MAT 458-Design of Experiments

(III) Experiments with a Single Factor: The Analysis of Variance

Fuxia Cheng

Outline

1. Introduction of ANOVA

2. ANOVA for the Fixed Effects Model with a Single Factor

- An Example
- Models for the Data
- ANOVA
- Model Adequacy Checking in the ANOVA
- Post-ANOVA Comparison of Means
- Sample Size Determination

3. The Random Effects Model with a Single Factor

- A Single Random Factor
- ANOVA for the Random Model
- Estimating the Model Parameters

1. Introduction of ANOVA

We have considered experiments to compare two treatments, i.e., simple comparative experiments. What if there are more than two treatments (or more than two factor levels)?

- The t-test does not directly apply.

- There are lots of practical situations where there are either more than two levels of interest, or there are several factors of simultaneous interest.

- The analysis of variance (ANOVA) is the appropriate analysis for these types of experiments.

- The ANOVA was developed by Fisher in the early 1920s, and initially applied to agricultural experiments.

- The ANOVA is used extensively today for industrial experiments.

Here we will consider ANOVA for the fixed effects models and the random effects models with a single-factor completely randomized experiment.

2. ANOVA for the Fixed Effects Model with a Single Factor

- An Example (See pg. 65) :

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

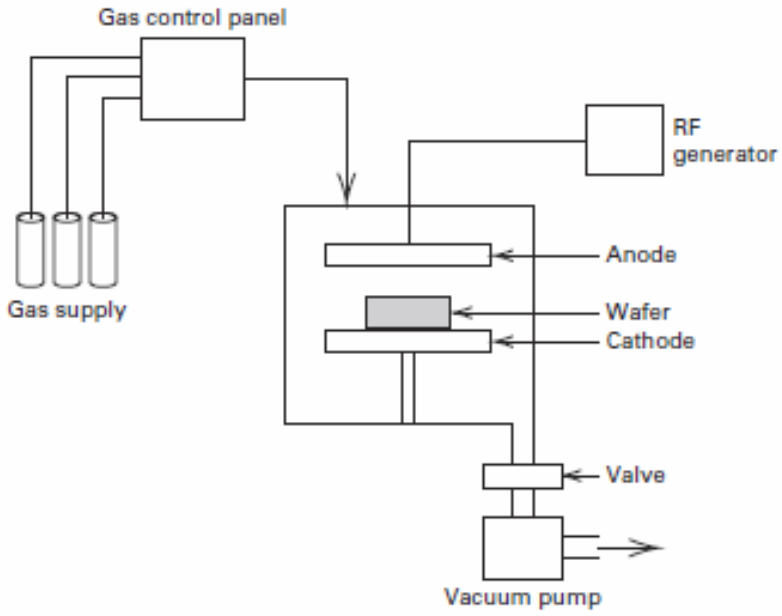
The response variable is etch rate.

She is interested in a particular gas (C2F6) and gap (0.80 cm), and wants to test four levels of RF power: 160W, 180W, 200W, and 220W.

She decided to test five wafers at each level of RF power.

The above experimenter is a single-factor with $a = 4$ levels of RF power 160W, 180W, 200W, and 220W

The experiment is replicated 5 times, i.e., $n =$ 5 replicates, and 20 runs are made in random order.

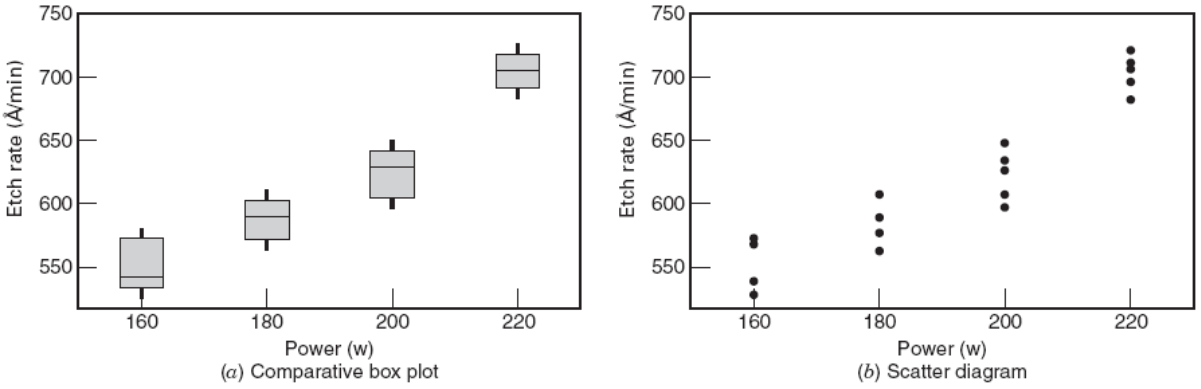


■ **FIGURE 3.1** **A single-wafer plasma etching tool**

■ **TABLE 3.1**

Etch Rate Data (in Å/min) from the Plasma Etching Experiment

Power (W)	Observations 1	2	3	4	5	Totals	Averages
160	575	542	530	539	570	2756	551.2
180	565	593	590	579	610	2937	587.4
200	600	651	610	637	629	3127	625.4
220	725	700	715	685	710	3535	707.0



■ **FIGURE 3.2** **Box plots and scatter diagram of the etch rate data**

- Some Questions

Question 1: Does changing the power change the mean etch rate?

Question 2: Is there an optimum level for power?

We would like to have an objective way to answer these questions. But, the t-test really doesn't apply here since there are more than two factor levels.

In general, there will be a levels of the factor, or a treatments, and n replicates of the experiment, run in random order, i.e., doing a completely randomized design (CRD).

$N = an$ is total runs.

Here we focus on the fixed effects case, and the random effects case will be discussed later.

Objective is to test hypotheses about the equality of the a treatment means.

Typical data for the single-factor ANOVA is as follows:

■ **TABLE 3.2**

Typical Data for a Single-Factor Experiment

Treatment (Level)	Observations				Totals	Averages
1	y_{11}	y_{12}	$\cdots$	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	$\cdots$	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots\ \vdots\ \vdots$	$\vdots$	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	$\cdots$	y_{an}	$y_{a.}$	$\bar{y}_{a.}$
					$y_{..}$	$\bar{y}_{..}$

- Models for the Data

The name analysis of variance stems from a partitioning of the total variability in the response variable into components that are consistent with a model for the experiment.

The basic single-factor ANOVA model is

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}\,, \quad \begin{cases} i = 1, 2, \cdots, a \\ j = 1, 2, \cdots, n \end{cases}$$

where μ is the overall mean; τ_i is the i^{th} treatment effect; ε_{ij} is experiment error and i.i.d. $N(0, \sigma^2)$.

The above model is called the effects model.

There is another way to write a model for the above data:

$$Y_{ij} = \mu_i + \varepsilon_{ij}\,, \quad \begin{cases} i = 1, 2, \cdots, a \\ j = 1, 2, \cdots, n \end{cases}$$

where $\mu_i = \mu + \tau_i$ is the i^{th} treatment mean. This model is called the means model.

- ANOVA

We are interested to test the following statistical hypotheses

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$$

H_a : At least one of the mean is different;

or equivalently

$$\begin{aligned} H_0 : &\quad \tau_1 = \tau_2 = \cdots = \tau_a = 0 \\ H_a : &\quad \tau_i \neq 0 \text{ for least one } i. \end{aligned}$$

Question: How to do the above tests ?

Total variability is measured by the total sum of squares:

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{n} (Y_{ij} - \bar{Y}_{..})^2$$

The basic ANOVA partitioning is:

$$\begin{aligned}
SS_T &= \sum_{i=1}^{a} \sum_{j=1}^{n} (Y_{ij} - \bar{Y}_{..})^2 \\
&= \sum_{i=1}^{a} \sum_{j=1}^{n} [(\bar{Y}_{i.} - \bar{Y}_{..}) + (Y_{ij} - \bar{Y}_{i.})]^2 \\
&= \sum_{i=1}^{a} \sum_{j=1}^{n} [(\bar{Y}_{i.} - \bar{Y}_{..})^2 + (Y_{ij} - \bar{Y}_{i.})^2] \\
&= n \sum_{i=1}^{a} (\bar{Y}_{i.} - \bar{Y}_{..})^2 + \sum_{i=1}^{a} \sum_{j=1}^{n} (Y_{ij} - \bar{Y}_{i.})^2 \\
&= SS_{Tr} + SS_E,
\end{aligned}$$

where the treatment sum of square

$$SS_{Tr} = n \sum_{i=1}^{a} (\bar{Y}_{i.} - \bar{Y}_{..})^2$$

and the error sum of squares

$$SS_E = \sum_{i=1}^{a} \sum_{j=1}^{n} (Y_{ij} - \bar{Y}_{i.})^2.$$

Note that

(i) a large value of SS_{Tr} reflects large differences in treatment means;

(ii) a small value of SS_{Tr} likely indicates no differences in treatment means.

While sums of squares cannot be directly compared to test the hypothesis of equal means, mean squares can be compared.

A **mean square** is a sum of squares divided by its degrees of freedom.

$$\begin{aligned} df_{Total} &= df_{Treatment} + df_{Error} \\ an - 1 &= (a-1) + a(n-1) \\ MS_{Tr} &= \frac{SS_{Tr}}{a-1} \text{ and } MS_E = \frac{SS_E}{a(n-1)} \end{aligned}$$

Proposition:

$$E(MS_E) = \sigma^2 \quad \text{and} \quad E(MS_{Tr}) = \sigma^2 + \frac{n\sum_{i=1}^{a}\tau_i^2}{a-1}.$$

If the treatment means are equal, the treatment and error mean squares will be (theoretically) equal. If treatment means differ, the treatment mean square will be larger than the error mean square.

How about the distributions of MS_E and MS_{Tr} ?

Cochran's theorem can tell us about the distributions of partitioned sums of squares of normally distributed random variables.

Cochran's Theorem: Let Z_i be $NID(0,1)$ for $i = 1, 2, \cdots, v$ and

$$\sum_{i}^{v} Z_i^2 = Q_1 + Q_2 + \cdots + Q_s,$$

where $s \leq v$, Q_i is a positive semi-definite quadratic form with v_i degrees of freedom ($i = 1, 2, \cdots, v$).

Then $Q_1, Q_2, \cdots, Q_s$ are independent Chi-square random variables with $v_1, v_2, \cdots, v_s$ degrees of freedom respectively, if and only if

$$v = v_1 + v_2 + \cdots + v_s.$$

Theorem 1

$$SS_E/\sigma^2 \sim \chi^2_{a(n-1)}.$$

Under the H_0, we have

$$SS_{Tr}/\sigma^2 \sim \chi^2_{a-1}$$

and SS_{Tr} and SS_E are independent random variables.

Theorem 2 The test statistic in single-factor ANOVA is

$$F = \frac{MS_{Tr}}{MS_E}.$$

We can rewrite F as follows:

$$F = \frac{\frac{SS_{Tr}}{\sigma^2}/(a-1)}{\frac{SS_E}{\sigma^2}/[a(n-1)]}$$

Note: When H_0 is true, by Theorem 1, we have that

$$F \sim F_{a-1,a(n-1)}$$

Let f denote the value of the test statistics. Then the **rejection region** at level α is

$$f \geq F_{\alpha,a-1,a(n-1)}.$$

We set up the ANOVA table

Source of variation	df	SS	Mean Square	f
Between Treatment	a-1	SS_{Tr}	MS_{Tr}	$\frac{MS_{Tr}}{MS_E}$
Error (within treatments)	N-a	SS_E	MS_E	
Total	N-1	SS_T		

Some short-cuts for computing.

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n} Y_{ij}^2 - \frac{Y_{..}^2}{N}$$

$$SS_{Tr} = \frac{1}{n}\sum_{i=1}^{a} Y_{i.}^2 - \frac{Y_{..}^2}{N}$$

and $SS_E = SS_T - SS_{Tr}$.

How about an unbalanced design ?

In some single-factor experiments, the number of observations taken within each treatment may be different. Then we say that the design is unbalanced.

The analysis of variance described above may still be used, but slight modifications must be made in the sum of squares formulas.

Let n_i observations be taken under treatment $i (i = 1, 2, \cdots, a)$ and $N = n_1 + n_2 + \cdots + n_a$.

The manual computational formulas for SS_T and SS_{Tr} become

$$\begin{aligned} SS_T &= \sum_{i=1}^{a} \sum_{j=1}^{n_i} Y_{ij}^2 - \frac{Y_{..}^2}{N} \\ SS_{Tr} &= \sum_{i=1}^{a} \frac{Y_{i.}^2}{n_i} - \frac{Y_{..}^2}{N} \end{aligned}$$

and $SS_E = SS_T - SS_{Tr}$.

No other changes are required in ANOVA for the unbalanced design.

But there are two advantages in choosing a balanced design.

(i) First, the test statistic is relatively insensitive to small departures from the assumption of equal variances for the treatments if the sample sizes are equal. This is not the case for unequal sample sizes.

(ii) Second, the power of the test is maximized if the samples are of equal sizes.

Question: How to estimate the model parameters μ, μ_i, τ_i and σ^2 ?

The overall mean, treatment means and the treatment effects are estimated by

$$\begin{aligned} \widehat{\mu} &= \bar{Y}_{..} \\ \widehat{\mu}_i &= \bar{Y}_{i.}, \quad i = 1, 2, \cdots, a \\ \widehat{\tau}_i &= \bar{Y}_{i.} - \bar{Y}_{..}, \quad i = 1, 2, \cdots, a \end{aligned}$$

$\widehat{\sigma}^2 = S_p^2 = \frac{SS_E}{N-a}$ is a pooled estimate of the common variance within each of the a treatments.

A $100(1-\alpha)\%$ confidence interval on the i^{th} treatment mean μ_i is

$$\left(\bar{Y}_{i.} - t_{\alpha/2,(N-a)}\sqrt{\frac{MS_E}{n}},\ \bar{Y}_{i.} + t_{\alpha/2,(N-a)}\sqrt{\frac{MS_E}{n}}\right)$$

We are also interested in the differences in treatments.

A $100(1-\alpha)\%$ confidence interval on the difference in any two treatments means, say $\mu_i - \mu_j$, would be

$$\left(\bar{Y}_{i.} - \bar{Y}_{j.} - t_{\alpha/2,(N-a)}\sqrt{\frac{2MS_E}{n}},\ \bar{Y}_{i.} - \bar{Y}_{j.} + t_{\alpha/2,(N-a)}\sqrt{\frac{2MS_E}{n}}\right)$$

The above confidence intervals are one-at-a-time confidence intervals.

How to obtain Simultaneous Confidence Intervals ?

By Bonferroni method, a $100(1-\alpha)\%$ simultaneous confidence interval on the i^{th} $(i = 1, 2, \cdots, r)$ treatment mean μ_i is

$$\left(\bar{Y}_{i.} - t_{\alpha/(2r),(N-a)}\sqrt{\frac{MS_E}{n}},\ \bar{Y}_{i.} + t_{\alpha/(2r),(N-a)}\sqrt{\frac{MS_E}{n}}\right)$$

- Model Adequacy Checking in the ANOVA

It is important to check model assumptions.

Do examination of residuals $\widehat{\varepsilon}_{ij} = Y_{ij} - \bar{Y}_{i.}$.

SAS can generate the residuals.

Normality

Independence

Durbin-Watson test

Constant variance

Bartlett's test

Have we fit the right model?

- Post-ANOVA Comparison of Treatment Means

Use ANOVA to test the hypothesis of equal treatment means, and assume that residual analysis is satisfactory. If that hypothesis is rejected, we don't know which specific means are different.

Determining which specific means differ following an ANOVA is called the **multiple comparisons** problem. There are lots of ways to do this.

Here we first introduce the idea of a **contrast**.

A **treatment contrast** is a linear combination of parameters of the form

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i$$

where the contrast constants $c_1, c_2, \cdots, c_a$ sum to zero, i.e.,

$$c_1 + c_2 + \cdots + c_a = 0.$$

Question: How to estimate a contrast $\Gamma = \sum_{i=1}^{a} c_i \mu_i$?

$$C = \sum_{i=1}^{a} c_i \bar{Y}_{i.}$$

$$E(C) = \Gamma = \sum_{i=1}^{a} c_i \mu_i$$

$$V(C) = \frac{\sigma^2}{n} \sum_{i=1}^{a} c_i^2$$

Question: How to estimate σ^2 ?

A $100(1-\alpha)\%$ confidence interval on the contrast $\sum_{i=1}^{a} c_i\mu_i$ is

$$\Big(\sum_{i=1}^{a} c_i\bar{y_{i.}} - t_{\alpha/2,N-a}\sqrt{\frac{MS_E}{n}\sum_{i=1}^{a} c_i^2},\ \sum_{i=1}^{a} c_i\bar{y_{i.}} + t_{\alpha/2,N-a}\sqrt{\frac{MS_E}{n}\sum_{i=1}^{a} c_i^2}\Big)$$

Question: How to test hypotheses on the contrast ?

$$H_o: \quad \sum_{i=1}^{a} c_i\mu_i = 0$$

$$H_1: \quad \sum_{i=1}^{a} c_i\mu_i \neq 0$$

Test statistic

$$T = \frac{\sum_{i=1}^{a} c_i\bar{Y}_{i.}}{\sqrt{\frac{MS_E}{n}\sum_{i=1}^{a} c_i^2}}$$

has t_{N-a} distribution under H_0.

Rejection H_0 at level α iff test statistic value satisfies $|t| \geq t_{\alpha/2,N-a}$.

Notice that

$$F = T^2 = \frac{(\sum_{i=1}^{a} c_i \bar{Y}_{i.})^2}{\frac{MS_E}{n} \sum_{i=1}^{a} c_i^2}$$

has $F_{1,N-a}$ distribution under H_0.

We also can do a F-test for the above hypothesis testing.

Rewrite F as follows

$$F = \frac{MS_C}{MS_E} = \frac{SS_C/1}{MS_E}$$

where

$$SS_C = \frac{(\sum_{i=1}^{a} c_i \bar{Y}_{i.})^2}{\frac{1}{n} \sum_{i=1}^{a} c_i^2}$$

is the single degree of freedom contrast sum of squares, and the contrast mean square $MS_C = SS_C$.

Standardized contrast: A contrast $\sum_{i=1}^{a} c_i \mu_i$ may be standardized as

$$\sum_{i=1}^{a} c_i^* \mu_i,$$

with $c_i^* = \frac{c_i}{\sqrt{\frac{1}{n}\sum_{i=1}^{a} c_i^2}}$.

For an unbalanced design, the definition of a contrast requires

$$\sum_{i=1}^{a} n_i c_i = 0$$

and correspondingly, we have $T = \frac{\sum_{i=1}^{a} c_i \bar{Y}_{i.}}{\sqrt{MS_E \sum_{i=1}^{a} \frac{c_i^2}{n_i}}}$

and the contrast sum of squares becomes

$$SS_C = \frac{(\sum_{i=1}^{a} c_i \bar{Y}_{i.})^2}{\sum_{i=1}^{a} \frac{c_i^2}{n_i}}.$$

Orthogonal contrast: Two contrasts with

coefficients $\{c_i\}$ and $\{d_i\}$ are orthogonal if

$$\sum_{i=1}^{a} c_i d_i = 0$$

or, for an unbalanced design, if

$$\sum_{i=1}^{a} n_i c_i d_i = 0.$$

For a treatments, we can have at most $a-1$ mutually orthogonal contrasts; and, any such set is called **complete**.

For a complete set of orthogonal contrasts, we have

$$SS_{Tr} = \sum_{k=1}^{a-1} SS_{C_k}$$

and tests performed on orthogonal contrasts are independent.

Now let's consider Scheff$\acute{e}$'s Method for comparing all contrasts, and, Tukey's Test for comparing all pairs of treatment means.

In many situations, experimenters may not know in advance which contrasts they wish to compare, or they may be interested in more than $a-1$ possible comparisons. In many exploratory experiments, the comparisons of interest are discovered only after preliminary examination of the data.

Scheff$\acute{e}$ (1953) has proposed a method for comparing any and all possible contrasts between treatment means.

In the Scheff$\acute{e}$'s method, the type I error is at most α for any of the possible comparisons.

In many practical situations, we will wish to compare only **pairs of means**.

Frequently, we can determine which means differ by testing the differences between all pairs of treatment means.

Thus, we are interested in contrasts of the form

$$\Gamma = \mu_i - \mu_j \quad \text{for all} \quad i \neq j$$

Although the above Scheff$\acute{e}$'s method could be easily applied to this problem, it is not the most sensitive procedure for such comparisons.

Here we consider Tukey's Test for comparing all pairs of treatment means.

- Sample Size Determination

In any experimental design problem, a critical decision is the choice of sample size, i.e., determining the number of replicates to run.

In general, if the experimenter is interested in detecting small effects, more replicates are required than if the experimenter is interested in detecting large effects.

Here, we discuss two approaches (Operating Characteristic (OC) curve method, and CI estimation method) to determining sample size.

Although our discussion focuses on the fixed effects model with a single factor design, most of the methods can be used in more complex experimental situations.

Operating characteristic (OC) curve method

An operating characteristic (OC) curve is a plot of the type II error probability of a statistical test for a particular sample size versus a parameter that reflects the extent to which the null hypothesis is false.

OC curves can be used to guide the experimenter in selecting the number of replicates

so that the design will be sensitive to important potential differences in the treatments.

For the following test

$$\begin{aligned} H_0: &\quad \tau_1 = \tau_2 = \cdots = \tau_a = 0 \\ H_a: &\quad \tau_i \neq 0 \text{ for least one } i, \end{aligned}$$

the type II error probability is

$$\begin{aligned} \beta &= 1 - P\{\text{Reject } H_0 \big| H_0 \text{ is false}\} \\ &= 1 - P\{F > F_{\alpha,a-1,N-a} \big| H_0 \text{ is false}\} \end{aligned}$$

and power = 1 -β.

In order to evaluate the probability statement in the above Equation, we need to know the distribution of the test statistic $F = \frac{MS_{Tr}}{MS_E}$ if the null hypothesis is false.

It can be shown that, if H_0 is false, the statistic $F = MS_{Tr}/MS_E$ is distributed as a **noncentral** F random variable with $a-1$ and $N-a$

degrees of freedom and the noncentrality parameter $\delta = \sum_{i=1}^{a} \frac{n\tau_i^2}{\sigma^2}$.

If $\delta = 0$, the noncentral F distribution becomes the usual (central) F distribution.

OC curves are used to evaluate the probability β. These curves plot the probability of type II error (β) against a parameter Φ, where

$$\Phi^2 = \frac{n\sum_{i=1}^{a}\tau_i^2}{a\sigma^2} = \frac{\delta}{a}$$

Curves are available for $\alpha = 0.05$ and $\alpha = 0.01$ and a range of degrees of freedom for numerator and denominator.

In using the OC curves, the experimenter must specify the parameter Φ and the value of σ^2, which is often difficult to do in practice.

One way to determine Φ is to choose the actual values of the treatment means for which we would like to reject the null hypothesis with high probability. Thus, if $\mu_1, \mu_2, \cdots, \mu_a$ are the specified treatment means, we find that $\tau_i = \mu_i - \bar{\mu}$, with $\bar{\mu}$ being the average of $\mu_1, \mu_2, \cdots, \mu_a$

The estimate of σ^2 may be available from prior experience, a previous experiment or a preliminary test, or a judgment estimate.

When we are uncertain about the value σ^2, sample sizes could be determined for a range of likely values of σ^2 to study the effect of this parameter on the required sample size before a final choice is made.

A very common way to use these OC curves is to define a difference in two means D of interest, then the minimum value of Φ^2 is

$$\Phi^2 = \frac{nD^2}{2a\sigma^2} = \frac{n}{2a}.(\frac{D}{\sigma})^2$$

Typically work in term of the ratio of D/σ and try values of n until the desired power is achieved.

Most statistics software packages can perform power and sample size calculations.

We can use SAS to do it.

CI estimation method.

3. The Random Effects Model with a Single Factor

If there are a large number of possible levels for the factor (theoretically an infinite number), the experimenter may choose a of these levels at random. Then we say that the factor is random.

Because the levels of the factor actually used in the experiment were chosen randomly, inference will be to the entire population of levels.

Here we assume that the population of factor levels is either of infinite size or is large enough to be considered infinite.

In fact, situations in which the population of factor levels is small enough to employ a finite population approach are not encountered frequently.

The linear statistical model is the following Random Effects Model.

- The Random Effects Model:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}\,, \quad \begin{cases} i = 1, 2, \cdots, a \\ j = 1, 2, \cdots, n, \end{cases}$$

where μ is the overall mean, both τ_i (the treatment effects) and ε_{ij} are random variables.

We will assume that the treatment effects τ_i are $NID(0, \sigma_\tau^2)$ random variables and that the errors ε_{ij} are $NID(0, \sigma^2)$, random variables, and that the τ_i and ε_{ij} are independent.

Question: What is the variance of Y_{ij} ?

$$V(Y_{ij}) = \sigma_\tau^2 + \sigma^2$$

The variances σ_τ^2 and σ^2 are called **variance components**, and the above **random effects model** is also called the **components of variance model**.

Question: What is the distribution of Y_{ij} ?

$$Y_{ij} \sim N(\mu, \sigma_\tau^2 + \sigma^2).$$

But, unlike the fixed effects case in which all Y_{ij} are independent, in the random model Y_{ij} are only independent if they come from different factor levels. Specifically, we have

$$\begin{aligned} Cov(Y_{ij}, Y_{ij'}) &= \sigma_\tau^2 \quad \text{if } j \neq j' \\ Cov(Y_{ij}, Y_{i'j'}) &= 0 \quad \text{if } i \neq i' \end{aligned}$$

Note that the observations within a specific factor level all have the same covariance, because before the experiment is conducted, we

expect the observations at that factor level to be similar because they all have the same random component.

Once the experiment has been conducted, we can assume that all observations can be assumed to be independent, because the parameter τ_i has been determined and the observations in that treatment differ only because of random error.

We can express the covariance structure of the Y_{ij} in the single-factor random effects model through the **covariance matrix**. To illustrate, suppose that we have $a = 3$ treatments and $n = 2$ replicates.

Observations (a = 3 and n = 2):

$$\mathbf{y} = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{bmatrix}$$

and the 6×6 covariance matrix of these observations is

$$Cov(\mathbf{y}) = \begin{bmatrix} \sigma_\tau^2 + \sigma^2 & \sigma_\tau^2 & 0 & 0 & 0 & 0 \\ \sigma_\tau^2 & \sigma_\tau^2 + \sigma^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_\tau^2 + \sigma^2 & \sigma_\tau^2 & 0 & 0 \\ 0 & 0 & \sigma_\tau^2 & \sigma_\tau^2 + \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_\tau^2 + \sigma^2 & \sigma^2 \\ 0 & 0 & 0 & 0 & \sigma_\tau^2 & \sigma_\tau^2 + \sigma^2 \end{bmatrix}$$

The main diagonals of this matrix are the variances of each individual observation and every off-diagonal element is the covariance of a pair of observations.

- ANOVA for the Random Model

The basic ANOVA sum of squares identity

$$SS_T = SS_{Tr} + SS_E$$

is still valid.

That is, we partition the total variability in the observations into a component that measures the variation between treatments (SS_{Tr}) and a component that measures the variation within treatments (SS_E).

Testing hypotheses about individual treatment effects is not very meaningful because they were selected randomly, we are more interested in the population of treatments, so we test hypotheses about the variance component σ_τ^2.

$$\begin{aligned} H_o: &\quad \sigma_\tau^2 = 0 \\ H_1: &\quad \sigma_\tau^2 > 0 \end{aligned}$$

Notice that, if $\sigma_\tau^2 = 0$, all treatments are identical; but if $\sigma_\tau^2 > 0$, variability exists between treatments.

As before, we still have the following theorem holds.

Theorem

$$SS_E/\sigma^2 \sim \chi^2_{a(n-1)}.$$

Under the H_0, we have

$$SS_{Tr}/\sigma^2 \sim \chi^2_{a-1}$$

and SS_{Tr} and SS_E are independent random variables.

Thus, under the null hypothesis $H_0 : \sigma_\tau^2 = 0$, the ratio

$$F = \frac{MS_{Tr}}{MS_E} \sim F_{(a-1),(N-a)}.$$

However, we need to examine the expected mean squares to fully describe the test procedure.

$$E(MS_{\text{Treatments}}) = \frac{1}{a-1} E(SS_{\text{Treatments}}) = \frac{1}{a-1} E\left[\sum_{i=1}^{a} \frac{y_{i.}^2}{n} - \frac{y_{..}^2}{N}\right]$$

$$= \frac{1}{a-1} E\left[\frac{1}{n}\sum_{i=1}^{a}\left(\sum_{j=1}^{n} \mu + \tau_i + \epsilon_{ij}\right)^2 - \frac{1}{N}\left(\sum_{i=1}^{a}\sum_{j=1}^{n} \mu + \tau_i + \epsilon_{ij}\right)^2\right]$$

$$= \frac{1}{a-1}\left[N\mu^2 + N\sigma_\tau^2 + a\sigma^2 - N\mu^2 - n\sigma_\tau^2 - \sigma^2\right]$$

$$E(MS_{\text{Treatments}}) = \sigma^2 + n\sigma_\tau^2$$

$$E(MS_E) = \sigma^2$$

ANOVA F-test is identical to the fixed-effects case.

The following is Example 3.10 on Page 114.

EXAMPLE 3.11

A textile company weaves a fabric on a large number of looms. It would like the looms to be homogeneous so that it obtains a fabric of uniform strength. The process engineer suspects that, in addition to the usual variation in strength within samples of fabric from the same loom, there may also be significant variations in strength between looms. To investigate this, she selects four looms at random and makes four strength determinations on the fabric manufactured on each loom. This experiment is run in random order, and the data obtained are shown in Table 3.17. The ANOVA is conducted and is shown in Table 3.18. From the ANOVA, we conclude that the looms in the plant differ significantly.

■ **TABLE 3.17**
Strength Data for Example 3.11

	Observations				
Looms	**1**	**2**	**3**	**4**	$y_{i\cdot}$
1	98	97	99	96	390
2	91	90	93	92	366
3	96	95	97	95	383
4	95	96	99	98	388
					$1527 = y_{\cdot\cdot}$

The variance components are estimated by $\hat{\sigma}^2 = 1.90$ and

$$\hat{\sigma}_\tau^2 = \frac{29.73 - 1.90}{4} = 6.96$$

Therefore, the variance of any observation on strength is estimated by

$$\hat{\sigma}_y = \hat{\sigma}^2 + \hat{\sigma}_\tau^2 = 1.90 + 6.96 = 8.86.$$

Most of this variability is attributable to differences *between* looms.

■ **TABLE 3.18**
Analysis of Variance for the Strength Data

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
Looms	89.19	3	29.73	15.68	<0.001
Error	22.75	12	1.90		
Total	111.94	15			

- Estimating the Model Parameters

For estimating the variance components (σ_τ^2 and σ^2) in the model, we introduce **the analysis of variance method**.

The procedure consists of equating the expected mean squares to their observed values in the ANOVA table and solving for the variance components. Thus the ANOVA variance component estimators are moment-method estimators.

By equating observed and expected mean squares in the single-factor random effects model, we obtain that the estimators of the variance components are

$$\begin{aligned} \widehat{\sigma^2} &= MS_E \\ \widehat{\sigma_\tau^2} &= \frac{MS_{Tr} - MS_E}{n} \end{aligned} \tag{1}$$

For unequal sample sizes, replace n in the equation (1) by

$$n_0 = \frac{1}{a-1}[\sum_{i=1}^{a} n_i - \frac{\sum_{i=1}^{a} n_i^2}{\sum_{i=1}^{a} n_i}]$$

Some Remarks:

(i) Normality isn't required in the above estimation.

(ii) They are unbiased estimators.

Confidence Intervals (CI's)

A $100(1-\alpha)\%$ CI for σ^2 is given by

$$\left(\frac{(N-a)MS_E}{\chi^2_{\alpha/2,(N-a)}}, \frac{(N-a)MS_E}{\chi^2_{1-\alpha/2,(N-a)}}\right)$$

The ratio $\sigma_\tau^2/(\sigma_\tau^2 + \sigma^2)$ is called the **intra-class correlation coefficient**, which reflects

the proportion of the variance of an observation (recall that $V(Y_{ij}) = \sigma_\tau^2 + \sigma^2$) that is the result of differences between treatments.

Confidence interval for the interclass correlation

The $100(1-\alpha)\%$ CI for $\sigma_\tau^2/(\sigma_\tau^2 + \sigma^2)$ is given by

$$\left(\frac{L}{1+L}, \frac{U}{1+U}\right),$$

where

$$L = \frac{1}{n}\left(\frac{MS_{Tr}}{MS_E}.\frac{1}{F_{\alpha/2,(a-1),(N-a)}} - 1\right)$$

and

$$U = \frac{1}{n}\left(\frac{MS_{Tr}}{MS_E}.\frac{1}{F_{1-\alpha/2,(a-1),(N-a)}} - 1\right).$$

In the end, let's consider the estimation of the overall mean.

Estimation of the Overall Mean μ. In many random effects experiments the experimenter is interested in estimating the overall mean μ. From the basic model assumptions it is easy to see that the expected value of any observation is just the overall mean. Consequently, an unbiased estimator of the overall mean is

$$\hat{\mu} = \bar{y}_{..}$$

So for Example 3.11 the estimate of the overall mean strength is

$$\hat{\mu} = \bar{y}_{..} = \frac{y_{..}}{N} = \frac{1527}{16} = 95.44$$

It is also possible to find a $100(1-\alpha)\%$ confidence interval on the overall mean. The variance of $\bar{y}$ is

$$V(\bar{y}_{..}) = V\left(\frac{\sum_{i-1}^{a}\sum_{j-1}^{n} y_{ij}}{an}\right) = \frac{n\sigma_\tau^2 + \sigma^2}{an}$$

The numerator of this ratio is estimated by the treatment mean square, so an unbiased estimator of $V(\bar{y})$ is

$$\hat{V}(\bar{y}_{..}) = \frac{MS_{\text{Treatments}}}{an}$$

Therefore, the $100(1-\alpha)\%$ CI on the overall mean is

$$\bar{y}_{..} - t_{\alpha/2,\, a(n-1)}\sqrt{\frac{MS_{\text{Treatments}}}{an}} \le \mu \le \bar{y}_{..} + t_{\alpha/2,\, a(n-1)}\sqrt{\frac{MS_{\text{Treatments}}}{an}} \qquad \textbf{(3.61)}$$

So, at 95% confidence the mean strength of the fabric produced by the looms in this facility is between 92.47 and 98.41.