MA500HW8

Chang Lu

April 06 2025

1 Question 2

2. Consider room temperature as an additional blocking factor for the cut-flower experiment described in Section 4.2.
   (a) What design could you use to incorporate the two blocking factors of flower type and room temperature?
   (b) Set up a randomization plan to conduct this experiment.
   (c) What is the model you would use to analyze the data resulting from this experiment?

Figure 1: This is the question 2

1.1 (a) What design could you use to incorporate the two blocking factors of the type of flower and room temperature?

To incorporate both flower type and room temperature as blocking factors, we can use a two-way randomized complete block design. In this setup:

- One blocking factor is the flower type (e.g., Rose, Carnation, Daisy, Tulip).
- The second blocking factor is room temperature (e.g. cold, medium, warm).

Each treatment will be applied once in each combination of flower type and room temperature. This ensures that the treatment effects are estimated after accounting for variability due to both blocking factors.

1.2 (b) Set up a randomization plan to carry out this experiment.

We begin by identifying all combinations of blocking factors. Suppose that we have:

- 4 flower types: Rose, Carnation, Daisy, Tulip

- 3 temperature levels: Cold, Medium, Warm

This results in $4 \times 3 = 12$ blocks. Within each block, we randomly assign the 4 treatments:

1. Tap water
2. Tap water with sugar
3. Tap water with carbonated water
4. Tap water with 7-Up

The R code that outlines the randomization process is shown in Rmd. file.

1.3 (c) What is the model you would use to analyze the data resulting from this experiment?

We use a linear model that includes treatment and the two blocking factors:

$$Y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + \epsilon_{ijk}$$

Where:

- Y_{ijk}: response variable (e.g., days until flower wilts)
- μ: overall mean
- τ_i: fixed effect of the i-th treatment
- β_j: effect of the j-th flower type (block)
- γ_k: effect of the k-th temperature level (block)
- ϵ_{ijk}: random error term

2 Question 4

4. Horiuchi *et al.* (2005) conducted an experiment during Japan's 2004 Upper House election. The purpose of the experiment was to test the hypothesis that voter turnout is influenced by the amount of information available

Table 4.6 *Total Yields of Barley Varieties in 12 Independent Trials*

		Variety				
Place	Year	Manchuria	Svansota	Velvet	Trebi	Peatland
1	1931	81.0	105.4	119.7	109.7	98.3
1	1932	80.7	82.3	80.4	87.2	84.2
2	1931	146.6	142.0	150.7	191.5	145.7
2	1932	100.4	115.5	112.2	147.7	108.1
3	1931	82.3	77.3	78.4	131.3	89.6
3	1932	103.1	105.1	116.5	139.9	129.6
4	1931	119.8	121.4	124.0	140.8	124.8
4	1932	98.9	61.9	96.2	125.5	75.7
5	1931	98.9	89.0	69.1	89.3	104.1
5	1932	66.4	49.9	96.7	61.9	80.3
6	1931	86.9	77.1	78.9	101.8	96.0
6	1932	67.7	66.7	67.4	91.8	94.1

to voters. Potential voters were contacted by e-mail and asked to fill out a screening survey and participate in the study. In the screening survey potential voters were asked what their gender was and whether they were planning to vote in the upcoming election (yes, no, or undecided). This information is known to be correlated with voter participation. Next, the respondents to the screening survey were randomly assigned to one of four groups. The first group was asked via e-mail to view the official Web site of the LDP party that shows their policy proposal. The second group was asked to view the Web site of the DPJ party that shows their policy proposal. A third group was asked to view both Web sites, and a fourth group was not asked to visit any Web site prior to the election. Each participant was again contacted after the election to find out whether he or she had voted in the election or not.

Figure 2: This is the question 4

(a) What was the treatment factor in this study?
(b) What was the experimental unit?
(c) What is the response?
(d) If potential voters were classified into the following six groups based on the screening survey

- Male—plans to vote
- Female—plans to vote
- Male—does not plan to vote
- Female—does not plan to vote
- Male—undecided
- Female—undecided

should this information be used in assigning respondents to the four groups? Why or why not?

EXERCISES 137

(e) Explain how you would analyze the resulting data. What model would you use, what is the distribution of the dependent variable, etc.?

Figure 3: This is the question 4

2.1 (a) What was the treatment factor in this study?

The treatment factor was the amount and type of political information provided to potential voters. There were four levels:

1. View the LDP party website
2. View the DPJ party website
3. View both websites
4. Did not view any website (control group)

2.2 (b) What was the experimental unit?

The experimental unit was each voter who participated in the study.

2.3 (c) What is the response?

The response is a binary outcome: whether the individual voted or not in the 2004 Upper House election.

2.4 (d) Should classification based on gender and voting intention be used in assigning respondents to the four groups? Why or why not?

Yes, this information should be used in the assignment process to ensure balanced randomization across strata. Because gender and voting intention are known to correlate with voter turnout, using a stratified randomization approach would help reduce variability and potential confounding. This means randomly assigning individuals within each of the six strata:

- Male — plans to vote
- Female — plans to vote
- Male — does not plan to vote
- Female — does not plan to vote
- Male — undecided
- Female — undecided

This ensures each treatment group is balanced with respect to important baseline characteristics.

2.5 (e) Explain how you would analyze the resulting data. What model would you use, what is the distribution of the dependent variable, etc.?

Since the response is binary (voted or not), a logistic regression model is appropriate. The model would be:

$$\text{logit}(\pi_i) = \beta_0 + \beta_1 T_1 + \beta_2 T_2 + \beta_3 T_3 + \gamma_1 \text{Male}_i + \gamma_2 \text{Intent}_i + \cdots$$

Where:

- π_i is the probability that individual i voted.
- T_1, T_2, T_3 are dummy variables for the treatment groups (with the control group as reference).
- Additional covariates such as gender and voting intention can be included to adjust for baseline differences.

The dependent variable follows a Bernoulli distribution, and the likelihood is binomial. Hypothesis tests on the treatment coefficients can be used to assess whether information exposure significantly influenced voter turnout.

3 Question 6

6. le Riche and Csima (1964) evaluated four hypnotic drugs and a placebo to determine their effect on the quality of sleep in elderly patients. The treatment levels were labeled (A=Placebo, B=Ethchlorvynol, C=Glutethimide, D=Chloral hydrate, and E=Secobarbitol sodium). Elderly patients were given one of the capsules for five nights in succession and their quality of sleep was rated by a trained nurse on a 4-point scale (0=poor to 3=excellent) each night. An average score was calculated for each patient over the five nights in a week. Each patient received all five treatments in successive weeks. A Latin-square design was used to account for patient-to-patient differences and week-to-week effects. The design and the response (mean quality of sleep rating) are shown in the table on the next page.

   (a) What is the appropriate model for this data?

   (b) Complete the ANOVA and determine if there are any significant differences among the treatments.

   (c) Use an appropriate method to determine if there is a significant difference between the placebo and the average of the other drugs, and if there are significant differences among the four drugs.

138 RANDOMIZED BLOCK DESIGNS

	Week				
Patient	1	2	3	4	5
1	B (2.92)	E (2.43)	A (2.19)	C (2.71)	D (2.71)
2	D (2.86)	A (1.64)	E (3.02)	B (3.03)	C (3.03)
3	E (1.97)	B (2.50)	C (2.47)	D (2.65)	A (1.89)
4	A (1.99)	C (2.39)	D (2.37)	E (2.33)	B (2.71)
5	C (2.64)	D (2.31)	B (2.44)	A (1.89)	E (2.78)

   (d) Use residual plots to check the assumptions for the model you fit.

Figure 4: This is the question 6

3.1 (a) What is the appropriate model for this data?

Since a Latin square design is used, the appropriate model includes effects for:

- Treatment (Drug A–E)
- Row (Patient)
- Column (Week)

The model can be written as:

$$Y_{ijk} = \mu + \tau_i + \rho_j + \gamma_k + \epsilon_{ijk}$$

Where:

- Y_{ijk} is the sleep rating for treatment i, patient j, and week k
- μ is the overall mean
- τ_i is the fixed effect of treatment i
- ρ_j is the row (patient) effect
- γ_k is the column (week) effect
- ϵ_{ijk} is the random error

3.2 (b) Complete the ANOVA and determine if there are any significant differences among the treatments.

```
           Df Sum Sq Mean Sq F value  Pr(>F)
Treatment   4 2.0498  0.5125  10.725 0.00062 ***
Patient     4 0.6073  0.1518   3.178 0.05358 .
Week        4 0.3689  0.0922   1.930 0.17000
Residuals  12 0.5734  0.0478
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 5: Caption

The ANOVA table shows that the effect of treatment is statistically significant:

$$F(4, 12) = 10.725, \quad p < 0.001$$

This indicates that there are significant differences among the five treatments (four drugs and the placebo).

The patient effect is marginally significant with $p = 0.0536$, suggesting that individual variability may influence sleep quality. The week effect is not significant ($p = 0.17$), indicating no strong evidence of week-to-week changes.

3.3 (c) Use an appropriate method to determine if there is a significant difference between the placebo and the average of the other drugs, and if there are significant differences among the four drugs.

- To test the placebo (A) versus the average of drugs B–E, perform a contrast test:

$$H_0 : \mu_A = \frac{1}{4}(\mu_B + \mu_C + \mu_D + \mu_E)$$

- To test differences among drugs B–E, use Tukey's HSD.

3.3.1 Placebo vs. Drugs (Contrast Test):

```
         Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: User-defined Contrasts


Fit: aov(formula = Score ~ Treatment + Patient + Week, data = sleep_data)

Linear Hypotheses:
                      Estimate Std. Error t value Pr(>|t|)
Placebo vs Others == 0  -0.6935     0.1093  -6.345 3.69e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)
```

Figure 6: Caption

A contrast test comparing the placebo (A) with the average of the other four drugs (B, C, D, E) yields:

$$\text{Estimate} = -0.6935, \quad t = -6.345, \quad p = 3.69 \times 10^{-5}$$

This result is highly significant, indicating that the average sleep quality score under placebo is significantly lower than the average across the four drugs. Thus, the drugs are effective in improving sleep compared to placebo.

3.3.2 Tukey HSD Among Drugs (B–E):

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Score ~ Treatment + Patient + Week, data = drug_data)

$Treatment
      diff        lwr       upr     p adj
C-B -0.072 -0.4878128 0.3438128 0.9427121
D-B -0.140 -0.5558128 0.2758128 0.7114739
E-B -0.214 -0.6298128 0.2018128 0.4069492
D-C -0.068 -0.4838128 0.3478128 0.9509577
E-C -0.142 -0.5578128 0.2738128 0.7029836
E-D -0.074 -0.4898128 0.3418128 0.9383205
```

Figure 7: Caption

Pairwise comparisons among the four drugs show that all adjusted p-values are well above 0.05. For example:

- B vs. C: $p = 0.94$
- B vs. D: $p = 0.71$
- C vs. E: $p = 0.70$

This suggests that there are no significant differences among the four drugs. They all appear to perform similarly in improving sleep.

3.4 (d) Use residual plots to check the assumptions for the model you fit.

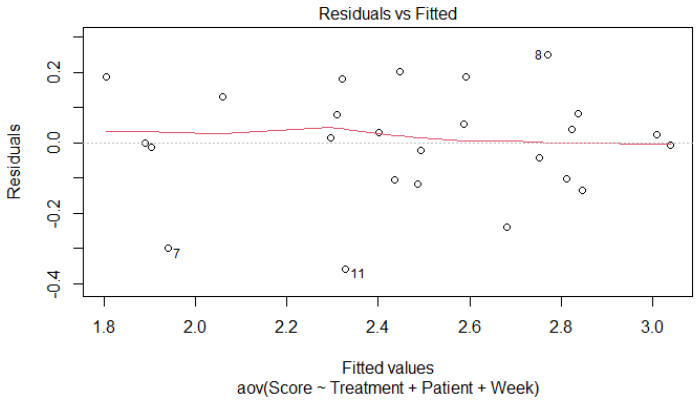

Figure 8: Caption

The Residuals vs Fitted plot is used to assess the assumptions of linearity and constant variance (homoscedasticity) in the ANOVA model.

The plot shows that the residuals are randomly scattered around the horizontal line at zero, with no clear pattern or trend. The red smooth line is relatively flat, indicating that the linearity assumption is met. Additionally, the spread of the residuals appears consistent across the range of fitted values, supporting the assumption of constant variance. The plot does not suggest any serious violations of the assumptions of linearity or homoscedasticity. The ANOVA model appears to be appropriate for the data based on this diagnostic.

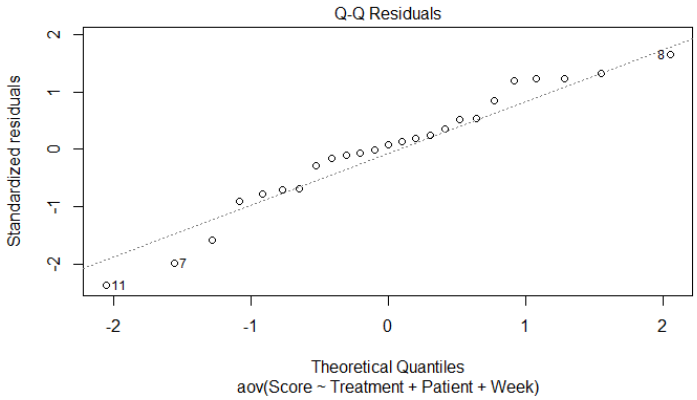

Figure 9: Caption

The Normal Q-Q Plot assesses whether the residuals are approximately normally distributed, which is an assumption of the ANOVA model.

In the plot, most residuals lie fairly close to the reference line, indicating that the normality assumption is largely satisfied. The residuals appear to be approximately normally distributed. The assumption of normality is reasonably met for the purpose of ANOVA.

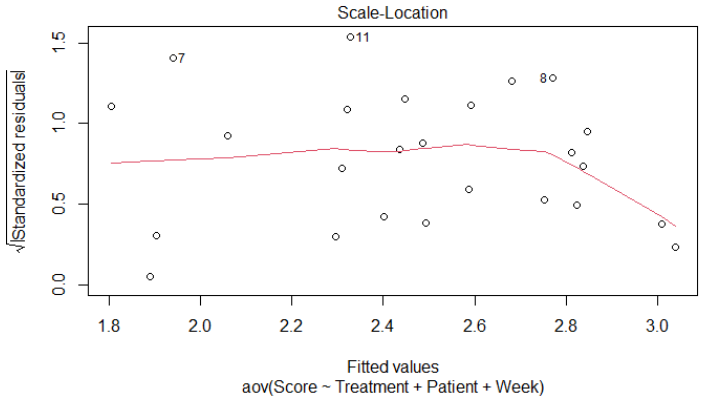

Figure 10: Caption

The Scale-Location plot checks for homoscedasticity, or constant variance of residuals, across levels of fitted values.

In the plot, the residuals appear to be fairly evenly spread across the range of fitted values, and the red trend line is relatively flat. Although there is a slight dip at higher fitted values, the variation remains reasonably constant throughout. The assumption of constant variance (homoscedasticity) appears to be met. There is no major indication of heteroscedasticity that would threaten the validity of the ANOVA model.

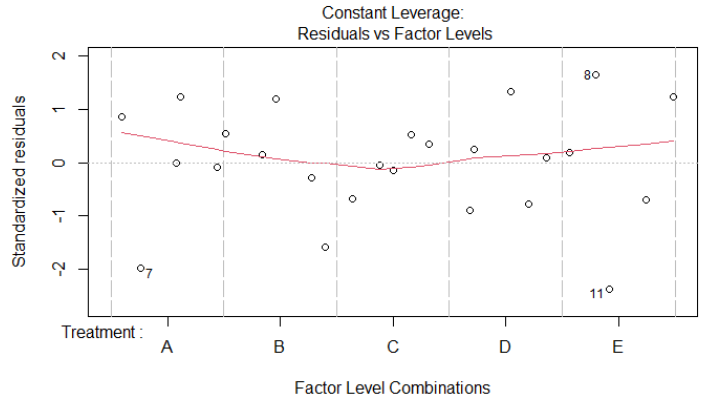

Figure 11: Caption

This plot displays the standardized residuals for each treatment level (A to E), allowing us to assess whether the variability or pattern of residuals differs

by treatment group.

The residuals appear to be fairly centered around zero across all treatment levels. While a few observations (e.g., 7 and 11) stand out as relatively extreme, there is no evidence of systematic bias or unequal spread of residuals across treatments. The red trend line remains relatively flat, suggesting no treatment-specific trend in the residuals.

The assumption of equal variance and independence of residuals across treatment groups is reasonably satisfied. No treatment group appears to introduce strong outliers or heteroscedasticity.

4 Question 8

8. Consider the situation described in Exercise 6 of Chapter 2 with $t = 4$ levels of the treatment factor, and $\Delta = 2.0\sigma$.

   (a) If by blocking the experimental units into blocks it was believed that the variance of the experimental error, σ^2, could be reduced by 50%, calculate the number of blocks that would be required to have a power of 0.90 for detecting a maximum difference in treatment means as large as Δ.

   (b) If by using a Latin-square design the variance of the experimental error, σ^2, could be reduced another 10%, determine the power for detecting treatment differences when using a 4×4 Latin-square design.

Figure 12: This is the question 2

Given:

- Number of treatment levels: $t = 4$
- Maximum difference in treatment means: $\Delta = 2\sigma$
- Desired power: 0.90

4.1 (a) Blocking reduces variance by 50%

When experimental units are blocked, the experimental error variance is reduced:

$$\sigma^2_{\text{new}} = 0.5 \cdot \sigma^2$$

```
      alpha nlev nreps Delta sigma     power
[1,]  0.05    4     2     2     1 0.1698028
[2,]  0.05    4     3     2     1 0.3390584
[3,]  0.05    4     4     2     1 0.5037050
[4,]  0.05    4     5     2     1 0.6442332
[5,]  0.05    4     6     2     1 0.7545861
[6,]  0.05    4     7     2     1 0.8361289
[7,]  0.05    4     8     2     1 0.8935978
[8,]  0.05    4     9     2     1 0.9325774
[9,]  0.05    4    10     2     1 0.9581855
```

Figure 13: Exercise 6

From Exercise 6, we found that with the original variance, a minimum of $r = 8$ replicates per treatment level is needed to achieve 90% power.

Using the relationship:

$$\lambda \propto \frac{r}{\sigma^2} \quad \Rightarrow \quad r_{\text{new}} = \frac{8}{2} = 4$$

Only 4 blocks are required to achieve 90% power when blocking reduces the variance by 50%.

4.2 (b) Latin Square Design with 60% Variance Reduction

A Latin square design provides additional control, reducing the error variance by another 10%, so:

$$\sigma^2_{\text{new}} = 0.4 \cdot \sigma^2 \quad \Rightarrow \quad \sigma_{\text{new}} = \sqrt{0.4} \approx 0.6325$$

Using the `Fpower1` function in R with:

- `nlev = 4`
- `nreps = 4`
- `Delta = 2`
- `sigma = 0.6325`

```
      alpha nlev nreps Delta     sigma     power
[1,]  0.05    4     4     2 0.6324555 0.9027768
```

Figure 14: Caption

The resulting power is:

$$\text{Power} = 0.9028$$

A 4×4 Latin square design achieves approximately 90.3% power for detecting treatment differences when the variance is reduced by 60%.