ST 5203: Experimental Design

(Semester 1, AY 2017/2018)

Text book: *Experiments: Planning, Analysis, and Optimization (2nd. edition)*

by Jeff Wu and Mike Hamada

Topic 3: Experiments with More than One Factor

- Paired comparison design.
- Randomized block (RB) design; additive model.
- Two-way and multiple-way layout experiments.
- Response transformation.
- Latin square design.
- Analysis of covariance (ANCOVA).

Example: Sewage Experiment

- To compare two methods MSI and SIB for determining chlorine content in sewage effluents; $y =$ residual chlorine reading.

Residual Chlorine Readings, Sewage Experiment

	Method		
Sample	MSI	SIB	d_i
1	0.39	0.36	-0.03
2	0.84	1.35	0.51
3	1.76	2.56	0.80
4	3.35	3.92	0.57
5	4.69	5.35	0.66
6	7.70	8.33	0.63
7	10.52	10.70	0.18
8	10.92	10.91	-0.01

- Experimental Design: Eight samples were collected at different doses and contact times. Two methods were applied to each of the eight samples. It is a paired comparison design because the pair of treatments are applied to the same samples (or units). d_i records the difference between MSI and SIB.

Paired Comparison Design vs. Unpaired Design

- Paired Comparison Design: Two treatments are randomly assigned to each block of two units. Can eliminate block-to-block variation and is effective if such variation is large.
  Examples: pairs of twins, eyes, kidneys, left and right feet. (Subject-to-subject variation much larger than within-subject variation).
- Unpaired Design: Each treatment is applied to a separate set of units, or called the two-sample problem. Useful if pairing is unnecessary; also it has more degrees of freedom for error estimation (why?).

Paired t-test

- Paired t-test: Let $y_{i,1}, y_{i,2}$ be the responses of treatments 1 and 2 for unit i, $i = 1, \ldots, N$. Let $d_i = y_{i,2} - y_{i,1}$, $\bar{d}$ and s_d^2 the sample mean and variance of d_i, then

$$T_{\text{paired}} = \frac{\bar{d}}{s_d/\sqrt{N}} \sim t_{N-1}$$

Unpaired t-test

- Assumption: $2N$ units are randomly grouped into N blocks of two units. Each treatment receives N out of the $2N$ units by random allocation.
- Let $\bar{y}_i, s_i^2$ be the sample mean and sample variance for the ith treatment, $i = 1, 2$, then,

$$T_{\text{unpaired}} = \frac{\bar{y}_2 - \bar{y}_1}{\sqrt{(s_1^2 + s_2^2)/N}} \sim t_{2N-2}.$$

  Question: detailed assumptions such that the above statement is rigorously true?
- Note that the degrees of freedom in paired and unpaired t-test are $N - 1$ and $2N - 2$ respectively. Question: $N - 1 < 2N - 2$, why do we bother to use paired t-test?

Example: t-test results for Sewage Experiment

- $t_{\text{paired}} = \dfrac{0.4138}{0.1135} = 3.645$. p-value $= 0.008$.
- $t_{\text{unpaired}} = \dfrac{0.4138}{2.0863} = 0.198$. p-value $= 0.848$.
- Unpaired t-test fails to declare significant difference because its denominator 2.0863 is too large.
- The reason is that the denominator contains the block-to-block variation component.

Example: ANOVA results for Sewage Experiment

ANOVA Table (Paired)

Source	D.F.	S.S.	M.S.	F
Sample	7	243.404	34,772	674.82
Method	1	0.6848	0.6848	13.29
Error	7	0.3607	0.0515	
Total	15	244.450		

ANOVA Table (Unpaired)

Source	D.F.	S.S.	M.S.	F
Method	1	0.6848	0.6848	0.04
Error	14	243.7649	17.412	
Total	15	244.450		

Example: ANOVA results for Sewage Experiment (cont.)

- In the correct analysis (top panel), the total variation is decomposed into three components; the largest one is the block-to-block variation (MS $= 34.77$).
- In the unpaired analysis (bottom panel), this block component is mistakenly included in the error SS, thus making the F-test powerless.

Example: Nitrogen Fertilization

- Objective: find a way to refine the wheat production by evaluating the effect of several different fertilization timing and nitrogen rate schedules.
- **Treatment factor:** fertilization schedules. **Levels:** six different nitrogen application timing and rate schedules: A, B, C, D, E, F.
- Due to the different irrigation conditions, 24 available field plots need to be grouped to 4 blocks (6 plots each), such that each block has the same irrigation conditions.
- **Blocking factor:** plot water irrigation condition. **Levels:** four different water irrigation conditions: B1, B2, B3, B4.
- Randomization: (1) randomly allocate treatment levels within each block; (2) randomly assign units.

Example: Nitrogen Fertilization (cont.)

- Table for experimental outputs.

Data Table for Timing of Nitrogen Fertilization Experiment

	Treatment Levels					
Block Levels	A	B	C	D	E	F
B1	34.98	40.89	42.07	37.18	37.99	34.89
B2	41.22	46.69	49.42	45.85	41.99	50.15
B3	36.94	46.65	52.68	40.23	37.61	44.57
B4	39.97	41.90	42.91	39.20	40.45	43.29

The Model for RB Design

- The **additive** model for randomized block design:

$$y_{ij} = \eta + \alpha_i + \tau_j + e_{ij}$$
$$i = 1, \ldots, b, \quad j = 1, \ldots, k,$$

  where

  - y_{ij} = observation of the jth treatment in the ith block;
  - τ_j = jth treatment effect;
  - α_i = ith block effect;
  - e_{ij} = errors, i.i.d. $N(0, \sigma^2)$.

- Constraint conditions?

Model Estimation

- Model estimation:

$$\begin{aligned} y_{ij} &= \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}) \\ &= \hat{\eta} + \hat{\alpha}_i + \hat{\tau}_j + \hat{e}_{ij} \end{aligned}$$

where

$$\hat{\eta} = \bar{y}_{..}, \quad \hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} \quad \hat{\tau}_j = \bar{y}_{.j} - \bar{y}_{..}$$

$$\hat{e}_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$$

$$\bar{y}_{i.} = \frac{1}{k}\sum_{j=1}^{k} y_{ij} \quad \bar{y}_{.j} = \frac{1}{b}\sum_{i=1}^{b} y_{ij} \quad \bar{y}_{..} = \frac{1}{bk}\sum_{i=1}^{b}\sum_{j=1}^{k} y_{ij}$$

Sum of Squares

- Move $\bar{y}_{..}$ to the left hand side, squaring both sides and summing over i and j yield

$$\sum_{i=1}^{b}\sum_{j=1}^{k}(y_{ij}-\bar{y}_{..})^2 = k\sum_{i=1}^{b}(\bar{y}_{i.}-\bar{y}_{..})^2 + b\sum_{j=1}^{k}(\bar{y}_{.j}-\bar{y}_{..})^2 + \sum_{i=1}^{b}\sum_{j=1}^{k}(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}+\bar{y}_{..})^2$$

$$\text{or} \quad \mathrm{SST} = \mathrm{SSB} + \mathrm{SSTr} + \mathrm{SSE}$$

ANOVA Table

ANOVA Table for Randomized Complete Block Design

Source	D.F.	S.S.	M.S.
Blocks	$b-1$	SSB	$\text{MSB} = \frac{\text{SSB}}{b-1}$
Treatments	$k-1$	SSTr	$\text{MSTr} = \frac{\text{SSTr}}{k-1}$
Error	$(b-1)(k-1)$	SSE	$\text{MSE} = \frac{\text{SSE}}{(b-1)(k-1)}$
Total	$bk-1$	SST	

F Tests

- $H_0 : \tau_1 = \tau_2 = \ldots = \tau_k = 0$ can be tested by F statistic

$$F = \frac{\text{MSTr}}{\text{MSE}} \sim F_{k-1,(b-1)(k-1)} \quad \text{under } H_0,$$

  The F test rejects H_0 at level α if $F > F_{k-1,(b-1)(k-1),\alpha}$.

- The similar test can be applied for the block effect:
  $H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$ can be tested by F statistic

$$F_{\text{block}} = \frac{\text{MSB}}{\text{MSE}} \sim F_{b-1,(b-1)(k-1)} \quad \text{under } H_0,$$

  The F test rejects H_0 at level α if $F_{\text{block}} > F_{b-1,(b-1)(k-1),\alpha}$.

Multiple Comparisons

- If F test of $H_0 : \tau_1 = \ldots = \tau_k = 0$ is rejected, multiple comparisons for τ_i can be performed.
  The t statistics for making multiple comparisons are given by:

$$T_{ij} = \frac{\bar{y}_{\cdot j} - \bar{y}_{\cdot i}}{\hat{\sigma}\sqrt{1/b + 1/b}},$$

  where $\hat{\sigma}^2 = \text{MSE}$ from ANOVA table.
- If the number of tests equals k', then at level α, Bonferroni method declare "treatments j and i as different" if

$$|T_{ij}| > t_{(b-1)(k-1),\alpha/(2k')}.$$

Tukey Method and Simultaneous C.I.s

- At level α, the Tukey multiple comparison method identifies "treatments i and j as different" if
$$|T_{ij}| > \frac{1}{\sqrt{2}} q_{k,(b-1)(k-1),\alpha}.$$

- By solving
$$\frac{|\bar{y}_{\cdot j} - \bar{y}_{\cdot i} - (\tau_j - \tau_i)|}{\hat{\sigma}\sqrt{2/b}} \leq \frac{1}{\sqrt{2}} q_{k,(b-1)(k-1),\alpha}$$
for $\tau_j - \tau_i$, the **simultaneous** confidence intervals for $\tau_j - \tau_i$ are given by
$$\bar{y}_{\cdot j} - \bar{y}_{\cdot i} \pm q_{k,(b-1)(k-1),\alpha} \cdot \frac{\hat{\sigma}}{\sqrt{b}}$$
for all $\binom{k}{2}$ pairs of (i, j).

Analysis of the Nitrogen Fertilization Experiment: F-test

ANOVA Table: Nitrogen Fertilization Experiment

Source	D.F.	S.S.	M.S.	F
Block	3	197.004	65.668	9.119
Treatment	5	201.316	40.263	5.592
Error	15	108.008	7.201	
Total	23	506.328		

- The F statistic has the value:

$$F = \frac{\text{MSTr}}{\text{MSE}} = \frac{40.263}{7.201} = 5.5917.$$

Therefore, the p-value for testing the difference between treatment effects is $P(F_{5,15} > 5.5917) = 0.0042$, the small p-value suggests that the six fertilization schedules are different.

Analysis of the Nitrogen Fertilization Experiment: Multiple Comparisons

(B,A)	(C,A)	(C,B)	(D,A)	(D,B)	(D,C)	(E,A)	(E,B)
3.03	4.48	1.44	1.23	-1.80	-3.24	0.65	-2.38
(E,C)	(E,D)	(F,A)	(F,B)	(F,C)	(F,D)	(F,E)	
-3.83	-0.58	2.61	-0.43	-1.87	1.38	1.96	

- The means for six treatments are $A : 38.278$, $B : 44.033$, $C : 46.770$, $D : 40.615$, $E : 39.510$, $F : 43.225$.
- The multiple comparison t statistics are displayed in above table. For example, the t statistic for B versus A is $t_{2,1} = \dfrac{44.033 - 38.278}{\sqrt{7.201}\sqrt{2/4}} = 3.03$.
- With $\alpha = 0.05$, $t_{15,0.05/(15\times 2)} = 3.48$ for Bonferroni method, since $t = 6$ and $\binom{6}{2} = 15$. $\frac{1}{\sqrt{2}} q_{6,15,0.05} = \frac{4.59}{1.414} = 3.25$ for Tukey method. Again, Tukey method is more powerful (why?)

Contrast Inference

- Testing $H_0: \sum_{j=1}^{k} c_j \tau_j = 0$ by

$$T = \frac{\sum_{j=1}^{k} c_j \hat{\tau}_j}{\sqrt{\sum_{j=1}^{k} c_j^2 \hat{\sigma}^2 / b}} \sim t_{(b-1)(k-1)}$$

- Likewise, the contrast for block effects can be defined and statistic can be obtained accordingly.

Two-way Layout Experiment

- Now, we consider the situation that we have two treatment factors, i.e. A and B. This is similar to randomize complete block design (both have two factors). The only difference is that we are now further interested in the "factor $A \times$ factor B interaction" (explain).

- Asphalt bonding strength experiment:
  - Two factors [Aggregate type (A), Compaction method (B)] are compared for the bonding strength of the asphalt pavement mixes.
  - Factor A with 2 levels: silicious (A_1) and basalt (A_2); Factor B with 4 levels: static pressure (B_1), regular kneading (B_2), low kneading (B_3) and very low kneading (B_4).
  - For each level of combinations, the experiment is repeated three times (**3 replicates for each cell**).

Data for Asphalt Bonding Strength Experiment

Aggregate Type	Compaction Method			
		Kneading		
	Static	Regular	Low	Very Low
Basalt	68	126	93	56
	63	128	101	59
	65	133	98	57
Silicious	71	107	63	40
	66	110	60	41
	66	116	59	44

Cell Means Models

- Cell means model: Assume factor A has I levels, factor B has J levels,

$$\begin{aligned} &y_{ij\ell} = \mu_{ij} + e_{ij\ell} \\ &i = 1, \ldots, I \quad j = 1, \ldots, J \quad \ell = 1, \ldots, n \end{aligned}$$

  - μ_{ij} is the mean of the treatment combination A at level i and B at level j.
  - $e_{ij\ell}$ are i.i.d. $N(0, \sigma^2)$.
  - Let $\bar{\mu}_{i\cdot} = \frac{1}{J}\sum_{j=1}^{J} \mu_{ij}$ and $\bar{\mu}_{\cdot j} = \frac{1}{I}\sum_{i=1}^{I} \mu_{ij}$.

Effects Models

- Effects model:

$$y_{ij\ell} = \eta + \alpha_i + \beta_j + \omega_{ij} + e_{ij\ell},$$

  - $\eta = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} \mu_{ij}$ is the grand mean;
  - $\alpha_i = \frac{1}{J} \sum_{j=1}^{J} \mu_{ij} - \eta$ and $\beta_j = \frac{1}{I} \sum_{i=1}^{I} \mu_{ij} - \eta$ are the effect of the ith level of factor A and the effect of the jth level of factor B, respectively;
  - $\omega_{ij} = \mu_{ij} - \bar{\mu}_{i\cdot} - \bar{\mu}_{\cdot j} + \eta$ is the interaction effect between the ith level of A and the jth level of B.
  - What kind of constraints should we impose on α_i, β_j and ω_{ij}?

Effects Models (cont.)

- Effects model:

$$y_{ij\ell} = \eta + \alpha_i + \beta_j + \omega_{ij} + e_{ij\ell},$$

- For the model identification purpose and the nature of their definitions, we need
  - $\sum_{i=1}^{I} \alpha_i = 0;$
  - $\sum_{j=1}^{J} \beta_j = 0;$
  - $\sum_{i=1}^{I} \omega_{i,j} = \sum_{j=1}^{J} \omega_{i,j} = 0.$

Model Estimation

- Estimation by decomposition:

$$
\begin{aligned}
y_{ij\ell} &= \bar{y}_{\cdot\cdot\cdot} + (\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot}) + (\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot}) \\
&\quad + (\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot}) + (y_{ij\ell} - \bar{y}_{ij\cdot}), \\
&= \hat{\eta} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\omega}_{ij} + \hat{e}_{ij\ell},
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{\eta} &= \bar{y}_{\cdot\cdot\cdot} \\
\hat{\alpha}_i &= \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot} \\
\hat{\beta}_j &= \bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot} \\
\hat{\omega}_{ij} &= \bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdot\cdot\cdot} \\
\hat{e}_{ij\ell} &= y_{ij\ell} - \bar{y}_{ij\cdot}
\end{aligned}
$$

ANOVA Table

Source	D.F.	S.S.	M.S.
Factor A	$I-1$	SSA	$\text{MSA} = \frac{\text{SSA}}{I-1}$
Factor B	$J-1$	SSB	$\text{MSB} = \frac{\text{SSB}}{J-1}$
Interaction	$(I-1)(J-1)$	SSAB	$\text{MSAB} = \frac{\text{SS(AB)}}{(I-1)(J-1)}$
Residual	$IJ(n-1)$	SSE	$\text{MSE} = \frac{\text{SSE}}{IJ(n-1)}$
Total	$IJn-1$	SST	

$$\text{SST} = \sum_i \sum_j \sum_\ell (y_{ij\ell} - \bar{y}_{\cdots})^2 \quad \text{SSE} = \sum_i \sum_j \sum_\ell (y_{ij\ell} - \bar{y}_{ij\cdot})^2$$

$$\text{SSA} = nJ \sum_i (\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots})^2 \quad \text{SSB} = nI \sum_j (\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots})^2$$

$$\text{SSAB} = n \sum_i \sum_j (\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdots})^2$$

F Tests

Consider the following hypothesis tests:

(1) $H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_I = 0$, then

$$F_A = \frac{\text{MSA}}{\text{MSE}} \sim F_{I-1,\ IJ(n-1)}, \text{ under } H_0.$$

(2) $H_0 : \beta_1 = \beta_2 = \ldots = \beta_J = 0$, then

$$F_B = \frac{\text{MSB}}{\text{MSE}} \sim F_{J-1,\ IJ(n-1)}, \text{ under } H_0.$$

(3) $H_0 : \omega_{11} = \omega_{12} = \ldots = \omega_{IJ} = 0$, then

$$F_{AB} = \frac{\text{MSAB}}{\text{MSE}} \sim F_{(I-1)(J-1),\ IJ(n-1)}, \text{ under } H_0.$$

Analysis of Asphalt Bonding Strength Experiment

Source	D.F.	S.S.	M.S.	F
Factor A	1	1,734.00	1734.00	182.53
Factor B	3	16,243.50	5414.50	569.95
Interaction AB	3	1,145.00	381.67	40.18
Error	16	152.00	9.50	
Total	23	19,274.50		

- Conclusions: Both factors and their interactions are significant.
- Multiple comparisons and contrast tests are argued in the next slide.

Multiple Comparisons and Contrast Tests

- If interactions are significant, we don't bother to construct tests or C.I.'s for linear combinations of effects (why?). Instead, we refer back to our cell mean model. All inference techniques in single factor experiments can be applied.
- If the interactions are not significant, then the model is reduced to additive model. All discussions in additive model of randomized block design can apply.

Multiple-Way Layout Model

- Three factor experiment models:

$$\begin{aligned} y_{ijk\ell} &= \mu_{ijk} + e_{ijk\ell} \\ &= \eta + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} \\ &\quad + (\alpha\beta\gamma)_{ijk} + e_{ijk\ell} \end{aligned}$$

  where $i = 1, \ldots, I$, $j = 1, \ldots, J$, $k = 1, \ldots, K$, $\ell = 1, \ldots, n$.

- Questions: parameter explanation, constraint conditions, ANOVA table, tests of hypothesis?

Latin Square Design: Wheat Yield Experiment

- **Wheat yield experiment:** determine the optimum seeding rate for a newly introduced durum wheat. "$y =$ The gain yield for each plot in hundred-pound (100lb)".
- One treatment factor: seeding rates ("A, B, C, D, E" stand for "30,80,130,180,230 lb/acre" respectively).
- Two blocking factors: (1). irrigation gradient (r1,r2,r3,r4,r5); (2). soil gradient (c1,c2,c3,c4,c5).
- **Latin square design of order k:** In a $k \times k$ matrix, "row" and "column" correspond to the two block factors. Each of the k Latin Letters (i.e.: treatments) appear once in each row and each column.
- It is an extension of randomized complete block design to accommodate **two** blocking factors. Randomization applied to assignments to rows, columns, treatments. (Collection of Latin Square Tables given in Appendix 3A of our textbook).

Wheat Yield Experiment: Design and Data

Latin Square Design, Wheat Yield Experiment

Irr. Grad.	Soil Gradient				
	c1	c2	c3	c4	c5
r1	E	A	C	B	D
r2	C	D	B	E	A
r3	B	C	D	A	E
r4	A	B	E	D	C
r5	D	E	A	C	B

Wheat Yield Data

Irr. Grad.	Soil Gradient				
	c1	c2	c3	c4	c5
r1	59.45	47.28	54.44	50.14	59.45
r2	55.16	60.89	56.59	60.17	48.71
r3	44.41	53.72	55.87	47.99	59.45
r4	42.26	50.14	55.87	58.74	55.87
r5	60.89	59.45	49.43	59.45	57.31

Statistical Model for Latin Square Design

- Model:

$$y_{ij\ell} = \eta + \alpha_i + \beta_j + \tau_\ell + e_{ij\ell}$$
$$i, j = 1, 2, \ldots, k$$

  - $\ell =$ Latin letter in the (i, j) cell of the Latin Square.
  - α_i and β_j are the row and column block effects respectively. $\sum_i \alpha_i = \sum_j \beta_j = \sum_\ell \tau_\ell = 0.$
  - τ_ℓ is the ℓth treatment (i.e., Latin letter) effect.
  - $e_{ij\ell}$ are i.i.d. $N(0, \sigma^2)$.

Parameter Estimation

- Note that we only have k^2 observations for the triplet (i, j, ℓ) determined by the particular Latin Square; We denote this set by S.
- Similar to previous, the estimation can be obtained as follows

$$
\begin{aligned}
y_{ij\ell} &= \bar{y}_{\dots} + (\bar{y}_{i..} - \bar{y}_{\dots}) + (\bar{y}_{.j.} - \bar{y}_{\dots}) + (\bar{y}_{..\ell} - \bar{y}_{\dots}) \\
&\quad + (y_{ij\ell} - \bar{y}_{i..} - \bar{y}_{.j.} - \bar{y}_{..\ell} + 2\bar{y}_{\dots}) \\
&= \hat{\eta} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\tau}_\ell + \hat{e}_{ij\ell}.
\end{aligned}
$$

Sum of Squares

- Sum of squares:

$$\mathrm{SST} = \sum_{(ij\ell)\in S} (y_{ij\ell} - \bar{y}_{\dots})^2 \qquad \mathrm{SS}_{\mathrm{trt}} = k\sum_{\ell}(\bar{y}_{\cdot\cdot\ell} - \bar{y}_{\dots})^2$$

$$\mathrm{SS}_{\mathrm{row}} = k\sum_{i}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\dots})^2 \qquad \mathrm{SS}_{\mathrm{col}} = k\sum_{j}(\bar{y}_{\cdot j\cdot} - \bar{y}_{\dots})^2$$

$$\mathrm{SSE} = \sum_{(ij\ell)\in S} (y_{ij\ell} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\ell} + 2\bar{y}_{\dots})^2$$

$$\mathrm{SST} = \mathrm{SS}_{\mathrm{row}} + \mathrm{SS}_{\mathrm{col}} + \mathrm{SS}_{\mathrm{trt}} + \mathrm{SSE}$$

- S is the set of all possible values for the triplet (i, j, ℓ).

ANOVA Table for Latin Square Design

Source	D.F.	Sum of Squares
row	$k-1$	SS_{row}
column	$k-1$	SS_{col}
treatment	$k-1$	SS_{trt}
Error	$(k-1)(k-2)$	SSE
Total	k^2-1	SST

ANOVA Table for Wheat Yield Experiment

Source	D.F.	Sum of Squares	Mean Square	F
row (irri. grad.)	4	99.20	24.80	5.25
column (soil grad.)	4	38.48	9.62	2.04
trt (seed rate)	4	522.30	130.57	27.67
Error	12	56.63	4.72	
Total	24	716.61		

F Test and Multiple Comparisons

- $H_0 : \tau_1 = \tau_2 = \ldots = \tau_k = 0$ can be tested by F statistic

$$F = \frac{\mathrm{SS}_{\mathrm{trt}}/(k-1)}{\mathrm{SSE}/\{(k-1)(k-2)\}} \sim F_{k-1,(k-1)(k-2)}.$$

- If H_0 is rejected, multiple comparisons of the τ_j can be performed. Similar to previous lectures, t statistic for making multiple comparisons is

$$T_{ij} = \frac{\bar{y}_{\cdot\cdot i} - \bar{y}_{\cdot\cdot j}}{\hat{\sigma}\sqrt{1/k + 1/k}} \quad \text{with } \hat{\sigma} = \sqrt{\mathrm{MSE}},\ 1 \leq i < j \leq k.$$

- At level α, the Tukey multiple comparison method identifies "treatment i and j as different" if

$$|T_{ij}| > \frac{1}{\sqrt{2}} q_{k,(k-1)(k-2),\alpha}.$$

Results for Wheat Yield Experiment

- The p-values for "Irrigation Gradient" and "Soil Gradient" are 0.011 [$=P(F_{4,12} > 5.25)$] and 0.152 [$=P(F_{4,12} > 2.04)$], which indicates that the "Irrigation Gradient" blocking is significant, but "Soil Gradient" is not significant.
- The treatment factor (seed rate) has p-value 5.62e-06 [$=P(F_{4,12} > 27.67)$] indicates significant treatment effect.

Results for Wheat Yield Experiment (cont.)

- Since $k = 5$, $(k-1)(k-2) = 12$, the critical value for the Tukey method is $\frac{1}{\sqrt{2}} q_{5,12,0.05} = \frac{4.199}{\sqrt{2}} = 2.97$ at level 0.05.
- Compare the multiple comparison t statistic presented in the table of next slide with 2.97, seed rate "B vs A", "C vs A", "D vs A", "D vs B", "E vs A", "E vs B" are identified as significantly different at level 0.05.

BA	CA	CB	DA	DB	DC	EA	EB	EC	ED
3.34	6.26	2.92	8.76	5.42	2.50	8.55	5.21	2.29	-0.21

Results for Wheat Yield Experiment (cont.)

Original ANOVA Table

Source	D.F.	Sum of Squares	Mean Square	F
row (irri. grad.)	4	99.20	24.80	5.25
column (soil grad.)	4	38.48	9.62	2.04
trt (seed rate)	4	522.30	130.57	27.67
Error	12	56.63	4.72	
Total	24	716.61		

ANOVA Table, Ignoring Soil Gradient

Source	D.F.	S.S.	M.S.	F	p-value
row (irri. grad.)	4	99.20	24.80	4.36	0.014
trt (seed rate)	4	522.30	130.57	22.95	1.84e-06
Error	16	91.11	5.69		
Total	24	716.61			

Results for Wheat Yield Experiment (cont.)

ANOVA Table, Ignoring Both Block Factors

Source	D.F.	S.S.	M.S.	F	p-value
trt (seed rate)	4	522.30	130.57	13.72	1.54e-05
Error	20	190.31	9.52		
Total	24	716.61			

Effectiveness of blocking. We consider the p-values of treatment effects: With two blocking factors p-value = 5.62e-06; ignoring the less significant blocking (Soil Gradient) p-value = 1.84e-06; ignoring the significant blocking (Irrigation Gradient) p-value = 1.54e-05. Therefore, significant blocking factors play important roles in decision.

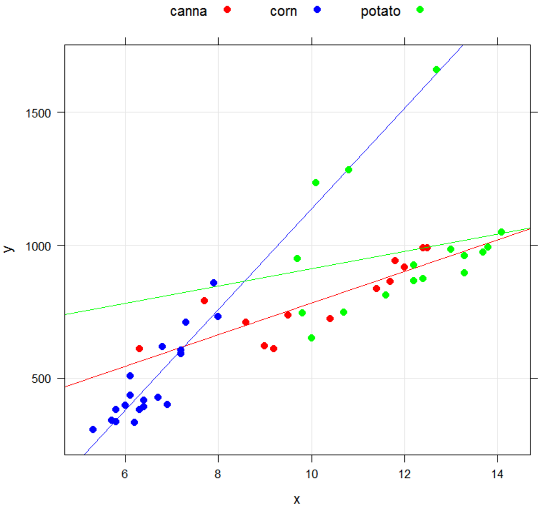
Analysis of Covariance: Starch Experiment

- Goal: To compare the three treatments (canna, corn, potato) for making starch film; Response y = break strength of film; continuous covariate x = film thickness.
- Known that x affects y (thicker films are stronger); thickness x cannot be controlled but is measurable after films are made.
- Question: How to perform treatment comparisons by incorporating the effect of the covariate x?

Starch Experiment Data

Canna		Corn		Potato	
y	x	y	x	y	x
791.7.	7.7	731.0	8.0	983.3	13.0
610.0	6.3	710.0	7.3	958.8	13.3
710.0	8.6	604.7	7.2	747.8	10.7
940.7	11.8	508.8	6.1	866.0	12.2
990.0	12.4	393.0	6.4	810.8	11.6
916.2	12.0	416.0	6.4	950.0	9.7
835.0	11.4	400.0	6.9	1282.0	10.8
724.3	10.4	335.6	5.8	1233.8	10.1
611.1	9.2	306.4	5.3	1660.0	12.7
621.7	9.0	426.0	6.7	746.0	9.8
735.4	9.5	382.5	5.8	650.0	10.0
990.0	12.5	340.8	5.7	992.5	13.8
862.7	11.7	436.7	6.1	896.7	13.3
		333.3	6.2	873.9	12.4
		382.3	6.3	924.4	12.2
		397.7	6.0	1050.0	14.1
		619.1	6.8	973.3	13.7
		857.3	7.9		
		592.5	7.2		

Starch Experiment Data



The Model

- Model:

$$y_{ij} = \eta + \tau_i + \gamma x_{ij} + e_{ij}, \tag{1}$$

  where
  - $i = 1, \ldots, k$, $j = 1, \ldots, n_i$;
  - τ_i: ith treatment effect;
  - x_{ij}: covariate value;
  - γ: regression coefficient for x_{ij};
  - e_{ij}: independent $N(0, \sigma^2)$.
- Special cases:
  - When $\gamma x_{ij} = 0$ (i.e. x_{ij} not available or no x effect), it is a one-way layout model.
  - When $\tau_i = 0$ (no treatment effect), it is a simple linear regression model.
- Model (1) can be easily extended to experiments with more than one factors or/and more than one covariates.

Regression Model Approach: Starch Experiment

- If we assume baseline constraint condition $\tau_1 = 0$, Model (1) can be rewritten as

$$
\begin{aligned}
y_{1j} &= \eta + \gamma x_{1j} + e_{1j}, \quad j = 1, \ldots, 13, i = 1 \text{ canna} \\
y_{2j} &= \eta + \tau_2 + \gamma x_{2j} + e_{2j}, \quad j = 1, \ldots, 19, i = 2 \text{ corn} \\
y_{3j} &= \eta + \tau_3 + \gamma x_{3j} + e_{3j}, \quad j = 1, \ldots, 17, i = 3 \text{ potato} \qquad (2)
\end{aligned}
$$

  where
  - η: intercept.
  - γ: regression coefficient for thickness.
  - τ_2: corn vs. canna.
  - τ_3: potato vs. canna.
- Questions:
  - Model matrix, $\mathbf{X}$, and model parameters, $\boldsymbol{\beta}$, for (2)?
  - We don't consider the interaction between x and treatments. Is it possible to consider? If yes, what does it mean? (explain)

Regression Analysis of Starch Experiment

- Regression analysis can be performed in the usual way.
- Coefficients table:

Effect	Estimate	S.E.	t	p-value
intercept	158.261	179.775	0.88	0.38
thickness	62.501	17.060	3.66	0.00
corn vs. canna	-83.666	86.095	-0.97	0.34
potato vs. canna	70.360	67.781	1.04	0.30
potato vs. corn	154.026	107.762	1.43	0.16

where, potato vs. corn $=$
$\hat{\tau}_3 - \hat{\tau}_2 = 70.360 - (-83.666) = 154.026.$

- No pair of film types has any significant difference after adjusting for thickness effect. (So, how should the choice be made between the three film types?) Most of the variation is explained by the covariate thickness.

Statistical Inference on Model Parameters

- Statistical inference on $\boldsymbol{\beta}$ can be based upon the fact from multiple linear regression:

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

- For instance, if we would like to make inference on $\tau_3 - \tau_2$, the key issue is $Var(\hat{\tau}_3 - \hat{\tau}_2)$, which can be evaluated as

$$\mathrm{Var}(\hat{\tau}_3 - \hat{\tau}_2) = \mathrm{Var}(\hat{\tau}_3) + \mathrm{Var}(\hat{\tau}_2) - 2\,\mathrm{Cov}(\hat{\tau}_3, \hat{\tau}_2),$$

  where $\mathrm{Var}(\hat{\tau}_3)$, $\mathrm{Var}(\hat{\tau}_2)$ and $\mathrm{Cov}(\hat{\tau}_3, \hat{\tau}_2)$ can be read from the according elements of $\hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}$ with $\hat{\sigma}^2 = \mathrm{MSE}$ in ANCOVA table.

Testing the Treatment Effect

Suppose we want to test the treatment effect, i.e.
$H_0 : \tau_1 = \ldots = \tau_k = 0$. We can compare the following two models:

$$y_{ij} = \eta + \gamma x_{ij} + e_{ij} \quad \cdots (\mathrm{I})$$
$$y_{ij} = \eta + \tau_i + \gamma x_{ij} + e_{ij} \quad \cdots (\mathrm{II})$$

- Model I is called the reduced model. Model II is called the full model.
- The F test for $H_0 : \tau_1 = \ldots = \tau_k = 0$ is given by

$$F_{\mathrm{trt}} = \frac{[\mathrm{SSE(I)} - \mathrm{SSE(II)}]/(k-1)}{\mathrm{MSE(II)}} \sim F_{k-1,N-k-1},$$

where SSE(I) and SSE(II) are the error sum of squares under the reduced model (Model I) and the full model (Model II), respectively, and MSE(II) is the mean squared error under the full model (Model II).

ANCOVA Table for Treatment Effect

Source	df	S.S.	M.S.	F
Covariate(x)	1	SS_x	MS_x	$F_x = \frac{MS_x}{MSE(II)}$
Treatment	$k-1$	SS_{trt}	MS_{trt}	$F_{trt} = \frac{MS_{trt}}{MSE(II)}$
Error	$N-k-1$	SSE(II)	MSE(II)	
Total	$N-1$	SST		

- In the table, $SS_{trt} = SSE(I) - SSE(II)$.
- Alternatively, if SSReg(I) and SSReg(II) are the regression sum of squares for the reduced model (Model I) and the full model (Model II) respectively, then we also have $SS_{trt} = SSReg(II) - SSReg(I)$.
- $SS_x = SSReg(I)$.
- The F test in this table for x (using F_x) should not be used for testing the significance of the covariate x (why?)

Testing the Covariate Effect

Suppose we want to test the covariate effect, i.e. $H_0 : \gamma = 0$. We can compare the following two models:

$$y_{ij} = \eta + \tau_i + e_{ij} \quad \cdots \text{(III)}$$
$$y_{ij} = \eta + \tau_i + \gamma x_{ij} + e_{ij} \quad \cdots \text{(II)}$$

- Model III is the new reduced model (for testing the covariate effect). Model II is the full model.
- The F test for $H_0 : \gamma = 0$ is given by

$$F'_x = \frac{[\text{SSE(III)} - \text{SSE(II)}]/1}{\text{MSE(II)}} \sim F_{1,N-k-1},$$

where SSE(III) and SSE(II) are the error sum of squares under the new reduced model (Model III) and the full model (Model II), respectively, and MSE(II) is the mean squared error under the full model (Model II).

ANCOVA Table for Covariate

Source	df	S.S.	M.S.	F
Treatment	$k-1$	$\mathrm{SS}'_{\mathrm{trt}}$	$\mathrm{MS}'_{\mathrm{trt}}$	$F'_{\mathrm{trt}} = \frac{\mathrm{MS}'_{\mathrm{trt}}}{\mathrm{MSE(II)}}$
Covariate(x)	1	SS'_x	MS'_x	$F'_x = \frac{\mathrm{MS}'_x}{\mathrm{MSE(II)}}$
Error	$N-k-1$	SSE(II)	MSE(II)	
Total	$N-1$	SST		

- In the table, $\mathrm{SS}'_x = \mathrm{SSE(III)} - \mathrm{SSE(II)}$.
- Alternatively, if SSReg(III) and SSReg(II) are the regression sum of squares for the reduced model (Model III) and the full model (Model II) respectively, then we also have $\mathrm{SS}'_x = \mathrm{SSReg(II)} - \mathrm{SSReg(III)}$.
- $\mathrm{SS}'_{\mathrm{trt}} = \mathrm{SSReg(III)}$.
- The order of the covariate and the treatment matters! $\mathrm{SS}'_x \neq \mathrm{SS}_x$, $\mathrm{SS}'_{\mathrm{trt}} \neq \mathrm{SS}_{\mathrm{trt}}$, $F'_{\mathrm{trt}} \neq F_{\mathrm{trt}}$, $F'_x \neq F_x$.

ANCOVA Table: Starch Experiment

If we test $H_0 : \tau_1 = \tau_2 = \tau_3$, the ANCOVA table is

Source	df	SS	MS	F
Thickness (x)	1	2553357	2553357	94.19
Starch (trt)	2	56725	28362	1.05
Error	45	1219940	27110	
Total	48	3830022		

$F_{\text{starch}} = 1.05$ with p-value 0.3597. Therefore we conclude that there is no difference between the three starch types.

ANCOVA Table: Starch Experiment

If we test $H_0 : \gamma = 0$, the ANCOVA table is

Source	df	SS	MS	F
Starch (trt)	2	2246204	1123102	41.428
Thickness (x)	1	363878	363878	13.422
Error	45	1219940	27110	
Total	48	3830022		

$F'_x = 13.422$ with p-value 0.00065. Therefore we conclude that the film thickness has significant impact on the break strength of the film.

Response Transformation

- Purpose: to make the resulted model roughly satisfy the equal variance assumption.
- Transform y before fitting a regression model.
- The most popularly used one is the power (Box-Cox) transformation:

$$z = f(y) = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log y, & \lambda = 0. \end{cases}$$

- A simple method for determining λ is to try a few selected values of λ (see the table in the next slide). In each transformation, analyze the z-data and choose the transformation (i.e., λ value) such that
  (a). it gives a parsimonious model.
  (b). no unusual patterns in the residual plots.
  (c). good interpretation on the transformation.

Variance Stabilizing Transformations

$\sigma_y \propto \mu^\alpha$	α	$\lambda = 1 - \alpha$	Transformation
$\sigma_y \propto \mu^3$	3	-2	reciprocal squared
$\sigma_y \propto \mu^2$	2	-1	reciprocal
$\sigma_y \propto \mu^{3/2}$	3/2	-1/2	reciprocal square root
$\sigma_y \propto \mu$	1	0	log
$\sigma_y \propto \mu^{1/2}$	1/2	1/2	square root
$\sigma_y \propto$ constant	0	1	no transformation
$\sigma_y \propto \mu^{-1/2}$	-1/2	3/2	3/2 power
$\sigma_y \propto \mu^{-1}$	-1	2	square