Week 1 Quiz

Back to Week 1

7/7 points earned (100%)

Quiz passed!

1 / 1 points

1.
We want to estimate the average coffee intake of Coursera students, measured in cups of coffee. A survey of 1,000 students yields an average of 0.55 cups per day, with a standard deviation of 1 cup per day. Which of the following is **not necessarily true**?

- (selected) $\mu = 0.55, \sigma = 1$

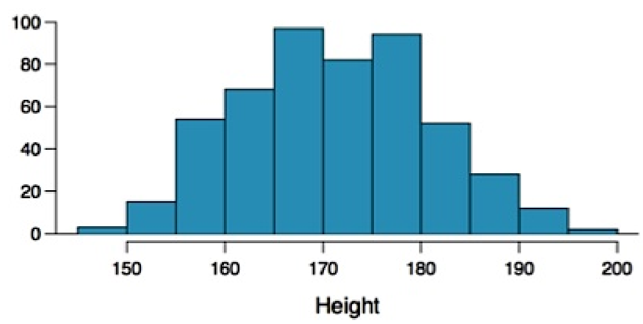
Correct Response
This question refers to the following learning objective(s): Define sample statistic as a point estimate for a population parameter, for example, the sample mean is used to estimate the population mean, and note that point estimate and sample statistic are synonymous.

Just because the sample statistics are these values doesn't mean the population values will be exactly equal to them, therefore it's not necessarily true that $\mu = 0.55, \sigma = 1$.

- The sample distribution is right skewed.
- 0.55 is a point estimate for the population mean.
- $\bar{x} = 0.55, s = 1$

1 / 1
points

2.
Researchers studying anthropometry collected various body and skeletal measurements for 507 physically active individuals. The histogram below shows the sample distribution of heights in centimeters. If the 507 individuals are a simple random sample - and let's assume they are - then the sample mean is a point estimate for the mean height of all active individuals. What measure do we use to quantify the variability of such an estimate? Compute this quantity using the data from this sample and choose the **best** answer below.



Min	147.2
Q1	163.8
Median	170.3
Mean	171.1
SD	9.4
Q3	177.8
Max	198.1

- standard error = 0.417

Correct Response
This question refers to the following learning objective(s): Calculate the sampling variability of the mean, the standard error, as $SE = \sigma/\sqrt{n}$.

We quantify variability in the sample mean by calculating the **standard error** (of the mean) $SE = \sigma/\sqrt{n}$. In this case we do not know the population standard deviation σ so in the formula we use the sample standard deviation $s = 9.4$ instead. The result is $9.4/\sqrt{507} = 0.417$.

- standard error = 0.019
- standard deviation = 0.417
- mean squared error = 0.105
- standard deviation = 0.019

1 / 1
points

3.
The standard error measures:

- the variability of population parameters
- the variability in the population
- the variability of sample statistics

Correct Response
This question refers to the following learning objective(s): Distinguish standard deviation (σ or s) and standard error (SE): standard deviation measures the variability in the data, while standard error measures the variability in point estimates from different samples of the same size and from the same population, i.e. measures the sampling variability.

- the variability of the sampled observations

---

1 / 1
points

4.
Suppose you took a large number of random samples of size n from a large population and calculated the mean of each sample. Then suppose you plotted the distribution of your sample means in a histogram. Now consider the following possible attributes of your collected data and the population from which they were sampled. For which of the following sets of attributes would you not expect your histogram of your sample means to follow a nearly normal distribution?

- n = 120. The population distribution is slightly skewed.
- n = 10. The population distribution is unknown, but the distribution of data in each sample is heavily skewed.

Correct Response
This question refers to the following learning objective(s):

Recognize that the Central Limit Theorem (CLT) is about the

Recognize that the Central Limit Theorem (CLT) is about the distribution of point estimates, and that given certain conditions, this distribution will be nearly normal.

- In the case of the mean the CLT tells us that if

(1a) the sample size is sufficiently large $(n \geq 30)$ and the data are not extremely skewed or

(1b) the population is known to have a normal distribution, and

(2) the observations in the sample are independent,

then the distribution of the sample mean will be nearly normal, centered at the true population mean and with a standard error of $\frac{\sigma}{\sqrt{n}}$.

$$\bar{x} \sim N\left(mean = \mu, SE = \frac{\sigma}{\sqrt{n}}\right)$$

- When the population distribution is unknown, condition (1a) can be checked using a histogram or some other visualization of the distribution of the observed data in the sample.
- The larger the sample size (n), the less important the shape of the distribution becomes, i.e. when n is very large the sampling distribution will be nearly normal regardless of the shape of the population distribution.

Sample size is small **and** the population distribution might be skewed, hence likely not going to yield nearly normal sampling distribution.

○ *n* = 20. The population distribution is nearly normal.

○ *n* = 120. The population distribution is unknown, but the distribution of data in each sample is slightly skewed.

---

1 / 1
points

5.

The General Social Survey (GSS) is a sociological survey used to collect data on demographic characteristics and attitudes of residents of the United States. In 2010, the survey collected responses from over a thousand US residents. The survey is conducted face-to-face with an in-person interview of a randomly-selected sample of adults. One of the questions on the survey is “For how many days during the past 30 days was your mental health, which includes stress, depression, and problems with emotions, not good?”

Based on responses from 1,151 US residents, the survey reported a 95% confidence interval of 3.40 to 4.24 days in 2010. Given this information, which of the following statements would be most appropriate to make regarding the true average number of days of “not good” mental health in 2010 for US residents?

- For all US residents in 2010, based on this 95% confidence interval, we would reject a null hypothesis stating that the true average number of days of “not good” mental health is 5 days.

Correct Response
This question refers to the following learning objective(s):

- Interpret a confidence interval as “We are XX% confident that the true population parameter is in this interval”, where XX% is the desired confidence level.
- Define margin of error as the distance required to travel in either direction away from the point estimate when constructing a confidence interval.

- For these 1,151 residents in 2010, we are 95% confident that the average number of days of “not good” mental health is between 3.40 and 4.24 days.
- For all US residents in 2010, there is a 95% probability that the true average number of days of “not good” mental health is between 3.40 and 4.24 days.
- There is not sufficient information to calculate the margin of error of this confidence interval.

---

1 / 1
points

6.

All but one of the following confidence intervals has a margin of error of 0.7.

Which is the confidence interval with the different margin of error?

- (−0.5,0.9)
- (1.6,4.4)

Correct Response

This question refers to the following learning objective(s):

- Recognize that when the sample size increases we would expect the sampling variability to decrease.
- Define margin of error as the distance required to travel in either direction away from the point estimate when constructing a confidence interval, i.e. $z^\star \times SE$.

The width of a confidence interval is 2 times the margin of error, since we add and subtract the same margin of error to the sample statistics to obtain the bounds of the confidence interval. To solve this question we need to calculate the margin of error using this rule for each choice:

$$|(1.6 - 4.4)/2| = 1.4$$

- (20.3,21.7)
- (−4.7,−3.3)

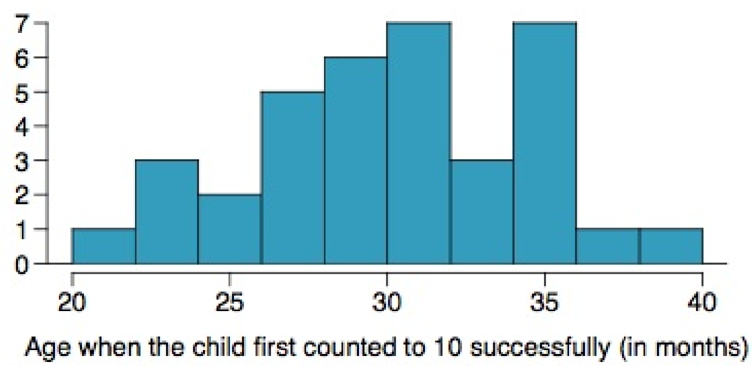
---

1 / 1
points

7.

Researchers investigating characteristics of gifted children collected data from schools in a large city on a random sample of thirty-six children who were identified as gifted children soon after they reached the age of four. The following histogram shows the distribution of the ages (in months) at which these children first counted to 10 successfully. Also provided are some sample statistics.

Calculate a 90% confidence interval for the average age at which gifted children first count to 10 successfully. Choose the closest answer.



n	36
min	21
mean	30.69
sd	4.31
max	39

- (29.28, 32.10)
- (30.12, 31.26)
- (30.49, 30.89)
- (29.50, 31.88)

Correct Response

This question refers to the following learning objective(s): Recognize that the nearly normal distribution of the point estimate (as suggested by the CLT) implies that a confidence interval can be calculated as

$$point\ estimate \pm z^{\star} \times SE,$$

where $z^{\star}$ corresponds to the cutoff points in the standard normal distribution to capture the middle XX\% of the data, where XX\% is the desired confidence level.

- For means this is: $\bar{x} \pm z^{\star} \frac{s}{\sqrt{n}}$
- Note that $z^{\star}$ is always positive.

The 90% confidence interval can be calculated as follows:

$$
\begin{aligned}
\bar{x} \pm z^{\star}\, se(\bar{x}) &= 30.69 \pm 1.65 \times \frac{4.31}{\sqrt{36}} \\
&= 30.69 \pm 1.19 \\
&= (29.50, 31.88)
\end{aligned}
$$