[PRINT]

Math 325 - Fall 2020 - Section 2,
Kyle Tolliver, 12/11/20 at 12:18:12 AM MST

Question1: *Score 0/1*

Use an appropriate test and the **starwars** dataset in R to determine if, on average, the species of Wookiees, Gungans, or Kaminoans are taller.

Your response	Correct response
There is evidence to suggest that Gungans are taller, on average, than Wookiees and both are taller, on average, than Kaminoans.	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question2: *Score 0/1*

Use the **Highway1** dataset in R to answer this question.

Suppose someone wanted to perform a t Test of the hypotheses

$$H_0 : \mu_{2\text{ Lanes}} = \mu_{4\text{ Lanes}}$$

$$H_a : \mu_{2\text{ Lanes}} \neq \mu_{4\text{ Lanes}}$$

where $\mu_{2\text{ Lanes}}$ represents the 1973 accident rate (per million vehicle miles) on 2 lane highways, and $\mu_{4\text{ Lanes}}$ represents the 1973 accident rate (per million vehicle miles) on 4 lane highways.

Which of the following statements would be most correct concerning the appropriateness of performing this test for this data?

Your response	Correct response
This test is appropriate for these data because even though the data is not normal, the sample size is over 30.	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question3: *Score 0/1*

Run the following code in R.

> plot(height ~ age, data=Loblolly)

Is it appropriate to perform a simple linear regression on this data?

Your response	Correct response
Yes, the regression requirements all appear to be well satisfied.	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question4: *Score 1/1*

Which analysis would allow you to predict whether or not it will rain on a given day based on the low temperature of that day?

Your response	Correct response
Logistic Regression	

Auto graded **Grade:** 1/1.0

Total grade: 1.0×1/1 = 100%

Feedback:

Because we are using temperature, which is quantitative, as an explanatory variable to try to predict whether or not it will rain (a 0, 1 outcome) this is a logistic regression.

Question5: *Score 0/1*

Use the sample of children's feet contained in the **KidsFeet** dataset to answer this question.

Why might a Wilcoxon Rank Sum Test not be appropriate for testing to see if the median width of a child's foot differs for boys and girls?

Your response	Correct response
The data is paired so a Wilcoxon Signed-Rank Test should be used instead.	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question6: *Score 1/1*

Use the **Highway1** dataset in R for this question.

Compute the average accident rate (per million vehicle miles) in 1973 for each type of roadway: `MC`, `FAI`, `PA`, and `MA`.

Your response	Correct response
$\bar{x}_{FAI} = 2.872$, $\bar{x}_{MA} = 4.870$, $\bar{x}_{MC} = 3.585$, and $\bar{x}_{PA} = 3.608$	

Auto graded **Grade:** 1/1.0

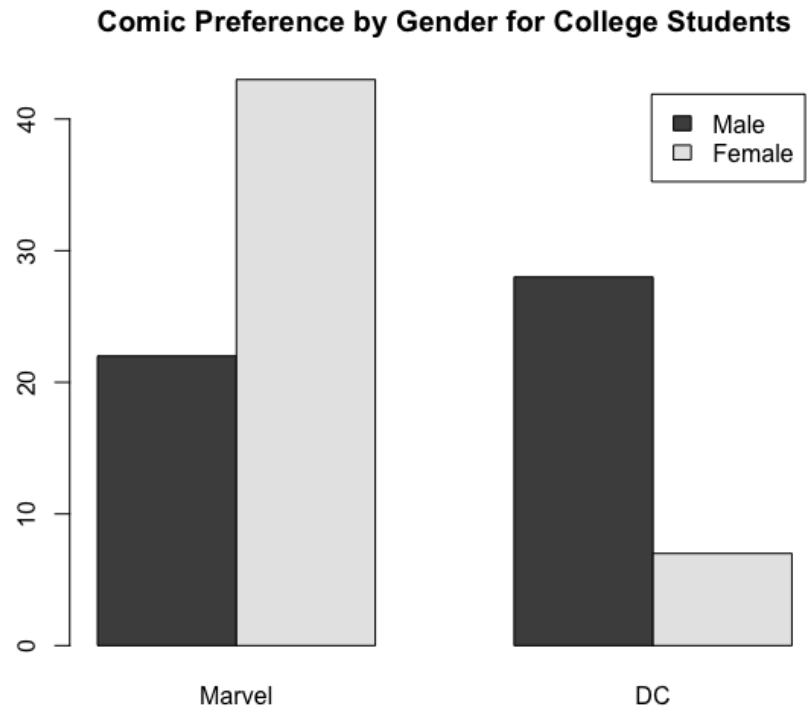
Total grade: 1.0×1/1 = 100%

Question7: *Score 0/1*

Consider the barplot below and the hypothesis that gender and preference for DC or Marvel comics is independent.

If this hypothesis is tested with an appropriate test using the data of this plot, what is the p-value of the test?

Note: DC comics include Superman, Batman, Green Latern, The Flash, and Wonder Woman. Marvel comics include Spiderman, Ironman, Captain America, Thor, and The X-Men.



Your response	Correct response
0.0821	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Feedback:

While you could technically load this data into R and test it manually to get a p-value with the code:

comic <- cbind(Male = c(Marvel = 22, DC = 28), Female = c(Marvel = 44, DC = 6))

chisq.test(comic)

This question is actually conceptual. In a chi squared test, independence (the null) would imply that the pattern in the bars is exactly the same for each group. As the pattern in the bars starts to differ between groups, the p-value gets smaller and smaller. With the graph shown, the pattern between the Marvel and DC bars is fundamentally different. Thus, the p-value is very small, so the correct answer would be 0.000009289 because all other p-values listed are not significant.

Question8: *Score 0/1*

In the **KidsFeet** dataset in R, there are three names of the children that occur twice, while all the other names occur only once. Compute the average foot width for the three names that occur more than once. Select the answer below that shows these three averages.

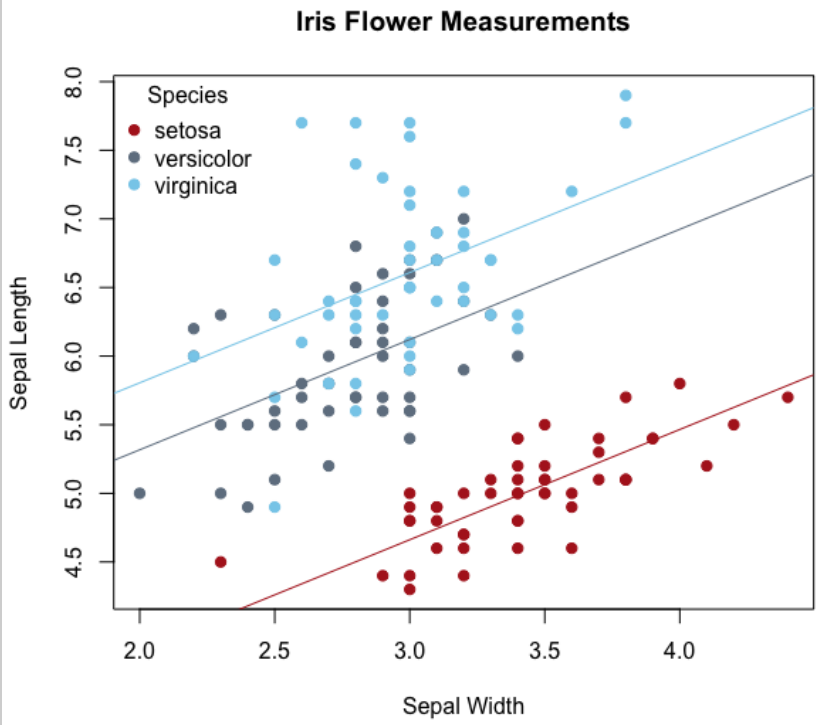
Your response	Correct response
8.7, 8.43, and 9.17	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question9: *Score 0/1*

Use the **iris** dataset in R and the image below to come up with the regression equation for the versicolor species.



Your response	Correct response
$\hat{Y} = 2.25 + 1.459x$	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question10: *Score 0/1*

Select the test that would be most appropriate to use in answering the following question for students here at BYU-Idaho.

Are Business majors more likely to be married than Data Analysis, Biostatistics, or Mathematics majors?

Your response	Correct response
Independent Samples t Test	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question11: *Score 0/1*

Use the **cars** dataset in R for this question.

Predict how fast a car (from 1920, see ?cars) was going if it took 130 feet for it to come to a complete stop after applying the brakes.

Your response	Correct response
Roughly 25 mph	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question12: *Score 0/1*

Use the **Galton** dataset to answer this question.

Perform a permutation test to determine if the average height of Males is different that the average height of Females.

How does the observed test statistic compare to the distribution of test statistics from the permutated data (i.e., the "permuted test statistics")?

Your response	Correct response
The observed test statistic is on the high end of the distribution of permuted test statistics, p-value $\approx$ 0.022.	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question13: *Score 1/1*

Use the **Highway1** dataset in R for this question.

Use an appropriate Wilcoxon Test to test if the median 1973 accident rate (per million vehicle miles) on 2 lane highways is different than the median 1973 accident rate (per million vehicle miles) on 4 lane highways.

Select the answer showing the test statistic and p-value of your test.

Your response	Correct response
W = 176, p-value = 0.8669	

Auto graded **Grade:** 1/1.0

Total grade: 1.0×1/1 = 100%

Question14: *Score 0/1*

Run the following code in R.

```
> plot(rate ~ slim, data=Highway1, pch=16, xlab="", ylab="", main="")
```

Which set of plot labels would provide the most correct and most useful information for the graphic produced by this code?

Your response	Correct response
xlab="Speed Limit", ylab="Accident Rate", main="Speed Limit and Accident Rates"	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question15: *Score 1/1*

Use an appropriate test and graphic for the **singer** dataset in R (?singer) to determine if and how the median height differs for different voice parts.

Your response	Correct response
The median heights differ significantly for the different voice parts with bases being the tallest and sopranos being the shortest ($p < 0.001$).	

Auto graded **Grade:** 1/1.0

Total grade: 1.0×1/1 = 100%

Question16: *Score 0/1*

Run the following code in R.

```
> plot(speed > 15 ~ dist, data=cars, ylab="Probability Speed > 15 mph", xlab="Stopping Distance (feet)")
```

What is the estimated probability that the speed of the car is greater than 15 mph if the car takes 70 feet to stop? (Note this data is from the 1920's, see ?cars for details.)

Your response	Correct response
0.9632993	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question17: *Score 1/1*

Which of the following is a nonparametric test?

Your response	Correct response
Permutation Test	

Auto graded **Grade:** 1/1.0

Total grade: 1.0×1/1 = 100%

Question18: *Score 1/1*

Perform a two-way ANOVA using the **KidsFeet** dataset in R that will allow you to test if the pattern shown in the following plot is real. Use the $\alpha = 0.05$ level and report the p-value of the test.

```
> xyplot(length ~ sex, data=KidsFeet, group=domhand, type=c("p","a"), auto.key=TRUE)
```

Your response	Correct response
Yes, the pattern is real (p = 0.0487)	

Auto graded **Grade:** 1/1.0

Total grade: 1.0×1/1 = 100%

Question19: *Score 0/1*

Run the following code in R.

```
> hist(islands, xlab="Area in Thousands of Square Miles", main="Areas of the World's Major Landmasses")
```

Which set of statistics would be most meaningful to include in an analysis containing this histogram?

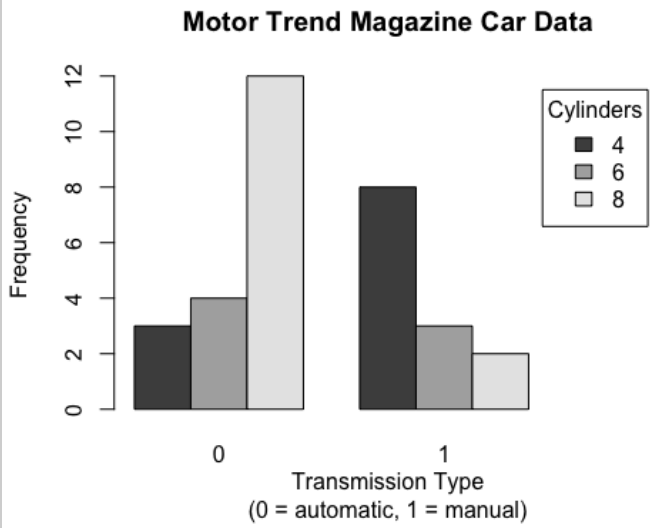
Your response	Correct response
median(islands) and sd(islands)	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question20: *Score 1/1*

Which of the following plots can be created using the **mtcars** dataset in R?

Your response	Correct response
	

Auto graded **Grade:** 1/1.0

Total grade: 1.0×1/1 = 100%

Question21: *Score 1/1*

What two things are needed to compute a p-value?

Your response	Correct response
A test statistic and a distribution of the test statistic.	

Auto graded **Grade:** 1/1.0

Total grade: 1.0×1/1 = 100%

Question22: *Score 0/1*

A logistic regression was performed to determine how cholesterol levels of adult men (X) impact their probability of having a heart attack (Y=1 denoting a heart attack). The value of β_0 in the model was estimated to be 56.4146 while the value of β_1 in the model was estimated to be 0.251.

Interpret the estimate for β_1 in context of the odds of an adult male having a heart attack.

Your response	Correct response
The odds increase by roughly 25.1% for every one unit increase in cholesterol level.	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question23: *Score 0/1*

An ANOVA is performed with the following output, but the F value is missing.

```
            Df Sum Sq Mean Sq F value   Pr(>F)
feed         5 231129   46226         5.94e-10 ***
Residuals   65 195556    3009
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

What must be the value of the F-statistic for this test?

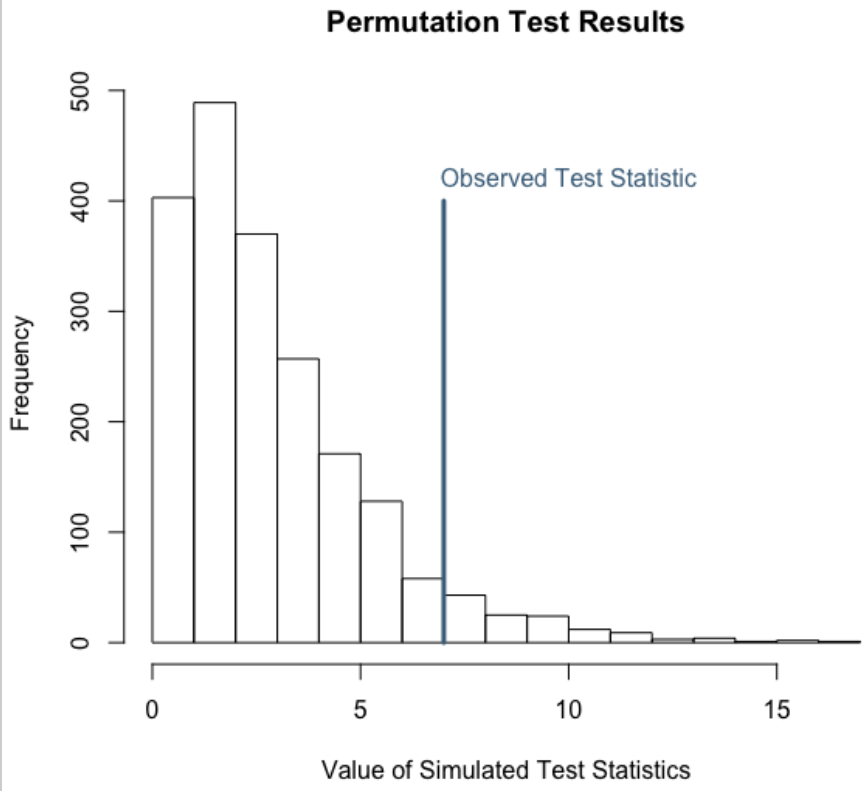
Your response	Correct response
F = 6.88349	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question24: *Score 0/1*

Suppose a permutation test is performed and the following histogram of 2,000 permuted test statistics is obtained. The observed test statistic is denoted by the vertical line. What is the p-value of the test?



Your response	Correct response
p = 0.005	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%

Question25: *Score 0/1*

Use an appropriate t Test and the **ToothGrowth** dataset in R to answer this question.

Is the average length of tooth growth in guinea pigs different for guinea pigs receiving an Orange Juice supplement than it is for guinea pigs receiving a Vitamin C supplement?

Your response	Correct response
Yes, there is evidence to conclude that the average length of tooth growth is longer in guinea pigs receiving the Orange Juice supplement (t=1.9153, df=55.309, p=0.06063).	

Auto graded **Grade:** 0/1.0

Total grade: 0.0×1/1 = 0%