Section 7.1
One-sample mean with the t-distribution

Stats 7 Summer Session II 2022

Mercury content of dolphin muscle

Let's get our first taste of inference with a mean and applying the t-distribution in the context of an example about the mercury content of dolphin muscle. Elevated mercury concentrations are an important problem for both dolphins and other animals, like humans, who occasionally eat them.

Figure 7.5: A Risso's dolphin.

We want to learn about the average mercury content in dolphin muscle using a sample of 19 Risso's dolphins from the Taiji area in Japan. They were found to have an average mercury content of 4.4 (µg / wet g) with a standard deviation of 2.3.

Mercury content of dolphin muscle

We want to learn if the average mercury content in dolphin muscle in this region have increased from 3 (μg / wet g).

What are our hypotheses?

H_0: The average dolphin muscle mercury content equals 3

(H_0: $\mu = 3$)

H_A: The average dolphin muscle mercury content is greater than 3

(H_A: $\mu > 3$)

What do we know about means?

CLT for sampling distribution of the sample mean

When we collect a sufficiently large sample of n independent observations from a population with mean μ and standard deviation σ, then the sample mean is distributed as:

$$\bar{x} \dot{\sim} N(\text{Mean} = \mu, \text{SE} = \frac{\sigma}{\sqrt{n}})$$

What purpose does a large sample server?

As long as observations are independent, and the population distribution is not extremely skewed, a large sample would ensure that...

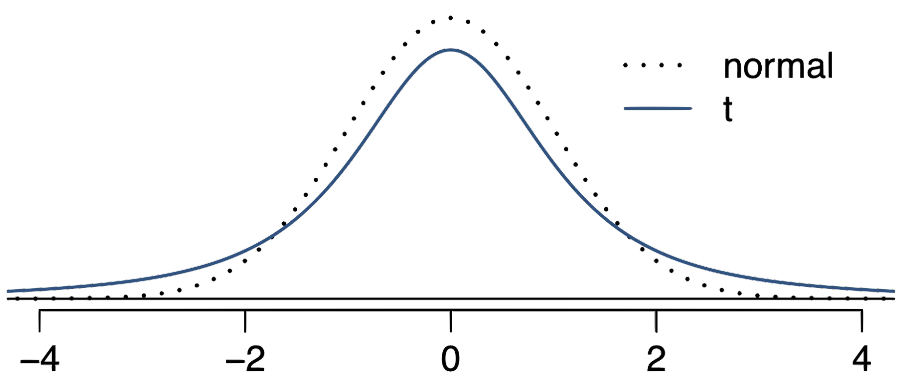
- the sampling distribution of the mean is nearly normal
- the estimate of the standard error, as $\frac{s}{\sqrt{n}}$ in place of $\frac{\sigma}{\sqrt{n}}$, is reliable

The normality condition

- The CLT, which states that sampling distributions will be nearly normal, hold true for *any* sample size as long as the population distribution is nearly normal
- While this is a helpful special case, it’s inherently difficult to verify normality in small data sets
- We should exercise caution when verifying the normality condition for small samples. It is important to not only examine the data but also think about where the data come from
  - For example, ask: would I expect this distribution to be symmetric, and am I confident that outliers are rare?
  - Rule of thumb: 2.5 * sample standard deviation separates “unusual” observations

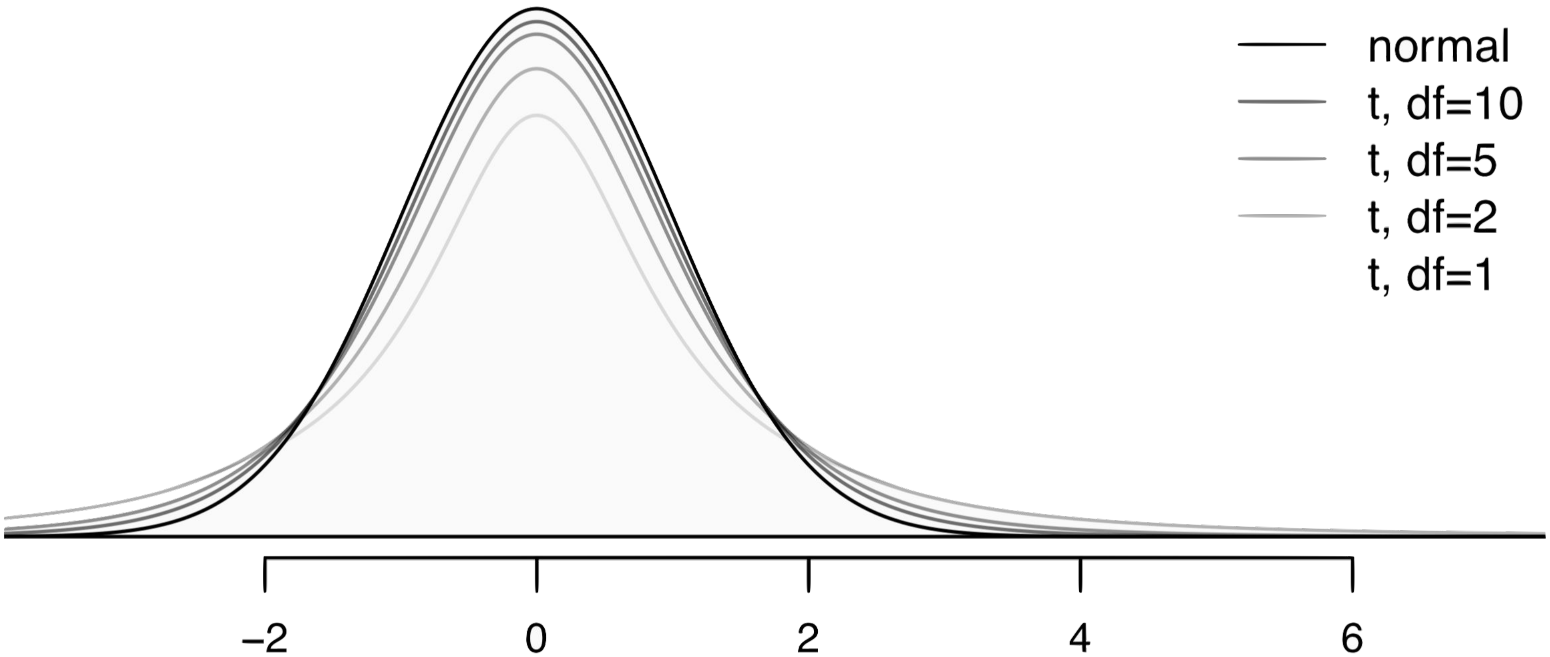
The t distribution

- When the population standard deviation, σ, is unknown (almost always), the uncertainty of the standard error estimate is addressed by using a new distribution: the t distribution.
- This distribution also has a bell shape, but its tails are thicker than the normal model's.
- Therefore observations are more likely to fall beyond two SDs from the mean than under the normal distribution
- These extra thick tails are helpful for resolving our problem with a less reliable estimate the standard error (since *n* is small)



The t distribution (cont.)

- Always centered at zero, like the standard normal (z) distribution
- Has a single parameter: *degrees of freedom* (*df*).



What happens to the shape of the t distribution as *df* increases?

Approaches normal

Find the test statistic

The test statistic for inference on a small sample (n < 50) mean is the T statistic with $df = n - 1$

Test statistic for inference on a small sample mean

$$T = \frac{\text{Point estimate} - \text{null value } \mu_0}{\text{SE}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

We use s in place of σ

For our problem

… sample of 19 Risso’s dolphins from the Taiji area in Japan. They were found to have an average mercury content of 4.4 (µg / wet g) with a standard deviation of 2.3

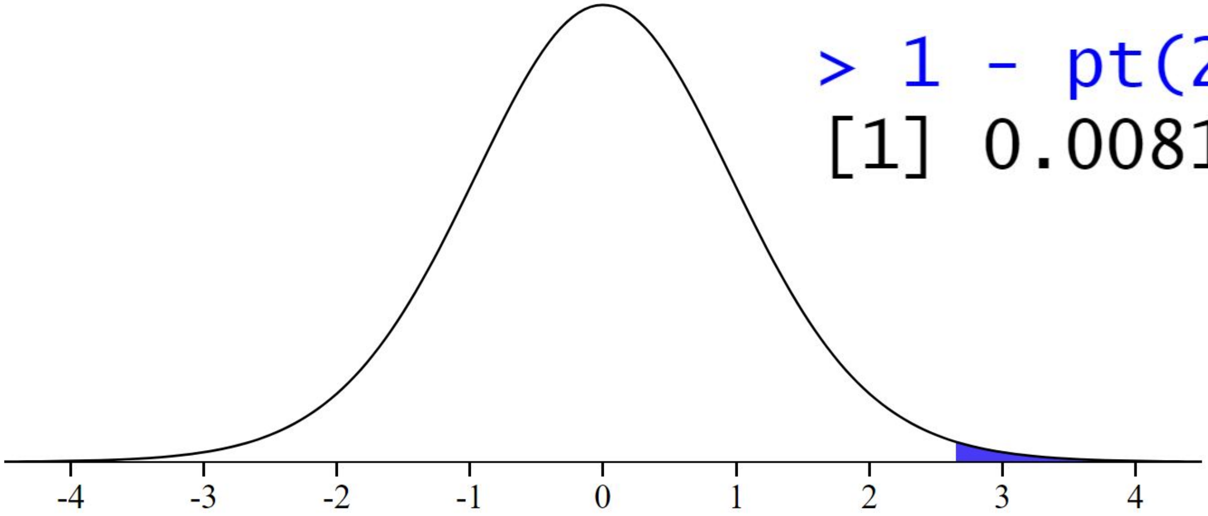
$$\text{Point estimate} = \bar{x} = 4.4$$

$$\text{Test statistic} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{4.4 - 3}{2.3/\sqrt{19}} = 2.65$$

$$df = n - 1 = 19 - 1 = 18$$

Finding the p-value

- The p-value is, once again, calculated as the area under the tail of the t distribution
  - This is similar to the normal distribution p-values with 3 possible tails (left tail, right tail, two-tails) dependent upon H_A




```
> 1 - pt(2.65, df = 18)
[1] 0.008144441
```

There is a 0.81% probability of observing data as or more extreme than we observed if the dolphin muscle mercury content was 3 (µg / wet g).

Conclusion of the test

What is the conclusion of this hypothesis test?

P-value is less than significance level so reject the null hypothesis.

The data provide very strong evidence (p-value = 0.0081 $< \alpha = 0.05$) that the average mercury content in dolphin muscle is greater than 3 μg / wet g.

What is the new average concentration?

- We concluded that there is an increase in concentration from 3.
- But it would be more interesting to find out what exactly this new average concentration is.
- We can use a confidence interval to estimate this mean

Confidence interval for a small sample mean

- Confidence intervals are always of the form

$$point\ estimate \pm ME$$

- ME is always calculated as the product of a critical value and SE
- Since small sample means follow a ***t*** distribution (and not a z distribution), the critical value is a ***t****** (as opposed to a z^*).

$$point\ estimate \pm t^*\ x\ SE$$

- We can compute ***t****** similarly to the computation of z^*.
  - We use qt( ) with an area of (1 - confidence level) / 2

We want to make a 90% confidence interval

We calculate our t^* with 18 degrees of freedom as:

```
> qt((1 - 0.90) / 2, df = 18)
[1] -1.734064
```

So our confidence interval is:

$$\text{point estimate} \pm t^* \times SE$$

$$4.4 \pm 1.73 * \frac{2.3}{\sqrt{19}}$$

$$4.4 \pm 0.913$$

$$(3.49, 5.31)$$

Interpreting the CI

Which of the following is the *best* interpretation for the confidence interval we just calculated? (3.49, 5.31)

We are 95% confident that…

A. the average mercury content of muscles in this sample of Risso's dolphins is between 3.29 and 5.51 µg/wet gram, which is considered extremely high.

B. there is a 95% change of the average mercury content of muscles in Risso's dolphins is between 3.29 and 5.51 µg/wet gram, which is considered extremely high.

C. the average mercury content of muscles in Risso's dolphins is between 3.29 and 5.51 µg/wet gram, which is considered extremely high.

Synthesis

Does the conclusion from the hypothesis test agree with the findings of the confidence interval?

Yes, the hypothesis test found a significant increase, and the CI does not contain the null value of 3

Recap: Inference using the *t*-distribution

- If σ is unknown, use the *t*-distribution with $SE = \frac{s}{\sqrt{n}}$
- Conditions:
  - independence of observations (often verified by a random sample, and if n < 10% of population)
  - Normality:
    - For n < 30, need data to be approximately normally distributed with no clear outliers
    - For n ≥ 30, need no extreme outliers
- Hypothesis Testing:

$$T_{df} = \frac{point\ estimate\ - null\ value}{SE}, where\ df = n - 1$$

- Confidence interval:

$$point\ estimate \pm t^*_{df} \times SE$$