Re: standard deviation and N−1

Source: http://sci.tech−archive.net/Archive/sci.math/2004−07/1752.html

---

From: Dimitris (*dimitris_at_dyn.unipg.it*)
Date: 07/08/04

Date: 8 Jul 2004 01:56:34 −0700

> In high school, I learned that the formula for standard deviation has n in
> the denominator, but in college the book has N−1 in the denominator. What
> is the reason for this?

Suppose sig^2_s the variance of your sample, and sig^2_t the true variance
of the population.

Then it can be worked out that the sample variance, sig^2_s has an average (over
an infinite amount of samples) equal to

sig^2_s = (N−1)/N sig^2_t

where N is the size of the sample and, as mentioned, sig^2_t is the true
variance of the population.

You can work this out from the formula for sig^2_s, and from the fact that
members within a sample are independent from each other.

Therefore, if, instead of sig^2_s, you take N/(N−1) sig^2_s, (using the 'N−1' formula) you will have a quantity whose average over many samples will be the true variance of the population.

(For a trivial case, suppose N=1. Then, obviously, sig^2_s will always be 0 and this will not tell you a thing about the true variance of the population.)

Bear in mind that this is exact only for *variances*, NOT for *standard deviations* (their square roots). But, to a good approximation, one can normally
carry it over to standard deviations.

Hope this helps,

Dimitris