Lecture 9. Analysis of Variance (Chapter 15)

Ping Yu

HKU Business School
The University of Hong Kong

Plan of This Lecture

- Comparison of Several Population Means
- One-Way ANOVA
- (**) The Kruskal-Wallis Test
- Two-Way ANOVA: One Observation per Cell, Randomized Blocks
- Two-Way ANOVA: More Than One Observation per Cell

Comparison of Several Population Means

Comparison of Several Population Means

- In Lecture 6, we tested whether two means are equal; sometimes we need to test whether more than two means are equal.
- In the two-level case, say, we want to compare the fuel consumption for two types of cars, A-cares and B-cars, we can either let each driver drive both types of cars (matched pair) or randomly assign drivers to drive each type of cars (independent samples).
- If there are more than two types of cars, how to test whether their mean fuel consumptions are equal?

Table 15.1 Fuel-Consumption Figures from Three Independent Random Samples, in Miles per Gallon

	A-Cars	B-Cars	C-Cars
	22.2	24.6	22.7
	19.9	23.1	21.9
	20.3	22.0	23.2
	21.4	23.5	24.1
	21.2	23.6	22.1
	21.0	22.1	23.4
	20.3	23.5	—
Sums	**146.3**	**162.4**	**137.4**

continue

- The mean fuel consumptions for the three types of cars are 20.9, 23.2 and 22.9, respectively, which are different, but do such differences arise by chance or not?

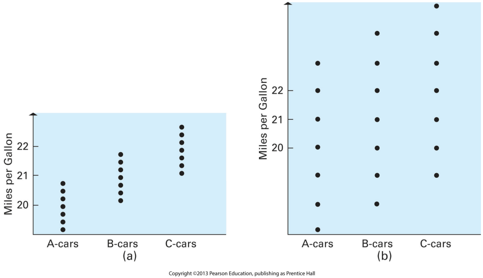


Figure: Two Sets of Sample Fuel-Consumption Data on Three Types of Cars

- Although the sample means in part (a) and part (b) are equal, part (a) suggests different population means while part (b) suggests the same population mean. The key point is the variability around the sample means vs. the variability among the sample means. If the former is small compared with the latter, then we doubt the population means are equal.

One-Way ANOVA

The Setup of One-Way ANOVA

Population			
1	2	$\cdots$	K
x_{11}	x_{21}	$\cdots$	x_{K1}
x_{12}	x_{22}	$\cdots$	x_{K2}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_{1n_1}	x_{2n_2}	$\cdots$	x_{Kn_K}

Table 15.2: Independent (not paired) Random Samples from K Populations

- In the K populations of Table 15.2, we assume each population has the same variance.
- H_0: $\mu_1 = \mu_2 = \cdots = \mu_K$ vs. H_1: $\mu_k \neq \mu_l$ for at least one pair (k, l), where μ_i is the population mean of population i.
- $\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$ is the sample mean of group i.
- $\overline{\overline{x}} = \frac{\sum_{i=1}^{K} \sum_{j=1}^{n_i} x_{ij}}{n} = \frac{\sum_{i=1}^{K} n_i \bar{x}_i}{n}$ is the common mean of all groups, where $n = \sum_{i=1}^{K} n_i$ is the total number of observations.
  - In the example above, $n_1 = n_2 = 7$, $n_3 = 6$, $n = 20$, $K = 3$, and $\overline{\overline{x}} = 22.3$.

Sum of Squares

- From the example above, an appropriate test should be based on assessment of two types of variabilities: within-groups variablity and between-groups variability.
- From Lecture 1, we know variability is well measured by the sum of squared deviations of the observations about the sample mean.
- As a result, we define the sum of squares within group i as

$$SS_i = \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_i\right)^2,$$

the sum of squares within (all K) groups as

$$SSW = \sum_{i=1}^{K} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_i\right)^2 = \sum_{i=1}^{K} SS_i,$$

the sum of squares between groups as

$$SSG = \sum_{i=1}^{K} n_i \left(\bar{x}_i - \bar{\bar{x}}\right)^2,$$

and the total sum of squares as

$$SST = \sum_{i=1}^{K} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{\bar{x}}\right)^2,$$

where in SSG, a weight n_i is imposed on group i, i.e., larger groups are given larger weights.

Sum of Squares Decomposition

- It turns out that

$$SST = SSW + SSG.$$

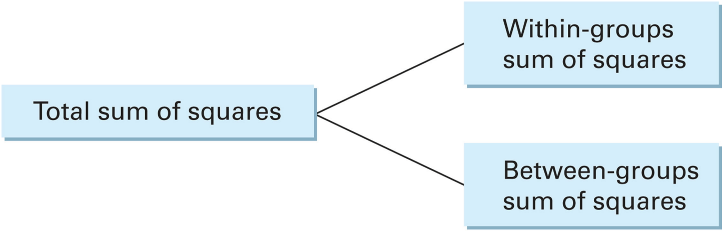


Figure: Sum of Squares Decomposition for One-Way ANOVA

Justification of Sum of Squares Decomposition

- Note that

$$\begin{aligned} SST &= \sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(x_{ij}-\overline{\overline{x}}\right)^2 \\ &= \sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(x_{ij}-\bar{x}_i+\bar{x}_i-\overline{\overline{x}}\right)^2 \\ &= \sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2+\sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(\bar{x}_i-\overline{\overline{x}}\right)^2+2\sum_{i=1}^{K}\left(\bar{x}_i-\overline{\overline{x}}\right)\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i) \\ &= \sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2+\sum_{i=1}^{K}n_i\left(\bar{x}_i-\overline{\overline{x}}\right)^2 \\ &= SSW+SSG, \end{aligned}$$

where $\sum_{j=1}^{n_i}\left(x_{ij}-\bar{x}_i\right)=0$ in the second-to-last equality.

Testing Procedure

- The basic idea is that under H_0, both SSW and SSG can be used to estimate the common population variance.
- Like $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ is an unbiased estimator of σ^2, we need to adjust the dfs for SSW and SSG. The df of SSW is $n - K$ because there are n summands and K estimated parameters $(\mu_1, \cdots, \mu_K)$, and the df of SSW is $K - 1$ because there are K summands and one estimate parameter (the common μ).
- Test Statistic:

$$f = \frac{MSG}{MSW} := \frac{SSG/(K-1)}{SSW/(n-K)},$$

which follows the $F_{K-1,n-K}$ distribution under H_0 and (i) the population variances are equal; (ii) the population distributions are normal. (i.e., $x_{ij} \sim N\left(\mu_i, \sigma^2\right)$)
- $MSW = SSW/(n-K)$ is called the within-groups mean square and is an unbiased estimator of σ^2 regardless whether H_0 holds or not, and
$MSG = SSG/(K-1)$ is called the between-groups mean square and is an unbiased estimator of σ^2 only under H_0. [see Appendices 2 and 3 of Chapter 15 for detailed analyses]

continue

- Why? Recall that $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$ from Lecture 5. Similarly,

$$\begin{aligned} \frac{(K-1)\,MSG}{\sigma^2} &\sim \chi^2_{K-1}, \\ \frac{(n-K)\,MSW}{\sigma^2} &\sim \chi^2_{n-K}, \end{aligned}$$

  and these two chi-square distributions are independent. From the definition of the F distribution, we have

$$\frac{\frac{(K-1)MSG}{\sigma^2} / (K-1)}{\frac{(n-K)MSW}{\sigma^2} / (n-K)} = \frac{MSG}{MSW} \sim F_{K-1,n-K}.$$

- Decision Rule: reject H_0 if $f > F_{K-1,n-K,\alpha}$. [why? Under H_1, $E[MSG] > \sigma^2$]
- The p-value is $P(F > f)$, where $F \sim F_{K-1,n-K}$.

Example Continued

- The test can be conveniently summarized in a one-way anova table:

Table 15.3
General Format of One-Way Analysis of Variance Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F Ratio
Between groups	SSG	$K-1$	$MSG = \frac{SSG}{K-1}$	$\frac{MSG}{MSW}$
Within groups	SSW	$n-K$	$MSW = \frac{SSW}{n-K}$	
Total	SST	$n-1$		

- In the example above, the one-way anova table is as follows:

Table 15.4
One-Way Analysis of Variance Table for Fuel-Consumption Data

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F Ratio
Between groups	21.55	2	10.78	15.05
Within groups	12.18	17	0.7165	
Total	33.73	19		

- Because $F_{2,17,0.01} = 6.112$ and $15.05 > 6.112$, we reject H_0 at the 1% level.

(**) Multiple Comparisons Between Subgroup Means

- If we reject H_0, a natural question followed is that which subgroup means are different from others. Specifically, we want a minimal interval that could be used to decide if any two subgroup means are different in a statistical sense.
- In the two-level case, we know when $|\bar{x} - \bar{y}|$ is larger than a critical value, called a minimum significant difference (MSD), we can conclude that $\mu_x \neq \mu_y$ in a statistical sense, e.g., if $\sigma_x^2 = \sigma_y^2$, then

$$MSD = t_{\alpha/2} s_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}.$$

- This MSD does not work for the $K > 2$ case because there are C_2^K comparisons of sample means such that the probability of error α would no longer hold.
- Intuitively, the correct MSD should be increasing in K, i.e., a larger MSD is required than the two-level case when $K > 2$.

continue

- We use a procedure proposed by John Tukey to deal with this multiple-comparisons question:

$$MSD(K) = Q\frac{s_p}{\sqrt{n}},$$

where the factor Q depends on α, K and $n-K$ (df of SSW) and is listed in Appendix Table 13, and s_p is the pooled standard deviation, i.e., $s_p = \sqrt{MSW}$.
- Q is also decreasing in $n-K$ and α besides increasing in K.

Population Model for One-Way ANOVA

- Let X_{ij} be the r.v. corresponding to the jth observation from the ith population, and decompose X_{ij} as

$$X_{ij} = \mu_i + \varepsilon_{ij},$$

where the error $\varepsilon_{ij} \overset{iid}{\sim} N\left(0, \sigma^2\right)$.

- Let μ be the overall mean of the K combined populations, and $G_i := \mu_i - \mu$; then

$$X_i = \mu + G_i + \varepsilon_{ij},$$

i.e., an observation is made up of the sum of an overall mean μ, a group-specific term G_i, and a random error ε_{ij}, [figure here] and H_0 can be equivalently stated as

$$H_0 : G_1 = \cdots = G_K = 0.$$

- SSW is also called the error sum of squares because it can be used to estimate the "error" variance σ^2.

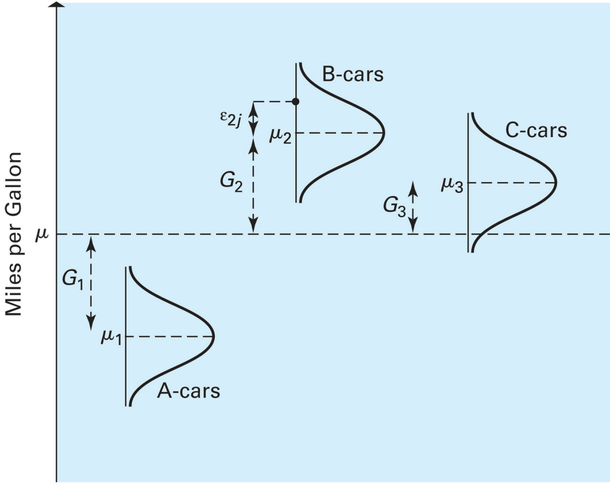


Figure: Illustration of the Population Model for the One-Way ANOVA

(**) The Kruskal-Wallis Test

The Kruskal-Wallis H Test

- This Kruskal-Wallis H test is a nonparametric counterpart of the one-way ANOVA where normality is not assumed; it is also an extension of the Mann-Whitney test in Lecture 8 to the $K > 2$ case.
- The null is that all subgroups have the same distribution which will reduce to the same median when all other aspects of the K distributions except the central location are the same.
- Like the Mann-Whitney test, we pool all samples and rank them in ascending order with the rank of x_{ij} being r_{ij}, and define

$$R_i = \sum_{j=1}^{n_i} r_{ij}, i = 1, \cdots, K,$$

  as the sum of the ranks for subgroup i.
- Also define

$$\bar{r}_i = \frac{R_i}{n_i}$$

  and

$$\bar{\bar{r}} = \frac{1}{n} \sum_{i=1}^{K} \sum_{j=1}^{n_i} r_{ij},$$

  which are the counterparts of $\bar{x}_i$ and $\bar{\bar{x}}$.

continue

- The test statistic is

$$H = (n-1)\frac{\sum_{i=1}^{K} n_i\left(\bar{r}_i - \bar{\bar{r}}\right)^2}{\sum_{i=1}^{K}\sum_{j=1}^{n_i}\left(r_{ij} - \bar{\bar{r}}\right)^2} := (n-1)\frac{SSG}{SST}.$$

  - Since $SST = SSG + SSW$, H is an increasing function of SSG and a large H will induce rejection of H_0,

- Where there are no ties, it is not hard to show that the denominator of H is equal to $\frac{(n-1)n(n+1)}{2}$ and $\bar{\bar{r}} = \frac{n+1}{2}$, which implies the W on Page 663:

$$H = \frac{12}{n(n+1)}\sum_{i=1}^{K}\frac{R_i^2}{n_i} - 3(n+1).$$

- Under H_0, $\frac{SST}{n-1} \to \sigma^2$ and $\frac{SSG}{\sigma^2} \to \chi^2_{K-1}$, where σ^2 is the variance of r_{ij}, so H approximately follow the χ^2_{K-1} distribution. As a result, the decision rule is to reject H_0 if $H > \chi^2_{K-1,\alpha}$.
- (**) The counterpart of MSD is Dunn's test which can be used to detect which of the sample pairs are different.

Example Continued

Table 15.6 Fuel-Consumption Figures (in Miles per Gallon) and Ranks from Three Independent Random Samples

A-Cars	Rank	B-Cars	Rank	C-Cars	Rank
22.2	11	24.6	20	22.7	12
19.9	1	23.1	13	21.9	7
20.3	2.5	22.0	8	23.2	14
21.4	6	23.5	16.5	24.1	19
21.2	5	23.6	18	22.1	9.5
21.0	4	22.1	9.5	23.4	15
20.3	2.5	23.5	16.5		
Rank sum	**32**		**101.5**		**76.5**

- $R_1 = 32, R_2 = 101.5$ and $R_3 = 76.5$, so

$$H = \frac{12}{20 \times 21}\left[\frac{32^2}{7} + \frac{101.5^2}{7} + \frac{76.5^2}{6}\right] - 3 \times 21 = 11.10.$$

Since $\chi^2_{2,0.01} = 9.210$, we reject H_0 at the 1% level, same conclusion as the one-way ANOVA.

Two-Way ANOVA: One Observation per Cell, Randomized Blocks

Example Continued

- If we can isolate other factors from ε_{ij}, e.g., driver habits, then we can reduce the variance of the error term and improve the test power.

Table 15.7 Sample Observations on Fuel Consumption Recorded for Three Types of Automobiles Driven by Drivers in Six Classes

	Automobile Type			
Driver Class	α-Cars	β-Cars	γ-Cars	Sum
1	25.1	23.9	26.0	75.0
2	24.7	23.7	25.4	73.8
3	26.0	24.4	25.8	76.2
4	24.3	23.3	24.4	72.0
5	23.9	23.6	24.2	71.7
6	24.2	24.5	25.4	74.1
Sum	148.2	143.4	151.2	442.8

- The data in Table 15.7 are like matched pairs in Lecture 6, where the additional variable, drivers (or driver ages), is called a block variable and the experiment is said to be arranged in six blocks.
- This kind of design is called a randomized blocks design because a driver is randomly selected from each (randomized) age class to drive each type of car (i.e., 18 drivers are randomly drawn but belong to only 6 classes).

The Setup of Two-Way ANOVA

Table 15.8
Sample Observation on K Groups and H Blocks

	Group			
Block	1	2	...	K
1	x_{11}	x_{21}	...	x_{K1}
2	x_{12}	x_{22}	...	x_{K2}
.	.	.		.
.	.	.		.
.	.	.		.
H	x_{1H}	x_{2H}	...	x_{KH}

- H_0^G: K group means $\{\mu_i\}_{i=1}^K$ are the same or H_0^B: H block means $\left\{\mu_j\right\}_{j=1}^H$ are the same vs. H_1: negation of H_0
- $\bar{x}_{i\cdot} = \frac{\sum_{j=1}^H x_{ij}}{H}$ is the sample mean of group i (estimator of $\mu_i = E\left[X_{ij}|\text{group } i\right]$).
- $\bar{x}_{\cdot j} = \frac{\sum_{i=1}^K x_{ij}}{K}$ is the sample mean of block j (estimator of $\mu_j = E\left[X_{ij}|\text{block } j\right]$).
- $\bar{\bar{x}} = \frac{\sum_{i=1}^K \sum_{j=1}^H x_{ij}}{n} = \frac{\sum_{i=1}^K \bar{x}_{i\cdot}}{K} = \frac{\sum_{j=1}^H \bar{x}_{\cdot j}}{H}$ is the common mean of all groups (estimator of the "overall" mean $\mu = E\left[X_{ij}\right]$), where $n = HK$ is the total number of observations.
  - In the example above, $K = 3, H = 6, n = 18, \bar{x}_{1\cdot} = \frac{148.2}{6} = 24.7, \bar{x}_{\cdot 1} = \frac{75}{3} = 25$, etc. and $\bar{\bar{x}} = 24.6$.

Population Model for Two-Way ANOVA

- Let X_{ij} be the r.v. corresponding to the observation for the ith group and jth block, and decompose X_{ij} as
$$X_{ij} = \mu + G_i + B_j + \varepsilon_{ij},$$
where $G_i = \mu_i - \mu$, $B_j = \mu_j - \mu$, and the error $\varepsilon_{ij} \overset{iid}{\sim} N\left(0, \sigma^2\right)$.
- Rewrite this as
$$X_{ij} - \mu = G_i + B_j + \varepsilon_{ij},$$
where μ, G_i and B_j are estimated by $\overline{\overline{x}}, \bar{x}_{i\cdot} - \overline{\overline{x}}$ and $\bar{x}_{\cdot j} - \overline{\overline{x}}$, respectively, so
$$\left(x_{ij} - \overline{\overline{x}}\right) - \left(\bar{x}_{i\cdot} - \overline{\overline{x}}\right) - \left(\bar{x}_{\cdot j} - \overline{\overline{x}}\right) = x_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \overline{\overline{x}},$$
which implies
$$x_{ij} - \overline{\overline{x}} = \left(\bar{x}_{i\cdot} - \overline{\overline{x}}\right) + \left(\bar{x}_{\cdot j} - \overline{\overline{x}}\right) + \left(x_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \overline{\overline{x}}\right),$$
where the last term is an estimator of ε_{ij}.
- We have decomposed the deviation of x_{ij} from $\overline{\overline{x}}$ into the group effect $\bar{x}_{i\cdot} - \overline{\overline{x}}$, the block effect $\bar{x}_{\cdot j} - \overline{\overline{x}}$, and the random error due to chance variability or experimental error.

Sum of Squares Decomposition

- By similar arguments as in one-way ANOVA, we can show

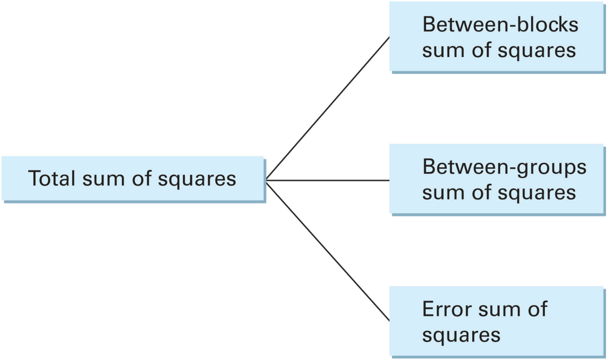
$$SST = SSG + SSB + SSE,$$

where

$$SST = \sum_{i=1}^{K}\sum_{j=1}^{H}\left(x_{ij} - \overline{\overline{x}}\right)^2 \text{ is the total sum of squares,}$$
$$SSG = H\sum_{i=1}^{K}\left(\bar{x}_{i\cdot} - \overline{\overline{x}}\right)^2 \text{ is the between-groups sum of squares}$$
$$SSB = K\sum_{j=1}^{H}\left(\bar{x}_{\cdot j} - \overline{\overline{x}}\right)^2 \text{ is the between-blocks sum of squares,}$$
$$SSE = \sum_{i=1}^{K}\sum_{j=1}^{H}\left(x_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \overline{\overline{x}}\right)^2 \text{ is the error sum of squares.}$$

- The analysis is called two-way ANOVA because the data are categorized in two ways, accordingly to groups and blocks.
- In one-way ANOVA, $x_{ij} - \overline{\overline{x}} = \left(\bar{x}_i - \overline{\overline{x}}\right) + \left(x_{ij} - \bar{x}_i\right)$, so $SST = SSG + SSE$, where SSE is termed as SSW, and n_i in SSG need not be the same so cannot be taken out of the summation.

continue



- Compared with one-way ANOVA, the extra component arises because we can extract from the data about differences among blocks.

Testing Procedure

- The basic idea is the same as in one-way ANOVA.

Table 15.9 General Format of Two-Way Analysis of Variance Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F Ratio
Between groups	SSG	$K-1$	$MSG = \frac{SSG}{K-1}$	$\frac{MSG}{MSE}$
Between blocks	SSB	$H-1$	$MSB = \frac{SSB}{H-1}$	$\frac{MSB}{MSE}$
Error	SSE	$(K-1)(H-1)$	$MSE = \frac{SSE}{(K-1)(H-1)}$	
Total	SST	$n-1$		

Note: the df of SSE is $(n-1)-(K-1)-(H-1)=(K-1)(H-1)$.

- Test Statistic: If $\varepsilon_{ij} \overset{iid}{\sim} N\left(0,\sigma^2\right)$, then

$$f^G = \frac{MSG}{MSE} := \frac{SSG/(K-1)}{SSE/(K-1)(H-1)}$$

will follow the $F_{K-1,(K-1)(H-1)}$ distribution under H_0^G, and

$$f^B = \frac{MSB}{MSE} := \frac{SSB/(H-1)}{SSE/(K-1)(H-1)}$$

will follow the $F_{H-1,(K-1)(H-1)}$ distribution under H_0^B.

continue

- Decision Rule: reject H_0^G if $f^G > F_{K-1,(K-1)(H-1),\alpha}$ and reject H_0^B if $f^B > F_{H-1,(K-1)(H-1),\alpha}$.
- The p-value in testing H_0^G is $P\left(F^G > f^G\right)$ and in testing H_0^B is $P\left(F^B > f^B\right)$, where $F^G \sim F_{K-1,(K-1)(H-1)}$ and $F^B \sim F_{H-1,(K-1)(H-1)}$.
- In the example above, the two-way anova table is as follows:

Two-way ANOVA: Mileage versus Car, Driver

Source	DF	SS	MS	F	P
Car	2	5.16	2.580	14.83	0.001
Driver	5	4.98	0.996	5.72	0.009
Error	10	1.74	0.174		
Total	17	11.88			

- Because both p values are less than 1%, we reject both H_0^G and H_0^B at the 1% level, where note that $K = 3, H = 6$ so $(K-1)(H-1) = 10$.

More Comments

- This two-way ANOVA is often called one-way ANOVA with block effects.
- Usually, i indexes the treatments, j indexes the blocks, and we are mainly interested in whether the treatment effects are equal, i.e., H_0^G, given that the block effects are different by construction.
- ε_{ij} is assumed to be independent across i, but need not be independent across i in general.
- Anyway, $Corr\left(\varepsilon_{ij}, \varepsilon_{i'j}\right)$ should be the same for any pair of (i, i'). In this case, we need to be sure the order in which treatments are administered to subjects is ranomized in order to assume equal correlation.
  - Of course, the same

Two-Way ANOVA: More Than One Observation per Cell

Example Continued

- If there are more than one observation for each group and block (or each cell), i.e., each type of care are driven by more than one drivers from each age class, we have two advantages: (i) more data mean more precise estimation and thus higher power; (ii) allow the isolation of a further source of variability – the interaction between groups and blocks.

Table 15.10 Sample Observations on Fuel Consumption Recorded for Three Types of Automobiles Driven by Five Classes of Drivers; Three Observations per Cell

	Automobile Type								
Driver Class	X-Cars			Y-Cars			Z-cars		
1	25.0	25.4	25.2	24.0	24.4	23.9	25.9	25.8	25.4
2	24.8	24.8	24.5	23.5	23.8	23.8	25.2	25.0	25.4
3	26.1	26.3	26.2	24.6	24.9	24.9	25.7	25.9	25.5
4	24.1	24.4	24.4	23.9	24.0	23.8	24.0	23.6	23.5
5	24.0	23.6	24.1	24.4	24.4	24.1	25.1	25.2	25.3

More on Interaction

- The interactions occur when differences in block (group) effects are not distributed uniformly across groups (blocks).
- Mathematically, this means that $\mu_{ij} - \mu_i \neq \mu_j - \mu$ for some i and j, where μ_{ij} is the population mean in the ijth cell, and μ_i, μ_j and μ are defined similarly as in the last section.
- Note that $\mu_j = \frac{\sum_{i=1}^K \mu_{ij}}{K}$ and $\mu = \frac{\sum_{i=1}^K \mu_i}{K}$, so $\mu_j - \mu = \frac{\sum_{i=1}^K \left(\mu_{ij} - \mu_i\right)}{K}$.
- $\mu_{ij} - \mu_i \neq \mu_j - \mu$ for some i and j implies there exist j, i and i' such that $\left(\mu_{ij} - \mu_i\right) - \left(\mu_j - \mu\right) > 0 > \left(\mu_{i'j} - \mu_i\right) - \left(\mu_j - \mu\right)$.
- That is what the following sentence in the textbook means, "drivers who achieve better-than-average fuel consumption figures may be considerably more successful in getting better fuel economy than other drivers when driving an α-car than when driving a β-car".

The Setup of Two-Way ANOVA: More Than One Observation per Cell

Table 15.11
Sample Observations on K Groups and H Blocks; m Observations per Cell

	Group			
Block	1	2	...	K
1	$x_{111}x_{112}\dots x_{11m}$	$x_{211}x_{212}\dots x_{21m}$	...	$x_{K11}x_{K12}\dots x_{K1m}$
2	$x_{121}x_{122}\dots x_{12m}$	$x_{221}x_{222}\dots x_{22m}$		$x_{K21}x_{K22}\dots x_{K2m}$
.	.	.		.
.	.	.		.
.	.	.		.
H	$x_{1H1}x_{1H2}\dots x_{1Hm}$	$x_{2H1}x_{2H2}\dots x_{2Hm}$	...	$x_{KH1}x_{KH2}\dots x_{KHm}$

Note: there are m observations in the ijth cell, $\{x_{ijl}\}_{l=1}^m$

- H_0^G: $\{\mu_i\}_{i=1}^K$ are the same or H_0^B: $\{\mu_j\}_{j=1}^H$ are the same or H_0^I: $\mu_{ij}-\mu_i=\mu_j-\mu$ for all i and j (i.e., no group-block interaction) vs. H_1: negation of H_0.
- Group means: $\bar{x}_{i\cdot\cdot}=\frac{\sum_{j=1}^H\sum_{l=1}^m x_{ijl}}{Hm}$ is the sample mean of group i.
- Block means: $\bar{x}_{\cdot j\cdot}=\frac{\sum_{i=1}^K\sum_{l=1}^m x_{ijl}}{Km}$ is the sample mean of block j.
- Cell means: $\bar{\bar{x}}_{ij\cdot}=\frac{\sum_{l=1}^m x_{ijl}}{m}$ is the sample mean of cell (i,j).
- Overall mean: $\bar{\bar{x}}=\frac{\sum_{i=1}^K\sum_{j=1}^H\sum_{l=1}^m x_{ijl}}{n}=\frac{\sum_{i=1}^K\bar{x}_{i\cdot\cdot}}{K}=\frac{\sum_{j=1}^H\bar{x}_{\cdot j\cdot}}{H}=\frac{\sum_{i=1}^K\sum_{j=1}^H\bar{\bar{x}}_{ij\cdot}}{KH}$ is the sample mean of all observations, where $n=HKm$ is the total number of observations.

Population Model for Two-Way ANOVA with More Than One Observation per Cell

- Let X_{ijl} be the r.v. corresponding to the lth observation in the ijth cell, and decompose X_{ijl} as

$$X_{ijl} = \mu + G_i + B_j + L_{ij} + \varepsilon_{ijl},$$

where $L_{ij} = \mu_{ij} - \mu_i - \mu_j + \mu$, G_i and B_j are defined similarly as in the last section, and the error $\varepsilon_{ijl} \overset{iid}{\sim} N\left(0, \sigma^2\right)$.

- Rewrite this as

$$X_{ijl} - \mu = G_i + B_j + \varepsilon_{ijl};$$

the corresponding sample version is

$$x_{ijl} - \overline{\overline{x}} = \left(\bar{x}_{i\cdot\cdot} - \overline{\overline{x}}\right) + \left(\bar{x}_{\cdot j\cdot} - \overline{\overline{x}}\right) + \left(\bar{x}_{ij} - \bar{x}_{i\cdot\cdot} - \bar{x}_{\cdot j\cdot} + \overline{\overline{x}}\right) + \left(x_{ijl} - \bar{x}_{ij}\right),$$

where the last term is an estimator of ε_{ijl}.

- Note: the number of observations in each cell need not be the same, but the formulae below would then be much more involved.

Sum of Squares Decomposition

- By similar arguments as in two-way ANOVA with one observation per cell, we can show

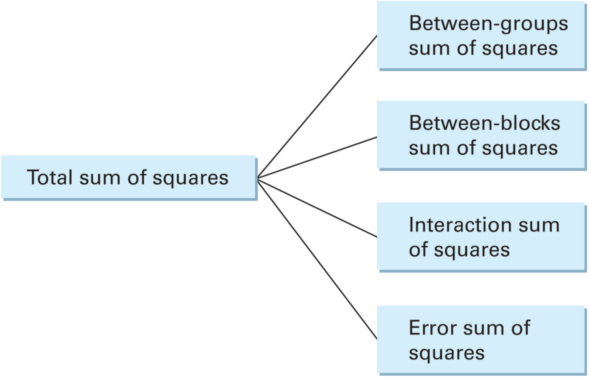
$$SST = SSG + SSB + SSI + SSE,$$

where

	Sum of Squares	Degrees of Freedom	
Total:	$\text{SST} = \sum_i \sum_j \sum_l (x_{ijl} - \overline{\overline{x}})^2$	$KHm - 1$	(15.20)
Between groups:	$\text{SSG} = Hm \sum_{i=1}^{K} (\overline{x}_{i\cdot\cdot} - \overline{\overline{x}})^2$	$K - 1$	(15.21)
Between blocks:	$\text{SSB} = Km \sum_{j=1}^{H} (\overline{x}_{\cdot j\cdot} - \overline{\overline{x}})^2$	$H - 1$	(15.22)
Interaction:	$\text{SSI} = m \sum_{i=1}^{K} \sum_{j=1}^{H} (\overline{x}_{ij\cdot} - \overline{x}_{i\cdot\cdot} - \overline{x}_{\cdot j\cdot} + \overline{\overline{x}})$	$(K-1)(H-1)$	(15.23)
Error:	$\text{SSE} = \sum_i \sum_j \sum_l (x_{ijl} - \overline{x}_{ij\cdot})^2$	$HK(m-1)$	(15.24)

- The dfs of all sums of squares can be derived similarly as before.

continue



- Compared with the sum of squares decomposition for the two-way ANOVA with one observation per cell, the extra component arises because we can isolate an interaction sum of squares.

Testing Procedure

- The basic idea is the same as in two-way ANOVA with one observation per cell.

Table 15.12
General Format of the Two-Way Analysis of Variance Table with m Observations per Cell

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F Ratio
Between groups	SSG	$K-1$	$MSG = \frac{SSG}{K-1}$	$\frac{MSG}{MSE}$
Between blocks	SSB	$H-1$	$MSB = \frac{SSB}{H-1}$	$\frac{MSB}{MSE}$
Interaction	SSI	$(K-1)(H-1)$	$MSI = \frac{SSI}{(K-1)(H-1)}$	$\frac{MSI}{MSE}$
Error	SSE	$KH(m-1)$	$MSE = \frac{SSE}{KH(m-1)}$	
Total	SST			

- Test Statistic: If $\varepsilon_{ijl} \overset{iid}{\sim} N\left(0,\sigma^2\right)$, then

$$f^G = \frac{MSG}{MSE} := \frac{SSG/(K-1)}{SSE/KH(m-1)}, f^B = \frac{MSB}{MSE} := \frac{SSB/(H-1)}{SSE/KH(m-1)},$$
$$f^I = \frac{MSI}{MSE} := \frac{SSI/(K-1)(H-1)}{SSE/KH(m-1)}$$

will follow the $F_{K-1,KH(m-1)}$, $F_{H-1,KH(m-1)}$ and $F_{(K-1)(H-1),KH(m-1)}$ distributions under H_0^G, H_0^B and H_0^I, respectively.

continue

- Decision Rule: reject H_0^G, H_0^B and H_0^I if $f^G > F_{K-1,KH(m-1),\alpha}, f^B > F_{H-1,KH(m-1),\alpha}$ and $f^I > F_{(K-1)(H-1),KH(m-1),\alpha}$, respectively.
- The p-values in testing H_0^G, H_0^B and H_0^I are $P\left(F^G > f^G\right)$, $P\left(F^B > f^B\right)$, and $P\left(F^I > f^I\right)$, where $F^G \sim F_{K-1,KH(m-1)}$, $F^B \sim F_{H-1,KH(m-1)}$ and $F^I \sim F_{(K-1)(H-1),KH(m-1)}$.
- In the example above, the two-way anova table is as follows:

Two-way ANOVA: Mileage versus Car, Driver

Source	DF	SS	MS	F	P
Car	2	7.156	3.57800	92.53	0.000
Driver	4	13.148	3.28700	85.01	0.000
Interaction	8	6.604	0.82550	21.35	0.000
Error	30	1.160	0.03867		
Total	44	28.068			

- Because all p values are zero, we strongly reject all nulls, especially, we conclude that there exists an interaction effect, where note that $K = 3, H = 5$ and $m = 3$.