Chapter 6. Experiments with One Factor

An Introduction to Experimental Design

- Statistical studies can be classified as being either experimental or observational.
- In an experimental study, one or more factors are controlled so that data can be obtained about how the factors influence the variables of interest.
- In an observational study, no attempt is made to control the factors.
- Cause-and-effect relationships are easier to establish in experimental studies than in observational studies.

Basic Concepts

- A factor is a variable that the experimenter has selected for investigation.
- A treatment is a level of a factor.
- Experimental units are the objects of interest in the experiment.
- A completely randomized design is an experimental design in which the treatments are randomly assigned to the experimental units.

Completely Randomized Design

Suppose we have 4 different diets which we want to compare. The diets are labeled Diet A, Diet B, Diet C, and Diet D. We are interested in how the diets affect the coagulation rates of rabbits. The coagulation rate is the time in seconds that it takes for a cut to stop bleeding. We have 16 rabbits available for the experiment, so we will use 4 on each diet. How should we use randomization to assign the rabbits to the four treatment groups? The 16 rabbits arrive and are placed in a large compound until you are ready to begin the experiment, at which time they will be transferred to cages.

Possible Assignment Plans

Method 1: We assume that rabbits will be caught "at random". Catch four rabbits and assign them to Diet A. Catch the next four rabbits and assign them to Diet B. Continue with Diets C and D. Since the rabbits were "caught at random", this would produce a completely randomized design. Analyze the results as a completely randomized design.

Method 2: Catch all the rabbits and label them 1-16. Select four numbers 1-16 at random (without replacement) and put them in a cage to receive Diet A. Then select another four numbers at random and put them in a cage to receive Diet B. Continue until you have four cages with four rabbits each. Each cage receives a different diet, and the experiment is analyzed as a completely randomized experiment.

Method 3: Have a bowl with the letters A, B, C, and D printed on separate slips of paper. Catch the first rabbit, pick a slip at random from the bowl and assign the rabbit to the diet letter on the slip. Do not replace the slip. Catch the second rabbit and select another slip from the remaining three slips. Assign that diet to the second rabbit. Continue until the first four rabbits are assigned one of the four diets. In this way, all of the slow rabbits have different diets. Replace the slips and repeat the procedure until all 16 rabbits are assigned to a diet. Analyze the results as a completely randomized design.

Method 4: Catch all the rabbits and label them 1-16. Put 16 slips of paper in a bowl, four each with the letters A, B, C, and D. Put another 16 slips of paper numbered 1-16 in a second bowl. Pick a slip from each bowl. The rabbit with the selected number is given the selected diet. To make it easy to remember which rabbit gets which diet, the cages are arranged as shown below.

A	A	A	A
B	B	B	B
C	C	C	C
D	D	D	D

A Completely Randomized Design

Label the cages 1-16. In a bowl put 16 strips of paper each with one of the integers 1-16 written on it. In a second bowl put 16 strips of paper, four each labeled A, B, C, and D. Catch a rabbit. Select a number and a letter from each bowl. Place the rabbit in the location indicated by the number and feed it the diet assigned by the letter. Repeat without replacement until all rabbits have been assigned a diet and cage.

1	5	9	13
2	6	10	14
3	7	11	15
4	8	12	16

1	5	9	13
2	6	10	14
3	7 B	11	15
4	8	12	16

1 C	5 A	9 B	13 D
2 D	6 B	10 D	14 C
3 C	7 B	11 A	15 D
4 A	8 A	12 C	16 B

One-Way ANOVA Table

One-Way ANOVA

To analyze the results of the experiment, we use a one-way analysis of variance. The measured coagulation times for each diet are given below:

	Diet A	Diet B	Diet C	Diet D
	62	63	68	56
	60	67	66	62
	63	71	71	60
	59	64	67	61
Mean	61	66.25	68	59.75

The null hypothesis is

$$H_0: \mu_A = \mu_B = \mu_C = \mu_D \text{ (all treatment means the same)}$$

and the alternative is

$$H_a: \text{ at least one mean different.}$$

One Factor Experiment

Trt.	Observations	Sample Mean	Sample Variance	DF
1	$Y_{11}, Y_{12}, \ldots, Y_{1n_1}$	$\overline{Y}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1} Y_{1j}$	$s_1^2 = \frac{1}{n_1 - 1}\sum_{j=1}^{n_1}\left(Y_{1j} - \overline{Y}_1\right)^2$	$n_1 - 1$
$\vdots$				
i	$Y_{i1}, Y_{i2}, \ldots, Y_{in_i}$	$\overline{Y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}$	$s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}_i\right)^2$	$n_i - 1$
$\vdots$				
k	$Y_{k1}, Y_{k2}, \ldots, Y_{kn_k}$	$\overline{Y}_k = \frac{1}{n_k}\sum_{j=1}^{n_k} Y_{ij}$	$s_k^2 = \frac{1}{n_k - 1}\sum_{j=1}^{n_k}\left(Y_{ij} - \overline{Y}_i\right)^2$	$n_k - 1$

Section 6. 1 Completely Randomized One-factor Design

- Model (with k treatments or k factor levels)

$$Y_{ij} = \mu_i + \varepsilon_{ij}, i = 1,\dots,k;\ j = 1,\dots,n_i.$$

$$\Leftrightarrow Y_{ij} = \mu + \tau_i + \varepsilon_{ij},$$

where iid error $\varepsilon_{ij} \sim N(0, \sigma^2), N = \sum_{i=1}^{k} n_i$.

Y_{ij} : response of j-th replicate at i-th treatment level

μ_i : mean of observations at i-th treatment (τ_i : treatment effect)

k : number of treatment levels

n_i : number of observations at i-th treatment level

$N = \sum_{i=1}^{k} n_i$: total number of observations

Least square estimator for $\mu = (\mu_1, \ldots, \mu_k)'$:

$$\hat{\mu} = (X'X)^{-1} X'Y = \begin{pmatrix} \bar{Y}_1 \\ \vdots \\ \bar{Y}_k \end{pmatrix} \text{ or } \hat{\mu}_i = \bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$$

- Sum Square Decomposition: *SSTO = SSTR + SSE*

$$SSTO = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \bar{Y}\right)^2, \bar{Y} = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n_i} Y_{ij}$$

$$SSTR = \sum_{i=1}^{k} n_i \left(\bar{Y}_i - \bar{Y}\right)^2, \bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}, i = 1, \ldots, k$$

$$SSE = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \bar{Y}_i\right)^2 = \sum_{i=1}^{k} (n_i - 1) s_i^2, s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} \left(Y_{ij} - \bar{Y}_i\right)^2$$

Testing main effects

- Null hypothesis to compare k treatment effects:

$$H_0 : \mu_1 = \ldots = \mu_k \Leftrightarrow H_0 : \tau_1 = \ldots = \tau_k = 0$$

given that $\sum_{i=1}^{k} n_i \tau_i = 0$. It can be shown that

$$E(MSE) = \sigma^2, E(MSTR) = \sigma^2 + \frac{1}{k-1} \sum_{i=1}^{k} n_i \tau_i^{\,2},$$

Under H_0, $E(MSTR) = E(MSE)$, and the F-statistic is

$$F = \frac{MSTR}{MSE} = \frac{SSTR/(k-1)}{SSE/(N-k)} \sim F(k-1, N-k)$$

- Reject H_0 (treatment effects are significantly different) when p-value is less than given level α.

or reject H_0 if $F > F_\alpha(k\text{-}1, N\text{-}k)$

Completely Randomized Design

- Analysis of Variance Table (ANOVA)

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F
Treatments	SSTR	k - 1	$MSTR = \dfrac{SSTR}{k-1}$	$F = \dfrac{MSTR}{MSE}$
Error	SSE	N - k	$MSE = \dfrac{SSE}{N-k}$	
Total	SST	N - 1		

Example - Etch Rate and RF Experiment

An engineer is interested in investigating the relationship between the RF power setting and the etch rate for his tool.

The objective of an experiment like this is to model the relationship between etch rate and RF power and to specify the power setting that will give a desired target etch rate. He wants to test four levels of RF power: 160W, 180W, 200W, and 220W. He decided to test five wafers at each level of RF power.

Power (W)	Observations 1	2	3	4	5	Total	Averages
160	575	542	530	539	570		
180	565	593	590	578	610		
200	600	651	610	637	629		
220	725	700	715	685	710		

SAS code (with boxplot) and R code

SAS Code - ANOVA

```
ods html;   /* Output Delivery System */

data ratedata;
     input Power Rate @@;
     datalines ;
160  575
160  542
.
.
.
220  685
220  710
;

proc print data=ratedata;
run;

proc anova data=ratedata;
    class Power;                    /* Specify factor(s) */
    model Rate=Power;
run;

title 'Box Plot for Etch Data';

proc boxplot data=ratedata;
  plot Rate * Power ;           /* Compare Boxplots at different power levels  */
run;

ods html close;
```

Section 6.2 Inferences in One-Factor Experiments

- ANVOA model $Y_{ij} = \mu_i + \varepsilon_{ij}, \varepsilon_{ij} \sim^{i.i.d.} N(0, \sigma^2)$

$$\Leftrightarrow Y_{ij} = 0 \cdot \mu_1 + ... + 1 \cdot \mu_i + ... + 0 \cdot \mu_k + \varepsilon_{ij}$$

$$Y_{ij} = e_i' \mu + \varepsilon_{ij}, i = 1, ..., k; \ j = 1, ..., n_i$$

$$Y = X\mu + \varepsilon$$

Least Square Estimator : $\hat{\mu} = (X'X)^{-1} X'Y, \hat{\mu}_i = \bar{Y}_i = \frac{1}{n_i} \sum_{i=1}^{n_i} Y_{ij}$

Reference Distribution :

$$\bar{Y}_i \sim N\left(\mu_i, \frac{\sigma^2}{n_i}\right) \Rightarrow \frac{\bar{Y}_i - \mu_i}{\sqrt{\frac{MSE}{n_i}}} \sim t(N-k),, i = 1, ..., k$$

- 100(1-α)% Confidence Interval of μ_i : $\bar{Y}_i \pm t_{\frac{\alpha}{2}}(N-k) \cdot \sqrt{\frac{MSE}{n_i}}$

Pairwise Comparison $H_0: \mu_i = \mu_j$ vs $H_1: \mu_i \neq \mu_j$

- 100(1-α)% Confidence Interval of $(\mu_i - \mu_j)$:

$$\left(\bar{Y}_i - \bar{Y}_j\right) \pm t_{\frac{\alpha}{2}}(N-k) \cdot \sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

- Fisher's Least Significant Difference (LSD)
- For a balanced design ($n_1 = n_2 = \ldots = n_k = n$),

$$LSD = t_{\frac{\alpha}{2}}(N-k) \cdot \sqrt{\frac{2 \cdot MSE}{n}}$$

- Reject H_0 if $|\bar{Y}_i - \bar{Y}_j| > LSD = t_{\frac{\alpha}{2}}(N-k) \cdot \sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$

Bonferroni Test (Multiple Pairwise Comparison)

- Test H_0: $\mu_i = \mu_j$ for any pair (I, j) $_{vs}$ H_1: at least one pair is not the same
- Additive Law $$\alpha = P\left(\bigcup_{j=1}^{m} R_j\right) \le \sum_{j=1}^{m} P(R_j), \alpha^* = \frac{\alpha}{m}$$
- A factor has k levels, # of pairwise comparisons is k(k-1)/2
- Bonferroni Confidence Interval for $(\mu_i - \mu_j)$

$$\left(\overline{Y}_i - \overline{Y}_j\right) \pm t_{\frac{\alpha*}{2}}(N-k) \cdot \sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}, \alpha* = \frac{\alpha}{k(k-1)/2}, \forall i, j = 1, \ldots, k$$

- Reject H_0 if there is at least one C.I. doesn't include 0.

Tukey's Multiple Comparison of Pairwise Difference

- Test H_0: $\mu_i = \mu_j$ for any pair (I, j) $_{vs}$ H_1: at least one pair is not the same
- Studentized range statistic

$$T_u = \max_{1 \le i, j \le k} \left| \frac{(\bar{Y}_i - \mu_i) - (\bar{Y}_j - \mu_j)}{\sqrt{\frac{MSE}{2}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}} \right| \sim q(k, N-k)$$

- Under H_0, for balanced design

$$T_u = \frac{\bar{Y}_{\max} - \bar{Y}_{\min}}{\sqrt{\frac{MSE}{n}}}$$

- Tukey's Simultaneous Confidence Interval for $(\mu_i - \mu_j)$

$$(\bar{Y}_i - \bar{Y}_j) \pm q_\alpha(k, N-k) \cdot \sqrt{\frac{MSE}{2}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}, \forall i, j = 1, \ldots, k$$

- Reject H_0 if there is at least one simultaneous C.I. doesn't include 0.

Table C.8 Upper Percentage Points of the Studentized Range Distribution: Values of $q(0.05; k, v)$

Degrees of Freedom ν	Number of Treatments k								
	2	3	4	5	6	7	8	9	10
1	18.0	27.0	32.8	37.2	40.5	43.1	45.4	47.3	49.1
2	6.09	8.33	9.80	10.89	11.73	12.43	13.03	13.54	13.99
3	4.50	5.91	6.83	7.51	8.04	8.47	8.85	9.18	9.46
4	3.93	5.04	5.76	6.29	6.71	7.06	7.35	7.60	7.83
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99
6	3.46	4.34	4.90	5.31	5.63	5.89	6.12	6.32	6.49
7	3.34	4.16	4.68	5.06	5.35	5.59	5.80	5.99	6.15
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92
9	3.20	3.95	4.42	4.76	5.02	5.24	5.43	5.60	5.74
10	3.15	3.88	4.33	4.66	4.91	5.12	5.30	5.46	5.60
11	3.11	3.82	4.26	4.58	4.82	5.03	5.20	5.35	5.49
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.40
13	3.06	3.73	4.15	4.46	4.69	4.88	5.05	5.19	5.32
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25
15	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20
16	3.00	3.65	4.05	4.34	4.56	4.74	4.90	5.03	5.15
17	2.98	3.62	4.02	4.31	4.52	4.70	4.86	4.99	5.11
18	2.97	3.61	4.00	4.28	4.49	4.67	4.83	4.96	5.07
19	2.96	3.59	3.98	4.26	4.47	4.64	4.79	4.92	5.04
20	2.95	3.58	3.96	4.24	4.45	4.62	4.77	4.90	5.01
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92
30	2.89	3.48	3.84	4.11	4.30	4.46	4.60	4.72	4.83
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.74
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65
120	2.80	3.36	3.69	3.92	4.10	4.24	4.36	4.47	4.56
∞	2.77	3.32	3.63	3.86	4.03	4.17	4.29	4.39	4.47

Source: Reproduced with permission from Table 29 of E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, Vol. 1 (Cambridge: Cambridge University Press, 1954).

Example (Deflection of Beams)

Type	Observations	Average
A	82,86,79,83,85,84,86,87	84
B	74,82,78,75,76,77	77
C	79,79,77,78,82,79	79

- k, N, n_1 , n_2 , n_3 , MSE
- 95% C.I. for individual treatment μ_i
- Pairwise C.I. for $\mu_i - \mu_i$
- Bonferroni C.I. for all ($\mu_i - \mu_i$)
- Tukey's Simultaneous C.I. for all ($\mu_i - \mu_i$)

Model Checking

- Normality Check -- histogram and QQ plot

  SAS code (proc univariate normal) R code (qqnorm)

  Note: Use nonparametric alternative if normality is violated, e.g. a rank test like the Kruskal-Wallis Test (follows a Chi-square distribution under condition that there is no treatment difference).

- Homogeneity of Variance – Levene's Test

$$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 \text{ vs. } H_1 : \text{not all } \sigma_i^2 \text{ are the same.}$$

  Levene's Test is not sensitive to the normality.

- Other comparison – Contrast

$$H_0 : \sum_{i=1}^{k} c_i \tau_i = 0 \text{ vs. } H_1 : \sum_{i=1}^{k} c_i \tau_i \neq 0 \text{ where } \sum_{i=1}^{k} c_i = 0.$$

  For example, $\tau_1 + \tau_2 - (\tau_3 + \tau_4) = 0.$

Example:

Suppose you are comparing the time to relief of three headache medicines -- brands 1, 2, and 3. The time to relief data is reported in minutes. For this experiment, 15 subjects were randomly placed on one of the three medicines. Which medicine (if any) is the most effective? (SAS output)

```
DATA ACHE;
INPUT BRAND RELIEF;
CARDS;
1 24.5
1 23.5
1 26.4
1 27.1
1 29.9
2 28.4
2 34.2
2 29.5
2 32.2
2 30.1
3 26.1
3 28.3
3 24.3
3 26.2
3 27.8
;
```

```
PROC ANOVA DATA=ACHE;
CLASS BRAND;
MODEL RELIEF=BRAND;
MEANS BRAND / BON TUKEY LSD CLDIFF HOVTEST=LEVENE;
TITLE 'COMPARE RELIEF ACROSS MEDICINES ';
RUN;


PROC BOXPLOT DATA=ACHE;
PLOT RELIEF*BRAND;
TITLE 'ANOVA RESULTS';
RUN;
```