question **115** views

Center and Spread Standard Deviation Sample

One of the questions asks for the standard deviation of the sample,
TWELVE COWORKERS LOG THEIR HOURS WORKED OVERTIME IN THE PAST MONTH

With the standard deviation, you need to divide by n-1 for a sample and n for a population?

comprehension_check2

11 days ago by Anonymous

the students' answer, *where students collectively construct a single answer*

just use the R function they talked about in the video
sd (x)

11 days ago by Full Name:

the instructors' answer, *where instructors collectively construct a single answer*

Yes, you answered your question correctly in your initial question.

9 days ago by Scott Carlton

followup discussions *for lingering questions and comments*

Resolved

Unresolved

panos makris 11 days ago
you will need the concat command

x <- c(10, 2, 6, 12, 14, 15, 15, 24, 15, 25, 3, 12)

Resolved

Unresolved

Karla McKee 10 days ago
Yes, you would use *n* - 1 for a sample and *n* for a population standard deviation calculation.

Jack 10 days ago Ok,that solved my problem

Stephanie 7 days ago Where is the explanation of this in the readings or videos? I see it for variance, but not for standard deviation.

Karla McKee 7 days ago You are correct, *n* and *n*-1 are used to calculate variance, then the square root of the variance provides the standard deviation:

$$s = \sqrt{\frac{\sum(x - \bar{x})^2)}{n - 1}}$$

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2)}{n}}$$

Karen West 7 days ago That solved my problem too - however - the chapter 4 provided notes do NOT say that.
They say for standard deviation you use divide by n, and it said nothing about dividing by n-1 for a sample
instead of dividing by n for a population. Where did you see that in that chapter? Perhaps I missed it.

Karla McKee 6 days ago I did not see this explicitly stated in the textbook, just inferred it from the variance discussion. Some of the subsequent standard deviation examples were a bit confusing, though, because they used the *s* symbol (the notation for sample standard deviation) but only used *n* in the calculation.

The optional Advanced High School Statistics textbook chapter provides a more explicit example on page 12 of the Chapter 2 PDF (page 48 of the textbook).

Elizabeth Rush 1 day ago I was really confused by this too. I am not sure I understand what the difference between a sample and population and when to use n and when to use n-1. I will read the optional reading again and hope that helps. Thanks for the update.

Karla McKee 6 minutes ago I find it helpful to think about the group I want to understand, then determine whether every member of that group is included in my data set. For example, if I want to know the average education level reached by adults living in a rural village, and only 1,000 adults live in that village, I might be able to ask everyone. I therefore would have the data for the entire population of interest. I would use *n* to calculate variance.

However, if I want to know the average education level reached by all adults who live in rural villages, interviewing all adults in all villages would be impracticable (and costly). I therefore might randomly select 1,000 adults from a random selection of villages. The 1,000 observations in this data set would comprise a *sample* of the population of interest. I therefore would use *n* - 1 to calculate the variance.

The idea, as I understand it, is that subtracting 1 results in a slightly larger variance than would be calculated for a population. Fewer observations usually result in greater variance, so we want to protect against the likelihood that we will interpret the subset data too narrowly to gain an accurate understanding of the whole.

Karen West Just now That was very helpful actually! I'm glad I joined this discussion.