**Question 1**(1 point)

*Saved*

There is a growing "speedrunning" community in Minecraft. Speedrunning means to attempt to complete the game as fast as possible. In order to do so in Minecraft, a player needs to obtain "Blaze rods". The principal source of these Blaze rods are "Blazes". When the player kills a Blaze, there is a 0.5 chance that a Blaze rod is yielded. You are part of the Minecraft speedrunning moderator team and a player submits a run with a record breaking time. In this run, this player obtained 211 Blaze rods from killing 305 Blazes.

This outcome is consistent with chance (assuming a conservative alpha of 0.005) and thus you should validate this run.

Question 1 options:

|  |  |
| --- | --- |
|  | True |
|  | False |

**Question 2**(1 point)

*Saved*

Corporate needs you to determine how much people like the song "No diggity". You sample 100 people and the standard error of the ratings is 0.5. This is not good enough. For an accurate estimate, the standard error needs to be 0.25. Assuming the standard deviation does not change, you need to sample - in total - this many people

Question 2 options:

|  |  |
| --- | --- |
|  | 400 |
|  | 125 |
|  | 250 |
|  | 200 |
|  | 2000 |

**Question 3**(1 point)

*Saved*

The significance level alpha is conventionally set at 0.05. The rationale for this choice is

Question 3 options:

|  |  |
| --- | --- |
|  | Pearson and Kolmogorov got together and proved that one can derive this from the first moment of any RV (assuming iid) |
|  | Sir Ronald Fisher had inscribed it on stone tablets, as he descended from the mountain |
|  | The international guild of frequentist statisticians (IGFS) decided to adopt this standard at their 1953 convention in Stockholm |
|  | It was a convenient choice for Fisher, as he was lacking a digital computer |
|  | It can be shown that this choice minimizes false positives. |

**Question 4**(1 point)

*Saved*

You want to know if the background color of the website you are developing has an emotional impact on your users. So in an A-B test design, you randomly assign a large group of users to versions of your site with either red or blue backgrounds (every user only sees one of these versions). You find that users who were assigned to the condition with the red background report much higher levels of anxiety and also spend considerably less time on the website than those who experienced the blue version. This difference is statistically significant. Can you conclude from this that the color of the background of your website has an emotional impact?

Question 4 options:

|  |  |
| --- | --- |
|  | No, correlation is not causation. |
|  | Yes, sure. Why not? |
|  | No, socioeconomic status is a possible confound |
|  | No, gender is a possible confound |
|  | No, both gender and socioeconomic status are possible confounds and there are others we haven't even considered that might also matter |

**Question 5**(1 point)

*Saved*

You want to estimate the value of a population mean from the sample mean. The best possible estimate comes from (select all that apply)

Question 5 options:

|  |  |
| --- | --- |
|  | Sampling independently |
|  | Sampling randomly |
|  | A large sample size |
|  | A small sample size |
|  | Allowing people in the population to self-select whether they want to be in the sample. |
|  | Hand-picking the people from the population that go into the sample. |

**Question 6**(1 point)

*Saved*

Find the matching pairs

Question 6 options:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | |  | Placeholder | |  | Treatment | |  | Measurement | |  | Function | |  | |  |  | | --- | --- | | **1**. | Variable | | **2**. | Random Variable | | **3**. | Independent Variable | | **4**. | Dependent Variable | |

**Question 7**(1 point)

*Saved*

A lady claims that she has a special gift: According to her, she can - blindfolded - determine the brand of softdrinks. You put this to the test, presenting her with 6 cans (3 coke, 3 pepsi, in random order). The lady identifies all of them correctly. This statement best describes the outcome of our test (assuming alpha = 0.05):

Question 7 options:

|  |  |
| --- | --- |
|  | We conclude that the lady has this ability, as such an outcome could be expected from chance (if she does not have the ability) about 4% of the time, and 0.04<0.05. |
|  | We conclude that the lady has this ability, as such an outcome could be expected from chance (if she does not have the ability) less than 2% of the time, and 0.02<0.05. |
|  | We conclude that the lady does not have this ability, as such an outcome could be expected from chance (if she does not have the ability) less than 2% of the time, and 0.02<0.05. |
|  | We conclude that the lady does not have this ability, as such an outcome could be expected from chance (if she does not have the ability) about 50% of the time, and 0.5>0.05. |
|  | We conclude that the lady does not have this ability, as such an outcome could be expected from chance (if she does not have the ability) about 6% of the time, and 0.06>0.05. |

**Question 8**(1 point)

*Saved*

The central limit theorem (check all that apply)

Question 8 options:

|  |  |
| --- | --- |
|  | is so important because it always applies, so we don't have to worry about sampling bias any more |
|  | is the basis of inferential statistics |
|  | allows us to take care of sampling error (assuming a large enough sample that was drawn randomly and independently) |
|  | ensures that sample means of sufficiently large samples drawn randomly and independently distribute normally |

**Question 9**(1 point)

*Saved*

The results of your study are statistically significant.

Question 9 options:

|  |  |
| --- | --- |
|  | You studied a big effect |
|  | You showed that the alternative hypothesis is correct |
|  | You showed that the null hypothesis is false |
|  | You showed that the alternative hypothesis is correct |
|  | You found something important |
|  | The observed pattern of data is unlikely due to chance alone |
|  | You proved that the null hypothesis is false |

**Question 10**(1 point)

*Saved*

Assuming that you sample randomly and independently, as you increase the sample size, the distribution of sample means approaches a

Question 10 options:

|  |  |
| --- | --- |
|  | uniform distribution |
|  | standard deviation |
|  | distribution that always depends on the population distribution that the samples are drawn from. |
|  | normal distribution |
|  | gamma distribution |
|  | null distribution |
|  | cauchy distribution |

**Question 11**(1 point)

*Saved*

You run a study on the average lifespan of the arctic snow lizard. You measure the lifespan of 20 of these animals. The standard error is 1.5 years. By increasing the total sample size to 180 (assuming the standard deviation of the sample stays the same), you can expect the new standard error to be about

Question 11 options:

|  |  |
| --- | --- |
|  | 0.75 years |
|  | 3 years |
|  | 0.5 years |
|  | 1.25 years |
|  | 1.5 years |

**Question 12**(1 point)

*Saved*

You have been constipated for a week. You then consume a vial of badger fat (certified natural, food grade), and shortly afterwards, the constipation lifts. A couple of months later, when you are again severely constipated for a long time, you try the badger fat again, and again the condition abates shortly thereafter (to your great relief). Does this mean that badger fat is an effective treatment to relieve constipation?

Question 12 options:

|  |  |
| --- | --- |
|  | Yes, the proof of the pudding is in the eating. Whatever works (or heals, in this case) is correct. We have to keep an open mind. |
|  | Yes, because we replicated the effect |
|  | No |
|  | Yes, the p-value we compute from this experiment shows that the effect is statistically significant |
|  | Yes, as we experimented with the badger fat, so this is an experiment. And as we discussed in class, experiments establish causality. |

**Question 13**(1 point)

*Saved*

The null hypothesis significance testing framework is an attempt to implement and standardize a \_\_\_\_\_\_\_ framework

Question 13 options:

|  |  |
| --- | --- |
|  | confirmation |
|  | correlational |
|  | deductive |
|  | falsification |
|  | proof |

**Question 14**(1 point)

*Saved*

You run a study on the virtues of virtue. You measure the lifetime earnings of 100 virtuous people and want a good estimate of average lifetime earnings. The standard error is $5,000. By doubling the sample size (assuming the standard deviation of the sample stays the same), you can expect the new standard error to be about

Question 14 options:

|  |  |
| --- | --- |
|  | $10,000 |
|  | $2,500 |
|  | $3,535 |
|  | $3,999 |
|  | $5,000 |

**Question 15**(1 point)

*Saved*

Corporate needs you to determine how high people rate the product "Model 9b multi mixer". You sample 25 people and the standard error of the ratings is 2. This is not good enough. For an accurate estimate, the standard error needs to be 0.4. Assuming the standard deviation does not change, you need to sample - in total - this many people

Question 15 options:

|  |  |
| --- | --- |
|  | 50 |
|  | 100 |
|  | 250 |
|  | 625 |
|  | 1000 |

**Question 16**(1 point)

*Saved*

You are a professor and are creating a multiple choice test. The test has 50 questions with 5 answer choices each. A student answers 20 questions correctly. Assuming an alpha level of 0.05, is it plausible that the student achieved this result by guessing randomly?

Question 16 options:

|  |  |
| --- | --- |
|  | Yes, the p-value of this outcome is 0.5, which is consistent with chance, so we don't reject the null hypothesis that the student is just guessing. |
|  | No, the p-value of this outcome is just below alpha - 0.049 - but a win is a win, so we reject the null hypothesis that the student is just guessing. |
|  | Yes, the p-value of this outcome is 0.20, which is consistent with chance, so we don't reject the null hypothesis that the student is just guessing. |
|  | Yes, the p-value of this outcome is 0.051, which is consistent with chance, so we don't reject the null hypothesis that the student is just guessing. |
|  | No, the p-value of this outcome is vanishingly small - less than 0.001, so we can confidently reject the null hypothesis that the student is just guessing. |

Problemset

Here, we do computational questions. You can use any tool (code, simulations, calculators, venn diagrams, etc.) to solve these problems. Just do not copy the solutions from your fellow students or look them up on cheating sites.

Quiz

In the quiz portion, we ask conceptual questions

**Question 17**(1 point)

*Saved*

As you increase the size of a sample, the standard error of the mean (SEM)

Question 17 options:

|  |  |
| --- | --- |
|  | This is a trick question. It is the standard deviation of the data itself that decreases. |
|  | increases as a function of sample size |
|  | decreases as a function of the square root of the sample size |
|  | increases as a function of the log of the sample size |
|  | increases as a function of the square of the sample size |
|  | neither increases nor decreases, as SEM is independent of sample size |
|  | decreases as a function of sample size |

**Question 18**(1 point)

*Saved*

Corporate needs you to determine how much people like the movie "The Matrix". You sample 20 people and the standard error of the ratings is 1. This is not good enough. For an accurate estimate, the standard error needs to be 0.1. Assuming the standard deviation does not change, you need to sample - in total - this many people

Question 18 options:

|  |  |
| --- | --- |
|  | 200 |
|  | 180 |
|  | 5000 |
|  | 2000 |
|  | 100 |