

# Confidence intervals based on normal data

## Class 21, 18.05

### Jeremy Orloff and Jonathan Bloom

## 1 Learning Goals

1. Be able to determine whether an expression defines a valid interval statistic.
2. Be able to compute  $z$  and  $t$  confidence intervals for the mean given normal data.
3. Be able to compute the  $\chi^2$  confidence interval for the variance given normal data.
4. Be able to define the confidence level of a confidence interval.
5. Be able to explain the relationship between the  $z$  confidence interval (and confidence level) and the  $z$  non-rejection region (and significance level) in NHST.

## 2 Introduction

We continue to survey the tools of frequentist statistics. Suppose we have a model (probability distribution) for observed data with an unknown parameter. We have seen how NHST uses data to test the hypothesis that the unknown parameter has a particular value.

We have also seen how point estimates like the MLE use data to provide an estimate of the unknown parameter. On its own, a point estimate like  $\bar{x} = 2.2$  carries no information about its accuracy; it's just a single number, regardless of whether its based on ten data points or one million data points.

For this reason, statisticians augment point estimates with confidence intervals. For example, to estimate an unknown mean  $\mu$  we might be able to say that our best estimate of the mean is  $\bar{x} = 2.2$  with a 95% confidence interval  $[1.2, 3.2]$ . Another way to describe the interval is:  $\bar{x} \pm 1$ .

Here is what a 95% confidence interval means. We can use the statistic  $\bar{x}$  in a NHST for *any* null hypothesis  $\text{mean} = \mu_0$ . If  $\bar{x}$  is close to  $\mu_0$ , the NHST will tell us *not to reject* the null hypothesis (with a significance of 5%). If  $\bar{x}$  is far from  $\mu_0$ , the NHST will tell us to *reject* the null hypothesis (with a significance of 5%). The 95% confidence interval for the statistic  $\bar{x}$  consists of those hypothesized means  $\mu_0$  that would not be rejected.

Note that taken together the width of the interval and the confidence level provide a measure on the strength of the evidence supporting the hypothesis that the  $\mu$  is close to our estimate  $\bar{x}$ . The confidence level of an interval as corresponds to the significance level of a NHST. A confidence level of  $0.95 = 1 - \alpha$  is constructed using an NHST of significance level  $\alpha = 0.05$ .

We will first explore confidence intervals in situations where you will easily be able to compute by hand:  $z$  and  $t$  confidence intervals for the mean and  $\chi^2$  confidence intervals for the variance. We will use R to handle all the computations in more complicated cases. Indeed, the challenge with confidence intervals is not their computation, but rather interpreting them correctly and knowing how to use them in practice.

### 3 Interval statistics

Recall that our definition of a statistic is anything that can be computed from data. In particular, the **formula for a statistic cannot include unknown quantities.**

**Example 1.** Suppose  $x_1, \dots, x_n$  is drawn from  $N(\mu, \sigma^2)$  where  $\mu$  and  $\sigma$  are unknown.

- (i)  $\bar{x}$  and  $\bar{x} - 5$  are statistics.
- (ii)  $\bar{x} - \mu$  is not a statistic since  $\mu$  is unknown.
- (iii) If  $\mu_0$  a known value, then  $\bar{x} - \mu_0$  is a statistic. This case arises when we consider the null hypothesis  $\mu = \mu_0$ . For example, if the null hypothesis is  $\mu = 5$ , then the statistic  $\bar{x} - \mu_0$  is just  $\bar{x} - 5$  from (i).

We can play the same game with intervals to define **interval statistics**

**Example 2.** Suppose  $x_1, \dots, x_n$  is drawn from  $N(\mu, \sigma^2)$  where  $\mu$  is unknown.

- (i) The interval  $[\bar{x} - 2.2, \bar{x} + 2.2] = \bar{x} \pm 2.2$  is an interval statistic.
- (ii) If  $\sigma$  is **known**, then  $\left[\bar{x} - \frac{2\sigma}{\sqrt{n}}, \bar{x} + \frac{2\sigma}{\sqrt{n}}\right]$  is an interval statistic.
- (iii) On the other hand, if  $\sigma$  is **unknown** then  $\left[\bar{x} - \frac{2\sigma}{\sqrt{n}}, \bar{x} + \frac{2\sigma}{\sqrt{n}}\right]$  is **not an interval statistic**.
- (iv) If  $s^2$  is the sample variance, then  $\left[\bar{x} - \frac{2s}{\sqrt{n}}, \bar{x} + \frac{2s}{\sqrt{n}}\right]$  is an interval statistic because  $s^2$  is computed from the data.

We will return to (ii) and (iv), as these are respectively the  $z$  and  $t$  confidence intervals for estimating  $\mu$ .

An **interval statistic** is nothing more than a **pair of point statistics giving the lower and upper bounds of the interval**. Our reason for emphasizing that the interval is a statistic is to highlight the following:

1. The interval is **random**—new random data will produce a new interval.
2. As frequentists we are perfectly happy using it because it **doesn't depend on the value of an unknown parameter or hypothesis**.
3. **As usual with frequentist statistics we have to assume a certain hypothesis, e.g. value of  $\mu$ , before we can compute probabilities about the interval.**

**Example 3.** Suppose we draw  $n$  samples  $x_1, \dots, x_n$  from a  $N(\mu, 1)$  distribution, where  $\mu$  is unknown. Suppose we wish to know the probability that 0 is in the interval  $[\bar{x} - 2, \bar{x} + 2]$ . Without knowing the value of  $\mu$  this is impossible. However, we can compute this probability for any given (hypothesized) value of  $\mu$ .

4. **A warning which will be repeated:** Be careful in your thinking about these probabilities. Confidence intervals are a frequentist notion. Since frequentists do not compute probabilities of hypotheses, the **confidence level is never a probability that the unknown parameter is in the confidence interval**.

## 4 $z$ confidence intervals for the mean

Throughout this section we will assume that we have normally distributed data:

$$x_1, x_2, \dots, x_n \sim N(\mu, \sigma^2).$$

As we often do, we will introduce the main ideas through examples, building on what we know about rejection and non-rejection regions in NHST until we have constructed a confidence interval.

### 4.1 Definition of $z$ confidence intervals for the mean

We start with  $z$  confidence intervals for the mean. First we'll give the formula. Then we'll walk through the derivation in one entirely numerical example. This will give us the basic idea. Then we'll repeat this example, replacing the explicit numbers by symbols. Finally we'll work through a computational example.

**Definition:** Suppose the data  $x_1, \dots, x_n \sim N(\mu, \sigma^2)$ , with unknown mean  $\mu$  and known variance  $\sigma^2$ . The  $(1 - \alpha)$  confidence interval for  $\mu$  is

$$\left[ \bar{x} - \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}, \bar{x} + \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}} \right], \quad (1)$$

where  $z_{\alpha/2}$  is the **right critical value**  $P(Z > z_{\alpha/2}) = \alpha/2$ .

For example, if  $\alpha = 0.05$  then  $z_{\alpha/2} = 1.96$  so the 0.95 (or 95%) confidence interval is

$$\left[ \bar{x} - \frac{1.96\sigma}{\sqrt{n}}, \bar{x} + \frac{1.96\sigma}{\sqrt{n}} \right].$$

We've created an applet that generates normal data and displays the corresponding  $z$  confidence interval for the mean. It also shows the  $t$ -confidence interval, as discussed in the next section. Play around to get a sense for random intervals!

<http://mathlets.org/mathlets/confidence-intervals/>

**Example 4.** Suppose we collect 100 data points from a  $N(\mu, 3^2)$  distribution and the sample mean is  $\bar{x} = 12$ . Give the 95 % confidence interval for  $\mu$ .

**answer:** Using the formula this is trivial to compute: the 95% confidence interval for  $\mu$  is

$$\left[ \bar{x} - \frac{1.96\sigma}{\sqrt{n}}, \bar{x} + \frac{1.96\sigma}{\sqrt{n}} \right] = \left[ 12 - \frac{1.96 \cdot 3}{10}, 12 + \frac{1.96 \cdot 3}{10} \right]$$

### 4.2 Explaining the definition part 1: rejection regions

Our next goal is to explain the definition (1) starting from our knowledge of rejection/non-rejection regions. The phrase '**non-rejection region**' is not pretty, but we will discipline ourselves to use it instead of the inaccurate phrase 'acceptance region'.

**Example 5.** Suppose that  $n = 12$  data points are drawn from  $N(\mu, 5^2)$  where  $\mu$  is unknown. Set up a two-sided significance test of  $H_0 : \mu = 2.71$  using the statistic  $\bar{x}$  at significance level  $\alpha = 0.05$ . Describe the rejection and non-rejection regions.

**answer:** Under the null hypothesis  $\mu = 2.71$  we have  $x_i \sim N(2.71, 5^2)$  and thus

$$\bar{x} \sim N(2.71, 5^2/12)$$

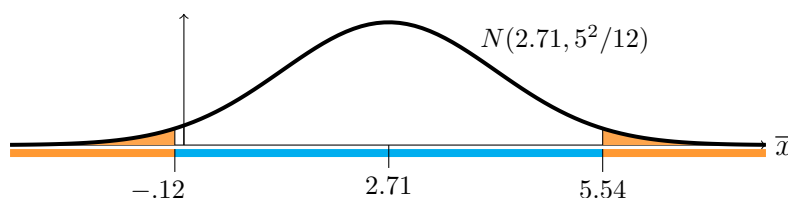
where  $5^2/12$  is the variance  $(\sigma_{\bar{x}})^2$  of  $\bar{x}$ . We know that significance  $\alpha = 0.05$  corresponds to a rejection region outside 1.96 standard deviations from the hypothesized mean. That is, the non-rejection and rejection regions are separated by the critical values  $\bar{x} \pm 1.96 \sigma_{\bar{x}}$ .

Non-rejection region:

$$\left[ 2.71 - \frac{1.96 \cdot 5}{\sqrt{12}}, 2.71 + \frac{1.96 \cdot 5}{\sqrt{12}} \right] = [-0.12, 5.54].$$

Rejection region:  $\left( -\infty, 2.71 - \frac{1.96 \cdot 5}{\sqrt{12}} \right] \cup \left[ 2.71 + \frac{1.96 \cdot 5}{\sqrt{12}}, \infty \right) = (-\infty, -0.12] \cup [5.54, \infty)$

The following figure shows the rejection and non-rejection regions for  $\bar{x}$ . The regions represent ranges of  $\bar{x}$  so they are represented by the colored bars on the  $\bar{x}$  axis. The area of the shaded region is the significance level.



The rejection (orange) and non-rejection (blue) regions for  $\bar{x}$ .

Let's redo the previous example using symbols for the known quantities as well as for  $\mu$ .

**Example 6.** Suppose that  $n$  data points are drawn from  $N(\mu, \sigma^2)$  where  $\mu$  is unknown and  $\sigma$  is known. Set up a two-sided significance test of  $H_0 : \mu = \mu_0$  using the statistic  $\bar{x}$  at significance level  $\alpha = 0.05$ . Describe the rejection and non-rejection regions.

**answer:** Under the null hypothesis  $\mu = \mu_0$  we have  $x_i \sim N(\mu_0, \sigma^2)$  and thus

$$\bar{x} \sim N(\mu_0, \sigma^2/n),$$

where  $\sigma^2/n$  is the variance  $(\sigma_{\bar{x}})^2$  of  $\bar{x}$  and  $\mu_0$ ,  $\sigma$  and  $n$  are all known values.

Let  $z_{\alpha/2}$  be the critical value:  $P(Z > z_{\alpha/2}) = \alpha/2$ . Then the non-rejection and rejection regions are separated by the values of  $\bar{x}$  that are  $z_{\alpha/2} \cdot \sigma_{\bar{x}}$  from the hypothesized mean.

Since  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$  we have

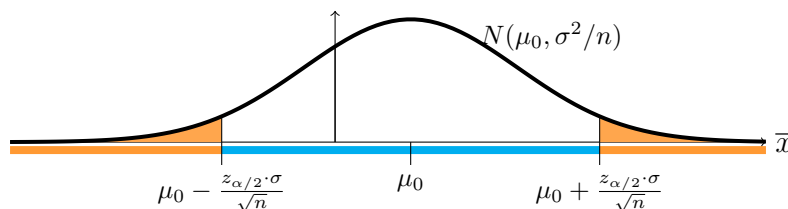
Non-rejection region:

$$\left[ \mu_0 - \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}, \mu_0 + \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}} \right] \quad (2)$$

Rejection region:

$$\left(-\infty, \mu_0 - \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}\right] \cup \left[\mu_0 + \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}, \infty\right).$$

We get the same figure as above, with the explicit numbers replaced by symbolic values.



The rejection (orange) and non-rejection (blue) regions for  $\bar{x}$ .

### 4.3 Manipulating intervals: pivoting

We need to get comfortable manipulating intervals. In general, we will make use of the type of ‘obvious’ statements that are very hard to get across. One key is to be clear about the various items.

Here is a quick summary of intervals around  $\bar{x}$  and  $\mu_0$  and what is called **pivoting**. Pivoting is the idea that  $\bar{x}$  is in  $\mu_0 \pm a$  says exactly the same thing as  $\mu_0$  is in  $\bar{x} \pm a$ .

**Example 7.** Suppose we have the sample mean  $\bar{x}$  and hypothesized mean  $\mu_0 = 2.71$ . Suppose also that the null distribution is  $N(\mu_0, 3^2)$ . Then with a significance level of 0.05 we have:

- $\mu_0 + 1.96\sigma = 2.71 + 1.96(3) = 2.71 + 5.88$  is the 0.025 critical value
- $\mu_0 - 1.96\sigma = 2.71 - 1.96(3) = 2.71 - 5.88$  is the 0.975 critical value
- The non-rejection region is centered on  $\mu_0 = 2.71$ . That is, we don’t reject  $H_0$  if  $\bar{x}$  is in the interval

$$[\mu_0 - 1.96\sigma, \mu_0 + 1.96\sigma] = [2.71 - 5.88, 2.71 + 5.88]$$

- The confidence interval is centered on  $\bar{x}$ . The 0.95 confidence interval uses the same width as the non-rejection region. It is the interval

$$[\bar{x} - 1.96\sigma, \bar{x} + 1.96\sigma] = [\bar{x} - 5.88, \bar{x} + 5.88]$$

There is a symmetry here:  $\bar{x}$  is in the interval  $[2.71 - 1.96\sigma, 2.71 + 1.96\sigma]$  is equivalent to 2.71 is in the interval  $[\bar{x} - 1.96\sigma, \bar{x} + 1.96\sigma]$ .

This symmetry is called pivoting. Here are some simple numerical examples of pivoting.

**Example 8.** (i) 1.5 is in the interval  $[0 - 2.3, 0 + 2.3]$ , so 0 is in the interval  $[1.5 - 2.3, 1.5 + 2.3]$   
(ii) Likewise 1.5 is **not** in the interval  $[0 - 1, 0 + 1]$ , so 0 is **not** in the interval  $[1.5 - 1, 1.5 + 1]$ .

The symmetry might be clearer if we talk in terms of distances: the statement

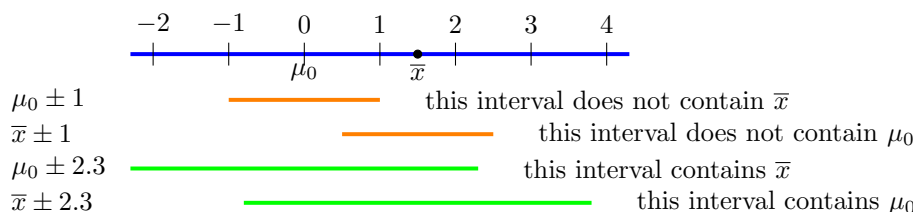
‘1.5 is in the interval  $[0 - 2.3, 0 + 2.3]$ ’

says that the distance from 1.5 to 0 is at most 2.3. Likewise, the statement

‘0 is in the interval  $[1.5 - 2.3, 1.5 + 2.3]$ ’

says exactly the same thing, i.e., the distance from 0 to 1.5 is less than 2.3.

Here is a visualization of *pivoting* from intervals around  $\mu_0$  to intervals around  $\bar{x}$ .



The distance between  $\bar{x}$  and  $\mu$  is 1.5. Now, since  $1 < 1.5$ ,  $\mu \pm 1$ , does not stretch far enough to contain  $\bar{x}$ . Likewise the interval  $\bar{x} \pm 1$  does not stretch far enough to contain  $\mu_0$ . In contrast, since  $2.3 > 1.5$ , we have  $\bar{x}$  is in the interval  $\mu_0 \pm 2.3$  and  $\mu_0$  is in the interval  $\bar{x} \pm 2.3$ .

#### 4.4 Explaining the definition part 2: translating the non-rejection region to a confidence interval

The previous examples are nice if we happen to have a null hypothesis. But what if we don't have a null hypothesis? In this case, we have the point estimate  $\bar{x}$  but we still want to use the data to estimate an interval range for the unknown mean. That is, we want an interval statistic. This is given by a confidence interval.

Here we will show how to translate the notion of a non-rejection region to that of a confidence interval. The confidence level will control the rate of certain types of errors in much the same way the significance level does for NHST.

The trick is to give a little thought to the non-rejection region. Using the numbers from Example 5 we would say that at significance level 0.05 we don't reject if

$$\bar{x} \text{ is in the interval } 2.71 \pm \frac{1.96 \cdot 5}{\sqrt{12}} = 2.71 \pm 1.96 \cdot 5/\sqrt{12}. \quad (3)$$

The roles of  $\bar{x}$  and 2.71 are symmetric. Equation (3) can be read as  $\bar{x}$  is within  $1.96 \cdot 5/\sqrt{12}$  of 2.71. This is exactly equivalent to saying that we don't reject if

$$2.71 \text{ is in the interval } \bar{x} \pm \frac{1.96 \cdot 5}{\sqrt{12}}, \quad (4)$$

i.e. 2.71 is within  $1.96 \cdot 5/\sqrt{12}$  of  $\bar{x}$ .

Now we have magically arrived at our goal of an interval statistic estimating the unknown mean. We can rewrite equation (4) as: at significance level 0.05 we don't reject if

$$\text{the interval } \left[ \bar{x} - \frac{1.96 \cdot 5}{\sqrt{12}}, \bar{x} + \frac{1.96 \cdot 5}{\sqrt{12}} \right] \text{ contains 2.71.} \quad (5)$$

Thus, different values of  $\bar{x}$  generate different intervals.

The interval in equation (5) is exactly the **confidence interval** defined in Equation (1). We make a few observations about this confidence interval.

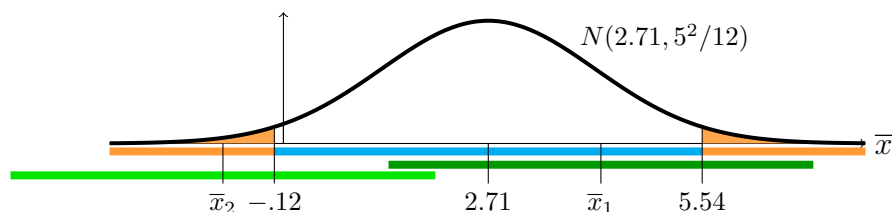
1. It only depends on  $\bar{x}$ , so it is a statistic.
2. The significance level  $\alpha = 0.05$  means that, **assuming the null hypothesis that  $\mu = 2.71$  is true**, random data will lead us to reject the null hypothesis 5% of the time (a Type I error).
3. Again **assuming** that  $\mu = 2.71$ , then 5% of the time the confidence interval will not contain 2.71, and conversely, 95% of the time it will contain 2.71.

The following figure illustrates how we don't reject  $H_0$  if the confidence interval around it contains  $\mu_0$  and we reject  $H_0$  if the confidence interval doesn't contain  $\mu_0$ . There is a lot in the figure so we will list carefully what you are seeing:

1. We started with the figure from Example 5 which shows the null distribution for  $\mu_0 = 2.71$  and the rejection and non-rejection regions.
2. We added two possible values of the statistic  $\bar{x}$ , i.e.  $\bar{x}_1$  and  $\bar{x}_2$ , and their confidence intervals. Note that the width of each interval is exactly the same as the width of the non-rejection region since both use  $\pm \frac{1.96 \cdot 5}{\sqrt{12}}$ .

The first value,  $\bar{x}_1$ , is in the non-rejection region and its interval includes the null hypothesis  $\mu_0 = 2.71$ . This illustrates that **not rejecting**  $H_0$  corresponds to the confidence interval **containing**  $\mu_0$ .

The second value,  $\bar{x}_2$ , is in the rejection region and its interval does not contain  $\mu_0$ . This illustrates that **rejecting**  $H_0$  corresponds to the confidence interval **not** containing  $\mu_0$ .



The non-rejection region (blue) and two confidence intervals (green).

We can still wring one more essential observation out of this example. Our choice of null hypothesis  $\mu = 2.71$  was completely arbitrary. If we replace  $\mu = 2.71$  by any other hypothesis  $\mu = \mu_0$  then the interval (5) will come out the same.

We call the interval (5) a 95% **confidence interval** because, **assuming  $\mu = \mu_0$** , on average it will contain  $\mu_0$  in 95% of random trials.

#### 4.5 Explaining the definition part 3: translating a general non-rejection region to a confidence interval

Note that the specific values of  $\sigma$  and  $n$  in the preceding example were of no particular consequence, so they can be replaced by their symbols. In this way we can take Example (6) quickly through the same steps as Example (5).

In words, Equation (2) and the corresponding figure say that we don't reject if

$$\bar{x} \text{ is in the interval } \mu_0 \pm \frac{z_{\alpha/2}\sigma}{\sqrt{n}}.$$

This is exactly equivalent to saying that we don't reject if

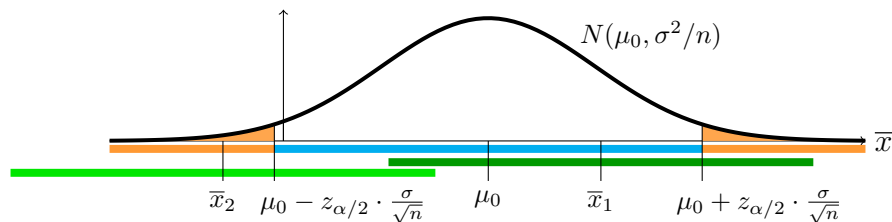
$$\mu_0 \text{ is in the interval } \bar{x} \pm \frac{z_{\alpha/2}\sigma}{\sqrt{n}}. \quad (6)$$

We can rewrite equation (6) as: at significance level  $\alpha$  we don't reject if

$$\text{the interval } \left[ \bar{x} - \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}, \bar{x} + \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}} \right] \text{ contains } \mu_0. \quad (7)$$

We call the interval (7) a  $(1 - \alpha)$  confidence interval because, assuming  $\mu = \mu_0$ , on average it will contain  $\mu_0$  in the fraction  $(1 - \alpha)$  of random trials.

The following figure illustrates the point that  $\mu_0$  is in the  $(1 - \alpha)$  confidence interval around  $\bar{x}$  is equivalent to  $\bar{x}$  is in the non-rejection region (at significance level  $\alpha$ ) for  $H_0 : \mu_0 = \mu$ .



$\bar{x}_1$  is in non-rejection region for  $\mu_0 \Leftrightarrow$  the confidence interval around  $\bar{x}_1$  contains  $\mu_0$ .

## 4.6 Computational example

**Example 9.** Suppose the data 2.5, 5.5, 8.5, 11.5 was drawn from a  $N(\mu, 10^2)$  distribution with unknown mean  $\mu$ .

(a) Compute the point estimate  $\bar{x}$  for  $\mu$  and the corresponding 50%, 80% and 95% confidence intervals.

(b) Consider the null hypothesis  $\mu = 1$ . Would you reject  $H_0$  at  $\alpha = 0.05$ ?  $\alpha = 0.20$ ?  $\alpha = 0.50$ ? Do these two ways: first by checking if the hypothesized value of  $\mu$  is in the relevant confidence interval and second by constructing a rejection region.

**answer:** (a) We compute that  $\bar{x} = 7.0$ . The critical points are

$$z_{0.025} = \text{qnorm}(0.975) = 1.96, \quad z_{0.1} = \text{qnorm}(0.9) = 1.28, \quad z_{0.25} = \text{qnorm}(0.75) = 0.67.$$

Since  $n = 4$  we have  $\bar{x} \sim N(\mu, 10^2/4)$ , i.e.  $\sigma_{\bar{x}} = 5$ . So we have:

$$\begin{aligned} 95\% \text{ conf. interval} &= [\bar{x} - z_{0.025}\sigma_{\bar{x}}, \bar{x} + z_{0.025}\sigma_{\bar{x}}] = [7 - 1.96 \cdot 5, 7 + 1.96 \cdot 5] = [-2.8, 16.8] \\ 80\% \text{ conf. interval} &= [\bar{x} - z_{0.1}\sigma_{\bar{x}}, \bar{x} + z_{0.1}\sigma_{\bar{x}}] = [7 - 1.28 \cdot 5, 7 + 1.28 \cdot 5] = [0.6, 13.4] \\ 50\% \text{ conf. interval} &= [\bar{x} - z_{0.25}\sigma_{\bar{x}}, \bar{x} + z_{0.25}\sigma_{\bar{x}}] = [7 - 0.67 \cdot 5, 7 + 0.67 \cdot 5] = [3.65, 10.35] \end{aligned}$$

Each of these intervals is a range estimate of  $\mu$ . Notice that the higher the confidence level, the wider the interval needs to be.

(b) Since  $\mu = 1$  is in the 95% and 80% confidence intervals, we would not reject the null hypothesis at the  $\alpha = 0.05$  or  $\alpha = 0.20$  levels. Since  $\mu = 1$  is not in the 50% confidence interval, we would reject  $H_0$  at the  $\alpha = 0.5$  level.



We construct the rejection regions using the same critical values as in part (a). The difference is that rejection regions are intervals centered on the hypothesized value for  $\mu$ :  $\mu_0 = 1$  and confidence intervals are centered on  $\bar{x}$ . Here are the rejection regions.

$$\begin{aligned}\alpha = 0.05 &\Rightarrow (-\infty, \mu_0 - z_{0.025}\sigma_{\bar{x}}] \cup [\mu_0 + z_{0.025}\sigma_{\bar{x}}, \infty) = (-\infty, -8.8] \cup [10.8, \infty) \\ \alpha = 0.20 &\Rightarrow (-\infty, \mu_0 - z_{0.1}\sigma_{\bar{x}}] \cup [\mu_0 + z_{0.1}\sigma_{\bar{x}}, \infty) = (-\infty, -5.4] \cup [7.4, \infty) \\ \alpha = 0.25 &\Rightarrow (-\infty, \mu_0 - z_{0.25}\sigma_{\bar{x}}] \cup [\mu_0 + z_{0.25}\sigma_{\bar{x}}, \infty) = (-\infty, -2.35] \cup [4.35, \infty)\end{aligned}$$

To do the NHST we must check whether or not  $\bar{x} = 7$  is in the rejection region.

$\alpha = 0.05$ :  $7 < 10.8$  is not in the rejection region.

We do not reject the hypothesis that  $\mu = 1$  at a significance level of 0.05.

$\alpha = 0.2$ :  $7 < 7.4$  is not in the rejection region.

We do not reject the hypothesis that  $\mu = 1$  at a significance level of 0.2.

$\alpha = 0.5$ :  $7 > 4.35$  is in the rejection region.

We reject the hypothesis that  $\mu = 1$  at a significance level 0.5.

We get the same answers using either method.

## 5 $t$ -confidence intervals for the mean

This will be nearly identical to normal confidence intervals. In this setting  $\sigma$  is not known, so we have to make the following replacements.

1. Use  $s_{\bar{x}} = \frac{s}{\sqrt{n}}$  instead of  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ . Here  $s$  is the sample variance we used before in  $t$ -tests
2. Use  $t$ -critical values instead of  $z$ -critical values.

### 5.1 Definition of $t$ -confidence intervals for the mean

**Definition:** Suppose that  $x_1, \dots, x_n \sim N(\mu, \sigma^2)$ , where the values of the mean  $\mu$  and the standard deviation  $\sigma$  are both unknown. The  $(1 - \alpha)$  confidence interval for  $\mu$  is

$$\left[ \bar{x} - \frac{t_{\alpha/2} \cdot s}{\sqrt{n}}, \bar{x} + \frac{t_{\alpha/2} \cdot s}{\sqrt{n}} \right], \quad (8)$$

here  $t_{\alpha/2}$  is the right critical value  $P(T > t_{\alpha/2}) = \alpha/2$  for  $T \sim t(n-1)$  and  $s^2$  is the sample variance of the data.

### 5.2 Construction of $t$ confidence intervals

Suppose that  $n$  data points are drawn from  $N(\mu, \sigma^2)$  where  $\mu$  and  $\sigma$  are unknown. We'll derive the  $t$  confidence interval following the same pattern as for the  $z$  confidence interval.

Under the null hypothesis  $\mu = \mu_0$ , we have  $x_i \sim N(\mu_0, \sigma^2)$ . So the studentized mean follows a Student  $t$  distribution with  $n - 1$  degrees of freedom:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n-1).$$

Let  $t_{\alpha/2}$  be the critical value:  $P(T > t_{\alpha/2}) = \alpha/2$ , where  $T \sim t(n-1)$ . We know from running one-sample  $t$ -tests that the non-rejection region is given by

$$|t| \leq t_{\alpha/2}$$

Using the definition of the  $t$ -statistic to write the rejection region in terms of  $\bar{x}$  we get: at significance level  $\alpha$  we don't reject if

$$\frac{|\bar{x} - \mu_0|}{s/\sqrt{n}} \leq t_{\alpha/2} \quad \Leftrightarrow \quad |\bar{x} - \mu_0| \leq t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}.$$

Geometrically, the right hand side says that we don't reject if

$$\mu_0 \text{ is within } t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \text{ of } \bar{x}.$$

This is exactly equivalent to saying that we don't reject if

$$\text{the interval } \left[ \bar{x} - \frac{t_{\alpha/2} \cdot s}{\sqrt{n}}, \bar{x} + \frac{t_{\alpha/2} \cdot s}{\sqrt{n}} \right] \text{ contains } \mu_0.$$

This interval is the confidence interval defined in (8).

**Example 10.** Suppose the data 2.5, 5.5, 8.5, 11.5 was drawn from a  $N(\mu, \sigma^2)$  distribution with  $\mu$  and  $\sigma$  both unknown.

Give interval estimates for  $\mu$  by finding the 95%, 80% and 50% confidence intervals.

**answer:** By direct computation we have  $\bar{x} = 7$  and  $s^2 = 15$ . The critical points are  $t_{0.025} = \text{qt}(0.975) = 3.18$ ,  $t_{0.1} = \text{qt}(0.9) = 1.64$ , and  $t_{0.25} = \text{qt}(0.75) = 0.76$ .

$$\begin{aligned} 95\% \text{ conf. interval} &= \left[ \bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}} \right] = [0.84, 13.16] \\ 80\% \text{ conf. interval} &= \left[ \bar{x} - t_{0.1} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{0.1} \cdot \frac{s}{\sqrt{n}} \right] = [3.82, 10.18] \\ 50\% \text{ conf. interval} &= \left[ \bar{x} - t_{0.25} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{0.25} \cdot \frac{s}{\sqrt{n}} \right] = [5.53, 8.47] \end{aligned}$$

All of these confidence intervals give interval estimates for the value of  $\mu$ . Again, notice that the higher the confidence level, the wider the corresponding interval.

## 6 Chi-square confidence intervals for the variance

We now turn to an interval estimate for the unknown variance.

**Definition:** Suppose the data  $x_1, \dots, x_n$  is drawn from  $N(\mu, \sigma^2)$  with mean  $\mu$  and standard deviation  $\sigma$  both unknown. The  $(1 - \alpha)$  confidence interval for the variance  $\sigma^2$  is

$$\left[ \frac{(n-1)s^2}{c_{\alpha/2}}, \frac{(n-1)s^2}{c_{1-\alpha/2}} \right]. \quad (9)$$

Here  $c_{\alpha/2}$  is the right critical value  $P(X^2 > c_{\alpha/2}) = \alpha/2$  for  $X^2 \sim \chi^2(n-1)$  and  $s^2$  is the sample variance of the data.

The derivation of this interval is nearly identical to that of the previous derivations, now starting from the chi-square test for variance. The basic fact we need is that, for data drawn from  $N(\mu, \sigma^2)$  with known  $\sigma$ , the statistic

$$\frac{(n-1)s^2}{\sigma^2}$$

follows a chi-square distribution with  $n-1$  degrees of freedom. So given the null hypothesis  $H_0 : \sigma = \sigma_0$ , the test statistic is  $(n-1)s^2/\sigma_0^2$  and the non-rejection region at significance level  $\alpha$  is

$$c_{1-\alpha/2} < \frac{(n-1)s^2}{\sigma_0^2} < c_{\alpha/2}.$$

A little algebra converts this to

$$\frac{(n-1)s^2}{c_{1-\alpha/2}} > \sigma_0^2 > \frac{(n-1)s^2}{c_{\alpha/2}}.$$

This says we don't reject if

$$\text{the interval } \left[ \frac{(n-1)s^2}{c_{\alpha/2}}, \frac{(n-1)s^2}{c_{1-\alpha/2}} \right] \text{ contains } \sigma_0^2$$

This is our  $(1-\alpha)$  confidence interval.

A difference from the  $z$  and  $t$  confidence intervals is that this one **is not exactly symmetric around the estimator  $s^2$** . The reason is that the chi-square distribution (with  $n-1$  degrees of freedom) is not symmetric around  $n-1$ .

We will continue our exploration of confidence intervals next class. In the meantime, truly the best way is to internalize the meaning of the confidence level is to experiment with the confidence interval applet:

<http://mathlets.org/mathlets/confidence-intervals/>