Two factor factorial design

Factorial design: treatments represent all combinations of levels of A and B.

Paper Towel: Amount of liquid absorbed (mL)

	Water	Detergent	Oil
Coronet	26,22,22	19,16,15	22,25,29
Kleenex	43,41,41	33,38,38	39,41,45
Scott	27,26,25	21,20,21	27,25,25

Factor A: paper towel type (3 levels)
Factor B: liquid type (3 levels).
9 treatments, 3 replications at each treatment ($n = 3$).

Model assumptions

Notation: Factor A has a levels, factor B has b levels. So there are ab treatments.

$y_{ijk} = \mu_{ij} + \epsilon_{ijk}, i = 1, \cdots, a; j = 1, \cdots, b; k = 1, \cdots, n.$
Assume independent normal error terms, $\epsilon_{ijk} \sim N(0, \sigma^2)$.
(equilavent to assuming independent, normal observations $y_{ijk} \sim N(\mu_{ij}, \sigma^2)$.)
Note we also assume equal number of replications n at each treatment. This is called a **balanced design**. With **unbalanced design** there are unequal number of replications for different treatments.

Decompose μ_{ij}: Factor effects model

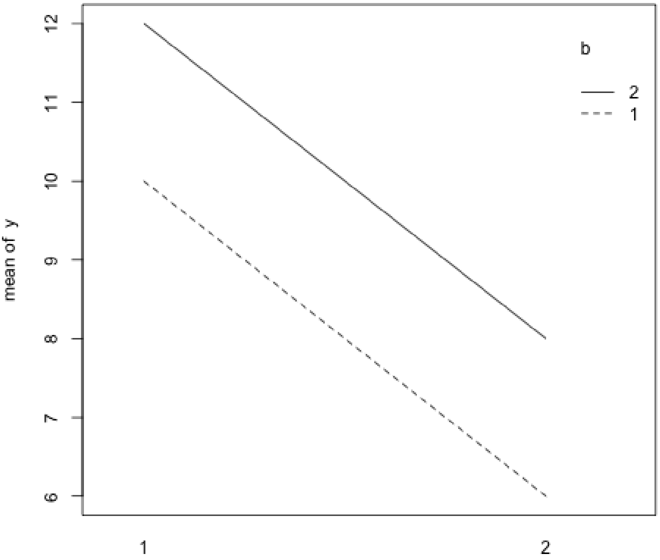
	Watering Regimen B_1	B_2		
Fertilizer				
A_1	$\mu_{11} = 10$	$\mu_{12} = 12$	$\mu_{1\cdot} = 11$	$\alpha_1 = 2$
A_2	$\mu_{21} = 6$	$\mu_{22} = 8$	$\mu_{2\cdot} = 7$	$\alpha_2 = -2$
	$\mu_{\cdot 1} = 8$	$\mu_{\cdot 2} = 10$	$\mu_{\cdot\cdot} = 9$	
	$\beta_1 = -1$	$\beta_2 = 1$		

average mean across the levels of B:
$\mu_{1\cdot} = (10 + 12)/2 = 11, \mu_{2\cdot} = 7$
grand mean $\mu_{\cdot\cdot} = 9$.

True main effect of A_1 is $\alpha_1 = \mu_{1\cdot} - \mu_{\cdot\cdot} = 11 - 9 = 2$,
similarly, main effect of A_2 is $\alpha_2 = -2$.
In general, main effect of A_i is $\alpha_i = \mu_{i\cdot} - \mu_{\cdot\cdot}$

Also we can get true main effect for B: $\beta_1 = -1, \beta_2 = 1$.
In general, main effect of B_j is $\beta_j = \mu_{\cdot j} - \mu_{\cdot\cdot}$

Interaction plot



No Interaction Model

Each μ_{ij} can be written as a sum of grand mean, factor A effect and factor B effect: $\mu_{ij} = \mu_{..} + \alpha_i + \beta_j, i = 1 \cdots , a; j = 1, \cdots , b$.

$$\mu_{11} = 10 = \mu_{..} + \alpha_1 + \beta_1 = 9 + 2 + (-1) = 10$$

$$\mu_{12} = 12 = \mu_{..} + \alpha_1 + \beta_2 = 9 + 2 + 1 = 12$$

$$\mu_{21} = 6 = \mu_{..} + \alpha_2 + \beta_1 = 9 + (-2) + (-1) = 6$$

$$\mu_{22} = 12 = \mu_{..} + \alpha_2 + \beta_2 = 9 + (-2) + 1 = 8$$

This is the no interaction model: The effect of one factor does not depend upon levels of the other factor. The lines are parallel in interaction plots which are plots of treatment means.

Interaction model

	B_1	B_2		
A_1	$\mu_{11} = 10$	$\mu_{12} = 16$	$\mu_{1\cdot} = 13$	$\alpha_1 = 3$
A_2	$\mu_{21} = 6$	$\mu_{22} = 8$	$\mu_{2\cdot} = 7$	$\alpha_2 = -3$
	$\mu_{\cdot 1} = 8$	$\mu_{\cdot 2} = 12$	$\mu_{\cdot\cdot} = 10$	
	$\beta_1 = -2$	$\beta_2 = 2$		

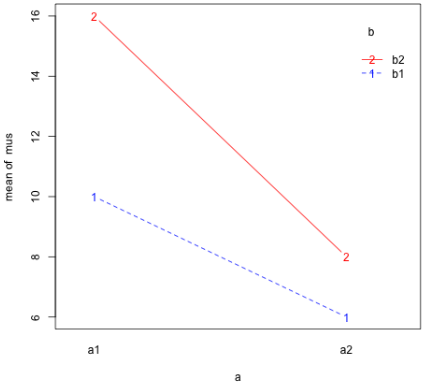
Need extra $\alpha\beta_{ij}$ term to decompose μ_{ij}.

$\mu_{ij} = \mu_{\cdot\cdot} + \alpha_i + \beta_j + \alpha\beta_{ij}$

where $\alpha\beta_{ij} = \mu_{ij} - (\mu_{\cdot\cdot} + \alpha_i + \beta_j)$ are called **interaction terms**.

$$\mu_{11} = \mu_{\cdot\cdot} + \alpha_1 + \beta_1 + [\mu_{11} - (\mu_{\cdot\cdot} + \alpha_1 + \beta_1)] \quad \text{or} \quad 10 = 10 + 3 + (-2) + [-1]$$
$$\mu_{12} = \mu_{\cdot\cdot} + \alpha_1 + \beta_2 + [\mu_{12} - (\mu_{\cdot\cdot} + \alpha_1 + \beta_2)] \quad \text{or} \quad 16 = 10 + 3 + 2 + [1]$$
$$\mu_{21} = \mu_{\cdot\cdot} + \alpha_2 + \beta_1 + [\mu_{21} - (\mu_{\cdot\cdot} + \alpha_2 + \beta_1)] \quad \text{or} \quad 6 = 10 + (-3) + (-2) + [1]$$
$$\mu_{22} = \mu_{\cdot\cdot} + \alpha_2 + \beta_2 + [\mu_{22} - (\mu_{\cdot\cdot} + \alpha_2 + \beta_2)] \quad \text{or} \quad 8 = 10 + (-3) + 2 + [-1]$$

Interaction plot



For models with interaction effects, the lines are not parallel in interaction plots.
The effect of one factor depends on the level of the other factor.

Estimate the main and interaction effects

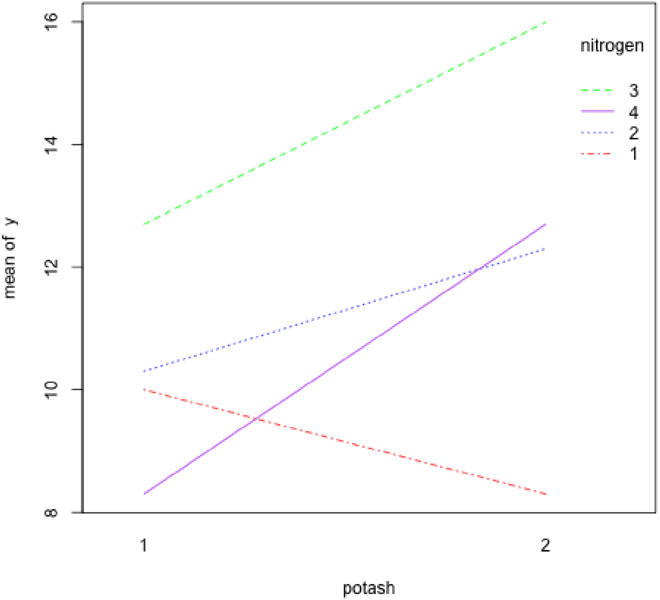
Replace the population parameters with sample statistics:
$\hat{\alpha}_i = \bar{y}_{i..} - \bar{y}_{...}$ (row mean-grand mean)
$\hat{\beta}_j = \bar{y}_{.j.} - \bar{y}_{...}$ (column mean-grand mean)
$\widehat{\alpha\beta}_{ij} = \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}$
(cell mean-row mean-column mean+grand mean)

Treatment means for 8 treatments

		Nitrogen				
		5%	10%	15%	20%	
Potash	10%	10.0	10.3	12.7	8.3	mean: 10.325
	20%	8.3	12.3	16.0	12.7	mean: 12.325
	mean:	9.15	11.3	14.35	10.5	grand mean: 11.325

MSE= 3.625.

a. Find $\hat{\alpha}_i, i = 1, 2$.
b. Find $\hat{\beta}_j, j = 1, 2, 3, 4$.
c. Find $\widehat{\alpha\beta}_{ij}, i = 1, 2; j = 1, 2, 3, 4$.



mean of y
16
14
12
10
8
1
2
potash
nitrogen
3
4
2
1

$\bar{y}_{1\cdot\cdot} = 10.325, \bar{y}_{2\cdot\cdot} = 12.325,$
$\bar{y}_{\cdot\cdot\cdot} = 11.325,$
Hence the estimated Potash effects are:
$\hat{\alpha}_1 = 10.325 - 11.325 = -1, \hat{\alpha}_2 = 1.$

similarly,
$\bar{y}_{\cdot 1 \cdot} = 9.15, \bar{y}_{\cdot 2 \cdot} = 11.3,$
$\bar{y}_{\cdot 3 \cdot} = 14.35, \bar{y}_{\cdot 4 \cdot} = 10.5,$
so the estimated Nitrogen effects are
$\hat{\beta}_1 = -2.175, \hat{\beta}_2 = -0.025, \hat{\beta}_3 = 3.025, \hat{\beta}_4 = -0.825.$

The estimated interaction effects (8 terms):
$\widehat{\alpha\beta}_{11} = 10.0 - 10.325 - 9.15 + 11.325 = 1.85.$
$\cdots\cdots$
$\widehat{\alpha\beta}_{24} = 12.7 - 12.325 - 10.5 + 11.325 = 1.2.$

R code to make interaction plot

Using data on the next page:

```
y11=c(8,8,9,12);y12=c(11,11,12,13)
y21=c(5,6,6,6); y22=c(7,8,8,9)
y= c(y11, y12, y21, y22)
a = c(rep(1,8), rep(2,8))
b= c(rep(1,4), rep(2,4), rep(1, 4), rep(2, 4))
a=factor(a)
b=factor(b)
interaction.plot(a,b,y)
```

Sum of squares due to factor A

```
                  B
             1              2
-----|------------|--------
A 1  |  8,8,9,12  |11,11,12,13      mean:10.5
-----|------------|----------
  2  |  5,6,6,6   |7,8,8,9          mean: 6.875

                           grand mean: 8.6875
```

Compute *SSA*.
$\bar{y}_{1\cdot\cdot} = 10.5, \bar{y}_{2\cdot\cdot} = 6.875, \bar{y}_{\cdot\cdot\cdot} = 8.6875$.
$b = 2, n = 4,$
SSA$= 2 * 4 * [(10.5 - 8.6875)^2 + (6.875 - 8.6875)^2] = 52.56$
(each row mean is the average of 8 observations)

Decompose SST

$$SST_c = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \bar{y}_{\cdots})^2,$$
$$SSA = bn \sum_{i=1}^{a} (\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots})^2,$$
$$SSB = an \sum_{j=1}^{b} (\bar{y}_{\cdot j \cdot} - \bar{y}_{\cdots})^2,$$
$$SSAB = n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j \cdot} + \bar{y}_{\cdots})^2,$$
$$SSE = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \bar{y}_{ij\cdot})^2.$$

Fact: $SST_c = SSA + SSB + SSAB + SSE$.

ANOVA table

Source of Variation	df	SS	MS	F	P-value
A	a-1	SSA	MSA	MSA/MSE	
B	b-1	SSB	MSB	MSB/MSE	
A*B	(a-1)*(b-1)	SSAB	MSAB	MSAB/MSE	
Error	ab(n-1)	SSE	MSE		
Total	N-1	SST_c			

$SST_c = SSA + SSB + SSAB + SSE$
$N - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1)$
where $N = abn$.

F test

Factor A effect:
$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$
$F = \frac{MSA}{MSE} \sim F_{a-1,ab(n-1)}$ under H_0.

Fator B effect:
$H_0 : \beta_1 = \beta_2 = \cdots = \beta_b = 0$
$F = \frac{MSB}{MSE} \sim F_{b-1,ab(n-1)}$ under H_0.

AB interaction effect:
$H_0 : \alpha\beta_{11} = \cdots = \alpha\beta_{ab} = 0$
$F = \frac{MSAB}{MSE} \sim F_{(a-1)(b-1),ab(n-1)}$.

Paper Towel: Amount of liquid absorbed (mL)

	Water	Detergent	Oil	
Coronet	26,22,22	19,16,15	22,25,29	mean:21.78
Kleenex	43,41,41	33,38,38	39,41,45	mean:39.89
Scott	27,26,25	21,20,21	27,25,25	

Paper towel example

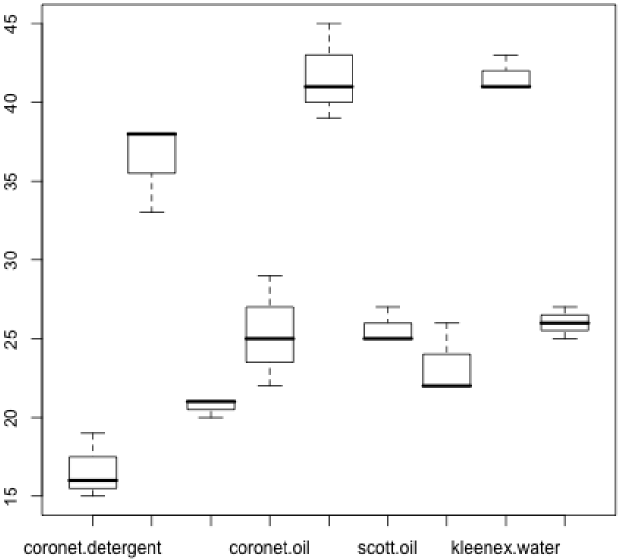
```
y = c(26,22,22,19,16,15,22,25,29,43,41,41,33,38,38,39,41,45,
27,26,25,21,20,21,27,25,25)
a = c(rep("coronet",9),rep("kleenex",9),rep("scott",9))
b1 =c(rep("water",3),rep("detergent",3),rep("oil",3))
b = c(b1,b1,b1)
a =factor(a)
b=factor(b)
interaction.plot(a,b,y)
output = aov(y~a+b+a*b)
anova(output)
boxplot(y~a+b)
output = aov (y~a+b)
TukeyHSD(output)
```

papertowel data

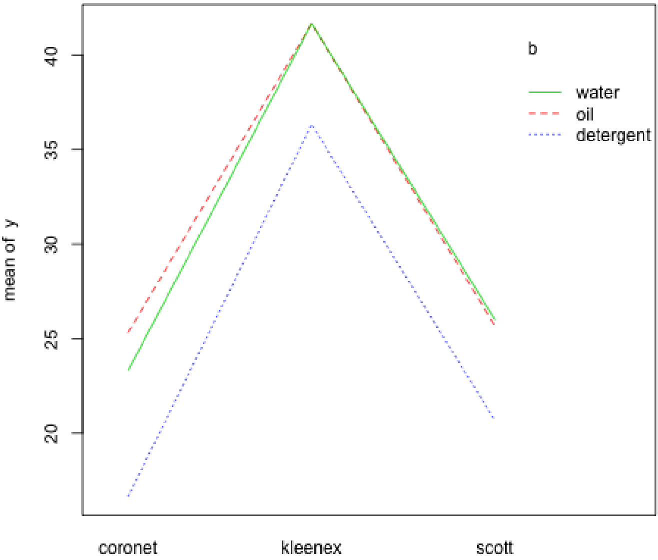
	y	a	b
1	26	coronet	water
2	22	coronet	water
3	22	coronet	water
4	19	coronet	detergent
5	16	coronet	detergent
6	15	coronet	detergent
7	22	coronet	oil
8	25	coronet	oil
9	29	coronet	oil
10	43	kleenex	water
11	41	kleenex	water
12	41	kleenex	water
13	33	kleenex	detergent
14	38	kleenex	detergent
15	38	kleenex	detergent

16	39	kleenex	oil
17	41	kleenex	oil
18	45	kleenex	oil
19	27	scott	water
20	26	scott	water
21	25	scott	water
22	21	scott	detergent
23	20	scott	detergent
24	21	scott	detergent
25	27	scott	oil
26	25	scott	oil
27	25	scott	oil

Boxplot


15
20
25
30
35
40
45
coronet.detergent
coronet.oil
scott.oil
kleenex.water

Interaction plot



Model with interaction terms

```
> output <- aov(y~a+b+a*b)
> summary(output)
Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq  F value    Pr(>F)
a          2 1747.19  873.59 180.0534 1.256e-12 ***
b          2  221.41  110.70  22.8168 1.160e-05 ***
a:b        4   12.59    3.15   0.6489     0.635
Residuals 18   87.33    4.85
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

Model without interaction

```
> output <- aov(y~a+b)
> summary(output)

Analysis of Variance Table

Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
a          2 1747.19  873.59 192.333 1.162e-14 ***
b          2  221.41  110.70  24.373 2.630e-06 ***
Residuals 22   99.93    4.54
```

```
> TukeyHSD(output)
  Tukey multiple comparisons of means
    95% family-wise confidence level
Fit: aov(formula = y ~ a + b)
$a
                      diff         lwr         upr     p adj
kleenex-coronet  18.111111  15.5873279  20.634894 0.0000000
scott-coronet     2.333333  -0.1904499   4.857117 0.0734828
scott-kleenex   -15.777778 -18.3015610 -13.253995 0.0000000

$b
                       diff       lwr      upr     p adj
oil-detergent     6.3333333  3.809550 8.857117 0.0000070
water-detergent   5.7777778  3.253995 8.301561 0.0000252
water-oil        -0.5555556 -3.079339 1.968228 0.8460364
```

check: qtukey(0.95,3,22)/sqrt(2)=2.512,
$39.89 - 21.78 \pm 2.512 * \sqrt{4.54} * \sqrt{1/9 + 1/9} = 18.11 \pm 2.52 =$
$(15.59, 20.63)$.