IES302 2011/2 Part II.4 Dr.Prapun

13.7. Summary: If $\overline{X}$ is the sample mean of a random sample of size n from a normal population with known variance σ^2, a $100(1-\alpha)\%$ CI on μ is given by

$$\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Like

where $z_{\alpha/2}$ is the upper $100\alpha/2$ percentage point of the standard normal distribution.

For 95% CI on μ, $z_{\alpha/2} = 1.96$

Example 13.8. From past experience, the population standard deviation of rod diameters produced by a machine has been found to be $\sigma = 0.053$ inches. For a simple random sample of $n = 30$ rods, the average diameter is found to be $\overline{x} = 1.400$ inches. What is the 95% confidence interval for the population mean, μ?

$\to$ Can apply CLT to assume $\overline{X} \sim \mathcal{N}$

$\overline{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 1.4 \pm 1.96 \times \frac{0.053}{\sqrt{30}} \Rightarrow$ between 1.381 and 1.419 inches.

13.9. Confidence Level and Precision of Estimation: The length of the $100(1-\alpha)\%$ confidence interval is

$$2 \times z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

Generally, for a fixed sample size n and standard deviation σ, the higher the confidence level, the longer the resulting CI.

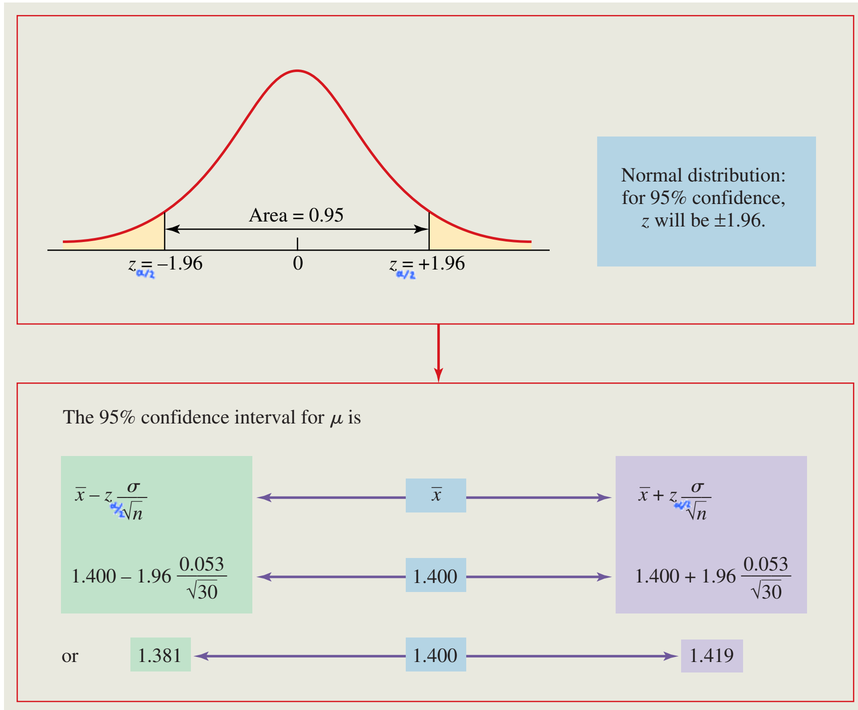


Figure 21: Construction of the 95% confidence interval

The length of a confidence interval is a measure of the ***precision*** of estimation. Observe that precision is inversely related to the confidence level. It is desirable to obtain a confidence interval that is short enough for decision-making purposes and that also has adequate confidence. One way to achieve this is by choosing the sample size n to be large enough to give a CI of specified length or precision with prescribed confidence.

13.2 Confidence Interval on the Mean of a Normal Distribution, Variance Unknown

It is rare that we know the standard deviation of a population but have no knowledge about its mean. For this reason, the techniques of the previous section are much less likely to be used in practice than those discussed here. assume this if not explicitly stated.

In this subsection, we assume that $X_1, X_2, \ldots, X_n$ is a random sample from a normal distribution with ***unknown*** mean μ and ***unknown*** variance σ^2. (Population is normally distributed.)

Let $\overline{X}$ and S^2 be the sample mean and variance, respectively.

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

$\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$

13.10. We wish to construct a CI on μ. Recall that if the variance σ^2 is known, we know that

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \quad \sim \mathcal{N}(0,1)$$

has a standard normal distribution. When σ is unknown, a logical procedure is to replace σ with the sample standard deviation S. The random variable Z now becomes

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

$\sim$ t distribution with degree of freedom $= n-1$

13.11. Question: What effect does replacing σ by S have on the distribution of the random variable T?

(a) If n is large, the answer to this question is "very little"; that is, the distribution of T will be approximately standard normal.

- In which case, we can proceed to use the confidence interval based on the normal distribution from the previous subsection.

(b) If n is small, it is better to consider the actual distribution of T which is discussed below.

Theorem 13.12. The random variable

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

has a t distribution with $n - 1$ degrees of freedom[22].

[22] The term degrees of freedom refers to the number of values that remain free to vary once some information about them is already known. For example, if four items have a mean of 10.0, and three of these items are known to have values of 8, 12, and 7, there is no choice but for the fourth item to have a value of 13. In effect, one degree of freedom has been lost.

The number of degrees of freedom associated with S^2 is the divisor $(n-1)$ used to calculate the sample variance S^2; that is, $df = n - 1$. The sample variance is the mean of the squared deviations. The number of degrees of freedom is the "number of unrelated deviations" available for use in estimating σ^2. Observe that the sum of the deviations, $\sum_i (X_i - \overline{X})$, must be zero. From a sample of size n, only the first of these deviations has freedom of value. That is, the last, or nth, value must make the sum of the n deviations total exactly zero. As a result, variance is said to average $n - 1$ unrelated squared deviation values, and this number, $n - 1$, was named "degrees of freedom."

Definition 13.13. The Student's t distribution[23] or simply the t distribution is a family of continuous, unimodal, bell-shaped distributions. It has the following properties

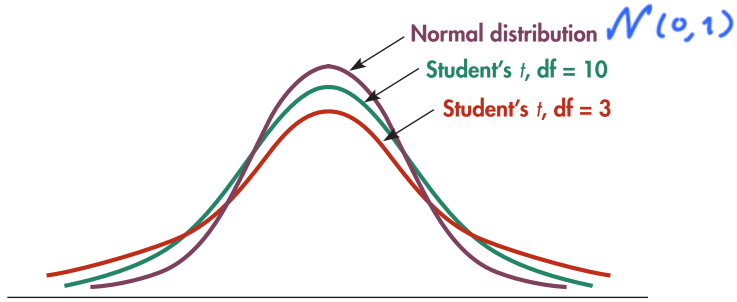


Figure 22: Student's t-Distributions (pdf)

(a) The expected value is 0

(b) The pdf is symmetric about the mean.

(c) Shape of the pdf is determined by what is called the number of ***degrees of freedom*** (df).

(d) The pdf approaches the standard normal distribution as the number of degrees of freedom increases.

(e) The variance is given by $\frac{df}{df-2}$ >1 for $df > 2$. So, the variance is greater than 1, but as the degrees of freedom increase, the variance approaches 1.

(f) The pdf is less peaked (flatter) at the mean and thicker at the tails than is the normal distribution.

13.14. As the number of degrees of freedom $k \to \infty$, the limiting form of the t distribution is the standard normal distribution.

Definition 13.15. The t-Distribution Table: A table for t values that correspond to selected areas beneath the t distribution is shown in Figure 23.

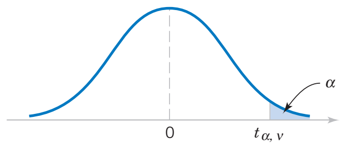
[23]It was developed in the early 1900s by W. S. Gossett, who used the pen name "Student" because his company did not permit employees to publish their research results.

For Φ table, use the whole table(s) to give $\Phi(z)$ for many values of z.

$P[Z \le z]$

For t table, use a row to give info. about $P[T > t]$ for a particular df.

t



This table gives the value of t such that $P[T > t] = \alpha$

Table V Percentage Points $t_{\alpha,\nu}$ of the t Distribution

df $=\nu$ \ α	.40	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	.289	.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	.271	.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	.265	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	.262	.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

ν = degrees of freedom.

Figure 23: The Student's t distribution table

In general, it is used in the same way as the standard normal table, but there are two exceptions:

(a) the areas provided are for the right tail only

(b) it is necessary to refer to the appropriate degrees of freedom (df) row in finding the appropriate t value.

$df = 14$

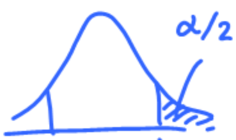
Example 13.16. For a sample size of $n = 15$, what t values would correspond to an area centered at $t = 0$ and having an area beneath the curve of 95%?

$t = 2.145.$

Excel: tinv(0.05, 14) = 2.144787

13.17. Confidence intervals using the t distribution:

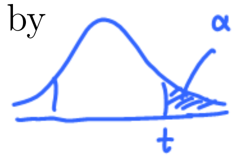
Aside from the use of the t distribution, the basic procedure for estimating confidence intervals is similar to that of the previous subsection. The appropriate t value is used instead of z, and s replaces σ.



If $\overline{X}$ and S are the sample mean and sample variance of a random sample of size n from a normal population with unknown variance σ^2, a $100(1-\alpha)\%$ CI on μ is given by

$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \xrightarrow{\sigma \text{ unknown}}$

$$\overline{X} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$$



where $t_{\alpha/2,n-1}$ is the upper $100\alpha/2$ percentage point of the t distribution with $n-1$ degree of freedom.

$n = 20 \Rightarrow df = 20 - 1 = 19$

Example 13.18. A random sample of 20 weights is taken from babies born at Northside Hospital. A mean ($\bar{x} =$) of 6.87 lb and a standard deviation ($s =$) of 1.76 lb were found for the sample. Estimate, with 95% confidence, the mean weight of all babies born in this hospital. Based on past information, it is assumed that weights of newborns are normally distributed.

$\alpha = 0.05$

$\frac{\alpha}{2} = .025$

$\bar{x} \pm \underbrace{t_{\alpha/2,n-1}}_{2.093}\frac{s}{\sqrt{n}} = 6.05 \text{ to } 7.69$

6.87   1.76   20

Excel: tinv(0.05, 19) = 2.093024

14 Tests of Hypotheses

We have seen how a parameter (e.g., the mean μ) of a population can be estimated from sample data, using either a *point estimate* or an interval of likely values called a *confidence interval*.

However, many problems in engineering require that we *decide* which of two competing claim or statements about some parameter is true. The statements are called ***hypotheses***, and the decision-making procedure is called ***hypothesis testing***.

Definition 14.1. A ***statistical hypothesis*** is a statement about the parameters of one or more populations.

- It is important to remember that hypotheses are ***not*** statements about the sample.

Definition 14.2. ***Null hypothesis*** ***(*** H_0 ***)*** vs. ***alternative hypothesis*** ***(*** H_1 ***)***

$H_0: \mu = 50$ $\quad$ $H_0: \mu = 50$ $\quad$ $H_0: \mu = 50$

$H_1: \mu \neq 50$ $\quad$ $H_1: \mu > 50$ $\quad$ $H_1: \mu < 50$

- For us, the null hypothesis will always be stated so that it specifies an exact value of the parameter.

14.3. The value of the population parameter specified in the null hypothesis is usually determined in one of three ways.

(a) It may result from past experience or knowledge of the process, or even from previous tests or experiments.

- The objective of hypothesis testing, then, is usually to determine whether the parameter value has changed.

(b) It may be determined from some theory or model regarding the process under study.

- The objective of hypothesis testing, then, is to verify the theory or model.

(c) A third situation arises when the value of the population parameter results from external considerations, such as design or engineering specifications, or from contractual obligations.

- In this situation, the usual objective of hypothesis testing is conformance testing.

Definition 14.4. A procedure leading to a decision about a particular hypothesis is called a ***test of a hypothesis***.

(a) Rely on using the information in a random sample from the population of interest.

(b) Decision:

- If the information is *consistent* with the hypothesis, we will *not reject* the hypothesis
- If the information is *inconsistent* with the hypothesis, we will conclude that the hypothesis is *false*.

(c) The truth or falsity of a particular hypothesis can never be known with certainty, unless we can examine the entire population.

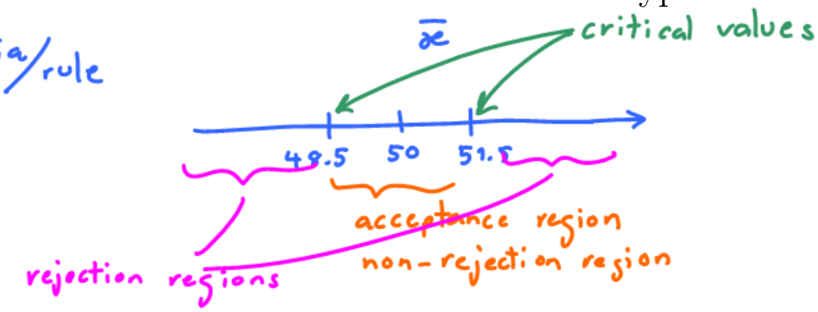
Testing the hypothesis involves taking a random sample, computing a ***test statistic*** from the sample data, and then using the test statistic to make a decision about the null hypothesis.

Example 14.5. Suppose we wish to test

$$H_0 : \mu = 50$$
$$H_1 : \mu \neq 50$$

It is customary to state conclusions relative to the null hypothesis

Decision criteria/rule



H_0. Therefore, we reject H_0 in favor of H_1 if the test statistic falls in the critical region, and fail to reject H_0 otherwise.

Definition 14.6. This decision procedure can lead to either of two wrong conclusions.

(a) ***Type I error***: Rejecting H_0 when it is true.

- The probability of making a type I error is denoted by the Greek letter α. Sometimes the type I error probability is called the ***significance level***, or the α-***error***, or the ***size*** of the test.

(b) ***Type II error***: Failing to reject H_0 when it is false.

- The probability of making a type I error is denoted by the Greek letter β. Sometimes the type II error probability is called the β-***error***.

14.7. In testing any statistical hypothesis, four different situations determine whether the final decision is correct or in error.

Reality — Healthy — Have disease — Probability $= \beta$

Decision	H_0 is true	H_0 is false
Fail to reject H_0	no error	type II error
Reject H_0	type I error	no error

"False negative"

test says you're healthy

test says you have disease

"False positive"

Probability $= \alpha$

Probability $= 1-\beta$
$=$ power

14.8. The size of the critical region, and consequently the probability of a type I error α, can be controlled by appropriate selection of the critical values.

Question: what value should be used?

Answer: the value of α should be chosen to reflect the consequences (economic, social, etc.) of incorrectly rejecting the null hypothesis. Smaller values of α would reflect more serious consequences and larger values of α would be consistent with less severe consequences. This is often hard to do, and what has evolved in much of scientific and engineering practice is to use the value $\alpha = 0.05$ in most situations,