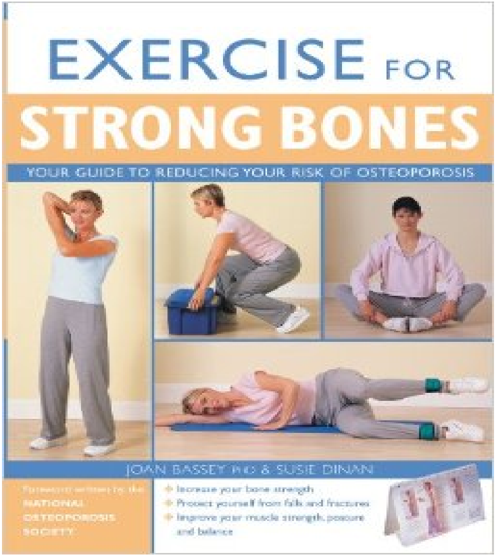
STAT234: Lecture 6 - Inference methods

Kushal K. Dey

Example: Does exercise make strong bones?



Example: Does exercise make strong bones?

Can a 6-month exercise program increase the total body bone mineral content (TBBMC) of young women? A team of researchers is planning a study to examine this question.

Example: Does exercise make strong bones?

Can a 6-month exercise program increase the total body bone mineral content (TBBMC) of young women? A team of researchers is planning a study to examine this question.

Based on the results of a previous study, they are willing to assume that pop. standard error $\sigma = 2$ for the percent change in TBBMC over the 6-month period. Assume the sample size is 25.

Example: Does exercise make strong bones?

Can a 6-month exercise program increase the total body bone mineral content (TBBMC) of young women? A team of researchers is planning a study to examine this question.

Based on the results of a previous study, they are willing to assume that pop. standard error $\sigma = 2$ for the percent change in TBBMC over the 6-month period. Assume the sample size is 25.

- What are the hypotheses?

Example: Does exercise make strong bones?

Can a 6-month exercise program increase the total body bone mineral content (TBBMC) of young women? A team of researchers is planning a study to examine this question.

Based on the results of a previous study, they are willing to assume that pop. standard error $\sigma = 2$ for the percent change in TBBMC over the 6-month period. Assume the sample size is 25.

- What are the hypotheses?
- A change in TBBMC of 1% change would be considered important. How to check if that is the case or not?
- What's the rejection region under 5% significance level?

Example: Does exercise make strong bones?

Can a 6-month exercise program increase the total body bone mineral content (TBBMC) of young women? A team of researchers is planning a study to examine this question.

Based on the results of a previous study, they are willing to assume that pop. standard error $\sigma = 2$ for the percent change in TBBMC over the 6-month period. Assume the sample size is 25.

- What are the hypotheses?
- A change in TBBMC of 1% change would be considered important. How to check if that is the case or not?
- What's the rejection region under 5% significance level?

Framing the hypothesis testing

Let μ denote the mean percent change of the population of young women. We do not know μ.

Framing the hypothesis testing

Let μ denote the mean percent change of the population of young women. We do not know μ.

We test

$$H_0 : \mu = 0 \qquad H_1 : \mu > 0$$

Framing the hypothesis testing

Let μ denote the mean percent change of the population of young women. We do not know μ.

We test

$$H_0 : \mu = 0 \qquad H_1 : \mu > 0$$

Assume we have 25 values $X_1, X_2, \cdots, X_{25}$ of observed percent changes in TBBMC for 25 women.

Framing the hypothesis testing

Let μ denote the mean percent change of the population of young women. We do not know μ.

We test

$$H_0 : \mu = 0 \qquad H_1 : \mu > 0$$

Assume we have 25 values $X_1, X_2, \cdots, X_{25}$ of observed percent changes in TBBMC for 25 women.
Suppose we assume normality

$$X_i \sim N(\mu, \sigma^2)$$

Framing the hypothesis testing

Let μ denote the mean percent change of the population of young women. We do not know μ.

We test

$$H_0 : \mu = 0 \qquad H_1 : \mu > 0$$

Assume we have 25 values $X_1, X_2, \cdots, X_{25}$ of observed percent changes in TBBMC for 25 women.
Suppose we assume normality

$$X_i \sim N(\mu, \sigma^2)$$

But from previous study, we know $\sigma^2 = 4$. So,

Framing the hypothesis testing

Let μ denote the mean percent change of the population of young women. We do not know μ.

We test

$$H_0 : \mu = 0 \qquad H_1 : \mu > 0$$

Assume we have 25 values $X_1, X_2, \cdots, X_{25}$ of observed percent changes in TBBMC for 25 women.
Suppose we assume normality

$$X_i \sim N(\mu, \sigma^2)$$

But from previous study, we know $\sigma^2 = 4$. So,

Example: Definitions

What is *rejection region*?

Example: Definitions

What is *rejection region*?

The region such that if the data occurs in that region, we will reject the null hypothesis.

Example: Definitions

What is *rejection region*?

The region such that if the data occurs in that region, we will reject the null hypothesis.

What is *significance level*?

Example: Definitions

What is *rejection region*?

The region such that if the data occurs in that region, we will reject the null hypothesis.

What is *significance level*?

It is a threshold for the p-value, or the probability of observing something as extreme as what you have observed from the data. If the p-value is less than this threshold, we reject.

Framing the hypothesis testing

We can write

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

Framing the hypothesis testing

We can write

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

here

$$\bar{X} \sim N(\mu, \frac{4}{25})$$

Framing the hypothesis testing

We can write

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

here

$$\bar{X} \sim N(\mu, \frac{4}{25})$$

Under H_0 or if H_0 is true, we can assume $\mu_0 = 0$. Then we can write

$$\bar{X} \sim N(0, \frac{4}{25}) \qquad \textit{under } H_0$$

Framing the hypothesis testing

We can write

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

here

$$\bar{X} \sim N(\mu, \frac{4}{25})$$

Under H_0 or if H_0 is true, we can assume $\mu_0 = 0$. Then we can write

$$\bar{X} \sim N(0, \frac{4}{25}) \qquad \textit{under } H_0$$

$$\sqrt{25}\ \bar{X} \sim N(0, 1) \qquad \textit{under } H_0$$

Hypothesis testing

If we wanted to test at 5% level of significance, meaning if we wanted to reject whenever p-value is less than 0.05, then we would choose a value c so that if we had observed c as the realization of $\bar{X}$, we reject the null hypothesis.

$$Pr_{H_0}\left[\bar{X} > c\right] < 0.05$$

Hypothesis testing

If we wanted to test at 5% level of significance, meaning if we wanted to reject whenever p-value is less than 0.05, then we would choose a value c so that if we had observed c as the realization of $\bar{X}$, we reject the null hypothesis.

$$Pr_{H_0}\left[\bar{X} > c\right] < 0.05$$

The left hand side is definition of p-value when c is observed and for the c to be an extreme value, we want that p-value to be less than 0.05.

Hypothesis testing

If we wanted to test at 5% level of significance, meaning if we wanted to reject whenever p-value is less than 0.05, then we would choose a value c so that if we had observed c as the realization of $\bar{X}$, we reject the null hypothesis.

$$Pr_{H_0}\left[\bar{X} > c\right] < 0.05$$

The left hand side is definition of p-value when c is observed and for the c to be an extreme value, we want that p-value to be less than 0.05.

Why is it one-sided here?

Hypothesis testing

If we wanted to test at 5% level of significance, meaning if we wanted to reject whenever p-value is less than 0.05, then we would choose a value c so that if we had observed c as the realization of $\bar{X}$, we reject the null hypothesis.

$$Pr_{H_0}\left[\bar{X} > c\right] < 0.05$$

The left hand side is definition of p-value when c is observed and for the c to be an extreme value, we want that p-value to be less than 0.05.

Why is it one-sided here?

What is the smallest c for which we get this property to hold. Get c so that

Hypothesis testing

If we wanted to test at 5% level of significance, meaning if we wanted to reject whenever p-value is less than 0.05, then we would choose a value c so that if we had observed c as the realization of $\bar{X}$, we reject the null hypothesis.

$$Pr_{H_0}\left[\bar{X} > c\right] < 0.05$$

The left hand side is definition of p-value when c is observed and for the c to be an extreme value, we want that p-value to be less than 0.05.

Why is it one-sided here?

What is the smallest c for which we get this property to hold. Get c so that

Hypothesis testing

That c would form the boundary between acceptance region $\bar{X} < c$, the set of all realizations for which p-value would be greater than 0.05 and we would accept null, and the rejection region $\bar{X} > c$, the region for which we will reject the null hypothesis.

Hypothesis testing

That c would form the boundary between acceptance region $\bar{X} < c$, the set of all realizations for which p-value would be greater than 0.05 and we would accept null, and the rejection region $\bar{X} > c$, the region for which we will reject the null hypothesis.

$$Pr_{H_0}\left[\bar{X} > c\right] = 0.05$$

$$Pr_{H_0}\left[\frac{\sqrt{25}}{\sqrt{4}}\bar{X} > \frac{\sqrt{25}}{\sqrt{4}}c\right] = 0.05$$

$$Pr_{H_0}\left[Z > \frac{5}{2}c\right] = 0.05$$

Hypothesis testing

That c would form the boundary between acceptance region $\bar{X} < c$, the set of all realizations for which p-value would be greater than 0.05 and we would accept null, and the rejection region $\bar{X} > c$, the region for which we will reject the null hypothesis.

$$Pr_{H_0}\left[\bar{X} > c\right] = 0.05$$

$$Pr_{H_0}\left[\frac{\sqrt{25}}{\sqrt{4}}\bar{X} > \frac{\sqrt{25}}{\sqrt{4}}c\right] = 0.05$$

$$Pr_{H_0}\left[Z > \frac{5}{2}c\right] = 0.05$$

We know that

$$Pr_{H_0}\left[Z > 1.645\right] = 0.05$$

Hypothesis testing

That c would form the boundary between acceptance region $\bar{X} < c$, the set of all realizations for which p-value would be greater than 0.05 and we would accept null, and the rejection region $\bar{X} > c$, the region for which we will reject the null hypothesis.

$$Pr_{H_0}\left[\bar{X} > c\right] = 0.05$$

$$Pr_{H_0}\left[\frac{\sqrt{25}}{\sqrt{4}}\bar{X} > \frac{\sqrt{25}}{\sqrt{4}}c\right] = 0.05$$

$$Pr_{H_0}\left[Z > \frac{5}{2}c\right] = 0.05$$

We know that

$$Pr_{H_0}\left[Z > 1.645\right] = 0.05$$

Hypothesis testing

So we get

$$\frac{5}{2}c = 1.645$$

Hypothesis testing

So we get

$$\frac{5}{2}c = 1.645$$

implies

$$c = 1.645 \times \frac{2}{5} = 0.658$$

Hypothesis testing

So we get

$$\frac{5}{2}c = 1.645$$

implies

$$c = 1.645 \times \frac{2}{5} = 0.658$$

So the rejection region under 5% level of significance is $\bar{X} > 0.658$.

Hypothesis testing

So we get

$$\frac{5}{2}c = 1.645$$

implies

$$c = 1.645 \times \frac{2}{5} = 0.658$$

So the rejection region under 5% level of significance is $\bar{X} > 0.658$.
That was all model......

Hypothesis testing

So we get

$$\frac{5}{2}c = 1.645$$

implies

$$c = 1.645 \times \frac{2}{5} = 0.658$$

So the rejection region under 5% level of significance is $\bar{X} > 0.658$.
That was all model......
Now look at the data, the realization $\bar{x}$ of the random variable $\bar{X}$.

Hypothesis testing

So we get

$$\frac{5}{2}c = 1.645$$

implies

$$c = 1.645 \times \frac{2}{5} = 0.658$$

So the rejection region under 5% level of significance is $\bar{X} > 0.658$.
That was all model......
Now look at the data, the realization $\bar{x}$ of the random variable $\bar{X}$.

Hypothesis testing

If say the observed value $\bar{x} = 0.6$, then we will accept the null, meaning there is no effect of the program.

Hypothesis testing

If say the observed value $\bar{x} = 0.6$, then we will accept the null, meaning there is no effect of the program.

If say the observed value $\bar{x} = 0.9$, then we will reject the null, meaning there is a positive effect of the program.

Hypothesis testing

If say the observed value $\bar{x} = 0.6$, then we will accept the null, meaning there is no effect of the program.

If say the observed value $\bar{x} = 0.9$, then we will reject the null, meaning there is a positive effect of the program.

The test is one-sided because as per research the alternative is that the program is effective and there is a positive change or increase in TBBMC. Thats why the alternative was $\mu > 0$.

Hypothesis testing

If say the observed value $\bar{x} = 0.6$, then we will accept the null, meaning there is no effect of the program.

If say the observed value $\bar{x} = 0.9$, then we will reject the null, meaning there is a positive effect of the program.

The test is one-sided because as per research the alternative is that the program is effective and there is a positive change or increase in TBBMC. Thats why the alternative was $\mu > 0$.

Hypothesis testing

Okay, we define some terminologies now....

Hypothesis testing

Okay, we define some terminologies now....

Type I error: The probability we reject the null when the null is true.

Hypothesis testing

Okay, we define some terminologies now....

Type I error: The probability we reject the null when the null is true.
For example, if we reject when $\bar{X} > 0.658$, then under the null, $\mu = 0$, we have

$$\textit{Type I error} : Pr\left[\bar{X} > 0.658 | \mu = 0\right] = 0.05$$

Hypothesis testing

Okay, we define some terminologies now....

Type I error: The probability we reject the null when the null is true.
For example, if we reject when $\bar{X} > 0.658$, then under the null, $\mu = 0$, we have

$$\textit{Type I error} : Pr\left[\bar{X} > 0.658 | \mu = 0\right] = 0.05$$

If we reject when $\bar{X} > 0.8$, then

$$\textit{Type I error} : Pr\left[\bar{X} > 0.8 | \mu = 0\right] = 0.0006$$

Hypothesis testing

Okay, we define some terminologies now....

Type I error: The probability we reject the null when the null is true.
For example, if we reject when $\bar{X} > 0.658$, then under the null, $\mu = 0$, we have

$$\textit{Type I error} : Pr\left[\bar{X} > 0.658 | \mu = 0\right] = 0.05$$

If we reject when $\bar{X} > 0.8$, then

$$\textit{Type I error} : Pr\left[\bar{X} > 0.8 | \mu = 0\right] = 0.0006$$

As we increase c, we make Type I error small.

Hypothesis testing

$$\textit{Type I error} : Pr\left[Z > \frac{\sqrt{n}}{\sigma}c|\mu = 0\right] = 0.05 \qquad Z \sim N(0,1)$$

Hypothesis testing

$$\text{Type I error} : Pr\left[Z > \frac{\sqrt{n}}{\sigma}c|\mu = 0\right] = 0.05 \qquad Z \sim N(0,1)$$

So we can reduce this Type I error by increasing c, but does it solve the problem?

Hypothesis testing

$$Type\ I\ error : Pr\left[Z > \frac{\sqrt{n}}{\sigma}c|\mu = 0\right] = 0.05 \qquad Z \sim N(0,1)$$

So we can reduce this Type I error by increasing c, but does it solve the problem?

Should we take c to be infinity then?

Hypothesis testing

$$\mathit{Type\ I\ error} : Pr\left[Z > \frac{\sqrt{n}}{\sigma} c | \mu = 0\right] = 0.05 \qquad Z \sim N(0,1)$$

So we can reduce this Type I error by increasing c, but does it solve the problem?

Should we take c to be infinity then?

NO....there is another error to worry about

Hypothesis testing

$$\textit{Type I error} : Pr\left[Z > \frac{\sqrt{n}}{\sigma}c|\mu = 0\right] = 0.05 \qquad Z \sim N(0,1)$$

So we can reduce this Type I error by increasing c, but does it solve the problem?

Should we take c to be infinity then?

NO....there is another error to worry about

What is the probability we accept the null when the null is not true.

Hypothesis testing

$$\textit{Type I error} : Pr\left[Z > \frac{\sqrt{n}}{\sigma} c | \mu = 0\right] = 0.05 \qquad Z \sim N(0,1)$$

So we can reduce this Type I error by increasing c, but does it solve the problem?

Should we take c to be infinity then?

NO....there is another error to worry about

What is the probability we accept the null when the null is not true.

This is called the **Type II error**.

Hypothesis testing

What do you mean by null not being TRUE. What is μ then?

Hypothesis testing

What do you mean by null not being TRUE. What is μ then?

Well, $\mu > 0$, so lets fix one such μ, say $\mu = 1$.

Hypothesis testing

What do you mean by null not being TRUE. What is μ then?

Well, $\mu > 0$, so lets fix one such μ, say $\mu = 1$.

Type II error can be written as

$$\textit{Type II error} : Pr\left[\bar{X} < 0.658 | \mu = 1\right]$$

Hypothesis testing

What do you mean by null not being TRUE. What is μ then?

Well, $\mu > 0$, so lets fix one such μ, say $\mu = 1$.

Type II error can be written as

$$\textit{Type II error} : Pr\left[\bar{X} < 0.658|\mu = 1\right]$$

Under this set up, since we know

$$\bar{X} \sim N(\mu = 1, \frac{4}{25})$$

Hypothesis testing

$$Type\ II\ error : Pr\left[\frac{5}{2}(\bar{X}-1) < \frac{5}{2}(0.658-1)|\mu = 1\right]$$

$$Type\ II\ error : Pr\left[Z < -0.855|\mu = 1\right]$$

$$Type\ II\ error : Pr\left[Z < -0.855\right]\ Z \sim N(0,1)$$

This can be calculated using Normal table that Type II error is 0.1963.
The important thing is for any c we choose as the boundary between acceptance and rejection

$$Type\ II\ error : Pr\left[Z < \frac{5}{2}(c-1)\right]\ Z \sim N(0,1)$$

Hypothesis testing

Note that as c increases, the Type II error increases.

Hypothesis testing

Note that as c increases, the Type II error increases.

So, as c increases, Type I error decreases and Type II error increases

Hypothesis testing

Note that as c increases, the Type II error increases.

So, as c increases, Type I error decreases and Type II error increases

Statisticians use an alternative to Type II error called *power*.

Hypothesis testing

Note that as c increases, the Type II error increases.

So, as c increases, Type I error decreases and Type II error increases

Statisticians use an alternative to Type II error called *power*.

Power: Probability of rejecting the null when the null is not true. What is the power when $\mu = 1$?

$$Power : Pr\left[\bar{X} > 0.658 | \mu = 1\right]$$

Hypothesis testing

$$Power : Pr\left[Z > \frac{\sqrt{n}}{\sigma}(c-1)\right] \; Z \sim N(0,1)$$

where $\sigma = 2$ and $n = 25$.

So power decreases as c increases.

Hypothesis testing

$$Power : Pr\left[Z > \frac{\sqrt{n}}{\sigma}(c-1)\right] \; Z \sim N(0,1)$$

where $\sigma = 2$ and $n = 25$.
So power decreases as c increases.

So, as c increase Type I error and power increase, the latter implying Type II error decreases because

$$Power := 1 - Type\ II\ error$$

Ways to Increase the Power

- Increase α. A 5% test of significance will have a greater chance of rejecting the null than a 1% test because the strength of evidence required for rejection is less.

Ways to Increase the Power

- Increase α. A 5% test of significance will have a greater chance of rejecting the null than a 1% test because the strength of evidence required for rejection is less.
- Consider a particular alternative that is farther away from $\mu_0 = 0$. Values of μ that are in H_a but lie close to the hypothesized value μ_0 are harder to detect than values of μ that are far from μ_0.

Ways to Increase the Power

- Increase α. A 5% test of significance will have a greater chance of rejecting the null than a 1% test because the strength of evidence required for rejection is less.
- Consider a particular alternative that is farther away from $\mu_0 = 0$. Values of μ that are in H_a but lie close to the hypothesized value μ_0 are harder to detect than values of μ that are far from μ_0.
- Increase the sample size. More data will provide more information about $\bar{x}$ so we have a better chance of distinguishing values of μ.

Ways to Increase the Power

- Increase α. A 5% test of significance will have a greater chance of rejecting the null than a 1% test because the strength of evidence required for rejection is less.
- Consider a particular alternative that is farther away from $\mu_0 = 0$. Values of μ that are in H_a but lie close to the hypothesized value μ_0 are harder to detect than values of μ that are far from μ_0.
- Increase the sample size. More data will provide more information about $\bar{x}$ so we have a better chance of distinguishing values of μ.
- Decrease σ. This has the same effect as increasing the sample size: it provides more information about μ. Improving the measurement process and restricting attention to a subpopulation are two common ways to decrease σ.

Example: Does exercise make strong bones? (cont.)

- Change the significance level to $\alpha = 0.01$.

Example: Does exercise make strong bones? (cont.)

- Change the significance level to $\alpha = 0.01$.

  The z test rejects H_0 at the $\alpha = 0.01$ level whenever

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x}}{2/\sqrt{25}} \geq 2.32$$

  that is, we reject H_0 when $\bar{x} \geq 2.32 \frac{2}{\sqrt{25}} = 0.928$

Example: Does exercise make strong bones? (cont.)

- Change the significance level to $\alpha = 0.01$.

  The z test rejects H_0 at the $\alpha = 0.01$ level whenever

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x}}{2/\sqrt{25}} \geq 2.32$$

  that is, we reject H_0 when $\bar{x} \geq 2.32\frac{2}{\sqrt{25}} = 0.928$

  The power of the test at alternative $\mu = 1$ is

$$\begin{aligned} P(\bar{x} \geq 0.928|\mu = 1) &= P(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \geq \frac{0.928 - 1}{2/\sqrt{25}}|\mu = 1) \\ &= P(Z \geq -0.18) = 0.57 \end{aligned}$$

Example: Does exercise make strong bones? (cont.)

- Change the significance level to $\alpha = 0.01$.

  The z test rejects H_0 at the $\alpha = 0.01$ level whenever

  $$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x}}{2/\sqrt{25}} \geq 2.32$$

  that is, we reject H_0 when $\bar{x} \geq 2.32\frac{2}{\sqrt{25}} = 0.928$

  The power of the test at alternative $\mu = 1$ is

  $$P(\bar{x} \geq 0.928|\mu = 1) = P(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \geq \frac{0.928 - 1}{2/\sqrt{25}}|\mu = 1)$$
  $$= P(Z \geq -0.18) = 0.57$$

- Change the alternative to $\mu = 2$.

Example: Does exercise make strong bones? (cont.)

- Change the significance level to $\alpha = 0.01$.

  The z test rejects H_0 at the $\alpha = 0.01$ level whenever

  $$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x}}{2/\sqrt{25}} \geq 2.32$$

  that is, we reject H_0 when $\bar{x} \geq 2.32\frac{2}{\sqrt{25}} = 0.928$

  The power of the test at alternative $\mu = 1$ is

  $$\begin{aligned} P(\bar{x} \geq 0.928|\mu = 1) &= P(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \geq \frac{0.928 - 1}{2/\sqrt{25}}|\mu = 1) \\ &= P(Z \geq -0.18) = 0.57 \end{aligned}$$

- Change the alternative to $\mu = 2$. The power of the test at alternative $\mu = 2$ at $\alpha = 0.01$ level is

  $$\begin{aligned} P(\bar{x} \geq 0.928|\mu = 2) &= P(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \geq \frac{0.928 - 2}{2/\sqrt{25}}|\mu = 2) \\ &= P(Z \geq -2.68) = 0.996 \end{aligned}$$

Example: Does exercise make strong bones? (cont.)

- Increase the sample size $n = 100$.

Example: Does exercise make strong bones? (cont.)

- Increase the sample size $n = 100$.
  The z test rejects H_0 at the $\alpha = 0.01$ level whenever

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x}}{2/\sqrt{100}} \geq 2.32$$

  that is, we reject H_0 when $\bar{x} \geq 2.32\frac{2}{\sqrt{100}} = 0.464$

Example: Does exercise make strong bones? (cont.)

- Increase the sample size $n = 100$.
  The z test rejects H_0 at the $\alpha = 0.01$ level whenever

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x}}{2/\sqrt{100}} \geq 2.32$$

  that is, we reject H_0 when $\bar{x} \geq 2.32\frac{2}{\sqrt{100}} = 0.464$
  The power of the test at alternative $\mu = 1$ is

$$P(\bar{x} \geq 0.464|\mu = 1) = P(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \geq \frac{0.464 - 1}{2/\sqrt{100}} \mid \mu = 1)$$

$$= P(Z \geq -2.68) = 0.996$$

Example: Does exercise make strong bones? (cont.)

- Decrease $\sigma = 1$

Example: Does exercise make strong bones? (cont.)

- Decrease $\sigma = 1$

  The z test rejects H_0 at the $\alpha = 0.01$ level whenever

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x}}{1/\sqrt{100}} \geq 2.32$$

  that is, we reject H_0 when $\bar{x} \geq 2.32\frac{1}{\sqrt{100}} = 0.232$

Example: Does exercise make strong bones? (cont.)

- Decrease $\sigma = 1$

  The z test rejects H_0 at the $\alpha = 0.01$ level whenever

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x}}{1/\sqrt{100}} \geq 2.32$$

  that is, we reject H_0 when $\bar{x} \geq 2.32\frac{1}{\sqrt{100}} = 0.232$ The power of the test at alternative $\mu = 1$ is

$$P(\bar{x} \geq 0.232|\mu = 1) = P(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \geq \frac{0.232 - 1}{1/\sqrt{100}}|\mu = 1)$$
$$= P(Z \geq -7.68) \approx 1$$

Two Types of Errors Revisited

Recall that the following four outcomes are possible when conducting a test:

Reality	Our Decision H_0	Our Decision H_a
H_0	$\surd$ (Prob $= 1 - \alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1 - \beta$)

Two Types of Errors Revisited

Recall that the following four outcomes are possible when conducting a test:

Reality	Our Decision H_0	Our Decision H_a
H_0	$\surd$ (Prob $= 1-\alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1-\beta$)

Type I error is generally **more serious** than type II error.

Two Types of Errors Revisited

Recall that the following four outcomes are possible when conducting a test:

Reality	Our Decision H_0	Our Decision H_a
H_0	$\surd$ (Prob $= 1 - \alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1 - \beta$)

Type I error is generally **more serious** than type II error.

For example: H_0 : the person is not guilty, H_1 : the person is guilty

Two Types of Errors Revisited

Recall that the following four outcomes are possible when conducting a test:

Reality	Our Decision	
	H_0	H_a
H_0	$\surd$ (Prob $= 1 - \alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1 - \beta$)

Type I error is generally **more serious** than type II error.

For example: H_0 : the person is not guilty, H_1 : the person is guilty

Type I error

Two Types of Errors Revisited

Recall that the following four outcomes are possible when conducting a test:

Reality	Our Decision H_0	Our Decision H_a
H_0	$\surd$ (Prob $= 1 - \alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1 - \beta$)

Type I error is generally **more serious** than type II error.

For example: H_0 : the person is not guilty, H_1 : the person is guilty

Type I error $=$ convicting an innocent person

Two Types of Errors Revisited

Recall that the following four outcomes are possible when conducting a test:

Reality	Our Decision H_0	Our Decision H_a
H_0	$\surd$ (Prob $= 1 - \alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1 - \beta$)

Type I error is generally **more serious** than type II error.

For example: H_0 : the person is not guilty, H_1 : the person is guilty

Type I error $=$ convicting an innocent person

Type II error

Two Types of Errors Revisited

Recall that the following four outcomes are possible when conducting a test:

Reality	Our Decision H_0	Our Decision H_a
H_0	$\surd$ (Prob $= 1 - \alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1 - \beta$)

Type I error is generally **more serious** than type II error.

For example: H_0 : the person is not guilty, H_1 : the person is guilty

Type I error $=$ convicting an innocent person

Type II error $=$ failing to convict a guilty person

Two Types of Errors Revisited

Recall that the following four outcomes are possible when conducting a test:

Reality	Our Decision H_0	H_a
H_0	$\surd$ (Prob $= 1 - \alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1 - \beta$)

Type I error is generally **more serious** than type II error.

For example: H_0 : the person is not guilty, H_1 : the person is guilty

Type I error $=$ convicting an innocent person

Type II error $=$ failing to convict a guilty person

In practice, we first choose an α and consider only tests with probability of Type I error no greater than α.

Two Types of Errors Revisited

Recall that the following four outcomes are possible when conducting a test:

Reality	Our Decision H_0	H_a
H_0	$\surd$ (Prob $= 1 - \alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1 - \beta$)

Type I error is generally **more serious** than type II error.

For example: H_0 : the person is not guilty, H_1 : the person is guilty

Type I error $=$ convicting an innocent person

Type II error $=$ failing to convict a guilty person

In practice, we first choose an α and consider only tests with probability of Type I error no greater than α. Then we select one that makes the probability of Type II error as small as possible

Two Types of Errors Revisited

Recall that the following four outcomes are possible when conducting a test:

Reality	Our Decision H_0	Our Decision H_a
H_0	$\surd$ (Prob $= 1-\alpha$)	Type I Error (Prob $= \alpha$)
H_a	Type II Error (Prob $= \beta$)	$\surd$ (Prob $= 1-\beta$)

Type I error is generally **more serious** than type II error.

For example: H_0 : the person is not guilty, H_1 : the person is guilty

Type I error $=$ convicting an innocent person

Type II error $=$ failing to convict a guilty person

In practice, we first choose an α and consider only tests with probability of Type I error no greater than α. Then we select one that makes the probability of Type II error as small as possible

Another important continuous distribution

Chi-Square Distribution

The chi-square distributions: χ^2_k

$$f(x) = \frac{x^{(k/2)-1} e^{-x/2}}{2^{k/2}\Gamma(k/2)} \qquad x > 0,\ k \geq 1$$

Chi-Square Distribution

The chi-square distributions: χ^2_k

$$f(x) = \frac{x^{(k/2)-1} e^{-x/2}}{2^{k/2}\Gamma(k/2)} \qquad x > 0,\ k \geq 1$$

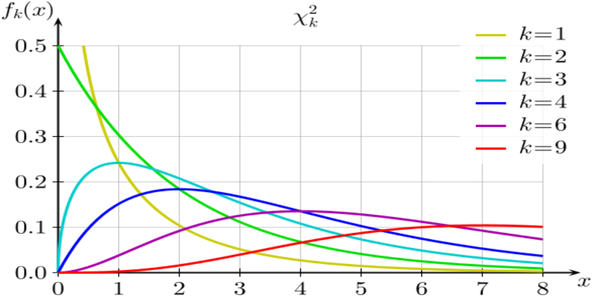
$$E(X) = k \qquad Var(X) = 2k$$

Chi-Square Distribution

The chi-square distributions: χ^2_k

$$f(x) = \frac{x^{(k/2)-1} e^{-x/2}}{2^{k/2}\Gamma(k/2)} \qquad x > 0,\ k \geq 1$$

$$E(X) = k \qquad Var(X) = 2k$$



Chi-Square Distribution

The chi-square distributions: χ_1^2

$$f(x) = \frac{x^{-1/2} e^{-x/2}}{\sqrt{2\pi}} \qquad x > 0$$

Chi-Square Distribution

The chi-square distributions: χ^2_1

$$f(x) = \frac{x^{-1/2} e^{-x/2}}{\sqrt{2\pi}} \qquad x > 0$$

Origin of the Chi-square distribution:
Let $Z \sim N(0,1)$ and define $Y = Z^2$. Then, $Y \sim \chi^2_1$.

Chi-Square Distribution

The chi-square distributions: χ_1^2

$$f(x) = \frac{x^{-1/2} e^{-x/2}}{\sqrt{2\pi}} \qquad x > 0$$

Origin of the Chi-square distribution:
Let $Z \sim N(0,1)$ and define $Y = Z^2$. Then, $Y \sim \chi_1^2$.

In general, χ_n^2 is a sum of n i.i.d. standard normal distributions.

Chi-Square Distribution

Properties:

- We will need the **reproductive property of** χ^2**:**

Chi-Square Distribution

Properties:

- We will need the **reproductive property of χ^2:**
  Let $X_1 \sim \chi^2_{n_1}, X_2 \sim \chi^2_{n_2}, \ldots, X_k \sim \chi^2_{n_k}$ with $X_1, X_2, \ldots, X_k$ independent random variables. Define $Y = X_1 + X_2 + \cdots + X_k$. Then, $Y \sim \chi^2_n$ where $n = n_1 + n_2 + \cdots + n_k$.

Chi-Square Distribution

Properties:

- We will need the **reproductive property of χ^2:** Let $X_1 \sim \chi^2_{n_1}, X_2 \sim \chi^2_{n_2}, \ldots, X_k \sim \chi^2_{n_k}$ with $X_1, X_2, \ldots, X_k$ independent random variables. Define $Y = X_1 + X_2 + \cdots + X_k$. Then, $Y \sim \chi^2_n$ where $n = n_1 + n_2 + \cdots + n_k$.
- Conversely, $X_1 \sim \chi^2_{n_1}, X_1 + X_2 \sim \chi^2_{n_1+n_2} \Rightarrow X_2 \sim \chi^2_{n_2}$.

Chi-Square Distribution

Properties:

- We will need the **reproductive property of χ^2:**
  Let $X_1 \sim \chi^2_{n_1}, X_2 \sim \chi^2_{n_2}, \ldots, X_k \sim \chi^2_{n_k}$ with $X_1, X_2, \ldots, X_k$ independent random variables. Define $Y = X_1 + X_2 + \cdots + X_k$. Then, $Y \sim \chi^2_n$ where $n = n_1 + n_2 + \cdots + n_k$.
- Conversely, $X_1 \sim \chi^2_{n_1}, X_1 + X_2 \sim \chi^2_{n_1+n_2} \Rightarrow X_2 \sim \chi^2_{n_2}$.
- If $X_i \sim N()$, then $\bar{X}$ and S^2 are independent.

Chi-Square Distribution

Properties:

- We will need the **reproductive property of χ^2:**
  Let $X_1 \sim \chi^2_{n_1}, X_2 \sim \chi^2_{n_2}, \ldots, X_k \sim \chi^2_{n_k}$ with $X_1, X_2, \ldots, X_k$ independent random variables. Define
  $Y = X_1 + X_2 + \cdots + X_k$. Then, $Y \sim \chi^2_n$ where
  $n = n_1 + n_2 + \cdots + n_k$.
- Conversely, $X_1 \sim \chi^2_{n_1}, X_1 + X_2 \sim \chi^2_{n_1+n_2} \Rightarrow X_2 \sim \chi^2_{n_2}$.
- If $X_i \sim N()$, then $\bar{X}$ and S^2 are independent.
  Reason: $Cov(\bar{X}, X_i - \bar{X}) = 0 \Rightarrow \bar{X}$ and $X_i - \bar{X}$ are uncorrelated
  $\Rightarrow \bar{X}$ and $X_i - \bar{X}$ are independent.

Chi-Square Distribution

To get started, consider a random sample $X_1, X_2, \ldots, X_n$ from a $N(\mu, \sigma^2)$ population.

Chi-Square Distribution

To get started, consider a random sample $X_1, X_2, \ldots, X_n$ from a $N(\mu, \sigma^2)$ population. Let

$$Z_i = \frac{X_i - \mu}{\sigma} \qquad \text{so that} \qquad Z_i^2 \sim \chi_1^2 \quad \text{and } Z_i's \text{ are independent.}$$

Chi-Square Distribution

To get started, consider a random sample $X_1, X_2, \ldots, X_n$ from a $N(\mu, \sigma^2)$ population. Let

$$Z_i = \frac{X_i - \mu}{\sigma} \qquad \text{so that} \qquad Z_i^2 \sim \chi_1^2 \quad \text{and } Z_i's \text{ are independent.}$$

Then,

$$\sum_{i=1}^{n} \left[\frac{X_i - \mu}{\sigma}\right]^2 = \sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} \chi_1^2 = \chi_n^2.$$

Chi-Square Distribution

To get started, consider a random sample $X_1, X_2, \ldots, X_n$ from a $N(\mu, \sigma^2)$ population. Let

$$Z_i = \frac{X_i - \mu}{\sigma} \qquad \text{so that} \qquad Z_i^2 \sim \chi_1^2 \quad \text{and } Z_i's \text{ are independent.}$$

Then,

$$\sum_{i=1}^{n} \left[\frac{X_i - \mu}{\sigma}\right]^2 = \sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} \chi_1^2 = \chi_n^2.$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

Mean of Sampling Distribution of s^2

Since we have,

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

for a random sample from a $N(\mu, \sigma^2)$ population, we have that

$$E\left[\frac{(n-1)S^2}{\sigma^2}\right] = E(\chi^2_{(n-1)}) = n-1$$

Mean of Sampling Distribution of s^2

Since we have,

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

for a random sample from a $N(\mu, \sigma^2)$ population, we have that

$$E\left[\frac{(n-1)S^2}{\sigma^2}\right] = E(\chi^2_{(n-1)}) = n-1$$

It follows that,

$$n-1 = E\left[\frac{(n-1)S^2}{\sigma^2}\right] = \frac{(n-1)}{\sigma^2}E(S^2) \Rightarrow E(s^2) = \sigma^2.$$

So, when sampling from a normal population,
s^2 is on target (σ^2) on average.

Mean of Sampling Distribution of s^2

What is the mean of the sampling distribution for S^2 for samples from *any* population?

Mean of Sampling Distribution of s^2

What is the mean of the sampling distribution for S^2 for samples from *any* population?

First, show that $E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = (n-1)\sigma^2$

Mean of Sampling Distribution of s^2

What is the mean of the sampling distribution for S^2 for samples from *any* population?

First, show that $E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = (n-1)\sigma^2$

Then, we can conclude that

$$E(S^2) = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = \sigma^2$$

for samples from *any* population.

Mean of Sampling Distribution of s^2

What is the mean of the sampling distribution for S^2 for samples from *any* population?

First, show that $E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = (n-1)\sigma^2$

Then, we can conclude that

$$E(S^2) = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = \sigma^2$$

for samples from *any* population.

This is one reason why sample variance is defined by dividing by $n-1$ instead of n. If we divide by n, the resulting sample variance will tend to underestimate σ^2 on average.

Mean of Sampling Distribution of s^2

What is the mean of the sampling distribution for S^2 for samples from *any* population?

First, show that $E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = (n-1)\sigma^2$

Then, we can conclude that

$$E(S^2) = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = \sigma^2$$

for samples from *any* population.

This is one reason why sample variance is defined by dividing by $n-1$ instead of n. If we divide by n, the resulting sample variance will tend to underestimate σ^2 on average.

The "Student's" T Probability Distribution

Let $Z \sim N(0,1)$ and $Y \sim \chi^2_\nu$ with Z, Y **independent**. Then, $T = \frac{Z}{\sqrt{Y/\nu}}$ is said to have the $T-$distribution with ν degrees of freedom (df): $T \sim t_n$.

The "Student's" T Probability Distribution

Let $Z \sim N(0,1)$ and $Y \sim \chi^2_\nu$ with Z, Y **independent**. Then, $T = \frac{Z}{\sqrt{Y/\nu}}$ is said to have the $T-$distribution with ν degrees of freedom (df): $T \sim t_n$. p.d.f:

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})}(1 + \frac{t^2}{\nu})^{-\frac{\nu+1}{2}}, -\infty < t < \infty$$

The “Student’s” T Probability Distribution

Let $Z \sim N(0,1)$ and $Y \sim \chi^2_\nu$ with Z, Y **independent**. Then, $T = \frac{Z}{\sqrt{Y/\nu}}$ is said to have the $T-$distribution with ν degrees of freedom (df): $T \sim t_n$. p.d.f:

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})}(1 + \frac{t^2}{\nu})^{-\frac{\nu+1}{2}}, -\infty < t < \infty$$

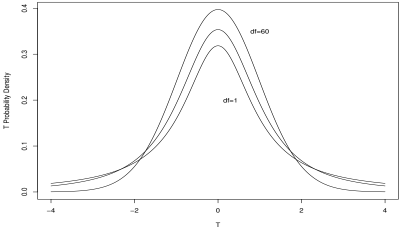
It can be shown that $E(T) = 0$ and $Var(T) = \nu/(\nu-2)$. $\nu > 2$

The "Student's" T Probability Distribution

Let $Z \sim N(0,1)$ and $Y \sim \chi^2_\nu$ with Z, Y **independent**. Then, $T = \frac{Z}{\sqrt{Y/\nu}}$ is said to have the $T-$distribution with ν degrees of freedom (df): $T \sim t_n$. p.d.f:

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})}(1+\frac{t^2}{\nu})^{-\frac{\nu+1}{2}}, -\infty < t < \infty$$

It can be shown that $E(T) = 0$ and $Var(T) = \nu/(\nu-2)$. $\nu > 2$

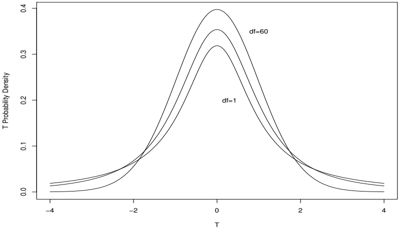


The "Student's" T Probability Distribution

Let $Z \sim N(0,1)$ and $Y \sim \chi^2_\nu$ with Z, Y **independent**. Then, $T = \frac{Z}{\sqrt{Y/\nu}}$ is said to have the $T-$distribution with ν degrees of freedom (df): $T \sim t_n$. p.d.f:

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})}(1+\frac{t^2}{\nu})^{-\frac{\nu+1}{2}}, -\infty < t < \infty$$

It can be shown that $E(T) = 0$ and $Var(T) = \nu/(\nu-2)$. $\nu > 2$



Question: What's the T distribution when $\nu \to \infty$?

How is the $T-$distribution Related to Estimation?

Let $X_1, X_2, \ldots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ popn.

We know that $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$.

How is the $T-$distribution Related to Estimation?

Let $X_1, X_2, \ldots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ popn.

We know that $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$.

But what if σ^2 is not known?

How is the $T-$distribution Related to Estimation?

Let $X_1, X_2, \ldots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ popn.

We know that $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$.

But what if σ^2 is not known? Use $s \approx \sigma$ to estimate it.

Consider

$$
\begin{aligned}
T &= \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \cdot \frac{\sigma}{s} = Z\frac{\sigma}{s} \quad (\text{where } Z \sim N(0,1)) \\
&= \frac{Z}{s/\sigma} = \frac{Z}{\sqrt{s^2/\sigma^2}} \\
&= \frac{Z}{\sqrt{\frac{(n-1)s^2}{\sigma^2(n-1)}}} = \frac{Z}{\sqrt{\chi^2_{(n-1)}/(n-1)}} \sim t_{(n-1)}
\end{aligned}
$$

How is the $T-$distribution Related to Estimation?

Let $X_1, X_2, \ldots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ popn.

We know that $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$.

But what if σ^2 is not known? Use $s \approx \sigma$ to estimate it.

Consider

$$
\begin{aligned}
T &= \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \cdot \frac{\sigma}{s} = Z\frac{\sigma}{s} \quad \text{(where } Z \sim N(0,1)\text{)} \\
&= \frac{Z}{s/\sigma} = \frac{Z}{\sqrt{s^2/\sigma^2}} \\
&= \frac{Z}{\sqrt{\frac{(n-1)s^2}{\sigma^2(n-1)}}} = \frac{Z}{\sqrt{\chi^2_{(n-1)}/(n-1)}} \sim t_{(n-1)}
\end{aligned}
$$

How is the $T-$distribution Related to Estimation?

$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$

How is the $T-$distribution Related to Estimation?

$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$

So, even when the popn variance σ^2 is not known, we can still find probabilities for the sample mean $\bar{X}$ for data from a normal popn.

How is the $T-$distribution Related to Estimation?

$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$

So, even when the popn variance σ^2 is not known, we can still find probabilities for the sample mean $\bar{X}$ for data from a normal popn.

We just replace σ by its estimate (s) in the standardization of $\bar{X}$ and then look up probabilities from the $t_{(n-1)}$ distribution instead of the $N(0,1)$ distribution.

How is the $T-$distribution Related to Estimation?

$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$

So, even when the popn variance σ^2 is not known, we can still find probabilities for the sample mean $\bar{X}$ for data from a normal popn.

We just replace σ by its estimate (s) in the standardization of $\bar{X}$ and then look up probabilities from the $t_{(n-1)}$ distribution instead of the $N(0, 1)$ distribution.

The $t_{(n-1)}$ distribution is similar to the $N(0, 1)$, but with more spread (to account for the extra uncertainty in the estimate of σ.)

How is the $T-$distribution Related to Estimation?

$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$

So, even when the popn variance σ^2 is not known, we can still find probabilities for the sample mean $\bar{X}$ for data from a normal popn.

We just replace σ by its estimate (s) in the standardization of $\bar{X}$ and then look up probabilities from the $t_{(n-1)}$ distribution instead of the $N(0,1)$ distribution.

The $t_{(n-1)}$ distribution is similar to the $N(0,1)$, but with more spread (to account for the extra uncertainty in the estimate of σ.)

As sample size increases, s gets closer to σ
and the $t_{(n-1)}$ distribution gets closer to $N(0,1)$.

How is the $T-$distribution Related to Estimation?

$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$

So, even when the popn variance σ^2 is not known, we can still find probabilities for the sample mean $\bar{X}$ for data from a normal popn.

We just replace σ by its estimate (s) in the standardization of $\bar{X}$ and then look up probabilities from the $t_{(n-1)}$ distribution instead of the $N(0,1)$ distribution.

The $t_{(n-1)}$ distribution is similar to the $N(0,1)$, but with more spread (to account for the extra uncertainty in the estimate of σ.)

As sample size increases, s gets closer to σ
and the $t_{(n-1)}$ distribution gets closer to $N(0,1)$.

t-tests

- Remember we discussed about the case σ not known very briefly?

t-tests

- Remember we discussed about the case σ not known very briefly?
- $H_0 : \mu = \mu_0$ vs $H_1 : \mu_0 \neq \mu_0$. Assume σ is known.
- $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ ..

t-tests

- Remember we discussed about the case σ not known very briefly?
- $H_0 : \mu = \mu_0$ vs $H_1 : \mu_0 \neq \mu_0$. Assume σ is known.
- $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ ..z-test

t-tests

- Remember we discussed about the case σ not known very briefly?
- $H_0 : \mu = \mu_0$ vs $H_1 : \mu_0 \neq \mu_0$. Assume σ is known.
- $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ ..z-test
- We reject H_0 if $|z| > z^* = z_{\alpha/2}$

t-tests

- Remember we discussed about the case σ not known very briefly?
- $H_0 : \mu = \mu_0$ vs $H_1 : \mu_0 \neq \mu_0$. Assume σ is known.
- $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ ..z-test
- We reject H_0 if $|z| > z^* = z_{\alpha/2}$
- If not known, replace σ by $\hat{\sigma} = s$

t-tests

- Remember we discussed about the case σ not known very briefly?
- $H_0 : \mu = \mu_0$ vs $H_1 : \mu_0 \neq \mu_0$. Assume σ is known.
- $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ ..z-test
- We reject H_0 if $|z| > z^* = z_{\alpha/2}$
- If not known, replace σ by $\hat{\sigma} = s$
- $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$

t-tests

- Remember we discussed about the case σ not known very briefly?
- $H_0 : \mu = \mu_0$ vs $H_1 : \mu_0 \neq \mu_0$. Assume σ is known.
- $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ ..z-test
- We reject H_0 if $|z| > z^* = z_{\alpha/2}$
- If not known, replace σ by $\hat{\sigma} = s$
- $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$.. t-test

t-tests

- Remember we discussed about the case σ not known very briefly?
- $H_0 : \mu = \mu_0$ vs $H_1 : \mu_0 \neq \mu_0$. Assume σ is known.
- $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ ..z-test
- We reject H_0 if $|z| > z^* = z_{\alpha/2}$
- If not known, replace σ by $\hat{\sigma} = s$
- $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$.. t-test
- If H_0 is true, observed value t follows t- distribution with degrees of freedom $n - 1$

t-tests

- Remember we discussed about the case σ not known very briefly?
- $H_0 : \mu = \mu_0$ vs $H_1 : \mu_0 \neq \mu_0$. Assume σ is known.
- $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$ ..z-test
- We reject H_0 if $|z| > z^* = z_{\alpha/2}$
- If not known, replace σ by $\hat{\sigma} = s$
- $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$.. t-test
- If H_0 is true, observed value t follows t- distribution with degrees of freedom $n - 1$
- We reject H_0 if $|t| > t^*_{(n-1)} = t_{(n-1),\alpha/2}$

Example: Mouse tumor experiment

Example: Growth of Tumor

Let X (in *mm*) denote the growth in 15 days of a tumor induced in a mouse. It is known from a previous experiment that the average tumor growth is $4mm$.

Example: Growth of Tumor

Let X (in *mm*) denote the growth in 15 days of a tumor induced in a mouse. It is known from a previous experiment that the average tumor growth is $4mm$.

A sample of 20 mice that have a genetic variant hypothesized to be involved in tumor growth yielded $\bar{x} = 3.8mm, s = 0.3mm$.

Example: Growth of Tumor

Let X (in *mm*) denote the growth in 15 days of a tumor induced in a mouse. It is known from a previous experiment that the average tumor growth is $4mm$.

A sample of 20 mice that have a genetic variant hypothesized to be involved in tumor growth yielded $\bar{x} = 3.8mm, s = 0.3mm$.

Test whether $\mu = 4$ or not, assuming growths are normally distributed.

Example: Growth of Tumor (Cont.)

1. State the hypotheses

$$H_0 : \mu = 4 \qquad H_a : \mu \neq 4$$

Example: Growth of Tumor (Cont.)

1. State the hypotheses

$$H_0 : \mu = 4 \qquad H_a : \mu \neq 4$$

2. Calculate the t-statistic

$$t = \frac{3.8 - 4.0}{0.3/\sqrt{20}} = -2.98$$

Example: Growth of Tumor (Cont.)

1. State the hypotheses

$$H_0 : \mu = 4 \qquad H_a : \mu \neq 4$$

2. Calculate the t-statistic

$$t = \frac{3.8 - 4.0}{0.3/\sqrt{20}} = -2.98$$

3. Determine the *p*-value

$$p = \mathrm{P}(t_{19} \geq 2.98) + \mathrm{P}(t_{19} \leq -2.98) = 0.008$$

Example: Growth of Tumor (Cont.)

1. State the hypotheses

$$H_0 : \mu = 4 \qquad H_a : \mu \neq 4$$

2. Calculate the t-statistic

$$t = \frac{3.8 - 4.0}{0.3/\sqrt{20}} = -2.98$$

3. Determine the p-value

$$p = \mathrm{P}(t_{19} \geq 2.98) + \mathrm{P}(t_{19} \leq -2.98) = 0.008$$

Since p is less than 0.01, we reject H_0 at significance level $\alpha = 0.01$. There is evidence that the population mean growth is not $4mm$.

Example: Growth of Tumor (Cont.)

1. State the hypotheses

$$H_0 : \mu = 4 \qquad H_a : \mu \neq 4$$

2. Calculate the t-statistic

$$t = \frac{3.8 - 4.0}{0.3/\sqrt{20}} = -2.98$$

3. Determine the *p*-value

$$p = \mathrm{P}(t_{19} \geq 2.98) + \mathrm{P}(t_{19} \leq -2.98) = 0.008$$

Since p is less than 0.01, we reject H_0 at significance level $\alpha = 0.01$. There is evidence that the population mean growth is not $4mm$.

What if we wanted a 99% CI for μ instead? The CI is given by

$$\left(\bar{x} - t^* \frac{s}{\sqrt{n}},\ \bar{x} + t^* \frac{s}{\sqrt{n}}\right) = \left(3.8 - 2.861 \times \frac{0.3}{\sqrt{20}},\ 3.8 + 2.861 \times \frac{0.3}{\sqrt{20}}\right)$$

$$= (3.61, 3.99)$$

Example: Growth of Tumor (Cont.)

1. State the hypotheses

$$H_0 : \mu = 4 \qquad H_a : \mu \neq 4$$

2. Calculate the t-statistic

$$t = \frac{3.8 - 4.0}{0.3/\sqrt{20}} = -2.98$$

3. Determine the p-value

$$p = \mathrm{P}(t_{19} \geq 2.98) + \mathrm{P}(t_{19} \leq -2.98) = 0.008$$

Since p is less than 0.01, we reject H_0 at significance level $\alpha = 0.01$. There is evidence that the population mean growth is not $4mm$.
What if we wanted a 99% CI for μ instead? The CI is given by

$$\left(\bar{x} - t^* \frac{s}{\sqrt{n}},\ \bar{x} + t^* \frac{s}{\sqrt{n}}\right) = \left(3.8 - 2.861 \times \frac{0.3}{\sqrt{20}},\ 3.8 + 2.861 \times \frac{0.3}{\sqrt{20}}\right)$$

$$= (3.61, 3.99)$$

Confidence intervals for t distribution

$$(\bar{x} - t^* \frac{s}{\sqrt{n}}, \bar{x} + t^* \frac{s}{\sqrt{n}})$$

is the confidence interval for a two-sided hypothesis. If the population is normal then this is exact, otherwise it is approximate.

Confidence intervals for t distribution

$$(\bar{x} - t^* \frac{s}{\sqrt{n}}, \bar{x} + t^* \frac{s}{\sqrt{n}})$$

is the confidence interval for a two-sided hypothesis. If the population is normal then this is exact, otherwise it is approximate.

- σ is known,

$$(\bar{x} - z^* \frac{\sigma}{\sqrt{n}}, \bar{x} + z^* \frac{\sigma}{\sqrt{n}})$$

Confidence intervals for t distribution

$$(\bar{x} - t^* \frac{s}{\sqrt{n}}, \bar{x} + t^* \frac{s}{\sqrt{n}})$$

is the confidence interval for a two-sided hypothesis. If the population is normal then this is exact, otherwise it is approximate.

- σ is known,

$$(\bar{x} - z^* \frac{\sigma}{\sqrt{n}}, \bar{x} + z^* \frac{\sigma}{\sqrt{n}})$$

- σ is not known but we still use z^*,

$$(\bar{x} - z^* \frac{s}{\sqrt{n}}, \bar{x} + z^* \frac{s}{\sqrt{n}})$$

Confidence intervals for t distribution

$$(\bar{x} - t^* \frac{s}{\sqrt{n}}, \bar{x} + t^* \frac{s}{\sqrt{n}})$$

is the confidence interval for a two-sided hypothesis. If the population is normal then this is exact, otherwise it is approximate.

- σ is known,
$$(\bar{x} - z^* \frac{\sigma}{\sqrt{n}}, \bar{x} + z^* \frac{\sigma}{\sqrt{n}})$$
- σ is not known but we still use z^*,
$$(\bar{x} - z^* \frac{s}{\sqrt{n}}, \bar{x} + z^* \frac{s}{\sqrt{n}})$$
-
$$(\bar{x} - t^* \frac{s}{\sqrt{n}}, \bar{x} + t^* \frac{s}{\sqrt{n}})$$

Confidence intervals for t distribution

$$(\bar{x} - t^* \frac{s}{\sqrt{n}}, \bar{x} + t^* \frac{s}{\sqrt{n}})$$

is the confidence interval for a two-sided hypothesis. If the population is normal then this is exact, otherwise it is approximate.

- σ is known,

$$(\bar{x} - z^* \frac{\sigma}{\sqrt{n}}, \bar{x} + z^* \frac{\sigma}{\sqrt{n}})$$

- σ is not known but we still use z^*,

$$(\bar{x} - z^* \frac{s}{\sqrt{n}}, \bar{x} + z^* \frac{s}{\sqrt{n}})$$

-

$$(\bar{x} - t^* \frac{s}{\sqrt{n}}, \bar{x} + t^* \frac{s}{\sqrt{n}})$$

Power of a t-test

- $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$

Power of a t-test

- $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$
- Find the rejection rule based on $\bar{x}$ and s

Power of a t-test

- $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$
- Find the rejection rule based on $\bar{x}$ and s
- Decide on a specific alternative where you want to compute the power . Say $\mu = 1$

Power of a t-test

- $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$
- Find the rejection rule based on $\bar{x}$ and s
- Decide on a specific alternative where you want to compute the power . Say $\mu = 1$
- Find $P(\text{Rejection rule}|\mu = 1)$. Difficulty?

Power of a t-test

- $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$
- Find the rejection rule based on $\bar{x}$ and s
- Decide on a specific alternative where you want to compute the power . Say $\mu = 1$
- Find $P(\text{Rejection rule}|\mu = 1)$. Difficulty?
- Assume s to be the true value of σ and use normal probabilities to get the *approximate* power.

When to use a t- test

- **Robust:** An inference procedure is called robust if the required probability calculations are insenssitive to violations of the assumptions made.

When to use a t- test

- **Robust:** An inference procedure is called robust if the required probability calculations are insenssitive to violations of the assumptions made.
- t involves $\bar{x}$ and s which are not robust to presence of outliers. So, t-test is *not* robust in the presence of outlier

When to use a t- test

- **Robust:** An inference procedure is called robust if the required probability calculations are insenssitive to violations of the assumptions made.
- t involves $\bar{x}$ and s which are not robust to presence of outliers. So, t-test is *not* robust in the presence of outlier
- If $n > 40$ then safely use t- test even if skewed

When to use a t- test

- **Robust:** An inference procedure is called robust if the required probability calculations are insenssitive to violations of the assumptions made.
- t involves $\bar{x}$ and s which are not robust to presence of outliers. So, t-test is *not* robust in the presence of outlier
- If $n > 40$ then safely use t- test even if skewed
- If $15 < n < 40$ and no outliers or strong skewness ..use t

When to use a t- test

- **Robust:** An inference procedure is called robust if the required probability calculations are insenssitive to violations of the assumptions made.
- t involves $\bar{x}$ and s which are not robust to presence of outliers. So, t-test is *not* robust in the presence of outlier
- If $n > 40$ then safely use t- test even if skewed
- If $15 < n < 40$ and no outliers or strong skewness ..use t
- If $n < 15$ and data looks normal and no outlier, use t

When to use a t- test

- **Robust:** An inference procedure is called robust if the required probability calculations are insenssitive to violations of the assumptions made.
- t involves $\bar{x}$ and s which are not robust to presence of outliers. So, t-test is *not* robust in the presence of outlier
- If $n > 40$ then safely use t- test even if skewed
- If $15 < n < 40$ and no outliers or strong skewness ..use t
- If $n < 15$ and data looks normal and no outlier, use t
- These are the rules in textbooks. In this course, we will use resampling techniques to be more sure before using t-tests.