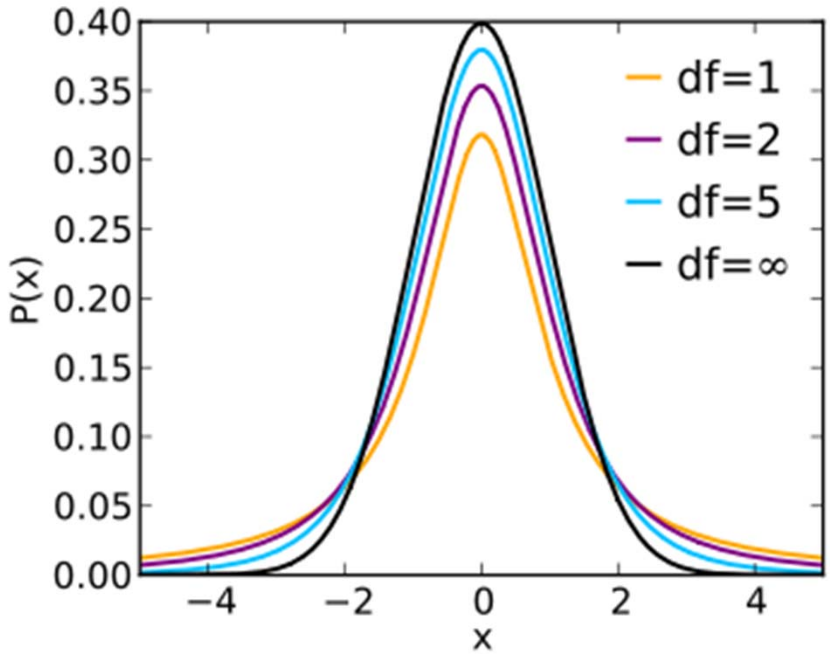
The *t*-distribution

The density of *t*-distribution is bell-shaped and depends on the number of degrees of freedom (df).



What are degrees of freedom?

Suppose we are free to choose n numbers. Then $df = n$. Suppose, for example, we have one constraint that these numbers must sum up to 5. Then we are free to choose n-1 numbers, and compute the last one by subtraction. In this case, $df = n$-1. In general, df is equal to the number of free choices minus the number of constraints.

Historical Note

William Sealy Gosset (1876 –1937) was an English statistician. He published under the pen name Student, and developed the Student's *t*-distribution in 1908.

9.3 Hypotheses Tests About μ When σ is Unknown

We are given sample size n, sample mean $\overline{x}$, and sample standard deviation

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

We test H_0: $\mu = \mu_0$ against H_1: $\mu \leq \mu_0$ or H_1: $\mu \geq \mu_0$ or H_1: $\mu \neq \mu_0$.

If

- the underlying distribution can be assumed <u>normal</u> and n is small (n<30),

or

- n is large ($n \geq 30$),

then we carry out a *t*-test based on *t*-distribution.

The **test statistic** is the observed value of

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

which under H_0 has a *t*-distribution with *n-1* degrees of freedom.

Denote by t_0 the observed value of the test statistic. The *P*-value for this test is defined as

- $P(t \geq t_0)$ if $H_1 : \mu_1 \geq \mu_2$
- $P(t \leq t_0)$ if $H_1 : \mu_1 \leq \mu_2$
- $2P(t \leq -|t_0|)$ if $H_1 : \mu_1 \neq \mu_2$.

Finally, we use the *t*-table to compare the *P*-value to α, and then we make a decision and draw conclusion.

Example. A company that manufactures light bulbs claims that its light bulbs last an average of 1150 hours. A sample of 25 light bulbs manufactured by this company gave a mean life of 1090 hours with a standard deviation of 85 hours. The significance level is 1%. A consumer research group wants to test the hypothesis that the mean life of light bulbs manufactured by this company is less than 1150 hours. Assume that the lifetime of a lightbulb is normally distributed.

Solution. Note that *n=25* is small, but it is given that the underlying distribution is normal, so we can use the *t*-test. We test $H_0\!:\ \mu = 1150$ against $H_1\!:\ \mu < 1150.$

The test statistic is $t = \dfrac{1090 - 1150}{85/\sqrt{25}} = -3.529.$

The number of degrees of freedom is *df=25-1=24.*

P-value $= P(t \leq -3.529) = P(t \geq 3.529)$ since the *t*-table gives the probability to be above. Using the *t*-table, we can find bounds for the P-value (for *df=24*):

$$P(t \geq 3.467) > P(t \geq 3.529) > P(t \geq 3.745)$$

or

$$0.001 > P(t \geq 3.529) > 0.0005.$$

Thus, P-value<0.001<0.01= α, therefore, we reject H_0, and conclude that the true mean life of light bulbs is less than 1150 hours.

Example. A researcher wants to test if the mean annual salary of all lawyers in a city is larger than \$110,000. A sample of 15 lawyers selected from this city reveals a mean annual salary of \$118,400 and a standard deviation of \$14,700. We can assume that the distribution of salary amounts is normal. Test the hypothesis at the 5% level of significance.

Solution. Note that $n=15$ is small, but it is given that the underlying distribution is normal, so we can use the *t*-test. We test $H_0\colon\ \mu = \$110{,}000$ against $H_1\colon\ \mu > \$110{,}000.$

The test statistic is $t = \dfrac{118{,}400 - 110{,}000}{14{,}700/\sqrt{15}} = 2.108.$

The number of degrees of freedom is $df=15-1=14.$

P-value = $P(t \geq 2.108)$.

From the t-table, for $df=14$,

$$2.145 \geq 2.108 \geq 1.761,$$

Thus,

$$P(t \geq 2.145) < P(t \geq 2.108) < P(t \geq 1.761),$$

or

$$0.025 < P-value < 0.05 = \alpha.$$

Therefore, we reject H_0, and conclude that the true average salary of all lawyers in a city is larger than $110,000.

Example. A company suspects that a 20-ounce coke bottles actually weigh less than that, and conducts a test of hypotheses based on a random sample of size 50 which produced a sample mean of 19.6 ounces and a sample standard deviation of 1 ounce. Carry out the test at significance level of 0.05, and draw conclusion.

Solution. Note that the sample size *n=50* is large, so we can use the *t*-test. We test

$H_0 : \mu = 20$ against $H_1 : \mu < 20.$

The test statistic is

$$t = \frac{19.6 - 20}{1/\sqrt{50}} = -2.83.$$

The number of degrees of freedom is *df=50-1 =49.* The *t*-table doesn't have 49 degrees of freedom, so we use the closest *df=40.*

P-value = $$P(t \le -2.83) = P(t \ge 2.83)$$
$$< 0.005 < 0.05 = \alpha$$

therefore, we reject H_0, and conclude that the true mean bottle content is less than 20 ounces.

Example. The mean flight delay was 15 minutes before the merger of two companies. The CEO wants to find out whether the mean flight delay has changed since the merger. A random sample of size 81 flights is drawn, and it is obtained that the sample average delay time is 14 minutes with the sample standard deviation of 5 minutes. Test the hypothesis and draw conclusion.

Solution. We test $H_0 : \mu = 15$ against $H_1 : \mu \neq 15.$ The test statistic is

$$t = \frac{14-15}{5/\sqrt{81}} = -1.8,$$

and $df=81-1=80$. The P-value is for a two-tailed test, that is,

P-value= $2P(t \leq -1.8) = 2P(t \geq 1.8)$

From the t-table, for $df=80$,

P-value > 0.05= α , therefore, we accept H_0, and conclude that the mean time of flight delays hasn't changed with the merger.