STAT 515 — Statistical Methods I (Total 40 points)

Quiz 2 : Lecture 3 Population vs. Sample — Solutions

6 questions · 15 minutes · Instructor answer key with explanations

1. What is the key difference between a population and a sample? (5 points)

A. Sample is the subset from which data is collected to learn about the population; population is the entire group of interest

B. They refer to the same thing in statistics

C. A population is always exactly 10 times larger than a sample

D. Sample refers only to categorical data

Correct answer: A

Explanation: *Population = the entire group for which information is wanted (e.g., all 18-year-old male college students). Sample = the subset actually collected (e.g., n=5 students) to learn about that population.*

2. A sample statistic such as the sample mean $\bar{x}$ is best described as: (5 points)

A. A fixed, known population value

B. An estimate, computed from the sample, of the corresponding population parameter

C. Always identical to the population parameter

D. A value used only to define the sampling frame

Correct answer: B

Explanation: *$\bar{x}$ and s are sample statistics — estimates of the population's true (but unknown) parameters μ and σ. They're "good" estimates when the sample is representative (e.g., randomly drawn), but they are not the parameters themselves.*

3. Why do we divide by $(n-1)$ instead of n when computing the sample variance? (5 points)

A. $\Sigma(x_i-\bar{x})^2$ tends to be smaller than $\Sigma(x_i-\mu)^2$, so dividing by a smaller number compensates

B. $n-1$ is always an even number, which simplifies calculation

C. It forces s^2 to exactly equal the population variance σ^2

D. It's purely a historical convention with no mathematical basis

Correct answer: A

Explanation: *Since we don't know μ, we use $\bar{x}$ instead — but $\Sigma(x_i-\bar{x})^2$ is generally smaller than $\Sigma(x_i-\mu)^2$ would be. Dividing by the smaller number $(n-1)$ compensates. $n-1$ is the "degrees of freedom": once you know $\bar{x}$ and $(n-1)$ of the deviations, the last deviation is determined (since all deviations sum to zero).*

4. Why is the standard deviation (s) generally preferred over the range as a measure of spread? (5 points)

A. The range is measured in different units than the original data

B. As sample size increases, the range tends to grow systematically (extreme values become more likely), while s better reflects the typical spread about the mean

C. The range is always numerically larger than s

D. s can never equal zero, unlike the range

Correct answer: B

Explanation: *As n increases, the sample maximum tends to increase and the minimum tends to decrease (extreme values become more likely with larger samples), systematically inflating the range. This makes the range an unstable, sample-size-dependent measure — unlike s, which measures spread about the mean directly.*

5. In the Mercury-in-Fish example, as the sample size grows from n=133, to n=1000, to the full population, what happens to the shape of the sample distribution (histogram)? (5 points)

A. It becomes progressively less representative of the population distribution

B. It has no systematic relationship to the population distribution

C. It increasingly resembles the shape of the underlying population distribution

D. It becomes perfectly uniform regardless of the population shape

Correct answer: C

Explanation: *This is the core takeaway of "sample distribution vs. population distribution": a random sample's distribution imperfectly mimics the population distribution, and that resemblance improves as sample size grows — visible in the fish-length histograms becoming smoother and closer to the population histogram's shape as n goes from 133 to 1000 to the full population.*

6. When we increase the sample size, which of the following is generally true. Select ALL that apply. (15 points)

A. The sample range tends to stay the same

B. The sample range tends to increase

C. The sample range tends to decrease

D. The sample standard deviation tends to increase substantially with n

E. The sample standard deviation tends to decrease substantially with n

F. The sample standard deviation remains roughly the same

G. The sample mean tends to increase

H. The sample mean remains roughly the same