For small sample sizes, the t-statistic is used, based on the student's t-distribution.

<u>Z statistic</u>

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{N}}$$

<u>T - statistic</u>

$$t = \frac{\bar{x} - \mu}{s / \sqrt{N-1}}$$

"s" = sample standard deviation.

$\sqrt{N-1}$ replaces $\sqrt{N}$ because s^2 is underestimating the true σ^2. Referred to as the degrees of freedom (more on how that is computed a bit later).

The t-statistic is dependent on the sample size, so as $N\uparrow$, the uncertainty decreases.

The t-distribution probability density function, unlike the bell curve, depends on the sample size.

$$f(t) = \frac{f_0(\nu)}{\left(1+\frac{t^2}{\nu}\right)^{\left(\frac{\nu+1}{2}\right)}}$$

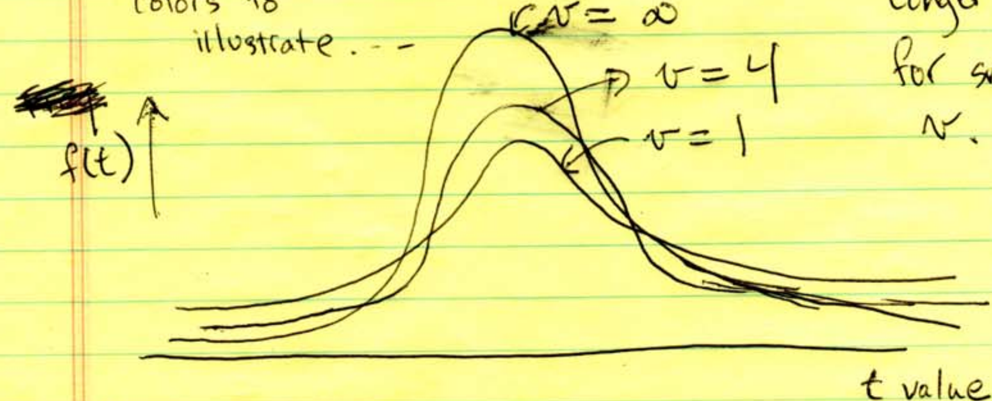
$\nu = N-1 \rightarrow$ degrees of freedom.

$t \Rightarrow t$ statistic

$f_0(\nu) \Rightarrow$ Constant that depends on ν and makes the area underneath the curve equal to unity.

Use different Colors to illustrate . . .

Longer tails for smaller ν.



As ν decreases, the distribution gets broader, reflecting the fact it is harder to get a t value near the tail of the distribution for a small sample size. Approach normal distribution for $N \sim 30$ or greater.

Student's t-table given in Appendix F of von Storch's text (or you can easily find this on line) or in Hartmann's notes.

Confidence intervals

Typically the Question being posed? What is the interval (within a specified percentage) that a given variable is different from the mean?

Normal dist. (z - statistic)

$$\mu = \bar{x} \pm z_c \frac{\sigma}{\sqrt{N}}$$

z_c = Critical z stat.

T-distribution (t-statistic)

$$\mu = \bar{x} \pm t_c \frac{s}{\sqrt{N-1}}$$

t_c = Critical t-stat. In this case, it is related to the degrees of freedom $(\nu - 1)$

12

Simple example: Mean temp.

5 Years of Jan. mean temperature

Sample mean $= -60°C = \mu$
Std. deviation $= 8°C = \sigma$ or s.

What is the 95% confidence interval?

z-stat $\mu = \bar{x} \pm z_c \frac{\sigma}{\sqrt{N}}$

$= -60°C \pm 1.96 \frac{8}{\sqrt{5}}$

$= -60°C \pm 7°C$

t-stat $\mu = \bar{x} \pm t_c \frac{s}{\sqrt{N-1}}$

$= +60 \pm 2.78 \frac{8}{\sqrt{5-1}}$

$= -60 \pm 11.1°C$
(larger!)

* Read z_c and t_c from tables.

We say that if the temperature falls outside of these ranges that it is statistically different from the mean at the given significance level.

Common significance levels

Outside Conf. interval 99%, 97.5%, 95%, and 90%

We'll talk about this more in discussion of hypothesis testing later.

13