Topic 18: t Procedures

November 19, 2009

The z-score is

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

taken under the assumption that the population standard deviation is known.

If we are forced to replace the unkown σ with its unbiased estimator s, then the statistic is known as t:

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

The term $s/\sqrt{n}$ which estimate the standard deviation of the sample mean is called the **standard error**.

The remarkable discovery by William Gossett is that the distribution of the t statistic can be determined **exactly**. Moreover, the t test is a likelihood ratio test. We will save showing this result into the last section and instead focus on the applications of this widely used set of procedures.

Write

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S}.$$

Then,

- The numerator is a standard normal random variable.

- The denominator is the square root of

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

  The sum is a chi-square random variable with $n-1$ degrees of freedom.

- The numerator and denominator are **independent**.

This is called the t distribution with $n - 1$ degrees of freedom.

1 Guidelines for Using the t Procedures

- Except in the case of small samples, the assumption that the data are a simple random sample from the population of interest is more important that the population distribution is normal.

- For sample sizes less than 15, use t procedures if the data are close to normal.

- For sample sizes at least 15m use t procedures except in the presence of outliers or strong skewness.

- The t procedures can be used even for clearly skewed distributions when the sample is large.
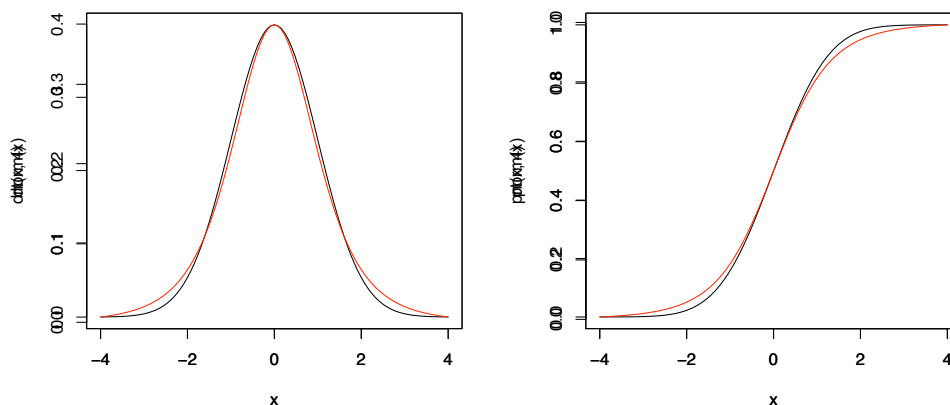
Figure 1: The density and distribution function for a standard normal random variable (black) and a t random variable with 4 degrees of freedom (red).

2   One Sample t Tests

We will later explain that the likelihood ratio test for the two sided hypothesis test

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu \neq \mu_0,$$

based on independent **normal** observations $X_1. \ldots, X_n$ with unknown mean μ and **unknown** variance σ^2 is a t-test.

So, compute the t statistic $T(\mathbf{x})$ from the data $\mathbf{x}$. Then, the critical region

$$C = \{|T(\mathbf{x})| > t_{n-1,\alpha/2}\}.$$

Where $t_{n-1,\alpha/2}$ is the upper $\alpha/2$ tail probability of the t distribution with $n-1$ degrees of freedom.

Example 1. *To check the reliability of radon detector, a university placed 12 detectors in a chamber having 105 picocuries of radon. The two sided hypothesis test*
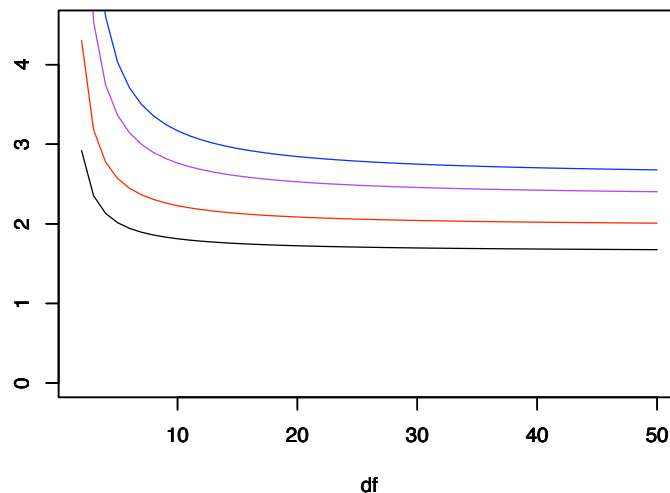
$$H_0 : \mu = 105 \quad \text{versus} \quad H_1 : \mu \neq 105,$$

The detector readings were:

91.9   97.8   111.4   122.3   105.4   95.0   103.8   99.6   96.6   119.3   104.8   101.7

Our basic issue is the loss of vitamin C contaIn R, *we find*

```
> radon<-c(91.9,97.8,111.4,122.3,105.4,95.0,103.8,99.6,96.6,119.3,104.8,101.7)
> hist(radon)
> mean(radon)
[1] 104.1333
> sd(radon)
[1] 9.39742
> length(radon)
[1] 12
> qt(0.975,11)
[1] 2.200985
```

Figure 2: Crticial values for the two sided t test with $\alpha = 0.10, 0.05.0.02$, and 0.01

Thus, the t-statistic is

$$t = \frac{105 - 104.1333}{9.39742/\sqrt{12}} = -0.3195.$$

Thus, for a 5% significance test, $|t| < 2.200985$, the critical value and we fail to reject H_0.

```
> t.test(radon,alternative=c("two.sided"),mu=105)

One Sample t-test

data:  radon
t = -0.3195, df = 11, p-value = 0.7554
alternative hypothesis: true mean is not equal to 105
95 percent confidence interval:
  98.1625 110.1042
sample estimates:
mean of x
 104.1333
```

The output also gives the 95% confidence interval

$$\bar{x} \pm \frac{s}{\sqrt{n}} t_{0.025,11}.$$

The power against an alternative of 110 is

```
> power.t.test(n=12,delta=5,sd=sd(radon),type=c("one.sample"))

    One-sample t test power calculation

            n = 12
        delta = 5
           sd = 9.39742
```

```
      sig.level = 0.05
          power = 0.3907862
    alternative = two.sided
```

If we want 80% power against an alternative of 110, then we need to make 30 measurements.

```
> power.t.test(power=0.80,delta=5,sd=sd(radon),type=c("one.sample"))

    One-sample t test power calculation

              n = 29.70383
          delta = 5
             sd = 9.39742
      sig.level = 0.05
          power = 0.8
    alternative = two.sided
```

3   Matched Pairs Procedure

A **matched pair procedure** is called for when a pair of quantitative measurements from a simple random sample

$$X_1, X_2, \ldots, X_n, \quad \text{and} \quad Y_1, Y_2, \ldots, Y_n$$

are made on the same subjects.

The null hypothesis says that the two populations are the same

$$H_0 : \mu_X = \mu_Y$$

The alternative can be either one-sided or two sided. Underlying this assumption is that the populations are the same under the null hypothesis. In particular, the standard deviations, although unknow, are equal. $\sigma_X = \sigma_Y = \sigma$.

Thus, when H_0 holds

$$E[\bar{X} - \bar{Y}] = \mu_X - \mu_Y \quad \text{and Var}(\bar{X} - \bar{Y}) = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}) = \frac{\sigma^2}{n} + \frac{\sigma^2}{n} = \frac{2\sigma^2}{n}.$$

In addition, if the data are normal, then $\bar{X} - \bar{Y}$ is also and so

$$T = \frac{\bar{X} - \bar{Y}}{S_{X-Y}/\sqrt{n}}$$

has a t distribution with $n - 1$ degrees of freedom.

The γ-level confidence interval for the difference in the population means is

$$\bar{x} - \bar{y} \pm \frac{s_{X-Y}}{n} t_{(1-\gamma)/2, n-1}.$$

Example 2. *Researchers are concerned about the impact of vitamin C content reduction due to and storage and shipment. To test this, researchers randomly chose a collection of bags of wheat soy blend bound for Haiti, marked them, and measured vitaman C from a sample of the contents. Five months later, the bags were opened and a second sample was measured for vitamin C content. The units are mg/100g.*
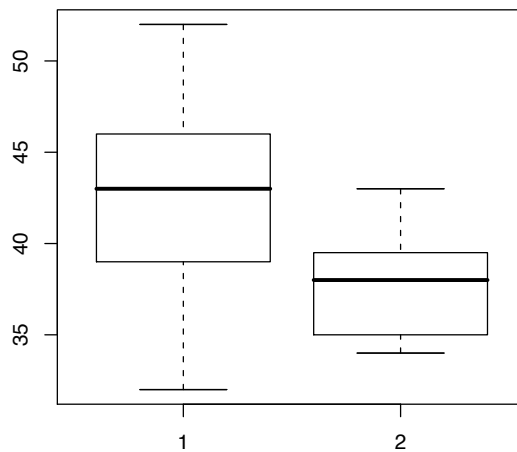
Factory	Haiti	Factory	Haiti	Factory	Haiti	Factory	Haiti
44	40	45	38	39	43	50	37
50	37	32	40	52	38	40	34
48	39	47	35	45	38	39	38
44	35	40	38	37	38	39	34
42	35	38	34	38	41	37	40
47	41	41	35	44	40	44	36
49	37	40	34	43	35		

Here is the R *output with the 95% confidence interval.*

```
> factory<-c(44,50,48,44,42,47,49,45,32,47,40,38,41,40,39,52,45,37,38,44,43,50,40,39,39,37,
> haiti<-c(40,37,39,35,35,41,37,38,40,35,38,34,35,34,43,38,38,38,41,40,35,37,34,38,34,40,36
> boxplot(factory,haiti)
> t.test(factory, haiti, alternative = c("two.sided"),mu = 0, paired = TRUE)

Paired t-test

data:  factory and haiti
t = 4.9589, df = 26, p-value = 3.745e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 3.122616 7.544050
sample estimates:
mean of the differences
              5.333333
```


Figure 3: Vitamin C content in milligrams per 100 grams, measured at the factory and measured 5 month later in Haiti.

The input

```
> t.test(factory - haiti, alternative = c("two.sided"),mu = 0)
```

gives essentially the same output.
In addition, the output

```
> t.test(haiti, alternative = c("less"),mu = 40)

One Sample t-test
```

```
data:  haiti
t = -5.3232, df = 26, p-value = 7.175e-06
alternative hypothesis: true mean is less than 40
95 percent confidence interval:
    -Inf 38.23811
sample estimates:
mean of x
 37.40741
```

shows that we would reject the one sided test

$$H_0 : \mu \geq 40 \quad versus \quad H_1 : \mu > 40,$$

based on a goal of having 40mg/100g vitamin C in the wheat consumed by the Haitians.

4   Two Sample Procedures

Now we consider the situation in which the two samples

$$X_1, X_2, \ldots, X_{n_X}, \quad \text{and} \quad Y_1, Y_2, \ldots, Y_{n_Y}$$

are not paired.

If the first sample has common mean μ_X and variance σ_X^2 and the second sample has common mean μ_Y and variance σ_Y^2, then

$$E[\bar{X} - \bar{Y}] = \mu_X - \mu_Y \quad \text{and} \quad \text{Var}(\bar{X} - \bar{Y}) = \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}.$$

For the two sided hypothesis test

$$H_0 : \mu_X = \mu_Y \quad \text{versus} \quad H_1 : \mu_X \neq \mu_Y,$$

The corresponding t-statistic is

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}}$$

with s_X^2 and s_Y^2 the unbiased sample variances. The distribution is not a t distribution. However, its tails can be approximated using an ordinary Student's t distribution with the degrees of freedom ν calculated using the **Welch-Satterthwaite** equation:

$$\nu = \frac{(s_X^2/n_X + s_Y^2/n_Y)^2}{(s_X^2/n_X)^2/(n_X - 1) + (s_Y^2/n_Y)^2/(n_Y - 1)}.$$

This also gives a γ-level confidence interval for the difference in the means μ_x and μ_Y.

$$\bar{x} - \bar{y} \pm t_{(1-\gamma)/2,\nu} \sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}.$$

Example 3. *To investigate the effect on blood pressure of added calcium in the diet, a researchers conducts a double blind randomized experiment. In the treatment group, each individual receives a calcium supplement. In the control group, the individual takes a placebo. The response variable is the decrease in systolic blood pressure, measured in millimeters of mercury, after 12 weeks. The test subjects are all male.*

```
> calcium<-c(7,-4,18,17,-3,-5,1,10,11,-2)
> mean(calcium)
[1] 5
> sd(calcium)
[1] 8.743251
> placebo<-c(-1,12,-1,-3,3,-5,5,2,-11,-1,-3)
> mean(placebo)
[1] -0.2727273
> sd(placebo)
[1] 5.900693
> boxplot(placebo,calcium)
```

    The null hypothesis is that the treatment did not reduce μ_t the mean blood pressure of the treatment any more than it did the mean μ_c for the control group. The alternative is that it did reduce blood pressure more. Formally the hypothesis test is

$$H_0 : \mu_c \leq \mu_t \quad \text{versus} \quad H_1 : \mu_c > \mu_t.$$
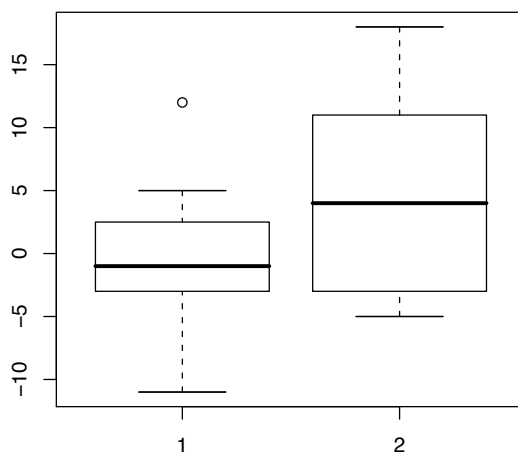
The t-statistic is

$$t = \frac{5.000 + 0.273}{\sqrt{\frac{8.743^2}{10} + \frac{5.901^2}{11}}} = 1.604.$$

```
> t.test(calcium,placebo,alternative = c("greater"))

Welch Two Sample t-test

data:  calcium and placebo
t = 1.6037, df = 15.591, p-value = 0.06442
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.476678        Inf
```

```
sample estimates:
 mean of x  mean of y
 5.0000000 -0.2727273
```

Thus. the evidence against the null hypothesis is modest with a p-value of about 6%.

To see a 90% confidence interval

```
> t.test(calcium,placebo,conf.level = 0.9)

Welch Two Sample t-test

data:  calcium and placebo
t = 1.6037, df = 15.591, p-value = 0.1288
alternative hypothesis: true difference in means is not equal to 0
90 percent confidence interval:
 -0.476678 11.022133
sample estimates:
 mean of x  mean of y
 5.0000000 -0.2727273
```

Example 4. *The life span in days of 88 wildtype and 99 transgenic mosquitoes is given in the following data set.*

```
> mosquitoes<-read.delim("http://math.arizona.edu/~jwatkins/mosquitoes.txt")
> boxplot(mosquitoes)
```
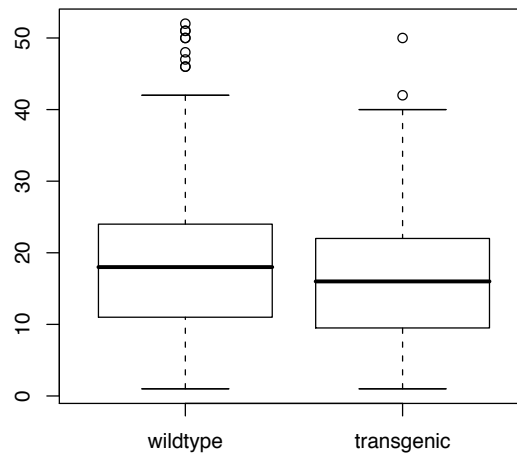

The goal is to see if overstimulation of the insulin signaling cascade in the mosquito midgut reduces the μ_t, the mean life span of these transgenic mosquitoes from that of the wild type μ_{wt}.

$$H_0 : \mu_{wt} \leq \mu_t \quad \text{versus} \quad H_1 : \mu_{wt} > \mu_t.$$

```
> wildtype<-mosquitoes[1:88,1]
> transgenic<-mosquitoes[,2]
> t.test(transgenic,wildtype,alternative = c("less"))

Welch Two Sample t-test

data:  transgenic and wildtype
t = -2.4106, df = 169.665, p-value = 0.008497
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
      -Inf -1.330591
sample estimates:
mean of x mean of y
 16.54545  20.78409

> t.test(transgenic,wildtype,conf.level=0.98)

Welch Two Sample t-test

data:  transgenic and wildtype
t = -2.4106, df = 169.665, p-value = 0.01699
alternative hypothesis: true difference in means is not equal to 0
98 percent confidence interval:
 -8.3680812 -0.1091915
sample estimates:
mean of x mean of y
 16.54545  20.78409
```

5 A Note on the Delta Method

For a one sample test hypothesizing a value for $g(\mu)$, we use the t statistic

$$t = \frac{g(\bar{x}) - g(\mu_0)}{|g'(\bar{x})| s/\sqrt{n}}$$

and base the test on the t distribution with $n - 1$ degrees of freedom.

For a test that compare a function of the mean of a two samples $g(\mu_X)$ and $g(\mu_Y)$ we can use the test statistic

$$t = \frac{g(\bar{x}) - g(\bar{y})}{\sqrt{\frac{(g'(\bar{x})s_X)^2}{n_X} + \frac{(g'(\bar{y})s_Y)^2}{n_Y}}}$$

The degrees of freedom ν can be computed from the Welch-Satterthwaite equation specialized to this circumstance.

$$\nu = \frac{(g(\bar{x})s_X)^2/n_X + (g'(\bar{y})s_Y)^2/n_Y)^2}{((g'(\bar{x})s_X)^2/n_X)^2/(n_X - 1) + ((g'(\bar{y})s_Y)^2/n_Y)^2/(n_Y - 1)}.$$

6 The t Test as a Likelihood Ratio Test

Again, we begin with independent normal observations $X_1, \ldots, X_n$ with unknown mean μ and unknown variance σ^2. The likelihood function

$$L(\mu, \sigma^2 | \mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp{-\frac{1}{2\sigma^2} \sum_{i=1}^{n}(x_i - \mu)^2}$$

$$\ln L(\mu, \sigma^2 | \mathbf{x}) = -\frac{n}{2}(\ln 2\pi + \ln \sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{\partial}{\partial \mu}\ln L(\mu, \sigma^2 | \mathbf{x}) = -\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu)$$

Thus, $\hat{\mu} = \bar{x}$.

$$\frac{\partial}{\partial \sigma^2}\ln L(\mu, \sigma^2 | \mathbf{x}) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \mu)^2.$$

Thus,

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

For the hypothesis

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu \neq \mu_0,$$

the **likelihood ratio test**

$$\Lambda(x) = \frac{L(\mu_0, \hat{\sigma}_0^2 | \mathbf{x})}{L(\hat{\mu}, \hat{\sigma}^2 | \mathbf{x})}$$

where the value

$$\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_0)^2$$

gives the maximum likelihood on the set $\mu = \mu_0$.

$$L(\mu_0, \hat{\sigma}_0^2 | \mathbf{x}) = \frac{1}{(2\pi\hat{\sigma}_0^2)^{n/2}}\exp{-\frac{1}{2\hat{\sigma}_0^2}\sum_{i=1}^{n}(x_i - \mu_0)^2} = \frac{1}{(2\pi\hat{\sigma}_0^2)^{n/2}}\exp{-\frac{2}{n}},$$

$$L(\hat{\mu}, \hat{\sigma}^2 | \mathbf{x})) = \frac{1}{(2\pi\hat{\sigma}^2)^{n/2}}\exp{-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{1}{(2\pi\hat{\sigma}^2)^{n/2}}, \exp{-\frac{2}{n}},$$

and

$$\Lambda(x) = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2} = \left(\frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{-n/2}$$

The critical region $\lambda)\mathbf{x}) \leq \lambda_0$ is equivalent to

$$c \leq \frac{\sum_{i=1}^{n}(x_i - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = 1 + \frac{n(\bar{x} - \mu_0)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

or

$$T(x)^2 \geq (c - 1)(n - 1)$$

where

$$T(x) = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

and we write s for the square root of the *unbiased* estimator of the variance

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$