Ch11. The Analysis of Variance

The quantity of information contained in a sample is affected by various factors that the experimenter may or may not be able to control. This chapter introduces tree different experimental designs, two of which are direct extensions of the unpaired and paired designs of Chapter 10. A new technique called the analysis of variance is used to determine how the different experimental factors affect the average response.

윤 연 옥

Experimental Design

- The **sampling plan** or **experimental design** determines the way that a sample is selected.
- In an **observational study**, the experimenter observes data that already exist. The **sampling plan** is a plan for collecting this data.
- In a **designed experiment**, the experimenter imposes one or more experimental conditions on the experimental units and records the response.

Definitions

- An **experimental unit** is the object on which a measurement or measurements is taken.
- A **factor** is an independent variable whose values are controlled and varied by the experimenter.
- A **level** is the intensity setting of a factor.
- A **treatment** is a specific combination of factor levels.
- The **response** is the variable being measured by the experimenter.

Example

- A group of people is randomly divided into an experimental and a control group. The control group is given an aptitude test after having eaten a full breakfast. The experimental group is given the same test without having eaten any breakfast. What are the factors, levels, and treatments in this experiment?

Experimental unit : person

Factor : meal

Levels(two) : “breakfast” and “no breakfast”

Treatments: “breakfast” and “no breakfast”

response : test score

Example

- Suppose that the experimenter from previous example began by randomly selecting 20 men and 20 women for the experiment. These two groups were then randomly divided into 10 each for the experimental and control groups. What are the factors, levels, and treatments in this experiment?

Experimental unit : person

Factor : meal, gender

Levels

- Meal : "breakfast" and "no breakfast"

- Gender: men and women

Treatments: men with breakfast, men without breakfast
women with breakfast, women without breakfast

The Analysis of Variance(ANOVA)

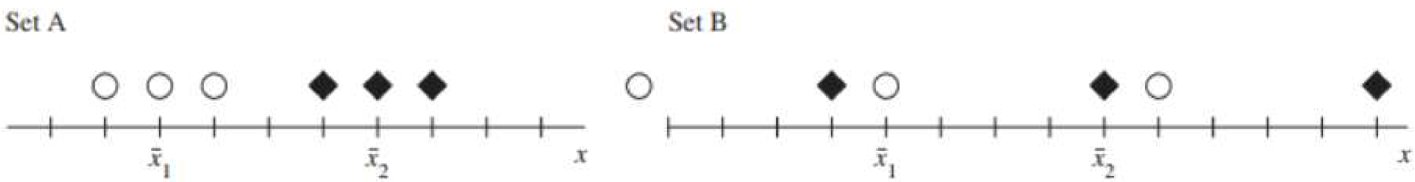
- All measurements exhibit **variability.**
- The total variation in the measurements is called the total sum of squares, given by

  Total SS = $\sum(x_i - \bar{x})^2$

- In an analysis of variance, the total variation in the response measurements is broken into portions that can be attributed to various **factors**, plus leftover variation that you cannot explain.

- If an experiment has been properly designed, we compare the variation due to any one factor to the typical random variation in the experiment.

- Example: Two sets of samples with the same means



✓ In set A, the variability of the measurements within the groups (◆s and ○s) is much smaller than the variability between the two groups.

✓ In set B, there is more variability within the groups (◆s and ○s), causing the two groups to "mix" together and making it more difficult to see the identical difference in the means.

Assumptions

- Similar to the assumptions required for the Student's t and F statistics in Chapter 10.

1. The observations within each treatment group are normally distributed with a common variance σ^2.
2. Assumptions regarding the sampling procedures are specified for each design.

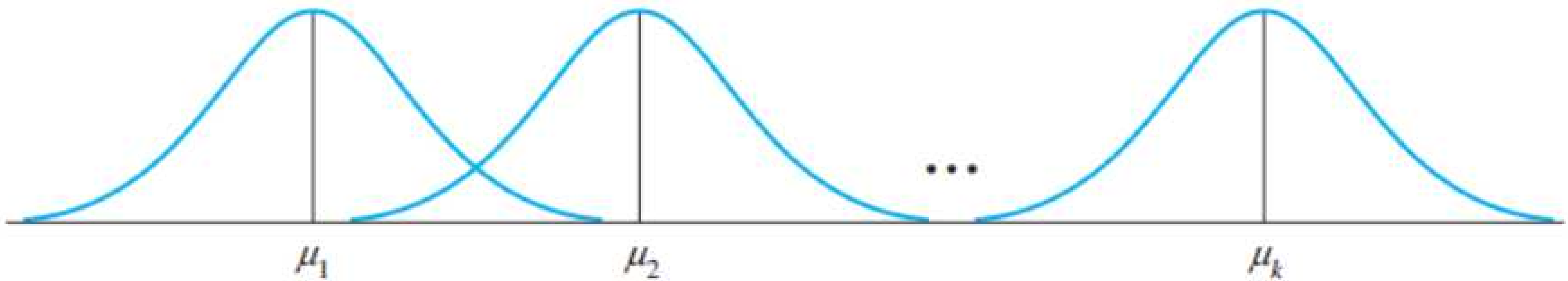
- Analysis of variance procedures are fairly robust when sample sizes are equal and when the data are fairly mound-shaped.

Three Designs

- **Completely randomized design:** an extension of the two independent sample *t*-test.
- **Randomized block design:** an extension of the paired difference test.
- $\boldsymbol{a \times b}$ **Factorial experiment:** we study two experimental factors and their effect on the response.

The Completely Randomized Design

- A **one-way classification** in which one factor is set at k different levels.
- The k levels correspond to k different normal populations, which are the **treatments.**
- Are the k population means the same, or is at least one mean different from the others?



Example

In an experiment to determine the effect of nutrition on the attention spans of elementary school students, a group of 15 students were randomly assigned to each of three meal plans: no breakfast, light breakfast, and full breakfast. Their attention spans (in minutes) were recorded during a morning reading period

Attention Spans of Students After Three Meal Plans

No Breakfast	Light Breakfast	Full Breakfast
8	14	10
7	16	12
9	12	16
13	17	15
10	11	12
$T_1 = 47$	$T_2 = 70$	$T_3 = 65$

✓ k = 3 treatments.

✓ Are the average attention spans different?

The Completely Randomized Design

- Random samples of size n_1, n_2, …,n_k are drawn from k populations with means μ_1, μ_2,…, μ_k and with common variance σ^2.
- Let x_{ij} be the j-th measurement in the i-th sample.
- The total variation in the experiment is measured by the **total sum of squares**:

$$\textbf{Total SS} = \sum (x_{ij} - \bar{x})^2$$

The Analysis of Variance

The **Total SS** is divided into two parts:

✓ **SST** (sum of squares for treatments): measures the variation among the k sample means.

✓ **SSE** (sum of squares for error): measures the variation within the k samples.

Total SS = SST + SSE

- Total SS $= \sum(x_{ij} - \bar{x})^2 = \sum x_{ij}^2 - \text{CM}$

with

$$\text{CM} = \frac{\text{G}^2}{n} \quad \text{where } \text{G} = \sum x_{ij} \quad \text{total of all n}$$

$$\text{SST} = \sum \frac{{T_i}^2}{n_i} - \text{CM} \quad , \quad \text{MST} = \frac{\text{SST}}{k-1}$$

where $T_i =$ total for treatment i

n_i = number of observations in sample i

$n = n_1 + n_2 + \ldots + n_k$

$$\text{SSE} = \text{Total SS} - \text{SST} \text{ , MSE} = \frac{\text{SSE}}{n-k}$$

Example(~ continued)

No Breakfast	Light Breakfast	Full Breakfast
8	14	10
7	16	12
9	12	16
13	17	15
10	11	12
$T_1 = 47$	$T_2 = 70$	$T_3 = 65$

$$\text{CM} = \frac{182^2}{15} = 2208.2667$$

$$Total\ SS = (8^2+7^2+\ldots+12^2) - CM \ = 2338 - 2208.2667 = 129.7333$$

$$SST = \frac{47^2}{5} + \frac{70^2}{5} + \frac{65^2}{5} - CM = 2266.8 - 2208.2667 = 58.5333$$

$$SSE = Total\ SS - SST = 129.7333 - 58.5333 = 71.2$$

Degrees of Freedom and Mean Squares

- These **sums of squares** behave like the numerator of a sample variance. When divided by the appropriate **degrees of freedom**, each provides a **mean square**, an estimate of variation in the experiment.
- **Degrees of freedom** are additive, just like the sums of squares.

$$\text{Total } df = \text{Trt } df + \text{Error } df$$

The ANOVA Table

Total *df* = $n_1 + n_1 + \cdots + n_k \ - 1 = n - 1$

Treatment *df* = $k - 1$

Error *df* = $n - 1 - (k - 1) = n - k$

Source	df	SS	MS	F
Treatments	k -1	SST	SST/(k-1)	MST/MSE
Error	n - k	SSE	SSE/(n-k)	
Total	n -1	Total SS		

Example(~ continued)

$$\mathrm{CM} = \frac{182^2}{15} = 2208.2667$$

$$Total\ SS = (8^2+7^2+\ldots+12^2) - CM\ = 2338 - 2208.2667 = 129.7333$$

$$SST = \frac{47^2}{5} + \frac{70^2}{5} + \frac{65^2}{5} - CM = 2266.8 - 2208.2667 = 58.5333$$

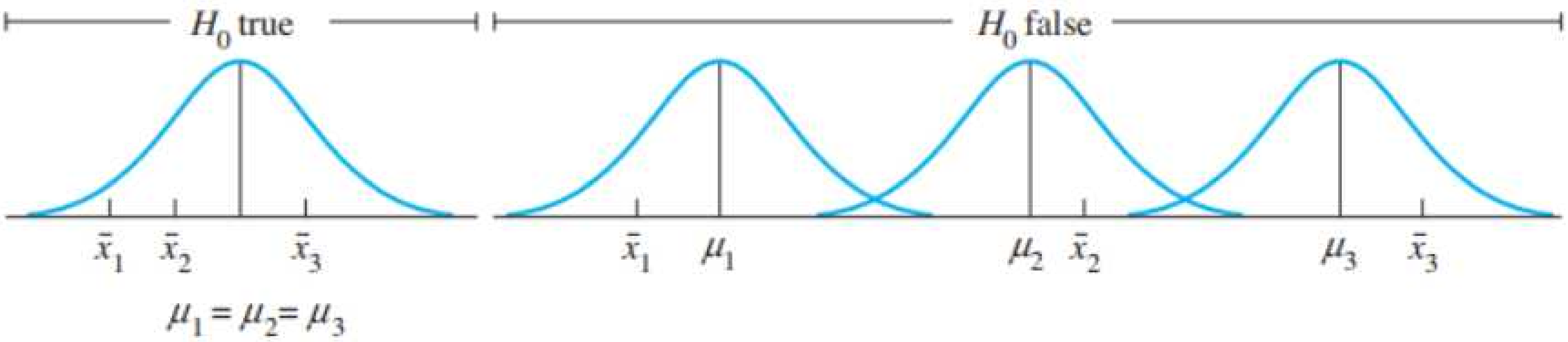
$$SSE = Total\ SS - SST = 129.7333 - 58.5333 = 71.2$$

Source	df	SS	MS	F
Treatments	*3-1=2*	58.533	29.267	4.933
Error	*15-3=12*	71.2	5.933	
Total	*15-1=14*	129.733		

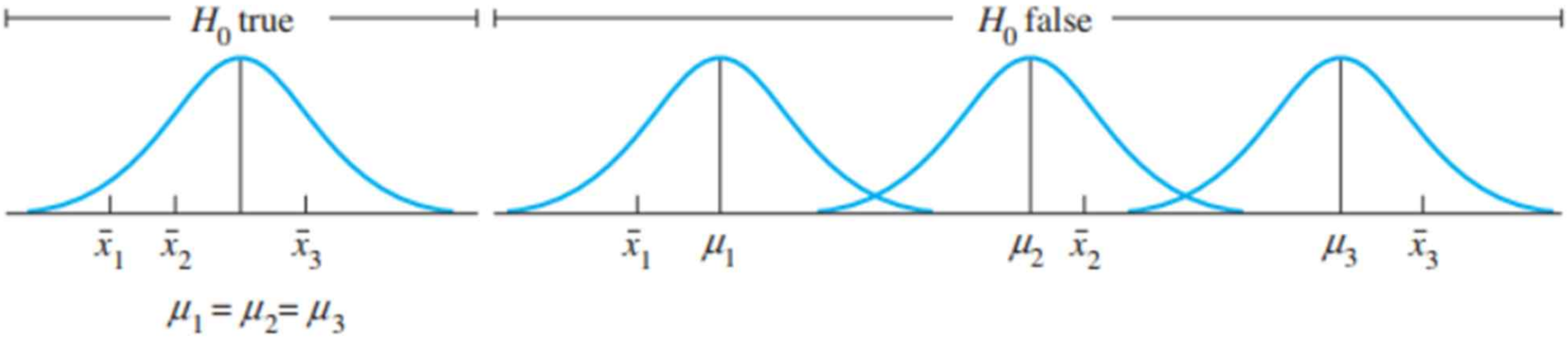
Testing the Treatment Means

$\mathbf{H_0} : \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$ *versus*

$\mathbf{H_a}$: at least one mean is different



Remember that σ^2 is the common variance for all k populations. The quantity MSE = SSE/(n - k) is a pooled estimate of σ^2, a weighted average of all k sample variances, whether or not H_0 is true.



- If H_0 is true, then the variation in the sample means, measured by MST = [SST/ $(k-1)$], also provides an unbiased estimate of σ^2.
- However, if H_0 is false and the population means are different, then MST— which measures the variance in the sample means — is unusually **large.**

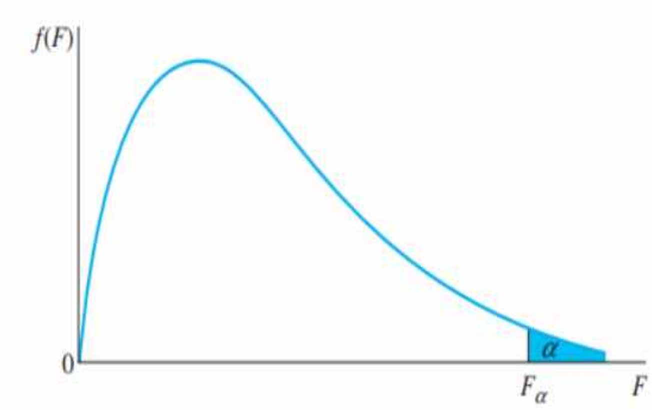
The test statistic **F = MST/ MSE** tends to be larger that usual.

The F test

- Reject H_0 for large values of F, using a **right-tailed** statistical test.
- When H_0 is true, this test statistic has an F distribution with $df_1 = (k-1)$ and $df_2 = (n-k)$ degrees of freedom and **right-tailed** critical values of the F distribution can be used.

- To test $H_0 : \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$
- Test Statistic: $F = \dfrac{MST}{MSE}$
- Reject H_0 if $F > F_\alpha$

  with $(k-1)$ and $(n-k)$ df



Example(~ continued)

Source	df	SS	MS	F
Treatments	2	58.533	29.267	4.933
Error	12	71.2	5.933	
Total	14	129.733		

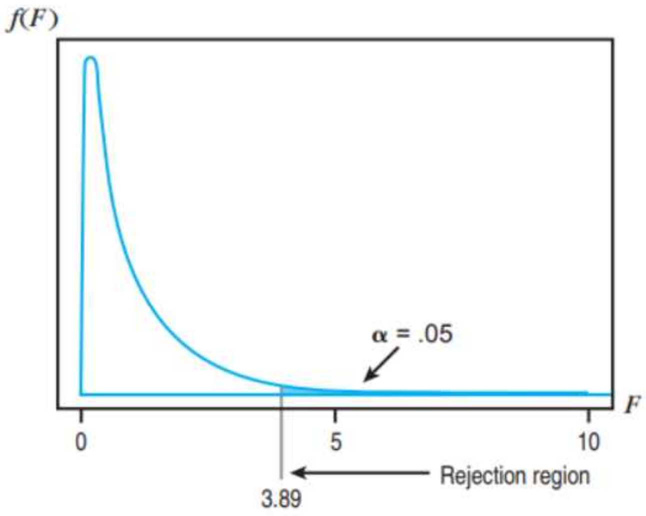
(1) $Hypo$ $H_0 : \mu_1 = \mu_2 = \mu_3$ versus H_a :at least one mean is different

(2) Test statistic $F = \frac{MST}{MSE} = \frac{29.267}{5.933} = 4.933$

(3) Rejection region: $F > F_{.05}(2, 12) = 3.89$ from table6 in Appendix I

(4) Since F=4.93 > 3.89, reject H_0.

There is sufficient evidence to indicate that at least one of the three average attention spans is different from at least one of the others.



Estimating Differences in the Treatment Means

- If a difference exists between the treatment means, we can explore it with confidence intervals.

- A single mean, μ_i : $\bar{x}_i \pm t_{\alpha/2} \frac{s}{\sqrt{n_i}}$

- Difference $\mu_i - \mu_j$: $(\bar{x}_i - \bar{x}_j) \pm t_{\alpha/2}\sqrt{s^2\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$

where $s = \sqrt{MSE} = \sqrt{\frac{SSE}{n-k}}$

and $t_{\alpha/2}$ is based on $(n-k)$ df.

Example(~ continued)

Find a 95% confidence interval for the average attention span for students who eat no breakfast, as well as a 95% confidence interval for the difference in the average attention spans for light versus full breakfast eaters

Sol) $s^2 = MSE = 5.9333$

$\Rightarrow s = \sqrt{5.9333} = 2.436 \;\; \text{with} \;\; df = (n-k) = 12$

- For no breakfast : $\bar{x}_1 \pm t_{\alpha/2}\frac{s}{\sqrt{n_1}} \quad \Rightarrow 9.4 \pm 2.179\,(\frac{2.436}{\sqrt{5}})$

  $\Rightarrow \;\; 9.4 \pm 2.37$

  between 7.03 and 11.77 minutes.

- For light versus full breakfast: $(\bar{x}_i - \bar{x}_j) \pm \; t_{\alpha/2}\sqrt{s^2\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$

  $\Rightarrow (14 - 13) \pm 2.179\sqrt{5.9333\left(\frac{1}{5}+\frac{1}{5}\right)} \quad \Rightarrow \quad 1 \pm 3.36$

  a difference of between -2.36 and 4.36 minutes. *(include 0)*

✓ *It does not indicate a difference in the average attention spans for students who ate light versus full breakfasts.*

Tukey's Method for Paired Comparisons

- Designed to test all pairs of population means simultaneously, with an **overall error rate of α.**
- Based on the **studentized range**, the difference between the largest and smallest of the k sample means.
- Assume that the **sample sizes are equal** and calculate a "ruler" that measures the distance required between any pair of means to declare a significant difference.

Tukey's Method

- Calculate: $\omega = q_{\alpha}(k, df)\dfrac{s}{\sqrt{n_t}}$

  where

  k = number of treatment means

  $s = \sqrt{MSE}$ df = error df

  n_t = common sample size

  (the number of observations in each of the k treatment means)

  $q_{\alpha}(k, df)$: Table 11(a) and 11(b) in Appendix I,

  for $\alpha = .05$ and $.01$, respectively , and for various combinations of k and df

Rule : Two population means are judged to differ if the corresponding sample means differ by ω or more.

Example(~ continued)

- Use Tukey's method to determine which of the three population means differ from the others.

	No Breakfast	**Light Breakfast**	**Full Breakfast**
	$T_1 = 47$	$T_2 = 70$	$T_3 = 65$
Means	47/5 = 9.4	70/5 = 14.0	65/5 = 13.0

$$ss = \sqrt{MSE} = 2.436,\ n_t = 5,\ df = (n-k) = 15-3 = 12$$

$$\omega = q_{.05}(3{,}12)\frac{s}{\sqrt{n_t}} = 3.77\frac{2.436}{\sqrt{5}} = 4.11$$

- List the sample means from smallest to largest

$$\begin{array}{ccc} \bar{x}_1 & \bar{x}_3 & \bar{x}_2 \\ \underline{9.4 \quad 13.0} & & \\ & \underline{13.0 \quad 14.0} & \end{array}$$

- check the difference between every pair of means.

- The only difference that exceeds $\omega = 4.11$ is the difference between no breakfast(9.4) and a light breakfast(14.0).

TABLE 11(a) Percentage Points of the Studentized Range, $q_{.05}(k, df)$; Upper 5% Points

df	k = 2	3	4	5	6	7	8	9	10	11
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51
13	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36

The Randomized Block Design

- If the design involves ***k* treatments** within each of ***b* blocks**, then the total number of observations is $n = bk$.
- The purpose of blocking is to remove or isolate the block-to-block variability that might hide the effect of the treatments.
- There are two factors—**treatments** and **blocks**, only one of which is of interest to the experimenter.

Example

- We want to investigate the affect of 3 methods of soil preparation on the growth of seedlings.
- Each method is applied to seedlings growing at each of 4 locations and the average first year growth is recorded.

	Location			
Soil Prep	**1**	**2**	**3**	**4**
A	11	13	16	10
B	15	17	20	12
C	10	15	13	10

✓ Treatment = soil preparation ($k = 3$)

✓ Block = location ($b = 4$)

✓ Is the average growth different for the 3 soil preps?

The Randomized Block Design

- Let x_{ij} be the response for the i-th treatment applied to the j-th block. $i = 1, 2, \ldots k$ $j = 1, 2, \ldots, b$
- The total variation in the experiment is measured by the **total sum of squares**: Total SS $= \sum(x_{ij} - \bar{x})^2$
- The **Total SS** is divided into 3 parts:
  - ✓ **SST** (sum of squares for treatments): measures the variation among the k treatment means
  - ✓ **SSB** (sum of squares for blocks): measures the variation among the b block means
  - ✓ **SSE** (sum of squares for error): measures the random variation or experimental error in such a way that:

$$\text{Total SS} = \text{SST} + \text{SSB} + \text{SSE}$$

Computing Formulas

- $\mathrm{CM} = \frac{\mathrm{G}^2}{n}$ where $\mathrm{G} = \sum x_{ij}$ total of all n
- Total SS $= \sum(x_{ij} - \bar{x})^2 = \sum x_{ij}^2 - \mathrm{CM}$
- $\mathrm{SST} = \frac{\sum {T_i}^2}{b} - \mathrm{CM}$ where $T_i =$ total for treatment i
- $\mathrm{SSB} = \frac{\sum {B_j}^2}{k} - \mathrm{CM}$ where $B_j =$ total for block j
- SSE $=$ Total SS−SST−SSB

The ANOVA table

Source	df	SS	MS	F
Treatments	k -1	SST	SST/(k-1)	MST/MSE
Blocks	b -1	SSB	SSB/(b-1)	MSB/MSE
Error	(b-1)(k-1)	SSE	SSE/(b-1)(k-1)	
Total	n -1	Total SS		

Example

The cell phone industry is involved in a fierce battle for customers, with each company devising its own complex pricing plan to lure customers. Since the cost of a cell phone minute varies drastically depending on the number of minutes per month used by the customer, a consumer watchdog group decided to compare the average costs for four cellular phone companies using three different usage levels as blocks. The monthly costs (in dollars) computed by the cell phone companies for peak-time callers at low (20 minutes per month), middle (150 minutes per month), and high (1000 minutes per month) usage levels are given in below Table. Construct the analysis of variance table for this experiment.

Monthly Phone Costs of Four Companies at Three Usage Levels

	Company				
Usage Level	A	B	C	D	Totals
Low	27	24	31	23	$B_1 = 105$
Middle	68	76	65	67	$B_2 = 276$
High	308	326	312	300	$B_3 = 1246$
Totals	$T_1 = 403$	$T_2 = 426$	$T_3 = 408$	$T_4 = 390$	$G = 1627$

Sol) ***randomized block design*** with ***b*= 3** usage levels (blocks) and ***k*= 4** companies (treatments), so there are $n= bk = 3*4 = 12$

$$G = \sum x_{ij} = 1627$$

$$CM = \frac{G^2}{n} = \frac{1627^2}{12} = 220{,}594.0833$$

$$\text{Total SS} = (27^2 + 24^2 + \cdots + 300^2) - CM = 189{,}798.9167$$

$$SST = \frac{(403^2 + .. + 390^2)}{3} - CM = 222.25$$

$$SSB = \frac{(105^2 + 276^2 + 1246^2)}{4} - CM = 189{,}335.1667$$

SSE = Total SS – SST - SSB = 241.5

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*
Usage	189335.167	2	94667.583	2351.990
Company	222.25	3	74.083	1.841
Error	241.5	6	40.25	
Total	189798.917	11		

Testing the Treatment and Block Means

- For either treatment or block means, we can test:
- $H_0: \mu_1 = \mu_2 = \mu_3 = \ ...$ Versus

H_0 : No difference among the *k* treatment(or b block) means

- H_a: at least one mean is different

✓ Remember that σ^2 is the common variance for all *bk* treatment-block combinations. MSE is an unbiased the best estimate of σ^2, whether or not H_0 is true.

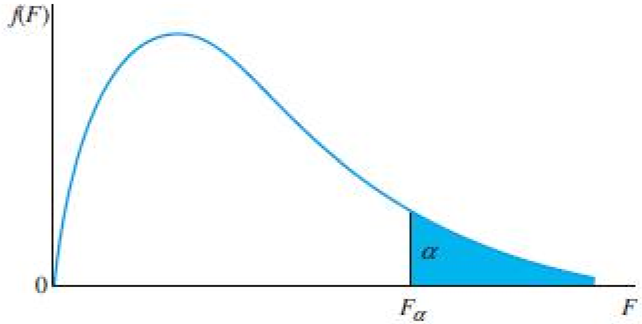
- If H_0 is false and the population means are different, then MST or MSB— whichever you are testing— will unusually **large.** The test statistic **F = MST/ MSE** (or **F = MSB/ MSE)** tends to be larger that usual.
- We use a right-tailed F test with the appropriate degrees of freedom.

- **To test H_0: treatment (or block) means are equal**

- Test Statistic: $F = \dfrac{MST}{MSE}$ (or $F = \dfrac{MSB}{MSE}$)

- Reject H_0

  if $F > F_\alpha$ with $k-1$ (or $b-1$) and $(b-1)(k-1)$ df



Example (Cellular phone)

H_0: No difference in the average cost among companies versus
H_a: the average cost is different for at least one of the four companies

Sol) Use the analysis of variance *F* statistic

$$F = \frac{\text{MST}}{\text{MSE}} = \frac{74.1}{40.3} = 1.84$$

Rejection region: F >$F_{.05}$(3, 6)=4.76

Since 1.84 < 4.76, do not reject H_0
There is insufficient evidence to indicate a difference in the average monthly costs for the four companies.

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*
Usage	189335.167	2	94667.583	2351.990
Company	222.25	3	74.083	1.841
Error	241.5	6	40.25	
Total	189798.917	11		

df_2	a	1	2	3
1	.100	39.86	49.50	53.59
	.050	161.4	199.5	215.7
	.025	647.8	799.5	864.2
	.010	4052	4999.5	5403
	.005	16211	20000	21615
2	.100	8.53	9.00	9.16
	.050	18.51	19.00	19.16
	.025	38.51	39.00	39.17
	.010	98.50	99.00	99.17
	.005	198.5	199.0	199.2
3	.100	5.54	5.46	5.39
	.050	10.13	9.55	9.28
	.025	17.44	16.04	15.44
	.010	34.12	30.82	29.46
	.005	55.55	49.80	47.47
4	.100	4.54	4.32	4.19
	.050	7.71	6.94	6.59
	.025	12.22	10.65	9.98
	.010	21.20	18.00	16.69
	.005	31.33	26.28	24.26
5	.100	4.06	3.78	3.62
	.050	6.61	5.79	5.41
	.025	10.01	8.43	7.76
	.010	16.26	13.27	12.06
	.005	22.78	18.31	16.53
6	.100	3.78	3.46	3.29
	.050	5.99	5.14	4.76
	.025	8.81	7.26	6.60
	.010	13.75	10.92	9.78
	.005	18.63	14.54	12.92

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Usage	189335.167	2	94667.583	2351.990	0.000	5.143
Company	222.25	3	74.083	1.841	0.240	4.757
Error	241.5	6	40.25			
Total	189798.917	11				

Identifying Difference in the Treatment Means

- If the *F*-test indicate a significant difference in the means, use Tukey's method of paired comparisons to determine which pairs of treatment or block means are significantly different from one another

- If you have a special interest in a particular *pair* of treatment or block means, you can estimate the difference using a $(1-\alpha)100\%$ confidence interval

✓ Remember that MSE always provides an unbiased estimator of σ^2 and uses information from the entire set of measurements.

✓ Use $s^2 =$ MSE with $df = (b - 1)(k - 1)$ to estimate σ^2 in comparing the treatment and block means.

Tukey's Method

- **For comparing treatment means:** $\omega = q_\alpha(k, df)\frac{s}{\sqrt{b}}$
- **For comparing block means:** $\omega = q_\alpha(b, df)\frac{s}{\sqrt{k}}$

$s = \sqrt{\text{MSE}}$ $\quad\quad df = \text{error } df$

$q_\alpha(k, df) = $ value from Table 11.

If any pair of means differ by more than ω,
they are declared different.

Confidence Intervals

- **Difference in treatment means:** $\left(\bar{T}_i - \bar{T}_j\right) \pm t_{\alpha/2}\sqrt{s^2\left(\frac{2}{b}\right)}$

- **Difference in block means:** $(\bar{B}_i - \bar{B}_j) \pm t_{\alpha/2}\sqrt{s^2\left(\frac{2}{k}\right)}$

where $\bar{T}_i = T_i/b$ and $\bar{B}_i = B_i/k$ are the necessary treatment or block means.
$s = \sqrt{\text{MSE}}$ and t is based on error df .

Example (Cellular phone)

- Since the F-test did not show any significant differences in the average costs for the four companies, there is no reason to use Tukey's method of paired comparisons.
- Suppose, however, that you are an executive for company B and your major competitor is company C. Can you claim a significant difference in the two average costs? Using a 95% confidence interval.

$$(\bar{T}_2 - \bar{T}_3) \pm t_{.025}\sqrt{\mathrm{MSE}\left(\frac{2}{b}\right)}$$

$$\left(\frac{426}{3} - \frac{408}{3}\right) \pm 2.447\sqrt{40.3\left(\frac{2}{3}\right)}$$

$$6 \pm 12.68$$

- ✓ the difference between the two average costs is estimated as between - - \$6.68 and \$18.68.
- ✓ Since 0 is contained in the interval, you do not have evidence to indicate a significant difference in your average costs.

Cautions about Blocking

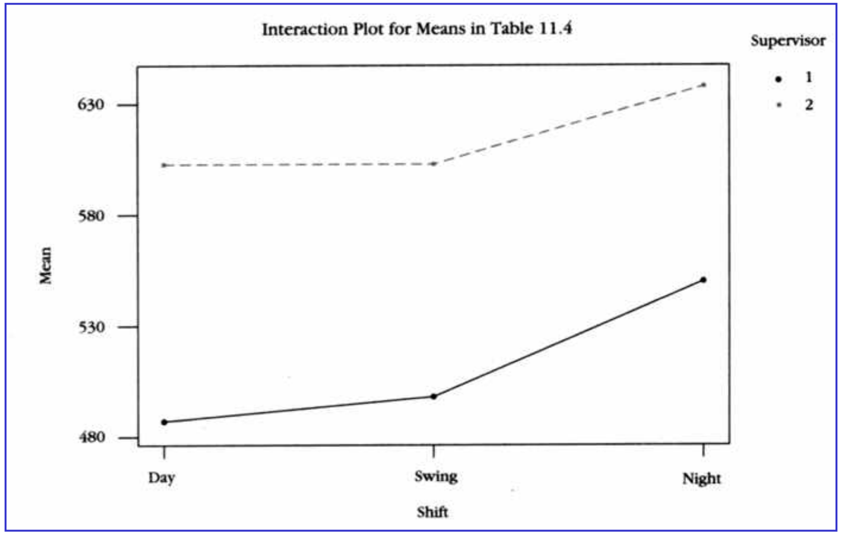
- ✓ A randomized block design should not be used when treatments and blocks **both** correspond to experimental factors of interest to the researcher
- ✓ Remember that **blocking may not always be beneficial**.
- ✓ Remember that you **cannot construct confidence intervals for individual treatment means** unless it is reasonable to assume that the b blocks have been randomly selected from a population of blocks.

An *a* x *b* Factorial Experiment

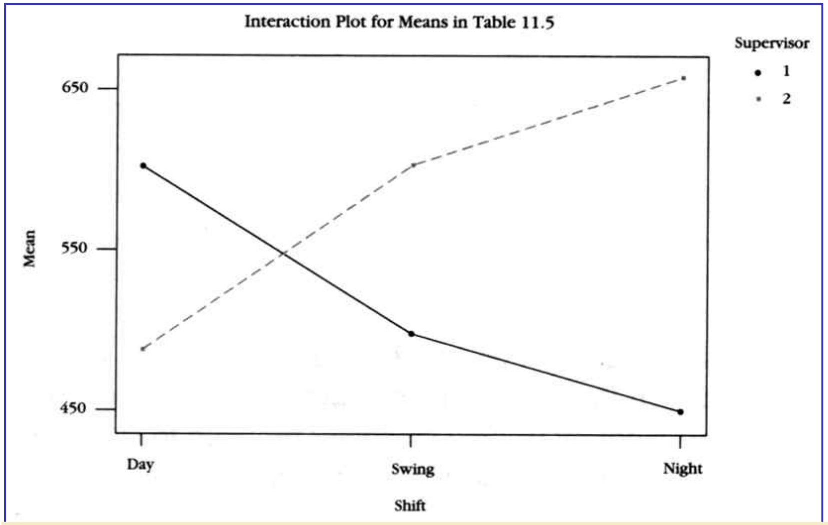
- A **two-way classification** in which involves two factors, both of which are of interest to the experimenter.
- There are ***a*** levels of factor A and ***b*** levels of factor B—the experiment is replicated ***r*** times at each factor-level combination.
- The replications allow the experimenter to investigate the **interaction** between factors A and B.
- The **interaction** between two factor A and B is the tendency for one factor to behave differently, depending on the particular level setting of the other variable.
- Interaction describes the effect of one factor on the behavior of the other. If there is **no interaction**, the two factors behave **independently**.

Example (drug manufacturer)

- A drug manufacturer has two supervisors who work at each of three different shift times. Do outputs of the supervisors behave differently, depending on the particular shift they are working?



Supervisor 2 always does better than 1, regardless of the shift.(No Interaction)



Supervisor 1 does better earlier in the day, while supervisor 2 does better at night.(Interaction)

The *a* x *b* Factorial Experiment

- Let x_{ijk} be the k-th replication
  at the i-th level of A and the j-th level of B.
  $i = 1, 2, \ldots, a \quad j = 1, 2, \ldots, b \quad k = 1, 2, \ldots, r$

- The total variation in the experiment is measured by the **total sum of squares**:

$$\text{Total SS} = \sum (x_{ijk} - \bar{x})^2$$

The Analysis of Variance

The **Total SS** is divided into 4 parts:

- ✓ **SSA** (sum of squares for factor A): measures the variation among the means for factor A
- ✓ **SSB** (sum of squares for factor B): measures the variation among the means for factor B
- ✓ **SS(AB)** (sum of squares for interaction): measures the variation among the *ab* combinations of factor levels
- ✓ **SSE** (sum of squares for error): measures experimental error in such a way that:

$$\text{Total SS} = \text{SSA} + \text{SSB} + \text{SS(AB)} + \text{SSE}$$

Computing Formulas

$$\mathrm{CM} = \frac{\mathrm{G}^2}{n} \quad \text{where } \mathrm{G} = \sum x_{ijk}$$

$$\text{Total SS} = \sum (x_{ijk} - \bar{x})^2 = \sum x_{ijk}^2 - \mathrm{CM}$$

$$\mathrm{SSA} = \frac{\sum {A_i}^2}{br} - \mathrm{CM} \quad \text{where } A_i = \text{ total for level } i \text{ of A}$$

$$\mathrm{SSB} = \frac{\sum {B_j}^2}{ar} - \mathrm{CM} \quad \text{where } B_j = \text{ total for level } j \text{ of B}$$

$$\mathrm{SS(AB)} = \frac{\sum {AB_{ij}}^2}{r} - \mathrm{CM} - \mathrm{SSA} - \mathrm{SSB}$$

where AB_{ij} = total for level i of A and level j of B

$$\mathrm{SSE} = \text{Total SS} - \mathrm{SSA} - \mathrm{SSB} - \mathrm{SS(AB)}$$

The ANOVA Table

Source	df	SS	MS	F
A	a -1	SST	SST/(a-1)	MST/MSE
B	b -1	SSB	SSB/(b-1)	MSB/MSE
Interaction	$(a$-1$)(b$-1$)$	SS(AB)	SS(AB)/$(a$-1$)(b$-1$)$	MS(AB)/MSE
Error	$ab(r$-$1)$	SSE	SSE/ab(r-1)	
Total	abr -1	Total SS		

Tests for a Factorial Experiment

- Test for the significance of both factors and the interaction using F-tests from the ANOVA table.

✓ *Remember that s^2 is the common variance for all ab factor-level combinations. MSE is the best estimate of s^2, whether or not H_0 is true.*

- The interaction is tested first using F = MS(AB)/MSE.
- If the interaction is not significant, the main effects A and B can be individually tested using F = MSA/MSE and F = MSB/MSE, respectively.
- If the interaction is significant, the main effects are NOT tested, and we focus on the differences in the *ab* factor-level means.

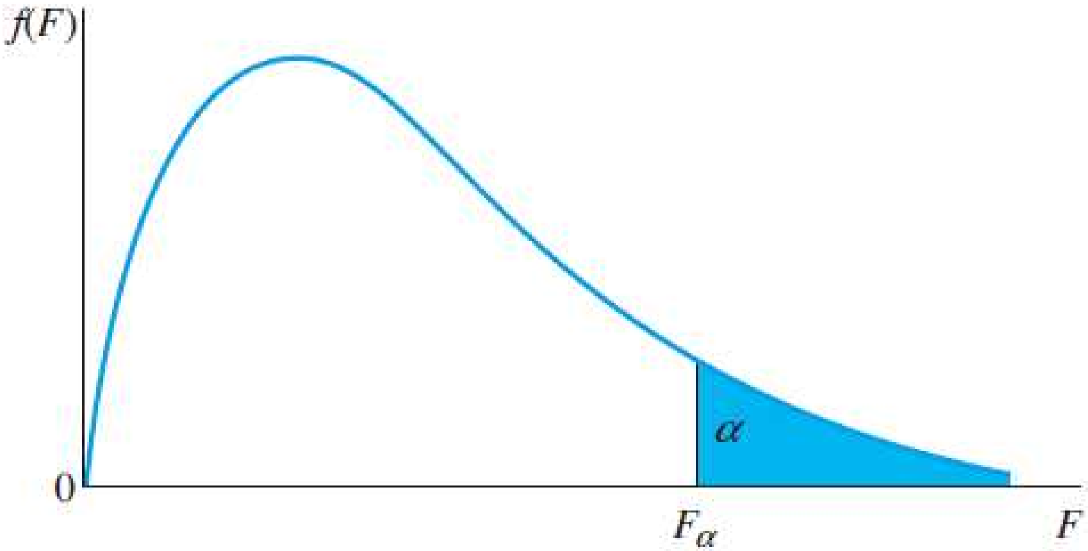
Tests for a Factorial Experiment

- **For interaction:**
  1. Null hypothesis: H_0 : Factors A and B do not interact
  2. Alternative hypothesis: H_a : Factors A and B interact
  3. Test statistic: F = MS(AB)/MSE, where F is based on $df_1 = (a - 1)(b - 1)$ and $df_2 = ab(r - 1)$
  4. Rejection region: Reject H_0 when $F > F_\alpha$, where F_α lies in the upper tail of the F distribution (see the figure), or when the p-value $< \alpha$

- **For main effects, factor A:**
  1. Null hypothesis: H_0 : There are no differences among the factor A means
  2. Alternative hypothesis: H_a : At least two of the factor A means differ
  3. Test statistic: F = MSA/MSE, where F is based on $df_1 = (a - 1)$ and $df_2 = ab(r - 1)$
  4. Rejection region: Reject H_0 when $F > F_\alpha$ (see the figure) or when the p-value $< \alpha$

Tests for a Factorial Experiment

- **For main effects, factor B:**
  1. Null hypothesis: H_0 : There are no differences among the factor B means
  2. Alternative hypothesis: H_a : At least two of the factor B means differ
  3. Test statistic: $F = \text{MSB/MSE}$, where F is based on $df_1 = (b - 1)$ and $df_2 = ab(r - 1)$
  4. Rejection region: Reject H_0 when $F > F_\alpha$ (see the figure) or when the p-value $< \alpha$



Example (drug manufacturer)

- Each supervisors works at each of three different shift times and the shift's output is measured on three randomly selected days.

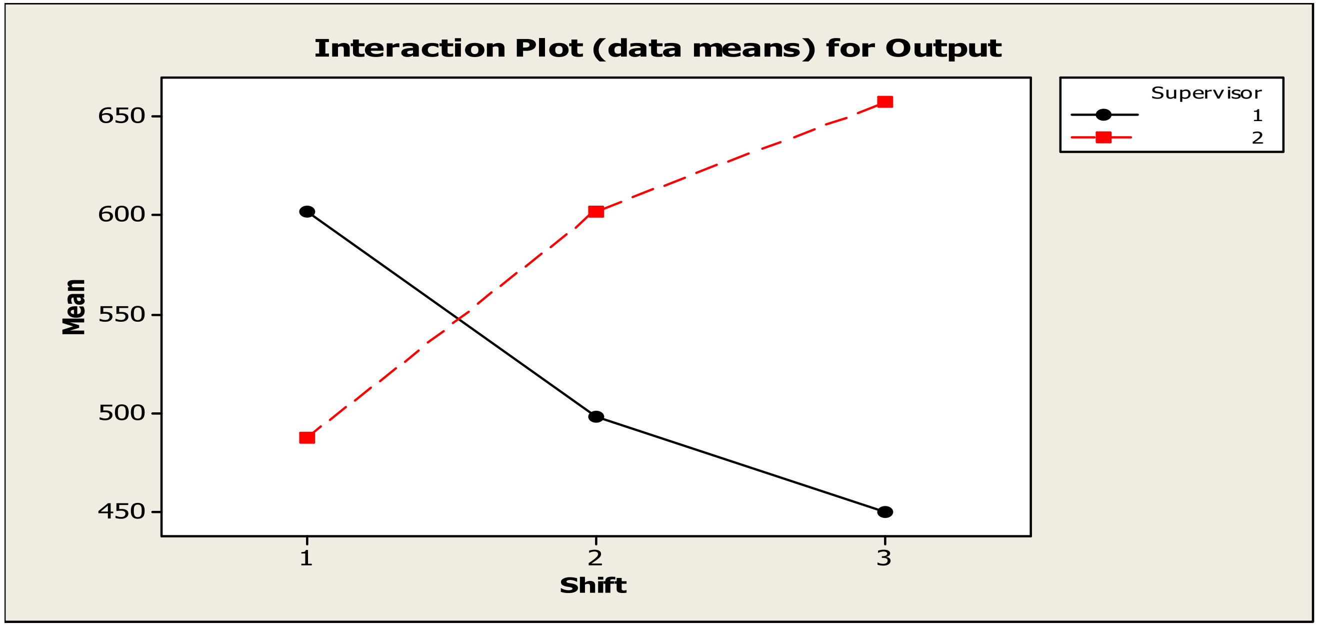
Supervisor	Day	Swing	Night	A_i
1	571 610 625	480 474 540	470 430 450	**4650**
2	480 516 465	625 600 581	630 680 661	**5238**
B_j	**3267**	**3300**	**3321**	**9888**

Example (drug manufacturer)

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Supervisor	19208	1	19208	26.678	0.000	4.747
Shift	247	2	123.5	0.172	0.844	3.885
Interaction	81127	2	40563.5	56.338	0.000	3.885
Within	8640	12	720			
Total	109222	17				

The test statistic for the interaction is F = 56.338 with *p*-value = .000. The interaction is highly significant, and the main effects are not tested. We look at the interaction plot to see where the differences lie.



Supervisor 1 does better earlier in the day, while supervisor 2 does better at night.

Supervisor	Day	Swing	Night
1	602	498	450
2	487	602	657

Revisiting the ANOVA Assumptions

1. The observations within each population are normally distributed with a common variance σ^2.
2. Assumptions regarding the sampling procedures are specified for each design.

✓ Remember that ANOVA procedures are fairly robust when sample sizes are equal and when the data are fairly mound-shaped.

Key Concepts

I. Experimental Designs

1. Experimental units, factors, levels, treatments, response variables.
2. Assumptions: Observations within each treatment group must be normally distributed with a common variance σ^2.
3. One-way classification—completely randomized design: Independent random samples are selected from each of k populations.
4. Two-way classification—randomized block design: k treatments are compared within b *blocks*.
5. Two-way classification — a x b factorial experiment: Two factors, A and B, are compared at several levels. Each factor–level combination is replicated r times to allow for the investigation of an interaction between the two factors.

Key Concepts

II. Analysis of Variance

1. The total variation in the experiment is divided into variation (sums of squares) explained by the various experimental factors and variation due to experimental error (unexplained).

2. If there is an effect due to a particular factor, its mean square(MS = SS/df) is usually large and F = MS(factor)/MSE is large.

3. Test statistics for the various experimental factors are based on F statistics, with appropriate degrees of freedom (df_2 = Error degrees of freedom).

Key Concepts

III. Interpreting an Analysis of Variance

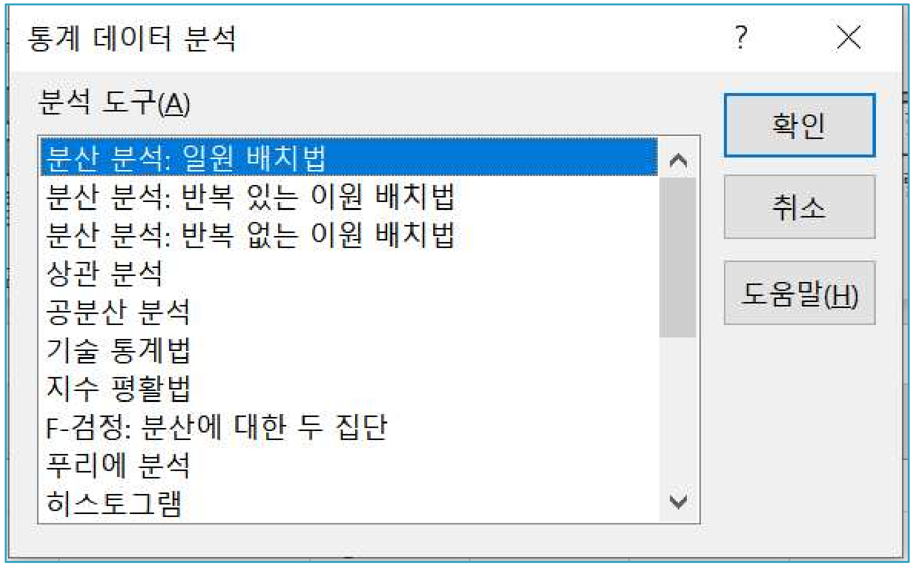
1. For the completely randomized and randomized block design, each factor is tested for significance.
2. For the factorial experiment, first test for a significant interaction. If the interactions is significant, main effects need not be tested. The nature of the difference in the factor–level combinations should be further examined.
3. If a significant difference in the population means is found, Tukey's method of pairwise comparisons or a similar method can be used to further identify the nature of the difference.
4. If you have a special interest in one population mean or the difference between two population means, you can use a confidence interval estimate. (For randomized block design, confidence intervals do not provide estimates for single population means).

Excel : Analysis of Variance

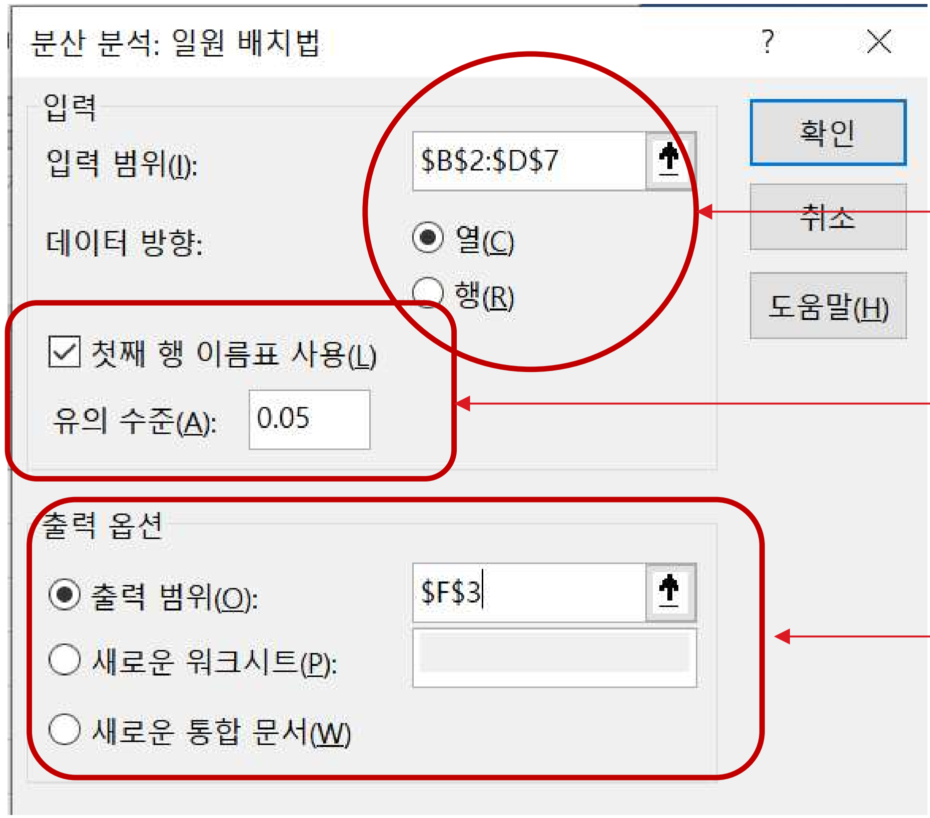
Example1(completely randomized design)

No Breakfast	Light Breakfast	Full Breakfast
8	14	10
7	16	12
9	12	16
13	17	15
10	11	12

1. 데이터 > 데이터분석> 분산분석: 일원배치법



	A	B	C	D
1				
2		no breakfast	Light	Full
3		8	14	10
4		7	16	12
5		9	12	16
6		13	17	15
7		10	11	12



1. 데이터 입력과 방향 입력

2. 데이터의 이름표 사용여부
유의수준

3. 출력 위치 선정

결과물(CRD)

분산 분석: 일원 배치법

요약표

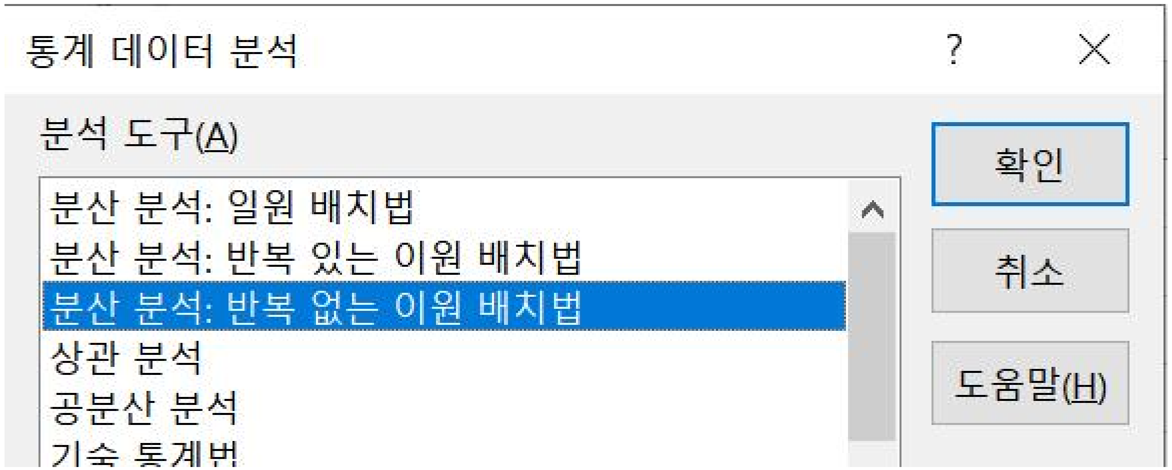
인자의 수준	관측수	합	평균	분산
no breakfast	5	47	9.4	5.3
Light	5	70	14	6.5
Full	5	65	13	6

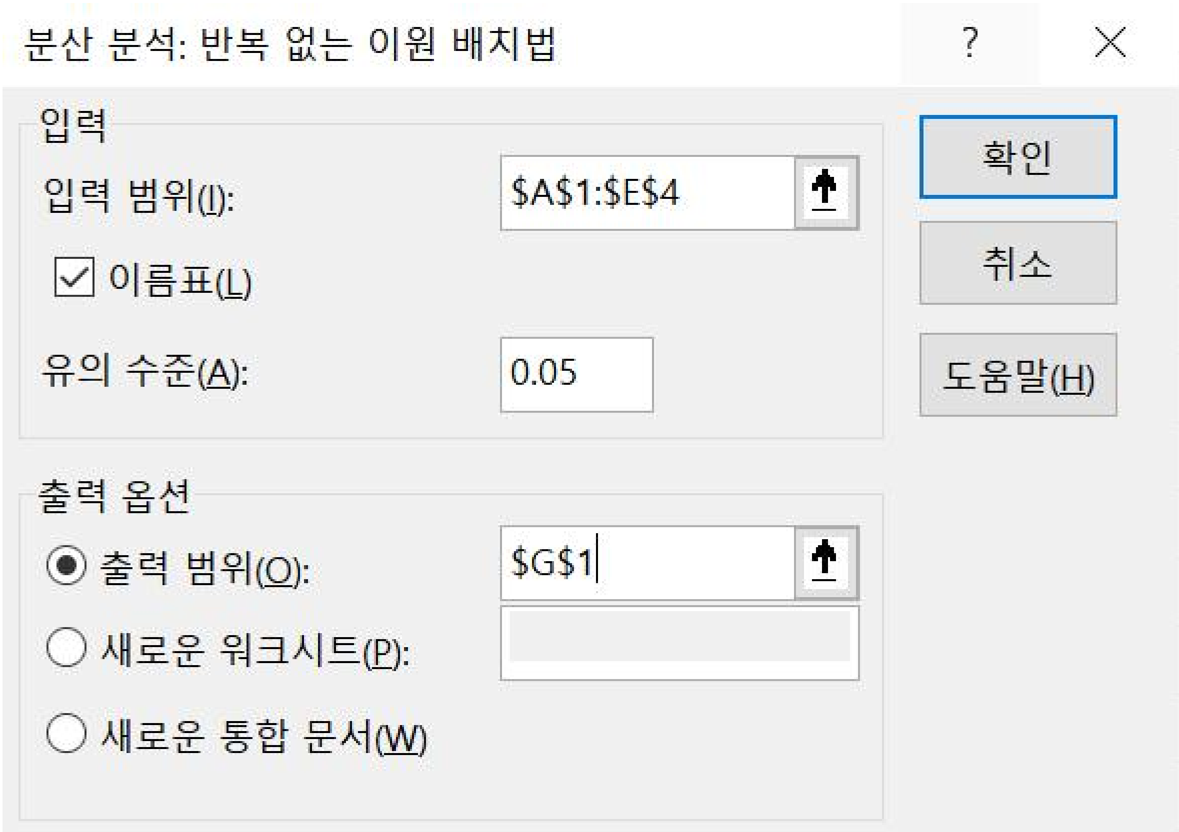
분산 분석

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
처리	58.53333	2	29.26667	4.932584	0.027326	3.885294
잔차	71.2	12	5.933333			
계	129.7333	14				

Example2(Randomized Block Design)

	A	B	C	D	E
1		A	B	C	D
2	Low	27	24	31	23
3	Middle	68	76	65	67
4	High	308	326	312	300




분산 분석: 반복 없는 이원 배치법
?
입력
입력 범위(I):
$A$1:$E$4
이름표(L)
유의 수준(A):
0.05
확인
취소
도움말(H)
출력 옵션
출력 범위(O):
$G$1
새로운 워크시트(P):
새로운 통합 문서(W)

결과물(RBD)

분산 분석: 반복 없는 이원 배치법

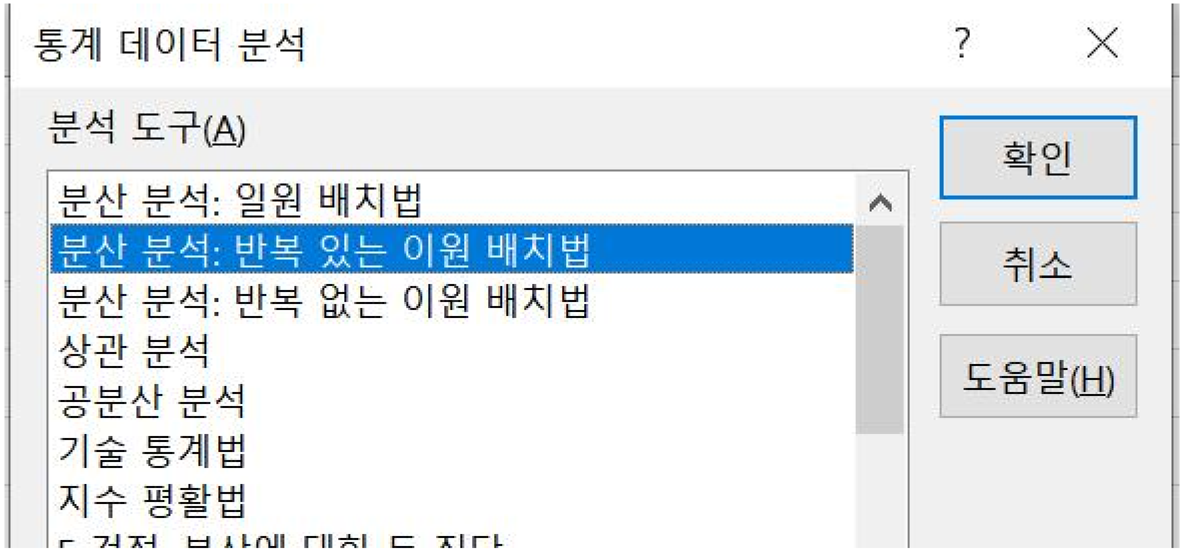
요약표	관측수	합	평균	분산
Low	4	105	26.25	12.91667
Middle	4	276	69	23.33333
High	4	1246	311.5	118.3333
A	3	403	134.3333	23040.33
B	3	426	142	26068
C	3	408	136	23521
D	3	390	130	22159

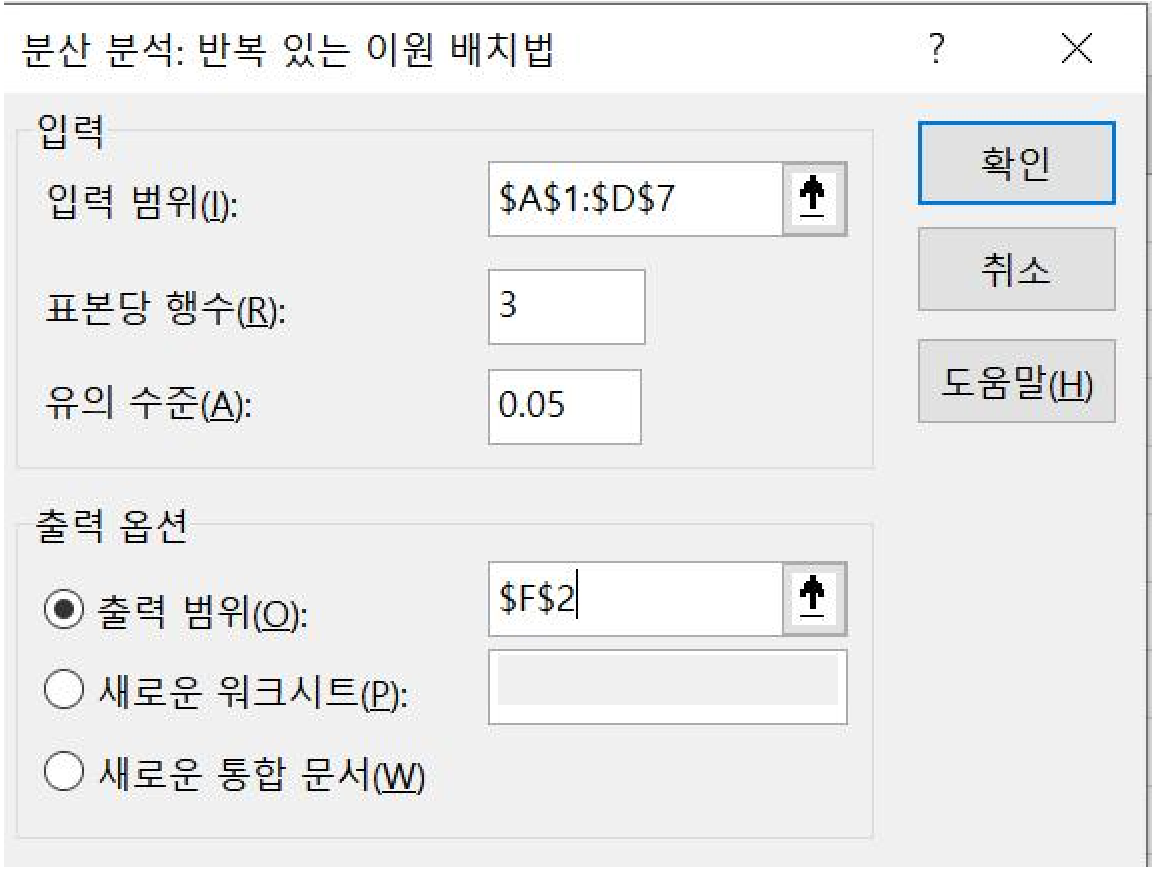
분산 분석

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(행)	189335.2	2	94667.58	2351.99	2.07E-09	5.143253
인자 B(열)	222.25	3	74.08333	1.84058	0.240378	4.757063
잔차	241.5	6	40.25			
계	189798.9	11				

Example3(Factorial Experiment)

	A	B	C	D
1	supervisor	day	swing	Night
2	1	571	480	470
3	1	610	474	430
4	1	625	540	450
5	2	480	625	630
6	2	516	600	680
7	2	465	581	661
8				




분산 분석: 반복 있는 이원 배치법
?
입력
입력 범위(I):
$A$1:$D$7
표본당 행수(R):
3
유의 수준(A):
0.05
출력 옵션
출력 범위(O):
$F$2
새로운 워크시트(P):
새로운 통합 문서(W)
확인
취소
도움말(H)

결과물(FA)

분산 분석: 반복 있는 이원 배치법

요약표	day	swing	Night	계
1				
관측수	3	3	3	9
합	1806	1494	1350	4650
평균	602	498	450	516.6667
분산	777	1332	400	5155.25
2				
관측수	3	3	3	9
합	1461	1806	1971	5238
평균	487	602	657	582
분산	687	487	637	6096.5
계				
관측수	6	6	6	
합	3267	3300	3321	
평균	544.5	550	553.5	
분산	4553.1	3972.4	13269.5	

분산 분석

변동의 요인	제곱합	자유도	제곱 평균	F 비	P-값	F 기각치
인자 A(행)	19208	1	19208	26.67778	0.000235	4.747225
인자 B(열)	247	2	123.5	0.171528	0.844406	3.885294
교호작용	81127	2	40563.5	56.33819	7.95E-07	3.885294
잔차	8640	12	720			
계	109222	17				