MAT 167: Statistics

Midterm Exam

Instructor: Anthony Tanbakuchi

Summer 2008

Name: _____

Computer / Seat Number: _____

No books, notes, or friends. **Show your work.** You may use the attached
equation sheet, R, and a calculator. No other materials. If you choose to use R,
write what you typed on the test or copy and paste your work into a word
document labeling the question number it corresponds to. When you are done
with the test print out the document. Be sure to save often on a memory stick just
in case. Using any other program or having any other documents open on the
computer will constitute cheating.

You have until the end of class to finish the exam, manage your time wisely.
If something is unclear quietly come up and ask me.
If the question is legitimate I will inform the whole class.

Express all final answers to 3 significant digits. Probabilities should be given as a
decimal number unless a percent is requested. Circle final answers, ambiguous or
multiple answers will not be accepted. Show steps where appropriate.

The exam consists of 21 questions for a total of 66 points on 10 pages.

This Exam is being given under the guidelines of our institution's
Code of Academic Ethics. You are expected to respect those guidelines.

Points Earned: _____ **out of 66 total points**

Exam Score: _____

1. The following is a partial list of statistical methods that we have discussed:

   1. mean
   2. median
   3. mode
   4. standard deviation
   5. z-score
   6. percentile
   7. coefficient of variation

   8. scatter plot
   9. histogram
   10. pareto chart
   11. box plot
   12. normal-quantile plot
   13. confidence interval for a mean
   14. confidence interval for a proportion

   For each situation below, which method is most applicable? **To get full points, give a short (1-3 sentence) description as to why your chosen method is appropriate!**
   - If it's a graphical method, **also describe what you would be looking for**.
   - If it's a statistic, how susceptible to outliers is it?

   (a) (2 points) A school board thinks that students might do better in morning math classes. They would like to compare the distribution of student math scores for 4 time categories: early morning classes, mid morning classes, early afternoon classes, and late afternoon classes.

   (b) (2 points) A clothing manufacturer wants to determine how much variability there is in toddler waist sizes.

   (c) (2 points) The Tucson realtors association is publishing their semi-annual newsletter and has an article discussing current housing prices. They would like to publish a single number

Instructor: Anthony Tanbakuchi   Points earned: _____ / 6 points

that is representative of the typical house price in Tucson.

    (d) (2 points) A political science researcher at the U of A wants to estimate the true percentage of Tucson residents who support Obama from a random survey of 1,000 people.

2. (1 point) Make a sketch of a normal distribution that has been positively skewed.

3. (1 point) What is the area under a positively skewed normal distribution?

4. (1 point) If the mean, median, and mode for a data set are not the same, what can you conclude about the data's distribution?

5. (2 points) Give an example of sampling error.

6. In regards to $\bar{x}$ and the Central Limit Theorem:
   (a) (2 points) What are the two conditions under which the CLT applies?

   (b) (2 points) If the conditions are met, what type of distribution will $\bar{x}$ have?

7. (2 points) Under what conditions can we approximate a binomial distribution as a normal distribution?

8. (2 points) Which distribution (normal, binomial, both, or neither) would be appropriate for describing:

   The sampling distribution of the mean for a random sample from a uniformly distributed population using a sample size of 10.

9. (2 points) Give **a clear specific example** of when you would use a population distribution.

10. (2 points) Give an example of when you would use a sampling distribution.

11. Super Fruity Chew candy has a mean weight of 500 g and a standard deviation of 10 g.

   (a) (2 points) Construct an interval using the Empirical Rule which you would expect 95% of
       the weights to fall within.

   (b) (2 points) Would you consider a candy with weight of 468 g unusual? (State why.)

12. Given the following frequency table summarizing data from a study:

age.years	frequency
0-9	5.00
10-19	8.00
20-29	12.00
30-39	2.00

   (a) (2 points) Construct a relative frequency table.

   (b) (1 point) What is the probability of randomly selecting someone from the study who is
       between 20-39 years old?

(c) (2 points) Calculate the approximate mean age of the subjects in the study from the data given.

13. Results from a randomized experiment to compare crop yields (as measured by dried weight of plants in grams) obtained under a control and two different treatment conditions are shown with a box plot of the data. The researcher who has developed the two new treatments hopes that at least one increases crop yield as compared to the control group.
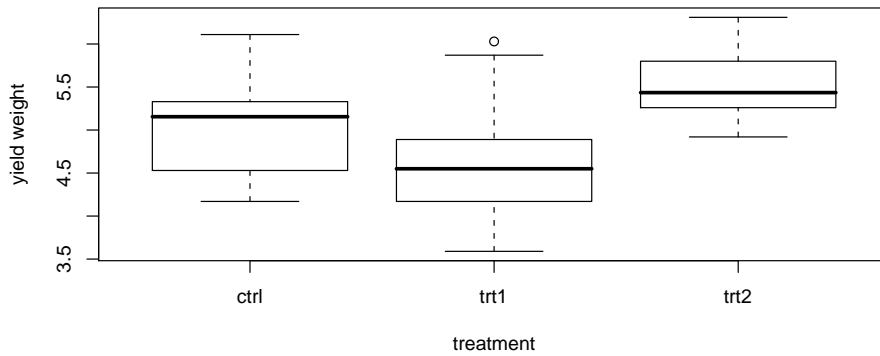
(a) (1 point) Did any groups contain an outlier? If so which ones?

(b) (1 point) What group had the least variability?

(c) (1 point) What was the approximate IQR for the control group (ctrl)?

(d) (1 point) Which group had the highest median crop yield? What was the median value?

14. (2 points) You would like to conduct a study to estimate (at the 90% confidence level) the mean weight of brown bears with a margin of error of 5 lbs. A preliminary study indicates that bear weights are normally distributed with a standard deviation of 22 lbs, what sample size should you use for this study?

15. (2 points) A random sample of 5 men was conducted to determine the mean resting heart rate. Below is the study data in beats per minute.

$$68.9,\ 71.9,\ 61.2,\ 77.9,\ 74.2$$

Construct a 99% confidence interval for the true population mean resting heart rate for men using the above data. (**Assume σ is unknown and the population is normally distributed.**)

16. A bag of M&M's contains 18 red, 12 blue, 8 green, and 7 brown candies.

    (a) (2 points) What is the probability of randomly selecting a red or brown M&M?

(b) (2 points) If 8 M&M's are randomly selected with replacement, what is the probability of getting exactly 6 red M&M's?

(c) (1 point) Would it be unusual to observe 6 red out of 8 randomly selected M&M's? (Why)

17. Answer the following question for a couple who had 4 children.

   (a) (1 point) Is the probability of a girl considered independent or dependent for each successive birth?

   (b) (2 points) What is the probability that the couple has at least 1 girl?

   (c) (2 points) What is the probability that all 4 children are born on different days of the year?

18. (2 points) With one method of a procedure called acceptance sampling, a sample of items is randomly selected without replacement and the entire batch is accepted if every item in the sample is okay. The Niko Electronics Company has just manufactured 10,000 CDs, and 500 are defective. If 10 of the CDs are randomly selected for testing without replacement, what is the probability that the entire batch will be accepted?

19. Engineers must consider the breadths of male heads when designing motorcycle helmets. Men have head breadths that are normally distributed with a mean of 6.0 in and a standard deviation of 1.0 in (based on anthropometric survey data from Gordon, Churchill, et al.).

    (a) (2 points) If 1 man is randomly selected, find the probability that his head breadth is between 6.1 and 6.6 inches.

    (b) (2 points) If 1 man is randomly selected, find the probability that his head breadth is less than 6.1 in.

    (c) (2 points) If 30 men are randomly selected, find the probability that their mean head

breadth is less than 6.1 in.

20. A craps table at a local casino has been losing more money than normal. It seems that bets involving a one on the face of the dice (such as "snake eyes") are appearing more than usual. The casino manager thinks that the dice have been weighted to cause the side with one to have a higher probability of occurring than a fair dice.

   (a) (2 points) The casino manager takes one of the dice from the table and flips it 100 times, the side with a value of one appears 22 times. Construct a 95% confidence interval for the true probability of getting a one with this die.

   (b) (1 point) If the die is fair, what should the true probability be for getting a one?

   (c) (1 point) Based on the casino manager's experiment, does the die appear to be unfair? Why?

21. (2 points) Given $x = \{4c, 2c, -2c\}$, where c is a constant, completely simplify the following expression:

$$\sqrt{\frac{\sum(x_i^2 - 2c)}{6c}}$$

****

End of exam. Reference sheets follow.

Instructor: Anthony Tanbakuchi   Points earned: _____ / 2 points

Basic Statistics: Quick Reference & R Commands

by Anthony Tanbakuchi. Version 1.7
http://www.tanbakuchi.com
ANTHONY@TANBAKUCHI.COM
Get R at: http://r-project.org
R commands: **bold typewriter text**

1 Misc R

To make a vector / store data: $\texttt{x=c(x1, x2, ...)}$
Get help on function: $\texttt{?functionName}$
Get column of data from table:
$\texttt{tableName\$columnName}$
List all variables: $\texttt{ls()}$
Delete all variables: $\texttt{rm(list=ls())}$

$$\sqrt{x} = \texttt{sqrt(x)} \tag{1}$$
$$x^n = \texttt{x^n} \tag{2}$$
$$n = \texttt{length(x)} \tag{3}$$
$$T = \texttt{table(x)} \tag{4}$$

2 Descriptive Statistics

2.1 NUMERICAL

Let $\texttt{x=c(x1, x2, x3, ...)}$

$$\text{total} = \sum_{i=1}^{n} x_i = \texttt{sum(x)} \tag{5}$$
$$\min = \texttt{min(x)} \tag{6}$$
$$\max = \texttt{max(x)} \tag{7}$$
six number summary : $\texttt{summary(x)}$ \qquad (8)
$$\mu = \frac{\sum x_i}{N} = \texttt{mean(x)} \tag{9}$$
$$\bar{x} = \frac{\sum x_i}{n} = \texttt{mean(x)} \tag{10}$$
$$\tilde{x} = P_{50} = \texttt{median(x)} \tag{11}$$
$$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N}} \tag{12}$$
$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = \texttt{sd(x)} \tag{13}$$
$$CV = \frac{\sigma}{\mu} = \frac{s}{\bar{x}} \tag{14}$$

2.2 RELATIVE STANDING

$$z = \frac{x - \mu}{\sigma} = \frac{x - \bar{x}}{s} \tag{15}$$

Percentiles:

$$P_k = x_i, \quad (\text{sorted } x)$$
$$k = \frac{i - 0.5}{n} \cdot 100\% \tag{16}$$

To find x_i given P_k, i is:
1. $L = (k/100\%)n$
2. if L is an integer: $i = L + 0.5$; otherwise i=L and round up.

2.3 VISUAL

All plots have optional arguments:
- $\texttt{main=""}$ sets title
- $\texttt{xlab="", ylab=""}$ sets x/y-axis label
- $\texttt{type="p"}$ for point plot
- $\texttt{type="l"}$ for line plot
- $\texttt{type="b"}$ for both points and lines

Ex: plot(x, y, type="b", main="My Plot")
Plot Types:
$\texttt{hist(x)}$ histogram
$\texttt{stem(x)}$ stem & leaf
$\texttt{boxplot(x)}$ box plot
$\texttt{plot(T)}$ bar plot, T=table(x)
$\texttt{plot(x,y)}$ scatter plot, x, y are ordered vectors
$\texttt{plot(t,y)}$ time series plot, t, y are ordered vectors
$\texttt{curve(expr, xmin,xmax)}$ plot expr involving x

2.4 ASSESSING NORMALITY

Q-Q plot: $\texttt{qqnorm(x); qqline(x)}$

3 Probability

Number of successes x with n possible outcomes.
(Don't double count!)

$$P(A) = \frac{x_A}{n} \tag{17}$$
$$P(\bar{A}) = 1 - P(A) \tag{18}$$
$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) \tag{19}$$
$$P(A \text{ or } B) = P(A) + P(B) \quad \text{if } A,B \text{ mut. excl.} \tag{20}$$
$$P(A \text{ and } B) = P(A) \cdot P(B|A) \tag{21}$$
$$P(A \text{ and } B) = P(A) \cdot P(B) \quad \text{if } A,B \text{ independent} \tag{22}$$
$$n! = n \cdot (n-1) \cdots 1 = \texttt{factorial(n)} \tag{23}$$
$$\,_nP_k = \frac{n!}{(n-k)!} \quad \text{Perm. no rep. alike} \tag{24}$$
$$= \frac{n!}{n_1! n_2! \cdots n_k!} \quad \text{Perm. } n_1 \text{ alike, } \dots \tag{25}$$
$$\,_nC_k = \frac{n!}{(n-k)! k!} = \texttt{choose(n,k)} \tag{26}$$

4 Discrete Random Variables

$$P(x_i) : \text{probability distribution} \tag{27}$$
$$E = \mu = \sum x_i \cdot P(x_i) \tag{28}$$
$$\sigma = \sqrt{\sum (x_i - \mu)^2 \cdot P(x_i)} \tag{29}$$

4.1 BINOMIAL DISTRIBUTION

$$\mu = n \cdot p \tag{30}$$
$$\sigma = \sqrt{n \cdot p \cdot q} \tag{31}$$
$$P(x) = \,_nC_x p^x q^{(n-x)} = \texttt{dbinom(x, n, p)} \tag{32}$$

4.2 POISSON DISTRIBUTION

$$P(x) = \frac{\mu^x \cdot e^{-\mu}}{x!} = \texttt{dpois(x, }\mu\texttt{)} \tag{33}$$

5 Continuous random variables

CDF $F(x)$ gives area to the left of x, $F^{-1}(p)$ expects p is area to the left.

$$f(x) : \text{probability density} \tag{34}$$
$$E = \mu = \int_{-\infty}^{\infty} x \cdot f(x) \, dx \tag{35}$$
$$\sigma = \sqrt{\int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) \, dx} \tag{36}$$
$$F(x) : \text{cumulative prob. density (CDF)} \tag{37}$$
$$F^{-1}(x) : \text{inv. cumulative prob. density} \tag{38}$$
$$F(x) = \int_{-\infty}^{x} f(x') \, dx' \tag{39}$$
$$p = P(x < x') = F(x') \tag{40}$$
$$x' = F^{-1}(p) \tag{41}$$
$$p = P(x > a) = 1 - F(a) \tag{42}$$
$$p = P(a < x < b) = F(b) - F(a) \tag{43}$$

5.1 UNIFORM DISTRIBUTION

$$p = P(x < u') = F(u') $$
$$= \texttt{punif(u', min=0, max=1)} \tag{44}$$
$$u' = F^{-1}(p) = \texttt{qunif(p, min=0, max=1)} \tag{45}$$

5.2 NORMAL DISTRIBUTION

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} \tag{46}$$
$$p = P(z < z') = F(z') = \texttt{pnorm(z')} \tag{47}$$
$$z' = F^{-1}(p) = \texttt{qnorm(p)} \tag{48}$$
$$p = P(x < x') = F(x')$$
$$= \texttt{pnorm(x', mean=}\mu\texttt{, sd=}\sigma\texttt{)} \tag{49}$$
$$x' = F^{-1}(p)$$
$$= \texttt{qnorm(p, mean=}\mu\texttt{, sd=}\sigma\texttt{)} \tag{50}$$

5.3 t-DISTRIBUTION

$$p = P(t < t') = F(t') = \texttt{pt(t', df)} \tag{51}$$
$$t' = F^{-1}(p) = \texttt{qt(p, df)} \tag{52}$$

5.4 χ^2-DISTRIBUTION

$$p = P(\chi^2 < \chi^{2'}) = F(\chi^{2'})$$
$$= \texttt{pchisq(}X^{2'}\texttt{, df)} \tag{53}$$
$$\chi^{2'} = F^{-1}(p) = \texttt{qchisq(p, df)} \tag{54}$$

5.5 F-DISTRIBUTION

$$p = P(F < F') = F(F')$$
$$= \texttt{pf(F', df1, df2)} \tag{55}$$
$$F' = F^{-1}(p) = \texttt{qf(p, df1, df2)} \tag{56}$$

6 Sampling distributions

$$\mu_{\bar{x}} = \mu \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \tag{57}$$
$$\mu_{\hat{p}} = p \qquad \sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} \tag{58}$$

7 Estimation

7.1 CONFIDENCE INTERVALS

proportion: $\hat{p} \pm E$, $\quad E = z_{\alpha/2} \cdot \sigma_{\hat{p}}$ \hfill (59)
mean (σ known): $\bar{x} \pm E$, $\quad E = z_{\alpha/2} \cdot \sigma_{\bar{x}}$ \hfill (60)
mean (σ unknown, use s): $\bar{x} \pm E$, $\quad E = t_{\alpha/2} \cdot \sigma_{\bar{x}}$, \hfill (61)
$df = n - 1$

variance: $\dfrac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \dfrac{(n-1)s^2}{\chi_L^2}$, \hfill (62)
$df = n - 1$

2 proportions: $\Delta\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$ \hfill (63)

2 means (indep): $\Delta\bar{x} \pm t_{\alpha/2} \cdot \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$, \hfill (64)
$df \approx \min(n_1 - 1, n_2 - 1)$

matched pairs: $\bar{d} \pm t_{\alpha/2} \cdot \dfrac{s_d}{\sqrt{n}}$, $\quad d_i = x_i - y_i$, \hfill (65)
$df = n - 1$

7.2 CI CRITICAL VALUES (TWO SIDED)

$$z_{\alpha/2} = F^{-1}(1 - \alpha/2) = \texttt{qnorm(1-alpha/2)} \tag{66}$$
$$t_{\alpha/2} = F^{-1}(1 - \alpha/2) = \texttt{qt(1-alpha/2, df)} \tag{67}$$
$$\chi_L^2 = F^{-1}(\alpha/2) = \texttt{qchisq(alpha/2, df)} \tag{68}$$
$$\chi_R^2 = F^{-1}(1 - \alpha/2) = \texttt{qchisq(1-alpha/2, df)} \tag{69}$$

7.3 REQUIRED SAMPLE SIZE

proportion: $n = \hat{p}\hat{q}\left(\dfrac{z_{\alpha/2}}{E}\right)^2$, \hfill (70)
($\hat{p} = 0.5$ if unknown)

mean: $n = \left(\dfrac{z_{\alpha/2} \cdot \sigma}{E}\right)^2$ \hfill (71)

8 Hypothesis Tests

Test statistic and R function (when available) are listed for each.

Optional arguments for **hypothesis tests**:
`alternative="two.sided"` can be:
  `"two.sided", "less", "greater"`
`conf.level=0.95` constructs a 95% confidence interval. Standard CI only when `alternative="two.sided"`.

Optional arguments for **power calculations & Type II error**:
`alternative="two.sided"` can be:
  `"two.sided" or "one.sided"`
`sig.level=0.05` sets the significance level α.

8.1 1-SAMPLE PROPORTION

$H_0 : p = p_0$
`prop.test(x, n, p=p0, alternative="two.sided")`

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0/n}} \tag{72}$$

8.2 1-SAMPLE MEAN (σ KNOWN)

$H_0 : \mu = \mu_0$

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \tag{73}$$

8.3 1-SAMPLE MEAN (σ UNKNOWN)

$H_0 : \mu = \mu_0$
`t.test(x, mu=μ0, alternative="two.sided")`
Where **x** is a vector of sample data.

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \quad df = n-1 \tag{74}$$

Required Sample size:
`power.t.test(delta=h, sd =σ, sig.level=α, power=1 −
β, type ="one.sample", alternative="two.sided")`

8.4 2-SAMPLE PROPORTION TEST

$H_0 : p_1 = p_2$ or equivalently $H_0 : \Delta p = 0$
`prop.test(x, n, alternative="two.sided")`
where: `x=c(`x_1`, `x_2`)` and `n=c(`n_1`, `n_2`)`

$$z = \frac{\Delta\hat{p} - \Delta p_0}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} \quad \Delta\hat{p} = \hat{p}_1 - \hat{p}_2 \tag{75}$$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}, \quad \bar{q} = 1 - \bar{p} \tag{76}$$

Required Sample size:
`power.prop.test(p1=`p_1`, p2=`p_2`, power=1 − β,
sig.level=α, alternative="two.sided")`

8.5 2-SAMPLE MEAN TEST

$H_0 : \mu_1 = \mu_2$ or equivalently $H_0 : \Delta\mu = 0$
`t.test(x1, x2, alternative="two.sided")`
where **x1** and **x2** are vectors of sample 1 and sample 2 data.

$$t = \frac{\Delta\bar{x} - \Delta\mu_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad df \approx \min(n_1 - 1, n_2 - 1), \quad \Delta\bar{x} = \bar{x}_1 - \bar{x}_2 \tag{77}$$

Required Sample size:
`power.t.test(delta=h, sd =σ, sig.level=α, power=1 −
β, type ="two.sample", alternative="two.sided")`

8.6 2-SAMPLE MATCHED PAIRS TEST

$H_0 : \mu_d = 0$
`t.test(x, y, paired=TRUE, alternative="two.sided")`
where: **x** and **y** are vectors of sample 1 and sample 2 data.

$$t = \frac{\bar{d} - \mu_0}{s_d/\sqrt{n}}, \quad d_i = x_i - y_i, \quad df = n-1 \tag{78}$$

Required Sample size:
`power.t.test(delta=h, sd =σ, sig.level=α, power=1 −
β, type ="paired", alternative="two.sided")`

8.7 TEST OF HOMOGENEITY, TEST OF INDEPENDENCE

$H_0 : p_1 = p_2 = \cdots = p_n$ (homogeneity)
$H_0 : X$ and Y are independent (independence)
`chisq.test(D)`
Enter table: `D=table(c1, c2, ...)`, where $c1$, $c2$, ... are column data vectors.
Or generate table: `D=table(x1, x2)`, where $x1$, $x2$ are ordered vectors of raw categorical data.

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}, \quad df = (\text{num rows - 1})(\text{num cols - 1}) \tag{79}$$

$$E_i = \frac{(\text{row total})(\text{column total})}{(\text{grand total})} = np_i \tag{80}$$

For 2×2 contingency tables, you can use the Fisher Exact Test:
`fisher.test(D, alternative= greater)`
(must specify alternative as greater)

9 Linear Regression

9.1 LINEAR CORRELATION

$H_0 : \rho = 0$
`cor.test(x, y)`
where: **x** and **y** are ordered vectors.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y}, \quad t = \frac{r - 0}{\sqrt{\frac{1-r^2}{n-2}}}, \quad df = n-2 \tag{81}$$

9.2 MODELS IN R

MODEL TYPE	EQUATION	R MODEL
linear 1 indep var	$y = b_0 + b_1 x_1$	y~x1
... 0 intercept	$y = 0 + b_1 x_1$	y~0+x1
linear 2 indep vars	$y = b_0 + b_1 x_1 + b_2 x_2$	y~x1+x2
... interaction	$y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2$	y~x1+x2+x1*x2
polynomial	$y = b_0 + b_1 x_1 + b_2 x_1^2$	y~x1+I(x2^2)

9.3 REGRESSION

Simple linear regression steps:
1. Make sure there is a significant linear correlation.
2. `results=lm(y~x)` Linear regression of y on x vectors
3. `results` View the results
4. `plot(x, y)`; `abline(results)` Plot regression line on data
5. `plot(x, results$residuals)` Plot residuals

$$y = b_0 + b_1 x_1 \tag{82}$$

$$b_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \tag{83}$$

$$b_0 = \bar{y} - b_1 \bar{x} \tag{84}$$

9.4 PREDICTION INTERVALS

To predict y when $x = 5$ and show the 95% prediction interval with regression model in results:
`predict(results, newdata=data.frame(x=5),
int="pred")`

10 ANOVA

10.1 ONE WAY ANOVA

1. `results=aov(depVarColName~indepVarColName,
   data=tableName)` Run ANOVA with data in TableName, factor data in indepVarColName column, and response data in depVarColName column.
2. `summary(results)` Summarize results
3. `boxplot(depVarColName~indepVarColName,
   data=tableName)` Boxplot of levels for factor

To find required sample size and power see `power.anova.test(...)`

11 Loading external data

- Export your table as a CSV file (comma separated file) from Excel.
- Import your table into `MyTable` in R using:
  `MyTable=read.csv(file.choose())`