Confidence intervals based on normal data
Class 22, 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 χ^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 μ we might be able to say that our best estimate of the mean is $\overline{x} = 2.2$ with a 95% confidence interval $[1.2, 3.2]$. Another way to describe the interval is: $\overline{x} \pm 1$.

We will leave to later the explanation of exactly what the 95% confidence level means. For now, we'll 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 μ is close to our estimate $\overline{x}$. You should think of the confidence level of an interval as analogous to the significance level of a NHST. As explained below, it is no accident that we often see significance level $\alpha = 0.05$ and confidence level $0.95 = 1 - \alpha$.

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 χ^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 $\mathrm{N}(\mu, \sigma^2)$ where μ and σ are unknown.

(i) $\overline{x}$ and $\overline{x} - 5$ are statistics.

(ii) $\overline{x} - \mu$ is **not** a statistic since μ is unknown.

(iii) If μ_0 a known value, then $\overline{x} - \mu_0$ 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 $\overline{x} - \mu_0$ is just $\overline{x} - 5$ 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 $\mathrm{N}(\mu, \sigma^2)$ where μ is unknown.

(i) The interval $[\overline{x} - 2.2,\ \overline{x} + 2.2] = \overline{x} \pm 2.2$ is an interval statistic.

(ii) If σ is known, then $\left[\overline{x} - \frac{2\sigma}{\sqrt{n}},\ \overline{x} + \frac{2\sigma}{\sqrt{n}}\right]$ is an interval statistic.

(iii) On the other hand, if σ is unknown then $\left[\overline{x} - \frac{2\sigma}{\sqrt{n}},\ \overline{x} + \frac{2\sigma}{\sqrt{n}}\right]$ is **not** an interval statistic.

(iv) If s^2 is the sample variance, then $\left[\overline{x} - \frac{2s}{\sqrt{n}},\ \overline{x} + \frac{2s}{\sqrt{n}}\right]$ 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 μ.

Technically 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 μ, before we can compute probabilities about the interval.

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

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 specific confidence interval computed from the given data**.

4 z confidence intervals for the mean

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

$$x_1, x_2, \ldots, x_n \sim \mathrm{N}(\mu, \sigma^2).$$

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 $x_1, \ldots, x_n \sim \mathrm{N}(\mu, \sigma^2)$, with unknown mean μ and known variance σ^2. The $(1-\alpha)$ confidence interval for μ is

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

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

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

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!

https://mathlets.org/mathlets/confidence-intervals/

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

Solution: Using formula 1, this is trivial to compute: the 95% confidence interval for μ is

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

4.2 Explaining the definition part 1: non-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 inacurate phrase 'acceptance region'.

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

Solution: Under the null hypothesis ($\mu = 2.71$) we have

$$z = \frac{\overline{x} - 2.71}{5/\sqrt{12}} \sim \mathrm{N}(0, 1)$$

We know that, for $\alpha = 0.05$, the non-rejection region for z is

$$[-1.96, 1.96].$$

That is, we do not reject if, assuming H_0, z is within two standard deviations of the standardized mean. By definition, this means

$$P(-1.96 \le z \le 1.96 \mid \mu = 2.71) = 0.95.$$

And, the rejection region is

$$(-\infty, -1.96) \cup (1.96, \infty).$$

For confidence intervals, we will want to unwind the definition of z and write the regions in terms of $\overline{x}$. This allows us to directly use the natural statistic $\overline{x}$.

Example 6. Redo the previous example using $\overline{x}$ as the test statistic.

Solution: Under the null hypothesis ($\mu = 2.71$) we have $x_i \sim \mathrm{N}(2.71, 5^2)$ and thus

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

where $5^2/12$ is the variance $\overline{x}$. We know that for normal data, significance $\alpha = 0.05$ corresponds to a rejection region starting 1.96 standard deviations from the hypothesized mean. That is,

Non-rejection region: We do not reject H_0 if $\overline{x}$ is in the interval

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

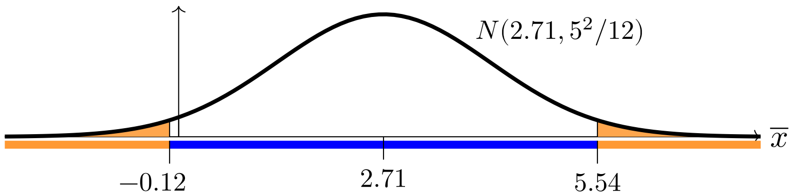
That is, we do not reject if, assuming H_0, $\overline{x}$ is within two standard deviations of the hypothesized mean. By definition, this means

$$P(-0.12 \le \overline{x} \le 5.54 \mid \mu = 2.71) = 0.95.$$

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 $\overline{x}$. The regions represent ranges of $\overline{x}$ so they are represented by the colored bars on the $\overline{x}$ axis. The area of the shaded region in the tails is the significance level.



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

Now, what about different data or null hypotheses. This is straight-forward, let's redo the previous example using symbols for all quantities.

Example 7. Suppose that n data points are drawn from $\mathrm{N}(\mu, \sigma^2)$ where μ is unknown and σ is known. Set up a two-sided significance test of $H_0 : \mu = \mu_0$ using the statistic $\overline{x}$ at significance level α. Describe the rejection and non-rejection regions.

Solution: Under the null hypothesis $\mu = \mu_0$ we have $x_i \sim \mathrm{N}(\mu_0, \sigma^2)$ and thus

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

where σ^2/n is the variance $(\sigma_{\overline{x}})^2$ of $\overline{x}$ and μ_0, σ 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 $\overline{x}$ that are $z_{\alpha/2} \cdot \sigma_{\overline{x}}$ from the hypothesized mean. Since $\sigma_{\overline{x}} = \dfrac{\sigma}{\sqrt{n}}$ we have

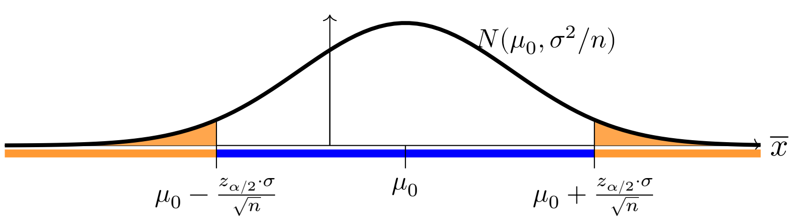
Non-rejection region: we do not reject H_0 if $\overline{x}$ is in the interval

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

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 $\overline{x}$.

4.3 Manipulating intervals: algebraic pivoting

We need to get comfortable manipulating intervals. In general, we will make use of the type of 'obvious' statements that can be hard to get across. First is the notion of pivoting. Stripping away the statistical terms, pivoting is the following algebraic maneuver.

Example 8. Algebraic pivoting. Suppose we have two variables a and b. Suppose also that a is in the interval $[b-4, b+6]$. Show that b is in the interval $[a-6, a+4]$.

Solution: We are given, $b-4 \le a \le b+6$. Therefore,

$$-4 \le a-b \le 6 \;\Rightarrow\; 4 \ge b-a \ge -6 \;\Rightarrow\; a+4 \ge b \ge a-6. \quad \text{QED}$$

This is called pivoting because the roles of a and b are reversed along with the direction of the inequalities.

In the example above, the ranges on either side of b are different. Quite often they will be the same. Here are some simple numerical examples of pivoting for symmetric intervals.

Example 9. **(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 $[1.5-1, 1.5+1]$.

4.4 Pivoting non-rejection intervals to confidence intervals

For normal data, the non-rejection region for $\overline{x}$ is an interval centered on μ_0. By pivoting, we get the confidence interval for μ centered on $\overline{x}$.

Example 10. Suppose we have n data points with a sample mean $\overline{x}$ and hypothesized mean $\mu_0 = 2.71$. Suppose also that the null distribution is $x_i \sim \mathrm{N}(\mu_0, 3^2)$. Then with a significance level of 0.05 we have:

(1a) The non-rejection region is centered on $\mu_0 = 2.71$. That is, we don't reject H_0 if $\overline{x}$ is in the interval

$$\left[\mu_0 - \frac{1.96\sigma}{\sqrt{n}},\ \mu_0 + \frac{1.96\sigma}{\sqrt{n}}\right]$$

(1b) Assuming the null hypothesis we have

$$P(\overline{x} \text{ is in the non-rejection region} \mid H_0) = 1-\alpha = 0.95.$$

That is,

$$P\left(\mu_0 - \frac{1.96\sigma}{\sqrt{n}} \le \overline{x} \le \mu_0 + \frac{1.96\sigma}{\sqrt{n}} \mid H_0\right) = 0.95$$

(2a) Pivoting (1a) gives: we don't reject H_0 if μ_0 is in the interval

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

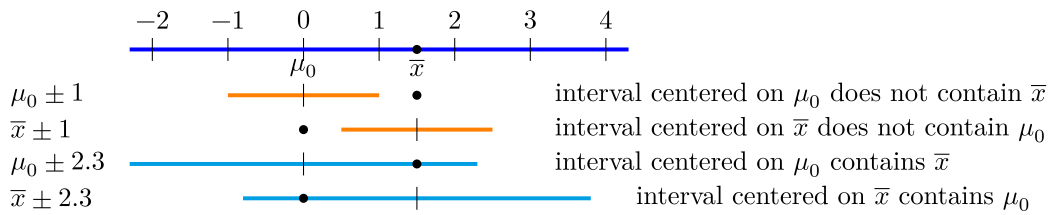
(2b) Pivoting (1b) gives: assuming the null hypothesis we have

$$P\left(\overline{x} - \frac{1.96\sigma}{\sqrt{n}} \le \mu_0 \le \overline{x} + \frac{1.96\sigma}{\sqrt{n}} \mid H_0\right) = 0.95$$

The interval in (2a) is called the 0.95 confidence interval for μ. It is centered on $\overline{x}$, it has the same width as the non-rejection region.

Again, notice the symmetry: the statement '$\overline{x}$ is in the non-rejection interval around μ_0' is equivalent to 'μ_0 is in the confidence interval $[\overline{x} - 1.96\sigma, \overline{x} + 1.96\sigma]$ around $\overline{x}$'.

Here is a visualization of *pivoting* from intervals around μ_0 to intervals around $\overline{x}$. In the figures, $\mu_0 = 1$ and $\overline{x} = 1.5$. The first pair of intervals have width 2 and the second pair have width 4.6.



The first pair of intervals shows the interval $\mu_0 \pm 1$ pivoted to the interval $\overline{x} \pm 1$. Since $\overline{x}$ is not in the first interval, μ_0 is not in the pivoted interval. In the second pair of intervals, since $\overline{x}$ is in the interval $\mu_0 \pm 2.3$, we see μ_0 is in the pivoted interval $\overline{x} \pm 2.3$.

4.5 Summary of normal confidence intervals: definition and properties

Suppose $x_1, x_2, \ldots, x_n$ are independent data from a $\mathrm{N}(\mu, \sigma^2)$ distribution. We assume μ is unknown, but σ is known.

- **Definition.** The $1 - \alpha$ confidence interval for μ is
$$\left[\overline{x} - \frac{z_{\alpha/2}\sigma}{\sqrt{n}}, \overline{x} + \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right],$$
where $z_{\alpha/2}$ is standard normal $\alpha/2$ critical value.
- The confidence interval only depends on $\overline{x}$ and known values, so it is a statistic.
- The confidence interval is random: different data generate different intervals.
- If the null hypothesis is $\mu = \mu_0$, then the confidence interval is found by pivoting the non-rejection region. If μ_0 is in the $1 - \alpha$ confidence interval, then we do not reject H_0 at significance level α. Likewise, we do reject H_0 at significance level α if μ_0 is *not* in the $1 - \alpha$ confidence interval.
- Assuming H_0, then in 95% of random trials the 95% confidence interval will contain μ_0.

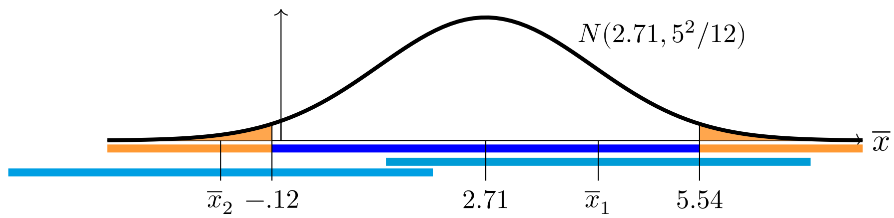
The following figure illustrates how we don't reject H_0 if the confidence interval around $\overline{x}$ contains μ_0 and we reject H_0 if the confidence interval doesn't contain μ_0. There is a lot in the figure so we will list carefully what you are seeing:

1. We started with the figure from Example 6 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 $\overline{x}$, i.e. $\overline{x}_1$ and $\overline{x}_2$, 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 $\pm\dfrac{1.96 \cdot 5}{\sqrt{12}}$.

The first value, $\overline{x}_1$, 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 μ_0.

The second value, $\overline{x}_2$, is in the rejection region and its interval does not contain μ_0. This illustrates that rejecting H_0 corresponds to the confidence interval not containing μ_0.



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

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 confidence interval is the same, i.e. it does not depend on any hypothesis.

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

Note that the specific values of σ 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 7 quickly through the same steps as Example6.

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

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

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

$$\mu_0 \text{ is in the interval } \overline{x} \pm \frac{z_{\alpha/2}\sigma}{\sqrt{n}}. \qquad (3)$$

We can rewrite equation 3 as: at significance level α we don't reject if

$$\text{the interval } \left[\overline{x} - \frac{z_{\alpha/2}\cdot\sigma}{\sqrt{n}},\ \ \overline{x} + \frac{z_{\alpha/2}\cdot\sigma}{\sqrt{n}}\right] \text{ contains } \mu_0. \qquad (4)$$

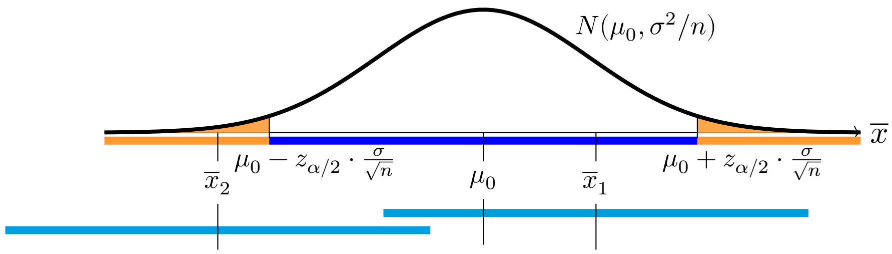
We call the interval 4 a $(1-\alpha)$ confidence interval because, assuming $\mu = \mu_0$, on average it will contain μ_0 in the fraction $(1-\alpha)$ of random trials.

The following figure illustrates the point that μ_0 is in the $(1-\alpha)$ confidence interval around $\overline{x}$ is equivalent to $\overline{x}$ is in the non-rejection region (at significance level α) for $H_0 : \mu_0 = \mu$.

The figure shows $\overline{x}_1$ is in the non-rejection region for μ_0, so the confidence interval around $\overline{x}_1$ contains μ_0.

Similarly, $\overline{x}_2$ is not in the non-rejection region for μ_0, so the confidence interval around $\overline{x}_2$ does not contain μ_0.

Note, that the confidence intervals and the non-rejection region all have the same width!



4.7 Computational example

Example 11. Suppose the data 2.5, 5.5, 8.5, 11.5 was drawn from a $\mathrm{N}(\mu, 10^2)$ distribution with unknown mean μ.

(a) Compute the point estimate $\overline{x}$ for μ 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 μ is in the relevant confidence interval and second by constructing a rejection region.

Solution: (a) We compute that $\overline{x} = 7.0$. The critical points are
$z_{0.025} =$ `qnorm(0.975)` $= 1.96$, $z_{0.1} =$ `qnorm(0.9)` $= 1.28$, $z_{0.25} =$ `qnorm(0.75)` $=$ 0.67.

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

$$\begin{array}{lclcl}
95\% \text{ conf. interval} & = & [\overline{x} - z_{0.025}\sigma_{\overline{x}}, \quad \overline{x} + z_{0.025}\sigma_{\overline{x}}] & = & [7 - 1.96 \cdot 5, \quad 7 + 1.96 \cdot 5] = [-2.8, \quad 16.8] \\
80\% \text{ conf. interval} & = & [\overline{x} - z_{0.1}\sigma_{\overline{x}}, \quad \overline{x} + z_{0.1}\sigma_{\overline{x}}] & = & [7 - 1.28 \cdot 5, \quad 7 + 1.28 \cdot 5] = [\ 0.6, \quad 13.4] \\
50\% \text{ conf. interval} & = & [\overline{x} - z_{0.75}\sigma_{\overline{x}}, \quad \overline{x} + z_{0.75}\sigma_{\overline{x}}] & = & [7 - 0.67 \cdot 5, \quad 7 + 0.67 \cdot 5] = [3.65, \quad 10.35]
\end{array}$$

Each of these intervals is a range estimate of μ. 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 μ: $\mu_0 = 1$ and confidence intervals are centered on $\overline{x}$. Here are the rejection regions.

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

To to do the NHST we must check whether or not $\overline{x} = 7$ 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 σ is not known, so we have to make the following replacements.

1. Use $s_{\overline{x}} = \dfrac{s}{\sqrt{n}}$ instead of $\sigma_{\overline{x}} = \dfrac{\sigma}{\sqrt{n}}$. 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 $x_1, \ldots, x_n \sim \mathrm{N}(\mu, \sigma^2)$, where the values of the mean μ and the standard deviation σ are both unknown. . The $(1 - \alpha)$ confidence interval for μ is

$$\left[\overline{x} \;-\; \frac{t_{\alpha/2} \cdot s}{\sqrt{n}},\;\; \overline{x} \;+\; \frac{t_{\alpha/2} \cdot s}{\sqrt{n}}\right], \tag{5}$$

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

For t confidence intervals we repeat the construction of normal confidence intervals with σ replaced by its estimate s.

Suppose that n data points are drawn from $\mathrm{N}(\mu, \sigma^2)$ where μ and σ 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 $x_i \sim \mathrm{N}(\mu_0, \sigma^2)$. So the studentized mean follows a Student t distribution with $n - 1$ degrees of freedom:

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

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 $\overline{x}$ we get: at significance level α we don't reject if

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

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

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

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

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

This interval is the confidence interval defined in 5.

Example 12. Suppose the data 2.5, 5.5, 8.5, 11.5 was drawn from a $\mathrm{N}(\mu,\sigma^2)$ distribution with μ and σ both unknown.

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

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

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

All of these confidence intervals give interval estimates for the value of μ. 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 $x_1, \ldots, x_n$ is drawn from $\mathrm{N}(\mu,\sigma^2)$ with mean μ and standard deviation σ both unknown. The $(1-\alpha)$ confidence interval for the variance σ^2 is

$$\left[\frac{(n-1)s^2}{c_{\alpha/2}},\ \frac{(n-1)s^2}{c_{1-\alpha/2}}\right]. \tag{6}$$

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 $\mathrm{N}(\mu,\sigma^2)$, 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 α is

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

Pivoting 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 chi-square confidence intervals are 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 its mean $n-1$.

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:

https://mathlets.org/mathlets/confidence-intervals/