Grant's Tutoring

BASIC STATISTICS 2

Volume 2 of 2

September 2014 edition

Because the book is so large,
the entire Basic Statistics 2 course
has been split into two volumes.

Learn What You Need to Know
Know What You Need to Learn

While studying this book, why not hear Grant explain it to you? Contact Grant for info about purchasing **Grant's Audio Lectures**. Some concepts make better sense when you hear them explained.

Better still, see Grant explain the key concepts in person. Sign up for **Grant's Weekly Tutoring** or attend **Grant's Exam Prep Seminars.** Text or Grant (204) 489-2884 or go to **www.grantstutoring.com** to find out more about all of Grant's services. **Seminar Dates will be finalized no later than Sep. 25 for first term and Jan. 25 for second term.**

HOW TO USE THIS BOOK

I have broken the course up into lessons. Do note that the numbering of my lessons do not necessarily correspond to the numbering of the units in your course outline. Study each lesson until you can do all of my lecture problems from start to finish without any help. If you are able to solve my Lecture Problems, then you should have nothing to fear about your exams.

Although NOT ESSENTIAL, you may want to purchase the *Multiple-Choice Problems Set for Basic Statistical Analysis II (Stat 2000)* by Dr. Smiley Cheng. This book is now out of print, but copies may be available at The Book Store. The appendices of my book include complete step-by-step solutions for all the problems and exams in Cheng's book. **Be sure to read the "Homework" section at the end of each lesson for important guidance from me on how to proceed in your studying, as there has been changes to the course since the Cheng book was published.**

You also need a good, but not expensive, scientific calculator. Any of the makes and models of calculators I discuss in Appendix A are adequate for this course. **Appendix A in this book shows you how to use all major models of calculators.**

I have presented the course in what I consider to be the most logical order. Although my books are designed to follow the course syllabus, it is possible your prof will teach the course in a different order or omit a topic. It is also possible he/she will introduce a topic I do not cover. **Make sure you are attending your class regularly! Stay current with the material, and be aware of what topics are on your exam. Never forget, it is your prof that decides what will be on the exam, so pay attention.**

If you have any questions or difficulties while studying this book, or if you believe you have found a mistake, do not hesitate to contact me. My phone number and website are noted at the bottom of every page in this book. "Grant's Tutoring" is also in the phone book. **I welcome your input and questions.**

Wishing you much success,

Grant Skene

Owner of Grant's Tutoring and author of this book

Have you signed up for

Grant's Homework Help yet?

No? Then what are you waiting for? IT'S FREE!

Go to *www.grantstutoring.com* right now,

and click the link to sign up for

Grant's Homework Help

IT'S FREE!

- Grant will send you extra study tips and questions of interest throughout the term.
- You are also welcome to contact Grant with any questions you have. Your question may even provide the inspiration for other tips to send.
- If there are any changes in the course work or corrections to this book, you will be the first to know.
- You will also be alerted to upcoming exam prep seminars and other learning aids Grant offers.
- If you sign up, you will also receive a **coupon** towards Grant's services.

And, it is all FREE!

Four ways Grant can help you:

- ## Grant's Study Books

  - **Basic Statistics 1 (Stat 1000)**
  - **Basic Statistics 2 (Stat 2000)**
  - **Linear Algebra and Vector Geometry (Math 1300)**
  - **Matrices for Management (Math 1310)**
  - **Intro Calculus (Math 1500 or Math 1510)**
  - **Calculus for Management (Math 1520)**
  - **Calculus 2 (Math 1700 or 1710)**

All these books are available at **UMSU Digital Copy Centre**, room 118 University Centre, University of Manitoba. **Grant's books can be purchased there all year round. You can also order a book from Grant directly.** Please allow one business day because the books are made-to-order.

- ## Grant's One-Day Exam Prep Seminars

These are one-day, 12-hour marathons designed to explain and review all the key concepts in preparation for an upcoming midterm or final exam. Don't delay! Go to www.grantstutoring.com right now to see the date of the next seminar. A seminar is generally held one or two weeks before the exam, but don't risk missing it just because you didn't check the date well in advance. You can also reserve your place at the seminar online. You are not obligated to attend if you reserve a place. You only pay for the seminar if and when you arrive.

- ## Grant's Weekly Tutoring Groups

This is for the student who wants extra motivation and help keeping on top of things throughout the course. Again, go to www.grantstutoring.com for more details on when the groups are and how they work.

- ## Grant's Audio Lectures

For less than the cost of 2 hours of one-on-one tutoring, you can listen to over 40 hours of Grant teaching this book. Hear Grant work through examples, and offer that extra bit of explanation beyond the written word. Go to www.grantstutoring.com for more details.

Table of Contents

TABLE OF CONTENTS (CONTINUED)

APPENDICES

FORMULA SHEET

A formula sheet is included in your exams. Check your course syllabus and compare it to the formula sheet I use below in case the formula sheet in your course has changed.

1. $SE(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ with $\text{df} = \dfrac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$

2. $SE(\bar{x}_1 - \bar{x}_2) = s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ with $\text{df} = n_1 + n_2 - 2$ if $\sigma_1^2 = \sigma_2^2$

where $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$

3. $SSG = \sum_{i=1}^{k} n_i \left(\bar{X}_i - \bar{\bar{X}}\right)^2$

4. Poisson Distribution: $P(X = k) = \dfrac{e^{-\lambda}\lambda^k}{k!}$ $k = 0, 1, 2, \ldots$

5. $t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$

6. $SE_{b_1} = \dfrac{s_e}{\sqrt{\sum (x_i - \bar{x})^2}}$, $s_e = \sqrt{MSE}$

7. $SE_{b_0} = s_e\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}}$

8. $SE_{\hat{\mu}} = s_e\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$

9. $SE_{\hat{y}} = s_e\sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$

10. $SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ if $p_1 = p_2$ where $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$

11. $SE(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ if $p_1 \neq p_2$

STEPS FOR TESTING A HYPOTHESIS

Step 1. State the null and alternative hypotheses (H_0 and H_a), and so determine if the test is 2-tailed, upper-tailed, or lower-tailed.

Step 2. Use the given α (always use α = 5% if none is given) to get the **critical value** (z^*, t^*, F^*, etc. depending on the hypothesis you are testing) from the appropriate table and state the **rejection region**.

Step 3. Compute the **test statistic** (z, t, F, etc. depending on the hypothesis you are testing) using the appropriate formula, and see if it lies in the rejection region.

Step 4. (Only if specifically asked to do so.) Compute the ***P*-value**.

Draw a density curve (z-bell curve, t-bell curve, F right-skewed curve, etc. depending on the test statistic you have computed), mark the test statistic (found in Step 3), and shade the area as instructed by H_a. That area is the P-value.

Remember, a P-value is very handy to know if you are asked to make decisions for more than one value of α.

Reject H_0 if P-value < α.

Step 5. State your conclusion.

Either: Reject H_0. There is statistically significant evidence that the alternative hypothesis is correct. (Replace the underlined part with appropriate wording from the problem that says H_a is correct.)

Or: Do not reject H_0. There is no statistically significant evidence that the alternative hypothesis is correct. (Replace the underlined part with appropriate wording from the problem that says we are not convinced that H_a is correct.)

SUMMARY OF KEY FORMULAS IN THIS COURSE

Lesson 1. The mean and standard deviation of $\bar{x}$ are $\mu_{\bar{x}} = \mu$ and $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.

The standard error of $\bar{x} = SE_{\bar{x}} = \frac{s}{\sqrt{n}}$.

Central Limit Theorem: If n is large, $\bar{x}$ is approximately normal.

Confidence Intervals for μ: $\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$ or $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$

Sample size determination: $n = \left(\frac{z^* \sigma}{m}\right)^2$

Lesson 2. Test statistics for H_0: $\mu = \mu_0$ are $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ or $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$.

Lesson 3. Standardizing formula for $\bar{x}$ bell curves: $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$.

To compute $\bar{x}^*$ for $\bar{x}$ decision rules: $\bar{x}^* = z^* \frac{\sigma}{\sqrt{n}} + \mu_0$.

Lesson 4. Properties for means of two random variables:

$$\mu_{X+Y} = \mu_X + \mu_Y$$
$$\mu_{X-Y} = \mu_X - \mu_Y$$

Properties for variance of two independent random variables:

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$$
$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y$$

Properties for variance of two dependent random variables with correlation ρ:

$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y + 2\rho\sigma_X\sigma_Y$$
$$\sigma^2_{X-Y} = \sigma^2_X + \sigma^2_Y - 2\rho\sigma_X\sigma_Y$$

Confidence interval for $\mu_1 - \mu_2$ is $(\bar{x}_1 - \bar{x}_2) \pm t^* SE(\bar{x}_1 - \bar{x}_2)$.

To test H_0: $\mu_1 = \mu_2$, the test statistic is $t = \frac{\bar{x}_1 - \bar{x}_2}{SE(\bar{x}_1 - \bar{x}_2)}$.

The formulas for the degrees of freedom and $SE(\bar{x}_1 - \bar{x}_2)$ are included on the Formula Sheet given on your exams (page 1 of this book).

SUMMARY OF KEY FORMULAS IN THIS COURSE (CONTINUED)

Lesson 5. $DFG = I - 1$ and $DFE = N - I$

$$\bar{\bar{x}} = \frac{\sum n_i \bar{x}_i}{N} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3 + \ldots + n_I\bar{x}_I}{N}$$

$MSG = \dfrac{SSG}{DFG}$ (The formula for SSG is included on the Formula Sheet given on your exams (page 1 of this book).)

$$MSE = \frac{SSE}{DFE} = \frac{\sum (n_i - 1)s_i^2}{N - I} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots + (n_I - 1)s_I^2}{N - I}$$

$s_p^2 = MSE$. (The Formula Sheet given on your exams (page 1 of this book) gives you the two-sample version of the s_p^2 formula; this formula can be generalized for three, four or more samples.)

The formula for the F test statistic is $F = \dfrac{MSG}{MSE}$ with df = DFG, DFE.

The coefficient of determination = $R^2 = \dfrac{SSG}{SST}$.

Lesson 6. If X is a discrete random variable:

$$\mu = \sum x\,p(x) \text{ and } \sigma^2 = \sum (x - \mu)^2 p(x) \text{ or } \sigma^2 = \left(\sum x^2 p(x)\right) - \mu^2$$

The Complement Rule: $P(A^C) = 1 - P(A)$

The General Addition Rule: $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

"Neither/nor" is 1 minus "or": $P(\text{neither } A \text{ nor } B) = 1 - P(A \text{ or } B)$

The General Multiplication Rule: $P(A \text{ and } B) = P(A) \times P(B|A)$

If A and B are disjoint, then $P(A \text{ and } B) = 0$.

If (and only if) A and B are independent, then $P(A \text{ and } B) = P(A) \times P(B)$.

Conditional Probability: $$P(B|A) = \frac{P(B \text{ and } A)}{P(A)}$$

SUMMARY OF KEY FORMULAS IN THIS COURSE (CONTINUED)

Lesson 7. If X is a discrete random variable:

$$\mu = \sum x\, p(x) \text{ and } \sigma^2 = \sum (x-\mu)^2 p(x) \text{ or } \sigma^2 = \left(\sum x^2 p(x)\right) - \mu^2$$

For a **binomial distribution** $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$. (This formula does not really have to be memorized as it is included at the top of Table C which you will be given on exams.)

For a binomial distribution with parameters n and p:

$$\boldsymbol{\mu_X = np} \text{ and } \boldsymbol{\sigma_X = \sqrt{np(1-p)}}.$$

If we are using the normal approximation to the binomial distribution, we can standardize the random variable X into z scores using $z = \frac{x-\mu}{\sigma}$.

For a **Poisson distribution** $P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}$. (This formula does not really have to be memorized since it is included on your formula sheet (#4).

The Poisson distribution has only one parameter, λ:

$$\boldsymbol{\mu_X = \lambda} \text{ and } \boldsymbol{\sigma_X^2 = \lambda}, \text{ so } \boldsymbol{\sigma_X = \sqrt{\lambda}}.$$

Lesson 8. To compute the sample proportion, $\hat{p} = \frac{x}{n}$.

The mean and standard deviation of $\hat{p}$ are $\mu_{\hat{p}} = p$ and $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$.

Confidence Interval for p : $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

Test statistic for H_0: $p = p_0$ is: $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$.

Sample size determination when estimating p: $n = \left(\frac{z^*}{m}\right)^2 p^*(1-p^*)$.

Confidence Interval for $p_1 - p_2$: $(\hat{p}_1 - \hat{p}_2) \pm z^* SE(\hat{p}_1 - \hat{p}_2)$.

Test statistic for H_0: $p_1 = p_2$ is: $z = \frac{\hat{p}_1 - \hat{p}_2}{SE(\hat{p}_1 - \hat{p}_2)}$.

The formulas for the standard error of $\hat{p}_1 - \hat{p}_2$, $SE(\hat{p}_1 - \hat{p}_2)$, are given on the Formula Sheet given on your exams (page 1 of my book).

SUMMARY OF KEY FORMULAS IN THIS COURSE (CONTINUED)

Lesson 9. In a two-way table df = $(r - 1) \times (c - 1)$.

In goodness-of-fit, df = $k - 1$ – 1 more for each estimated parameter.

In a two –way table:

$$\text{Expected count for a cell} = \frac{(\text{The cell's Row Total}) \times (\text{The cell's Column Total})}{\text{The Grand Total}}$$

In a goodness-of-fit test:

Expected count is found using the given probability distribution

Expected count = each probability × the total of the observed counts

To compute each cell's chi-square: $\chi^2_{\text{cell}} = \dfrac{(\text{Observed count} - \text{Expected count})^2}{\text{Expected count}}$

The chi-square test statistic is: $\chi^2 = \sum \chi^2_{\text{cell}}$

Lesson 10. slope $= b = r\dfrac{s_y}{s_x}$ intercept $= a = \bar{y} - b\bar{x}$

The least-squares regression equation is $\hat{y} = a + bx$.

a residual $= e = y - \hat{y}$

The variance of the residuals is $s_e^2 = MSE = \dfrac{\sum(\text{residuals})^2}{n-2}$.

The model of simple linear regression is $y_i = \alpha + \beta x_i + \varepsilon_i$.

We also know $\mu_y = \alpha + \beta x$.

SUMMARY OF KEY FORMULAS IN THIS COURSE (CONTINUED)

Lesson 11. t has df = $n - 2$ in simple linear regression; df = $n - k$ in multiple linear regression where k = the number of variables including y the response variable.

Confidence Intervals:

For the intercept: $a \pm t^* SE_a$

For the slope: $b \pm t^* SE_b$

For an individual prediction at x^*: $\hat{y} \pm t^* SE_{\hat{y}}$

For the mean value of y at x^*: $\hat{y} \pm t^* SE_{\hat{\mu}}$

Test statistics:

For zero correlation: $t = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ (given on the formula sheet)

For zero slope: $F = \dfrac{MSM}{MSE}$

For zero slope: $t = \dfrac{b}{SE_b}$

For zero intercept: $t = \dfrac{a}{SE_a}$

The model of multiple linear regression is $y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi} + \varepsilon_i$.

The coefficient of determination is $r^2 = \dfrac{SSM}{SST}$.

Four ways Grant can help you:

- ## Grant's Study Books

  - **Basic Statistics 1 (Stat 1000)**
  - **Basic Statistics 2 (Stat 2000)**
  - **Linear Algebra and Vector Geometry (Math 1300)**
  - **Matrices for Management (Math 1310)**
  - **Intro Calculus (Math 1500 or Math 1510)**
  - **Calculus for Management (Math 1520)**
  - **Calculus 2 (Math 1700 or 1710)**

  All these books are available at **UMSU Digital Copy Centre**, room 118 University Centre, University of Manitoba. **Grant's books can be purchased there all year round. You can also order a book from Grant directly.** Please allow one business day because the books are made-to-order.

- ## Grant's One-Day Exam Prep Seminars

  These are one-day, 12-hour marathons designed to explain and review all the key concepts in preparation for an upcoming midterm or final exam. Don't delay! Go to www.grantstutoring.com right now to see the date of the next seminar. A seminar is generally held one or two weeks before the exam, but don't risk missing it just because you didn't check the date well in advance. You can also reserve your place at the seminar online. You are not obligated to attend if you reserve a place. You only pay for the seminar if and when you arrive.

- ## Grant's Weekly Tutoring Groups

  This is for the student who wants extra motivation and help keeping on top of things throughout the course. Again, go to www.grantstutoring.com for more details on when the groups are and how they work.

- ## Grant's Audio Lectures

  For less than the cost of 2 hours of one-on-one tutoring, you can listen to over 40 hours of Grant teaching this book. Hear Grant work through examples, and offer that extra bit of explanation beyond the written word. Go to www.grantstutoring.com for more details.