

Institute for the Advancement of University Learning  
&  
Department of Statistics



Descriptive Statistics for Research  
(Hilary Term, 2002)

*Lecture 5: Confidence Intervals*

**(I.) Introduction**

Confidence intervals (or regions) are an important and often underused area of statistical inference, previously defined as the science of using a data sample to deduce various population attributes. As was true for our presentation of estimators in Lecture 5, the interpretation of confidence intervals and the methods of constructing them that we introduce in this lecture all come from the (classical) Frequentist paradigm. However, one should be aware that each of the other statistical paradigms (e.g., Bayesian) has its own techniques for constructing and its own interpretation of confidence intervals for population attributes.

**(II.) Definitions and Interpretation**

We can define a confidence interval (CI) as a region, constructed under the assumptions of our model, that contains the “true” value (the parameter of interest) with a specified probability. This region is constructed using particular properties of an estimator and, as we will see, is a statement regarding both the accuracy and precision of this estimate.

There are two quantities associated with confidence intervals that we need to define:

- *Coverage probability*: This term refers to the probability that a *procedure* for constructing random regions will produce an interval containing, or *covering*, the true value. It is a property of the interval producing procedure, and is independent of the particular sample to which such a procedure is applied. We can think of this quantity as the chance that the interval constructed by such a procedure will contain the parameter of interest.
- *Confidence level*: The interval produced for any particular sample, using a procedure with coverage probability  $p$ , is said to have a confidence level of  $p$ ; hence the term ‘confidence interval’. Note that, by this definition, the confidence level and coverage probability are equivalent *before* we have obtained our

sample. After, of course, the parameter is either in or not in the interval, and hence the 'chance' that the interval contains the parameter is either 0 or 1.

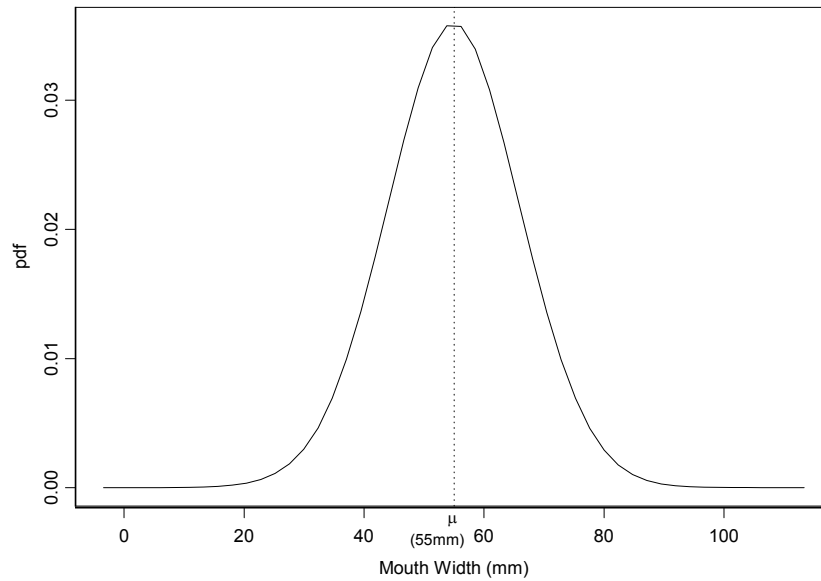
Thus, for example, 95% of 95% CIs we construct will cover the parameter under repeated sampling. Of course, for any particular sample we will not know if the CI produced contains the true value.

It is a common (and natural) mistake, when confronted with, for example, a 95% confidence interval, to say something like: "The probability that the average height of 10 year old girls lies between  $x$  and  $y$  is 0.95". Unfortunately, when we do this we are using a set of rules which we haven't been playing by! The correct interpretation, at least for Frequentist confidence intervals, requires us to invert our thinking: instead of focusing on the probability of the parameter being in the interval, we need to focus on the probability of the interval containing the parameter. The difference is subtle but important, and arises because, under Frequentist rules, parameters are regarded as fixed, unknown **constants** (i.e. they are *not* random quantities).

### **A Simulation Example**

To illustrate the above points, we will conduct a simulation. Suppose we actually know the population distribution of mouth widths of the Argentinean Wide Mouth Frog (not a real creature!) to be normal with mean 55mm and standard deviation 10mm, denoted  $N(55, 100)$ . The probability density function of this distribution is shown in figure 1. We will consider making inference on the mean mouth width ( $\mu = 55$  mm).

To simulate the inferential process, we simply draw a random sample of size  $n$  from the population using a random number generator on a computer (most statistical software packages have this capability). We pretend this is our data: each of the  $n$  points represents the carefully measured mouth width of a Wide Mouth Frog, obtained after painstakingly identifying, and then sampling from, the population of interest. From this sample we can construct  $(100 \times p)\%$  confidence intervals for the mean mouth width of the Wide Mouth Frog (we will see how to do this later in the lecture). We can then check to see if the interval contains  $\mu$ , which we know to be 55mm. By repeating this process  $N$  times, we are able to calculate the proportion of the  $N$  intervals containing  $\mu$ , and hence check that it is (roughly) equivalent to the confidence level,  $p$ .



**Figure 1:** The population distribution of the mouth width of the Argentinean Wide Mouth Frog

While we are making use of the computer, we will also repeat the procedure described above for different sample sizes ( $n = 10, 50, 100,$  and  $500$ ) and confidence levels ( $p = 0.95$  and  $0.99$ ). This will allow us to empirically determine:

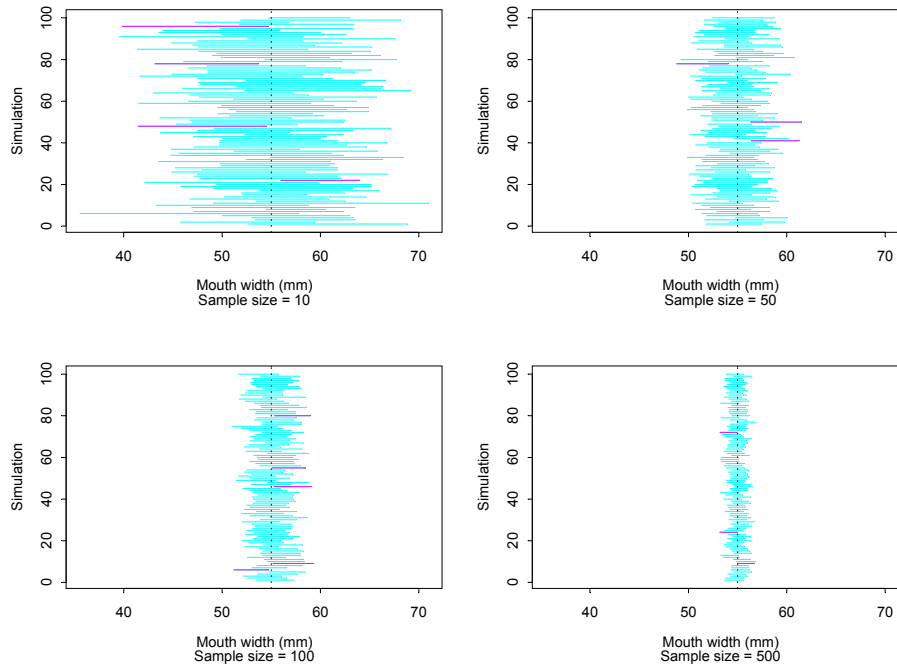
1. How, for a given confidence level, the sample size effects the width of an interval; and
2. How, for a given sample size, the confidence level effects the width of an interval.

The results of our simulation are given in table 1. The number of simulations was  $N = 1000$ . Graphical representations of the coverage are given in figures 2 and 3. Note that we use  $N = 100$  for these graphs, so that we can distinguish individual CIs.

Confidence Level (%)	Observed Coverage (%)			
	$n = 10$	$n = 50$	$n = 100$	$n = 500$
95	95.3	94.7	94.1	95.7
99	98.9	99	98.7	98.7

**Table 1:** Observed Coverage for 95- and 99- % CIs for a normal mean, with varying sample size. Figures based on  $N = 1000$  simulations

As seen in table 1, the observed coverage appears to be approximately equivalent to the confidence level, as we would expect from the definitions above. Note that the observed coverage *does not* depend on the sample size.



**Figure 2:** One hundred 95% CIs for mean mouth width of the Argentinean Wide Mouth Frog, based on samples of size 10, 50, 100, and 500. Pink (dark) lines represent CIs that do not include the true mean

Figures 2 and 3 show 100 95 - and 99 -% CIs, for each of the four sample sizes  $n = 10, 50, 100, 500$ , respectively. Each confidence interval is represented by a horizontal line. The blue (light) lines are those CIs that cover the true mean mouth width (the dotted vertical line); the pink (dark) lines are those that do not. We calculate the observed coverage from these graphs in the obvious way:

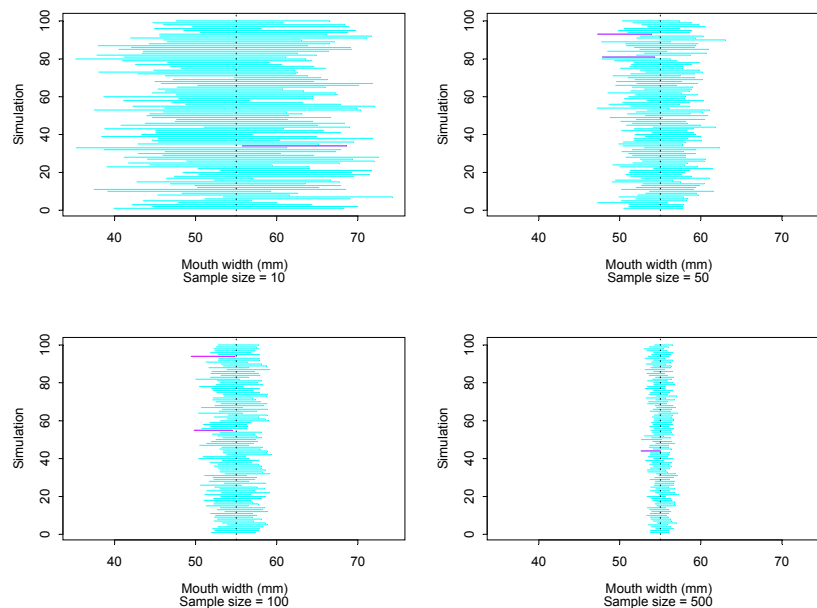
$$\frac{100 - \# \text{ pink (dark) lines}}{100}.$$

Note that this represents a measure of the *accuracy* of our estimation procedure.

We are also in a position to answer the two questions posed above:

1. For any given confidence level, an increase in sample size will yield narrower, or more *precise*, confidence intervals. We can see this phenomenon clearly in each figure: the lengths of the horizontal lines decrease dramatically as the sample size increases from 10 to 500 (the x - axis for each graph has been held fixed at the  $n=10$  range so we can see this more clearly). The reason for this will become apparent in a later section, but for now we will simply note that the length of a

CI depends on (among other things) the standard error of an appropriate estimator. As you know, the standard error has an inverse relationship with the sample size and so will *decrease* as  $n$  gets larger. At an intuitive level, a larger sample will contain more information about the population parameter of interest, and we can therefore be more precise in our estimation of it.



**Figure 3:** One hundred 99% CIs for the mean mouth width of the Argentinean Wide Mouth Frog, based on samples of size 10, 50, 100, and 500. Pink (dark) lines represent CIs that do not include the true mean

2. For any given sample size, an increase in confidence level will yield *wider* intervals. This is harder to see from the graphs, but if we look carefully at, for instance, the  $n = 50$  case, we notice that the intervals appear to lie roughly within the 50 – 60 mm range in figure 2, whereas they appear to lie slightly outside this range in figure 3. We will see the technical reason for this increase in interval width when we describe the procedure for constructing confidence intervals for a normal mean. However, we can see intuitively why this should be so: if we want to be more confident that an interval will contain the true value, we should make that interval wider. Indeed, in the extreme, a 100% CI for the mean mouth width of the Argentinean Wide Mouth Frog is  $(-\infty, \infty)$ ! In any event, the important point to note is that there is a trade – off between precision (interval length) and accuracy (coverage).

### One - Sided Confidence Intervals

In the above simulation, we constructed what are known as ‘two – sided’ confidence intervals. Two – sided confidence intervals are used when we are interested in inferring

two points between which the population quantity lies. This is almost always the case in practice.

If, however, we have very strong prior knowledge/beliefs regarding the process under investigation, we might consider constructing a “one – sided” confidence interval. These CIs are appropriate when we are interested in either an upper or lower bound for  $\mu$ , but not both. The following example illustrates the distinction.

Example: (Pagano and Gauvreau, 1993). Consider the distribution of haemoglobin levels for the population of children under the age of 6 years who have been exposed to high levels of lead. Suppose this population is normally distributed with mean  $\mu$  (and standard deviation  $\sigma$ ), on which we wish to make inference. Under usual circumstances, we simply want to ‘locate’  $\mu$ , with the focus of the inference being “between which two points does  $\mu$  lie?” In this circumstance, a 2 – sided CI is appropriate.

Suppose, however, that we have some extra information regarding the process under investigation. Specifically, suppose it is known that children with lead poisoning tend to have lower levels of haemoglobin than children who do not. We might therefore be interested in an *upper bound* on  $\mu$ . Hence, we would construct a one – sided confidence interval for  $\mu$  of the form  $(-\infty, \mu_u]$ , where  $\mu_u$  denotes this upper bound. Note that we would truncate this interval below at 0, giving  $[0, a]$ , as haemoglobin levels cannot be negative.

### (III.) Constructing Confidence Intervals

Because this is an introductory course, we will focus most of our attention on the construction of confidence intervals for means of normally distributed populations. Our approach will be to make the assumptions required by such constructions explicit. Typically, the most important assumption is that our data are independent and (approximately) normally distributed. The reason for this will be made clear presently.

Before continuing, we give a word of warning. If the assumptions of a procedure are not satisfied in a particular instance, the procedure is not valid and should not be used. However, we are free to explore ways in which to satisfy such assumptions. For example, a transformation may induce normality in non - normal data. As another example, taking differences of paired (dependent) data will yield independent observations.

For notational convenience, we introduce the quantity

$$\alpha = 1 - \text{confidence level.}$$

Rearranging, we can express this relationship, in its more usual form, as

confidence level =  $1 - \alpha$ .

We will see  $\alpha$  again when we discuss hypothesis tests. In the present context, note that it defines the probability that the confidence interval *does not* contain the true parameter, and is thus a measure of inaccuracy, or error. In fact, it is what we estimated by the proportion of pink lines in figures 2 & 3. We would obviously like  $\alpha$  to be quite small, although we should bear in mind the accuracy/precision trade-off discussed above.

### A Motivating Example

Recall that, by the central limit theorem, the sample mean,  $\bar{X}$ , can be regarded as being normally distributed with mean  $\mu$  and standard deviation  $\frac{\sigma}{\sqrt{n}}$  when  $n$  is large enough.

We can thus calculate the probability that  $\bar{X}$  lies within a certain distance of  $\mu$  (provided  $\sigma^2$ , the population variance, is known) by using the standard normal, or  $Z$ , transform. Although unlikely in most applications, for the purposes of the following illustration, we will assume we know  $\sigma^2$ .

Our example concerns the annual incomes of UK statisticians. Suppose that the distribution of annual incomes (in pounds) within this profession has a standard deviation of  $\sigma = \text{£}3250$ . For a random sample of  $n = 400$  statisticians, we would like to know the probability that the sample mean,  $\bar{X}$ , lies within  $\text{£}100$  of the true, or population, mean; that is:

$$\Pr(|\bar{X} - \mu| < 100) = ?.$$

We can think of the quantity  $|\bar{X} - \mu|$  as a measure of the error in estimating  $\mu$  by  $\bar{X}$ . The use of the absolute value reflects that error can arise because we either underestimate or overestimate the population mean. The probability statement above then becomes: "What is the probability that the (absolute) estimation error is less than  $\text{£}100$ ?"

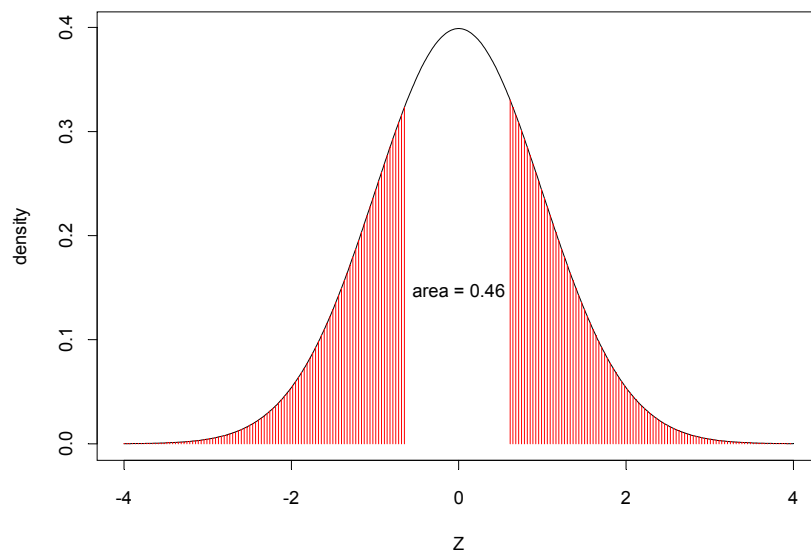
We know that  $\bar{X} \sim N(\mu, \sigma/\sqrt{n})$ . Recall from lecture 2 that we can answer probability questions regarding normally distributed random variables (in this case  $\bar{X}$ ) by using the  $Z$  transform. The  $Z$  transform for the sample mean is

$$Z = \frac{\bar{X} - \mu}{s.e.(\bar{X})} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$

The random variable  $Z$  has a standard normal distribution (i.e. has mean 0 and variance 1). We thus calculate the above probability to be

$$\begin{aligned}
 \Pr(|\bar{X} - \mu| < 100) &= \Pr\left(\left|\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right| < \frac{100}{3250/\sqrt{400}}\right) \\
 &= \Pr(|Z| < 0.615) \\
 &= \Pr(-0.615 < Z < 0.615) \\
 &= \Pr(Z < 0.615) - \Pr(Z < -0.615) \\
 &= 0.46,
 \end{aligned}$$

using standard normal tables. Figure 4 shows a standard normal pdf with the area corresponding to this probability.



**Figure 4:** Area under  $Z$  corresponding to an estimation error of £100

Note three important points:

1. The area under the standard normal curve between  $-0.615$  and  $0.615$  ( $0.46$ ) is the same as the area between  $\mu - 100$  and  $\mu + 100$  under the sampling distribution of  $\bar{X}$  (a normal distribution with mean  $\mu$  and standard deviation  $3250/\sqrt{400}$ ).
2. The population mean,  $\mu$ , is unknown.
3. If the sample size ( $400$ ) or the estimation error (£100) is made larger, the calculated probability will become larger.



We conclude that, prior to observing any data (but given that we plan to sample 400 statisticians), there is a 46% chance that the value of the estimate  $\bar{X}$  will lie at a distance no greater than £100 from the unknown quantity  $\mu$ .

Note that we have fixed the (absolute) distance of  $\bar{X}$  from  $\mu$ , and then calculated the probability of this event occurring using the properties of the sampling distribution of  $\bar{X}$ . A quick review of its definition makes it clear that to construct a confidence interval, we need to do the opposite: fix the probability (the confidence level,  $p$ ) and then determine the (absolute) distance  $\delta$  of  $\bar{X}$  from  $\mu$ . The following section examines this procedure.

### The General Construction Technique

Continuing the example, suppose we want to construct a 95% CI for the mean annual income of UK statisticians. As stated at the end of the previous section, this is equivalent to finding a value  $\delta$  such that

$$\Pr(|\bar{X} - \mu| < \delta) = 0.95.$$

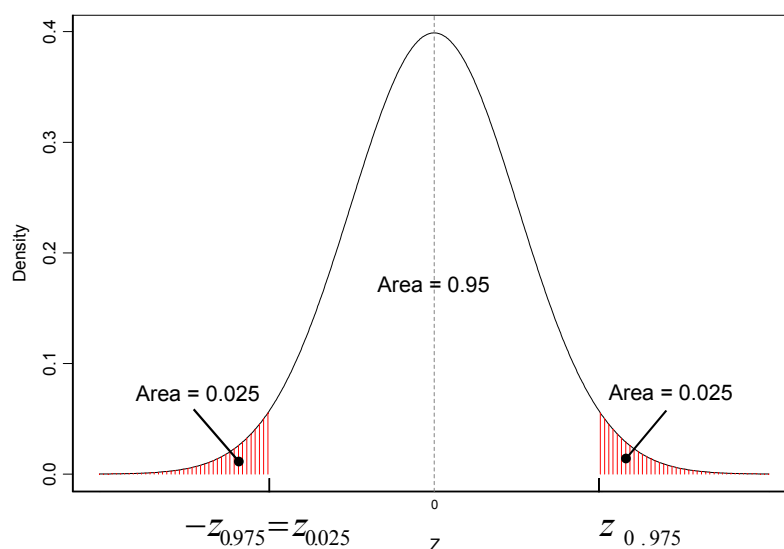
That is, we want there to be a 95% chance that  $\bar{X}$  lies *within*  $\delta$  units of  $\mu$ , and hence a 5% chance that  $\bar{X}$  lies *at least*  $\delta$  units *from*  $\mu$ . Recall that we denote the latter probability  $\alpha$ , and treat it as a measure of error. In this case, it is measuring the probability that  $\bar{X}$  does not lie within  $\delta$  units of  $\mu$ .

The procedure for determining  $\delta$  is similar to that set out in the above example. We again use the Z transform to give

$$\begin{aligned} \Pr[|\bar{X} - \mu| < \delta] &= \Pr\left[|Z| < \frac{\delta}{\sigma/\sqrt{n}}\right] \\ &= \Pr\left[\frac{-\delta}{\sigma/\sqrt{n}} < Z < \frac{\delta}{\sigma/\sqrt{n}}\right] \\ &= \Pr\left[Z < \frac{\delta}{\sigma/\sqrt{n}}\right] - \Pr\left[Z < \frac{-\delta}{\sigma/\sqrt{n}}\right] \\ &= 0.95. \end{aligned}$$

It is clear that we need to find a point  $\frac{\delta}{\sigma/\sqrt{n}}$  with the property that 95% of the area

under a standard normal curve lies between it and its negative value. This is equivalent to finding the point above which 2.5% of the probability lies: as stated in lecture 1, this value is the 0.975<sup>th</sup> quantile of  $Z$  and is denoted  $z_{0.975}$ . Because  $Z$  is symmetric about 0, we immediately know that 2.5% of the area under the standard normal curve lies to the left of  $-z_{0.975}$  (that is,  $-z_{0.975} = z_{0.025}$ ), and thus 95% of the area lies between these two points. A graphical representation of this argument is given in figure 5.



**Figure 5:** Points defining the central 95% of area under a standard normal curve

We can find  $z_{0.975}$  using standard normal tables: it is approximately 1.96. Thus

$$\frac{\delta}{\sigma/\sqrt{n}} = 1.96$$

or

$$\delta = 1.96 \times \sigma/\sqrt{n}.$$

Once  $\bar{X}$  is known (i.e. after we obtain a random sample of size  $n = 400$ ), a 95% CI for the mean annual income of UK statisticians would be given by

$$\left[ \bar{X} \pm 1.96 \times \frac{3250}{\sqrt{400}} \right].$$

This idea generalizes to accommodate any level of confidence we desire. Denoting the confidence level as  $1 - \alpha$ , and the  $(1 - \alpha/2)$ th quantile of  $Z$  as  $z_{1-\alpha/2}$ , we can write

$$\delta = z_{1-\alpha/2} \times \sigma / \sqrt{n}.$$

Hence,  $\bar{X}$  lies within  $\delta$  units of  $\mu$  with probability  $1 - \alpha$ . That is, when  $n$  is large and  $\sigma^2$  is known, a  $100 \times (1 - \alpha)$  % CI for  $\mu$  is

$$\left[ \bar{X} \pm z_{1-\alpha/2} \times \sigma / \sqrt{n} \right].$$

### Critical Values of the Standard Normal

It is convenient at this point to make a small diversion in order to discuss some important quantiles of the standard normal distribution. Such quantiles are also referred to as *critical values* of the random variable  $Z$ .

Figure 6 shows the (symmetric) values between which the standard Normal distribution accumulates 99%, 95% and 90% of probability. The corresponding intervals can be found from standard Normal tables, and are  $[-2.58, 2.58]$ ,  $[-1.96, 1.96]$  and  $[-1.645, 1.645]$  respectively. Note that if, for instance,  $1 - \alpha = 0.99$ , the area below  $-2.58$  is equal to  $\alpha/2 = 0.005$ , which is the same area above  $2.58$  by symmetry.

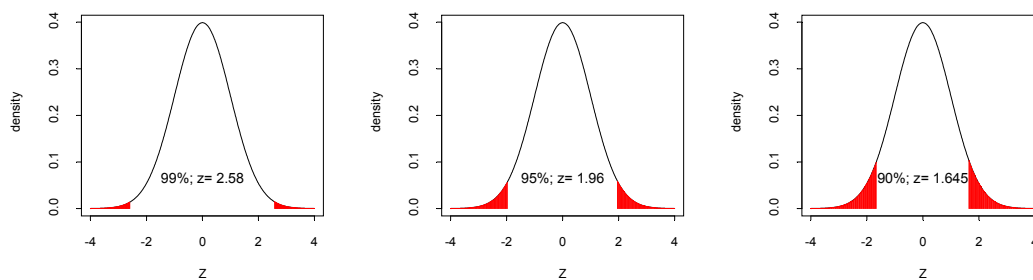


Figure 6: Some important values for the standard normal distribution

## (IV.) Specific Confidence Intervals: Confidence Intervals for Means

In the following sections we assume data have been procured via a random sample. That is, we assume observations are independent and identically distributed.

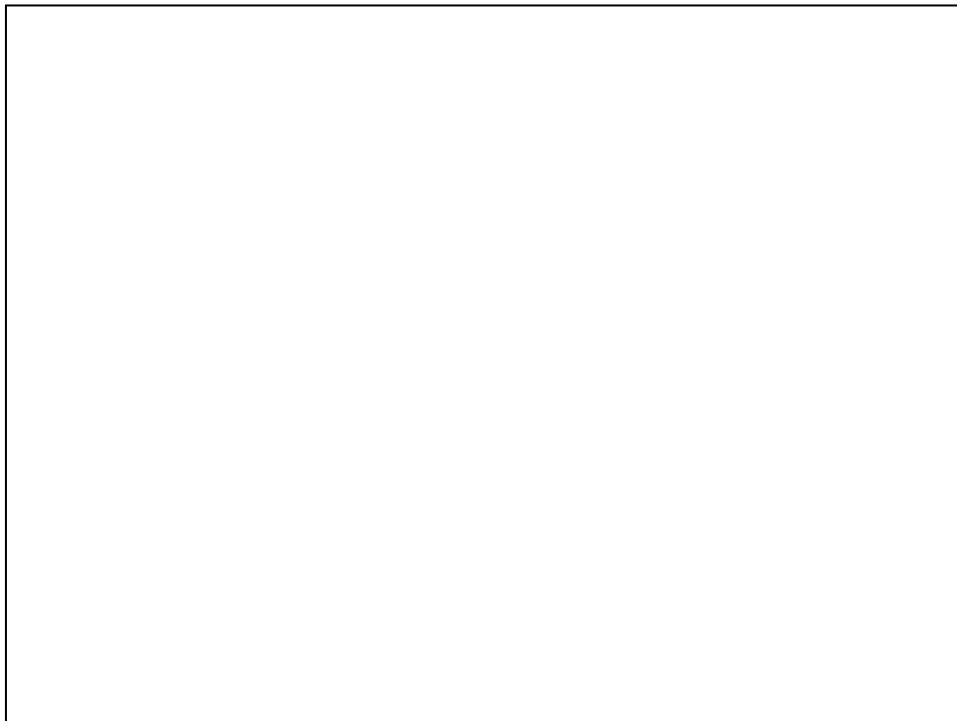
### (IV.a.) Known Variance

There are two possible situations we will consider in this section. Both yield identical expressions for constructing confidence intervals on a population mean.

The first is precisely the situation we considered in the previous example. When  $n$  is large and the population variance,  $\sigma^2$ , is known, the sampling distribution of  $\bar{X}$  is  $N(\mu, \sigma/\sqrt{n})$ , regardless of the distribution of the population from which the sample was taken. This is a direct consequence of the Central Limit Theorem.

The second situation arises when we sample from a normally distributed population. Again, given that we know  $\sigma^2$ , the sampling distribution of  $\bar{X}$  is  $N(\mu, \sigma/\sqrt{n})$ , regardless of the sample size.

### *Confidence Interval Formula*



**Example:** Consider the beard lengths (in cubits) of the population of Biblical Prophets. Suppose we know that the random variable which measures beard

length, denoted  $X$ , is distributed normally with unknown mean  $\mu$  cubits and standard deviation  $\frac{1}{2}$  cubits (i.e.  $X \sim N(\mu, \frac{1}{4})$ ). A sample of size  $n=30$  was obtained from this population, and the sample mean calculated to be  $\bar{X} = 1.3$  cubits. What is a 95% CI for the mean beard length  $\mu$  of this population, based on the information obtained from the sample?

**A:** First we need to check that either the assumption of normality is satisfied (using, for example, a true histogram or a goodness of fit test) or that the sample size is 'large'. Imagine we constructed a true histogram of the data, and it looked normal. Since we have satisfied one of the its assumptions, we may use the procedure defined above. (Note that, by a commonly asserted 'rule of thumb',  $n = 30$  classifies as being 'large' enough, so we have satisfied the other assumption as well).

Since the level of confidence required is 0.95,  $\alpha$  must be 0.05. Therefore,  $z_{1-\alpha/2} = z_{0.975} = 1.96$  (approximately) from standard normal tables. We know  $\sigma^2 = \frac{1}{4}$ ,  $n = 30$ , and  $\bar{X} = 1.3$ . Therefore, a 95% CI for mean beard lengths of Biblical Prophets is

$$\begin{aligned} \left[ \bar{X} \pm z_{1-\alpha/2} \times \frac{\sigma}{\sqrt{n}} \right] &= \left[ 1.3 \pm 1.96 \times \frac{.5}{\sqrt{30}} \right] \\ &= [1.12, 1.48]. \end{aligned}$$

### *Sample Size Calculation*

Prior to the collection of data, it would be useful to determine the sample size required to achieve a certain level of precision in our interval estimate for  $\mu$ . We are in a position to be able to write an explicit formula relating interval length (precision) to sample size. Recall that

$$\delta = z_{1-\alpha/2} \times \frac{\sigma}{\sqrt{n}}.$$

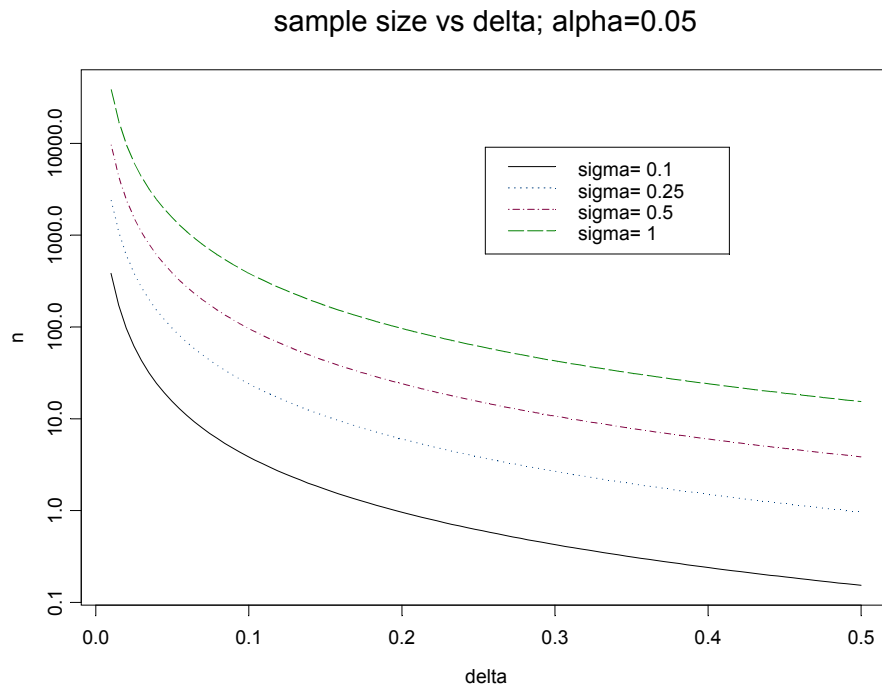
Solving this equation for  $n$ , we obtain

$$n = \left( \frac{z_{1-\alpha/2} \sigma}{\delta} \right)^2.$$

Note that  $n$  will increase if

- $\delta$  decreases (i.e. we require greater precision); or
- $\sigma$  increases (i.e. there is more dispersion in the population); or
- $\alpha$  decreases (i.e. we require greater accuracy).

Figure 7 illustrates these ideas.



**Figure 7:** Sample size as a function of  $\delta$  and  $\sigma$

**Example (continued):** Suppose that *before* we collected data from our population of hirsute Prophets, we decided that we wanted a 95% CI for mean beard length to be no greater than 0.1 cubits in length. Note that our previous 95% CI for  $\mu$ , based on a sample of size 30, was approximately 0.36 cubits in length. We are thus asking for a more precise (interval) estimate of  $\mu$ . How large should our sample size be to achieve this interval length (precision)?

**A:** We require  $\delta$  to be 0.05, because we add and subtract  $\delta$  from  $\bar{X}$  when we construct the CI: thus the width of a (two - sided) CI is  $2\delta$ . We know that  $\alpha$  is 0.05, so  $z_{1-\alpha/2} = 1.96$  as before. Again,  $\sigma^2 = 1/4$ . Therefore

$$n = \left( \frac{z_{1-\alpha/2} \sigma}{\delta} \right)^2$$

$$= \left( \frac{1.96 \times 0.5}{0.05} \right)^2 = 384.16.$$

We need a sample of at least 385 prophets in order to construct a 95% CI for mean beard length which is at most 0.1 cubits wide. Note that we *did not* require a data sample to make this calculation.

#### (IV.b.) Unknown Variance

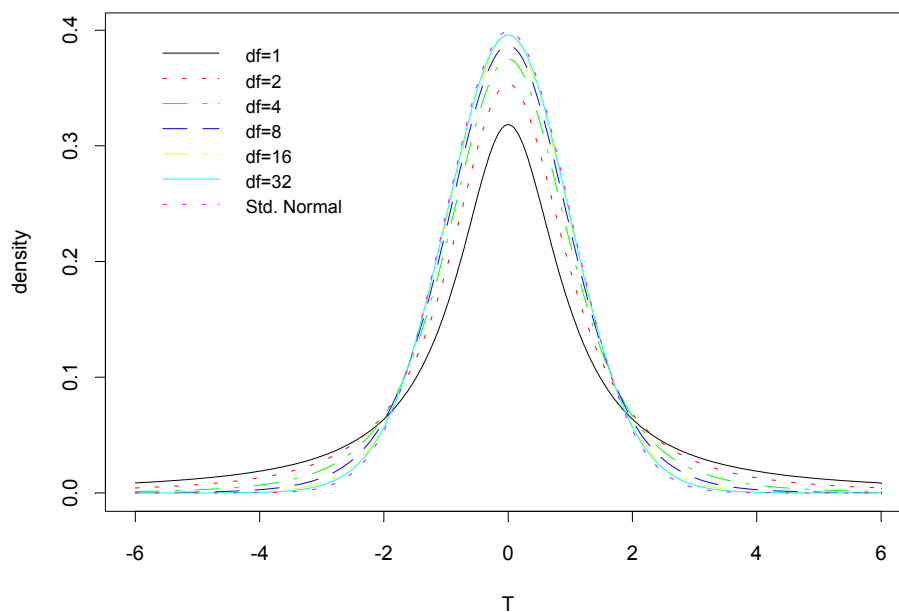
Because  $\sigma^2$  is unknown, we need to estimate it. The procedure for constructing the CI then proceeds as described above, but with  $\sigma^2$  replaced by its estimate. However, using an estimate of the population variance changes the sampling distribution of the transformed sample mean. In particular, when we use the sample variance,  $S^2$ , to estimate  $\sigma^2$ , the random variable

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

no longer has a standard normal distribution, but rather a *Student's t distribution with (n-1) degrees of freedom*, denoted  $t_{n-1}$ . Degrees of freedom will be discussed in a future lecture.

This distribution was named after its discoverer. "Student" was the pseudonym of the English statistician W.S. Gossett, who published this classic result in 1908. Gossett took a first in Chemistry from Oxford in 1899 and worked most of his adult life in the Guinness Brewery, in Dublin. It was his employer's request that he use a pseudonym in his statistical publications. He died in 1937.

Student's  $t$  distribution is symmetric with respect to 0 and has heavier tails than the standard normal pdf. It approaches the standard normal distribution as the number of degrees of freedom (and hence sample size) increase. Figure 8 shows several  $t$  pdfs together with a Normal pdf for comparison.



**Figure 8:** Student's  $t$  pdfs compared with a standard normal pdf

**Confidence Interval Formula**

When  $\sigma^2$ , the population variance, is unknown and the sample is normally distributed, a  $100(1-\alpha)\%$  CI for  $\mu$  is

$$\left[ \bar{X} \pm t_{n-1}(1-\alpha/2) \times S/\sqrt{n} \right],$$

where  $t_{n-1}(1-\alpha/2)$  denotes the  $(1-\alpha/2)$ th quantile of a Student's  $t$  random variable with  $n-1$  degrees of freedom, and

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

Critical values of Student's  $t$  distributions, for different degrees of freedom, can be found either by using statistical tables or computer software.

Note that we lose some information in not knowing  $\sigma^2$ . We thus introduce an added level of uncertainty by having to estimate this quantity. By noting that  $t_{n-1}(1-\alpha/2) \geq z_{1-\alpha/2}$ , we see that the  $t$  distribution attempts to compensate for this uncertainty by making the confidence interval wider, for given  $n$  and  $\alpha$ , than it would otherwise have been. Of course, this is not always true as the width of this CI also depends on  $S$ , the sample standard deviation, which will vary between samples.

Example:

The data given below are from two similar experiments designed to measure acceleration due to gravity. The data given are  $10^3 \times$  (measurement in cm/sec<sup>2</sup> - 980).

Experiment 1: 95 90 76 87 79 77 71

Experiment 2: 82 79 81 79 79 78 79 82 76 73 64

For this data, we can calculate the following 95% CIs:

	$df$	$\bar{X}$	$s.e.(\bar{X})$	$t_{n-1}(1-\alpha/2)$	Lower Bound	Upper Bound
<b>Experiment 1</b>	6	82.14	3.269	2.447	74.14	90.14
<b>Experiment 2</b>	10	77.45	1.557	2.228	73.98	80.92



### *Sample Size Calculation*

Unfortunately, calculating sample size in this particular instance is not as straightforward as in the previous section. An explicit formula cannot be given. However, several interactive sample size calculators exist on the web. See, for example, <http://members.aol.com/johnp71/javastat.html#Power>. Note that these procedures often make the simplifying assumption that the population variance is known.

## **(V.) Conclusion**

Confidence intervals are highly underrated and underused in many areas of scientific research. However, interval estimation is an important area of statistical inference, because a random interval simultaneously provides information regarding both estimate accuracy and precision. It is unfortunate that, due to ignorance of this area, journals prefer the ubiquitous p-value (we will deal with p-values in lecture 7).

The important points to take from this lecture are:

- The *interval* is random, *not* the parameter. Thus, we talk of the probability of the interval containing the parameter, not the probability of the parameter lying in the interval.
- The width of an interval is a measure of precision. The confidence level of an interval is a measure of accuracy.
- The width of a CI depends on two things:
  - i. the size of the estimator's standard error (which depends on the sample size); and
  - ii. the level of confidence we require (which depends on the sampling distribution of the particular statistic we use to construct the CI).
- Various formulae for CIs have been given. However, special care must be taken to ensure that the assumptions required by these formulas are satisfied. If these assumptions are not or cannot be satisfied, a different procedure must be used.

Note that we have only covered a very small subset of situations for which confidence intervals can be constructed. In particular, we have only discussed situations for which we can derive mathematical expressions for these CIs, and even then only for means of normally distributed univariate populations. The material in lecture 6 will not be presented in these seminars, but should be read as it contains ideas on what to do if we cannot derive mathematical formulae for, or satisfy the typical assumptions of, CIs for means and other population quantities (such as the median and variance).

## **(VI.) Reference**

PAGANO, M., and GAUVREAU, K. (1993) *Principles of Biostatistics*. Wadsworth Belmont California, USA.

*MCB (I-2000), KNJ (III-2001), JMcB (III-2001)*