PLSC 502 – Fall 2023
Linear Regression II

November 27, 2023

Inference

Variation in $\hat{\beta}_0$ and $\hat{\beta}_1$

$\hat{\beta}_0$ and $\hat{\beta}_1$ are *random variables*...

- Q: Where does their variation come from?
- A: From the *stochastic* variation in Y...
- ...that is, from u.

Next question: What does the random variation in Y "look like"?

Setup

For the linear model:

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

Note that we can rewrite the formula for $\hat{\beta}_1$:

$$
\begin{aligned}
\hat{\beta}_1 &= \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{N}(X_i - \bar{X})^2} \\
&= \frac{\sum_{i=1}^{N}(X_i - \bar{X})Y_i - \sum_{i=1}^{N}(X_i - \bar{X})\bar{Y}}{\sum_{i=1}^{N}(X_i - \bar{X})^2} \\
&= \frac{\sum_{i=1}^{N}(X_i - \bar{X})Y_i - \bar{Y}\sum_{i=1}^{N}(X_i - \bar{X})}{\sum_{i=1}^{N}(X_i - \bar{X})^2} \\
&= \frac{\sum_{i=1}^{N}(X_i - \bar{X})Y_i - \bar{Y}(0)}{\sum_{i=1}^{N}(X_i - \bar{X})^2} \\
&= \frac{\sum(X_i - \bar{X})Y_i}{\sum(X_i - \bar{X})^2}
\end{aligned}
$$

An assumption:

$$u_i \sim \text{i.i.d. } N(0, \sigma^2)$$

Implies:

$$\text{Var}(Y|X, \beta) = \sigma^2$$

so:

$$\begin{aligned}
\text{Var}(\hat{\beta}_1) &= \text{Var}\left[\frac{\sum_{i=1}^{N}(X_i - \bar{X})Y_i}{\sum_{i=1}^{N}(X_i - \bar{X})^2}\right] \\
&= \left[\frac{1}{\sum(X_i - \bar{X})^2}\right]^2 \sum(X_i - \bar{X})^2 \,\text{Var}(Y) \\
&= \left[\frac{1}{\sum(X_i - \bar{X})^2}\right]^2 \sum(X_i - \bar{X})^2 \,\sigma^2 \\
&= \frac{\sigma^2}{\sum(X_i - \bar{X})^2}.
\end{aligned}$$

$\mathrm{Var}(\hat{\beta}_0)$ and $\mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1)$

Similarly:

$$\mathrm{Var}(\hat{\beta}_0) = \frac{\sum X_i^2}{N \sum (X_i - \bar{X})^2} \sigma^2$$

and :

$$\mathrm{Cov}(\hat{\beta}_0, \hat{\beta}_1) = \frac{-\bar{X}}{\sum (X_i - \bar{X})^2} \sigma^2$$

Note that:

- Var($\hat{\beta}_0$) and Var($\hat{\beta}_1$) $\propto \sigma^2$
  $\hookrightarrow$ Var($\hat{\beta}$s) increases as Y gets "noisier"...

- Var($\hat{\beta}_0$) and Var($\hat{\beta}_1$) $\propto -\sum(X_i - \bar{X})$
  $\hookrightarrow$ Var($\hat{\beta}$s) decreases with greater variation in X...

- Var($\hat{\beta}_0$) and Var($\hat{\beta}_1$) $\propto -N$
  $\hookrightarrow$ Var($\hat{\beta}$s) decreases as N gets larger...

- sign[Cov($\hat{\beta}_0, \hat{\beta}_1$)] $= -$sign($\bar{X}$)
  $\hookrightarrow$ The sign of the covariance between $\hat{\beta}_0$ and $\hat{\beta}_1$ is the opposite of the sign of the mean of X

"Given the assumptions of the classical linear regression model, the least squares estimators are the minimum variance estimators among the class of unbiased linear estimators. (They are BLUE)."

Imagine:

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

Rewrite:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{N} (X_i - \bar{X}) Y_i}{\sum_{i=1}^{N} (X_i - \bar{X})^2}.$$

k are "weights":

$$\hat{\beta}_1 = \sum k_i Y_i$$

with $k_i = \frac{X_i - \bar{X}}{\sum (X_i - \bar{X})^2}$.

Gauss-Markov (continued)

Alternative (non-LS) estimator:

$$\tilde{\beta}_1 = \sum w_i Y_i$$

Unbiasedness requires:

$$\begin{aligned}
\mathrm{E}(\tilde{\beta}_1) &= \sum w_i \mathrm{E}(Y_i) \\
&= \sum w_i (\beta_0 + \beta_1 X_i) \\
&= \beta_0 \sum w_i + \beta_1 \sum w_i X_i
\end{aligned}$$

Variance:

$$\begin{aligned}
\mathrm{Var}(\tilde{\beta}_1) &= \mathrm{Var}\left(\sum w_i Y_i\right) \\
&= \sigma^2 \sum w_i^2 \\
&= \sigma^2 \sum \left[w_i - \frac{X_i - \bar{X}}{\sum(X_i - \bar{X})^2} + \frac{X_i - \bar{X}}{\sum(X_i - \bar{X})^2}\right]^2 \\
&= \sigma^2 \sum \left[w_i - \frac{X_i - \bar{X}}{\sum(X_i - \bar{X})^2}\right]^2 + \sigma^2 \left[\frac{1}{\sum(X_i - \bar{X})^2}\right]
\end{aligned}$$

Gauss-Markov (continued)

Because $\sigma^2 \left[\frac{1}{\sum(X_i - \bar{X})^2}\right]$ is a constant, $\min[\text{Var}(\tilde{\beta}_1)]$ minimizes

$$\sum \left[w_i - \frac{X_i - \bar{X}}{\sum(X_i - \bar{X})^2} \right]^2.$$

Minimized at:

$$w_i = \frac{X_i - \bar{X}}{\sum(X_i - \bar{X})^2}.$$

implying:

$$\begin{aligned} \text{Var}(\tilde{\beta}_1) &= \frac{\sigma^2}{\sum(X_i - \bar{X})^2} \\ &= \text{Var}(\hat{\beta}_1) \end{aligned}$$

For the Gauss-Markov theorem to hold, it must be the case that:

1. $\mathrm{E}(u) = 0$
2. $\mathrm{Cov}(X, u) = 0$

3a. $\mathrm{Var}(u) = \sigma^2 \; \forall \; i$

3b. $\mathrm{Cov}(u_i, u_j) = 0$

4. $\mathrm{Rank}(\mathbf{X}) = k$
5. $u \sim$ i.i.d. $N(0, \sigma^2)$

(...don't sweat these too much for now...)

BLUE, BUE, and Linearity

BLUE vs. BUE:

- OLS has been BLUE since about 1821 (see, e.g., Plackett 1949).
- Hansen (2022): OLS is "BUE" – most efficient among *all* unbiased estimators, linear or otherwise...
- Challenged by others; resolved by Portnoy (2022): Any unbiased estimator *must* be linear (so "BLUE" = "BUE").
- A pretty good nontechnical discussion of all this by Paul Allison is here.

If $u_i \sim N(0, \sigma^2)$, then:

$$\hat{\beta}_0 \sim N[\beta_0, \text{Var}(\hat{\beta}_0)]$$

and

$$\hat{\beta}_1 \sim N[\beta_1, \text{Var}(\hat{\beta}_1)]$$

Means:

$$\begin{aligned} z_{\hat{\beta}_1} &= \frac{(\hat{\beta}_1 - \beta_1)}{\sqrt{\text{Var}(\hat{\beta}_1)}} \\ &= \frac{(\hat{\beta}_1 - \beta_1)}{\text{s.e.}(\hat{\beta}_1)} \\ &= \sim N(0, 1) \end{aligned}$$

A Small Problem...

$\sigma^2 =???$

Solution: use

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

Gives:

$$\widehat{\mathrm{Var}(\hat{\beta}_1)} = \frac{\hat{\sigma}^2}{\sum (X_i - \bar{X})^2},$$

and

$$\widehat{\mathrm{Var}(\hat{\beta}_0)} = \frac{\sum X_i^2}{N \sum (X_i - \bar{X})^2} \hat{\sigma}^2$$

The estimated standard error:

$$\begin{aligned}
\widehat{\text{s.e.}(\hat{\beta}_1)} &= \sqrt{\widehat{\text{Var}(\hat{\beta}_1)}} \\
&= \sqrt{\frac{\hat{\sigma}^2}{\sum(X_i - \bar{X})^2}} \\
&= \frac{\hat{\sigma}}{\sqrt{\sum(X_i - \bar{X})^2}}
\end{aligned}$$

implies:

$$\begin{aligned}
t_{\hat{\beta}_1} \equiv \frac{(\hat{\beta}_1 - \beta_1)}{\widehat{\text{s.e.}(\hat{\beta}_1)}} &= \frac{(\hat{\beta}_1 - \beta_1)}{\frac{\hat{\sigma}}{\sqrt{\sum(X_i - \bar{X})^2}}} \\
&= \frac{(\hat{\beta}_1 - \beta_1)\sqrt{\sum(X_i - \bar{X})^2}}{\hat{\sigma}} \\
&\sim t_{N-k}
\end{aligned}$$

Note: We can derive a similar formula for $\widehat{\text{s.e.}(\hat{\beta}_0)}$...

In practice:

- ...we test the hypothesis $\hat{\beta}_1 = k$ by
  - ...calculating $\hat{t} = \frac{\hat{\beta}_1 - k}{\widehat{\text{s.e.}(\hat{\beta}_1)}}$, then
  - ...calculating the P-value associated with $\hat{t}$.

- ...we calculate $(1 - \alpha) \times 100$-percent confidence intervals around $\hat{\beta}_1$ by
  - ...calculating the t-value $\hat{t}$ associated with the $(1 - \alpha) \times 100$-percent confidence level,
  - multiplying $\hat{t}$ times $\widehat{\text{s.e.}(\hat{\beta}_1)}$ to get the width of the confidence interval, and
  - creating the confidence interval around $\hat{\beta}_1$ according to:

$$\hat{\beta}_1 - 1/2[\hat{t} \times \widehat{\text{s.e.}(\hat{\beta}_1)}] \ < \ \hat{\beta}_1 \ < \ \hat{\beta}_1 + 1/2[\hat{t} \times \widehat{\text{s.e.}(\hat{\beta}_1)}]$$

Point prediction:

$$\hat{Y}_k = \hat{\beta}_0 + \hat{\beta}_1 X_k$$

Y_k is unbiased:

$$\begin{aligned}
\mathrm{E}(\hat{Y}_k) &= \mathrm{E}(\hat{\beta}_0 + \hat{\beta}_1 X_k) \\
&= \mathrm{E}(\hat{\beta}_0) + X_k \mathrm{E}(\hat{\beta}_1) \\
&= \beta_0 + \beta_1 X_k \\
&= \mathrm{E}(Y_k)
\end{aligned}$$

Variability:

$$\begin{aligned}
\mathrm{Var}(\hat{Y}_k) &= \mathrm{Var}(\hat{\beta}_0 + \hat{\beta}_1 X_k) \\
&= \frac{\sum X_i^2}{N \sum (X_i - \bar{X})^2} \sigma^2 + \left[\frac{\sigma^2}{\sum (X_i - \bar{X})^2} \right] X_k^2 + 2 \left[\frac{-\bar{X}}{\sum (X_i - \bar{X})^2} \sigma^2 \right] X_k \\
&= \sigma^2 \left[\frac{1}{N} + \frac{(X_k - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right]
\end{aligned}$$

Variability of Predictions

$$\text{Var}(\hat{Y}_k) = \sigma^2 \left[\frac{1}{N} + \frac{(X_k - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right]$$

means that $\text{Var}(\hat{Y}_k)$:

- Decreases in N
- Decreases in $\text{Var}(X)$
- Increases in $|X - \bar{X}|$

Standard error of the prediction:

$$\widehat{\text{s.e.}(\hat{Y}_k)} = \sqrt{\sigma^2 \left[\frac{1}{N} + \frac{(X_k - \bar{X})^2}{\sum(X_i - \bar{X})^2} \right]}$$

$\rightarrow$ (e.g.) confidence intervals:

$$95\%\,\text{c.i.}(\hat{Y}_k) = \hat{Y}_k \pm [1.96 \times \widehat{\text{s.e.}(\hat{Y}_k)}]$$

Example Redux: SCOTUS Voting, OT1946-2021

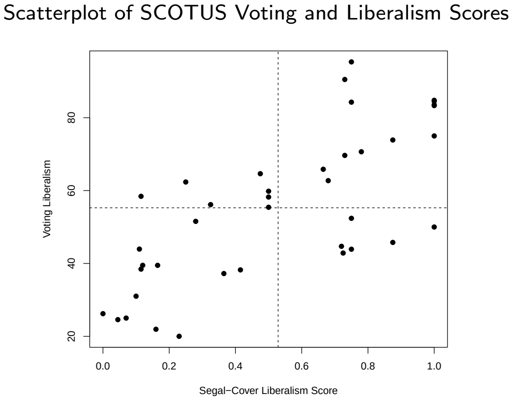
Data from the Supreme Court Database and the justices' Segal-Cover scores...

- Y is `CivLibs` = liberal voting percentage in civil rights & liberties cases
- X is `IdeologyScore` $\in [0, 1] \rightarrow$ SCOTUS justice liberalism

```
> describe(SCOTUS,skew=FALSE,trim=0)
                    vars  n    mean    sd     min     max range   se
justice                1 38   97.37 11.32   78.00  116.00 38.00 1.84
justiceName*           2 38   19.50 11.11    1.00   38.00 37.00 1.80
CivLibs                3 38   56.49 19.94   22.36   95.33 72.97 3.23
Nom.Order*             4 38   19.50 11.11    1.00   38.00 37.00 1.80
Nominee*               5 38   19.50 11.11    1.00   38.00 37.00 1.80
ChiefJustice*          6  4    1.00  0.00    1.00    1.00  0.00 0.00
SenateVote*            7 38   17.05  8.23    1.00   25.00 24.00 1.33
IdeologyScore          8 38    0.54  0.33    0.00    1.00  1.00 0.05
QualificationsScore*   9 38   16.45  7.91    1.00   25.00 24.00 1.28
Nominator (Party)*    10 38    7.03  3.72    1.00   13.00 12.00 0.60
Year                  11 38 1969.74 24.70 1937.00 2018.00 81.00 4.01
```

Example Redux: SCOTUS Voting

```
> with(SCOTUS, describe(CivLibs))
   vars  n mean   sd median trimmed  mad min  max range skew kurtosis   se
X1    1 39 55.3 20.7   55.4    55.1 23.6  20 95.3  75.3 0.13    -1.03 3.31

> with(SCOTUS, describe(IdeologyScore))
   vars  n mean   sd median trimmed mad min max range  skew kurtosis   se
X1    1 39 0.53 0.33    0.5    0.53 0.4   0   1     1 -0.06    -1.47 0.05
```



Example, Continued

```
> SCLib<-lm(CivLibs~IdeologyScore,data=SCOTUS)
> summary(SCLib)   # regression

Call:
lm(formula = CivLibs ~ IdeologyScore, data = SCOTUS)

Residuals:
   Min      1Q Median      3Q     Max
-26.70 -10.01   2.79   7.64  29.99

Coefficients:
              Estimate Std. Error t value    Pr(>|t|)
(Intercept)      31.24       4.33    7.21 0.000000015 ***
IdeologyScore    45.47       6.96    6.53 0.000000121 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14.3 on 37 degrees of freedom
Multiple R-squared:  0.535,Adjusted R-squared:  0.523
F-statistic: 42.6 on 1 and 37 DF,  p-value: 0.000000121
```

Example, Continued

```
> anova(SCLib)     # ANOVA

Analysis of Variance Table

Response: CivLibs
              Df Sum Sq Mean Sq F value     Pr(>F)
IdeologyScore  1   8693    8693    42.6 0.00000012 ***
Residuals     37   7543     204
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$\widehat{\text{Var}}(\hat{\beta})$:

```
> vcov(SCLib)
              (Intercept) IdeologyScore
(Intercept)          18.8         -25.6
IdeologyScore       -25.6          48.5
```

Estimated standard errors of $\hat{\beta}_0$ and $\hat{\beta}_1$:

```
> sqrt(vcov(SCLib))
              (Intercept) IdeologyScore
(Intercept)          4.33           NaN
IdeologyScore         NaN          6.96
Warning message:
In sqrt(vcov(SCLib)) : NaNs produced
```

95 percent c.i.s:

```
> confint(SCLib)
              2.5 % 97.5 %
(Intercept)   24.56  42.22
IdeologyScore 29.02  57.07
```

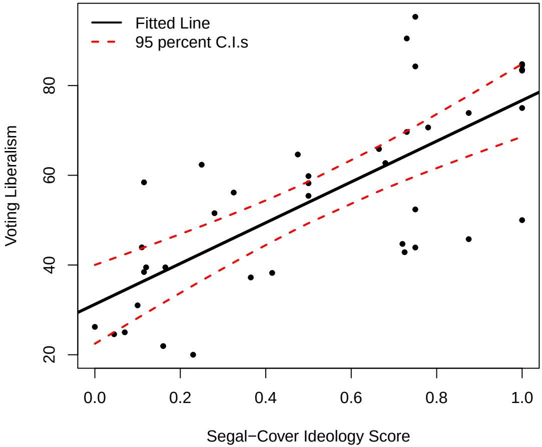
99 percent c.i.s:

```
> confint(SCLib,level=0.99)
              0.5 % 99.5 %
(Intercept)   21.55  45.23
IdeologyScore 24.23  61.85
```

```
> SEs<-predict(SCLib,interval="confidence")
> SEs
    fit  lwr  upr
1  71.0 64.3 77.8
2  64.2 58.8 69.6
3  61.5 56.5 66.5
4  64.4 59.0 69.9
5  76.7 68.6 84.8
.
.
<rows omitted>
.
.
36 64.4 59.0 69.9
37 36.2 28.7 43.7
38 34.4 26.5 42.4
39 41.7 35.4 48.0
```

A Plot, With CIs

Scatterplot of SCOTUS Voting and Ideology Scores, along with Least-Squares Line and 95% Prediction Confidence Intervals



Model Fit

Model fit is:

- The closeness of the mapping between model-based values of Y and actual values of Y...
- Can be *in-sample* or *out-of-sample* ($\rightarrow$ "overfitting")
- Is (in part) a function of *model specification* (choice of predictors, functional form, interactions, etc.)
- Related (but not identical) to prediction / predictive ability

R^2 Introduced

Recall that for

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

We have:

$$\begin{aligned} \text{“TSS”} &= \sum (Y_i - \bar{Y})^2 \\ \text{“MSS”} &= \sum (\hat{Y}_i - \bar{Y})^2 \\ \text{“RSS”} &= \sum (Y_i - \hat{Y}_i)^2 \equiv \sum \hat{u}_i^2 \end{aligned}$$

Then:

$$\begin{aligned} R^2 &= \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \\ &= \frac{\text{MSS}}{\text{TSS}} \\ &= 1 - \frac{\text{RSS}}{\text{TSS}} \\ &= 1 - \frac{\sum \hat{u}_i^2}{\sum (Y_i - \bar{Y})^2} \end{aligned}$$

R^2 Is...

R-squared:

- is "the proportion of variance explained"
- $\in [0, 1]$
  - $R^2 = 1.0 \equiv$ a "perfect (linear) fit"'
  - $R^2 = 0 \equiv$ no (linear) $X - Y$ association

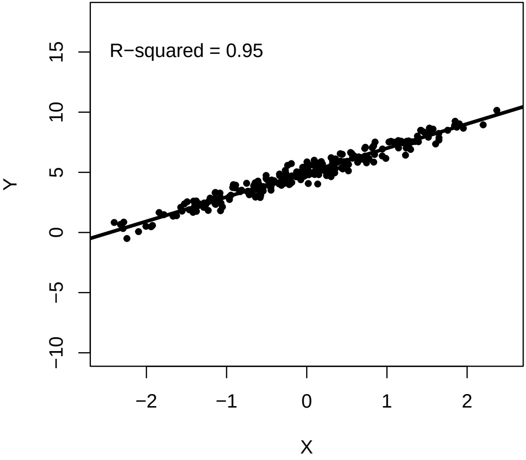
For a single X,

$$
\begin{aligned}
R^2 &= \hat{\beta}_1^2 \frac{\sum(X_i - \bar{X})^2}{\sum(Y_i - \bar{Y})^2} \\
&= (r_{XY})^2
\end{aligned}
$$

A (Simulated) Example

```
seed <- 7222009
set.seed(seed)
> X<-rnorm(250)
> Y1<-5+2*X+rnorm(250,mean=0,sd=sqrt(0.2))
> Y2<-5+2*X+rnorm(250,mean=0,sd=sqrt(20))
> fit<-lm(Y1~X)
> summary(fit)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.97712    0.02846  174.86   <2e-16 ***
X            2.02529    0.02785   72.73   <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 0.4491 on 248 degrees of freedom
Multiple R-squared:  0.9552,Adjusted R-squared:  0.955
F-statistic:  5290 on 1 and 248 DF,  p-value: < 2.2e-16
```

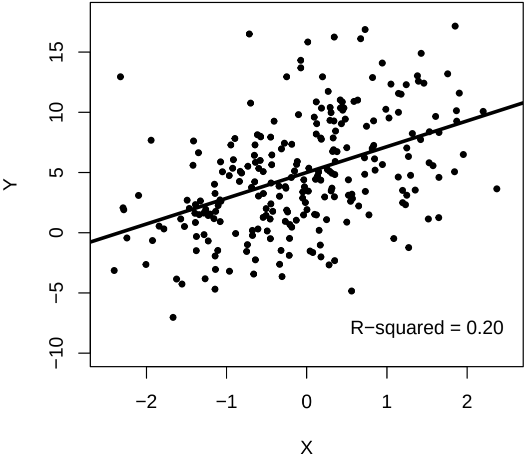
Regression of $Y_i = 5 + 2X_i + u_i$ ($R^2 = 0.95$)



Same Slope/Intercept, Different R^2

```
> fit2<-lm(Y2~X)
> summary(fit2)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.0048     0.2757  18.151  < 2e-16 ***
X             2.1402     0.2697   7.934 7.29e-14 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 4.351 on 248 degrees of freedom
Multiple R-squared:  0.2024,Adjusted R-squared:  0.1992
F-statistic: 62.95 on 1 and 248 DF,  p-value: 7.288e-14
```

Regression of $Y_i = 5 + 2X_i + u_i$ ($R^2 = 0.20$)



R^2 is Also an *Estimate*...

Luskin: Population analogue "P^2":

$$P^2 = 1 - \frac{\sigma^2}{\sigma_Y^2}$$

Then $\hat{P}^2 = R^2$ has variance:

$$\widehat{\text{Var}(R^2)} = \frac{4R^2(1-R^2)^2(N-k)^2}{(N^2-1)(N+3)}$$

and standard error:

$$\widehat{\text{s.e.}(R^2)} = \sqrt{\frac{4R^2(1-R^2)^2(N-k)^2}{(N^2-1)(N+3)}}.$$

Adjusted R^2

"Adjusted" R^2 is:

$$R^2_{adj.} = 1 - \frac{(1 - R^2)(N - c)}{(N - k)}$$

where $c = 1$ if there is a constant in the model and $c = 0$ otherwise.

$R^2_{adj.}$ characteristics:

- $R^2_{adj.} \to R^2$ as $N \to \infty$
- $R^2_{adj.}$ can be > 1, or < 0...
- $R^2_{adj.}$ increases with model "fit," but
- The extent of that increase is discounted by a factor proportional to the number of covariates.

Other R^2 / Goodness-Of-Fit Alternatives

- Standard Error of the Estimate:

$$\text{SEE} = \sqrt{\frac{\text{RSS}}{N-k}}$$

- F-statistic (bivariate regression, for $\beta_1 = 0$):

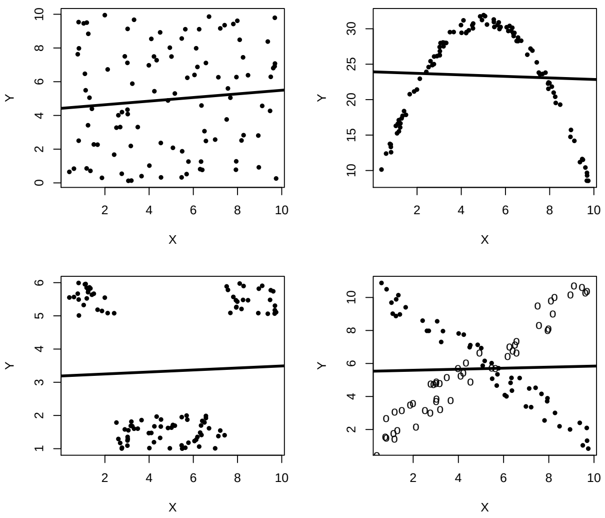
$$\begin{aligned} F &= \frac{\sum(Y_i - \bar{Y})^2 - \sum(Y_i - \hat{Y}_i)^2}{(N-1)-(N-2)} \div \frac{\sum(Y_i - \hat{Y}_i)^2}{(N-2)} \\ &= \frac{\text{“explained” variance}}{\text{“unexplained” variance}} \end{aligned}$$

which is $\sim F(1, N-2)$.

- ROC / AUC (later...)

- Graphical methods

Caution: Different Ways to get $R^2 \approx 0$



Remember This Regression?

```
> fit<-lm(CivLibs~IdeologyScore,data=SCOTUS)
> summary(fit)

Call:
lm(formula = CivLibs ~ IdeologyScore, data = SCOTUS)

Residuals:
   Min      1Q Median      3Q     Max
-26.70 -10.01   2.79   7.64  29.99

Coefficients:
              Estimate Std. Error t value    Pr(>|t|)
(Intercept)      31.24       4.33    7.21 0.000000015 ***
IdeologyScore    45.47       6.96    6.53 0.000000121 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14.3 on 37 degrees of freedom
Multiple R-squared:  0.535,Adjusted R-squared:  0.523
F-statistic: 42.6 on 1 and 37 DF,  p-value: 0.000000121
```

ANOVA Redux

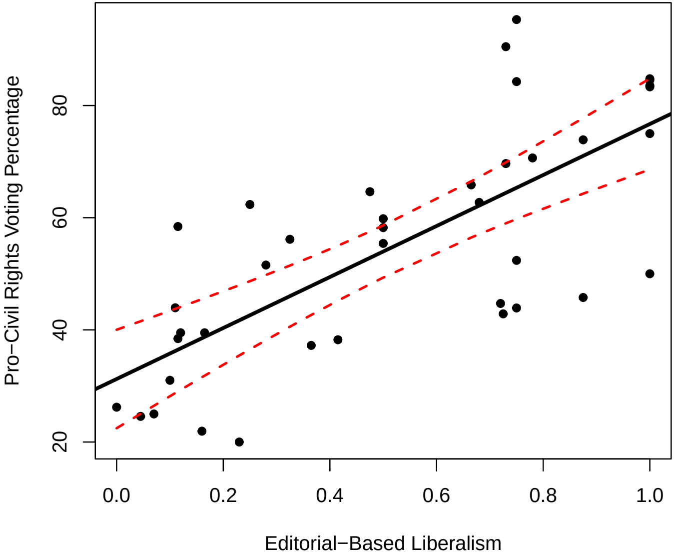
```
> anova(fit)

Analysis of Variance Table

Response: CivLibs
              Df Sum Sq Mean Sq F value     Pr(>F)
IdeologyScore  1   8693    8693    42.6 0.00000012 ***
Residuals     37   7543     204
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> # R-squared:
>
> anova(fit)$‘Sum Sq‘[1] / (anova(fit)$‘Sum Sq‘[1] + anova(fit)$‘Sum Sq‘[2])
[1] 0.535

> # F-statistic:
>
> anova(fit)$‘Mean Sq‘[1] / anova(fit)$‘Mean Sq‘[2]
[1] 42.6
```

Stupid Regression Tricks

SCOTUS Regression Redux

```
> fit<-lm(CivLibs~IdeologyScore,data=SCOTUS)
> summary(fit)

Call:
lm(formula = CivLibs ~ IdeologyScore, data = SCOTUS)

Residuals:
   Min     1Q Median     3Q    Max 
-26.70 -10.01   2.79   7.64  29.99 

Coefficients:
              Estimate Std. Error t value    Pr(>|t|)    
(Intercept)      31.24       4.33    7.21 0.000000015 ***
IdeologyScore    45.47       6.96    6.53 0.000000121 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14.3 on 37 degrees of freedom
Multiple R-squared:  0.535,Adjusted R-squared:  0.523 
F-statistic: 42.6 on 1 and 37 DF,  p-value: 0.000000121
```

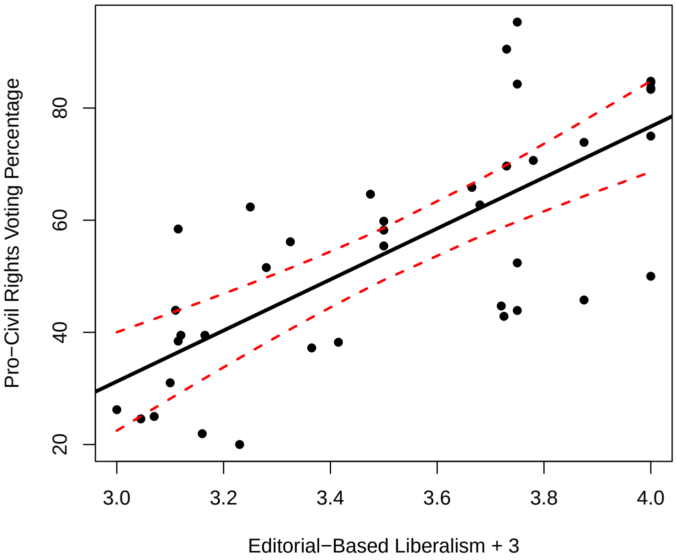
SCOTUS Regression Plot



Add Three to IdeologyScore

```
> SCOTUS$IdeoPlus3 <- SCOTUS$IdeologyScore + 3
>
> fit2<-lm(CivLibs~IdeoPlus3,data=SCOTUS)
> summary(fit2)

Call:
lm(formula = CivLibs ~ IdeoPlus3, data = SCOTUS)

Residuals:
   Min     1Q Median     3Q    Max
-26.70 -10.01   2.79   7.64  29.99

Coefficients:
            Estimate Std. Error t value   Pr(>|t|)
(Intercept)  -105.16      24.68   -4.26    0.00013 ***
IdeoPlus3      45.47       6.96    6.53 0.00000012 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14.3 on 37 degrees of freedom
Multiple R-squared:  0.535,Adjusted R-squared:  0.523
F-statistic: 42.6 on 1 and 37 DF,  p-value: 0.000000121
```

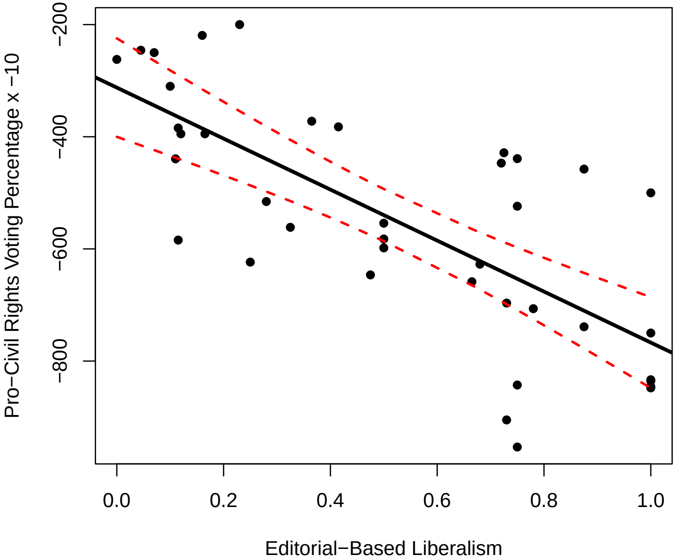
SCOTUS Plot With Rescaled X



Multiply CivLibs Times -10

```
> SCOTUS$CivLibNeg10 <- -10 * SCOTUS$CivLibs
>
> fit3<-lm(CivLibNeg10~IdeologyScore,data=SCOTUS)
> summary(fit3)

Call:
lm(formula = CivLibNeg10 ~ IdeologyScore, data = SCOTUS)

Residuals:
   Min     1Q Median     3Q    Max
-299.9  -76.4  -27.9  100.1  267.0

Coefficients:
              Estimate Std. Error t value    Pr(>|t|)
(Intercept)     -312.4       43.3   -7.21 0.000000015 ***
IdeologyScore   -454.7       69.6   -6.53 0.000000121 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 143 on 37 degrees of freedom
Multiple R-squared:  0.535,Adjusted R-squared:  0.523
F-statistic: 42.6 on 1 and 37 DF,  p-value: 0.000000121
```

SCOTUSplot With Rescaled Y



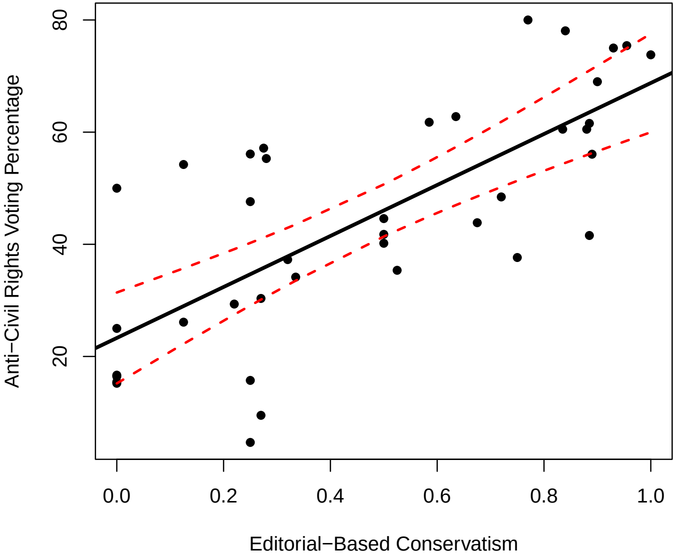
- Adding (subtracting) a positive constant to X shifts the X-axis to the left (right).
- Adding (subtracting) a positive constant to Y shifts the Y-axis downwards (upwards).
- Multiplying X (Y) times a positive constant greater than 1.0 stretches the X (Y) axis.
- Multiplying X (Y) times a positive constant less than 1.0 shrinks the X (Y) axis.
- Multiplying X (Y) times a negative constant inverts the X (Y) axis, and stretches / shrinks it as above.

Linear transformations do not alter the model in a statistically / substantively important way.

Application: Reversing The Scales

```
> SCOTUS$CivLibCons <- 100 - SCOTUS$CivLibs
> SCOTUS$IdeolCons <- 1 - SCOTUS$IdeologyScore

> fit4<-lm(CivLibCons~IdeolCons,data=SCOTUS)
> summary(fit4)

Call:
lm(formula = CivLibCons ~ IdeolCons, data = SCOTUS)

Residuals:
   Min     1Q Median     3Q    Max
-29.99  -7.64  -2.79  10.01  26.70

Coefficients:
            Estimate Std. Error t value   Pr(>|t|)
(Intercept)    23.30       4.00    5.82 0.00000108 ***
IdeolCons      45.47       6.96    6.53 0.00000012 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14.3 on 37 degrees of freedom
Multiple R-squared:  0.535,Adjusted R-squared:  0.523
F-statistic: 42.6 on 1 and 37 DF,  p-value: 0.000000121
```

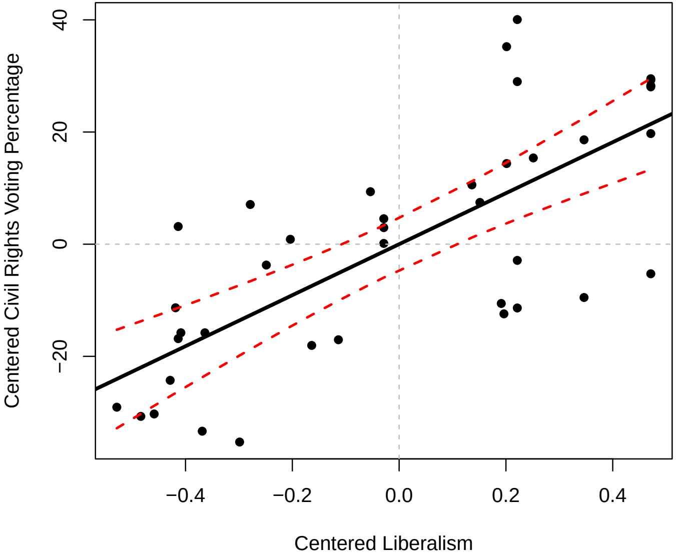
Plot of Civil Liberties Conservatism vs. Ideological Conservatism



Application: “Centering” Variables

```
> SCOTUS$CivLibCentered <- SCOTUS$CivLibs - mean(SCOTUS$CivLibs)
> SCOTUS$IdeolCentered <- SCOTUS$IdeologyScore - mean(SCOTUS$IdeologyScore)
>
> fit5<-lm(CivLibCentered~IdeolCentered,data=SCOTUS)
> summary(fit5)

Call:
lm(formula = CivLibCentered ~ IdeolCentered, data = SCOTUS)

Residuals:
   Min     1Q Median     3Q    Max
-26.70 -10.01   2.79   7.64  29.99

Coefficients:
              Estimate Std. Error t value   Pr(>|t|)
(Intercept)   -7.60e-15   2.29e+00    0.00          1
IdeolCentered  4.55e+01   6.96e+00    6.53 0.00000012 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14.3 on 37 degrees of freedom
Multiple R-squared:  0.535,Adjusted R-squared:  0.523
F-statistic: 42.6 on 1 and 37 DF,  p-value: 0.000000121
```

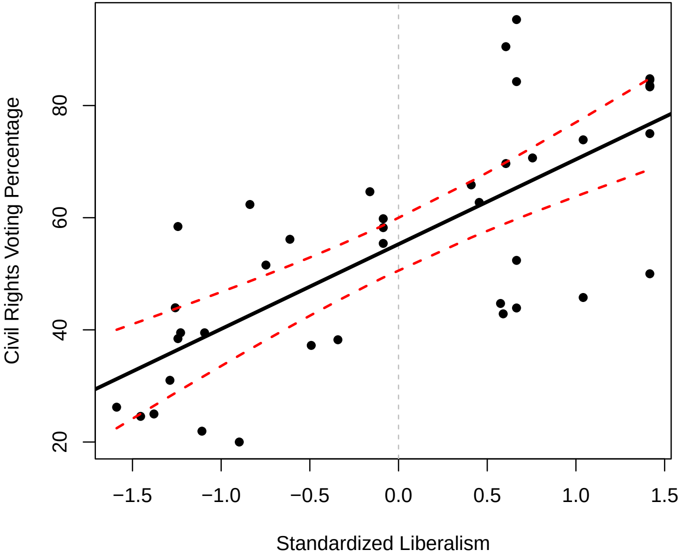
“Regression Through The Origin”



Application: "Standardizing" a Variable

```
> SCOTUS$IdeolStd <- scale(SCOTUS$IdeologyScore)
>
> fit6<-lm(CivLibs~IdeolStd,data=SCOTUS)
> summary(fit6)

Call:
lm(formula = CivLibs ~ IdeolStd, data = SCOTUS)

Residuals:
   Min     1Q Median     3Q    Max
-26.70 -10.01   2.79   7.64  29.99

Coefficients:
            Estimate Std. Error t value   Pr(>|t|)
(Intercept)    55.28       2.29   24.18    < 2e-16 ***
IdeolStd       15.12       2.32    6.53 0.00000012 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14.3 on 37 degrees of freedom
Multiple R-squared:  0.535,Adjusted R-squared:  0.523
F-statistic: 42.6 on 1 and 37 DF,  p-value: 0.000000121
```

OLS with Standardized X



Rescaling for Interpretability

```
> fit7<-lm(CivLibs~Year,data=SCOTUS)
> summary(fit7)

Call:
lm(formula = CivLibs ~ Year, data = SCOTUS)

Residuals:
   Min     1Q Median     3Q    Max
-30.36 -15.32  -3.05  16.15  37.92

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  752.804    234.451    3.21   0.0027 **
Year          -0.354      0.119   -2.98   0.0051 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 18.8 on 37 degrees of freedom
Multiple R-squared:  0.193,Adjusted R-squared:  0.171
F-statistic: 8.85 on 1 and 37 DF,  p-value: 0.00513
```

Rescaling for Interpretability (continued)

```
> SCOTUS$Year1900<-SCOTUS$Year-1900
> fit8<-lm(CivLibs~Year1900,data=SCOTUS)
> summary(fit8)

Call:
lm(formula = CivLibs ~ Year1900, data = SCOTUS)

Residuals:
   Min     1Q Median     3Q    Max 
-30.36 -15.32  -3.05  16.15  37.92 

Coefficients:
            Estimate Std. Error t value       Pr(>|t|)    
(Intercept)   80.413      8.969    8.97 0.000000000082 ***
Year1900      -0.354      0.119   -2.98         0.0051 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 18.8 on 37 degrees of freedom
Multiple R-squared:  0.193,Adjusted R-squared:  0.171 
F-statistic: 8.85 on 1 and 37 DF,  p-value: 0.00513
```

Binary $X \equiv t$-test

```
> SCOTUS$Chief<-ifelse(is.na(SCOTUS$ChiefJustice),0,1)
> fit9<-lm(CivLibs~Chief,data=SCOTUS)
> summary(fit9)

Call:
lm(formula = CivLibs ~ Chief, data = SCOTUS)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    55.71       3.53   15.76   <2e-16 ***
Chief          -4.18      11.03   -0.38     0.71
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 20.9 on 37 degrees of freedom
Multiple R-squared:  0.00387,Adjusted R-squared:  -0.0231
F-statistic: 0.144 on 1 and 37 DF,  p-value: 0.707


> t.test(CivLibs~Chief,data=SCOTUS,var.equal=TRUE)

Two Sample t-test

data:  CivLibs by Chief
t = 0.4, df = 37, p-value = 0.7
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -18.2  26.5
sample estimates:
mean in group 0 mean in group 1
           55.7            51.5
```

The results:

```
> summary(fit)

Call:
lm(formula = CivLibs ~ IdeologyScore, data = SCOTUS)

Residuals:
   Min     1Q Median     3Q    Max
-26.70 -10.01   2.79   7.64  29.99

Coefficients:
              Estimate Std. Error t value    Pr(>|t|)
(Intercept)      31.24       4.33    7.21 0.000000015 ***
IdeologyScore    45.47       6.96    6.53 0.000000121 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 14.3 on 37 degrees of freedom
Multiple R-squared:  0.535,Adjusted R-squared:  0.523
F-statistic: 42.6 on 1 and 37 DF,  p-value: 0.000000121
```

The table:

Table: OLS Regression Model of SCOTUS Voting

Variables	Model I
(Constant)	31.24
	(4.33)
Ideological Liberalism	45.47*
	(6.96)
Adjusted R^2	0.52

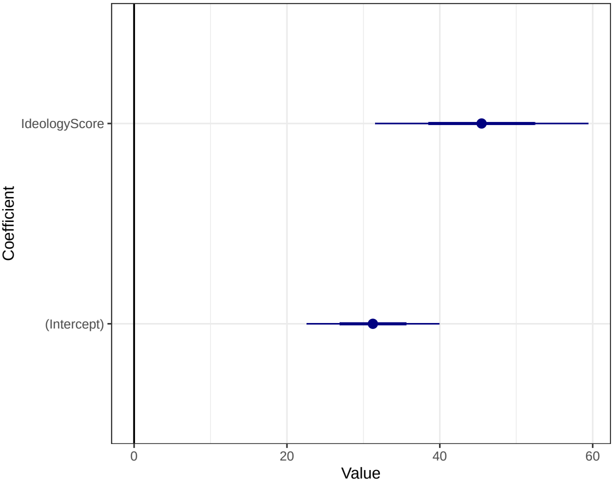
Note: $N = 39$. *Cell entries are coefficient estimates; numbers in parentheses are estimated standard errors. Asterisks indicate* $p < .05$ *(one-tailed). See text for details.*

Another Table (using default-y stargazer)

Table: OLS Regression Model of SCOTUS Voting

	Model I
(Constant)	31.20^{***}
	(4.33)
Ideological Liberalism	45.50^{***}
	(6.96)
Observations	39
R^2	0.54
Adjusted R^2	0.52
Residual Std. Error	14.30 (df = 37)
F Statistic	42.60^{***} (df = 1; 37)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Default-y Ladderplot -fitplot-



Some Guidelines ("Rules"?)

Tables:

- *Use column headings descriptively.*
- *Use multiple rows / columns rather than multiple tables.*
- *Learn about significant digits, and don't report more than 4-5 of them.*
- *Use a figure to replace a table when you can.*
- *Be aware of norms about *s.*

Figures:

- *Report the scale of axes, and label them.*
- *Use as much "space" as you need, but no more.*
- *Use color sparingly.*