Econ 103: Topics in Single Linear Regression

Manu Navjeevan

UCLA

July 25, 2021

Content Outline

Advanced Inference Topics

- Inference on Linear Combinations of Parameters
- Heteroskedasticity

Evaluating our Model

- R^2 and goodness of fit

Modeling Choices

- How do results change if we apply linear transformations?
- Useful non-linear transformations of X and Y

Table of Contents

Inference: Linear Combinations of Parameters

Recall that, approximately for large n

$$\sqrt{n}(\hat{\beta}_0 - \beta_0) \sim N\left(0, \mathbb{E}[X^2]\sigma_\epsilon^2/\sigma_X^2\right), \quad \sqrt{n}(\hat{\beta}_1 - \beta_1) \sim N\left(0, \sigma_\epsilon^2/\sigma_X^2\right)$$

and $\sigma_{\beta_{01}} = \mathrm{Cov}(\sqrt{n}\{\hat{\beta}_0 - \beta_0\}, \sqrt{n}\{\hat{\beta}_1 - \beta_1\}) = -\mathbb{E}[X]\frac{\sigma_\epsilon^2}{\sigma_X^2}$.

Inference: Linear Combinations of Parameters

Recall that, approximately for large n

$$\sqrt{n}(\hat{\beta}_0 - \beta_0) \sim N\left(0, \mathbb{E}[X^2]\sigma_\epsilon^2/\sigma_X^2\right), \quad \sqrt{n}(\hat{\beta}_1 - \beta_1) \sim N\left(0, \sigma_\epsilon^2/\sigma_X^2\right)$$

and $\sigma_{\beta_{01}} = \mathrm{Cov}(\sqrt{n}\{\hat{\beta}_0 - \beta_0\}, \sqrt{n}\{\hat{\beta}_1 - \beta_1\}) = -\mathbb{E}[X]\frac{\sigma_\epsilon^2}{\sigma_X^2}$.

These results were also often presented in the following equivalent manners

$$\frac{\hat{\beta}_0 - \beta_0}{\sigma_{\beta_0}/\sqrt{n}} \sim N(0,1) \text{ and } \hat{\beta}_0 \sim N(\beta_1, \sigma_{\beta_0}^2/n)$$
$$\frac{\hat{\beta}_1 - \beta_1}{\sigma_{\beta_1}/\sqrt{n}} \sim N(0,1) \text{ and } \hat{\beta}_1 \sim N(\beta_1, \sigma_{\beta_1}^2/n)$$

where $\sigma_{\beta_0}^2 = \mathbb{E}[X^2]\sigma_{\epsilon^2}/\sigma_X^2$ and $\sigma_{\beta_1}^2 = \sigma_\epsilon^2/\sigma_X^2$.

Inference: Linear Combinations of Parameters

Recall that, approximately for large n

$$\sqrt{n}(\hat{\beta}_0 - \beta_0) \sim N\left(0, \mathbb{E}[X^2]\sigma_\epsilon^2/\sigma_X^2\right), \quad \sqrt{n}(\hat{\beta}_1 - \beta_1) \sim N\left(0, \sigma_\epsilon^2/\sigma_X^2\right)$$

and $\sigma_{\beta_{01}} = \mathrm{Cov}(\sqrt{n}\{\hat{\beta}_0 - \beta_0\}, \sqrt{n}\{\hat{\beta}_1 - \beta_1\}) = -\mathbb{E}[X]\frac{\sigma_\epsilon^2}{\sigma_X^2}$.

These results were also often presented in the following equivalent manners

$$\frac{\hat{\beta}_0 - \beta_0}{\sigma_{\beta_0}/\sqrt{n}} \sim N(0,1) \text{ and } \hat{\beta}_0 \sim N(\beta_1, \sigma_{\beta_0}^2/n)$$
$$\frac{\hat{\beta}_1 - \beta_1}{\sigma_{\beta_1}/\sqrt{n}} \sim N(0,1) \text{ and } \hat{\beta}_1 \sim N(\beta_1, \sigma_{\beta_1}^2/n)$$

where $\sigma_{\beta_0}^2 = \mathbb{E}[X^2]\sigma_{\epsilon^2}/\sigma_X^2$ and $\sigma_{\beta_1}^2 = \sigma_\epsilon^2/\sigma_X^2$.

- Also went over how to estiamte these variances

Inference: Linear Combinations of Parameters

In the last lecture, we used these distributional results to compute objects like

$$\Pr\left(|\hat{\beta}_1| > 5 \mid \beta_1 = 0\right).$$

which in turn were useful for hypothesis testing

$$H_0 : \beta_1 = 0 \text{ vs. } H_1 : \beta_1 \neq 0.$$

Inference: Linear Combinations of Parameters

However, often we want to preform inference not just on one parameter, but on a linear combination of parameters, i.e we want to test

$$H_0 : \beta_0 + 5\beta_1 = 0 \text{ vs. } H_1 : \beta_0 + 5\beta_1 \neq 0.$$

Inference: Linear Combinations of Parameters

However, often we want to preform inference not just on one parameter, but on a linear combination of parameters, i.e we want to test

$$H_0 : \beta_0 + 5\beta_1 = 0 \text{ vs. } H_1 : \beta_0 + 5\beta_1 \neq 0.$$

This is useful, for example, if we are trying to test something like

$$H_0 : \mathbb{E}[Y|X = 5] = 0 \text{ vs. } H_1 : \mathbb{E}[Y|X = 5] \neq 0$$

and we view the linear regression model $Y = \beta_0 + \beta_1 X + \epsilon$ as a way of approximating the conditional mean function $\mathbb{E}[Y|X = x]$.

Inference: Linear Combinations of Parameters

In order to test such a hypothesis we want to know the distribution of a linear combination of our model parameters. That is, for $\lambda = a\beta_0 + b\beta_1$ we would like to know the approximate distribution of

$$\hat{\lambda} = a\hat{\beta}_0 + b\hat{\beta}_1$$

so that we can calculate objects like $\Pr(|\hat{\lambda}| > 0.5 \mid \lambda = 0)$.

Inference: Linear Combinations of Parameters

In order to test such a hypothesis we want to know the distribution of a linear combination of our model parameters. That is, for $\lambda = a\beta_0 + b\beta_1$ we would like to know the approximate distribution of

$$\hat{\lambda} = a\hat{\beta}_0 + b\hat{\beta}_1$$

so that we can calculate objects like $\Pr(|\hat{\lambda}| > 0.5 \mid \lambda = 0)$.

Note that

$$\begin{aligned}\sqrt{n}\left(\hat{\lambda} - \lambda\right) &= \sqrt{n}\left(a\hat{\beta}_0 + b\hat{\beta}_1 - a\beta_0 - b\beta_1\right) \\ &= a\sqrt{n}\left(\hat{\beta}_0 - \beta_0\right) + b\sqrt{n}\left(\hat{\beta}_1 - \beta_1\right).\end{aligned}$$

and that we know the (joint) distribution of $\sqrt{n}(\hat{\beta}_0 - \beta_0)$ and $\sqrt{n}(\hat{\beta}_1 - \beta_1)$.

Inference: Linear Combinations of Parameters

Recall from our Econ 41 Review that the sum of two jointly normal random variables is also normally distributed and that if X and Y are random variables then

$$\mathrm{Var}(aX + bY) = a^2 \,\mathrm{Var}(X) + b^2 \,\mathrm{Var}(Y) + 2ab \,\mathrm{Cov}(X, Y).$$

Inference: Linear Combinations of Parameters

Recall from our Econ 41 Review that the sum of two jointly normal random variables is also normally distributed and that if X and Y are random variables then

$$\mathrm{Var}(aX + bY) = a^2 \mathrm{Var}(X) + b^2 \mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X, Y).$$

Using this result along with $X = \sqrt{n}(\hat{\beta}_0 - \beta_0)$ and $Y = \sqrt{n}(\hat{\beta}_1 - \beta_1)$ gives us that, for large n:

$$\sqrt{n}\left(\hat{\lambda} - \lambda\right) \sim N(0, \sigma_\lambda^2) \implies \frac{\hat{\lambda} - \lambda}{\sigma_\lambda / \sqrt{n}} \sim N(0, 1),$$

where $\sigma_\lambda^2 = a^2\sigma_{\beta_0}^2 + b^2\sigma_{\beta_1}^2 + 2ab\sigma_{\beta_{01}}$

Inference: Linear Combinations of Parameters

As a reminder, we can estimate

$$\sigma^2_{\beta_0} = \mathbb{E}[X^2]\frac{\sigma^2_\epsilon}{\sigma^2_X} \iff \hat{\sigma}^2_{\beta_0} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \cdot \frac{\hat{\sigma}^2_\epsilon}{\hat{\sigma}^2_X}$$

$$\sigma^2_{\beta_1} = \frac{\sigma^2_\epsilon}{\sigma^2_X} \iff \hat{\sigma}^2_{\beta_1} = \frac{\hat{\sigma}^2_\epsilon}{\hat{\sigma}^2_X}$$

$$\sigma_{\beta_{01}} = -\mathbb{E}[X]\frac{\sigma^2_\epsilon}{\sigma^2_X} \iff \hat{\sigma}_{\beta_{01}} = \bar{X}\frac{\hat{\sigma}^2_\epsilon}{\hat{\sigma}^2_X}$$

As a reminder, we can estimate

$$\sigma^2_{\beta_0} = \mathbb{E}[X^2]\frac{\sigma^2_\epsilon}{\sigma^2_X} \iff \hat{\sigma}^2_{\beta_0} = \frac{1}{n}\sum_{i=1}^n X_i^2 \cdot \frac{\hat{\sigma}^2_\epsilon}{\hat{\sigma}^2_X}$$

$$\sigma^2_{\beta_1} = \frac{\sigma^2_\epsilon}{\sigma^2_X} \iff \hat{\sigma}^2_{\beta_1} = \frac{\hat{\sigma}^2_\epsilon}{\hat{\sigma}^2_X}$$

$$\sigma_{\beta_{01}} = -\mathbb{E}[X]\frac{\sigma^2_\epsilon}{\sigma^2_X} \iff \hat{\sigma}_{\beta_{01}} = \bar{X}\frac{\hat{\sigma}^2_\epsilon}{\hat{\sigma}^2_X}$$

So, we can use these to estimate $\sigma^2_\lambda = a^2\sigma^2_{\beta_0} + b^2\sigma^2_{\beta_1} + 2ab\sigma_{\beta_{01}}$ with

$$\hat{\sigma}^2_\lambda = a^2\hat{\sigma}^2_{\beta_0} + b^2\hat{\sigma}^2_{\beta_1} + 2ab\hat{\sigma}_{\beta_{01}}.$$

As $n \to \infty$, $\hat{\sigma}^2_{\beta_0} \to \sigma^2_{\beta_0}$, $\hat{\sigma}^2_{\beta_1} \to \sigma^2_{\beta_1}$, and $\hat{\sigma}_{\beta_{01}} \to \sigma_{\beta_{01}}$ by the Law of Large Numbers. This gives us that $\hat{\sigma}^2_\lambda \to \sigma^2_\lambda$ as $n \to \infty$ so that we can say (approximately for large n):

$$\frac{\hat{\lambda} - \lambda}{\hat{\sigma}_\lambda / \sqrt{n}} \sim N(0, 1).$$

As $n \to \infty$, $\hat{\sigma}^2_{\beta_0} \to \sigma^2_{\beta_0}$, $\hat{\sigma}^2_{\beta_1} \to \sigma^2_{\beta_1}$, and $\hat{\sigma}_{\beta_{01}} \to \sigma_{\beta_{01}}$ by the Law of Large Numbers. This gives us that $\hat{\sigma}^2_{\lambda} \to \sigma^2_{\lambda}$ as $n \to \infty$ so that we can say (approximately for large n):

$$\frac{\hat{\lambda} - \lambda}{\hat{\sigma}_{\lambda}/\sqrt{n}} \sim N(0,1).$$

As when considering just $\hat{\beta}_0$ or $\hat{\beta}_1$, this distributional result will be useful for hypothesis testing and creating confidence intervals.

Inference: Linear Combinations of Parameters

Using the distributional result:

$$\frac{\hat{\lambda} - \lambda}{\hat{\sigma}_\lambda/\sqrt{n}} \sim N(0, 1),$$

we can test a null hypothesis of the form $H_0 : \lambda \leq \ell$, $H_0 : \lambda \geq \ell$, or $H_0 : \lambda = \ell$ by first constructing our test statistic

$$t^* = \frac{\hat{\lambda} - \ell}{\hat{\sigma}_\lambda/\sqrt{n}}.$$

Inference: Linear Combinations of Parameters

Using the distributional result:

$$\frac{\hat{\lambda} - \lambda}{\hat{\sigma}_\lambda / \sqrt{n}} \sim N(0, 1),$$

we can test a null hypothesis of the form $H_0 : \lambda \leq \ell$, $H_0 : \lambda \geq \ell$, or $H_0 : \lambda = \ell$ by first constructing our test statistic

$$t^* = \frac{\hat{\lambda} - \ell}{\hat{\sigma}_\lambda / \sqrt{n}}.$$

As before, we want to reject our null hypothesis if the probability of obtaining our test statistic (or something even further from the null hypothesis) under the null hyptohesis is less than or equal to some pre-specified value λ.

Inference: Linear Combinations of Parameters

Using the distributional result:

$$\frac{\hat{\lambda} - \lambda}{\hat{\sigma}_\lambda / \sqrt{n}} \sim N(0,1),$$

we can test a null hypothesis of the form $H_0 : \lambda \leq \ell$, $H_0 : \lambda \geq \ell$, or $H_0 : \lambda = \ell$ by first constructing our test statistic

$$t^* = \frac{\hat{\lambda} - \ell}{\hat{\sigma}_\lambda / \sqrt{n}}.$$

As before, we want to reject our null hypothesis if the probability of obtaining our test statistic (or something even further from the null hypothesis) under the null hyptohesis is less than or equal to some pre-specified value λ.

- Recall that by the distributional result, under the null $t^* \sim N(0,1)$
- The quantity $\hat{\sigma}_\lambda / \sqrt{n}$ is called the standard error of $\hat{\lambda}$.

Now that we have constructed our test statistic t^* we can conduct our test in two (equivalent) ways, as before

1. Construct a p-value and reject if $p < \alpha$:
   - If $H_0 : \lambda \leq \ell$ and $H_1 : \lambda > \ell$:
   $$p = \Pr(Z \geq t^*).$$
   - If $H_0 : \lambda \geq \ell$ and $H_1 : \lambda < \ell$:
   $$p = \Pr(Z \leq t^*).$$
   - If $H_0 : \lambda = \ell$ and $H_1 : \lambda \neq \ell$:
   $$p = \Pr(|Z| \geq |t^*|) = 2\Pr(Z \geq |t^*|).$$

Inference: Linear Combinations of Parameters

Now that we have constructed our test statistic t^* we can conduct our test in two (equivalent) ways, as before

2. Compare the t statistic to the $1-\alpha$ or $1-\alpha/2$ quantile of the standard normal distribution: $z_{1-\alpha}$ or $z_{1-\alpha/2}$.
   - If $H_0 : \lambda \leq \ell$ and $H_1 : \lambda > \ell$ reject if
   $$t^* \geq z_{1-\alpha}.$$
   - If $H_0 : \lambda \geq \ell$ and $H_1 : \lambda < \ell$ reject if
   $$t^* \leq -z_{1-\alpha}.$$
   - If $H_0 : \lambda = \ell$ and $H_1 : \lambda \neq \ell$ reject if
   $$|t^*| \geq z_{1-\alpha/2}.$$

   As a reminder $z_{1-\alpha}$ and $z_{1-\alpha/2}$ are such that
   $$\Pr(Z \leq z_{1-\alpha}) = 1-\alpha \iff \Pr(Z > z_{1-\alpha}) = \alpha$$
   $$\Pr(Z \leq z_{1-\alpha/2}) = 1-\alpha/2 \iff \Pr(|Z| > z_{1-\alpha/2}) = \alpha$$

We can also construct a $100(1-\alpha)\%$ confidence interval for λ in the same way as before: by looking at the values of ℓ for which we would fail to reject the null hypothesis $H_0 : \lambda = \ell$ against a two-sided alternative $H_1 : \lambda \neq \ell$ at level α.

Inference: Linear Combinations of Parameters

We can also construct a $100(1-\alpha)\%$ confidence interval for λ in the same way as before: by looking at the values of ℓ for which we would fail to reject the null hypothesis $H_0 : \lambda = \ell$ against a two-sided alternative $H_1 : \lambda \neq \ell$ at level α.

This gives us a symmetric formula as before, a $100(1-\alpha)\%$ confidence interval for λ is given

$$\hat{\lambda} \pm z_{1-\alpha/2} \frac{\hat{\sigma}_\lambda}{\sqrt{n}}.$$

Inference: Linear Combinations of Parameters

As an aside, we can start to see a pattern here. Essentially anytime we have a distributional result like

$$\frac{\text{Estimator} - \text{True Value}}{\text{Standard Error of Estimator}} \sim N(0,1).$$

we can test a null hypothesis by constructing our test statistic

$$t^* = \frac{\text{Estimator} - \text{Null Hypothesis Value}}{\text{Standard Error of Estimate}}.$$

and then computing a p-value or directly compating this test statistic to $z_{1-\alpha}$, $-z_{1-\alpha}$, or $z_{1-\alpha/2}$ (depending on what alternate hypothesis we are testing).

We can also use this distributional result to generate $100(1-\alpha)\%$ confidence intervals for the true value via

$$\text{Estimator} \pm z_{1-\alpha/2} \cdot \text{Standard Error of Estimator.}$$

Linear Combinations of Parameters: Questions

Questions?

Inference: Linear Combinations of Parameters

Example: Suppose we are arguing with our professional colleague Kyle Kuzma about the relationship between number of mental health days taken in a month (X) and the average number of points per game scored in the NBA (Y). Kuzma claims that $\mathbb{E}[Y|X=3]=20$, we want to test this claim at level $\alpha=0.05$.

Example: Suppose we are arguing with our professional colleague Kyle Kuzma about the relationship between number of mental health days taken in a month (X) and the average number of points per game scored in the NBA (Y). Kuzma claims that $\mathbb{E}[Y|X = 3] = 20$, we want to test this claim at level $\alpha = 0.05$.

First we collect a random sample of 49 NBA players and ask them how many mental health days they took this month and their average points per game, $\{Y_i, X_i\}_{i=1}^{49}$. Then, since we believe the relationship between Y and X to be linear, we estimate the linear model

$$Y = \beta_0 + \beta_1 \cdot X + \epsilon.$$

Example: Suppose we are arguing with our professional colleague Kyle Kuzma about the relationship between number of mental health days taken in a month (X) and the average number of points per game scored in the NBA (Y). Kuzma claims that $\mathbb{E}[Y|X=3]=20$, we want to test this claim at level $\alpha=0.05$.

First we collect a random sample of 49 NBA players and ask them how many mental health days they took this month and their average points per game, $\{Y_i, X_i\}_{i=1}^{49}$. Then, since we believe the relationship between Y and X to be linear, we estimate the linear model

$$Y=\beta_0+\beta_1\cdot X+\epsilon.$$

We can then estimate $\mathbb{E}[Y|X=3]$ by $\hat{\beta}_0+3\hat{\beta}_1$.

Inference: Linear Combinations of Parameters

We can test Kuzma's claim that $\mathbb{E}[Y|X = 3] = 20$ by running the following hypothesis test

$$H_0 : \beta_0 + 3\beta_1 = 20 \text{ vs. } H_1 : \beta_0 + 3\beta_1 \neq 20.$$

Inference: Linear Combinations of Parameters

We can test Kuzma's claim that $\mathbb{E}[Y|X=3]=20$ by running the following hypothesis test

$$H_0: \beta_0 + 3\beta_1 = 20 \quad \text{vs.} \quad H_1: \beta_0 + 3\beta_1 \neq 20.$$

To test this claim we use our data (with $n = 49$) to estimate

$$\hat{\beta}_0 = 10, \quad \hat{\beta}_1 = 3$$
$$\hat{\sigma}^2_{\beta_0} = \hat{\sigma}^2_{\beta_0} = \hat{\sigma}_{\beta_{01}} = 1$$

Inference: Linear Combinations of Parameters

We can test Kuzma's claim that $\mathbb{E}[Y|X=3]=20$ by running the following hypothesis test

$$H_0 : \beta_0 + 3\beta_1 = 20 \text{ vs. } H_1 : \beta_0 + 3\beta_1 \neq 20.$$

To test this claim we use our data (with $n = 49$) to estimate

$$\hat{\beta}_0 = 10, \quad \hat{\beta}_1 = 3$$
$$\hat{\sigma}^2_{\beta_0} = \hat{\sigma}^2_{\beta_0} = \hat{\sigma}_{\beta_{01}} = 1$$

Using these estimates we get

$$\hat{\lambda} = \hat{\beta}_0 + 3\hat{\beta}_1 = 19$$
$$\hat{\sigma}^2_{\lambda} = \hat{\sigma}^2_{\beta_0} + 9\hat{\sigma}^2_{\beta_1} + 6\hat{\sigma}_{\beta_{01}} = 16$$

We can test Kuzma's claim that $\mathbb{E}[Y|X=3]=20$ by running the following hypothesis test

$$H_0 : \beta_0 + 3\beta_1 = 20 \text{ vs. } H_1 : \beta_0 + 3\beta_1 \neq 20.$$

To test this claim we use our data (with $n = 49$) to estimate

$$\hat{\beta}_0 = 10, \quad \hat{\beta}_1 = 3$$
$$\hat{\sigma}^2_{\beta_0} = \hat{\sigma}^2_{\beta_0} = \hat{\sigma}_{\beta_{01}} = 1$$

Using these estimates we get

$$\hat{\lambda} = \hat{\beta}_0 + 3\hat{\beta}_1 = 19$$
$$\hat{\sigma}^2_{\lambda} = \hat{\sigma}^2_{\beta_0} + 9\hat{\sigma}^2_{\beta_1} + 6\hat{\sigma}_{\beta_{01}} = 16$$

- Notice how much larger $\hat{\sigma}^2_{\lambda}$ is than $\hat{\sigma}^2_{\beta_0}$ or $\hat{\sigma}^2_{\beta_1}$.

Using $\hat{\lambda} = 19$, $\hat{\sigma}_\lambda^2 = 16$, and $n = 49$ we can construct our test statistic for $H_0 : \lambda = 20$ vs $H_1 : \lambda \neq 20$

$$t^* = \frac{\hat{\lambda} - 20}{\hat{\sigma}_\lambda / \sqrt{n}} = \frac{19 - 20}{\sqrt{16}/\sqrt{49}} = -\frac{1}{4/7} = -\frac{7}{4} = -1.75.$$

Inference: Linear Combinations of Parameters

Using $\hat{\lambda} = 19$, $\hat{\sigma}_\lambda^2 = 16$, and $n = 49$ we can construct our test statistic for $H_0 : \lambda = 20$ vs $H_1 : \lambda \neq 20$

$$t^* = \frac{\hat{\lambda} - 20}{\hat{\sigma}_\lambda / \sqrt{n}} = \frac{19 - 20}{\sqrt{16}/\sqrt{49}} = -\frac{1}{4/7} = -\frac{7}{4} = -1.75.$$

We'll run our test in two ways. First, let's compute our p-value

$$p = \Pr(|Z| \geq |-1.75|) = 2\Pr(Z \geq 1.75) = 2(1 - \Pr(Z \leq 1.75)) = 2 \cdot 0.04 = 0.08.$$

Since $0.08 > 0.05$ we fail to reject Kuzma's claim.

Inference: Linear Combinations of Parameters

Using $\hat{\lambda} = 19$, $\hat{\sigma}_\lambda^2 = 16$, and $n = 49$ we can construct our test statistic for $H_0 : \lambda = 20$ vs $H_1 : \lambda \neq 20$

$$t^* = \frac{\hat{\lambda} - 20}{\hat{\sigma}_\lambda / \sqrt{n}} = \frac{19 - 20}{\sqrt{16}/\sqrt{49}} = -\frac{1}{4/7} = -\frac{7}{4} = -1.75.$$

We'll run our test in two ways. Next, let's compare our test statistic to $z_{1-\alpha/2}$. Since $\alpha = 2$ we get that $z_{1-\alpha/2} = z_{0.975} = 1.96$. Because

$$|t^*| = 1.75 < 1.96 = z_{0.975}$$

we again fail to reject Kuzma's claim

Inference: Linear Combinations of Parameters

Let's use these same estimates, $\hat{\lambda} = 19$ and $\hat{\sigma}^2_\lambda = 16$, to construct a 95% confidence interval for the true parameter $\lambda = \beta_0 + 3\beta_1$.

Inference: Linear Combinations of Parameters

Let's use these same estimates, $\hat{\lambda} = 19$ and $\hat{\sigma}^2_{\lambda} = 16$, to construct a 95% confidence interval for the true parameter $\lambda = \beta_0 + 3\beta_1$.

- Since we have assumed that the true relationship between Y (points per game) and X (number of mental health days taken per month) is linear then $\lambda = \mathbb{E}[Y|X = 3]$.
    - By linear we mean that $\mathbb{E}[Y|X = x] = \beta_0 + \beta_1 \cdot x$
    - Otherwise we can view $\lambda = \beta_0 + 3\beta_1$ as an approximation of $\mathbb{E}[Y|X = 3]$.

Inference: Linear Combinations of Parameters

Let's use these same estimates, $\hat{\lambda} = 19$ and $\hat{\sigma}_\lambda^2 = 16$, to construct a 95% confidence interval for the true parameter $\lambda = \beta_0 + 3\beta_1$.

From above we have that a 95% confidence interval for λ can be constructed

$$\hat{\lambda} \pm z_{0.975} \frac{\hat{\sigma}_\lambda}{\sqrt{n}} = 19 \pm 1.96 \frac{7}{4}.$$

So that we are 95% confident that the true value of $\lambda = \beta_0 + 3\beta_1$ lies in the inteval $[15.57, 22.43]$.

- How could we use this interval to test the hypothesis $H_0 : \lambda = 20$ vs $H_1 : \lambda \neq 20$ at level $\alpha = 0.05$?

Let's use these same estimates, $\hat{\lambda} = 19$ and $\hat{\sigma}_{\lambda}^2 = 16$, to construct a 95% confidence interval for the true parameter $\lambda = \beta_0 + 3\beta_1$.

From above we have that a 95% confidence interval for λ can be constructed

$$\hat{\lambda} \pm z_{0.975}\frac{\hat{\sigma}_{\lambda}}{\sqrt{n}} = 19 \pm 1.96\frac{7}{4}.$$

So that we are 95% confident that the true value of $\lambda = \beta_0 + 3\beta_1$ lies in the inteval $[15.57, 22.43]$.

- How could we use this interval to test the hypothesis $H_0 : \lambda = 20$ vs $H_1 : \lambda \neq 20$ at level $\alpha = 0.05$?
- What about testing this hypothesis at level $\alpha = 0.01$?

Inference: Heteroskedasticitt

Table of Contents

Table of Contents