Here is a list of tests that are possible answers in Question 1–7:

1. One-sample t-test
2. Paired t-test
3. Two-sample t-test
4. One-sample proportion test
5. Two-sample proportion test
6. Chi-square test of goodness of fit
7. Chi-square test of independence
8. Linear regression

Question 1

You are testing a drug designed to treat the medical condition plantar fasciatis. 80 patients are split into two groups at random. One group receives a placebo, and one group receives the drug. After two weeks, patients are asked if they still have symptoms of plantar fasciatis, and either yes or no is recorded in the dataset. You want to determine if the drug is helpful.

• What test would you use to investigate this question? Just give the number.
  5, two-sample proportion test

• Write null and alternative hypotheses to investigate this question.

  Let \( p_1 \) denote the proportion out of all people with plantar fasciatis whose symptoms resolve after two weeks of treatment with the new drug. Let \( p_2 \) denote the proportion out of all people with plantar fasciatis whose symptoms resolve after two weeks of a placebo.

  \( H_0: p_1 = p_2 \)
  \( H_A: p_1 \neq p_2 \)

Question 2

According to physicists’ predictions, the subatomic particle called the higgs boson would have a mass of 125,100 MeV/c\(^2\). A group of researchers collects 1000 observations of the boson’s mass to check if it differs on average from the expected mass.

• What test would you use to investigate this question? Just give the number.
  1, one-sample t-test

• Write null and alternative hypotheses to investigate this question.

  Let \( \mu \) be the mean weight of the Higgs boson over all measurements ever taken.

  \( H_0: \mu = 125,100 \) MeV/c\(^2\)
$H_A$: $\mu \neq 125, 100 \text{ MeV/c}^2$
Question 3

Two different fertilizers are tested on different fields for growing corn. At the end of the season, the amount of corn grown (in pounds) is measured on 37 fields that received fertilizer A and on 34 fields that received fertilizer B. You would like to know whether there is a difference in corn production on average for the two fertilizers.

- **What test would you use to investigate this question?** Just give the number.
  3, two-sample t-test

- **Write null and alternative hypotheses to investigate this question.**
  \( H_0: \) mean corn grown is equal with fertilizer A and B
  \( H_A: \) mean corn grown is not equal with fertilizer A and B

Question 4

An economic research project tracks 400 people sampled at random across the country. The amount of money they spend on holiday gifts is recorded in 2018 and again in 2019. The researchers want to know if the different economic conditions in 2018 and 2019 are affecting holiday spending.

- **What test would you use to investigate this question?** Just give the number.
  2, paired t-test

- **Write null and alternative hypotheses to investigate this question.**
  \( H_0: \) the mean difference of a person’s 2018 and 2019 holiday spending is 0
  \( H_A: \) the mean difference of a person’s 2018 and 2019 holiday spending is not 0
Question 5

Flu shots can sometimes cause a minor but painful reaction at the injection site. A medical study seeks to determine whether the location of the shot is relevant. In the study, patients are given a flu shot either in the shoulder, the thigh, or the calf. For each patient, the location of the shot is recorded, and it is recorded whether or not they have a painful reaction. You wish to determine if there is any difference in the frequency of painful reactions depending on the location of the shot.

- What test would you use to investigate this question? Just give the number.
  7, chi-square test of independence

- Write null and alternative hypotheses to investigate this question.
  \( H_0 \): there is no association between location of shot and painful reaction at the injection site
  \( H_A \): there is an association between location of shot and painful reaction at the injection site

Question 6

You take a sample of city blocks. On each block, you record the number of cars driving on the block in an hour, and you count the number of pigeons seen on the block. You are interested in the relationship between these two counts, and specifically if the number of pigeons on average varies based on the number of cars.

- What test would you use to investigate this question? Just give the number.
  8, linear regression

- Write null and alternative hypotheses to investigate this question.
  Let \( \beta_1 \) be the slope of the least-squares regression line where number of cars driving on the block is the explanatory variable and number of pigeons on the block is the response variable.
  \( H_0 \): \( \beta_1 = 0 \)
  \( H_A \): \( \beta_1 \neq 0 \)
Question 7

You would like to know if people who run regularly differ in foot width from people who don’t run regularly. You measure the foot width of randomly sampled runners and nonrunners.

- **What test would you use to investigate this question?** Just give the number.
  3, two-sample t-test

- **Write null and alternative hypotheses to investigate this question.**
  
  $H_0$: mean foot width is the same in the populations of runners and nonrunners
  
  $H_A$: mean foot width is not the same in these populations
Question 8

You plan to measure the average density of taste receptors on the human tongue. You obtain a sample of 500 people and measure the density for each person. Here is a histogram and Q-Q plot of the sample data:
• List the conditions for valid inference of the mean density of taste receptors in the population using the t-distribution. Give your judgment as to whether they are satisfied.

The conditions to be satisfied are independence of observations and having a large enough sample. Independence of observations is satisfied if the sample is a simple random sample. A sample size of 500 is plenty in the absence of extreme outliers or skew, neither of which is present according to the histogram. We didn’t cover Q-Q plots in this class, so you should just ignore the Q-Q plot. It’s not relevant anyway (they’re a tool for testing normality of data, but we don’t even need the data to be close to the normal distribution for this sample size anyhow).
Question 9

You would like to determine if children in London and in Tel Aviv have different rates of peanut allergies. You take samples in both places and find that 23 of 1000 children in London and 16 of 2000 in Tel Aviv have peanut allergies. You want to do a two-sample proportion test with null hypothesis that an equal proportion of children in the two cities have peanut allergies and alternative hypothesis that a different proportion of children do.

- **List the conditions for the test of significance to provide valid results. Judge if they’re satisfied.** (You can assume that the samples are simple random samples of children from the two cities.)

  We need independence of observations and the success-failure condition to hold. Independence of observations holds assuming our sample is a simple random sample. We test the success-failure condition under the assumption of the null hypothesis when doing a hypothesis test. This means that we assume that \( p \) is equal to the null proportion, which for a two-sample test is the pooled proportion:

  \[
  p_0 = \frac{23 + 16}{1000 + 2000} = .013
  \]

  So, under this assumption, we have

  \[
  n_1p = (1000)(.013) = 13, \quad n_1(1 - p) = 987, \\
  \]

  All of these are 10 or above, so the condition is satisfied.
Question 10

You have a dataset \texttt{bdims}. Each observation represents one man, recording some physical measurements about him including the following two variables:

- \texttt{hgt}: the height of the man in inches
- \texttt{bic.gi}: the bicep girth of the man in inches (i.e., the distance around the person’s flexed bicep)

You would like to study the relationship of the two variables by fitting a simple linear regression model with \texttt{hgt} as the explanatory variable and \texttt{bic.gi} as the response variable.

Here is a scatterplot of the dataset:

![Scatterplot of dataset](image1)

Here is a scatterplot showing the residuals for the regression:

![Residuals plot](image2)
Here's a histogram and Q-Q plot of the residuals:

```r
ggplot(bdims, aes(x=residual)) + geom_histogram(bins=25)
ggplot(bdims, aes(sample=residual)) + stat_qq()
```

- **List the conditions for simple linear regression modeling. Judge if they’re satisfied.** (You may assume that the dataset is a simple random sample from the population. Do your best on judging whether the conditions are satisfied—I’ll be generous in my grading so long as you get the conditions right.)

We need independence of observations to hold, and we need the linear trend, normal residuals, and constant variation assumptions to hold. Independence of observations holds assuming the sample is a simple random sample. In the first scatterplot, there doesn’t appear to be a strong linear trend, but there are no signs of any nonlinear trend, and this is enough to make the assumption hold. The histogram of the residuals shows a good match to the normal distribution. In particular, there are no outliers. You can ignore the Q-Q plot since we didn’t cover them. And the variability of the residuals appears constant left to right in the residual plot.
Question 11

With the data from the previous problem, you run the following commands in R. Regardless of your answer to the previous problem, assume now that inference for the linear regression model is valid.

Your run the following commands in R to make this chart:

```r
results <- lm(bic.gi~hgt, data=bdims)
ggplot(bdims, aes(x=hgt, y=bic.gi)) + geom_point() + geom_smooth(method=lm)
```

`geom_smooth()` using formula 'y ~ x'
Then you run the following commands:

```r
coef(results)
```

```
## (Intercept)  hgt
## 7.65851046  0.08419799
```

```r
confint(results, level=.95)
```

```
## 2.5 % 97.5 %
## (Intercept) 4.073339 11.24368
## hgt 0.033046 0.13535
```

```r
xvals <- data.frame(hgt=c(70))
predict(results, xvals, interval="confidence", level=.95)
```

```
## fit lwr   upr
## 1 13.55237 13.40788 13.69686
```

```r
predict(results, xvals, interval="prediction", level=.95)
```

```
## fit  lwr   upr
## 1 13.55237 11.27702 15.82772
```

Answer the following multiple-choice questions:

- With 95% confidence, the best-fit linear model for the population predicts that every inch of height increases the mean bicep girth by:

  a) 4.073 to 11.244 inches  
  b) 0.033 to 0.135 inches  
  c) 13.408 to 13.697 inches  
  d) 11.277 to 15.828 inches  
  e) impossible to determine from information above

**ANSWER:** b

- According to the model, 95% of all people with height 70 inches will have bicep girth between:

  a) 4.073 to 11.244 inches  
  b) 0.033 to 0.135 inches  
  c) 13.408 to 13.697 inches  
  d) 11.277 to 15.828 inches  
  e) impossible to determine from information above

**ANSWER:** d
Question 12

Psychologists carry out an experiment to determine the effect of competition. Participants in the study are assigned to one of two groups at random. In both groups, participants are timed as they complete a short mathematical quiz and are told to go as quickly as they can. In the control group, the participants do this one at a time, in an empty room. In the competition group, the participants do this simultaneously, in a room together. The times of the participants are stored in seconds as control.times and competition.times, and the researchers run the following command in R.

```r
t.test(control.times, competition.times, mu=0, alternative="two.sided")
```

##
## Welch Two Sample t-test
##
## data: control.times and competition.times
## t = 1.5916, df = 40.339, p-value = 0.1193
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -4.059767 34.189876
## sample estimates:
## mean of x mean of y
## 69.85694 54.79188

- What were the null and alternative hypotheses tested?

Let $\mu_1$ and $\mu_2$ denote the mean time to do the quiz for people in an empty room and in a room with other people doing the same task, respectively.

$H_0$: $\mu_1 = \mu_2$ (or $\mu_1 - \mu_2 = 0$)

$H_A$: $\mu_1 \neq \mu_2$ (or $\mu_1 - \mu_2 \neq 0$)

- On average, how long did it take the participants in the control group to complete the quiz?

69.9 seconds

- On average, how long did it take the participants in the competition group to complete the quiz?

54.8 seconds

- What is the result of the test, at a 5% significance level?

Fail to reject the null hypothesis: there’s insufficient evidence to conclude that doing the task in an empty/full room makes a difference.
Question 13

The IRS would like to know if rates of tax evasion have increased from 2017 to 2018. They choose a random sample of tax returns from 2017 and from 2018 and investigate each one for signs of tax evasion. They find evidence of tax evasion in 33 out of 3521 tax returns from 2017 and in 63 out of 3965 tax returns from 2018. They run the following command in R:

\[
\text{prop.test(c(33, 63), n=c(3521, 3965), correct=FALSE, alternative="less")}
\]

---

**What were the null and alternative hypotheses tested?**

Let \( p_{2017} \) be the proportion of 2017 tax returns with signs of tax evasion and \( p_{2018} \) the proportion of 2018 tax returns with signs of tax evasion.

\[
H_0: p_{2017} = p_{2018} \\
H_A: p_{2017} < p_{2018}
\]

We didn’t cover one-sided tests in this class, so you probably would have just answered \( p_{2017} \neq p_{2018} \) for the alternative hypothesis.

**At a 5% significance level, what is the result of the test?**

Reject the null hypothesis in favor of the alternative.
Question 14

You have a fair die (i.e., it is equally likely to come up 1, 2, 3, 4, 5, or 6 when rolled). You roll it 1000 times and record the number of times each value is rolled. Then you perform a chi-square test of goodness of fit with significance level .05 and the following hypotheses:

- $H_0$: Each roll of the die is sampled uniformly from the values 1, 2, 3, 4, 5, and 6 (i.e., each value has probability 1/6 of being rolled).
- $H_A$: The rolls of the die are not sampled like this.

• Respond TRUE or FALSE to the following statements.
  - The null hypothesis is true.
    **TRUE**   **FALSE**
  
  - The p-value for the test will definitely be smaller than .05.
    **TRUE**   **FALSE**
  
  - The result of the test will definitely be to fail to reject the null hypothesis.
    **TRUE**   **FALSE**
  
  - If the result of the test is to reject the null hypothesis, then a type-1 error has occurred.
    **TRUE**   **FALSE**
  
  - If the result of the test is to fail to reject the null hypothesis, then a type-2 error has occurred.
    **TRUE**   **FALSE**
  
  - About 5% of the time, the result of the test will be to reject the null hypothesis.
    **TRUE**   **FALSE**
Question 15

- Respond TRUE or FALSE to the following statements.
  - The 99%-confidence interval for a parameter will be wider than the 95%-confidence interval.
    TRUE FALSE
  - If the 95%-confidence interval for mean monthly rent in Brooklyn is found to be $1893 to $2641, then approximately 95% of renters in Brooklyn pay between $1893 and $2641 per month.
    TRUE FALSE
  - When you compute the 95%-confidence interval for a mean based on a sample, the confidence interval will always contain the sample mean.
    TRUE FALSE
  - When you compute the 95%-confidence interval for a mean based on a sample, the confidence interval will always contain the population mean.
    TRUE FALSE

Question 16

A nutrition study collects a very large sample of U.S. individuals (almost 10,000) and then interviews them about their eating habits. The individuals are then monitored for the next ten years and their health outcomes are recorded. One variable egg-intake has the values "<= 1/week", "2-5/week", or ">= 6 week", depending on the typical number of eggs the individual eats per week. Another variable, heart-attack has the values "yes" or "no" depending on whether the individual experiences a heart attack during the ten years when health outcomes are tracked. The researchers carry out a chi-square test of independence between the egg-intake and heart-attack variables and find a p-value of .0000018.

- Respond TRUE or FALSE to the following statements.
  - If people in the different egg-eating categories in the population in fact have heart attacks at the same rate, the chance of observing differences between the groups as large as we did in the sample is only .0000018.
    TRUE FALSE
  - The tiny p-value provides extremely strong evidence that eating more eggs increases the risk of a heart attack.
    TRUE FALSE
  - This is an observational study, not an experiment.
    TRUE FALSE
  - A chi-square test is inappropriate for this data. Researchers should have done a paired t-test.
    TRUE FALSE