

Workbook For Statistics 514

Design of Experiments

Fall 2005

by

Jonathan Kuhn, Ph.D.
Associate Professor of Statistics,
Mathematics, Statistics and Physics,
Purdue University North Central

© by Jonathan Kuhn

Preface

This is a second course in statistics. The aim of this course is acquaint a student with analysis of variance. Numerical computation, (matrix) algebra, graphs and mathematical notation are used; calculus is also used. The statistical software package called SAS is used.

This workbook is a necessary component for a student to successfully complete this course. Without the workbook, a student will not be able to participate in the course.

- This attendance workbook is *based* on the text.
- Although the material covered in each is very similar, the *presentation* of the material in the workbook is quite different from the presentation given in the text. The text consists essentially of definitions, formulas, worked out examples and exercises; this workbook, on the other hand, consists *solely* of exercises to be worked out by the student.
- The overheads presented during each lecture are based *exclusively* on the workbook. A student is to use this workbook to follow along with during a lecture.
- There are different kinds of exercises, including multiple choice, true/false, matching and fill-in-the-blank.
- Each week, I recommend you read the text, answer the questions given here in the attendance workbook and then do either the quiz or homework assignment, in that order.

On the one hand, the workbook is, as you will see, quite a bit more elaborate than typical lecture notes, which are usually a summary of what the instructor finds important in a recommended course text. On the other hand, this workbook is not quite a text, because although it has many exercises, it does not have quite enough exercises to qualify it as a complete text. I should also point out that this workbook, unfortunately, possesses a number of typographical errors. In short, this workbook aspires to be text and, in the next few years, when enough exercises have been collected, and when most of the typographical errors have been weeded out, it will become a text.

Jon Kuhn
Associate Professor of Statistics,
Mathematics, Statistics and Physics,
Purdue University North Central,
May 2004

Part V
Analysis of Variance: I

Chapter 16

Single-Factor ANOVA Model and Tests

We begin our discussion of the design of experiments by focusing on the single-factor analysis of variance (ANOVA) model. This model is used to test whether or not the means from two or more populations are the same or different.

16.1 Relation between Regression and Analysis of Variance

The single-factor analysis of variance (ANOVA) model can be thought of as a special case of the regression model.

Exercise 16.1 (Relation between Regression and Analysis of Variance)

Twelve different patients are subjected to one drug at three dosages. Their individual and average responses to the three dosages are given below.

10 mg	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1. \approx 5.90$
20 mg	5.51	5.50	5.50	5.49	5.50	$\bar{Y}_2. \approx 5.50$
30 mg	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3. \approx 5.00$

Compare how a regression analysis would deal with this data versus how an analysis of variance would deal with this data.

1. Consider the following two scatterplots of data.

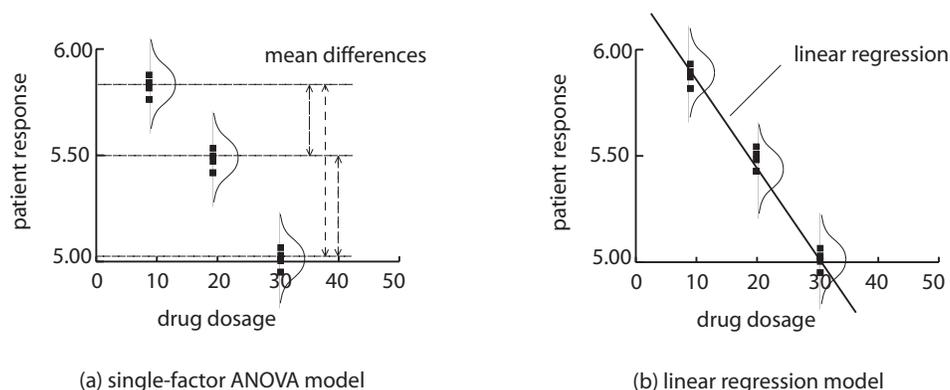


Figure 16.1 (Single-Factor ANOVA Model and Linear Regression Model)

The single-factor ANOVA procedure tests if (circle one)

- (a) at least two of the treatment averages are different,
 - (b) the slope of the “line” (regression) through the data is different from zero.
2. The single-factor ANOVA model seems to be (choose one) **more complicated** / **as complicated as** / **less complicated** than the regression model because the two models are the same, except the regression model also involves pre-supposing a (linear, in this case) relationship between the patient response and drug dosage.
 3. The single-factor ANOVA model seems to be a (choose one) **special case** / **generalization** of the regression model.
 4. **True** / **False** It is possible to recast many problems solved using ANOVA procedures as problems solved using regression procedures.
 5. **True** / **False** The explanatory or predictor variables in regression procedures are called *factors* in ANOVA procedures. The different values of a predictor variable are called *factor levels* in ANOVA procedures. Both regression and ANOVA procedures have a *response* variable.

16.2 Experimental and Observational Studies, Factors and Treatments

We take a first look at experimental and observational studies and discuss some useful terminology.

Exercise 16.2 (Terminology)

Section 2. Experimental and Observational Studies, Factors and Treatments (ATTENDANCE 1)5

1. *Experimental Study: Temperature and Mice Rate of Oxygen Consumption, A First Look.* The effect of air temperature on the rate of oxygen consumption (ROC) of mice is investigated.

temperature	room temperature (70° F)	10.3	14.0
	cold temperature (−10° F)	9.7	11.2

- (a) **True / False**
The ROC of one mouse at room temperature is 10.3 units.
- (b) Each mouse is best described as a
(circle one)
subject
experimental unit
response variable
of the experiment.
- (c) The four ROCs of the mice are best described as (circle one)
subjects
experimental units
responses
of the experiment.
- (d) The average ROC of the mice at *room* temperature is
 $\frac{10.3+14.0}{2} =$ (circle one) **10.45 / 11.45 / 11.95 / 12.15 / 12.65.**
- (e) The average ROC of the mice at *cold* temperature is
(circle one) **10.45 / 11.45 / 11.95 / 12.15 / 12.65.**
- (f) The *difference* in the average ROC of the mice at room temperature and the average ROC of the mice at cold temperature is
 $12.15 - 10.45 =$ (circle one) **0.7 / 1.2 / 1.7 / 2.1 / 2.5.**
- (g) The difference in the average ROC of the mice at room temperature and the average ROC of the mice at the cold temperature *could* be caused by
(circle none, one or more)
i. the different temperatures (we hope).
ii. the different mice (unfortunately, if different mice are used).
iii. the different lighting or noise conditions (unfortunately, if different lighting or noise conditions occur).
iv. other unknown extraneous variables.
- (h) We are *most* interested in determining the influence of (circle two)
temperature
amount of light
amount of noise
on the ROC of the mice in this experiment.

- (i) This is an *experiment* (as opposed to *observational study*) because (choose one)
- i. the mice decided at which temperature to be tested.
 - ii. the experimenters decided how to assign the temperature factor levels to each mouse.
- (j) The temperature is an example of (choose one)
an experimental factor
a classification factor
 because temperature is controlled by the experimenter and can be assigned to the mice (experimental unit) at random.
- (k) *Classification (observational) factors* are attributes of the experimental unit that are *not* manipulated by the experimenters such as the (choose none, one or more)
weight
health
gender
 of the mice. The experimenters either might not be able to manipulate these attributes or they may be able to manipulate them but have chosen not to because, say, it would be extraordinarily difficult to do so. For example, it might be possible to genetically manipulate the gender of mice, but, for purposes of the simple temperature experiment given here, it is simply not worth the effort.
- (l) The two *levels* of the experimental temperature factor are (circle one)
- i. room temperature and cold temperature
 - ii. bright lighting and dark lighting
 - iii. 10.3 and 9.7
- (m) The *status-quo* (or "do-nothing") level of a factor is often referred to as the *control level* or, more simply, as the *control*. In this case, the control is most likely the (circle one) **cold temperature / room temperature**.
- (n) **True / False**
 Although it makes sense to designate room temperature as the control, it is possible to designate cold temperature as the control. The control, then, can actually be any one of the factor levels.
- (o) **True / False**
 The control *group*, in this case, are the two mice subjected to the control, the room temperature, say.
- (p) It is necessary to have a control group because (circle none, one or more)
- i. although we could measure the average ROC of the cold temperature group of mice, we would not know if this average ROC was high or

low, unless we also had the average ROC of the room temperature group of mice (the control group, say) to compare against it.

- ii. we are interested in determining the influence of temperature on the ROC of mice and without subjecting mice to at least two different temperatures, we would not be able to determine this influence.
- iii. we need a group to *compare* against.

(q) **True / False**

The control (factor level) is an *effective* baseline for comparison only if all of the conditions in this experiment (age/weight/health of mice, lighting conditions, duration of measurements and so on) are as identical to one another as it possible, *except*, of course, for the effect under consideration; namely, temperature.

- (r) Since the temperature levels are described by the numerical values 70°F and -10°F, the temperature factor is
(choose one) **quantitative / qualitative**

- (s) How many mice *could* be used in this experiment? *Circle none, one or more.*

- i. Four: one mouse is subjected to room temperature, a second mouse to subjected to the same room temperature, a third mouse is subjected to the cold temperature and a fourth mouse is subjected to the same cold temperature.
- ii. Two: one mouse is subjected to both room temperature and the cold temperature and a second mouse is also subjected to both room temperature and the cold temperature.
- iii. One: one mouse is *twice* subjected to both room temperature and the cold temperature.

(t) **True / False**

The *design* of the experiment, in other words, how we assign the different temperature levels to the different mice, is a *good* one, if, at the end of the experiment, we are very certain that it is the *temperature*, and *not* any other variable, that is influencing the ROC of mice.

- (u) The *one* mouse design, described above, is probably (circle one) **better / worse** than the *four* mouse design in the sense we are certain the different ROCs are *not* being caused by different mice (there is only one mouse) and so, consequently, more certain the different ROCs are being caused by the different temperatures the mice are being subjected to.

(v) **True / False**

The design of the experiment is a *good* one if the difference in average ROCs under the two different temperatures measured from the four ROCs of the

particular mice used in the experiment is *representative* of the difference in the ROCs of *typical* mice. Representative results are more valuable to us because these results apply to not only to the few mice in the experiment, but also apply to all mice.

(w) The *two* mouse design, described above, is probably (circle one) **better** / **worse** than the *one* mouse design in the sense we are more certain that the *average* ROC of two *different* mice is more representative of the ROC of “typical” mice than the average ROC of one mouse measured twice.

(x) **True / False**

However many mice used in this experiment, it makes sense to assign the different temperatures to the mice at random because this leads to impartial allocation free from personal biases and supports the necessary independence assumptions required when applying statistical analyses to this data.

(y) *Statistic, Sample, Parameter and Population.*

It is (very) important to be able to distinguish between statistics and parameters. Match the statistical terms with the various parts of this example.

statistical terms	mice example
(i) statistics	(i) ROCs of mice in study
(ii) sample	(ii) average ROC for mice in study
(iii) parameters	(iii) ROCs of all mice
(iv) population	(iv) average ROC for all mice

statistical terms	(i)	(ii)	(iii)	(iv)
mice example				

The sample (observed) mean ROC of the mice at room temperature in the study, 12.15, could be used as an estimate for the unknown population (actual) mean ROC of all mice at room temperature.

2. *Experimental Study: Lighting and Mice ROC.*

The effect of *lighting* (not temperature!) on the ROC of deer mice is investigated. Four different mice are subjected to two lighting levels.

lighting	bright	10.3	14.0
	dim	9.7	11.2

(a) Match the statistical terms with the various parts of this example.

Section 2. Experimental and Observational Studies, Factors and Treatments (ATTENDANCE 1)9

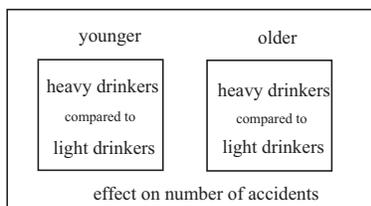
statistical terms	mice example
(i) factor	(i) lighting
(ii) control	(ii) ROC
(iii) response variable	(iii) dim light
(iv) experimental unit	(iv) deer mouse

statistical terms	(i)	(ii)	(iii)	(iv)
lighting and ROC				

- (b) The factor “lighting” has
(circle one) **one** / **two** / **three** / **four** levels.
- (c) The levels of “lighting” are (circle none, one or more)
- i. mice
 - ii. bright lighting
 - iii. dim lighting
 - iv. ROC of mice
- (d) This is an experiment because (choose one)
- i. the mice decided at which lighting level to be tested.
 - ii. the experimenters decided how to assign the lighting factor levels to each mouse.
- (e) Since the lighting levels are described by “dim” and “bright”, the lighting factor is
(choose one) **quantitative** / **qualitative**
- (f) If, in addition to the factor “lighting”, a second factor was also considered in this study, “temperature”, the model would be called a (choose two)
- i. single-factor ANOVA model
 - ii. two-factor ANOVA model
 - iii. multi-factor ANOVA model

3. *Observational Study: Alcohol, Traffic Accidents and Age Confounder.*

Indiana police records from 1999–2001 are analyzed to determine if there is an association between amount of drinking and traffic accidents. The age of a driver is also to be taken into account in this study.



alcohol vs accidents,
controlling for age confounder

Figure 16.2 (Observed Studies: Controlling For Confounding Variables)

- (a) This is an observed study because (circle one)
- the police decided who was going to drink and drive and who was not.
 - the drivers decided who was going to drink and drive and who was not.
 - drivers were assigned to drink and drive or be sober drivers at random.
- (b) The response variable is (choose one)
- age
 - amount of drinking
 - number of drivers
 - number of accidents
- (c) This is a multifactor study where the treatment has the two factors (choose two)
- age
 - amount of drinking
 - number of drivers
 - number of accidents
- where both factors each have two levels.
- (d) Both factors are (choose one) **experimental** / **classifications (observational)**
because the experimenters do not manipulate either of these factors.
- (e) Both factors are (choose one) **qualitative** / **quantitative**
- (f) Suppose age influences the number of traffic accidents. Age would be considered a confounding variable with amount of drinking on the number of traffic accidents of Indiana drivers if (circle one)
- young drivers had a greater number of traffic accidents than older drivers.
 - intoxicated drivers had a greater number of traffic accidents than sober drivers.
 - it was not clear at the end of the study whether the number of traffic accidents was a consequence of being intoxicated or not, or whether it was a consequence of age.
 - it was not clear at the end of the study whether the number of traffic accidents was a consequence of being intoxicated or not, or whether it was a consequence of *gender*.

- (g) One way to eliminate the confounding effect of age with amount of drinking on the number of traffic accidents (to control for age) in this observed study would be to (circle one)
- i. assign drivers to be either drunk or sober at random (Is this possible, since the data was collected from police records?)
 - ii. compare the number of traffic accidents of drunk drivers with sober driver who both have similar ages
 - iii. compare the number of traffic accidents of drunk drivers with sober driver who both have different ages
- (h) Although it is possible to assign treatments to observation units in experimental studies *at random*, it is not possible to do this in observational studies. Consequently, the effect of treatments on responses in observed studies are (choose one) **less accurately** / **more accurately** known than the effect of treatments on responses in experimental studies.
4. *Observational Study: Internet Learning and Average GPA Confounder.* A recent study was conducted to compare the academic achievement (measured by final examination scores) of Internet students with classroom students. Internet students did not attend class, but received all instruction over the Internet, whereas classroom students received instruction at a fixed time every week in a specified classroom.

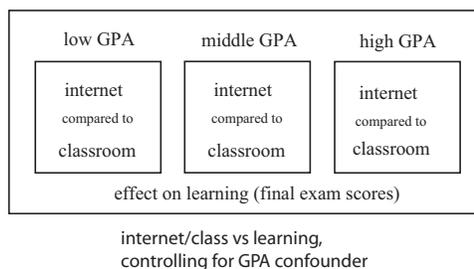


Figure 16.3 (Observed Studies: Controlling For Confounding Variables)

- (a) This is an observed study because (circle one)
- i. the instructor decided who was a classroom student and who was an Internet student.
 - ii. the students decided to be either a classroom student or an Internet student.
 - iii. students were assigned to be “classroom” students and “Internet” students at random.
- (b) This is a multifactor study where the treatment has the two factors (choose two)

- i. age
- ii. type of instruction (internet, class)
- iii. average GPA
- iv. academic achievement (final exam scores)

where the classification factor “GPA” has three levels and the classification factor “type of instruction” has two levels.

- (c) The response variable is (choose one)
- i. age
 - ii. type of instruction (internet, class)
 - iii. average GPA
 - iv. academic achievement (final exam scores)
- (d) The response variable is (choose one) **qualitative** / **quantitative**
- (e) Suppose average GPA influences academic achievement (measured by final examination scores). Average GPA would be considered a confounding variable with type of instruction on the academic achievement of students if (circle one)
- i. students with high average GPAs had better final examination scores than students with low average GPAs.
 - ii. temporary Internet students had better final examination scores than classroom students.
 - iii. it was not clear at the end of the study whether students’ academic achievement was a consequence of being either an Internet students or classroom students, or was as a consequence of average GPA.
 - iv. it was not clear at the end of the study whether students’ academic achievement was a consequence of being either a male or female, or was as a consequence of average GPA.
- (f) One way to eliminate the confounding effect of average GPA with type of instruction on the academic achievement of students (to control for average GPA) in this observed study would be to (circle one)
- i. assign students to be either “classroom” or “Internet” students at random (Is this possible, since these students choose between these two options themselves?)
 - ii. compare the academic achievement of classroom students with Internet students who both have similar average GPAs
 - iii. compare the academic achievement of classroom students with Internet students who both have different average GPAs

(g) **True / False**

Controlling for the confounder average GPA in this study does not control for any other confounder. Each confounder must be controlled for separately from every other confounder in an observed study, unlike in a randomized design, where randomization takes care of all confounders all at once.

(h) *Statistic, Sample, Parameter and Population.* Match the statistical terms with the various parts of this example.

statistical terms	example
(i) statistics	(i) final exam scores for students in study
(ii) sample	(ii) average final exam scores for students in study
(iii) parameters	(iii) final exam scores for all students
(iv) population	(iv) average final exam scores for all students

statistical terms	(i)	(ii)	(iii)	(iv)
example				

The sample (observed) mean final exam score of the students in the study could be used as an estimate for the unknown population (actual) mean score of all students.

16.3 Design of Analysis of Variance Studies

This material is covered in the previous section above, section 3.

16.4 Uses of Analysis of Variance Models

This material is covered in section 3.

16.5 Use of Computers

We will be using the statistical software package, SAS, extensively throughout this course.

16.6 ANOVA Model I—Fixed Factor Levels

We look at the notation associated with one kind of (single-factor) ANOVA model which has *fixed* factors. We will refer to this model as either ANOVA Model I, or,

simply, as the ANOVA model; we will look at both ANOVA Models II and III later in the semester¹.

Exercise 16.3 (ANOVA Model I—Fixed Factor Levels)

1. **Notation: Drug Responses.** Twelve different patients are subjected to three drugs. A higher number indicates the patient is better off than a patient with a lower number.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1. \approx 5.90$
drug 2	5.50	5.50				$\bar{Y}_2. = 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3. \approx 5.00$

- (a) $r =$ (circle one) **1 / 2 / 3** factor levels (treatments)
- (b) $n_1 =$ (circle one) **2 / 5 / 12** drug 1 cases
- (c) $n_2 =$ (circle one) **2 / 5 / 12** drug 2 cases
- (d) $n_T = \sum_{i=1}^r n_i =$ (circle one) **2 / 5 / 12** total cases
- (e) drug 1, case 1, $Y_{11} =$ (circle one) **5.50 / 5.90 / 5.92**
- (f) drug 3, case 4, $Y_{34} =$ (circle one) **4.98 / 5.90 / 5.92**
- (g) sum of drug 1 cases, $Y_{1.} = \sum_{j=1}^{n_1} Y_{1j} =$ (circle one) **5.50 / 5.90 / 29.5**
- (h) sum of drug 2 cases, $Y_{2.} = \sum_{j=1}^{n_2} Y_{2j} =$ (circle one) **4.98 / 5.90 / 11.00**
- (i) average of drug 1 cases,
 $\bar{Y}_{1.} = \frac{1}{n_1} \sum_{j=1}^{n_1} Y_{1j} =$ (circle one) **5.50 / 5.90 / 29.5**
- (j) average of drug 3 cases,
 $\bar{Y}_{3.} = \frac{1}{n_3} \sum_{j=1}^{n_3} Y_{3j} =$ (circle one) **4.98 / 5.00 / 11.00**
- (k) variance of drug 1 responses, $s_1^2 =$ (circle one) **0.00025 / 0.0158 / 29.5**
 (to get s_1^2 , use calculator: type responses into L_1 , then STAT CALC ENTER, then square S_x)
- (l) variance of drug 3 responses, $s_3^2 =$ (circle one) **0.00025 / 0.436 / 11.00**
- (m) To say the average responses are *not* the same (circle one) **is / is not**
 the same thing as saying the average responses are *all* different from one another.
- (n) The total of all observations is

$$Y_{..} = \sum_{i=1}^r \sum_{j=1}^{n_i} Y_{ij} = 65.5$$

¹These two other kinds of ANOVA model have *random* (rather than fixed) effects.

and so the overall mean,

$$\bar{Y}_{..} = \frac{1}{n_T} \sum_{i=1}^r \sum_{j=1}^{n_i} Y_{ij} \approx 5.46$$

is given by (circle none, one or more)

- i. adding all twelve cases and dividing by 12.
 - ii. $\frac{\bar{Y}_1 + \bar{Y}_2 + \bar{Y}_3}{12}$.
 - iii. $\frac{5\bar{Y}_1 + 2\bar{Y}_2 + 5\bar{Y}_3}{12}$.
- (o) If all the average responses to the three drugs are the same, then they (circle one) **are** / **are not** equal to the overall average.
- (p) The overall variance, $s^2 \approx 0.0857$, is given by (circle none, one or more)
- i. determining the variance of the twelve responses.
 - ii. $\frac{\bar{Y}_1 + \bar{Y}_2 + \bar{Y}_3}{12}$.
 - iii. $\frac{(5-1)s_1^2 + (2-1)s_2^2 + (5-1)s_3^2}{12-3}$.
- (q) *Statistic, Sample, Parameter and Population.* Match the statistical terms with the various parts of this example.

statistical terms	example
(i) statistics	(i) drug responses for patients in study, Y_{ij}
(ii) sample	(ii) mean drug i , $i = 1, 2, 3$, responses for patients in study, \bar{Y}_i .
(iii) parameters	(iii) drug responses for all patients, Y_{ij}
(iv) population	(iv) mean drug i , $i = 1, 2, 3$, responses for all patients, μ_i

statistical terms	(i)	(ii)	(iii)	(iv)
example				

The observed (sample) mean drug 1 response for the patients in the study, $\bar{Y}_1 \approx 5.90$, is often used as an estimate², $\bar{Y}_1 = \hat{Y}_{1j}$, for the unknown actual (population) mean drug 1 response for all patients, μ_1 .

- (r) *ANOVA Model I Versus ANOVA Model II.*

It makes sense in this case that we are interested in the patient responses to the three drugs in the study *and only these three drugs* in the study. This is called an ANOVA model I (which we will simply refer to as an ANOVA model, unless otherwise specified). Later in the semester, we will look at ANOVA Model II, where we assume that these three drugs are actually a random selection (a sample) of a larger set of drugs.

²The estimate, \bar{Y}_i , is a “good” estimate because it is both a *least squares estimate* and a *maximum likelihood estimate* of the parameter μ_i .

2. *Notation: Rats in Big City.* Where are the rats in Big City? The rat count per square meter is recorded and shown in the following table.

sewers	parks	city hall
3	1	5
5	3	8
7	2	9
4		8
		10

- (a) $r =$ (circle one) **1 / 2 / 3** factor levels (treatments)
 (b) $n_T =$ (circle one) **2 / 4 / 12** responses
 (c) $n_1 =$ (circle one) **3 / 4 / 12** responses
 (d) $n_2 =$ (circle one) **2 / 3 / 12** responses
 (e) sewer, replication 1 response, $Y_{11} =$ (circle one) **3 / 5.90 / 5.92**
 (f) city hall, replication 4 response, $Y_{34} =$ (circle one) **8 / 5.90 / 5.92**
 (g) sum of sewer responses, $Y_1 =$ (circle one) **5.50 / 19 / 29.5**
 (h) sum of parks responses, $Y_2 =$ (circle one) **4.98 / 6 / 11.00**
 (i) average of sewer responses, $\bar{Y}_1 =$ (circle one) **4.75 / 5.90 / 29.5**
 (j) average of city hall responses, $\bar{Y}_3 =$ (circle one) **4.98 / 5.00 / 8**
 (k) variance of sewer responses, $s_1^2 =$ (circle one) **0.00025 / 0.0158 / 2.92**
 (l) variance of city hall responses, $s_3^2 =$ (circle one) **0.19025 / 0.436 / 3.5**
 (m) *Statistic, Sample, Parameter and Population.*

Match the statistical terms with the various parts of this example.

statistical terms	example
(i) statistics	(i) rat count per m^2 for three areas in Big City, Y_{ij}
(ii) sample	(ii) mean rat count per m^2 for three areas in Big City, Y_i .
(iii) parameters	(iii) rat count per m^2 for all areas in Big City, Y_{ij}
(iv) population	(iv) mean rat count per m^2 for all areas in Big City, μ_i

statistical terms	(i)	(ii)	(iii)	(iv)
example				

The observed (sample) mean rat count per m^2 for the city hall area in Big City, $\bar{Y}_3 = 8$, is often used as an estimate, $\bar{Y}_3 = \hat{Y}_{3j}$, for the unknown actual (population) mean rat count per m^2 for the city hall area in Big City, μ_3 .

- (n) *ANOVA Model I Versus ANOVA Model II.* If only these three areas of Big City are of interest to us in this study, then we would be assuming (choose one) **ANOVA Model I / ANOVA Model II.**

16.7 Fitting of ANOVA Model

This material is partly covered in the previous section, section 6, and partly covered later on, in section 10.

16.8 Analysis of Variance

SAS program: att1-16-8-drugs,rats-singleANOVA

The single-factor analysis of variance (ANOVA) procedure is demonstrated in this section. This ANOVA test uses the data to decide if two or more of the population factor level (treatment) means are the same or different. An important part of this procedure involves the calculation of the ANOVA table,

Source	Degrees of Freedom, df	Sum Of Squares, SS	Mean Squares, MS
Treatment (between treatments)	$r - 1$	$SSTR = \sum n_i(\bar{Y}_i - \bar{Y}_{..})^2$	$MSTR = \frac{SSTR}{r-1}$
Error (within treatments)	$n_T - r$	$SSE = \sum \sum (Y_{ij} - \bar{Y}_i)^2$	$MSE = \frac{SSE}{n_T-r}$
Total	$N_T - 1$	$SSTO = \sum \sum (Y_{ij} - \bar{Y}_{..})^2$	

which leads to the calculation of the observed F ,

$$F = \frac{MSTR}{MSE}$$

A large F indicates that the population factor level means are different; a small F indicates the means are the same. We will use both the TI-83 calculator and the statistical software package, SAS, to do this.

Exercise 16.4 (Single-Factor Analysis of Variance)

1. Comparing Three Drugs

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1. \approx 5.90$
drug 2	5.50	5.50	5.50	5.49	5.50	$\bar{Y}_2. = 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3. \approx 5.00$

Test if at least two of the three average patient responses to the drug are different at $\alpha = 0.05$.

- (a) *Test Statistic Versus Critical Value.*

i. *Statement*

The statement of the test is (check none, one or more):

A. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus $H_a : \mu_1 \neq \mu_2, \mu_1 = \mu_3$.

B. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus $H_a : \mu_1 \neq \mu_3, \mu_1 \neq \mu_2$.

C. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus

$H_a : \text{at least one } \mu_i \neq \mu_j, i \neq j; i, j = 1, 2, 3.$

D. $H_0 : \text{means the same}$ versus $H_a : \text{means different}$

ii. *Test.*

From SAS³, the ANOVA table is given by,

Source	Degrees of Freedom	Sum Of Squares	Mean Squares
Treatment (Drugs)	2	2.033	1.0165
Error	12	0.0022	0.00018
Total	14	2.035	

and so the test statistic is

$$F \text{ test statistic} = \frac{1.0165}{0.00018} =$$

(circle one) **1.02 / 123 / 5545.45**.

The upper critical value at $\alpha = 0.05$, with $r - 1 = 3 - 1 = 2$ and $n_T - r = 15 - 3 = 12$ degrees of freedom, is

(circle one) **3.22 / 3.89 / 4.82**

(Use PRGM INVF ENTER 2 ENTER 12 ENTER 0.95 ENTER)

iii. *Conclusion.*

Since the test statistic, 5545.45, is larger than the critical value, 3.89, we (circle one) **accept / reject** the null hypothesis that the average patient responses to the three drugs are the same.

(b) *P-Value Versus Level of Significance.*i. *Statement*

The statement of the test is (check none, one or more):

A. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus $H_a : \mu_1 \neq \mu_2, \mu_1 = \mu_3$.

B. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus $H_a : \mu_1 \neq \mu_3, \mu_1 \neq \mu_2$.

C. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus

$H_a : \text{at least one } \mu_i \neq \mu_j, i \neq j; i, j = 1, 2, 3.$

D. $H_0 : \text{means the same}$ versus $H_a : \text{means different}$

³To use your TI-83, enter the data L_1, L_2, L_3 into STAT EDIT ENTER, then type perform the ANOVA by STAT TESTS ANOVA(ENTER L_1, L_2, L_3) ENTER.)

ii. *Test.*

Since the test statistic is $F = 5545.45$, the p-value, with $r - 1 = 3 - 1 = 2$ and $n_T - r = 15 - 3 = 12$ degrees of freedom, is given by

$$\text{p-value} = P(F \geq 5545.45)$$

which equals (circle one) **0.00** / **0.35** / **0.43**.

(Use 2nd DISTR 9:Fcdf(5545.45,E99,2,12).)

The level of significance is 0.05.

iii. *Conclusion.*

Since the p-value, 0.00, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that the average patient responses to the three drugs are the same.

(c) *Related Questions*i. *Factor level means and variability plot*

The SAS plot shows that the factor mean levels are (choose one)

different / **the same**,

and the variability appears to be (choose one)

different / **the same**

for the three different levels.

ii. *Fitted (predicted) values*

Match the statistical terms with the various parts of this example.

factor level means	values
(i) $\hat{Y}_{1j} = \bar{Y}_1. \approx$	(i) 5.498
(ii) $\hat{Y}_{2j} = \bar{Y}_2. \approx$	(ii) 5
(iii) $\hat{Y}_{3j} = \bar{Y}_3. \approx$	(iii) 5.9

factor level means	(i)	(ii)	(iii)
values			

iii. *Residuals*

A couple of the residuals (given on the output) are $e_{11} = 0$ and

$e_{23} =$ (choose one) **-0.002** / **0** / **0.002**

The residuals (choose one) **do** / **do not** sum to zero.

2. *Rats in New York.*

Where are the rats in New York city? The rat count per square meter is recorded and shown in the following table.

sewers	parks	city hall
3	1	5
5	3	8
7	2	9
4		8
		10

Test the claim that the average rat count per square meter in at least two of the three city areas is different at 5%.

(a) *Test Statistic Versus Critical Value.*

i. *Statement*

The statement of the test is (check none, one or more):

A. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus $H_a : \mu_1 \neq \mu_2, \mu_1 = \mu_3$.

B. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus $H_a : \mu_1 \neq \mu_3, \mu_1 \neq \mu_2$.

C. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus

$H_a : \text{at least one } \mu_i \neq \mu_j, i \neq j; i, j = 1, 2, 3.$

D. $H_0 : \text{means the same}$ versus $H_a : \text{means different}$

ii. *Test.*

From SAS, the ANOVA table is given by,

Source	Sum Of Squares	Degrees of Freedom	Mean Squares
Between	70.2	2	35.1
Within	24.8	9	2.75
Total	95.0	11	

and so the test statistic is

$$F \text{ test statistic} = \frac{35.1}{2.75} =$$

(circle one) **12.76 / 123 / 4612.3.**

The upper critical value at $\alpha = 0.05$, with $r - 1 = 3 - 1 = 2$ and $n_T - r = 12 - 3 = 9$ degrees of freedom, is

(circle one) **3.22 / 3.89 / 4.26**

(Use PRGM INVF ENTER 2 ENTER 9 ENTER 0.95 ENTER)

iii. *Conclusion.*

Since the test statistic, 12.76, is larger than the critical value, 4.26, we (circle one) **accept / reject** the null hypothesis that the average number of rats in the three locations are the same.

(b) *P-Value Versus Level of Significance.*

i. *Statement*

The statement of the test is (check none, one or more):

A. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus $H_a : \mu_1 \neq \mu_2, \mu_1 = \mu_3$.

B. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus $H_a : \mu_1 \neq \mu_3, \mu_1 \neq \mu_2$.

C. $H_0 : \mu_1 = \mu_2 = \mu_3$ versus

$H_a : \text{at least one } \mu_i \neq \mu_j, i \neq j; i, j = 1, 2, 3.$

D. $H_0 : \text{means the same}$ versus $H_a : \text{means different}$

ii. *Test.*

Since the test statistic is $F = 12.76$, the p-value, with $r - 1 = 3 - 1 = 2$ and $n_T - r = 12 - 3 = 9$ degrees of freedom, is given by

$$\text{p-value} = P(F \geq 12.76)$$

which equals (circle one) **0.002** / **0.35** / **0.43**.

(Use 2nd DISTR 9:Fcdf(12.76,E99,2,9).)

The level of significance is 0.05.

iii. *Conclusion.*

Since the p-value, 0.002, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that the average number of rats in the three locations are the same.

(c) *Related Questions*i. *Factor level means and variability plot*

The SAS plot shows that the factor mean levels are (choose one)

different / **the same**,

the variability appears to be (choose one)

different / **the same**

and there

is / **is not**

a linear relationship between the three different levels.

ii. *Fitted (predicted) values*

Match the statistical terms with the various parts of this example.

factor level means	values
(i) $\hat{Y}_{1j} = \bar{Y}_1. \approx$	(i) 4.75
(ii) $\hat{Y}_{2j} = \bar{Y}_2. \approx$	(ii) 2
(iii) $\hat{Y}_{3j} = \bar{Y}_3. \approx$	(iii) 8

factor level means	(i)	(ii)	(iii)
values			

iii. *Residuals*

A couple of the residuals (given on the output) are $e_{34} = 0$ and $e_{35} =$ (choose one) **-2** / **0** / **2**

16.9 F Test for Equality of Factor Level Means

This material is covered in the previous section, section 8.

16.10 Alternative Formulation of Model I

In this section, we discuss the mathematical notation, assumptions and formulation for a single-factor ANOVA model I. The ANOVA assumptions for a single-factor ANOVA model I are,

- populations have a normal distribution
- population variances are equal (are all equal to a constant)
- observed responses are independent random samples from r factor levels

More precisely, the *cell means* version of ANOVA model I is

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

where

Y_{ij} is the value of the response variable in the j th trial of the i th factor level (treatment)

μ_i are factor level (treatment) parameters

ε_{ij} are independent $N(0, \sigma^2)$

$i = 1, \dots, r; j = 1, \dots, n_i$

or the alternative but equivalent *factor means* version of ANOVA model I is

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

where

Y_{ij} is the value of the response variable in the j th trial of the i th factor level (treatment)

$\mu.$ is a (weighted or unweighted) of all factor level mean parameter

$\tau_i = \mu_i - \mu.$ are factor level *effects*

ε_{ij} are independent $N(0, \sigma^2)$

$i = 1, \dots, r; j = 1, \dots, n_i$

Exercise 16.5 (Notation for Single-Factor ANOVA Model I)

1. *Comparing Three Drugs, Common Number of Replications For Each Treatment.* Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{y}_{1+} \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{y}_{2+} \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{y}_{3+} \approx 5.00$

- (a) The observed patient responses for drug 1, Y_{1j} , $j = 1, 2, \dots, 5$, are given by
- $\{5.90, 5.92, 5.91, 5.89, 5.88\}$
 - $\{5.51, 5.50, 5.50, 5.49, 5.50\}$
 - $\{5.01, 5.00, 4.99, 4.98, 5.02\}$
- (b) The observed patient responses for drug 3, Y_{3j} , $j = 1, 2, \dots, 5$, are given by
- $\{5.90, 5.92, 5.91, 5.89, 5.88\}$
 - $\{5.51, 5.50, 5.50, 5.49, 5.50\}$
 - $\{5.01, 5.00, 4.99, 4.98, 5.02\}$
- (c) The *observed* mean patient response for drug 2 is
 $\bar{Y}_2 =$ (circle one) **5.00 / 5.50 / 5.90**.
 This observed mean patient response for drug 2 could be used to estimate the *expected* mean patient response for drug 2, μ_2 .
- (d) The difference between the 4th observation of the patient response for drug 2, $Y_{24} = 5.49$, and the observed mean patient response for drug 2, $\bar{Y}_2 = 5.50$, is the observed *error* (or *residual*),
 $e_{24} = Y_{24} - \bar{Y}_2 = 5.49 - 5.50 =$ (circle one) **-0.01 / 0.00 / 0.08**.
 The observed residual, e_{ij} , is an estimate of the actual error, ε_{ij} .
- (e) The difference between the *observed* patient response for drug 2 and the *observed* overall average is
 $\bar{Y}_1 - \bar{Y}_{..} = 5.50 - 5.50 =$ (circle one) **-0.03 / 0.00 / 0.08**.
- (f) The *expected* difference between the patient response for drug 2 and the overall mean is given by
 $\mu_2 - \mu =$ (circle one) **$\tau_1 / \tau_2 / \tau_3 / \tau_i$** .
 (Hint: Look at the two mathematics models given above.)
- (g) The *expected* difference between the patient response for drug i and the overall mean is given by
 $\mu_i - \mu =$ (circle one) **$\tau_1 / \tau_2 / \tau_3 / \tau_i$** .
- (h) Taking into account the influence of the different drugs, and making appropriate assumptions about the errors, the statistical model associated with this data is (circle two)
- $Y_{ij} = \mu_i + \varepsilon_{ij}$, $i = 1, 2, 3$; $j = 1, 2, \dots, 5$
 - $Y_{ij} = \mu + \tau_j + \varepsilon_{ij}$, $i = 1, 2, 3$; $j = 1, 2, \dots, 5$,

$$\text{iii. } Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, 3; \quad j = 1, 2, \dots, 5,$$

2. *Comparing Three Drugs, Different Number of Replications For Each Treatment.*
Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1. \approx 5.90$
drug 2	5.50	5.50				$\bar{Y}_2. = 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3. \approx 5.00$

- (a) The observed patient responses for drug 1, Y_{1j} , $j = 1, 2, \dots, 5$, are given by
- {5.90, 5.92, 5.91, 5.89, 5.88 }
 - {5.50, 5.50}
 - {5.01, 5.00, 4.99, 4.98, 5.02}
- (b) The observed patient responses for drug 2, Y_{2j} , $j = 1, 2$, are given by
- {5.90, 5.92, 5.91, 5.89, 5.88 }
 - {5.50, 5.50}
 - {5.01, 5.00, 4.99, 4.98, 5.02}
- (c) The observed mean patient response for drug 2 is
 $\bar{Y}_2. =$ (circle one) **5.00** / **5.50** / **5.90**.
- (d) The best (only?) way to calculate the *observed* overall mean, $\bar{Y}_{..}$, using the observed factor level means, $\bar{Y}_{i.}$, is to use the following calculations (circle one)
- unweighted mean: $\frac{5.90+5.50+5.00}{3} \approx 5.47$.
 - $\frac{5.90+5.50+5.00}{12} \approx 4.1$.
 - weighted mean: $\frac{5(5.90)+3(5.50)+5(5.00)}{12} \approx 5.92$.

Hint: The calculation chosen must equal all twelve observations divided by twelve,

$$\bar{Y}_{..} = \frac{1}{n_T} \sum \sum Y_{ij} = \frac{5.90 + \dots + 5.02}{12}$$

- (e) Although there is only one way to calculate the *observed* overall mean, there are a number of different sensible ways to define the *expected* overall mean, $\mu_{..}$, using the observed factor level means, $\mu_{i.}$, such as (circle none, one or more)
- unweighted mean: $\mu_{..} \equiv \frac{1}{r} \sum_{i=1}^r \mu_{i.}$
 - weighted mean: $\mu_{..} \equiv \sum_{i=1}^r w_i \mu_{i.}$, $\sum_{i=1}^r w_i = 1$.
 - weighted (according to number in each factor level) mean:
 $\mu_{..} \equiv \sum_{i=1}^r w_i \mu_{i.}$, $w_i = \frac{n_{i.}}{n_T}$, (notice that $\sum_{i=1}^r w_i = 1$).

These definitions are sensible because, whereas we know n_T when calculating $\bar{Y}_{..}$, we often do not know the number of observations when calculating $\mu_{..}$ and so we do not know how to “properly” weigh each μ_i .

- (f) The *observed* difference between the observed patient response for drug 2 and the observed grand average is
 $\bar{Y}_{1.} - \bar{Y}_{..} = 5.50 - 5.50 =$ (circle one) **-0.03 / 0.00 / 0.08**.
- (g) The *expected* difference between the patient response for drug 2 and the overall mean is given by
 (circle one) **$\tau_1 / \tau_2 / \tau_3 / \tau_i$** .
- (h) The *expected* difference between the patient response for drug i and the overall mean is given by
 (circle one) **$\tau_1 / \tau_2 / \tau_3 / \tau_i$** .
- (i) Taking into account the influence of the different drugs, and making appropriate assumptions about the errors, the statistical model associated with this data is (circle two)
- i. $Y_{ij} = \mu + \tau_i + \varepsilon_j, \quad i = 1, 2, 3; \quad j = 1, 2, \dots, 5$
 - ii. $Y_{ij} = \mu_i + \varepsilon_{ij}, \quad i = 1, 2, 3; \quad j = 1, 2, \dots, n_i,$
 where $n_1 = n_3 = 5$ and $n_2 = 2$
 - iii. $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, 3; \quad j = 1, 2, \dots, n_i,$
 where $n_1 = n_3 = 5$ and $n_2 = 2$

3. *Chlorophyll In Leaves.* Calculate F using the following results obtained from a study measuring the amount of chlorophyll in three different types of leaves.

Treatment	Oak	Birch	Maple
replications, n_i	5	3	7
mean, \bar{Y}_i	4.67	3.53	1.40
standard deviation, s_i	1.06	0.27	0.27

Taking into account the influence of the different leaves, and making appropriate assumptions about the errors, the statistical model associated with this data is (circle two)

- (a) $Y_{ij} = \mu_i + \varepsilon_{ij}, \quad i = 1, 2, \dots, n_i; \quad j = 1, 2, \dots, n_i,$
 where $n_1 = 5, n_2 = 3$ and $n_3 = 7$
- (b) $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, 3; \quad j = 1, 2, \dots, n_i,$
 where $n_1 = 5, n_2 = 3$ and $n_3 = 7$
- (c) $Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, 3; \quad j = 1, 2, \dots, 15,$
 where $n_1 = 5, n_2 = 3$ and $n_3 = 7$

16.11 Regression Approach to Single-Factor Analysis of Variance

SAS program: att1-16-11-drugs,rats-regressionANOVA

The factor means version of the single-factor ANOVA model,

$$Y_{ij} = \mu. + \tau_i + \varepsilon_{ij}$$

can be written using matrix notation, in the following way,

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon$$

where

$$\mathbf{Y}_{n_T \times 1} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ \vdots \\ Y_{T,n_T} \end{bmatrix}, \quad \beta_{r \times 1} = \begin{bmatrix} \mu. \\ \tau_1 \\ \vdots \\ \tau_{r-1} \end{bmatrix}, \quad \varepsilon_{n_T \times 1} = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{T,n_T} \end{bmatrix}$$

and

$$\mathbf{X}_{n_T \times r} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1,r-1} \\ \vdots & & \cdots & \vdots \\ 1 & X_{n_T,1} & \cdots & X_{n_T,r-1} \end{bmatrix}$$

Exercise 16.6 (Regression Approach to Single-Factor Analysis of Variance)

1. *Comparing Three Drugs, Common Number of Replications For Each Treatment.* Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{y}_{1+} \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{y}_{2+} \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{y}_{3+} \approx 5.00$

(a) *Test equality of means*

i. *Statement*

The statement of the test is (choose one)

- (i) $H_0 : \tau_1 = \tau_2 = 0$ versus $H_a : \text{at least one } \tau_i \text{ not zero}$
- (ii) $H_0 : \tau_1 = \tau_2 \neq 0$ versus $H_a : \text{at least one } \tau_i \text{ not zero}$
- (iii) $H_0 : \tau_1 \neq \tau_2 = 0$ versus $H_a : \text{at least one } \tau_i \text{ not zero}$

ii. *Test.*

Since the test statistic is $F = 12.76$, the p-value, with $r - 1 = 3 - 1 = 2$ and $n_T - r = 12 - 3 = 9$ degrees of freedom, is given by

$$\text{p-value} = P(F \geq 12.76)$$

which equals (circle one) **0.002** / **0.35** / **0.43**.

(Use 2nd DISTR 9:Fcdf(12.76,E99,2,9).)

The level of significance is 0.05.

iii. *Conclusion.*

Since the p-value, 0.002, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that the average number of rats in the three locations are the same.

(b) *ANOVA*

From SAS, the ANOVA table is

Source	Degrees of Freedom	Sum Of Squares	Mean Squares
Treatment (Drugs)	2	2.033	1.0165
Error	12	0.0022	0.00018
Total	14	2.035	

which is (choose one) **the same as** / **different from** the previous ANOVA table for this data set.

(c) *Regression version of model*

True / False

Since,

$$Y_{ij} = \mu. + \tau_1 X_{ij1} + \tau_2 X_{ij2} + \varepsilon_{ij}$$

because⁴ $\tau_3 = -\tau_1 - \tau_2$ since $\sum_{i=1}^3 \tau_i = 0$,

and so, as a consequence,

$$X_{ij1} = \begin{cases} 1, & \text{if case from level 1} \\ -1, & \text{if case from level 3} \\ 0, & \text{otherwise,} \end{cases}$$

$$X_{ij2} = \begin{cases} 1, & \text{if case from level 2} \\ -1, & \text{if case from level 3} \\ 0, & \text{otherwise,} \end{cases}$$

⁴If the regression were defined as, say,

$$Y_{ij} = \mu. + \tau_1 X_{ij1} + \tau_2 X_{ij2} + \tau_3 X_{ij3} + \varepsilon_{ij}$$

then the resulting \mathbf{X} matrix would be singular; in other words, unsolvable. The regression must be defined with only two of the three parameters.

(d) *More regression version of model*

True / False

Since,

$$Y_{ij} = \mu. + \tau_1 X_{ij1} + \tau_2 X_{ij2} + \varepsilon_{ij}$$

where the X_{ij} are defined above, then

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon$$

where

$$\mathbf{Y} = \begin{bmatrix} 5.90 \\ 5.92 \\ \vdots \\ 4.98 \\ 5.02 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{bmatrix}, \quad \beta = \begin{bmatrix} \mu. \\ \tau_1 \\ \tau_2 \end{bmatrix}$$

and so

$$\mathbf{X}\beta = \begin{bmatrix} \mu. + \tau_1 \\ \mu. + \tau_1 \\ \vdots \\ \mu. - \tau_1 - \tau_2 \\ \mu. - \tau_1 - \tau_2 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_1 \\ \vdots \\ \mu_3 \\ \mu_3 \end{bmatrix}$$

(e) *Regression*

From SAS, the approximate regression to

$$Y_{ij} = \mu. + \tau_1 X_{ij1} + \tau_2 X_{ij2} + \varepsilon_{ij}$$

is (choose one)

(i) $\hat{Y} = 5.466 - 0.434X_1 + 0.032X_2$

(ii) $\hat{Y} = 5.466 + 0.433X_1 + 0.033X_2$

(iii) $\hat{Y} = 5.466 + 0.434X_1 - 0.032X_2$

2. *Rats in New York.*

Where are the rats in New York city? The rat count per square meter is recorded and shown in the following table.

sewers	parks	city hall
3	1	5
5	3	8
7	2	9
4		8
		10

Test the claim that the average rat count per square meter in at least two of the three city areas is different at 5%.

(a) *Test equality of means*

i. *Statement*

The statement of the test is (choose one)

(i) $H_0 : \tau_1 = \tau_2 = 0$ versus $H_a : \text{at least one } \tau_i \text{ not zero}$

(ii) $H_0 : \tau_1 = \tau_2 \neq 0$ versus $H_a : \text{at least one } \tau_i \text{ not zero}$

(iii) $H_0 : \tau_1 \neq \tau_2 = 0$ versus $H_a : \text{at least one } \tau_i \text{ not zero}$

ii. *Test.*

Since the test statistic is $F = 5863.5$, the p-value, with $r-1 = 3-1 = 2$ and $n_T - r = 15 - 3 = 12$ degrees of freedom, is given by

$$\text{p-value} = P(F \geq 5863.5)$$

which equals (circle one) **0.00** / **0.35** / **0.43**.

(Use 2nd DISTR 9:Fcdf(5863.5,E99,2,12).)

The level of significance is 0.05.

iii. *Conclusion.*

Since the p-value, 0.00, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that the average patient responses to the three drugs are the same.

(b) *Regression*

From SAS, the approximate regression to

$$Y_{ij} = \mu. + \tau_1 X_{ij1} + \tau_2 X_{ij2} + \varepsilon_{ij}$$

is (choose one)

(i) $\hat{Y} = 4.92 - 0.166X_1 - 2.92X_2$

(ii) $\hat{Y} = 5.466 + 0.434X_1 + 0.032X_2$

$$\text{(iii)} \quad \hat{Y} = 5.466 + 0.434X_1 - 0.032X_2$$

If $X_1 = 1$ and $X_2 = 0$ (the effect of the first factor level)

$$\hat{Y} = 4.92 - 0.166(1) - 2.92(0) = 4.75$$

If $X_1 = 0$ and $X_2 = 1$ (the effect of the second factor level)

$$\hat{Y} = 4.92 - 0.166(0) - 2.92(1) = 2$$

If $X_1 = -1$ and $X_2 = -1$ (the effect of the third factor level)

$$\hat{Y} = 4.92 - 0.166(-1) - 2.92(-1) = 8$$

Chapter 17

Analysis of Factor Level Effects

17.1 Introduction

We will look at estimation and testing of factor level means.

17.2 Plots of Estimated Factor Level Means

Line plots and normal probability plots can be used to visually determine if the factor level means are different or not. The normal probability plot is calculated by plotting the observed mean factor levels and comparing them to the *expected* mean factor levels, given by

$$\hat{Y}_{..} + z \left(\frac{k - 0.375}{r + 0.25} \right) \sqrt{\frac{MSE}{n}}$$

and k is the k ranked mean, r is the total number of means and where MSE is the mean squared error from a single-factor ANOVA table.

Exercise 17.1 (Simultaneous Inference Procedures)

1. *Line Plot.*

Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1 \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{Y}_2 \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3 \approx 5.00$

What does the following line plot of this drug data indicate about the three mean factor levels?

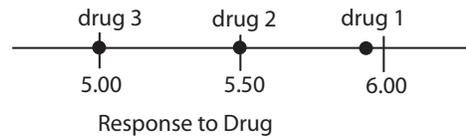


Figure 17.1 (Line Plot)

This line plot indicates (choose one)

- (a) average response drug 1 and average response drug 2 are different
 - (b) mean factor levels all different from one another
 - (c) mean factor levels all the same
 - (d) hard to tell
2. *Another Line Plot.* The tensile strength of fibers with varying amounts of cotton content is investigated.

5% cotton	7	7	15	11	9	10	$\bar{Y}_1. \approx 9.8$
10% cotton	12	17	12	18	18	16	$\bar{Y}_2. \approx 15.5$
15% cotton	14	18	18	19	19	17	$\bar{Y}_3. \approx 17.5$
20% cotton	19	25	22	19	23	24	$\bar{Y}_4. \approx 22$
25% cotton	7	10	11	15	11	14	$\bar{Y}_5. \approx 11.3$

What does the following line plot of this cotton data indicate about the three mean factor levels?

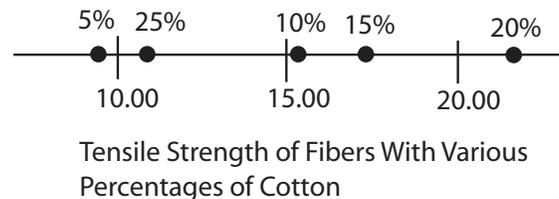


Figure 17.2 (Line Plot)

This line plot indicates (choose one)

- (a) mean factor levels 1, 2, 3 and 5 the same, but different from 4
 - (b) mean factor levels all different from one another
 - (c) mean factors 1, 5 the same; mean factors 2, 3 the same and mean factor 4 alone
3. *Normal Probability Plot.* Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1 \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{Y}_2 \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3 \approx 5.00$

(a) *Normal Probability Plot.*

Calculate and draw the normal probability plot of this drug data. What does it indicate about the three mean factor levels?

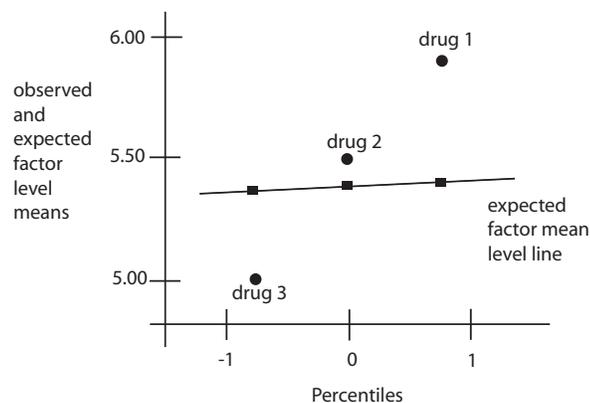


Figure 17.3 (Normal Probability Plot)

Since the observed mean factor levels do *not* fall along the line of expected mean factor levels, the normal probability plot indicates (choose one)

- average response drug 1 and average response drug 2 are different
- mean factor levels all different from one another
- mean factor levels all the same
- hard to tell

(b) *Plotting the normal probability plot.*

There are two plots on the normal probability plot. Both are constructed using the following table,

factor level	\hat{Y}_i	rank k	$z\left(\frac{k-0.375}{3+0.25}\right)$	$5.47 + z\left(\frac{k-0.375}{r+0.25}\right)(0.006)$
drug 1	5.90	3	0.869	5.475
drug 2	5.50	2	0	5.47
drug 3	5.00	1	-0.869	5.465

On the one hand, the *expected factor mean level line* is a plot of (choose one)

- \hat{Y}_i versus $z\left(\frac{k-0.375}{3+0.25}\right)$
- \hat{Y}_i versus rank k
- $5.47 + z\left(\frac{k-0.375}{r+0.25}\right)(0.006)$ versus $z\left(\frac{k-0.375}{3+0.25}\right)$

On the other hand, the plot of observed means is a plot of (choose one)

- i. \hat{Y}_i versus $z\left(\frac{k-0.375}{3+0.25}\right)$
 - ii. \hat{Y}_i versus rank k
 - iii. $5.47 + z\left(\frac{k-0.375}{r+0.25}\right)$ (0.006) versus $z\left(\frac{k-0.375}{3+0.25}\right)$
- (c) *Calculations for the normal probability plot.*

To determine

$$\hat{Y}_{..} + z\left(\frac{k-0.375}{r+0.25}\right) \sqrt{\frac{MSE}{n}},$$

first notice that the factor means are given in the table above, that the overall mean is $\bar{Y}_{..} \approx 5.47$, $n = 5$, and, from the following single-factor ANOVA table,

Source	Degrees of Freedom	Sum Of Squares	Mean Squares
Treatment (Drugs)	2	2.033	1.0165
Error	12	0.0022	0.00018
Total	14	2.035	

$MSE =$ (circle one) **0.00018** / **143.05** / **175.17**.

(Type data into L_1 , L_2 and L_3 and then type STAT TESTS F:ANOVA(L_1, L_2, L_3) ENTER.)

and so $\sqrt{\frac{MSE}{n}} = \sqrt{\frac{0.00018}{5}} \approx 0.006$ where, for example, for the mean response to drug 1, $\hat{Y}_i = 5.90$,

$$z\left(\frac{k-0.375}{3+0.25}\right) = z\left(\frac{3-0.375}{3+0.25}\right) \approx z(0.80769) \approx 0.869$$

(2nd DISTR invNorm(0.80769).)

17.3 Estimation and Testing of Estimated Factor Level Means

SAS program: att1-17-3-drugs-contrasts

In this section, we look at tests and estimation for the *linear combination* of factor level means,

$$L = \sum_{i=1}^r c_i \mu_i = c_1 \mu_1 + \cdots + c_r \mu_r$$

where we assume the ANOVA model I,

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

where μ_i are parameters and ε_{ij} are independent $N(0, \sigma^2)$.

Test The test is given by
 $H_o : L = c$ vs $H_a : L \neq c$
 with test statistic

$$t^* = \frac{\hat{L} - c}{s\{\hat{L}\}}$$

Confidence Interval The confidence interval is given by

$$\hat{L} \pm t\left(1 - \frac{\alpha}{2}; n_T - r\right)$$

where

$$\hat{L} = \sum_{i=1}^r c_i \bar{Y}_{i.} = c_1 \bar{Y}_{1.} + \dots + c_r \bar{Y}_{r.}$$

$$s\{\hat{L}\} = \sqrt{MSE \sum_{i=1}^r \frac{c_i^2}{n_i}}$$

and where MSE is the mean squared error from a single-factor ANOVA table. We look at four cases, including

- single factor mean μ_i ,
 where $c_i = 1$ and $c_j = 0$ for all $j \neq i, i, j = 1, \dots, r$
- difference between two factor level means,
 where $c_i = 1$ and $c_j = -1, i > j, i, j = 1, \dots, r$
- contrasts among factor level means,
 where $\sum_{i=1}^r c_i = 0, i = 1, \dots, r$
- linear combinations among factor level means,
 where there are no restrictions on $c_i, i = 1, \dots, r$

Exercise 17.2 (Tests And Estimation For The Linear Combination Of Factor Level Means)

1. *Examples of Linear Combinations: Drugs.*
 Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1. \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{Y}_2. \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3. \approx 5.00$

- (a) The difference between two factor level means, $\hat{L} = \bar{Y}_1 - \bar{Y}_2$, or,

$$\begin{aligned}\hat{L} &= c_1\bar{Y}_1 + c_2\bar{Y}_2 + c_3\bar{Y}_3 \\ &= (1)\bar{Y}_1 + (-1)\bar{Y}_2 + 0\bar{Y}_3 \\ &= (1)5.90 + (-1)5.50 + 0(5.00) \\ &= 0.40\end{aligned}$$

is a special case of a linear combination where

$$(c_1, c_2, c_3) = (\text{choose one}) \mathbf{(0, 1, -1)} / \mathbf{(1, -1, 0)} / \mathbf{(1, 1, 1)}$$

where because $c_1 + c_2 + c_3 = 1 - 1 + 0 = 0$,

(circle one) **is** / **is not** also a contrast.

- (b) The single factor level mean, $\hat{L} = \bar{Y}_1$, or,

$$\begin{aligned}\hat{L} &= c_1\bar{Y}_1 + c_2\bar{Y}_2 + c_3\bar{Y}_3 \\ &= (1)\bar{Y}_1 + (0)\bar{Y}_2 + 0\bar{Y}_3 \\ &= (1)5.90 + (0)5.50 + 0(5.00) \\ &= 5.90\end{aligned}$$

is a special case of a linear combination where

$$(c_1, c_2, c_3) = (\text{choose one}) \mathbf{(0, 1, -1)} / \mathbf{(1, -1, 0)} / \mathbf{(1, 0, 0)}$$

where because $c_1 + c_2 + c_3 = 1 + 0 + 0 \neq 0$,

(circle one) **is** / **is not** also a contrast.

- (c) If $(c_1, c_2, c_3) = (-1, 0, 1)$, then we are dealing with an example of a (choose none, one or more)

- i. single factor mean
- ii. difference between two factor level means
- iii. contrast among factor level means
- iv. linear combination among factor level means

$$\text{where } \hat{L} = (-1)5.90 + (0)5.50 + (1)(5.00) = (\text{choose one}) \mathbf{-0.90} / \mathbf{0} / \mathbf{0.90}$$

- (d) If $(c_1, c_2, c_3) = (-1, -1, 1)$, then we are dealing with an example of a (choose none, one or more)

- i. single factor mean
- ii. difference between two factor level means
- iii. contrast among factor level means
- iv. linear combination among factor level means

$$\begin{aligned}\text{where } \hat{L} &= (-1)5.90 + (-1)5.50 + (1)(5.00) = \\ &(\text{choose one}) \mathbf{-7.90} / \mathbf{-6.40} / \mathbf{-5.40}\end{aligned}$$

- (e) If $(c_1, c_2, c_3) = (-\frac{1}{3}, -\frac{1}{3}, \frac{2}{3})$, then we are dealing with an example of a (choose none, one or more)

Section 3. Estimation and Testing of Estimated Factor Level Means (ATTENDANCE 1)37

- i. single factor mean
- ii. difference between two factor level means
- iii. contrast among factor level means
- iv. linear combination among factor level means

where $\hat{L} = (-1/3)5.90 + (-1/3)5.50 + (2/3)(5.00) =$
 (choose one) $-\frac{1.4}{3} / -\frac{0.4}{3} / \frac{1.4}{3}$

2. More Examples of Linear Combinations: Fiber Strength.

The tensile strength of fibers with varying amounts of cotton content is investigated.

5% cotton	7	7	15	11	9	10	$\bar{Y}_1. \approx 9.8$
10% cotton	12	17	12	18	18	16	$\bar{Y}_2. \approx 15.5$
15% cotton	14	18	18	19	19	17	$\bar{Y}_3. \approx 17.5$
20% cotton	19	25	22	19	23	24	$\bar{Y}_4. \approx 22$
25% cotton	7	10	11	15	11	14	$\bar{Y}_5. \approx 11.3$

- (a) The difference between two factor level means,

$$\begin{aligned} \hat{L} &= c_1\bar{Y}_1. + c_2\bar{Y}_2. + c_3\bar{Y}_3. + c_4\bar{Y}_4. + c_5\bar{Y}_5. \\ &= (0)\bar{Y}_1. + (0)\bar{Y}_2. + (0)\bar{Y}_3. + (-1)\bar{Y}_4. + (1)\bar{Y}_5. \end{aligned}$$

is a special case of a linear combination where $(c_1, c_2, c_3, c_4, c_5) =$
 (choose one) $(\mathbf{0}, \mathbf{1}, -\mathbf{1}, \mathbf{1}, \mathbf{1}) / (\mathbf{1}, -\mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{0}) / (\mathbf{0}, \mathbf{0}, \mathbf{0}, -\mathbf{1}, \mathbf{1})$
 where, because $c_1 + c_2 + c_3 + c_4 + c_5 = 0 + 0 + 0 - 1 + 1 = 0$,
 (circle one) **is** / **is not** also a contrast.

- (b) If $(c_1, c_2, c_3, c_4, c_5) = (-\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, 1)$, then we are dealing with an example of a (choose none, one or more)

- i. single factor mean
- ii. difference between two factor level means
- iii. contrast among factor level means
- iv. linear combination among factor level means

where $\hat{L} = (-1/4)9.8 + (-1/4)15.5 + (-1/4)(17.5) + (-1/4)22 + (1)(11.3) =$
 (choose one) $-\mathbf{3.9} / -\mathbf{4.9} / -\mathbf{5.9}$

3. Test of Linear Combination, Single Factor Level: Drugs

Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1. \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{Y}_2. \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3. \approx 5.00$

Test if the mean patient response to drug 1 is different than 5.898 at $\alpha = 0.05$.

(a) *Statement.*

Let $L = \mu_1$; in other words,

$$\begin{aligned} L &= c_1\mu_1 + c_2\mu_2 + c_3\mu_3 \\ &= 0 + (1)\mu_1 + (0)\mu_2 + 0\mu_3 \end{aligned}$$

where $c_1 = 1$, $c_2 = 0$ and $c_3 =$ (circle one) $-1 / \mathbf{0} / \mathbf{1}$.

The statement of the test is (check none, one or more):

i. $H_0 : L = 5.898$ versus $H_a : L < 5.898$.

ii. $H_0 : L = 5.898$ versus $H_a : L > 5.898$.

iii. $H_0 : L = 5.898$ versus $H_a : L \neq 5.898$.

(b) *Test.*

From SAS, the test statistic is

$$t^* = \frac{\hat{L} - c}{s\{\hat{L}\}} = \frac{5.90 - 5.898}{0.006}$$

which is (circle one) **0.333** / **1.234** / **1.045**

which has p-value

$$\text{p-value} = 2 \times P(t \geq 0.33)$$

which equals (circle one) **0.00** / **0.39** / **0.75**.

The level of significance is 0.05.

(c) *Conclusion.*

Since the p-value, 0.75, is larger than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that mean patient response to drug 1 is equal to 5.898.

4. *Test of Linear Combination, Difference Between Two Factor Levels: Drugs*

Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1 \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{Y}_2 \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3 \approx 5.00$

Test if the mean patient response to drug 1 is different from the mean patient response to drug 2 at $\alpha = 0.05$.

(a) *Statement.*

Let $L = \mu_1 - \mu_2$; in other words,

$$\begin{aligned} L &= c_1\mu_1 + c_2\mu_2 + c_3\mu_3 \\ &= 0 + (1)\mu_1 + (-1)\mu_2 + 0\mu_3 \end{aligned}$$

Section 3. Estimation and Testing of Estimated Factor Level Means (ATTENDANCE 1)39

where $c_1 = 1$, $c_2 = -1$ and $c_3 =$ (circle one) $-1 / 0 / 1$. The statement of the test is (check none, one or more):

- i. $H_0 : L = 0$ versus $H_1 : L < 0$.
- ii. $H_0 : L = 0$ versus $H_1 : L > 0$.
- iii. $H_0 : L = 0$ versus $H_1 : L \neq 0$.

(b) *Test.*

From SAS, the test statistic is

$$t^* = \frac{\hat{L} - c}{s\{\hat{L}\}} = \frac{0.40 - 0}{0.0085}$$

which is (circle one) **46.71 / 63.23 / 73.45**.

which has a p-value given by

$$\text{p-value} = 2 \times P(t \geq 46.71)$$

which equals (circle one) **0.00 / 0.35 / 0.43**.

The level of significance is 0.05.

(c) *Conclusion.*

Since the p-value, 0.00, is smaller than the level of significance, 0.05, we (circle one) **accept / reject** the null hypothesis that mean patient response to drug 1 is equal to the mean patient response to drug 2.

5. *Test of Linear Combination, Contrast Among Factor Levels: Drugs*

Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1. \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{Y}_2. \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3. \approx 5.00$

Test if the mean patient response to drug 1 is different than the average of the mean patient response to drug 2 and drug 3 at $\alpha = 0.05$.

(a) *Statement.*

Let $L = \mu_1 - \frac{1}{2}(\mu_2 + \mu_3)$; in other words,

$$\begin{aligned} L &= c_1\mu_1 + c_2\mu_2 + c_3\mu_3 \\ &= (1)\mu_1 - \frac{1}{2}\mu_2 - \frac{1}{2}\mu_3 \end{aligned}$$

where $c_1 = 1$, $c_2 = -\frac{1}{2}$ and $c_3 =$ (circle one) $-1 / -\frac{1}{2} / 0$. The statement of the test is (check none, one or more):

- i. $H_0 : L = 0$ versus $H_1 : L < 0$.
- ii. $H_0 : L = 0$ versus $H_1 : L > 0$.
- iii. $H_0 : L = 0$ versus $H_1 : L \neq 0$.

(b) *Test.*

From SAS, the test statistic is

$$t^* = \frac{\hat{L} - c}{\hat{\sigma}_{\hat{L}}} = \frac{0.65 - 0}{0.00735}$$

which is (circle one) **58.93** / **63.23** / **88.65**.

which has a p-value given by

$$\text{p-value} = 2 \times P(t \geq 87.65)$$

which equals (circle one) **0.00** / **0.35** / **0.43**.

The level of significance is 0.05.

(c) *Conclusion.*

Since the p-value, 0.00, is smaller than the level of significance, 0.05, we (circle one) **accept** / **reject** the null hypothesis that mean patient response to drug 1 is equal to the average of the mean patient response to drug 2 and drug 3.

6. *95% Confidence Interval¹ of Linear Combination, Contrast Among Factor Levels, By Formula: Drugs*

Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1 \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{Y}_2 \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3 \approx 5.00$

A (two-sided) 95% CI for $L = \mu_1 - \frac{1}{2}(\mu_2 + \mu_3)$ is given by

$$\hat{L} \pm t(1 - \alpha/2; n_T - r)_s \left\{ \hat{L} \right\} = 0.65 \pm t(1 - 0.05/2; 15 - 3)(0.00735)$$

which is

(circle one) **(0.647, 0.659)** / **(0.657, 0.669)** / **(0.634, 0.666)**.

(To determine $t(12, 0.025)$, type PRGM INVT 12 ENTER 0.975 ENTER.)

Since this confidence interval does *not* include zero (0), this implies that this contrast (choose one) **is** / **is not** zero.

¹Although possible, it is awkward to get SAS to output confidence intervals and so, for this reason, is done “by formula” in this case.

7. Test of Linear Combination, Single Factor Level, By Formula: Drugs

Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1. \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{Y}_2. \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3. \approx 5.00$

Test if the mean patient response to drug 1 is different than 5.898 at $\alpha = 0.05$.

(a) Statement.

Let $L = \mu_1$; in other words,

$$\begin{aligned} L &= c_1\mu_1 + c_2\mu_2 + c_3\mu_3 \\ &= 0 + (1)\mu_1 + (0)\mu_2 + 0\mu_3 \end{aligned}$$

where $c_1 = 1$, $c_2 = 0$ and $c_3 =$ (circle one) $-1 / 0 / 1$.

The statement of the test is (check none, one or more):

- i. $H_0 : L = 5.898$ versus $H_a : L < 5.898$.
- ii. $H_0 : L = 5.898$ versus $H_a : L > 5.898$.
- iii. $H_0 : L = 5.898$ versus $H_a : L \neq 5.898$.

(b) Test.

Since

$$\begin{aligned} \hat{L} &= c_1\bar{Y}_1. + c_2\bar{Y}_2. + c_3\bar{Y}_3. \\ &= 0 + (1)\bar{Y}_1. + (0)\bar{Y}_2. + 0\bar{Y}_3. \\ &= 0 + (1)5.90 + (0)5.50 + 0(5.00) = \end{aligned}$$

(circle one) **5.00 / 5.50 / 5.90** and, from the ANOVA table,

Source	Degrees of Freedom	Sum Of Squares	Mean Squares
Treatment	2	2.033	1.0165
Error	12	0.0022	0.00018
Total	14	2.035	

where, in particular, $MSE = 0.00018$, then

(Type data into L_1 , L_2 and L_3 of TI-83, and then type STAT TESTS F:ANOVA(L_1, L_2, L_3) ENTER.)

$$\begin{aligned} s \{ \hat{L} \} &= \sqrt{MSE \sum_{i=1}^r \frac{c_i^2}{n_i}} \\ &= \sqrt{MSE \left\{ \frac{c_1^2}{n_1} + \frac{c_2^2}{n_2} + \frac{c_3^2}{n_3} \right\}} \\ &= \sqrt{0.00018 \left\{ \frac{(1)^2}{5} + \frac{(0)^2}{5} + \frac{0^2}{5} \right\}} \end{aligned}$$

which is (circle one) **0.006 / 0.0848 / 0.8485**
and so the test statistic is

$$t^* = \frac{\hat{L} - c}{s\{\hat{L}\}} = \frac{5.90 - 5.898}{0.006}$$

which is (circle one) **0.333 / 1.234 / 1.045**.

Since the test statistic is $t^* = 0.33$, the p-value, with $t - 1 = 3 - 1 = 2$ degrees of freedom, is given by

$$\text{p-value} = 2 \times P(t \geq 0.33)$$

which equals (circle one) **0.00 / 0.39 / 0.75**.

(Use 2nd DISTR 5:tcdf(0.33,E99,2).)

The level of significance is 0.05.

(c) *Conclusion.*

Since the p-value, 0.75, is larger than the level of significance, 0.05, we (circle one) **accept / reject** the null hypothesis that mean patient response to drug 1 is equal to 5.898^2 .

8. *A Word About Notation* The most general notation has been used throughout this exercise, the linear combination/contrast, $\hat{L} = \sum_{i=1}^r c_i \hat{Y}_i$, with corresponding $s\{\hat{L}\}$. Other specialized notation is possible, including,

- in the single factor mean μ_i case, with \hat{Y}_i and corresponding $s\{\hat{Y}_i\}$.
- in difference between two factor level means case, with \hat{D} and corresponding $s\{\hat{D}\}$.

17.4 Need For Simultaneous Inference Procedures

Estimations and tests must be altered, particularly, the level of significance, α is made *smaller*, to account for when either (simultaneous, multiple) series of estimates or tests or if data snooping occurs. When the level of significance is altered in this way, it is called a *family* level of significance, α . There are many different ways to determine the family level of significance, α . Four are discussed in *future* sections of this chapter and include,

- Tukey

²We could have used STAT TESTS T-Test here also.

- Scheffe
- Bonferroni
- Holm

Exercise 17.3 (Simultaneous Inference Procedures) Fifteen different patients are subjected to three drugs.

drug 1	5.90	5.92	5.91	5.89	5.88	$\bar{Y}_1. \approx 5.90$
drug 2	5.51	5.50	5.50	5.49	5.50	$\bar{Y}_2. \approx 5.50$
drug 3	5.01	5.00	4.99	4.98	5.02	$\bar{Y}_3. \approx 5.00$

1. *Simultaneous Tests*

Consider the following two situations.

- *One Test.*

A random sample is collected to test $L = \mu_1 - \mu_2$,

$H_0 : L = 0$ versus $H_1 : L \neq 0$

- *Two Simultaneous Tests.*

Two different random samples are collected,

one to test $L_1 = \mu_1 - \mu_2$,

$H_0 : L_1 = 0$ versus $H_1 : L_1 \neq 0$

and one to test $L_2 = \mu_1 - \mu_3$,

$H_0 : L_2 = 0$ versus $H_1 : L_2 \neq 0$

- Since *two* random samples were collected in the two-test case, there is a (choose one) **greater** / **lesser** chance, relative to the one-test case, that at least one of the two tests involved will be significant.
- Since a test is significant if the p-value is smaller than the level of significance, α , the α is “too big (inflated)” in the two-test case. Consequently, it seems sensible to make the level of significance, α , (choose one) **bigger** / **smaller** to account for the inflation that occurs when two simultaneous tests are conducted.
- If more than two tests are conducted simultaneously, then α should be made (choose one) **smaller** / **bigger** than the α used if only two simultaneous tests are conducted.

2. *Simultaneous Confidence Intervals*

Consider the following two situations.

- *One Confidence Interval (CI).*

A random sample is collected to determine the confidence interval for $L = \mu_1 - \mu_2$,

$$\hat{L} \pm t(1 - \alpha/2; n_T - r)s \left\{ \hat{L} \right\}$$

- *Two Simultaneous Confidence Intervals.*

Two different random samples are collected,

one to determine the confidence interval for $L_1 = \mu_1 - \mu_2$,

$$\hat{L}_1 \pm t(1 - \alpha/2; n_T - r)s \left\{ \hat{L}_1 \right\}$$

and to determine the confidence interval for $L_2 = \mu_1 - \mu_3$,

$$\hat{L}_2 \pm t(1 - \alpha/2; n_T - r)s \left\{ \hat{L}_2 \right\}$$

- (a) Since the width of a CI is determined by $1 - \alpha$, a *narrow* CI has (choose one) **big** / **small** α .
- (b) In the two-CI case, when the inflated (“too big”) α is made smaller, the two CIs are made (choose one) **narrower** / **wider**.

3. *Problem of data snooping*

Consider the following two situations.

- *One Test.*

A random sample is collected to test *one* (only one, chosen at random) of the following contrasts,

$$L_1 = \mu_1 - \mu_2,$$

$$L_2 = \mu_1 - \mu_3,$$

$$L_3 = \mu_2 - \mu_3,$$

$$H_0 : L = 0 \text{ versus } H_1 : L \neq 0$$

- *Data Snooping Test.*

A random sample is collected to test $L = \mu_i - \mu_j$, where μ_i and μ_j are chosen according to the observed sample means that are *most* different. In this case, since $(\hat{Y}_1, \hat{Y}_2, \hat{Y}_3) = (5.90, 5.50, 5.00)$, and $5.90 - 5.00 = 0.90$ is the largest difference of the three possible pairs, we will test $L_3 = \mu_1 - \mu_3$.

- (a) For a given α , a test of $L_3 = \mu_1 - \mu_3$ is (choose one) **most** / **least** likely to be significant, as compared to a test of either L_1 or L_2 .
- (b) **True** / **False**
Rather than decide which few contrasts to test after looking at the data, decide which contrasts to test before looking at the data. Or, test all possible contrasts to avoid favoring a select few contrasts.

4. *Order The Factor Level Means From Smallest To Largest.*

Many of the simultaneous tests require ordering the factor level means.

- (a) The factor level means given above are (circle none, one or more)
- set A: $\bar{Y}_1 \approx 9.8, \bar{Y}_2 \approx 15.5, \bar{Y}_3 \approx 17.5, \bar{Y}_4 \approx 22, \bar{Y}_5 \approx 11.3$
 - set B: $\bar{Y}_{(1)} \approx 9.8, \bar{Y}_{(2)} \approx 11.3, \bar{Y}_{(3)} \approx 15.5, \bar{Y}_{(4)} \approx 17.5, \bar{Y}_{(5)} \approx 22$
 - set C: $\bar{Y}_{(1)} \approx 8.8, \bar{Y}_{(2)} \approx 11.3, \bar{Y}_{(3)} \approx 15.5, \bar{y}_{(4)} \approx 17.5, \bar{Y}_{(5)} \approx 23$
- (b) Although sets A and B above contain the *same* five averages, (circle one) **set A / set B** is *ordered* from smallest to largest. Notice that the notation used to indicate an ordered set of averages involves using a bracket on the subscript; for example, $\bar{Y}_{(i)}$ instead of \bar{Y}_i .
- (c) If $\bar{Y}_1 \approx -2.8, \bar{Y}_2 \approx 5.5, \bar{Y}_3 \approx 1.5, \bar{Y}_4 \approx 2, \bar{Y}_5 \approx 4.3$, then $\bar{y}_{(3)} \approx$ (circle one) **-2.8 / 2 / 4.3**.
- (d) The average, $\bar{Y}_{(2)}$, is the (circle none, one or more)
- average one larger than the smallest average
 - average one smaller than the largest average
 - second smallest average
 - second largest average
- (e) *One Last Word On Notation.* Although a common notation, the bracket notation is *not* used in this text. Instead, the averages, if not ordered smallest to largest, are ordered in this way, before proceeding with the Tukey pairwise difference procedure.

17.5 Tukey Multiple Comparison Procedure

SAS program: att1-17-5-cotton-multiple-tukey

We look at confidence intervals and significance for the Tukey pairwise (multiple) comparison of means (which, recall, is a special important case of the linear combination of means).

Test The $\frac{r(r-1)}{2}$ pairwise tests are given by

$$H_o : \mu_i - \mu_j = 0 \text{ vs } H_a : \mu_i - \mu_j \neq 0, i > j, i, j = 1, \dots, r$$

with test statistic

$$q^* = \frac{\sqrt{2}\hat{D}}{s\{\hat{D}\}}$$

and the null is *rejected* if

$$|q^*| > q(1 - \alpha; r; n_T - r)$$

where the values of the *studentized range distribution*, $q(r, \nu)$, are given in table B.9, pages 1352–1354 of the Neter text.

Confidence Interval The confidence interval is given by

$$\hat{D} \pm \frac{1}{\sqrt{2}}q(1 - \alpha; r; n_T - r) s \left\{ \hat{D} \right\}$$

where

$$\begin{aligned} \hat{D} &= \bar{Y}_i - \bar{Y}_j. \\ s \left\{ \hat{D} \right\} &= \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \end{aligned}$$

and where MSE is the mean squared error from the associated single-factor ANOVA table.

Exercise 17.4 (Tukey Pairwise Comparisons: Fiber Tensile Strength) The tensile strength of fibers with varying amounts of cotton content is investigated.

5% cotton	7	7	15	11	9	10	$\bar{Y}_1. \approx 9.8$
10% cotton	12	17	12	18	18	16	$\bar{Y}_2. \approx 15.5$
15% cotton	14	18	18	19	19	17	$\bar{Y}_3. \approx 17.5$
20% cotton	19	25	22	19	23	24	$\bar{Y}_4. \approx 22$
25% cotton	7	10	11	15	11	14	$\bar{Y}_5. \approx 11.3$

1. *Tukey Pairwise Difference Tests*, $D = \mu_i - \mu_j$, $i > j$.

Use Tukey's test to decide if there are any significant differences in the i -th mean tensile strength of fiber and the j -th mean tensile strength of fiber, for all i and j , $i > j$, at $\alpha = 0.05$.

- (a) *Statement.*

Let $D = \mu_i - \mu_j$. The statement of the tests for the ten cases is (circle one):

- i. $H_0 : D \geq 0$ versus $H_1 : D < 0$.
- ii. $H_0 : D \leq 0$ versus $H_1 : D > 0$.
- iii. $H_0 : D = 0$ versus $H_1 : D \neq 0$.

- (b) *Test and Conclusions.*

From SAS,

the mean 20% cotton strength ($\bar{Y}_4. = 22$, μ_4)

is found to be significantly different

from mean 15% cotton strength ($\bar{Y}_3. = 17.5$, μ_3)

whereas

the mean 5% cotton strength ($\bar{Y}_1. = 9.8$, μ_1)

is found to be the same as

from mean 10% cotton strength ($\bar{Y}_2. = 11.3$, μ_2)

for example. All results are given in the graph below.

ordered mean	5% cotton, (1)	25% cotton, (5)	10% cotton, (2)	15% cotton, (3)	20% cotton, (4)
	9.8	11.3	15.5	17.5	22

For example, this shows μ_5 and μ_2 (choose one) **are** / **are not** significantly different from one another.

2. *Tukey Pairwise Difference Confidence Intervals*, $D = \mu_i - \mu_j$, $i > j$.

From SAS, the confidence interval for $\mu_4 - \mu_3$ is given by (0.012, 8.988), whereas for $\mu_5 - \mu_1$ it is

(choose one) **(-2.988, 5.988)** / **(1.178, 10.155)** / **(2.012, 10.988)**

3. *Tukey Pairwise Difference Tests*, $D = \mu_i - \mu_j$, $i > j$, in detail³.

Use Tukey's test to decide if there are any significant differences in the i -th mean tensile strength of fiber and the j -th mean tensile strength of fiber, for all i and j , $i > j$, at $\alpha = 0.05$.

- (a) *Statement*.

Let $D = \mu_i - \mu_j$. The statement of the tests for the ten cases is (circle one):

- i. $H_0 : D \geq 0$ versus $H_1 : D < 0$.
- ii. $H_0 : D \leq 0$ versus $H_1 : D > 0$.
- iii. $H_0 : D = 0$ versus $H_1 : D \neq 0$.

The $\frac{r(r-1)}{2} = \frac{5(4)}{2} = 10$ tests are

- $H_0 : \mu_1 - \mu_2 = 0$ versus $H_1 : \mu_1 - \mu_2 \neq 0$
- $H_0 : \mu_1 - \mu_3 = 0$ versus $H_1 : \mu_1 - \mu_3 \neq 0$
- $H_0 : \mu_1 - \mu_4 = 0$ versus $H_1 : \mu_1 - \mu_4 \neq 0$
- $H_0 : \mu_1 - \mu_5 = 0$ versus $H_1 : \mu_1 - \mu_5 \neq 0$
- $H_0 : \mu_2 - \mu_3 = 0$ versus $H_1 : \mu_2 - \mu_3 \neq 0$
- $H_0 : \mu_2 - \mu_4 = 0$ versus $H_1 : \mu_2 - \mu_4 \neq 0$
- $H_0 : \mu_2 - \mu_5 = 0$ versus $H_1 : \mu_2 - \mu_5 \neq 0$
- $H_0 : \mu_3 - \mu_4 = 0$ versus $H_1 : \mu_3 - \mu_4 \neq 0$
- $H_0 : \mu_3 - \mu_5 = 0$ versus $H_1 : \mu_3 - \mu_5 \neq 0$
- $H_0 : \mu_4 - \mu_5 = 0$ versus $H_1 : \mu_4 - \mu_5 \neq 0$

where the factor level means have been *ordered from smallest to largest*⁴,

³All of the means are *ordered* in this exercise, unlike the previous exercise—it makes it difficult to compare the results of the two exercises.

⁴This ordering was *not* done in SAS output results given above and so it will be difficult to compare the results here with the previous SAS output results.

5% cotton	$\bar{Y}_1 \approx 9.8$	(originally \bar{Y}_1)
25% cotton	$\bar{Y}_2 \approx 11.3$	(originally \bar{Y}_5)
10% cotton	$\bar{Y}_3 \approx 15.5$	(originally \bar{Y}_2)
15% cotton	$\bar{Y}_4 \approx 17.5$	(originally \bar{Y}_3)
20% cotton	$\bar{Y}_5 \approx 22$	(originally \bar{Y}_4)

(b) *Test.*

On the one hand, since the ANOVA table is given by,

Source	Degrees of Freedom	Sum Of Squares	Mean Squares	F
Treatment (% Cotton)	4	572.2	143.05	20.42
Error	25	175.17	7.007	
Total	29	747.37		

where $MSE =$ (circle one) **7.007 / 143.05 / 175.17**.(Type data into L_1, L_2, L_3, L_4 and L_5 and then type STAT TESTS F:ANOVA(L_1, L_2, L_3, L_4, L_5) ENTER.)and also $n_i = n_j = 5, i, j = 1, \dots, 5$, then

$$\begin{aligned}
 s\{\hat{D}\} &= \sqrt{MSE \left(\frac{1}{n_i} + \frac{1}{n_j} \right)} \\
 &= \sqrt{7.007 \left(\frac{1}{5} + \frac{1}{5} \right)}
 \end{aligned}$$

which is equal to (circle one) **1.52 / 4.52 / 5.18** and so the various test statistics, q^* , are given by

pairs	i	j	$\hat{D} = \bar{Y}_i - \bar{Y}_j$	$q^* = \frac{\sqrt{2}\hat{D}}{s\{\hat{D}\}} = \frac{\sqrt{2}\hat{D}}{1.52}$
25% and 5% cotton	2	1	1.5	1.39
10% and 5% cotton	3	1	5.7	5.30
15% and 5% cotton	4	1	7.7	7.16
20% and 5% cotton	5	1	12.2	11.35
10% and 10% cotton	3	2	4.2	3.91
15% and 10% cotton	4	2	6.2	5.77
25% and 10% cotton	5	2	10.7	9.96
15% and 15% cotton	4	3	2	1.86
20% and 15% cotton	5	3	6.5	6.05
20% and 15% cotton	5	4	4.5	4.19

On the other hand, the value of the studentized range (critical value) is given by

$$q(1 - \alpha; r; n_T - r) = q(1 - 0.05; 5, 30 - 5)$$

which is equal to (circle one) **2.54 / 4.17 / 5.11**.

(Hint: Table B.9, pages 1352–1354.)

(c) *Conclusions.*

One way to present the conclusions is to give a table such as the one below (fill in the blanks).

i	j	$ q^* $	$q(0.95; 5, 30 - 5)$	significant?
2	1	1.39	4.17	N
3	1	5.30	4.17	S
4	1	7.16	4.17	S
5	1	11.35	4.17	_____
3	2	3.91	4.17	N
4	2	5.77	4.17	S
5	2	9.96	4.17	S
4	3	1.86	4.17	_____
5	3	6.05	4.17	S
5	4	4.19	4.17	N

Another way to present conclusions are given in the graph below⁵.

ordered mean	5% cotton, (1)	25% cotton, (2)	10% cotton, (3)	15% cotton, (4)	20% cotton, (5)
	9.8	11.3	15.5	17.5	22

	_____			_____	
	_____		_____		
	_____	_____			

(d) *Some Further Comments.*

- i. The sample sizes for the five factor levels in the example are (choose one) **equal (to 6)** / **unequal**. It is still possible to perform the Tukey test if the sample sizes are unequal where $n_i \neq n_j$, $i \neq j$ using the formulas given above.
- ii. Tukey's test is used to test (choose one)
 - A. single factor mean μ_i , where $c_i = 1$ and $c_j = 0$ for all $j \neq i$, $i, j = 1, \dots, r$
 - B. difference between two factor level means, where $c_i = 1$ and $c_j = -1$, $i > j$, $i, j = 1, \dots, r$
 - C. contrasts among factor level means, where $\sum_{i=1}^r c_i = 0$, $i = 1, \dots, r$
 - D. linear combinations among factor level means, where there are no restrictions on c_i , $i = 1, \dots, r$

However, this test could be used to test contrasts or linear combinations. It is *not* used in these more general cases because other tests (Scheffe, for example) are better tests in these more general situations.

⁵These results do not match the previous results exactly because of round-off error.

4. Tukey Pairwise Difference Confidence Intervals, $D = \mu_i - \mu_j$, $i > j$, by formula⁶. The confidence intervals are given by

$$\hat{D} \pm \frac{1}{\sqrt{2}} q(1 - \alpha; r; n_T - r) s \left\{ \hat{D} \right\}$$

where, from above, $q(1 - \alpha; r; n_T - r) = q(0.95; 5; 25) \approx 4.17$ and $s \left\{ \hat{D} \right\} = 1.52$ and so

pairs	i	j	$\hat{D} = \bar{Y}_i - \bar{Y}_j$	$\hat{D} \pm \frac{1}{\sqrt{2}}(4.17)(1.52)$
25% and 5% cotton	2	1	1.5	(-2.988, 5.988)
10% and 5% cotton	3	1	5.7	(1.178, 10.1551)
15% and 5% cotton	4	1	7.7	(3.178, 12.155)
20% and 5% cotton	5	1	12.2	(7.678, 16.655)
10% and 10% cotton	3	2	4.2	(-0.322, 8.655)
15% and 10% cotton	4	2	6.2	(1.678, 10.655)
25% and 10% cotton	5	2	10.7	(6.178, 15.155)
15% and 15% cotton	4	3	2	(-2.488, 6.488)
20% and 15% cotton	5	3	6.5	(2.012, 10.988)
20% and 15% cotton	5	4	4.5	(0.012, 8.988)

where the four *nonsignificant* factor level mean differences are identified as the confidence intervals that (choose one) **include** / **exclude** zero (0).

17.6 Scheffe Multiple Comparison Procedure

SAS program: att1-17-6-cotton-multiple-scheffe

We look at confidence intervals and significance for the Scheffe multiple linear contrasts of factor level means.

Test The tests are given by

$$H_o : L = 0 \text{ vs } H_a : L \neq 0, L = \sum_{i=1}^r c_i \mu_i, \sum_{i=1}^r c_i = 0$$

with test statistic

$$F^* = \frac{\hat{L}^2}{(r - 1)s^2 \left\{ \hat{L} \right\}}$$

Confidence Interval The confidence interval is given by

$$\hat{L} \pm \sqrt{(r - 1)F(1 - \alpha; r - 1; n_T - r)s \left\{ \hat{L} \right\}}$$

⁶Again, the means have been ordered in this case.

where

$$\hat{L} = \sum_{i=1}^r c_i \bar{Y}_i.$$

$$s^2 \{ \hat{L} \} = MSE \sum_{i=1}^r \frac{c_i^2}{n_i}$$

and where MSE is the mean squared error from the associated single-factor ANOVA table.

Exercise 17.5 (Scheffe Contrast Comparisons: Fiber Tensile Strength) The tensile strength of fibers with varying amounts of cotton content is investigated.

5% cotton	7	7	15	11	9	10	$\bar{Y}_1 \approx 9.8$
10% cotton	12	17	12	18	18	16	$\bar{Y}_2 \approx 15.5$
15% cotton	14	18	18	19	19	17	$\bar{Y}_3 \approx 17.5$
20% cotton	19	25	22	19	23	24	$\bar{Y}_4 \approx 22$
25% cotton	7	10	11	15	11	14	$\bar{Y}_5 \approx 11.3$

We are interested in determining the simultaneous Scheffe tests and confidence intervals for the following four (4) contrasts.

$$L_1 = -\frac{1}{4}\mu_1 - \frac{1}{4}\mu_2 - \frac{1}{4}\mu_3 - \frac{1}{4}\mu_4 + \mu_5$$

$$L_2 = \mu_2 - \mu_3$$

$$L_3 = -\mu_1 + \mu_4$$

$$L_4 = -\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 + \frac{1}{2}\mu_3 - \frac{1}{2}\mu_4$$

1. *Scheffe Contrast Tests*, $L = \sum_{i=1}^r c_i \mu_i$.

Use Scheffe's tests to decide if any of the four contrasts are significant at $\alpha = 0.05$.

- (a) *Statement*.

The statement of the tests for the four cases is (circle one):

- i. $H_0 : L_i \geq 0$ versus $H_1 : L_i < 0$, $i = 1, \dots, 4$.
- ii. $H_0 : L_i \leq 0$ versus $H_1 : L_i > 0$, $i = 1, \dots, 4$.
- iii. $H_0 : L_i = 0$ versus $H_1 : L_i \neq 0$, $i = 1, \dots, 4$.

- (b) *Test*.

On the one hand, since

$$F^* = \frac{\hat{L}^2}{(r-1)s^2 \{ \hat{L} \}}$$

the different values of the four test statistics are given by

L	$(c_1, c_2, c_3, c_4, c_5)$	\hat{L}	$s \left\{ \hat{L} \right\}$	$F^* = \frac{\hat{L}^2}{(5-1)s^2 \left\{ \hat{L} \right\}}$
L_1	$(-\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, 1)$	-4.875	1.2082	3.2562
L_2	$(0, 1, -1, 0, 0)$	-2	1.5283	0.34253
L_3	$(-1, 0, 0, 1, 0)$	12.167	1.5283	12.676
L_4	$(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, 0)$	0.5833	1.0806	0.05828

On the other hand, the critical F value is given by

$$F(1 - \alpha; r - 1; n_T - r) = F(1 - 0.05; 5 - 1, 30 - 5)$$

which is equal to (circle one) **2.76** / **4.17** / **5.11**.

(PRGM INV F 4 ENTER 25 ENTER 0.95 ENTER)

(c) *Conclusions.*

Comparing the test statistics, F^* , with the critical value, F , (fill in the blanks),

L	F^*	F	significant?
L_1	3.26	2.76	S
L_2	0.34	2.76	N
L_3	12.68	2.76	_____
L_4	0.058	2.76	_____

where “S” is “significant” and “N” is “not significant”.

2. *Scheffe Contrast Confidence Intervals*, $L = \sum_{i=1}^r c_i \mu_i$.

The confidence intervals are given by

$$\hat{L} \pm \sqrt{(r-1)F(1-\alpha; r-1; n_T-r)} s \left\{ \hat{L} \right\}$$

where, from above, $F(1-\alpha; r-1; n_T-r) = F(0.95; 4; 25) \approx 2.76$ and so

L	$(c_1, c_2, c_3, c_4, c_5)$	\hat{L}	$s \left\{ \hat{L} \right\}$	$\hat{L} \pm \sqrt{4 \times 2.76} s \left\{ \hat{L} \right\}$
L_1	$(-\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, 1)$	-4.875	1.2082	(-8.889, -0.8606)
L_2	$(0, 1, -1, 0, 0)$	-2	1.5283	(-7.078, 3.0778)
L_3	$(-1, 0, 0, 1, 0)$	12.167	1.5283	(7.0888, 17.245)
L_4	$(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, 0)$	0.5833	1.0806	(-3.007, 4.1739)

where the one *nonsignificant* factor level mean differences are identified as the confidence intervals that (choose one) **include** / **exclude** zero (0).

3. *Scheffe Contrast Tests*, $L = \sum_{i=1}^r c_i \mu_i$, *in detail*. Use Scheffe’s tests to decide if any of the four contrasts are significant at $\alpha = 0.05$.

(a) *Statement.*

The statement of the tests for the four cases is (circle one):

- i. $H_0 : L_k \geq 0$ versus $H_1 : L_k < 0, k = 1, \dots, 4.$
- ii. $H_0 : L_k \leq 0$ versus $H_1 : L_k > 0, k = 1, \dots, 4.$
- iii. $H_0 : L_k = 0$ versus $H_1 : L_k \neq 0, k = 1, \dots, 4.$

(b) *Test.*

On the one hand, the different values of the four test statistics are given by

$$F^* = \frac{\hat{L}^2}{(r-1)s^2 \left\{ \hat{L} \right\}}$$

where, since the ANOVA table is given by,

Source	Degrees of Freedom	Sum Of Squares	Mean Squares	F
Treatment (% Cotton)	4	572.2	143.05	20.42
Error	25	175.17	7.007	
Total	29	747.37		

where $MSE =$ (circle one) **7.007 / 143.05 / 175.17.**

and also

$$\begin{aligned} s^2 \left\{ \hat{L} \right\} &= MSE \sum_{i=1}^r \frac{c_i^2}{n_i} \\ &= 7.007 \left(\frac{c_1^2}{6} + \frac{c_2^2}{6} + \frac{c_3^2}{6} + \frac{c_4^2}{6} + \frac{c_5^2}{6} \right) \end{aligned}$$

and so

L	$(c_1, c_2, c_3, c_4, c_5)$	\hat{L}	$s \left\{ \hat{L} \right\}$	$F^* = \frac{\hat{L}^2}{(5-1)s^2 \left\{ \hat{L} \right\}}$
L_1	$(-\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, 1)$	-4.875	1.2082	3.2562
L_2	$(0, 1, -1, 0, 0)$	-2	1.5283	0.34253
L_3	$(-1, 0, 0, 1, 0)$	12.167	1.5283	12.676
L_4	$(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, 0)$	0.5833	1.0806	0.05828

On the other hand, the critical F value is given by

$$F(1 - \alpha; r - 1; n_T - r) = F(1 - 0.05; 5 - 1, 30 - 5)$$

which is equal to (circle one) **2.76 / 4.17 / 5.11.**

(PRGM INVF 4 ENTER 25 ENTER 0.95 ENTER)

(c) *Conclusions.*

Comparing the test statistics, F^* , with the critical value, F , (fill in the blanks),

L	F^*	F	significant?
L_1	3.26	2.76	S
L_2	0.34	2.76	N
L_3	12.68	2.76	_____
L_4	0.058	2.76	_____

(d) *Some Further Comments.*

- i. The sample sizes for the five factor levels in the example are (choose one) **equal (to 5)** / **unequal**. It is still possible to perform the Scheffe test if the sample sizes are unequal where $n_i \neq n_j, i \neq j$ using the formulas given above.
- ii. Scheffe's test is used to test (choose one or more)
 - A. single factor mean μ_i , where $c_i = 1$ and $c_j = 0$ for all $j \neq i, i, j = 1, \dots, r$
 - B. difference between two factor level means, where $c_i = 1$ and $c_j = -1, i > j, i, j = 1, \dots, r$
 - C. contrasts among factor level means, where $\sum_{i=1}^r c_i = 0, i = 1, \dots, r$
 - D. linear combinations among factor level means, where there are no restrictions on $c_i, i = 1, \dots, r$

This test could be used to test pairwise difference in two factor level means. It is *not* used in this more specialized case because other tests (Tukey, for example) are better tests in this more special situation.

17.7 Bonferroni Multiple Comparison Procedure

SAS program: att1-17-7-cotton-multiple-bonferroni

We look at confidence intervals and significance for the Bonferroni multiple linear combinations of factor level means.

Test The tests are given by

$$H_o : L = 0 \text{ vs } H_a : L \neq 0, L = \sum_{i=1}^r c_i \mu_i$$

with test statistic

$$t^* = \frac{\hat{L}}{s\{\hat{L}\}}$$

and the null is *rejected* if

$$|t^*| > t(1 - \alpha/(2g); n_T - r)$$

where g is number of linear combinations considered in the problem.

Confidence Interval The confidence interval is given by

$$\hat{L} \pm t(1 - \alpha/(2g); n_T - r) s\{\hat{L}\}$$

where

$$\hat{L} = \sum_{i=1}^r c_i \bar{Y}_i.$$

$$s\{\hat{L}\} = \sqrt{MSE \sum_{i=1}^r \frac{c_i^2}{n_i}}$$

and where MSE is the mean squared error from the associated single-factor ANOVA table.

Exercise 17.6 (Bonferroni Linear Combination Comparisons: Fiber Tensile Strength) The tensile strength of fibers with varying amounts of cotton content is investigated.

5% cotton	7	7	15	11	9	10	$\bar{Y}_1. \approx 9.8$
10% cotton	12	17	12	18	18	16	$\bar{Y}_2. \approx 15.5$
15% cotton	14	18	18	19	19	17	$\bar{Y}_3. \approx 17.5$
20% cotton	19	25	22	19	23	24	$\bar{Y}_4. \approx 22$
25% cotton	7	10	11	15	11	14	$\bar{Y}_5. \approx 11.3$

We are interested in determining the simultaneous Bonferroni tests and confidence intervals for the following four ($g = 4$) contrasts/combinations.

$$L_1 = -\frac{1}{4}\mu_1 - \frac{1}{4}\mu_2 - \frac{1}{4}\mu_3 - \frac{1}{4}\mu_4 + \mu_5$$

$$L_2 = \mu_2 - \mu_3$$

$$L_3 = -\mu_1 + \mu_4$$

$$L_4 = -\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 + \frac{1}{2}\mu_3 - \frac{1}{2}\mu_4$$

1. *Bonferroni Linear Combination Tests*, $L = \sum_{i=1}^r c_i \mu_i$.

Use Bonferroni's tests to decide if any of the four contrasts are significant at $\alpha = 0.05$.

- (a) *Statement.*

The statement of the tests for the four cases is (circle one):

- i. $H_0 : L_k \geq 0$ versus $H_1 : L_k < 0$, $k = 1, \dots, 4$.
- ii. $H_0 : L_k \leq 0$ versus $H_1 : L_k > 0$, $k = 1, \dots, 4$.
- iii. $H_0 : L_k = 0$ versus $H_1 : L_k \neq 0$, $k = 1, \dots, 4$.

- (b) *Test and Conclusions.*

From SAS, on the one hand, since

$$t^* = \frac{\hat{L}}{s\{\hat{L}\}}$$

the different values of the four test statistics and associated two-sided p-values are given by,

L	$(c_1, c_2, c_3, c_4, c_5)$	\hat{L}	$s\{\hat{L}\}$	$t^* = \frac{\hat{L}}{s\{\hat{L}\}}$	p-value
L_1	$(-\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, 1)$	-4.875	1.2082	-4.03	0.0005
L_2	$(0, 1, -1, 0, 0)$	-2	1.5283	-1.31	0.2025
L_3	$(-1, 0, 0, 1, 0)$	12.167	1.5283	7.96	< 0.0001
L_4	$(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, 0)$	0.5833	1.0806	0.54	0.5941

In other words, the significant contrasts at $\alpha = 0.05$ are (choose none, one or more)

$L_1 / L_2 / L_3 / L_4$

(c) *One Further Comment.*

The *family* level of significance, $1 - \alpha/(2g)$, (choose one) **remains the same / changes** to account for the number g of contrasts/combinations tested. Scheffe, on the other hand, although it does not change to account for the number of contrasts/combinations tested, typically gives a *conservative* (large) family level of significance value to account for an “infinity” of possible tested contrasts/combinations.

2. *Bonferroni Combination Confidence Intervals*, $L = \sum_{i=1}^r c_i \mu_i$.

The confidence intervals are given by

$$\hat{L} \pm t(1 - \alpha/(2g); n_T - r) s \left\{ \hat{L} \right\}$$

where, $t(1 - \alpha/(2g); n_T - r) = t(0.99375, 25) \approx 2.69$

PRGM INVT ENTER 25 ENTER 0.99375 ENTER

and so

L	$(c_1, c_2, c_3, c_4, c_5)$	\hat{L}	$s\{\hat{L}\}$	$\hat{L} \pm t(1 - \alpha/(2g); n_T - r) s \left\{ \hat{L} \right\}$
L_1	$(-\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, 1)$	-4.875	1.2082	(-8.125, -1.625)
L_2	$(0, 1, -1, 0, 0)$	-2	1.5283	(-6.111, 2.111)
L_3	$(-1, 0, 0, 1, 0)$	12.167	1.5283	(8.0557, 16.278)
L_4	$(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, 0)$	0.58333	1.0806	(-2.324, 3.4902)

where the significant factor level mean differences are (choose none, one or more)

$L_1 / L_2 / L_3 / L_4$

17.8 Hohm Simultaneous Testing Procedure

SAS program: att1-17-8-cotton-multiple-holm

We look at confidence intervals and significance for the Holm multiple linear combinations of factor level means is a refinement of the Bonferroni simultaneous inference procedure.

Test The tests are given by

$$H_o : L = 0 \text{ vs } H_a : L \neq 0, L = \sum_{i=1}^r c_i \mu_i$$

with test statistic

$$t^* = \frac{\hat{L}}{s\{\hat{L}\}}$$

and the various null tests are accepted or rejected in the following sequential way,

$$\text{if } P_k < \frac{\alpha}{g - k + 1}, \text{ conclude } H_a \text{ and proceed}$$

$$\text{if } P_k \geq \frac{\alpha}{g - k + 1}, \text{ conclude } H_o \text{ for this and remaining tests}$$

where P_k is the p-value for the k th ranked test and g is number of simultaneous linear combinations in the problem.

Confidence Interval The confidence interval is given by

$$\hat{L} \pm t(1 - \alpha; n_T - r) s\{\hat{L}\}$$

where α is that α is last significant.

where

$$\hat{L} = \sum_{i=1}^r c_i \bar{Y}_i$$

$$s\{\hat{L}\} = \sqrt{MSE \sum_{i=1}^r \frac{c_i^2}{n_i}}$$

and where MSE is the mean squared error from the associated single-factor ANOVA table.

Exercise 17.7 (Holm Linear Combination Comparisons: Fiber Tensile Strength) The tensile strength of fibers with varying amounts of cotton content is investigated.

5% cotton	7	7	15	11	9	10	$\bar{Y}_1 \approx 9.8$
10% cotton	12	17	12	18	18	16	$\bar{Y}_2 \approx 15.5$
15% cotton	14	18	18	19	19	17	$\bar{Y}_3 \approx 17.5$
20% cotton	19	25	22	19	23	24	$\bar{Y}_4 \approx 22$
25% cotton	7	10	11	15	11	14	$\bar{Y}_5 \approx 11.3$

We are interested in determining the simultaneous Holm tests and confidence intervals for the following four ($g = 4$) contrasts/combinations.

$$\begin{aligned} L_1 &= -\frac{1}{4}\mu_1 - \frac{1}{4}\mu_2 - \frac{1}{4}\mu_3 - \frac{1}{4}\mu_4 + \mu_5 \\ L_2 &= \mu_2 - \mu_3 \\ L_3 &= -\mu_1 + \mu_4 \\ L_4 &= -\frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 + \frac{1}{2}\mu_3 - \frac{1}{2}\mu_4 \end{aligned}$$

1. *Holm Linear Combination Tests*, $L = \sum_{i=1}^r c_i \mu_i$, $i > j$.

Use Holm's tests to decide if any of the four contrasts are significant at $\alpha = 0.05$.

- (a) *Statement.*

The statement of the tests for the four cases is (circle one):

- i. $H_0 : L_k \geq 0$ versus $H_1 : L_k < 0$, $k = 1, \dots, 4$.
- ii. $H_0 : L_k \leq 0$ versus $H_1 : L_k > 0$, $k = 1, \dots, 4$.
- iii. $H_0 : L_k = 0$ versus $H_1 : L_k \neq 0$, $k = 1, \dots, 4$.

- (b) *Test and Conclusions.*

From SAS, on the one hand, since

$$t^* = \frac{\hat{L}}{s \left\{ \hat{L} \right\}}$$

the different values of the four test statistics and associated two-sided p-values, ranked smallest to largest, are given by,

rank (k)	L	$(c_1, c_2, c_3, c_4, c_5)$	\hat{L}	$s \left\{ \hat{L} \right\}$	$t^* = \frac{\hat{L}}{s \left\{ \hat{L} \right\}}$	p-value	$\frac{\alpha}{g-k+1} = \frac{0.05}{4-k+1}$	significant?
1	L_3	$(-1, 0, 0, 1, 0)$	12.167	1.5283	7.96	< 0.0001	$\frac{0.05}{4-1+1} = 0.0125$	S
2	L_1	$(-\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, 1)$	-4.875	1.2082	-4.03	0.0005	0.01667	S
3	L_2	$(0, 1, -1, 0, 0)$	-2	1.5283	-1.31	0.2025	0.025	N
4	L_4	$(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, 0)$	0.5833	1.0806	0.54	0.5941	0.05	N

In other words, the significant contrasts are (choose none, one or more)

\mathbf{L}_1 / \mathbf{L}_2 / \mathbf{L}_3 / \mathbf{L}_4

2. *Holm Combination Confidence Intervals*, $L = \sum_{i=1}^r c_i \mu_i$.

The confidence intervals are given by

$$\hat{L} \pm t(1 - g\alpha'; n_T - r) s \left\{ \hat{L} \right\}$$

where, since the *last significant* p-value occurs for $\alpha' = \frac{0.05}{4-2+1} = 0.01667$ and so $t(1 - g\alpha'; n_T - r) = t(1 - 4(0.01667), 25) \approx 1.55$ and so

L	$(c_1, c_2, c_3, c_4, c_5)$	\hat{L}	$s \left\{ \hat{L} \right\}$	$\hat{L} \pm t(1 - g\alpha'; n_T - r)s \left\{ \hat{L} \right\}$
L_3	$(-1, 0, 0, 1, 0)$	12.1667	1.53	$(-6.748, -3.002)$
L_1	$(-\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}, 1)$	-4.875	1.21	$(9.757, 14.535)$

where the confidence intervals for only the (choose one) **nonsignificant** / **significant** factor level mean differences are calculated.

17.9 Analysis of Factor Levels When Factor Quantitative

SAS program: att1-17-9-cotton-multiple-quadregression

When a factor is quantitative, is sometimes worthwhile going beyond a single factor ANOVA procedure and performing a regression.

Exercise 17.8 (Analysis of Factor Levels When Factor Quantitative: Fiber Tensile Strength) The tensile strength of fibers with varying amounts of cotton content is investigated.

5% cotton	7	7	15	11	9	10	$\bar{Y}_1. \approx 9.8$
10% cotton	12	17	12	18	18	16	$\bar{Y}_2. \approx 15.5$
15% cotton	14	18	18	19	19	17	$\bar{Y}_3. \approx 17.5$
20% cotton	19	25	22	19	23	24	$\bar{Y}_4. \approx 22$
25% cotton	7	10	11	15	11	14	$\bar{Y}_5. \approx 11.3$

We are interested in performing a quadratic regression for this data and assessing the results.

1. *Quadratic regression*

From SAS, the quadratic regression is (choose one)

(a) $\hat{Y} = -2.7 + 2.777X - 0.086X^2$

(b) $\hat{Y} = -2.7 - 2.777X - 0.086X^2$

(c) $\hat{Y} = -2.7 + 2.777X + 0.086X^2$

2. *Residuals*

One residual, for observation 1, is $e_1 = -2.02381$ and another is $e_{28} =$ (choose one) **-0.22889** / **1.2234** / **2.17619**

3. *Lack of Fit Test*

preliminary work

REGRESSION	Degrees of Freedom	Sum Of Squares	Mean Squares
Regression	2	444.16	222.08
Error	27	303.20	11.23
Total	29	747.37	

ANOVA	Degrees of Freedom	Sum Of Squares	Mean Squares
Treatment	4	572.2	143.05
Error	25	175.17	7.006
Total	29	747.37	

ANOVA for lack of fit	Degrees of Freedom	Sum Of Squares	Mean Squares
Regression	2	444.16	222.08
Error	27	303.20	11.23
Lack of fit	2	128.03	64.015
Error	25	175.17	7.006
Total	29	747.37	

statement

$$H_0 : E\{Y\} = \beta_0 + \beta_1x + \beta_{11}x^2 \quad \text{versus}$$

$$H_a : E\{Y\} \neq \beta_0 + \beta_1x + \beta_{11}x^2$$

test

$$F \text{ test statistic, } F^* = \frac{\text{MSPE}}{\text{MSLF}} = \frac{64.015}{7.006} =$$

9.04 / 9.14 / 10.11

(choose one)

$$F \text{ critical value, } F(0.95; 2, 25) = 3.39$$

conclusions

since $F^* > F$

reject null; that is, quadratic model does *not* fit data