Interval Estimation and Confidence Intervals

From Point Estimates to Ranges of Plausible Values

Jake Anderson

March 21, 2026

Outline

Where We Left Off

Last time we estimated the food expenditure model and reported:

$$\widehat{\text{food}} = \underset{(43.41)}{83.42} + \underset{(2.09)}{10.21}\,\text{income}$$

Where We Left Off

Last time we estimated the food expenditure model and reported:

$$\widehat{\text{food}} = \underset{(43.41)}{83.42} + \underset{(2.09)}{10.21}\,\text{income}$$

We know $b_2 = 10.21$ with $\text{se}(b_2) = 2.09$.

Where We Left Off

Last time we estimated the food expenditure model and reported:

$$\widehat{\text{food}} = \underset{(43.41)}{83.42} + \underset{(2.09)}{10.21}\,\text{income}$$

We know $b_2 = 10.21$ with $\text{se}(b_2) = 2.09$.

But what does that *tell* us? How confident are we in 10.21?

Where We Left Off

Last time we estimated the food expenditure model and reported:

$$\widehat{\text{food}} = \underset{(43.41)}{83.42} + \underset{(2.09)}{10.21}\,\text{income}$$

We know $b_2 = 10.21$ with $\text{se}(b_2) = 2.09$.

But what does that *tell* us? How confident are we in 10.21?

- Could the true return to income be \$5 per \$100? *(5.21 / 2.09 = 2.49 se's away)*
- Could it be \$15? *(4.79 / 2.09 = 2.29 se's away)*
- Could it be \$0? *(10.21 / 2.09 = 4.89 se's away)*

Where We Left Off

Last time we estimated the food expenditure model and reported:

$$\widehat{\text{food}} = \underset{(43.41)}{83.42} + \underset{(2.09)}{10.21}\,\text{income}$$

We know $b_2 = 10.21$ with $\text{se}(b_2) = 2.09$.

But what does that *tell* us? How confident are we in 10.21?

- Could the true return to income be \$5 per \$100? *(5.21 / 2.09 = 2.49 se's away)*
- Could it be \$15? *(4.79 / 2.09 = 2.29 se's away)*
- Could it be \$0? *(10.21 / 2.09 = 4.89 se's away)*

$\Longrightarrow$ A point estimate and a standard error are the ingredients. A **confidence interval** is the recipe that turns them into a range of plausible values for β_2.

Outline

Step 1: b_2 Is Normally Distributed

Under assumptions SR1–SR6 (including normality of errors), we showed:

$$b_2 \sim N\left(\beta_2,\ \frac{\sigma^2}{\sum(x_i - \bar{x})^2}\right)$$

Step 1: b_2 Is Normally Distributed

Under assumptions SR1–SR6 (including normality of errors), we showed:

$$b_2 \sim N\left(\beta_2, \ \frac{\sigma^2}{\sum(x_i - \bar{x})^2}\right)$$

Standardize by subtracting the mean and dividing by the standard deviation:

$$Z = \frac{b_2 - \beta_2}{\sqrt{\sigma^2 / \sum(x_i - \bar{x})^2}} \sim N(0, 1)$$

Step 1: b_2 Is Normally Distributed

Under assumptions SR1–SR6 (including normality of errors), we showed:

$$b_2 \sim N\left(\beta_2, \; \frac{\sigma^2}{\sum(x_i - \bar{x})^2}\right)$$

Standardize by subtracting the mean and dividing by the standard deviation:

$$Z = \frac{b_2 - \beta_2}{\sqrt{\sigma^2 / \sum(x_i - \bar{x})^2}} \sim N(0, 1)$$

If we knew σ^2, we could write:

$$P(-1.96 \leq Z \leq 1.96) = 0.95$$

Step 1: b_2 Is Normally Distributed

Under assumptions SR1–SR6 (including normality of errors), we showed:

$$b_2 \sim N\left(\beta_2,\ \frac{\sigma^2}{\sum(x_i - \bar{x})^2}\right)$$

Standardize by subtracting the mean and dividing by the standard deviation:

$$Z = \frac{b_2 - \beta_2}{\sqrt{\sigma^2 / \sum(x_i - \bar{x})^2}} \sim N(0, 1)$$

If we knew σ^2, we could write:

$$P(-1.96 \leq Z \leq 1.96) = 0.95$$

and rearrange to get an interval for β_2. But we do not know σ^2.

Step 2: Replace σ^2 with $\hat{\sigma}^2$

We **estimated** σ^2 in Topic 8 using the residuals:

$$\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{N - 2}$$

Step 2: Replace σ^2 with $\hat{\sigma}^2$

We **estimated** σ^2 in Topic 8 using the residuals:

$$\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{N-2}$$

Plugging this into the denominator changes Z into a new statistic:

$$t = \frac{b_2 - \beta_2}{\underbrace{\sqrt{\hat{\sigma}^2 / \sum (x_i - \bar{x})^2}}_{\text{se}(b_2)}} = \frac{b_2 - \beta_2}{\text{se}(b_2)}$$

Step 2: Replace σ^2 with $\hat{\sigma}^2$

We **estimated** σ^2 in Topic 8 using the residuals:

$$\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{N-2}$$

Plugging this into the denominator changes Z into a new statistic:

$$t = \frac{b_2 - \beta_2}{\underbrace{\sqrt{\hat{\sigma}^2 / \sum (x_i - \bar{x})^2}}_{\text{se}(b_2)}} = \frac{b_2 - \beta_2}{\text{se}(b_2)}$$

Replacing the known σ with $\hat{\sigma}$ in the denominator adds estimation uncertainty. The resulting statistic follows a t**-distribution** rather than the standard normal.

Step 3: The t-Distribution

$$t = \frac{b_2 - \beta_2}{\text{se}(b_2)} \sim t_{(N-2)}$$

Step 3: The t-Distribution

$$t = \frac{b_2 - \beta_2}{\text{se}(b_2)} \sim t_{(N-2)}$$

Where does $N - 2$ come from?

- We used N residuals to estimate $\hat{\sigma}^2$
- But we spent 2 degrees of freedom estimating b_1 and b_2
- $\implies N - 2$ degrees of freedom remain

Step 3: The t-Distribution

$$t = \frac{b_2 - \beta_2}{\text{se}(b_2)} \sim t_{(N-2)}$$

Where does $N - 2$ come from?

- We used N residuals to estimate $\hat{\sigma}^2$
- But we spent 2 degrees of freedom estimating b_1 and b_2
- $\implies N - 2$ degrees of freedom remain

The t-distribution looks like the standard normal, but with **heavier tails**:

- Extra tail weight reflects our **uncertainty about** σ
- As $N \to \infty$, $\hat{\sigma}^2 \to \sigma^2$ and $t_{(N-2)} \to N(0, 1)$

Step 3: The t-Distribution

$$t = \frac{b_2 - \beta_2}{\text{se}(b_2)} \sim t_{(N-2)}$$

Where does $N - 2$ come from?

- We used N residuals to estimate $\hat{\sigma}^2$
- But we spent 2 degrees of freedom estimating b_1 and b_2
- $\implies N - 2$ degrees of freedom remain

The t-distribution looks like the standard normal, but with **heavier tails**:

- Extra tail weight reflects our **uncertainty about** σ
- As $N \to \infty$, $\hat{\sigma}^2 \to \sigma^2$ and $t_{(N-2)} \to N(0,1)$

$\implies$ With small samples, we need wider intervals to compensate for estimating σ.

What If You Ignore the Extra Uncertainty?

Suppose you use the normal critical value $z_{0.025} = 1.96$ instead of $t_c = 2.024$:

What If You Ignore the Extra Uncertainty?

Suppose you use the normal critical value $z_{0.025} = 1.96$ instead of $t_c = 2.024$:

	Using $z = 1.96$	**Using** $t_c = 2.024$
Margin	$1.96 \times 2.09 = 4.10$	$2.024 \times 2.09 = 4.23$
Interval	$[6.11,\ 14.31]$	$[5.98,\ 14.44]$

What If You Ignore the Extra Uncertainty?

Suppose you use the normal critical value $z_{0.025} = 1.96$ instead of $t_c = 2.024$:

	Using $z = 1.96$	**Using** $t_c = 2.024$
Margin	$1.96 \times 2.09 = 4.10$	$2.024 \times 2.09 = 4.23$
Interval	$[6.11,\ 14.31]$	$[5.98,\ 14.44]$

The normal-based interval is **too narrow**. It acts as if σ were known exactly, ignoring the additional uncertainty from estimating it.

What If You Ignore the Extra Uncertainty?

Suppose you use the normal critical value $z_{0.025} = 1.96$ instead of $t_c = 2.024$:

	Using $z = 1.96$	**Using** $t_c = 2.024$
Margin	$1.96 \times 2.09 = 4.10$	$2.024 \times 2.09 = 4.23$
Interval	$[6.11,\ 14.31]$	$[5.98,\ 14.44]$

The normal-based interval is **too narrow**. It acts as if σ were known exactly, ignoring the additional uncertainty from estimating it.

$\Longrightarrow$ In small samples, this interval would fail to capture β_2 more than 5% of the time. The t-distribution corrects for this.

Step 4: Rearranging for the Interval

Start with the probability statement:

$$P\left(-t_c \;\le\; \frac{b_2 - \beta_2}{\mathrm{se}(b_2)} \;\le\; t_c\right) = 1 - \alpha$$

where $t_c = t_{(1-\alpha/2,\ N-2)}$ is the critical value from the t-table.

Step 4: Rearranging for the Interval

Start with the probability statement:

$$P\left(-t_c \;\le\; \frac{b_2 - \beta_2}{\mathrm{se}(b_2)} \;\le\; t_c\right) = 1 - \alpha$$

where $t_c = t_{(1-\alpha/2,\; N-2)}$ is the critical value from the t-table.

Multiply through by $\mathrm{se}(b_2)$:

$$P\Big(-\,t_c \cdot \mathrm{se}(b_2) \;\le\; b_2 - \beta_2 \;\le\; t_c \cdot \mathrm{se}(b_2)\Big) = 1 - \alpha$$

Step 4: Rearranging for the Interval

Start with the probability statement:

$$P\left(-t_c \;\leq\; \frac{b_2 - \beta_2}{\mathrm{se}(b_2)} \;\leq\; t_c\right) = 1 - \alpha$$

where $t_c = t_{(1-\alpha/2,\, N-2)}$ is the critical value from the t-table.

Multiply through by $\mathrm{se}(b_2)$:

$$P\Big(-t_c \cdot \mathrm{se}(b_2) \;\leq\; b_2 - \beta_2 \;\leq\; t_c \cdot \mathrm{se}(b_2)\Big) = 1 - \alpha$$

Rearrange (subtract b_2, multiply by -1, flip inequalities):

$$\boxed{P\Big[b_2 - t_c \cdot \mathrm{se}(b_2) \;\leq\; \beta_2 \;\leq\; b_2 + t_c \cdot \mathrm{se}(b_2)\Big] = 1 - \alpha}$$

Step 4: Rearranging for the Interval

Start with the probability statement:

$$P\left(-t_c \;\leq\; \frac{b_2 - \beta_2}{\text{se}(b_2)} \;\leq\; t_c\right) = 1 - \alpha$$

where $t_c = t_{(1-\alpha/2,\ N-2)}$ is the critical value from the t-table.

Multiply through by $\text{se}(b_2)$:

$$P\Big(-t_c \cdot \text{se}(b_2) \;\leq\; b_2 - \beta_2 \;\leq\; t_c \cdot \text{se}(b_2)\Big) = 1 - \alpha$$

Rearrange (subtract b_2, multiply by -1, flip inequalities):

$$\boxed{P\Big[b_2 - t_c \cdot \text{se}(b_2) \;\leq\; \beta_2 \;\leq\; b_2 + t_c \cdot \text{se}(b_2)\Big] = 1 - \alpha}$$

This is the $100(1-\alpha)\%$ **confidence interval** for β_2.

The Confidence Interval Formula

$$\text{CI for } \beta_k : \qquad b_k \pm t_c \cdot \text{se}(b_k)$$

The Confidence Interval Formula

$$\text{CI for } \beta_k : \qquad b_k \pm t_c \cdot \text{se}(b_k)$$

where:

- b_k = OLS point estimate (center of the interval)
- $\text{se}(b_k)$ = standard error (measures precision)
- $t_c = t_{(1-\alpha/2,\ N-2)}$ = critical value from the t-distribution
- α = significance level (e.g., $\alpha = 0.05$ for a 95% CI)

The Confidence Interval Formula

$$\text{CI for } \beta_k : \qquad b_k \pm t_c \cdot \text{se}(b_k)$$

where:

- b_k = OLS point estimate (center of the interval)
- $\text{se}(b_k)$ = standard error (measures precision)
- $t_c = t_{(1-\alpha/2,\ N-2)}$ = critical value from the t-distribution
- α = significance level (e.g., $\alpha = 0.05$ for a 95% CI)

Three moving parts, and you already know two of them (b_k and $\text{se}(b_k)$). The only new ingredient is t_c, which you look up in a table or compute in software.

Outline

The Repeated Sampling Interpretation

Before we compute our first interval, we need to understand what it will tell us.

The Repeated Sampling Interpretation

Before we compute our first interval, we need to understand what it will tell us.

“95% confidence” does **not** mean: “there is a 95% probability that β_2 is in the interval.”

The Repeated Sampling Interpretation

Before we compute our first interval, we need to understand what it will tell us.

"95% confidence" does **not** mean: "there is a 95% probability that β_2 is in the interval."

Why not? Because β_2 is a **fixed, unknown constant**. It is either inside the interval or it is not. There is nothing random about β_2.

The Repeated Sampling Interpretation

Before we compute our first interval, we need to understand what it will tell us.

“95% confidence” does **not** mean: “there is a 95% probability that β_2 is in the interval.”

Why not? Because β_2 is a **fixed, unknown constant**. It is either inside the interval or it is not. There is nothing random about β_2.

What *is* random? **The interval.**

The Repeated Sampling Interpretation

Before we compute our first interval, we need to understand what it will tell us.

“95% confidence” does **not** mean: “there is a 95% probability that β_2 is in the interval.”

Why not? Because β_2 is a **fixed, unknown constant**. It is either inside the interval or it is not. There is nothing random about β_2.

What *is* random? **The interval.**

- b_2 changes from sample to sample $\implies$ the center moves
- $\text{se}(b_2)$ changes from sample to sample $\implies$ the width changes
- Each sample produces a different interval

The Repeated Sampling Interpretation

Before we compute our first interval, we need to understand what it will tell us.

"95% confidence" does **not** mean: "there is a 95% probability that β_2 is in the interval."

Why not? Because β_2 is a **fixed, unknown constant**. It is either inside the interval or it is not. There is nothing random about β_2.

What *is* random? **The interval.**

- b_2 changes from sample to sample $\implies$ the center moves
- $\text{se}(b_2)$ changes from sample to sample $\implies$ the width changes
- Each sample produces a different interval

The correct interpretation: if we drew many samples and built a 95% CI from each one, **approximately 95% of those intervals would contain the true β_2.**

The Repeated Sampling Interpretation

Before we compute our first interval, we need to understand what it will tell us.

"95% confidence" does **not** mean: "there is a 95% probability that β_2 is in the interval."

Why not? Because β_2 is a **fixed, unknown constant**. It is either inside the interval or it is not. There is nothing random about β_2.

What *is* random? **The interval.**

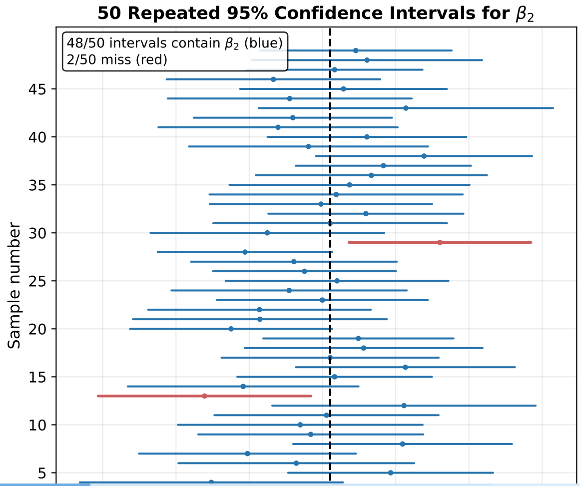
- b_2 changes from sample to sample $\implies$ the center moves
- $\text{se}(b_2)$ changes from sample to sample $\implies$ the width changes
- Each sample produces a different interval

The correct interpretation: if we drew many samples and built a 95% CI from each one, **approximately 95% of those intervals would contain the true** β_2.

$\implies$ Our confidence is in the **procedure**, not in any single interval.

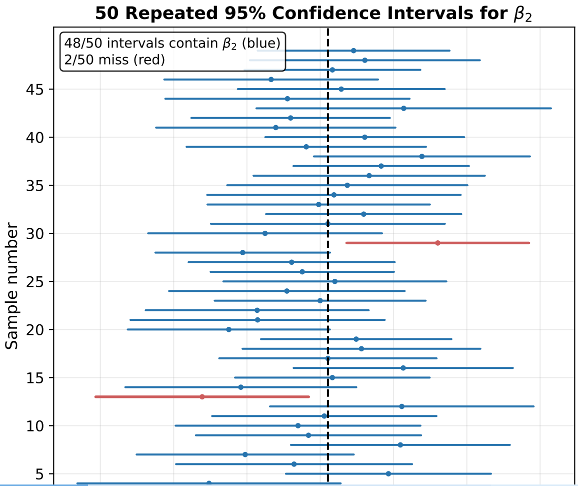
Visualizing Repeated Sampling

Simulation: draw 50 random samples from the food expenditure DGP, compute a 95% CI for β_2 from each.



Visualizing Repeated Sampling

Simulation: draw 50 random samples from the food expenditure DGP, compute a 95% CI for β_2 from each.



The Interval Is Random, Not β_2

Common misinterpretation

"There is a 95% probability that β_2 lies in the interval."

The Interval Is Random, Not β_2

Common misinterpretation

"There is a 95% probability that β_2 lies in the interval."

Correct interpretation

"The procedure we used captures the true β_2 in 95% of repeated samples. We do not know whether this particular interval is one of the 95% or the 5%."

The Interval Is Random, Not β_2

Common misinterpretation

"There is a 95% probability that β_2 lies in the interval."

Correct interpretation

"The procedure we used captures the true β_2 in 95% of repeated samples. We do not know whether this particular interval is one of the 95% or the 5%."

Analogy: a factory produces light bulbs, and 95% work. You buy one. You do not say "there is a 95% probability this bulb works." It either works or it does not. The 95% describes the **factory's track record**, not your particular bulb.

Outline

95% CI for the Return to Income

Data: $N = 40$ households. Degrees of freedom: $N - 2 = 38$.

95% CI for the Return to Income

Data: $N = 40$ households. Degrees of freedom: $N - 2 = 38$.

Estimates: $b_2 = 10.21$, $\text{se}(b_2) = 2.09$.

95% CI for the Return to Income

Data: $N = 40$ households. Degrees of freedom: $N - 2 = 38$.

Estimates: $b_2 = 10.21$, $\text{se}(b_2) = 2.09$.

Critical value: For a 95% CI, $\alpha = 0.05$:

$$t_c = t_{(0.975,\ 38)} = 2.024$$

95% CI for the Return to Income

Data: $N = 40$ households. Degrees of freedom: $N - 2 = 38$.

Estimates: $b_2 = 10.21$, $\text{se}(b_2) = 2.09$.

Critical value: For a 95% CI, $\alpha = 0.05$:

$$t_c = t_{(0.975,\ 38)} = 2.024$$

Interval:

$$\begin{aligned} b_2 \pm t_c \cdot \text{se}(b_2) &= 10.21 \pm 2.024 \times 2.09 \\ &= 10.21 \pm 4.23 \\ &= [5.98,\ 14.44] \end{aligned}$$

95% CI for the Return to Income

Data: $N = 40$ households. Degrees of freedom: $N - 2 = 38$.

Estimates: $b_2 = 10.21$, $\text{se}(b_2) = 2.09$.

Critical value: For a 95% CI, $\alpha = 0.05$:

$$t_c = t_{(0.975,\ 38)} = 2.024$$

Interval:

$$\begin{aligned} b_2 \pm t_c \cdot \text{se}(b_2) &= 10.21 \pm 2.024 \times 2.09 \\ &= 10.21 \pm 4.23 \\ &= [5.98,\ 14.44] \end{aligned}$$

We estimate with 95% confidence that an additional $100 of weekly income increases food expenditure by between $5.98 and $14.44.

Changing the Confidence Level

Same sample, same data. What happens if we change α?

Changing the Confidence Level

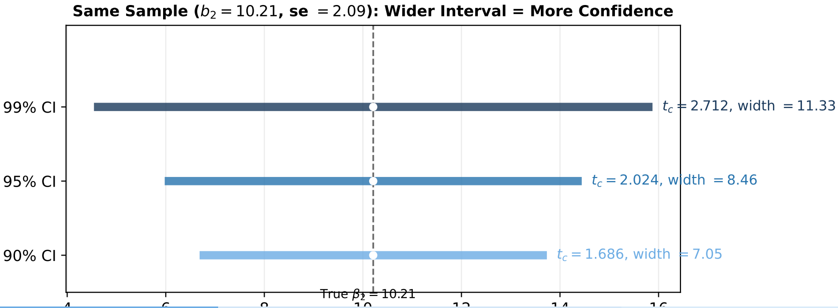
Same sample, same data. What happens if we change α?

Level	t_c	Interval	Width
90% ($\alpha = 0.10$)	$t_{(0.95,38)} = 1.686$	$[6.69,\ 13.73]$	7.04
95% ($\alpha = 0.05$)	$t_{(0.975,38)} = 2.024$	$[5.98,\ 14.44]$	8.46
99% ($\alpha = 0.01$)	$t_{(0.995,38)} = 2.712$	$[4.54,\ 15.88]$	11.34

Changing the Confidence Level

Same sample, same data. What happens if we change α?

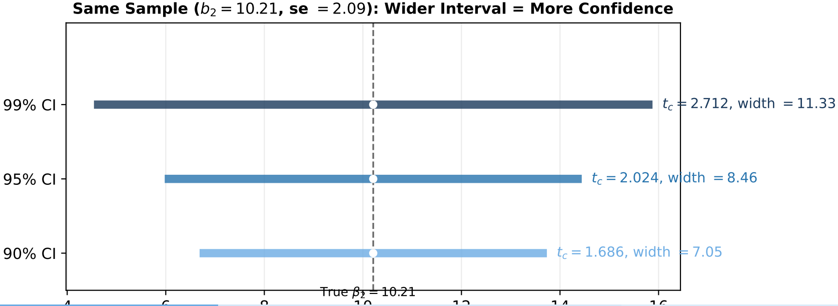
Level	t_c	Interval	Width
90% ($\alpha = 0.10$)	$t_{(0.95,38)} = 1.686$	$[6.69,\ 13.73]$	7.04
95% ($\alpha = 0.05$)	$t_{(0.975,38)} = 2.024$	$[5.98,\ 14.44]$	8.46
99% ($\alpha = 0.01$)	$t_{(0.995,38)} = 2.712$	$[4.54,\ 15.88]$	11.34



Changing the Confidence Level

Same sample, same data. What happens if we change α?

Level	t_c	Interval	Width
90% ($\alpha = 0.10$)	$t_{(0.95,38)} = 1.686$	$[6.69,\ 13.73]$	7.04
95% ($\alpha = 0.05$)	$t_{(0.975,38)} = 2.024$	$[5.98,\ 14.44]$	8.46
99% ($\alpha = 0.01$)	$t_{(0.995,38)} = 2.712$	$[4.54,\ 15.88]$	11.34



Outline

Why Linear Combinations?

We now have a CI for a single parameter. But what if we want a CI for a **prediction** that depends on both β_1 and β_2?

Why Linear Combinations?

We now have a **CI** for a single parameter. But what if we want a **CI** for a **prediction** that depends on both β_1 and β_2?

Example: What is the expected food expenditure for a household earning $2,000/week ($x_0 = 20$)?

$$\mathrm{E}(\text{food} \mid x = 20) = \beta_1 + 20\,\beta_2$$

Why Linear Combinations?

We now have a CI for a single parameter. But what if we want a CI for a **prediction** that depends on both β_1 and β_2?

Example: What is the expected food expenditure for a household earning $2,000/week ($x_0 = 20$)?

$$\mathrm{E}(\text{food} \mid x = 20) = \beta_1 + 20\,\beta_2$$

This is a **linear combination**: $\lambda = c_1\beta_1 + c_2\beta_2$ with $c_1 = 1$, $c_2 = 20$.

Why Linear Combinations?

We now have a CI for a single parameter. But what if we want a CI for a **prediction** that depends on both β_1 and β_2?

Example: What is the expected food expenditure for a household earning $2,000/week ($x_0 = 20$)?

$$E(\text{food} \mid x = 20) = \beta_1 + 20\,\beta_2$$

This is a **linear combination:** $\lambda = c_1\beta_1 + c_2\beta_2$ with $c_1 = 1$, $c_2 = 20$.

Point estimate:

$$\hat{\lambda} = b_1 + 20 \cdot b_2 = 83.42 + 20 \times 10.21 = 287.61$$

Why Linear Combinations?

We now have a CI for a single parameter. But what if we want a CI for a **prediction** that depends on both β_1 and β_2?

Example: What is the expected food expenditure for a household earning $2,000/week ($x_0 = 20$)?

$$E(\text{food} \mid x = 20) = \beta_1 + 20\,\beta_2$$

This is a **linear combination**: $\lambda = c_1\beta_1 + c_2\beta_2$ with $c_1 = 1$, $c_2 = 20$.

Point estimate:

$$\hat{\lambda} = b_1 + 20 \cdot b_2 = 83.42 + 20 \times 10.21 = 287.61$$

We predict $287.61 per week. But how precise is this?

Variance of a Linear Combination

For $\hat{\lambda} = c_1 b_1 + c_2 b_2$, the variance of a sum of correlated random variables is:

$$\mathrm{Var}(\hat{\lambda}) = c_1^2 \ \mathrm{Var}(b_1) + c_2^2 \ \mathrm{Var}(b_2) + 2\, c_1 c_2 \ \mathrm{Cov}(b_1, b_2)$$

Variance of a Linear Combination

For $\hat{\lambda} = c_1 b_1 + c_2 b_2$, the variance of a sum of correlated random variables is:

$$\mathrm{Var}(\hat{\lambda}) = c_1^2 \ \mathrm{Var}(b_1) + c_2^2 \ \mathrm{Var}(b_2) + 2\, c_1 c_2 \ \mathrm{Cov}(b_1, b_2)$$

Do not forget the **covariance term**. The estimators b_1 and b_2 are correlated (both come from the same data), so you cannot just add the individual variances.

Variance of a Linear Combination

For $\hat{\lambda} = c_1 b_1 + c_2 b_2$, the variance of a sum of correlated random variables is:

$$\text{Var}(\hat{\lambda}) = c_1^2 \ \text{Var}(b_1) + c_2^2 \ \text{Var}(b_2) + 2\, c_1 c_2 \ \text{Cov}(b_1, b_2)$$

Do not forget the **covariance term**. The estimators b_1 and b_2 are correlated (both come from the same data), so you cannot just add the individual variances.

Where do the numbers come from? Software reports the estimated covariance matrix:

$$\widehat{\text{Cov}}(b_1, b_2) = \begin{pmatrix} \widehat{\text{Var}}(b_1) & \widehat{\text{Cov}}(b_1, b_2) \\ \widehat{\text{Cov}}(b_1, b_2) & \widehat{\text{Var}}(b_2) \end{pmatrix} = \begin{pmatrix} 1884.44 & -85.90 \\ -85.90 & 4.38 \end{pmatrix}$$

Variance of a Linear Combination

For $\hat{\lambda} = c_1 b_1 + c_2 b_2$, the variance of a sum of correlated random variables is:

$$\text{Var}(\hat{\lambda}) = c_1^2 \ \text{Var}(b_1) + c_2^2 \ \text{Var}(b_2) + 2\, c_1 c_2 \ \text{Cov}(b_1, b_2)$$

Do not forget the **covariance term**. The estimators b_1 and b_2 are correlated (both come from the same data), so you cannot just add the individual variances.

Where do the numbers come from? Software reports the estimated covariance matrix:

$$\widehat{\text{Cov}}(b_1, b_2) = \begin{pmatrix} \widehat{\text{Var}}(b_1) & \widehat{\text{Cov}}(b_1, b_2) \\ \widehat{\text{Cov}}(b_1, b_2) & \widehat{\text{Var}}(b_2) \end{pmatrix} = \begin{pmatrix} 1884.44 & -85.90 \\ -85.90 & 4.38 \end{pmatrix}$$

For $c_1 = 1$, $c_2 = 20$:

$$\begin{aligned} \widehat{\text{Var}}(\hat{\lambda}) &= (1)^2 \times 1884.44 + (20)^2 \times 4.38 + 2(1)(20)(-85.90) \\ &= 1884.44 + 1752.72 - 3436.13 \\ &= 201.03 \end{aligned}$$

Variance of a Linear Combination

For $\hat{\lambda} = c_1 b_1 + c_2 b_2$, the variance of a sum of correlated random variables is:

$$\text{Var}(\hat{\lambda}) = c_1^2 \ \text{Var}(b_1) + c_2^2 \ \text{Var}(b_2) + 2\, c_1 c_2 \ \text{Cov}(b_1, b_2)$$

Do not forget the **covariance term**. The estimators b_1 and b_2 are correlated (both come from the same data), so you cannot just add the individual variances.

Where do the numbers come from? Software reports the estimated covariance matrix:

$$\widehat{\text{Cov}}(b_1, b_2) = \begin{pmatrix} \widehat{\text{Var}}(b_1) & \widehat{\text{Cov}}(b_1, b_2) \\ \widehat{\text{Cov}}(b_1, b_2) & \widehat{\text{Var}}(b_2) \end{pmatrix} = \begin{pmatrix} 1884.44 & -85.90 \\ -85.90 & 4.38 \end{pmatrix}$$

For $c_1 = 1$, $c_2 = 20$:

$$\begin{aligned} \widehat{\text{Var}}(\hat{\lambda}) &= (1)^2 \times 1884.44 + (20)^2 \times 4.38 + 2(1)(20)(-85.90) \\ &= 1884.44 + 1752.72 - 3436.13 \\ &= 201.03 \end{aligned}$$

$\text{se}(\hat{\lambda}) = \sqrt{201.03} = 14.18$

CI for Expected Food Expenditure

The CI formula itself is unchanged: $\hat{\lambda} \pm t_c \cdot \text{se}(\hat{\lambda})$. The only new ingredient is the variance formula for the linear combination.

CI for Expected Food Expenditure

The CI formula itself is unchanged: $\hat{\lambda} \pm t_c \cdot \text{se}(\hat{\lambda})$. The only new ingredient is the variance formula for the linear combination.

Same $t_c = 2.024$ (95% CI, $df = 38$):

$$\begin{aligned} 287.61 \pm 2.024 \times 14.18 &= 287.61 \pm 28.69 \\ &= [258.91,\ 316.31] \end{aligned}$$

CI for Expected Food Expenditure

The CI formula itself is unchanged: $\hat{\lambda} \pm t_c \cdot \text{se}(\hat{\lambda})$. The only new ingredient is the variance formula for the linear combination.

Same $t_c = 2.024$ (95% CI, $df = 38$):

$$\begin{aligned} 287.61 \pm 2.024 \times 14.18 &= 287.61 \pm 28.69 \\ &= [258.91,\ 316.31] \end{aligned}$$

We estimate with 95% confidence that a household earning \$2,000/week spends between \$258.91 and \$316.31 on food.

CI for Expected Food Expenditure

The CI formula itself is unchanged: $\hat{\lambda} \pm t_c \cdot \text{se}(\hat{\lambda})$. The only new ingredient is the variance formula for the linear combination.

Same $t_c = 2.024$ (95% CI, $df = 38$):

$$
\begin{aligned}
287.61 \pm 2.024 \times 14.18 &= 287.61 \pm 28.69 \\
&= [258.91,\ 316.31]
\end{aligned}
$$

We estimate with 95% confidence that a household earning \$2,000/week spends between \$258.91 and \$316.31 on food.

This interval reflects our uncertainty about **both** β_1 and β_2 simultaneously, plus their covariance.

Outline

What We Covered

A point estimate without a confidence interval is incomplete. The formula is one line, but the interpretation requires understanding that confidence is about the procedure, not the interval.

What We Covered

A point estimate without a confidence interval is incomplete. The formula is one line, but the interpretation requires understanding that confidence is about the procedure, not the interval.

The derivation: b_2 is normal (SR6) $\implies$ standardize $\implies$ replace σ with $\hat{\sigma}$ $\implies$ t-distribution with $N-2$ df.

What We Covered

A point estimate without a confidence interval is incomplete. The formula is one line, but the interpretation requires understanding that confidence is about the procedure, not the interval.

The derivation: b_2 is normal (SR6) $\implies$ standardize $\implies$ replace σ with $\hat{\sigma}$ $\implies$ t-distribution with $N-2$ df.

The formula: $b_k \pm t_c \cdot \text{se}(b_k)$, where $t_c = t_{(1-\alpha/2,\ N-2)}$.

What We Covered

A point estimate without a confidence interval is incomplete. The formula is one line, but the interpretation requires understanding that confidence is about the procedure, not the interval.

The derivation: b_2 is normal (SR6) $\implies$ standardize $\implies$ replace σ with $\hat{\sigma}$ $\implies$ t-distribution with $N - 2$ df.

The formula: $b_k \pm t_c \cdot \text{se}(b_k)$, where $t_c = t_{(1-\alpha/2,\ N-2)}$.

The interpretation: confidence is in the procedure, not in any single interval. Approximately $100(1 - \alpha)\%$ of intervals constructed this way will contain the true parameter.

What We Covered

A point estimate without a confidence interval is incomplete. The formula is one line, but the interpretation requires understanding that confidence is about the procedure, not the interval.

The derivation: b_2 is normal (SR6) $\implies$ standardize $\implies$ replace σ with $\hat{\sigma}$ $\implies$ t-distribution with $N-2$ df.

The formula: $b_k \pm t_c \cdot \text{se}(b_k)$, where $t_c = t_{(1-\alpha/2,\ N-2)}$.

The interpretation: confidence is in the procedure, not in any single interval. Approximately $100(1-\alpha)\%$ of intervals constructed this way will contain the true parameter.

The tradeoff: more confidence $\implies$ wider interval. No free lunch.

What We Covered

A point estimate without a confidence interval is incomplete. The formula is one line, but the interpretation requires understanding that confidence is about the procedure, not the interval.

The derivation: b_2 is normal (SR6) $\implies$ standardize $\implies$ replace σ with $\hat{\sigma}$ $\implies$ t-distribution with $N - 2$ df.

The formula: $b_k \pm t_c \cdot \text{se}(b_k)$, where $t_c = t_{(1-\alpha/2,\ N-2)}$.

The interpretation: confidence is in the procedure, not in any single interval. Approximately $100(1-\alpha)\%$ of intervals constructed this way will contain the true parameter.

The tradeoff: more confidence $\implies$ wider interval. No free lunch.

Linear combinations: for $\lambda = c_1\beta_1 + c_2\beta_2$, use the full variance formula (do not forget the covariance term), then apply the same CI recipe.

What We Covered

A point estimate without a confidence interval is incomplete. The formula is one line, but the interpretation requires understanding that confidence is about the procedure, not the interval.

The derivation: b_2 is normal (SR6) $\implies$ standardize $\implies$ replace σ with $\hat{\sigma}$ $\implies$ t-distribution with $N-2$ df.

The formula: $b_k \pm t_c \cdot \text{se}(b_k)$, where $t_c = t_{(1-\alpha/2,\ N-2)}$.

The interpretation: confidence is in the procedure, not in any single interval. Approximately $100(1-\alpha)\%$ of intervals constructed this way will contain the true parameter.

The tradeoff: more confidence $\implies$ wider interval. No free lunch.

Linear combinations: for $\lambda = c_1\beta_1 + c_2\beta_2$, use the full variance formula (do not forget the covariance term), then apply the same CI recipe.

Next: now that we can build intervals, what if we want a yes-or-no answer? $\implies$ hypothesis testing (Topic 10).

Thank you!

jakeanderson@g.ucla.edu