Questions

1. Investigate whether using the known-σ procedure works when you *don't* know σ by simulation in StatCrunch, as follows.
   (a) Generate a population of 1000 values from a normal distribution with mean 0 and SD 1, using Data and Simulate Data.
   (b) Generate 100 samples of size 5 from this population, using Stat and Sample Columns. Store them "stacked with a sample ID".
   (c) Calculate a chapter 6-style confidence interval for each sample, by selecting Stat, Z-statistics, 1-sample Z, With Data. Enter your column of sampled values as your data, and Group By sample number. Select a 95% CI. Store the results in the data table.
   (d) This produces 100 confidence intervals. Count how many of them actually contain the population mean of 0. (Or investigate Data, Bin Column and Stat, Table, Contingency, With Data.
   (e) How many of your intervals actually contain 0? How many should contain 0? Does this mean that the Z procedure gives intervals that are too wide, too narrow or about right?

Summary: too few intervals contain the pop mean, so have to make intervals bigger.

2. Instead of using z^* for a confidence interval, we use t^*, from table D. First find the df as $n-1$ (row), then find your confidence level along the bottom (column). This gives t^*.
   (a) Find t^* for the following intervals:
      (i) 95% CI with $n=5$
      (ii) 99% CI with $n=5$
      (iii) 95% CI with $n=51$
      (iv) 95% CI with $n=1001$
      (v) z^* for 95% CI.
   (b) What appears to happen as the confidence level gets bigger?
   (c) What appears to happen as the sample size gets bigger?

a:
I: 2.776 ii: 4.604 iii: 2.009 iv: 1.962 v: 1.960
(you check these)

b: t* gets bigger (like z*)
c: t* gets closer and closer to z*

3. What are sensible t^* values to use for a 95% CI when $n = 48$, $n = 120$? What, in general, should you do if your df is not in the table?

Round df *down* to next lowest value in table.

4. When people buy bicycles, they often buy other accessories too (helmet, water bottle, etc.) A bike store took a random sample of 12 bike-buying customers and found that the sample mean amount spent on accessories was \$77.83 and the sample SD was \$33.51.
   (a) Find a 95% confidence interval for the mean sales of accessories.
   (b) Check your answer using StatCrunch.

t*= 2.201   (95%, 11 df)

CI: $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$ or 77.83 +/- 21.29,

56.54 to 99.12

StatCrunch: t-statistics- 1-sample – with summary (get CI rather than test)

95% confidence interval results:

μ : population mean

Mean	Sample Mean	Std. Err.	DF	L. Limit	U. Limit
μ	77.83	9.673504	11	56.53876	99.12124

5. As for confidence intervals, we do t tests by replacing the unknown σ by s, calling the test statistic t, and using Table D to get a P-value. What can you say about the P-value in each of the following cases?
   - (a) $n = 10, t = 2.37, H_a : \mu > 10$
   - (b) $n = 10, t = 2.37, H_a : \mu \neq 10$
   - (c) $n = 20, t = 3.3, H_a : \mu \neq 10$
   - (d) $n = 20, t = 5.1, H_a : \mu \neq 10$
   - (e) $n = 20, t = 2.5, H_a : \mu < 10$

a: between 0.02 and 0.025
b: multiply a: by 2: between 0.04 and 0.05
you do the rest

6. At a second bike shop, mean accessory sales per bike sold are \$90. Is there evidence that mean accessory sales for all bikes sold at the first bike shop are less than this?
   (a) Write down suitable hypotheses.
   (b) Use the sample data $\bar{x} = 77.83, s = 33.51, n = 12$ to obtain the test statistic.
   (c) Obtain a P-value (as accurately as Table D will let you). What is your conclusion?
   (d) Check your work with StatCrunch. Is the P-value you obtained from StatCrunch consistent with the one you got from Table D?

Ha: mu<90
H0: mu=90

$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$ = (77.83-90)/(33.51/sqrt(12))=-1.258

P-value: throw away minus sign, between 0.10 and 0.15: do not reject H0, no evidence that mean accessory sales at this bike shop are less.

StatCrunch:

Hypothesis test results:
μ : population mean
$H_0 : \mu = 90$
$H_A : \mu < 90$

Mean	Sample Mean	Std. Err.	DF	T-Stat	P-value
μ	77.83	9.673504	11	-1.2580757	0.1172

P-value 0.1172 is between 0.10 and 0.15.

7. The t procedures are most useful for small samples. They are derived from populations having normal distributions. What if the populations are not normal?

T procedures are “robust” meaning that it doesn't matter too much what the population distribution is (as long as not too far from normal).

8. Make a table summarizing how you use Table D to obtain a P-value for each of the 3 possible types of H_a and the 2 cases of whether t is positive or negative.

Below, $|t|$ means t without its minus sign.

H_a	$t<0$	$t>0$
$H_a: \mu \neq \mu_0$	2 x P-val for $\lvert t \rvert$	2 x P-val for t
$H_a: \mu < \mu_0$	P-val for $\lvert t \rvert$	large
$H_a: \mu > \mu_0$	large	P-val for t

9. Back in Chapter 3, we learned about two different ways of doing a comparative study, matching and randomizing.
   (a) Briefly review these two ways.
   (b) How would matching fit into the 1-sample t framework that we just learned? (That is, how could we analyze a matched pairs experiment using methods that we already know?)

skip this one today

10. Sleep apnea is a condition in which a sleeping person stops breathing for a moment. It is especially serious in premature infants. A drug is tested on 13 premature infants, and the number of apneic episodes per hour recorded for each before and after the drug is given. Some results are shown below.

	n	mean	SD
Before	13	1.751	0.855
After	13	0.984	0.833
Difference	13	0.767	0.524

(a) Explain how this is a matched pairs experiment.
(b) Test whether the drug reduces the mean number of apneic episodes.
(c) Calculate a 90% confidence interval for the mean difference in apneic episodes per hour, before minus after. What is your confidence interval telling you about how effective the drug is?
(d) Check your results using StatCrunch.
(e) (optional) Get hold of the data, and check your answers again with StatCrunch.

a: 2 measurements on each of the 13 infants

One-sample t on difference as shown:
t=(0.767-0)/(0.524/sqrt(13))=5.278,
P-value less than 0.0005:
evidence the drug helps. (StatCrunch gives P-value less than 0.0001.)

CI: t*=1.782, interval 0.767 +/- 1.782 (0.524/sqrt(13)) = from 0.51 to 1.03.

11. Table 7.3 on page 421 of the text gives lengths in seconds of 50 audio files sampled from an iPod. We are interested in whether the "average" length of audio file is 240 seconds or not. Use StatCrunch to answer the following:

   (a) Why should you have doubts about using a t-test for these data? Hint: look at a picture.

   (b) One approach is to do a "transformation" of the data: take the logarithms (ln) of each value, and test whether the mean of the log-values could be $\ln(240) = 5.480$ or not:

      (i) Calculate a column of log file lengths using Data and Transform Data.

      (ii) Check that the new column is much less skewed.

      (iii) Do a 1-sample t-test on the new column, testing $H_0 : \mu = 5.480$ vs. $H_a : \mu \neq 5.480$. What do you conclude?

skip this one and the remaining questions

12. Another possibility, for the same data as above, is the "sign test". Let m be the population *median*. This uses the original data to test $H_0 : m = 240$ vs. $H_a : m \neq 240$:
    - (a) If H_0 is true, what is the probability that an individual observation will be above 240? (Ignore the possibility that an observation could be exactly equal to 240.)
    - (b) Out of the 50 values in the sample, what is the distribution of the number X of values bigger than 240, if H_0 is true?
    - (c) In the data, 32 of the 50 values were bigger than 240. What is $P(X \geq 32)$ in your distribution above? Double it to obtain a P-value for your test. What do you conclude now?
    - (d) (optional) Compare StatCrunch's built-in sign test (Stat, Nonparametrics, Sign Test). Do you get a similar P-value? Why is the P-value not exactly the same?

skip this one

13. For the sleep apnea example, what is the power of the matched pairs test to detect a reduction of 0.5 apneic episode per hour? Assume that the differences have SD 0.6. Follow the steps below:
    (a) Generate a population of 1000 values from a normal distribution with mean 0.5 and SD 0.6 (the differences).
    (b) Obtain 100 samples each of size 13 from this population, "stacked with column ID".
    (c) For each sample, test $H_0 : \mu = 0$ vs. $H_a : \mu > 0$, obtaining a P-value. Save these P-values in the worksheet.
    (d) How many of the P-values are less than 0.05? What, then, is the (simulated) power of the test?

    Note that the power of any t-test can be obtained in the same way, not just a matched-pairs one.

Skip this one.