the actual distribution depends on the sample size. This problem led him to "discover" what is called the **Student's t-distribution**. The name comes from the fact that Gosset wrote under the pen name "Student."

Up until the mid-1970s, some statisticians used the **normal distribution** approximation for large sample sizes and used the Student's t-distribution only for sample sizes of at most 30. With graphing calculators and computers, the practice now is to use the Student's t-distribution whenever *s* is used as an estimate for *σ*.

If you draw a simple random sample of size *n* from a population that has an approximately normal distribution with mean *μ* and unknown population standard deviation *σ* and calculate the *t*-score $t = \frac{\overline{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)}$, then the *t*-scores follow a **Student's t-distribution with *n* – 1 degrees of freedom**. The *t*-score has the same interpretation as the ***z*-score**. It measures how far $\overline{x}$ is from its mean *μ*. For each sample size *n*, there is a different Student's t-distribution.

The **degrees of freedom**, ***n* – 1**, come from the calculation of the sample standard deviation ***s***. This calculation requires *n* deviations ($x - \overline{x}$ values). Because the sum of the deviations is zero, we can find the last deviation once we know the other ***n* – 1** deviations. The other ***n* – 1** deviations can change or vary freely. **We call the number *n* – 1 the degrees of freedom (df).**

Properties of the Student's t-Distribution

- The graph for the Student's t-distribution is similar to the standard normal curve.
- The mean for the Student's t-distribution is zero and the distribution is symmetric about zero.
- The Student's t-distribution has more probability in its tails than the standard normal distribution because the spread of the t-distribution is greater than the spread of the standard normal. So the graph of the Student's t-distribution will be thicker in the tails and shorter in the center than the graph of the standard normal distribution.
- The exact shape of the Student's t-distribution depends on the degrees of freedom. As the degrees of freedom increases, the graph of Student's t-distribution becomes more like the graph of the standard normal distribution.
- The underlying population of individual observations is assumed to be normally distributed with unknown population mean *μ* and unknown population standard deviation *σ*. The size of the underlying population is generally not relevant unless it is very small. If it is bell shaped (normal) then the assumption is met and doesn't need discussion. Random sampling is assumed, but that is a completely separate assumption from normality.

Calculators and computers can easily calculate any Student's t-probabilities. The TI-83,83+, and 84+ have a tcdf function to find the probability for given values of *t*. The grammar for the tcdf command is tcdf(lower bound, upper bound, degrees of freedom). However for confidence intervals, we need to use **inverse** probability to find the value of *t* when we know the probability.

For the TI-84+ you can use the invT command on the DISTRibution menu. The invT command works similarly to the invnorm. The invT command requires two inputs: **invT(area to the left, degrees of freedom)** The output is the t-score that corresponds to the area we specified.

The TI-83 and 83+ do not have the invT command. (The TI-89 has an inverse T command.)

A probability table for the Student's t-distribution can also be used. The table gives t-scores that correspond to the confidence level (column) and degrees of freedom (row). (The TI-86 does not have an invT program or command, so if you are using that calculator, you need to use a probability table for the Student's t-Distribution.) When using a *t*-table, note that some tables are formatted to show the confidence level in the column headings, while the column headings in some tables may show only corresponding area in one or both tails.

A Student's t table (See Appendix H Tables) gives *t*-scores given the degrees of freedom and the right-tailed probability. The table is very limited. **Calculators and computers can easily calculate any Student's t-probabilities.**

The notation for the Student's t-distribution (using *T* as the random variable) is:

- $T \sim t_{df}$ where $df = n - 1$.
- For example, if we have a sample of size $n = 20$ items, then we calculate the degrees of freedom as $df = n - 1 = 20 - 1 = 19$ and we write the distribution as $T \sim t_{19}$.

If the population standard deviation is not known, the **error bound for a population mean** is:

- $EBM = \left(t_{\frac{\alpha}{2}}\right)\left(\frac{s}{\sqrt{n}}\right)$,
- $t_{\frac{\sigma}{2}}$ is the *t*-score with area to the right equal to $\frac{\alpha}{2}$,
- use $df = n - 1$ degrees of freedom, and