m)
\end{aligned}
$$

because the average of the $\bar{y}_j$'s is $\bar{y}_{\cdot\cdot}$.

MSA as a measure of across-group heterogeneity

$$
\begin{aligned}
\mathrm{E}[\bar{y}_j] &= \mu + a_j \\
\bar{y}_j &\approx \mu + a_j
\end{aligned}
$$

$$
\begin{aligned}
\text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) &\approx \text{sample variance}(\mu + a_1, \ldots, \mu + a_m) \\
&= \text{sample variance}(a_1, \ldots, a_m) \\
&= \frac{1}{m-1} \sum a_j^2
\end{aligned}
$$

Intuitively,

$$H_0 \text{ true} \Leftrightarrow \frac{1}{m-1} \sum a_j^2 = 0 \Leftrightarrow \text{small MSA}$$

$$H_1 \text{ true} \Leftrightarrow \frac{1}{m-1} \sum a_j^2 > 0 \Leftrightarrow \text{large MSA}$$

MSA as a measure of across-group heterogeneity

$$
\begin{aligned}
\mathrm{E}[\bar{y}_j] &= \mu + a_j \\
\bar{y}_j &\approx \mu + a_j
\end{aligned}
$$

$$
\begin{aligned}
\text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) &\approx \text{sample variance}(\mu + a_1, \ldots, \mu + a_m) \\
&= \text{sample variance}(a_1, \ldots, a_m) \\
&= \frac{1}{m-1} \sum a_j^2
\end{aligned}
$$

Intuitively,

$$H_0 \text{ true} \Leftrightarrow \frac{1}{m-1} \sum a_j^2 = 0 \Leftrightarrow \text{small MSA}$$

$$H_1 \text{ true} \Leftrightarrow \frac{1}{m-1} \sum a_j^2 > 0 \Leftrightarrow \text{large MSA}$$

MSA as a measure of across-group heterogeneity

$$
\begin{aligned}
\mathrm{E}[\bar{y}_j] &= \mu + a_j \\
\bar{y}_j &\approx \mu + a_j
\end{aligned}
$$

$$
\begin{aligned}
\text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) &\approx \text{sample variance}(\mu + a_1, \ldots, \mu + a_m) \\
&= \text{sample variance}(a_1, \ldots, a_m) \\
&= \frac{1}{m-1} \sum a_j^2
\end{aligned}
$$

Intuitively,

$$
H_0 \text{ true} \Leftrightarrow \frac{1}{m-1} \sum a_j^2 = 0 \Leftrightarrow \text{small MSA}
$$

$$
H_1 \text{ true} \Leftrightarrow \frac{1}{m-1} \sum a_j^2 > 0 \Leftrightarrow \text{large MSA}
$$

MSA as a measure of across-group heterogeneity

$$\begin{aligned} \mathrm{E}[\bar{y}_j] &= \mu + a_j \\ \bar{y}_j &\approx \mu + a_j \end{aligned}$$

$$\begin{aligned} \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) &\approx \text{sample variance}(\mu + a_1, \ldots, \mu + a_m) \\ &= \text{sample variance}(a_1, \ldots, a_m) \\ &= \frac{1}{m-1} \sum a_j^2 \end{aligned}$$

Intuitively,

$$H_0 \text{ true} \Leftrightarrow \frac{1}{m-1} \sum a_j^2 = 0 \Leftrightarrow \text{small MSA}$$

$$H_1 \text{ true} \Leftrightarrow \frac{1}{m-1} \sum a_j^2 > 0 \Leftrightarrow \text{large MSA}$$

MSA as a measure of across-group heterogeneity

$$\begin{aligned} \mathrm{E}[\bar{y}_j] &= \mu + a_j \\ \bar{y}_j &\approx \mu + a_j \end{aligned}$$

$$\begin{aligned} \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) &\approx \text{sample variance}(\mu + a_1, \ldots, \mu + a_m) \\ &= \text{sample variance}(a_1, \ldots, a_m) \\ &= \frac{1}{m-1} \sum a_j^2 \end{aligned}$$

Intuitively,

$$H_0 \text{ true} \Leftrightarrow \frac{1}{m-1} \sum a_j^2 = 0 \Leftrightarrow \text{small MSA}$$

$$H_1 \text{ true} \Leftrightarrow \frac{1}{m-1} \sum a_j^2 > 0 \Leftrightarrow \text{large MSA}$$

Expected mean squares

$$
\begin{aligned}
MSA &= n \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \\
&\approx n \times \text{sample variance}(a_1, \ldots, a_m) \\
&= n \times \frac{1}{m-1} \sum a_j^2
\end{aligned}
$$

More precisely, one can show that

$$\mathrm{E}[MSA] = \sigma^2 + n \times \frac{1}{m-1} \sum a_j^2,$$

where the σ^2 comes from the fact that $\bar{y}_j$ only approximates a_j.

Letting $\tau^2 = \frac{1}{m-1} \sum a_j^2$, we have

$$\mathrm{E}[MSA] = \sigma^2 + n \times \tau^2,$$

where τ^2 is the *across-group variability* - the "empirical" variance of $a_1, \ldots, a_m$.

Expected mean squares

$$
\begin{aligned}
MSA &= n \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \\
&\approx n \times \text{sample variance}(a_1, \ldots, a_m) \\
&= n \times \frac{1}{m-1} \sum a_j^2
\end{aligned}
$$

More precisely, one can show that

$$
\mathrm{E}[MSA] = \sigma^2 + n \times \frac{1}{m-1} \sum a_j^2,
$$

where the σ^2 comes from the fact that $\bar{y}_j$ only approximates a_j.

Letting $\tau^2 = \frac{1}{m-1} \sum a_j^2$, we have

$$
\mathrm{E}[MSA] = \sigma^2 + n \times \tau^2,
$$

where τ^2 is the *across-group variability* - the “empirical” variance of $a_1, \ldots, a_m$.

Expected mean squares

$$\begin{aligned} MSA &= n \times \text{sample variance}(\bar{y}_1, \ldots, \bar{y}_m) \\ &\approx n \times \text{sample variance}(a_1, \ldots, a_m) \\ &= n \times \frac{1}{m-1} \sum a_j^2 \end{aligned}$$

More precisely, one can show that

$$\mathrm{E}[MSA] = \sigma^2 + n \times \frac{1}{m-1} \sum a_j^2,$$

where the σ^2 comes from the fact that $\bar{y}_j$ only approximates a_j.

Letting $\tau^2 = \frac{1}{m-1} \sum a_j^2$, we have

$$\mathrm{E}[MSA] = \sigma^2 + n \times \tau^2,$$

where τ^2 is the *across-group variability* - the "empirical" variance of $a_1, \ldots, a_m$.

Testing across-group variability

How can we use MSA to evaluate $H_0 : \tau^2 = 0$?

Idea:

$$MSA \approx \sigma^2 \;\Rightarrow\; \tau^2 \text{ is small or zero} \;\Rightarrow\; \text{accept } H_0$$

$$MSA > \sigma^2 \;\Rightarrow\; \tau^2 \text{ is not zero} \;\Rightarrow\; \text{accept } H_1$$

Problem: We don't know what σ^2 is.

Solution: Compare MSA to an estimate of σ^2.

Testing across-group variability

How can we use MSA to evaluate $H_0 : \tau^2 = 0$?

Idea:

$$MSA \approx \sigma^2 \;\Rightarrow\; \tau^2 \text{ is small or zero} \;\Rightarrow\; \text{accept } H_0$$

$$MSA > \sigma^2 \;\Rightarrow\; \tau^2 \text{ is not zero} \;\Rightarrow\; \text{accept } H_1$$

Problem: We don't know what σ^2 is.

Solution: Compare MSA to an estimate of σ^2.

Testing across-group variability

How can we use MSA to evaluate $H_0 : \tau^2 = 0$?

Idea:

$$MSA \approx \sigma^2 \;\Rightarrow\; \tau^2 \text{ is small or zero} \;\Rightarrow\; \text{accept } H_0$$

$$MSA > \sigma^2 \;\Rightarrow\; \tau^2 \text{ is not zero} \;\Rightarrow\; \text{accept } H_1$$

Problem: We don't know what σ^2 is.

Solution: Compare MSA to an estimate of σ^2.

Testing across-group variability

How can we use MSA to evaluate $H_0 : \tau^2 = 0$?

Idea:

$$MSA \approx \sigma^2 \;\Rightarrow\; \tau^2 \text{ is small or zero} \;\Rightarrow\; \text{accept } H_0$$

$$MSA > \sigma^2 \;\Rightarrow\; \tau^2 \text{ is not zero} \;\Rightarrow\; \text{accept } H_1$$

Problem: We don't know what σ^2 is.

Solution: Compare MSA to an estimate of σ^2.

Testing across-group variability

How can we use MSA to evaluate $H_0 : \tau^2 = 0$?

Idea:

$$MSA \approx \sigma^2 \;\Rightarrow\; \tau^2 \text{ is small or zero} \;\Rightarrow\; \text{accept } H_0$$

$$MSA > \sigma^2 \;\Rightarrow\; \tau^2 \text{ is not zero} \;\Rightarrow\; \text{accept } H_1$$

Problem: We don't know what σ^2 is.

Solution: Compare MSA to an estimate of σ^2.

Testing across-group variability

We have shown that

$$MSW = SSW/m(n-1) = \frac{1}{m(n-1)}\sum_j \sum_i (y_{i,j} - \bar{y}_j)^2$$

$$= \frac{1}{m}\sum_j s_j^2$$

$$\mathrm{E}[MSW] = \sigma^2.$$

Under H_0 and H_1:

$$\mathrm{E}[MSA] = \sigma^2 + n \times \tau^2$$

$$\mathrm{E}[MSW] = \sigma^2$$

Under H_0 only:

$$\mathrm{E}[MSA] = \sigma^2$$

$$\mathrm{E}[MSW] = \sigma^2$$

Testing across-group variability

We have shown that

$$MSW = SSW/m(n-1) = \frac{1}{m(n-1)} \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2$$

$$= \frac{1}{m} \sum_j s_j^2$$

$$\mathrm{E}[MSW] = \sigma^2.$$

Under H_0 and H_1:

$$\mathrm{E}[MSA] = \sigma^2 + n \times \tau^2$$

$$\mathrm{E}[MSW] = \sigma^2$$

Under H_0 only:

$$\mathrm{E}[MSA] = \sigma^2$$

$$\mathrm{E}[MSW] = \sigma^2$$

Testing across-group variability

We have shown that

$$MSW = SSW/m(n-1) = \frac{1}{m(n-1)}\sum_j \sum_i (y_{i,j} - \bar{y}_j)^2$$
$$= \frac{1}{m}\sum_j s_j^2$$
$$\mathrm{E}[MSW] = \sigma^2.$$

Under H_0 and H_1:

$$\mathrm{E}[MSA] = \sigma^2 + n \times \tau^2$$
$$\mathrm{E}[MSW] = \sigma^2$$

Under H_0 only:

$$\mathrm{E}[MSA] = \sigma^2$$
$$\mathrm{E}[MSW] = \sigma^2$$

Testing across-group variability

We have shown that

$$MSW = SSW/m(n-1) = \frac{1}{m(n-1)} \sum_j \sum_i (y_{i,j} - \bar{y}_j)^2$$
$$= \frac{1}{m} \sum_j s_j^2$$
$$\mathrm{E}[MSW] = \sigma^2.$$

Under H_0 and H_1:

$$\mathrm{E}[MSA] = \sigma^2 + n \times \tau^2$$
$$\mathrm{E}[MSW] = \sigma^2$$

Under H_0 only:

$$\mathrm{E}[MSA] = \sigma^2$$
$$\mathrm{E}[MSW] = \sigma^2$$

The F-statistic and distribution

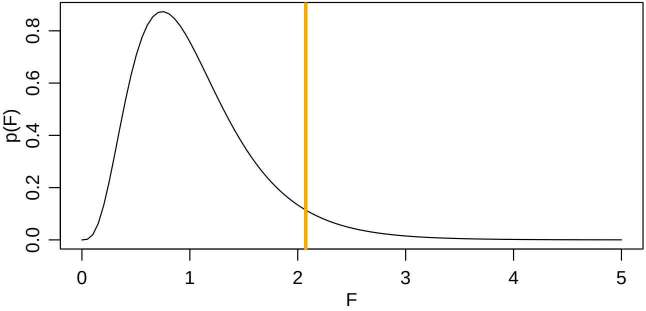
Let $F = MSA/MSW$. Then

under H_0, MSA/MSW should be around 1,

under H_1, MSA/MSW should be bigger than 1.

Under the normal model $y_{1,1} \ldots, y_{n,m} \sim$ i.i.d. $N(\mu, \sigma^2)$,

$$MSA/MSW = F \sim F_{m-1,m(n-1)}.$$



The F-statistic and distribution

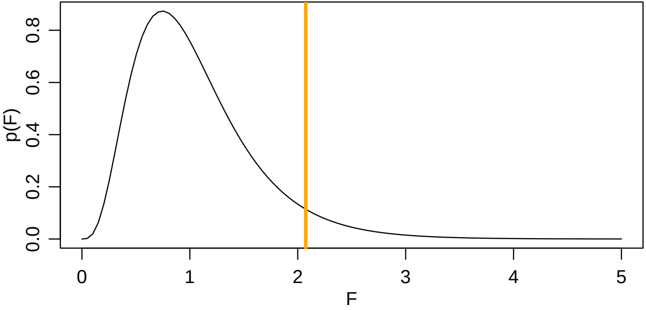
Let $F = MSA/MSW$. Then

under H_0, MSA/MSW should be around 1,

under H_1, MSA/MSW should be bigger than 1.

Under the normal model $y_{1,1} \ldots, y_{n,m} \sim$ i.i.d. $N(\mu, \sigma^2)$,

$$MSA/MSW = F \sim F_{m-1,m(n-1)}.$$



The F-statistic and distribution

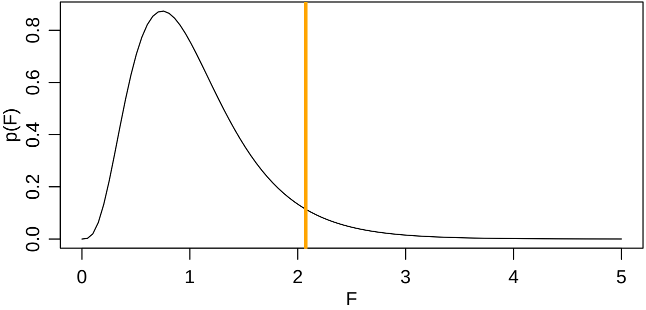
Let $F = MSA/MSW$. Then

under H_0, MSA/MSW should be around 1,

under H_1, MSA/MSW should be bigger than 1.

Under the normal model $y_{1,1} \ldots, y_{n,m} \sim$ i.i.d. $N(\mu, \sigma^2)$,

$$MSA/MSW = F \sim F_{m-1,m(n-1)}.$$



The F-statistic and distribution

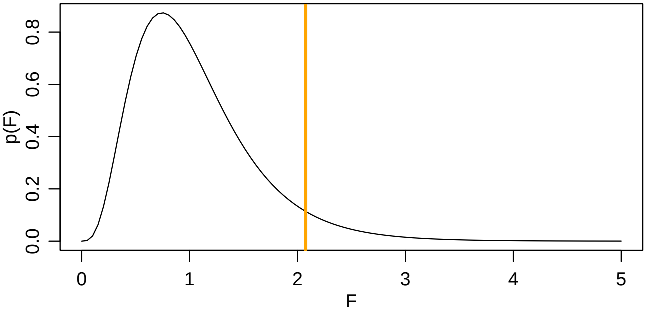
Let $F = MSA/MSW$. Then

under H_0, MSA/MSW should be around 1,

under H_1, MSA/MSW should be bigger than 1.

Under the normal model $y_{1,1} \ldots, y_{n,m} \sim$ i.i.d. $N(\mu, \sigma^2)$,

$$MSA/MSW = F \sim F_{m-1,m(n-1)}.$$



Classical testing for across-group heterogeneity

- We expect an $F_{m-1,m(n-1)}$-distribution under H_0.
- We observe $F(\boldsymbol{y}) = MSA/MSW$.
- Discrepancy between $F_{m-1,m(n-1)}$ and $F(\boldsymbol{y})$ is evidence against H_0.

$$p\text{-value} = \Pr(F_{m-1,m(n-1)} \geq F(\boldsymbol{y}))$$

```
MSA<-SSA/(m-1)
MSW<-SSW/(m*(n-1))
MSA/MSW

## [1] 2.074518

1-pf( MSA/MSW, m-1,m*(n-1))

## [1] 0.05550019
```

Classical testing for across-group heterogeneity

- We expect an $F_{m-1,m(n-1)}$-distribution under H_0.
- We observe $F(\boldsymbol{y}) = MSA/MSW$.
- Discrepancy between $F_{m-1,m(n-1)}$ and $F(\boldsymbol{y})$ is evidence against H_0.

$$p\text{-value} = \Pr(F_{m-1,m(n-1)} \geq F(\boldsymbol{y}))$$

```
MSA<-SSA/(m-1)
MSW<-SSW/(m*(n-1))
MSA/MSW

## [1] 2.074518

1-pf( MSA/MSW, m-1,m*(n-1))

## [1] 0.05550019
```

Classical testing for across-group heterogeneity

- We expect an $F_{m-1,m(n-1)}$-distribution under H_0.
- We observe $F(\boldsymbol{y}) = MSA/MSW$.
- Discrepancy between $F_{m-1,m(n-1)}$ and $F(\boldsymbol{y})$ is evidence against H_0.

$$p\text{-value} = \Pr(F_{m-1,m(n-1)} \geq F(\boldsymbol{y}))$$

```
MSA<-SSA/(m-1)
MSW<-SSW/(m*(n-1))
MSA/MSW

## [1] 2.074518

1-pf( MSA/MSW, m-1,m*(n-1))

## [1] 0.05550019
```

Classical testing for across-group heterogeneity

- We expect an $F_{m-1,m(n-1)}$-distribution under H_0.
- We observe $F(\boldsymbol{y}) = MSA/MSW$.
- Discrepancy between $F_{m-1,m(n-1)}$ and $F(\boldsymbol{y})$ is evidence against H_0.

$$p\text{-value} = \Pr(F_{m-1,m(n-1)} \geq F(\boldsymbol{y}))$$

```
MSA<-SSA/(m-1)
MSW<-SSW/(m*(n-1))
MSA/MSW

## [1] 2.074518

1-pf( MSA/MSW, m-1,m*(n-1))

## [1] 0.05550019
```

Classical testing for across-group heterogeneity

- We expect an $F_{m-1,m(n-1)}$-distribution under H_0.
- We observe $F(\boldsymbol{y}) = MSA/MSW$.
- Discrepancy between $F_{m-1,m(n-1)}$ and $F(\boldsymbol{y})$ is evidence against H_0.

$$p\text{-value} = \Pr(F_{m-1,m(n-1)} \geq F(\boldsymbol{y}))$$

```
MSA<-SSA/(m-1)
MSW<-SSW/(m*(n-1))
MSA/MSW

## [1] 2.074518

1-pf( MSA/MSW, m-1,m*(n-1))

## [1] 0.05550019
```

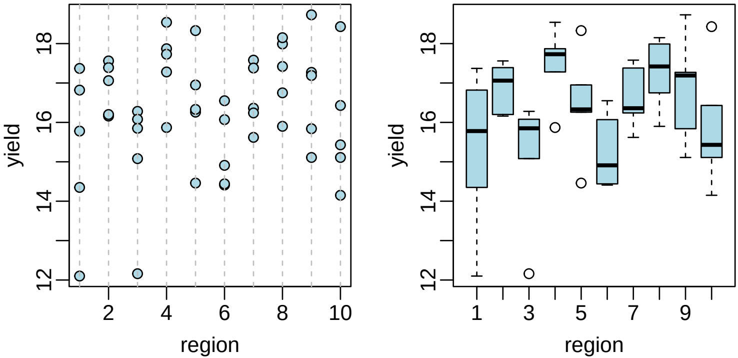
ANOVA table

```
anova(lm(y~as.factor(g)))

## Analysis of Variance Table
##
## Response: y
##              Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(g)  9 33.368  3.7076  2.0745 0.0555 .
## Residuals    40 71.488  1.7872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Group comparisons

If $H : \tau^2 = 0$ is rejected, which groups are likely different from each other?



Confidence intervals and pairwise comparisons

Under the normal model,

$$\bar{y}_j \sim N(\theta_j, \sigma^2/n)$$

Based on this result,

$$c_j(\mathbf{y}) = \bar{y}_j \pm t_{1-\alpha/2,m(n-1)} \times \sqrt{\hat{\sigma}^2/n}$$

is a $1-\alpha$ confidence interval for θ_j. This means

$$\Pr(\theta_j \in c_j(\mathbf{y})|\theta_1,\ldots,\theta_m,\sigma^2) = 1-\alpha,$$

where the probability is over the data $\mathbf{y}$.

Confidence intervals and pairwise comparisons

Under the normal model,

$$\bar{y}_j \sim N(\theta_j, \sigma^2/n)$$

Based on this result,

$$c_j(\mathbf{y}) = \bar{y}_j \pm t_{1-\alpha/2, m(n-1)} \times \sqrt{\hat{\sigma}^2/n}$$

is a $1-\alpha$ confidence interval for θ_j. This means

$$\Pr(\theta_j \in c_j(\mathbf{y}) | \theta_1, \ldots, \theta_m, \sigma^2) = 1 - \alpha,$$

where the probability is over the data $\mathbf{y}$.

Confidence intervals and pairwise comparisons

Similarly, under the normal model,

$$\bar{y}_j - \bar{y}_k \sim N(\theta_j - \theta_k, 2\sigma^2/n)$$

Confidence interval: A $1-\alpha$ confidence interval for $\theta_j - \theta_k$ is

$$c_{j,k}(\mathbf{y}) = (\bar{y}_j - \bar{y}_k) \pm t_{1-\alpha/2, m(n-1)} \times \sqrt{2\hat{\sigma}^2/n}$$

Two-sample t-test: The hypothesis $H_{j,k} : \theta_j = \theta_k$ is rejected at level α if

$$t_{j,k} = \frac{|\bar{y}_j - \bar{y}_k|}{\sqrt{2\hat{\sigma}^2/n}} > t_{1-\alpha/2, m(n-1)}.$$

Confidence intervals and pairwise comparisons

Similarly, under the normal model,

$$\bar{y}_j - \bar{y}_k \sim N(\theta_j - \theta_k, 2\sigma^2/n)$$

Confidence interval: A $1-\alpha$ confidence interval for $\theta_j - \theta_k$ is

$$c_{j,k}(\mathbf{y}) = (\bar{y}_j - \bar{y}_k) \pm t_{1-\alpha/2,m(n-1)} \times \sqrt{2\hat{\sigma}^2/n}$$

Two-sample t-test: The hypothesis $H_{j,k} : \theta_j = \theta_k$ is rejected at level α if

$$t_{j,k} = \frac{|\bar{y}_j - \bar{y}_k|}{\sqrt{2\hat{\sigma}^2/n}} > t_{1-\alpha/2,m(n-1)}.$$

Confidence intervals and pairwise comparisons

Similarly, under the normal model,

$$\bar{y}_j - \bar{y}_k \sim N(\theta_j - \theta_k, 2\sigma^2/n)$$

Confidence interval: A $1-\alpha$ confidence interval for $\theta_j - \theta_k$ is

$$c_{j,k}(\mathbf{y}) = (\bar{y}_j - \bar{y}_k) \pm t_{1-\alpha/2, m(n-1)} \times \sqrt{2\hat{\sigma}^2/n}$$

Two-sample t-test: The hypothesis $H_{j,k} : \theta_j = \theta_k$ is rejected at level α if

$$t_{j,k} = \frac{|\bar{y}_j - \bar{y}_k|}{\sqrt{2\hat{\sigma}^2/n}} > t_{1-\alpha/2, m(n-1)}.$$

Wheat yield example

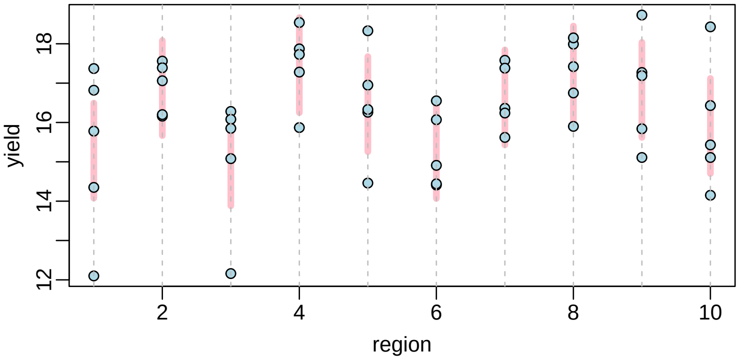
```
## using confint
fit<-lm( y ~ -1+as.factor(g) )
ciTheta<-confint(fit)
ciTheta

##                    2.5 %   97.5 %
## as.factor(g)1  14.07567 16.49233
## as.factor(g)2  15.66567 18.08233
## as.factor(g)3  13.88167 16.29833
## as.factor(g)4  16.24967 18.66633
## as.factor(g)5  15.25767 17.67433
## as.factor(g)6  14.06767 16.48433
## as.factor(g)7  15.42767 17.84433
## as.factor(g)8  16.03367 18.45033
## as.factor(g)9  15.61967 18.03633
## as.factor(g)10 14.70167 17.11833

## "by hand"
s2hat<-anova(fit)[2,3]
mean(y[g==3]) + c(-1,1)*qt(.975,m*(n-1))*sqrt(s2hat/n)

## [1] 13.88167 16.29833
```

Wheat yield example



Wheat yield example

```
wheat.compare<-agricolae::LSD.test(aov(y~as.factor(g)),"as.factor(g)")
wheat.compare$means

##          y       std r        se      LCL      UCL   Min   Max   Q25   Q50   Q75
## 1   15.284 2.1193702 5 0.5978639 14.07567 16.49233 12.10 17.37 14.35 15.78 16.82
## 10  15.910 1.6270218 5 0.5978639 14.70167 17.11833 14.15 18.43 15.11 15.43 16.43
## 2   16.874 0.6586957 5 0.5978639 15.66567 18.08233 16.16 17.56 16.20 17.06 17.39
## 3   15.090 1.6999118 5 0.5978639 13.88167 16.29833 12.16 16.28 15.08 15.85 16.08
## 4   17.458 0.9959769 5 0.5978639 16.24967 18.66633 15.87 18.54 17.28 17.73 17.87
## 5   16.466 1.3958617 5 0.5978639 15.25767 17.67433 14.46 18.33 16.26 16.33 16.95
## 6   15.276 0.9793263 5 0.5978639 14.06767 16.48433 14.41 16.55 14.44 14.91 16.07
## 7   16.636 0.8230917 5 0.5978639 15.42767 17.84433 15.62 17.58 16.24 16.36 17.38
## 8   17.242 0.9298763 5 0.5978639 16.03367 18.45033 15.90 18.15 16.75 17.42 17.99
## 9   16.828 1.4028257 5 0.5978639 15.61967 18.03633 15.11 18.73 15.84 17.19 17.27

wheat.compare$groups

##          y groups
## 4   17.458      a
## 8   17.242      a
## 2   16.874     ab
## 9   16.828     ab
## 7   16.636    abc
## 5   16.466    abc
## 10  15.910    abc
## 1   15.284     bc
## 6   15.276     bc
## 3   15.090      c
```