Final Review

(If you catch any errors in the solutions, let me know!)

1. Twenty-year-old men's weights have a population mean of 155 pounds and a population standard deviation of 22 pounds (www.kidsgrowth.com). The distribution is right-skewed. Suppose we take a random sample of 100 twenty-year-old men's weights.[1]

   (a) Explain why the Central Limit Theorem is applicable.

   (b) Sketch the sampling distribution of means. Label the mean, the mean $\pm$ 1, 2, and 3 standard errors, and indicate what percent of the distribution falls in each region.

   (c) What percent of random samples of 100 twenty-year-old men have means between 150.6 and 161.6 pounds?

2. It is believed that nearsightedness affects about 8% of all children. Would it be considered unusual if 21 our of 194 randomly sampled children are nearsighted? Explain your reasoning.

3. A clinical trial with 400 subjects was conducted to test whether the average weight loss using a new diet pill is significant. You know that the standard deviation of the test subject's weight loss was 5lbs. Give an example of sample mean weight loss which would result in rejecting the null hypothesis at the 5% level but not rejecting it at the 1% level. Hint: We only care if the subjects lost weight and not gained it, think about what tails are appropriate and start by writing the appropriate hypotheses.

4. In June 2002, the *Journal of Applied Psychology* reported on a study that examined whether the content of TV shows influenced the ability of viewers to recall brand names of items featured in the commercials. The researchers randomly assigned volunteers to watch one of three programs, each containing the same nine commercials. One of the programs had a violent content, another sexual content, and the third neutral content. After the shows ended, the subjects were asked to recall the brands of products what were advertised. Results are summarized in the table below.[2]

	Program Type		
	Violent	Sexual	Neutral
No. of subjects	101	106	103
Brands recalled			
Mean	3.02	2.72	4.65
SD	1.61	1.85	1.62

   For the following questions, you may assume that all assumptions and conditions necessary for inference are satisfied.
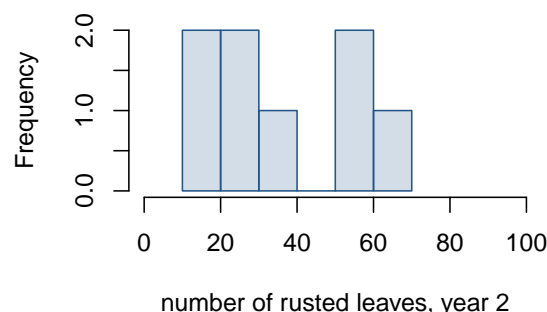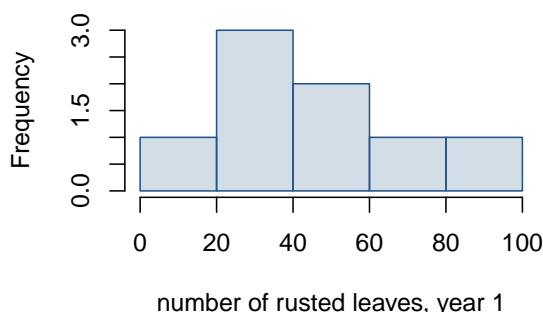
   (a) Is there a significant difference in viewers' abilities to remember brands advertised in shows with **violent** vs. **neutral** content? Use a hypothesis test to answer this question and make sure to interpret your conclusion in context.

   (b) Construct a 95% confidence interval for the difference in number of brand names recalled between the groups watching shows with **sexual** content and those watching **neutral** shows. Interpret your interval in context.

   (c) What type of test would we use if we wanted to compare all three means simultaneously.

---

[1]Adapted from *Introductory Statistics* by Gould and Ryan.
[2]Adapted from *Intro Stats*, by De Veaux, Velleman, and Bock.

5. Cedar-apple rust is a (non-fatal) disease that affects apple trees. Its most obvious symptom is rust-colored spots on apple leaves. Red cedar trees are the immediate source of the fungus that infects the apple trees. If you could remove all red cedar trees within a few miles of the orchard, you should eliminate the problem. In the first year of an experiment the number of affected leaves on 8 randomly sampled trees were counted; the following winter all red cedar trees within 100 yards of the orchard were removed and the following year the same trees were examined for affected leaves. The results are recorded below:[3]

tree	number of rusted leaves: year 1	number of rusted leaves: year 2	difference year 1 - year 2
1	38	32	6
2	10	16	-6
3	84	57	27
4	36	28	8
5	50	55	-5
6	35	12	23
7	73	61	12
8	48	29	19
average	46.8	36.2	10.5
standard dev	23	19	12

number of rusted leaves, year 1

number of rusted leaves, year 2

(a) Write the hypotheses for testing for a difference between the average number of rusted leaves between years 1 and 2.

(b) Check the assumptions and conditions necessary for inference and determine the most appropriate statistical method for evaluating these hypotheses.

(c) Calculate the test statistic and the p-value.

(d) What is the conclusion of the hypothesis test?

(e) What is the p-value for testing if the number of rusted leaves has <u>decreased</u> from year 1 to 2. Give an interpretation of this new p-value.

6. In November 2009, 11% of Americans thought the country representing "greatest danger" to the US was China. In Jan 2011, a survey conducted by the Pew Research Foundation showed that out of the 1,503 randomly sampled Americans 20% agree with that statement. Do these data provide convincing evidence that the proportion of Americans who think China is the greatest danger to the US has increased over the last couple years? Conduct a hypothesis test to answer this question and also interpret the p-value in context of the question.

7. A USA Today/Gallup poll conducted between Dec. 21, 2010 - Jan. 9, 2011 asked a group of

---

[3]Adapted from http://www.physics.csbsju.edu/stats/t-test.html

unemployed and underemployed Americans if they have had major problems in their relationships with their spouse or another close family member as a result of not having a job (if unemployed) or not having a full-time job (if underemployed). 1,145 unemployed and 675 underemployed people were randomly sampled and surveyed. 27% of the unemployed and 25% of the underemployed people said they had major problems in relationships as a result of their employment status.

(a) Are conditions for inference satisfied?

(b) Construct a 95% confidence interval to estimate the difference between the true population proportions of unemployed and underemployed people who had such problems.

(c) Conduct a hypothesis test to evaluate if these data provide convincing evidence to suggest a difference between the true population proportions of unemployed and underemployed people who had such problems. Use a 5% significance level.

8. As recently as 2008, 70% of users of social networking sites such as Facebook were 35 years old or younger. Now the age distribution is much more spread out. The table below shows the age distribution of 975 users of social networking sires from a Pew Research survey from June 2011.[4]

Age	18 - 22	23 - 35	36 - 49	50 - 65	65 +
Frequency	156	312	253	195	59

(a) Test an assumption that users are equally likely to be in each of the five age groups listed. Show all details of the test.

(b) Which age group contributes the largest amount to the test statistic? For the age group, is the observed count smaller or larger than the expected count?

9. "Further observations on the manner of clasping the hands", Downey (1926), found 190 right-thumb and 149 left-thumb-claspers among right-handed women, and 42 right-thumb and 49 left-thumb-claspers among left-handed women.[5]

(a) Make a contingency table summarizing the data.
(b) Use a χ^2 statistics to evaluate if handedness and choice of thumb are related.

10. A random sample of 3,052 cell phone users were asked "Do you send or receive text messages on your cell phone?" Of the 800 teens 696 said they did, and of the 2252 adults 1621 said they did. The histogram below shows the bootstrap distribution for the difference between the proportions of teens and adults who send or receive text messages on their cell phones. The plot below shows the distribution for the difference between the proportions of teens and adults ($\hat{p}_{teen} - \hat{p}_{adult}$) in 10,000 **bootstrap** samples.[6]

[4]From *Unlocking the Power of Data* by Lock, Lock, Lock, Lock, and Lock.
[5]Adapted from *Handbook of Biological Statistics* by McDonald.
[6]Adapted from *Unlocking the Power of Data* by Lock, Lock, Lock, Lock, and Lock.

(a) Estimate the mean of the bootstrap distribution.

(b) Which of the below is the most reasonable estimate of the standard error of the difference in proportions? Explain your reasoning for your choice.

  (a) 0.015   (c) 0.05

  (b) 0.03   (d) 0.15

(c) Using your choice for SE in part (b), find an interval estimate for the difference in proportions of teen and adult cell phone users who send/receive text messages.
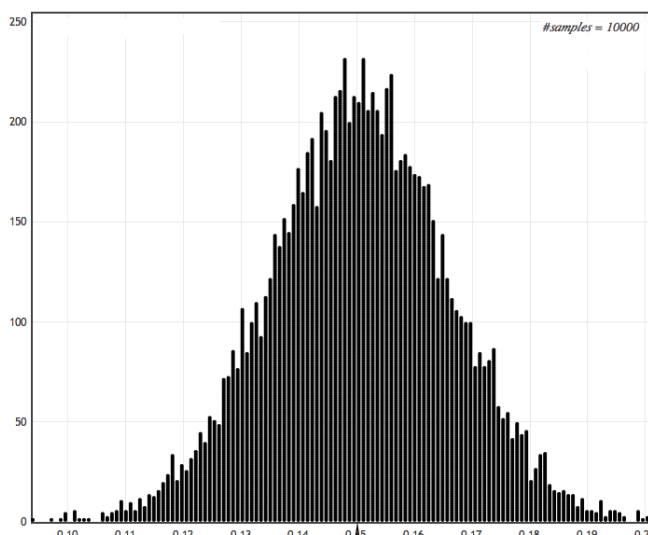
11. In a state where the political race is tight you have the job of conducting a survey on a simple random sample of voters from the state to determine who has the edge. As a pollster, you know that to give your poll credit, you need to ensure the estimate is within 4 percentage points (with 99% confidence). Determine the minimum sample size that will ensure this accuracy.

12. The mean fat content of the 30 Burger King menu items is 23.5g with a standard deviation of 16.4g, and the mean protein content of these items is 17.2g with a standard deviation of 14g. If the correlation between protein and fat contents is 0.83. A scatterplot of the data is shown below.
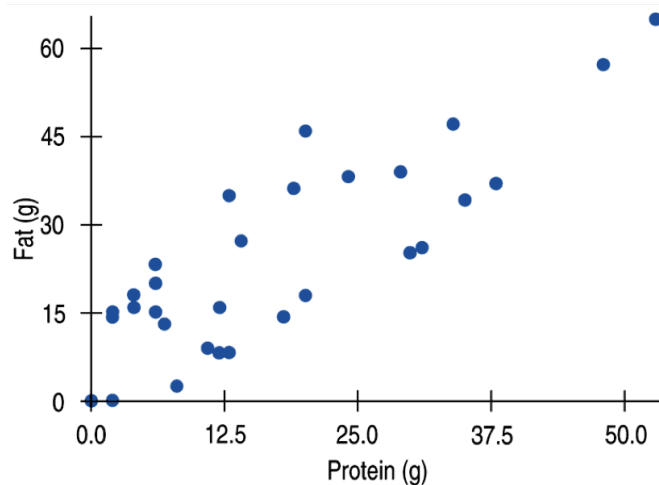

(a) Write the linear model for predicting fat content.

(b) Interpret the slope and the intercept in context.

(c) A new BK menu item has 30 grams of protein. Can we use this model to predict its fat content? If so, what is it?

(d) Another new BK menu item has 75 grams of protein. Can we use this model to predict its fat content? If so, what is it?

(e) Based on the residuals plot shown below, does the linear model appear to be appropriate for these data?
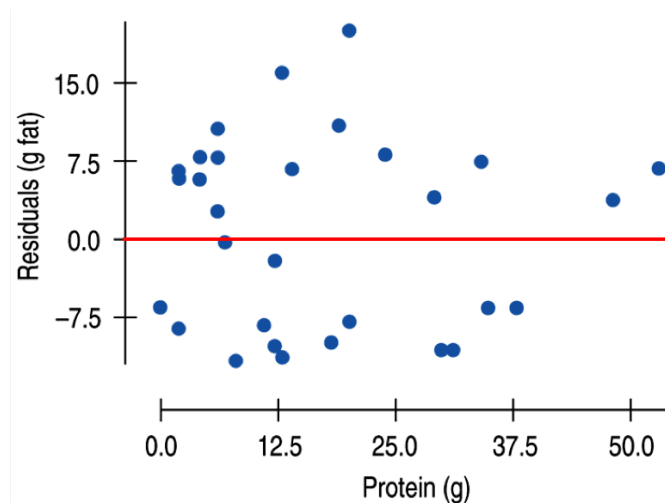
(f) Calculate R^2 and interpret it in context.

13. Data was collected on nave heights (ft) and total lengths (ft) of 25 English medieval cathedrals (excerpt shown below).

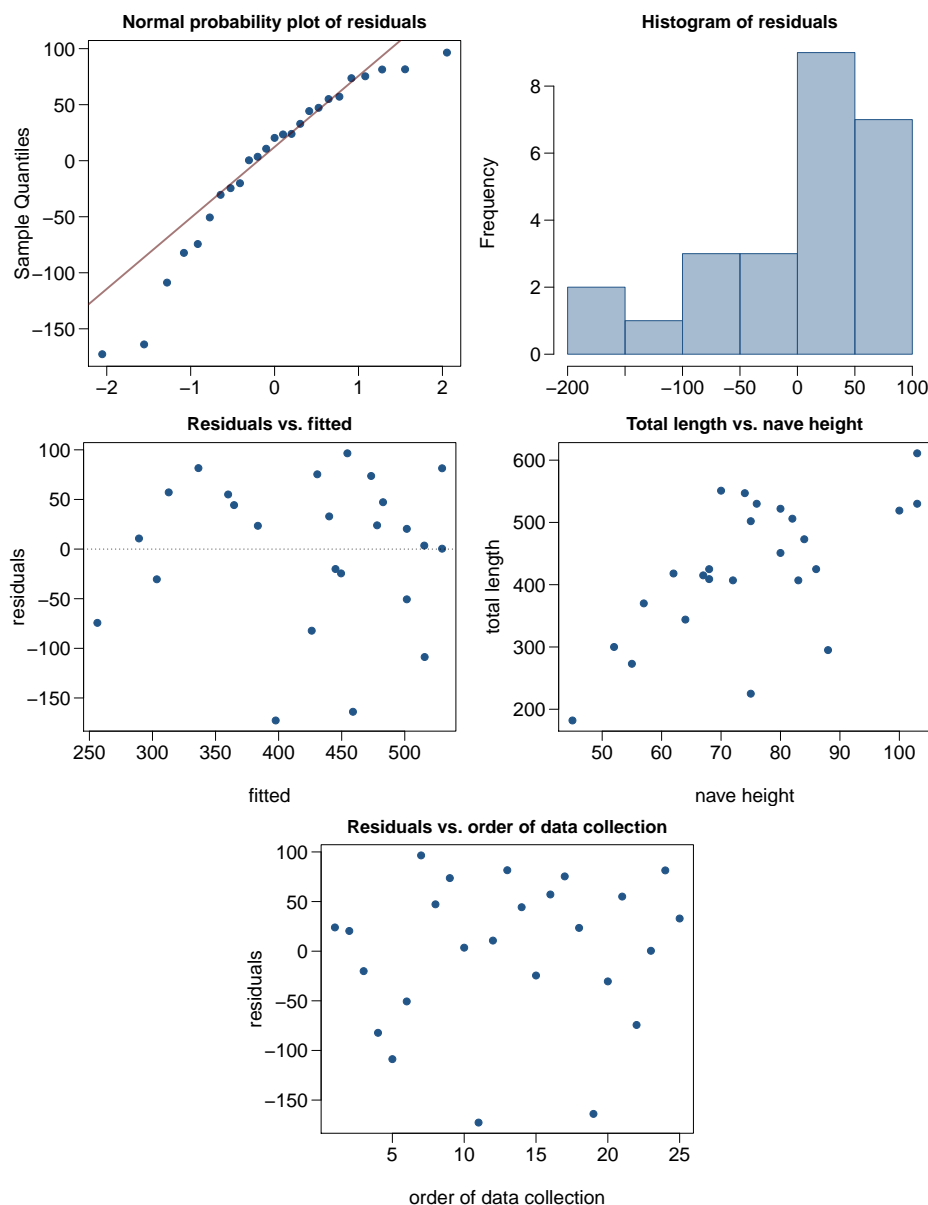| name	style	nave height	total length
Durham	roman	75.00	502.00
Canterbury	roman	80.00	522.00
⋮			
Old St Paul	gothic	103.00	611.00
Salisbury	gothic	84.00	473.00
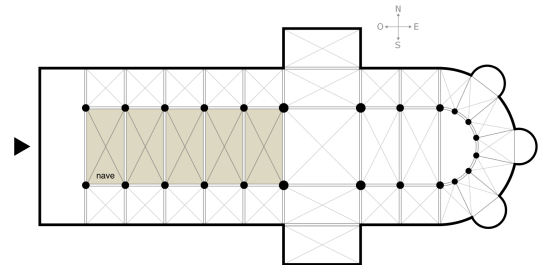
(a) The regression model for predicting total length from name height and style is also given below. Which of the following is false?

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	44.2979	81.6477	0.54	0.5929
naveheight	4.7116	1.0584	4.45	0.0002
style_roman	80.3925	32.3063	2.49	0.0209

$R^2_{adj} = 49.64\%$

(A) Each additional foot in nave height is associated with a 4.7116 foot increase in total length.

(B) Roman cathedrals with 0 nave height are expected on average to be 44.2979 feet in total length.

(C) Roman cathedrals are expected on average to be 80.3925 feet longer than Gothic cathedrals.

(D) Both nave height and style of cathedral are significant predictors of total length.

(b) Using the plots provided below check if the assumptions and conditions for MLR is satisfied.

Normal probability plot of residuals

Histogram of residuals

Residuals vs. fitted

Total length vs. nave height

Residuals vs. order of data collection

14. The Nielsen organization did a poll to determine whether men and women in different age groups watched different amounts of comedy television. The table below shows some summary statistics on each age/gender group.[7]

level	n	mean	sd
men 18-34	10	287	65.05
men 55+	10	171	40.81
women 18-34	10	353.70	20.78
women 55	10	356.90	39.37

(a) What method can we use to evaluate if different age/gender groups watch different amounts of comedy television on average. Explain your reasoning.

(b) Do assumptions and conditions for this technique appear to be satisfied?

15. National Health and Nutrition Examination Survey (NHANES) collects data on people's cholesterol levels and marital status, among many other variables. The data come from 940 respondents

[7]Adapted from *Introductory Statistics* by Gould and Ryan.

and marital status has 6 levels (divorced, living with partner, married, never married, separated, widowed). Below is the relevant ANOVA table.[8]

(a) Write the hypotheses for evaluating if average cholesterol level varies among people with different marital statuses.

(b) Below is the relevant ANOVA table. Fill in the blanks.

	df	SS	MS	F	p-value
marital_st		89,082			<0.0001
Residuals					
Total		1,909,292			

(c) What is the conclusion of the analysis?

(d) Assuming you did find an association between marital status and cholesterol levels, would this association mean that marital status caused different cholesterol levels? Can you think of a confounding factor?

16. Suppose the true population proportion were $p = 0.1$. The figure below shows what the distribution of a sample proportion looks like when the sample size is $n = 20$, $n = 100$, and $n = 500$. What does each observation in each distribution represent? Describe how the distribution of the sample proportion, $\hat{p}$, changes as n becomes larger.
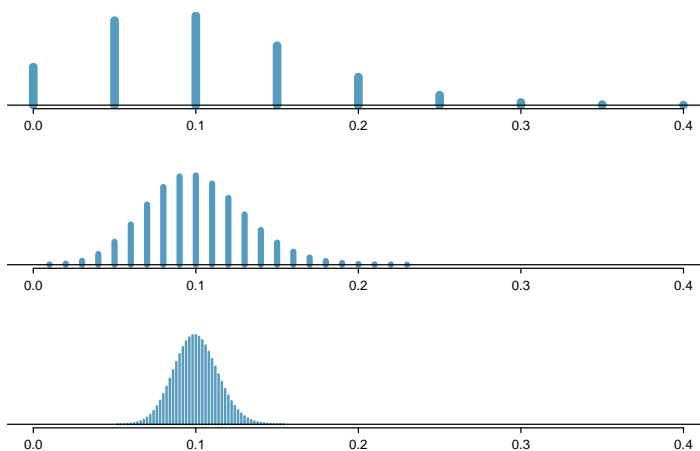

17. Suppose the true population proportion were $p = 0.5$. The figure below shows what the distribution of a sample proportion looks like when the sample size is $n = 20$, $n = 100$, and $n = 500$. What does each observation in each distribution represent? Describe how the distribution of the sample proportion, $\hat{p}$, changes as n becomes larger.
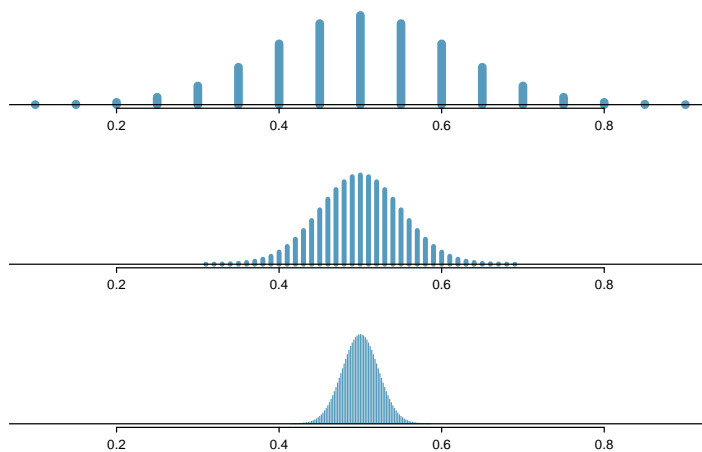
18. Suppose the true population proportion were $p = 0.95$. The figure below shows what the distribution of a sample proportion looks like when the sample size is $n = 20$, $n = 100$, and $n = 500$. What does each observation in each distribution represent? Describe how the distribution of the sample proportion, $\hat{p}$, changes as n becomes larger.
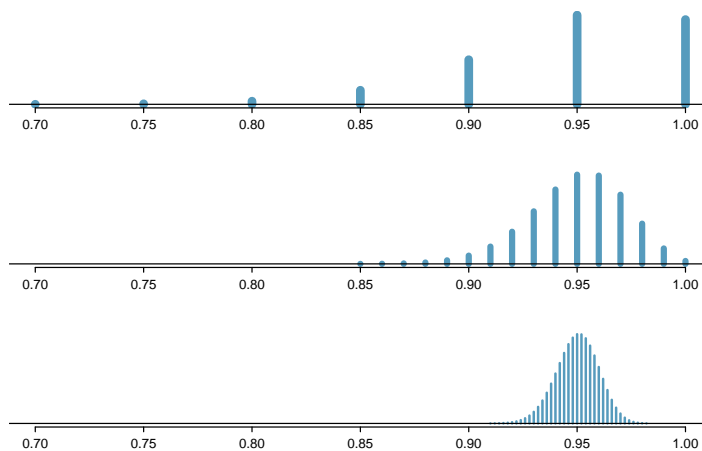
19. In 2013, the Pew Research Foundation reported that "45% of U.S. adults report that they live with one or more chronic conditions". However, this value was based on a sample, so it isnt perfect. The study reported a standard error of about 1.2%, and a normal model may be reasonably be used in this setting.

    (a) Create a 95% confidence interval for the proportion of U.S. adults who live with one or more chronic conditions. Also interpret the confidence interval in the context of the study.

    (b) Identify each of the following statements as true or false. Provide an explanation to justify each of your answers.

        i. We can say with certainty that the confidence interval contains the true percentage of U.S. adults who suffer from a chronic illness.

        ii. If we repeated this study 1,000 times and constructed a 95% confidence interval for each study, then approximately 950 of those confidence intervals would contain the true fraction of U.S. adults who suffer from chronic illnesses.

        iii. The poll provides statistically significant evidence (at $\alpha = 0.025$) that the percentage of U.S. adults who suffer from chronic illnesses is below 50%.

iv. Since the standard error is 1.2%, only 1.2% of people in the study communicated uncertainty about their answer.

20. A "social experiment" conducted by a TV program questioned what people do when they see a very obviously bruised woman getting picked on by her boyfriend. On two different occasions at the same restaurant, the same couple was depicted. In one scenario the woman was dressed "provocatively" and in the other scenario the woman was dressed "conservatively". The table below shows how many restaurant diners were present under each scenario, and whether or not they intervened. We are interested in evaluating whether people are equally likely to intervene when the woman is wearing a provocative or conservative outfit.

		Scenario		
		Provocative	Conservative	Total
Intervene	Yes	5	15	20
	No	15	10	25
	Total	20	25	45

Explain why the sampling distribution of the difference between the proportions of interventions under provocative and conservative scenarios does not follow an approximately normal distribution.