Bootstrap and Bivariate Regression

Kirk Bansak

February 14, 2023

Problem Sets

- Problem Set 2 is now due.
- Problem Set 3 is due next Thursday, February 23, by 12:29PM PT.
- Reminder: You should not use any `R` packages or functionalities beyond what is already loaded when you open up `RStudio`, unless a package is explicitly mentioned.
- The packages automatically loaded by default (and hence allowed) are: `base`, `datasets`, `grDevices`, `graphics`, `methods`, `stats`, `utils`
- You may always use `ggplot2`.

First, revisiting the bootstrap...

Uncertainty

When conducting statistical analyses, we will often want to not only produce estimates/predictions, but also quantify the uncertainty surrounding those estimates/predictions.

Two possible methods:

1. Classic statistical inference
2. **Bootstrap**

Welcome to the Bootstrap World

Real World

Unknown probability distribution

Observed random sample

$$P \longrightarrow \mathrm{Z} = (\mathrm{Z}_1, \ldots, \mathrm{Z}_n)$$

$$\downarrow$$

$$\hat{\theta} = s(\mathrm{Z})$$

Statistic of interest

Welcome to the Bootstrap World

Real World

Unknown probability distribution

Observed random sample

$$P \longrightarrow \mathrm{Z} = (\mathrm{Z}_1, \ldots, \mathrm{Z}_n)$$

$$\downarrow$$

$$\hat{\theta} = s(\mathrm{Z})$$

Statistic of interest

- Given no other information, the observed sample Z contains all the available information about the underlying population distribution P

Welcome to the Bootstrap World

Real World		**Bootstrap World**
Unknown probability distribution — Observed random sample		Empirical distribution — Bootstrap sample
$P \longrightarrow \mathrm{Z} = (\mathrm{Z}_1, \ldots, \mathrm{Z}_n)$	$\Longrightarrow$	$\hat{P} \longrightarrow \mathrm{Z}^* = (\mathrm{Z}_1^*, \ldots, \mathrm{Z}_n^*)$
$\downarrow$		$\downarrow$
$\hat{\theta} = s(\mathrm{Z})$		$\hat{\theta}^* = s(\mathrm{Z}^*)$
Statistic of interest		Bootstrap replication

- Given no other information, the observed sample Z contains all the available information about the underlying population distribution P
- Thus, resampling from Z is the best guide to what can be expected from resampling from P

Welcome to the Bootstrap World

Real World			**Bootstrap World**	
Unknown probability distribution	Observed random sample		Empirical distribution	Bootstrap sample
$P \longrightarrow \mathrm{Z} = (\mathrm{Z}_1, \ldots, \mathrm{Z}_n)$		$\Longrightarrow$	$\hat{P} \longrightarrow \mathrm{Z}^* = (\mathrm{Z}_1^*, \ldots, \mathrm{Z}_n^*)$	
$\downarrow$			$\downarrow$	
$\hat{\theta} = s(\mathrm{Z})$			$\hat{\theta}^* = s(\mathrm{Z}^*)$	
Statistic of interest			Bootstrap replication	

- Given no other information, the observed sample Z contains all the available information about the underlying population distribution P
- Thus, resampling from Z is the best guide to what can be expected from resampling from P
- Just like we drew Z from P, let's draw a resample Z^* from Z

Welcome to the Bootstrap World

Real World

Unknown probability distribution — Observed random sample

$$P \longrightarrow \mathrm{Z} = (\mathrm{Z}_1, \ldots, \mathrm{Z}_n)$$

$$\downarrow$$

$$\hat{\theta} = s(\mathrm{Z})$$

Statistic of interest

$$\Longrightarrow$$

Bootstrap World

Empirical distribution — Bootstrap sample

$$\hat{P} \longrightarrow \mathrm{Z}^* = (\mathrm{Z}_1^*, \ldots, \mathrm{Z}_n^*)$$

$$\downarrow$$

$$\hat{\theta}^* = s(\mathrm{Z}^*)$$

Bootstrap replication

- Given no other information, the observed sample Z contains all the available information about the underlying population distribution P
- Thus, resampling from Z is the best guide to what can be expected from resampling from P
- Just like we drew Z from P, let's draw a resample Z^* from Z
- If n is sufficiently large, the observed sample Z should be a good approximation of P (i.e. treat Z as $\hat{P}$)

(Nonparametric) Bootstrap

1. Decide on a summary statistic s of interest.
   - e.g. if Z contains only a single variable, perhaps the mean of that variable

(Nonparametric) Bootstrap

1. Decide on a summary statistic s of interest.
   - e.g. if Z contains only a single variable, perhaps the mean of that variable
2. Draw B resamples of n observations from Z **with replacement**.
   - Note: If Z contains multiple variables, resampled observations should be kept intact (i.e. full rows should be drawn).

   Call the bth resample Z_b^*.

(Nonparametric) Bootstrap

1. Decide on a summary statistic s of interest.
   - e.g. if Z contains only a single variable, perhaps the mean of that variable
2. Draw B resamples of n observations from Z **with replacement**.
   - Note: If Z contains multiple variables, resampled observations should be kept intact (i.e. full rows should be drawn).

   Call the bth resample Z_b^*.
3. For each Z_b^*, compute $s(Z_b^*)$ and store it, creating the set $\{s(Z_1^*), ..., s(Z_B^*)\}$

(Nonparametric) Bootstrap

1. Decide on a summary statistic s of interest.
   - e.g. if Z contains only a single variable, perhaps the mean of that variable
2. Draw B resamples of n observations from Z **with replacement**.
   - Note: If Z contains multiple variables, resampled observations should be kept intact (i.e. full rows should be drawn).

   Call the bth resample Z_b^*.
3. For each Z_b^*, compute $s(Z_b^*)$ and store it, creating the set $\{s(Z_1^*), ..., s(Z_B^*)\}$
4. You can assess the distribution of $\{s(Z_1^*), ..., s(Z_B^*)\}$; for instance, its standard deviation represents the standard error of your statistic.

(Nonparametric) Bootstrap

1. Decide on a summary statistic s of interest.
   - e.g. if Z contains only a single variable, perhaps the mean of that variable
2. Draw B resamples of n observations from Z **with replacement**.
   - Note: If Z contains multiple variables, resampled observations should be kept intact (i.e. full rows should be drawn).

   Call the bth resample Z_b^*.
3. For each Z_b^*, compute $s(Z_b^*)$ and store it, creating the set $\{s(Z_1^*), ..., s(Z_B^*)\}$
4. You can assess the distribution of $\{s(Z_1^*), ..., s(Z_B^*)\}$; for instance, its standard deviation represents the standard error of your statistic.
5. Or to compute 95% CI, use 2.5/97.5 percentiles of $\{s(Z_1^*), ..., s(Z_B^*)\}$ as the lower/upper bounds (bootstrap percentile CI)

Does this look familiar?

Does this look familiar?

You just programmed a bootstrap on the problem set!

To *R*!

Quick Review

Supervised Learning

We assume some relationship between Y and $X = (X_1, X_2, ..., X_p)$, such that:

$$Y = f(X) + \epsilon$$

- f is some fixed but unknown function of $X_1, X_2, ..., X_p$
- ϵ is a random irreducible error term

Supervised Learning

Using observed data (X and Y) to estimate f with $\hat{f}$

Ultimate Goal: Build a $\hat{f}$ that is close to f

Using a Linear Model for f

For each election i, where i is used to index different observations,

Let:

$Vote_i$ = Incumbent Vote Share in election i.
$Approval_i$ = Incumbent Net Approval in election i.

We want to predict incumbent vote share using approval as our input.

Using a Linear Model for f

For each election i, where i is used to index different observations,

Let:

$Vote_i =$ Incumbent Vote Share in election i.
$Approval_i =$ Incumbent Net Approval in election i.

We want to predict incumbent vote share using approval as our input.

Thus, we need a function f that relates $Approval_i$ to $Vote_i$, and one fundamental approach is to employ a linear function:

Using a Linear Model for f

For each election i, where i is used to index different observations,

Let:

$Vote_i$ = Incumbent Vote Share in election i.
$Approval_i$ = Incumbent Net Approval in election i.

We want to predict incumbent vote share using approval as our input.

Thus, we need a function f that relates $Approval_i$ to $Vote_i$, and one fundamental approach is to employ a linear function:

$$\begin{aligned} \text{Vote}_i &= f(\text{Approval}_i) + \epsilon_i \\ \text{Vote}_i &= \beta_0 + \beta_1 \text{Approval}_i + \epsilon_i \end{aligned}$$

Bivariate Regression: Function Perspective

Employing linear function f to relate $Approval_i$ to $Vote_i$:

$$\text{Vote}_i \quad = \quad \beta_0 + \beta_1 \text{Approval}_i + \epsilon_i$$

β_0 and β_1 are two unknown quantities known as the model coefficients or parameters.

Bivariate Regression: Function Perspective

Employing linear function f to relate $Approval_i$ to $Vote_i$:

$$\text{Vote}_i \quad = \quad \beta_0 + \beta_1 \text{Approval}_i + \epsilon_i$$

β_0 and β_1 are two unknown quantities known as the model coefficients or parameters.

Now we need to use our training data to estimate β_0 and β_1.

Bivariate Regression: Function Perspective

Employing linear function f to relate $Approval_i$ to $Vote_i$:

$$\text{Vote}_i = \beta_0 + \beta_1 \text{Approval}_i + \epsilon_i$$

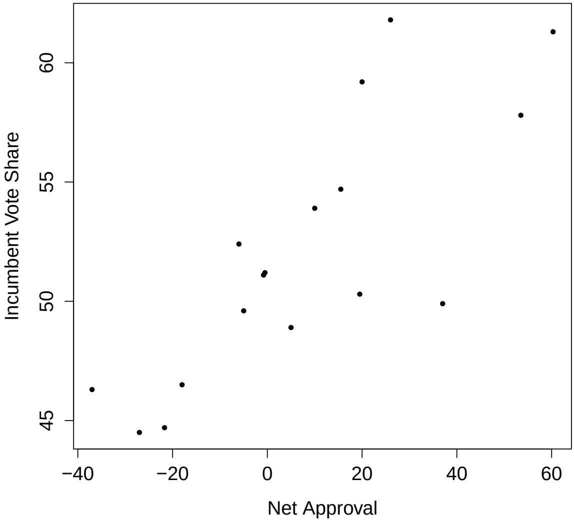
β_0 and β_1 are two unknown quantities known as the model coefficients or parameters.

Now we need to use our training data to estimate β_0 and β_1.

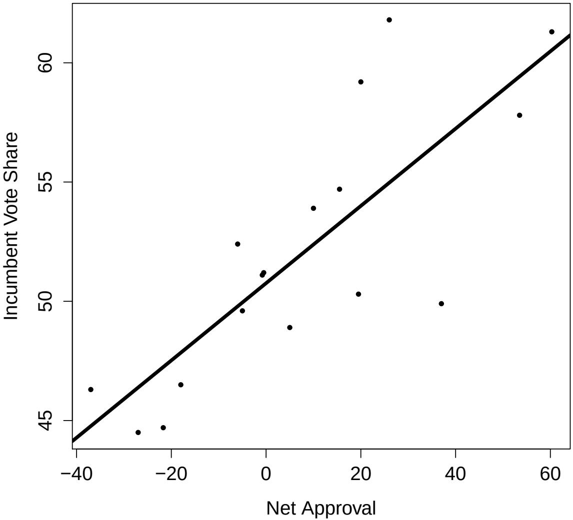
By using our estimates, $\hat{\beta}_0$ and $\hat{\beta}_1$, we can then make predictions for $Vote_i$:

$$\widehat{Vote}_i = \hat{\beta}_0 + \hat{\beta}_1 Approval_i$$

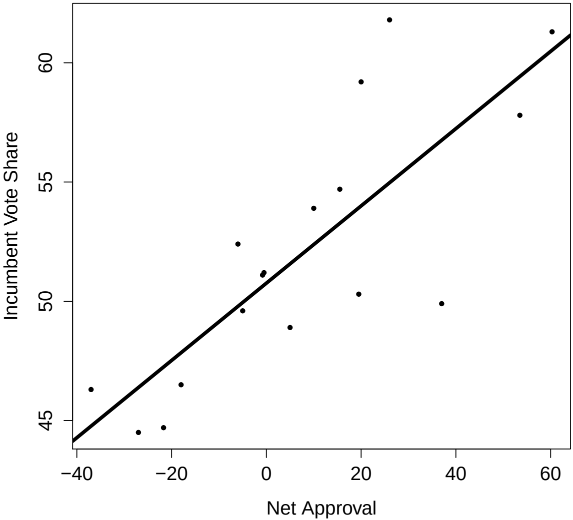
Bivariate Regression: Geometric Perspective



Bivariate Regression: Geometric Perspective



Test Your Knowledge



1. What corresponds to $\hat{\beta}_0$?
2. What corresponds to $\hat{\beta}_1$?
3. What corresponds to $\hat{\epsilon}_i$?
4. What's the difference between Y and $\hat{Y}$?
5. Bonus: where are we interpolating and where are we extrapolating?

Fitting a Bivariate Regression

Regression Model:

$$\text{Vote}_i = \beta_0 + \beta_1 \text{Approval}_i + \epsilon_i$$

Fitting a Bivariate Regression

Regression Model:

$$\text{Vote}_i = \beta_0 + \beta_1 \text{Approval}_i + \epsilon_i$$

Fitted Regression:

$$\text{Vote}_i = \underbrace{\hat{\beta}_0 + \hat{\beta}_1 \text{Approval}_i}_{\widehat{\text{Vote}}_i} + \hat{\epsilon}_i$$

Fitting a Bivariate Regression

Regression Model:

$$\text{Vote}_i = \beta_0 + \beta_1 \text{Approval}_i + \epsilon_i$$

Fitted Regression:

$$\text{Vote}_i = \underbrace{\hat{\beta}_0 + \hat{\beta}_1 \text{Approval}_i}_{\widehat{\text{Vote}_i}} + \hat{\epsilon}_i$$

Residuals:

$$\hat{\epsilon}_i = \text{Vote}_i - \widehat{\text{Vote}_i}$$

Fitting a Bivariate Regression

Regression Model:

$$\text{Vote}_i = \beta_0 + \beta_1 \text{Approval}_i + \epsilon_i$$

Fitted Regression:

$$\text{Vote}_i = \underbrace{\hat{\beta}_0 + \hat{\beta}_1 \text{Approval}_i}_{\widehat{\text{Vote}_i}} + \hat{\epsilon}_i$$

Residuals:

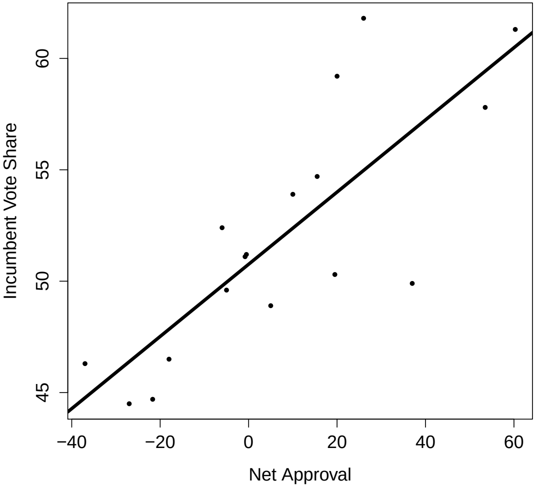
$$\hat{\epsilon}_i = \text{Vote}_i - \widehat{\text{Vote}_i}$$

Goal:

Choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to 'make residuals small' (i.e. get $\widehat{\text{Vote}_i}$ 'close to' Vote_i)

Bivariate Regression: Geometric Perspective

Choose $\hat{\beta}_0$ and $\hat{\beta}_1$ such that the regression line is 'close to' the data points.



How Should We Choose the Parameter Values?

Preliminary Ideas:

Choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to minimize the following?

How Should We Choose the Parameter Values?

Preliminary Ideas:

Choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to minimize the following?

$$\sum_{i=1}^{N} |\hat{\epsilon}_i| = \sum_{i=1}^{N} \left| \text{Vote}_i - \widehat{\text{Vote}_i} \right|$$

How Should We Choose the Parameter Values?

Preliminary Ideas:
Choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to minimize the following?

$$\sum_{i=1}^{N} |\hat{\epsilon}_i| = \sum_{i=1}^{N} \left| \text{Vote}_i - \widehat{\text{Vote}_i} \right|$$

Good idea, but not the most common approach.

Fitting a Bivariate Regression

The classic approach is to choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to minimize the **sum of squared residuals**:

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2$$

Fitting a Bivariate Regression

The classic approach is to choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to minimize the **sum of squared residuals**:

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2$$

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2 \quad = \quad \sum_{i=1}^{N} \left(\text{Vote}_i - \widehat{\text{Vote}_i} \right)^2$$

Fitting a Bivariate Regression

The classic approach is to choose $\hat{\beta}_0$ and $\hat{\beta}_1$ to minimize the **sum of squared residuals**:

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2$$

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2 = \sum_{i=1}^{N} \left(\text{Vote}_i - \widehat{\text{Vote}_i}\right)^2$$

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2 = \sum_{i=1}^{N} \left(\text{Vote}_i - \hat{\beta}_0 - \hat{\beta}_1 \text{Approval}_i\right)^2$$

Terminology

In statistics, the quantity we want to minimize is often called the:
Loss Function (also sometimes called the **Objective Function**)

Terminology

In statistics, the quantity we want to minimize is often called the: **Loss Function** (also sometimes called the **Objective Function**)

The following loss function

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2$$

is the **Sum of Squared Residuals**.

Terminology

In statistics, the quantity we want to minimize is often called the:
Loss Function (also sometimes called the **Objective Function**)

The following loss function

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2$$

is the **Sum of Squared Residuals**.
It is also often called the **Residual Sum of Squares**,
Sum of Squared Error, or **Residual Squared Error**

Terminology

In statistics, the quantity we want to minimize is often called the:
Loss Function (also sometimes called the **Objective Function**)

The following loss function

$$\sum_{i=1}^{N} \hat{\epsilon}_i^2$$

is the **Sum of Squared Residuals**.
It is also often called the **Residual Sum of Squares**,
Sum of Squared Error, or **Residual Squared Error**

Ordinary Least Squares (OLS) Regression: Choosing parameter values (in bivariate case, $\hat{\beta}_0$ and $\hat{\beta}_1$) to minimize the sum of squared residuals

OLS Regression: Choose the coefficients to **minimize the sum of squared residuals**.

In our bivariate case:

Bivariate Ordinary Least Squares Regression

OLS Regression: Choose the coefficients to **minimize the sum of squared residuals**.

In our bivariate case:

$$\underset{\hat{\beta}_0, \hat{\beta}_1}{\arg\min} \quad \sum_{i=1}^{N} \left(\text{Vote}_i - \hat{\beta}_0 - \hat{\beta}_1 \text{Approval}_i \right)^2$$

Bivariate Ordinary Least Squares Regression

OLS Regression: Choose the coefficients to **minimize the sum of squared residuals**.

In our bivariate case:

$$\underset{\hat{\beta}_0, \hat{\beta}_1}{\arg\min} \quad \sum_{i=1}^{N} \left(\text{Vote}_i - \hat{\beta}_0 - \hat{\beta}_1 \text{Approval}_i \right)^2$$

Minimization of loss function can be accomplished via:

1. Calculus
2. Computer optimization algorithms

Bivariate Ordinary Least Squares Solution

$$\underset{\hat{\beta}_0,\hat{\beta}_1}{\arg\min} \sum_{i=1}^{N} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \right)^2$$

Bivariate Ordinary Least Squares Solution

$$\underset{\hat{\beta}_0,\hat{\beta}_1}{\arg\min} \quad \sum_{i=1}^{N} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i \right)^2$$

Recall that finding an extreme (minimum or maximum) involves taking the derivative and setting to zero.

Bivariate Ordinary Least Squares Solution

$$S(\hat{\beta}_0, \hat{\beta}_1) = \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$$

Bivariate Ordinary Least Squares Solution

$$
\begin{aligned}
S(\hat{\beta}_0, \hat{\beta}_1) &= \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \\
&= \sum_{i=1}^{N} (y_i^2 - 2y_i\hat{\beta}_0 - 2y_i\hat{\beta}_1 x_i + \hat{\beta}_0^2 + 2\hat{\beta}_0\hat{\beta}_1 x_i + \hat{\beta}_1^2 x_i^2)
\end{aligned}
$$

Bivariate Ordinary Least Squares Solution

$$
\begin{aligned}
S(\hat{\beta}_0, \hat{\beta}_1) &= \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \\
&= \sum_{i=1}^{N} (y_i^2 - 2y_i\hat{\beta}_0 - 2y_i\hat{\beta}_1 x_i + \hat{\beta}_0^2 + 2\hat{\beta}_0\hat{\beta}_1 x_i + \hat{\beta}_1^2 x_i^2)
\end{aligned}
$$

$$
\frac{\partial S(\hat{\beta}_0, \hat{\beta}_1)}{\partial \hat{\beta}_0} =
$$

Bivariate Ordinary Least Squares Solution

$$
\begin{aligned}
S(\hat{\beta}_0, \hat{\beta}_1) &= \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \\
&= \sum_{i=1}^{N} (y_i^2 - 2y_i\hat{\beta}_0 - 2y_i\hat{\beta}_1 x_i + \hat{\beta}_0^2 + 2\hat{\beta}_0\hat{\beta}_1 x_i + \hat{\beta}_1^2 x_i^2)
\end{aligned}
$$

$$
\frac{\partial S(\hat{\beta}_0, \hat{\beta}_1)}{\partial \hat{\beta}_0} = \sum_{i=1}^{N} (-2y_i + 2\hat{\beta}_0 + 2\hat{\beta}_1 x_i)
$$

Bivariate Ordinary Least Squares Solution

$$
\begin{aligned}
S(\hat{\beta}_0, \hat{\beta}_1) &= \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \\
&= \sum_{i=1}^{N} (y_i^2 - 2y_i\hat{\beta}_0 - 2y_i\hat{\beta}_1 x_i + \hat{\beta}_0^2 + 2\hat{\beta}_0\hat{\beta}_1 x_i + \hat{\beta}_1^2 x_i^2)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial S(\hat{\beta}_0, \hat{\beta}_1)}{\partial \hat{\beta}_0} &= \sum_{i=1}^{N} (-2y_i + 2\hat{\beta}_0 + 2\hat{\beta}_1 x_i) \\
\frac{\partial S(\hat{\beta}_0, \hat{\beta}_1)}{\partial \hat{\beta}_1} &=
\end{aligned}
$$

Bivariate Ordinary Least Squares Solution

$$
\begin{aligned}
S(\hat{\beta}_0, \hat{\beta}_1) &= \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 \\
&= \sum_{i=1}^{N} (y_i^2 - 2y_i\hat{\beta}_0 - 2y_i\hat{\beta}_1 x_i + \hat{\beta}_0^2 + 2\hat{\beta}_0\hat{\beta}_1 x_i + \hat{\beta}_1^2 x_i^2)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial S(\hat{\beta}_0, \hat{\beta}_1)}{\partial \hat{\beta}_0} &= \sum_{i=1}^{N} (-2y_i + 2\hat{\beta}_0 + 2\hat{\beta}_1 x_i) \\
\frac{\partial S(\hat{\beta}_0, \hat{\beta}_1)}{\partial \hat{\beta}_1} &= \sum_{i=1}^{N} (-2y_i x_i + 2\hat{\beta}_0 x_i + 2\hat{\beta}_1 x_i^2)
\end{aligned}
$$

Bivariate Least Squares Solution

The first order conditions are

$$0 = \sum_{i=1}^{N} (-2y_i + 2\hat{\beta}_0 + 2\hat{\beta}_1 x_i)$$

and

$$0 = \sum_{i=1}^{N} (-2y_i x_i + 2\hat{\beta}_0 x_i + 2\hat{\beta}_1 x_i^2)$$

Bivariate Least Squares Solution

The first order conditions are

$$0 = \sum_{i=1}^{N}(-2y_i + 2\hat{\beta}_0 + 2\hat{\beta}_1 x_i)$$

and

$$0 = \sum_{i=1}^{N}(-2y_i x_i + 2\hat{\beta}_0 x_i + 2\hat{\beta}_1 x_i^2)$$

Now solving for $\hat{\beta}_0$ and $\hat{\beta}_1$ yields the following equations:

$$\hat{\beta}_0 N = \left(\sum_{i=1}^{N} y_i\right) - \hat{\beta}_1 \left(\sum_{i=1}^{N} x_i\right)$$

$$\hat{\beta}_1 \sum_{i=1}^{N} x_1^2 = \left(\sum_{i=1}^{N} x_i y_i\right) - \hat{\beta}_0 \left(\sum_{i=1}^{N} x_i\right)$$

Bivariate Least Squares Solution

And these equations

$$\hat{\beta}_0 N = \left(\sum_{i=1}^{N} y_i\right) - \hat{\beta}_1 \left(\sum_{i=1}^{N} x_i\right)$$

$$\hat{\beta}_1 \sum_{i=1}^{N} x_1^2 = \left(\sum_{i=1}^{N} x_i y_i\right) - \hat{\beta}_0 \left(\sum_{i=1}^{N} x_i\right)$$

Bivariate Least Squares Solution

And these equations

$$\hat{\beta}_0 N = \left(\sum_{i=1}^{N} y_i\right) - \hat{\beta}_1 \left(\sum_{i=1}^{N} x_i\right)$$

$$\hat{\beta}_1 \sum_{i=1}^{N} x_1^2 = \left(\sum_{i=1}^{N} x_i y_i\right) - \hat{\beta}_0 \left(\sum_{i=1}^{N} x_i\right)$$

rearranged yield the bivariate ordinary least squares (OLS) estimators:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} = \frac{Cov(x, y)}{Var(x)}$$

Our Estimated Prediction Function

$$\text{Vote}_i = \widehat{\beta}_0 + \widehat{\beta}_1 \text{Approval}_i + \hat{\epsilon}_i$$

Our Estimated Prediction Function

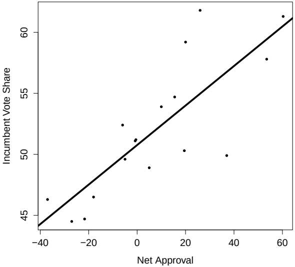
$$\begin{aligned}
\text{Vote}_i &= \widehat{\beta}_0 + \widehat{\beta}_1 \text{Approval}_i + \hat{\epsilon}_i \\
\text{Vote}_i &= 50.76 + 0.16 \times \text{Approval}_i + \hat{\epsilon}_i
\end{aligned}$$

Our Estimated Prediction Function

$$\text{Vote}_i = \widehat{\beta}_0 + \widehat{\beta}_1 \text{Approval}_i + \hat{\epsilon}_i$$
$$\text{Vote}_i = 50.76 + 0.16 \times \text{Approval}_i + \hat{\epsilon}_i$$

Our Estimated **Prediction Function:**

$$\widehat{\text{Vote}_i} = 50.76 + 0.16 \times \text{Approval}_i$$



Test Your Knowledge

$$\widehat{\text{Vote}}_i = 50.76 + 0.16 \times \text{Approval}_i$$

Interpret the parameter values:

Test Your Knowledge

$$\widehat{\text{Vote}}_i = 50.76 + 0.16 \times \text{Approval}_i$$

Interpret the parameter values:

$$\hat{\beta}_0 = 50.76$$

Test Your Knowledge

$$\widehat{\text{Vote}}_i = 50.76 + 0.16 \times \text{Approval}_i$$

Interpret the parameter values:

$$\hat{\beta}_0 = 50.76$$

$$\hat{\beta}_1 = 0.16$$

Predicting Election Result

Consider the following Gallup polling numbers:

Predicting Election Result

Consider the following Gallup polling numbers:

- 46% Approve, 50% Disapprove
- -4% Net Approval

Predicting Election Result

Consider the following Gallup polling numbers:

- 46% Approve, 50% Disapprove
- -4% Net Approval

$$\widehat{\text{Vote}}_{2024} \quad = \quad \hat{\beta}_0 + \hat{\beta}_1 \times \text{Approval}_{2024}$$

Predicting Election Result

Consider the following Gallup polling numbers:

- 46% Approve, 50% Disapprove
- -4% Net Approval

$$
\begin{aligned}
\widehat{\text{Vote}}_{2024} &= \hat{\beta}_0 + \hat{\beta}_1 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times \text{Approval}_{2024}
\end{aligned}
$$

Predicting Election Result

Consider the following Gallup polling numbers:

- 46% Approve, 50% Disapprove
- -4% Net Approval

$$
\begin{aligned}
\widehat{\text{Vote}}_{2024} &= \hat{\beta}_0 + \hat{\beta}_1 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times -4
\end{aligned}
$$

Predicting Election Result

Consider the following Gallup polling numbers:

- 46% Approve, 50% Disapprove
- -4% Net Approval

$$
\begin{aligned}
\widehat{\text{Vote}}_{2024} &= \hat{\beta}_0 + \hat{\beta}_1 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times -4 \\
&= 50.76 - 0.64 = 50.12
\end{aligned}
$$

Predicting Election Result

Consider the following Gallup polling numbers:

- 46% Approve, 50% Disapprove
- -4% Net Approval

$$\begin{aligned}
\widehat{\text{Vote}}_{2024} &= \hat{\beta}_0 + \hat{\beta}_1 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times -4 \\
&= 50.76 - 0.64 = 50.12
\end{aligned}$$

What about the following:

- Net approval: 105%

Predicting Election Result

Consider the following Gallup polling numbers:

- 46% Approve, 50% Disapprove
- -4% Net Approval

$$
\begin{aligned}
\widehat{\text{Vote}}_{2024} &= \hat{\beta}_0 + \hat{\beta}_1 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times -4 \\
&= 50.76 - 0.64 = 50.12
\end{aligned}
$$

What about the following:

- Net approval: 105%
- Net approval: -50%

Predicting Election Result

Consider the following Gallup polling numbers:

- 46% Approve, 50% Disapprove
- -4% Net Approval

$$
\begin{aligned}
\widehat{\text{Vote}}_{2024} &= \hat{\beta}_0 + \hat{\beta}_1 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times \text{Approval}_{2024} \\
\widehat{\text{Vote}}_{2024} &= 50.76 + 0.16 \times -4 \\
&= 50.76 - 0.64 = 50.12
\end{aligned}
$$

What about the following:

- Net approval: 105%
- Net approval: -50%

Caution: Function is defined for all values!