Instructions.

- DON’T PANIC! If you get stuck, take a deep breath and go on to the next question.
- You are allowed a two-sided sheet of notes in your own handwriting.
- You may use any calculator, but you cannot use a phone or computer.

<table>
<thead>
<tr>
<th>Question</th>
<th>Points</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td><strong>132</strong></td>
<td></td>
</tr>
</tbody>
</table>

Good luck!
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.

• Write null and alternative hypotheses to investigate this question.

Question 2

According to physicists’ predictions, the subatomic particle called the higgs boson would have a mass of 125,100 MeV/c². 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.

• Write null and alternative hypotheses to investigate this question.
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.

- Write null and alternative hypotheses to investigate this question.

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.

- Write null and alternative hypotheses to investigate this question.
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.

- **Write null and alternative hypotheses to investigate this question.**

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.

- **Write null and alternative hypotheses to investigate this question.**
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.

- Write null and alternative hypotheses to investigate this question.
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.
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.)
Question 10

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

- `hgt`: the height of the man in inches
- `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 `hgt` as the explanatory variable and `bic.gi` as the response variable.

Here is a scatterplot of the dataset:

![Scatterplot of hgt vs. bic.gi](image)

Here is a scatterplot showing the residuals for the regression:

![Residuals Scatterplot](image)
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.)
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)
```
Then you run the following commands:

```r
coef(results)
## (Intercept)  hgt
##  7.65851046  0.08419799

confint(results, level=.95)
##  2.5 %  97.5 %
## (Intercept) 4.073339 11.24368
##   hgt  0.033046  0.13535

xvals <- data.frame(hgt=c(70))
predict(results, xvals, interval="confidence", level=.95)
##  fit   lwr   upr
## 1 13.55237 13.40788 13.69686

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:** __________

- 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:** __________
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.

```
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?

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

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

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

```r
prop.test(c(33, 63), n=c(3521, 3965), correct=FALSE, alternative="less")
```

```
## 2-sample test for equality of proportions without continuity correction
## data: c(33, 63) out of c(3521, 3965)
## X-squared = 6.256, df = 1, p-value = 0.006189
## alternative hypothesis: less
## 95 percent confidence interval:
## -1.00000000 -0.00229722
## sample estimates:
## prop 1 prop 2
## 0.009372337 0.015889029
```

- What were the null and alternative hypotheses tested?

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

<table>
<thead>
<tr>
<th>TRUE</th>
<th>FALSE</th>
</tr>
</thead>
</table>

- The p-value for the test will definitely be smaller than .05.

<table>
<thead>
<tr>
<th>TRUE</th>
<th>FALSE</th>
</tr>
</thead>
</table>

- The result of the test will definitely be to fail to reject the null hypothesis.

<table>
<thead>
<tr>
<th>TRUE</th>
<th>FALSE</th>
</tr>
</thead>
</table>

- If the result of the test is to reject the null hypothesis, then a type-1 error has occurred.

<table>
<thead>
<tr>
<th>TRUE</th>
<th>FALSE</th>
</tr>
</thead>
</table>

- If the result of the test is to fail to reject the null hypothesis, then a type-2 error has occurred.

<table>
<thead>
<tr>
<th>TRUE</th>
<th>FALSE</th>
</tr>
</thead>
</table>

- About 5% of the time, the result of the test will be to reject the null hypothesis.

<table>
<thead>
<tr>
<th>TRUE</th>
<th>FALSE</th>
</tr>
</thead>
</table>
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