# Denominator $n$ or $(n-1)$?

## ¿What is the denominator when we do a paired comparisons test?

Related to the last problem in handout #13 I am asked this question, whose answer may well be of interest to persons other than the asker. (See also "Comments on approximate tests and paired comparisons tests", in e-Gela).

We have to remember where the $t$ statistic comes from. When we want to test a hypothesis about the mean and we do not know the variance, the value of

$$Z = \frac{\overline{X} - m_0}{\sqrt{\sigma^2/n}}$$

cannot be computed (if it were, we would compare it to the appropriate quantile of the $N(0,1)$ distribution, the distribution under $H_0$).

Since this approach is barred, we have to do something else: get rid of $\sigma^2$ so that we obtain a statistic which we **can** compute and has known distribution under $H_0$. The usual procedure is to take the ratio

$$T = \frac{\frac{\overline{X}-m_0}{\sqrt{\sigma^2/n}}}{\sqrt{\frac{ns^2/\sigma^2}{n-1}}} = \frac{\overline{X} - m_0}{\sqrt{\frac{s^2}{n-1}}}$$

We know that $ns^2/\sigma^2 \sim \chi^2_{n-1}$, so the ratio above is distributed as $t_{n-1}$.

One further detail that I may not have stressed sufficiently is that $s^2$ is the máximum likelihood estimator,

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

**not** the unbiased estimator

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

When we do a paired comparisons test, the observations are the differences between the pre- and post-measures, $D_i = X_i - Y_i$, which under the null hypothesis would have mean zero. So we use the test statistic,

$$T == \frac{\overline{D} - 0}{\sqrt{\frac{s^2}{n-1}}}$$

where

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(D_i - \overline{D})^2.$$