Theorem 8.4. Let X_1, X_2, $\cdots$, X_n be i.i.d. $N(\mu, \sigma)$ random variables. Also, let S^2 be the standard variance for this random sample. Then, the random variable T defined as

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

has a t-distribution with $n-1$ degrees of freedom, i.e., $T \sim T(n-1)$.

Proof:
Define the random variable Z as

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}.$$

Then, $Z \sim N(0,1)$. Also, define the random variable Y as

$$Y = \frac{(n-1)S^2}{\sigma^2}.$$

Then by Theorem Theorem 8.3, $Y \sim \chi^2(n-1)$. We conclude that the random variable

$$T = \frac{Z}{\sqrt{\frac{Y}{n-1}}} = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

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

Confidence Intervals for the Mean of Normal Random Variables

Here, we assume that X_1, X_2, X_3, $\ldots$, X_n is a random sample from a normal distribution $N(\mu, \sigma^2)$, and our goal is to find an interval estimator for μ. We no longer require n to be large. Thus, n could be any positive integer. There are two possible scenarios depending on whether σ^2 is known or not. If the value of σ^2 is known, we can easily find a confidence interval for μ. This can be done using exactly the same method that we used to estimate μ for a general distribution for the case of large n. More specifically, we know that the random variable

$$Q = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

has $N(0,1)$ distribution. In particular, Q is a function of the X_i's and μ, and its distribution does not depend on μ. Thus, Q is a pivotal quantity, and we conclude that $\left[\overline{X} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$ is $(1-\alpha)100\%$ confidence interval for μ.

Assumptions: A random sample X_1, X_2, X_3, $\ldots$, X_n is given from a $N(\mu, \sigma^2)$ distribution, where $\mathrm{Var}(X_i) = \sigma^2$ is known.

Parameter to be Estimated: $\mu = EX_i$.

Confidence Interval: $\left[\overline{X} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$ is a $(1-\alpha)100\%$ confidence interval for μ.

The more interesting case is when we do not know the variance σ^2. More specifically, we are given X_1, X_2, X_3, $\ldots$, X_n, which is a random sample from a normal distribution $N(\mu, \sigma^2)$, and our goal is to find an interval estimator for μ. However, σ^2 is also unknown. In this case, using Theorem 8.4, we conclude that the random variable T defined as

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

has a t-distribution with $n-1$ degrees of freedom, i.e., $T \sim T(n-1)$. Here, the random variable T is a pivotal quantity, since it is a function of the X_i's and μ, and its distribution does not depend on μ or any other unknown parameters. Now that we have a pivot, the next step is to find a $(1-\alpha)$ interval for T. Using the definition of $t_{p,n}$, a $(1-\alpha)$ interval for T can be stated as

$$P\left(-t_{\frac{\alpha}{2},n-1} \leq T \leq t_{\frac{\alpha}{2},n-1}\right) = 1-\alpha.$$

Therefore,

$$P\left(-t_{\frac{\alpha}{2},n-1} \leq \frac{\overline{X} - \mu}{S/\sqrt{n}} \leq t_{\frac{\alpha}{2},n-1}\right) = 1-\alpha,$$

which is equivalent to

$$P\left(\overline{X} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right) = 1-\alpha.$$

We conclude that $\left[\overline{X} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}, \overline{X} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right]$ is $(1-\alpha)100\%$ confidence interval for μ.

Assumptions: A random sample X_1, X_2, X_3, $\ldots$, X_n is given from a $N(\mu, \sigma^2)$ distribution, where $\mu = EX_i$ and $\mathrm{Var}(X_i) = \sigma^2$ are unknown.

Parameter to be Estimated: $\mu = EX_i$.

Confidence Interval: $\left[\overline{X} - t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}, \overline{X} + t_{\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right]$ is a $(1-\alpha)$ confidence interval for μ.

Example 8.20

A farmer weighs 10 randomly chosen watermelons from his farm and he obtains the following values (in lbs):

$$7.72 \quad 9.58 \quad 12.38 \quad 7.77 \quad 11.27 \quad 8.80 \quad 11.10 \quad 7.80 \quad 10.17 \quad 6.00$$

Assuming that the weight is normally distributed with mean μ and and variance σ, find a 95% confidence interval for μ.

Solution

Using the data we obtain

$$\begin{aligned} \overline{X} &= 9.26, \\ S^2 &= 3.96 \end{aligned}$$

Here, $n = 10$, $\alpha = 0.05$, so we need

$$t_{0.025,9} \approx 2.262$$

The above value can be obtained in MATLAB using the command `tinv(0.975,9)`. Thus, we can obtain a 95% confidence interval for μ as