

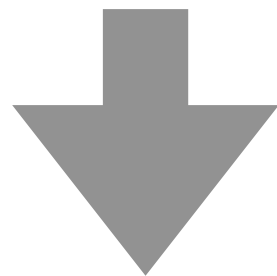
The Normal Model

An Approximate Histogram for Many Data Sets

Data Reduction

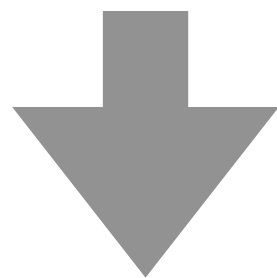
huge list of numbers

-0.39, -0.6, -0.17, -0.45, -0.19, -0.29, -0.49, -0.35, -0.51, -0.19, -0.44, -0.26, -0.38, -0.49, -0.31, -0.58, -0.43, -0.23, -0.37, -0.29, -0.38, -0.24, -0.47, -0.45, -0.4, -0.18, -0.53, -0.3, -0.37, -0.36, -0.52, -0.3, -0.28, -0.56, -0.3, -0.58, -0.39, -0.37, -0.19, -0.35, -0.66, -0.38, -0.47, -0.4, -0.34, -0.68, -0.4, -0.45, -0.28, -0.28, -0.38, -0.43, -0.34, -0.24, -0.36, -0.42, -0.4, -0.33, -0.42, -0.44, -0.4, -0.48, -0.4, -0.22, -0.33, -0.26, -0.16, -0.38, NA, -0.56, -0.3, -0.47, -0.27, -0.36, -0.35, -0.28, -0.2, -0.41, -0.23, -0.32, -0.18, -0.24, -0.46, -0.32, -0.57, -0.49, -0.54, -0.62, -0.48, -0.49, -0.6, -0.43, -0.27, -0.24, -0.26, -0.38, -0.49, -0.34, -0.42, -0.11, -0.38, -0.45, -0.3, -0.47, -0.34, -0.54, -0.26, -0.42, -0.35, -0.33, -0.48, -0.28, -0.31, -0.47, -0.29, -0.54, -0.39, -0.41, -0.45, -0.48, -0.31, -0.34, -0.26, -0.42, -0.31, -0.29, -0.26, -0.52, -0.45, -0.24, -0.3, -0.35, -0.28, -0.26, -0.28, -0.4, -0.4, -0.43, -0.62, -0.15, -0.3, -0.27, -0.2, -0.37, -0.52, -0.33, -0.4, -0.6, -0.35, -0.61, -0.42, -0.38, -0.49, -0.24, -0.29, -0.62, -0.22, -0.36, -0.23, -0.24, -0.31, -0.41, -0.54, -0.51, -0.55, -0.39, -0.41, -0.34, -0.4, -0.47, -0.26, -0.38, -0.35, -0.4, -0.58, -0.43, -0.59, -0.43, -0.18, -0.39, -0.43, -0.36, -0.4, -0.46, -0.13, -0.47, -0.31, -0.49, -0.31, -0.33, -0.38, -0.39, -0.46, -0.16, -0.4, -0.44, -0.23, -0.4, -0.4, -0.31, -0.38, -0.44, -0.4, -0.39, -0.82, -0.4, -0.46, -0.48, -0.41, -0.31, -0.39, -0.38, -0.2, -0.17, -0.35, -0.45, -0.36, -0.31, -0.69, -0.33, -0.4, -0.28, -0.28, -0.53, -0.54, -0.26

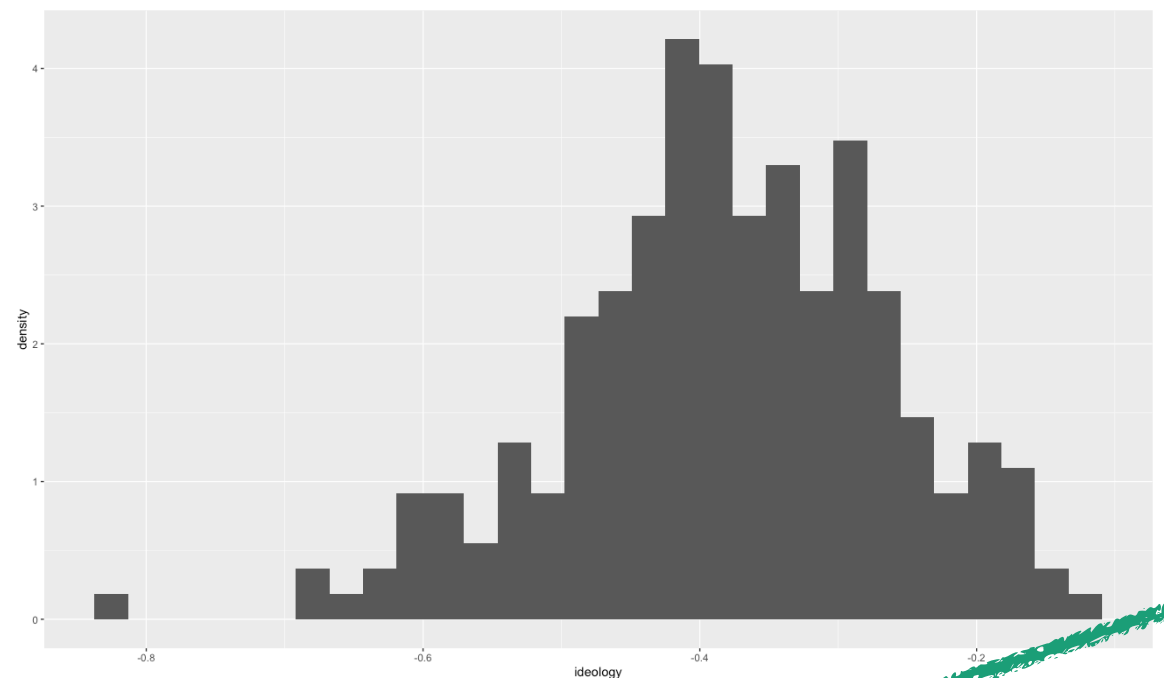


histograms

(bins and densities within those bins)



average and SD



avg = -0.38
SD = 0.12

How much have we lost?

Why are the average and SD
an effective summary?

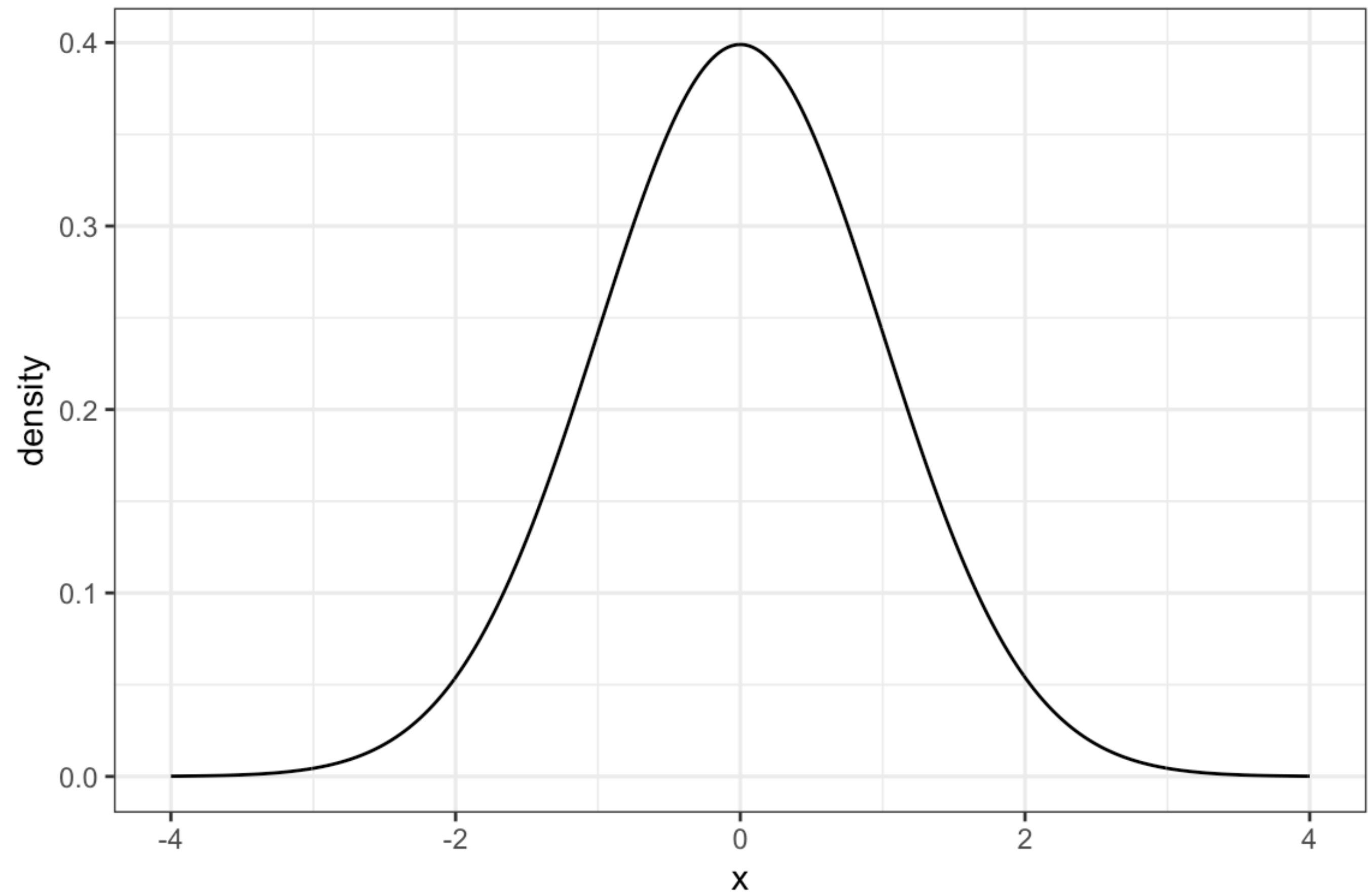
ch. 4

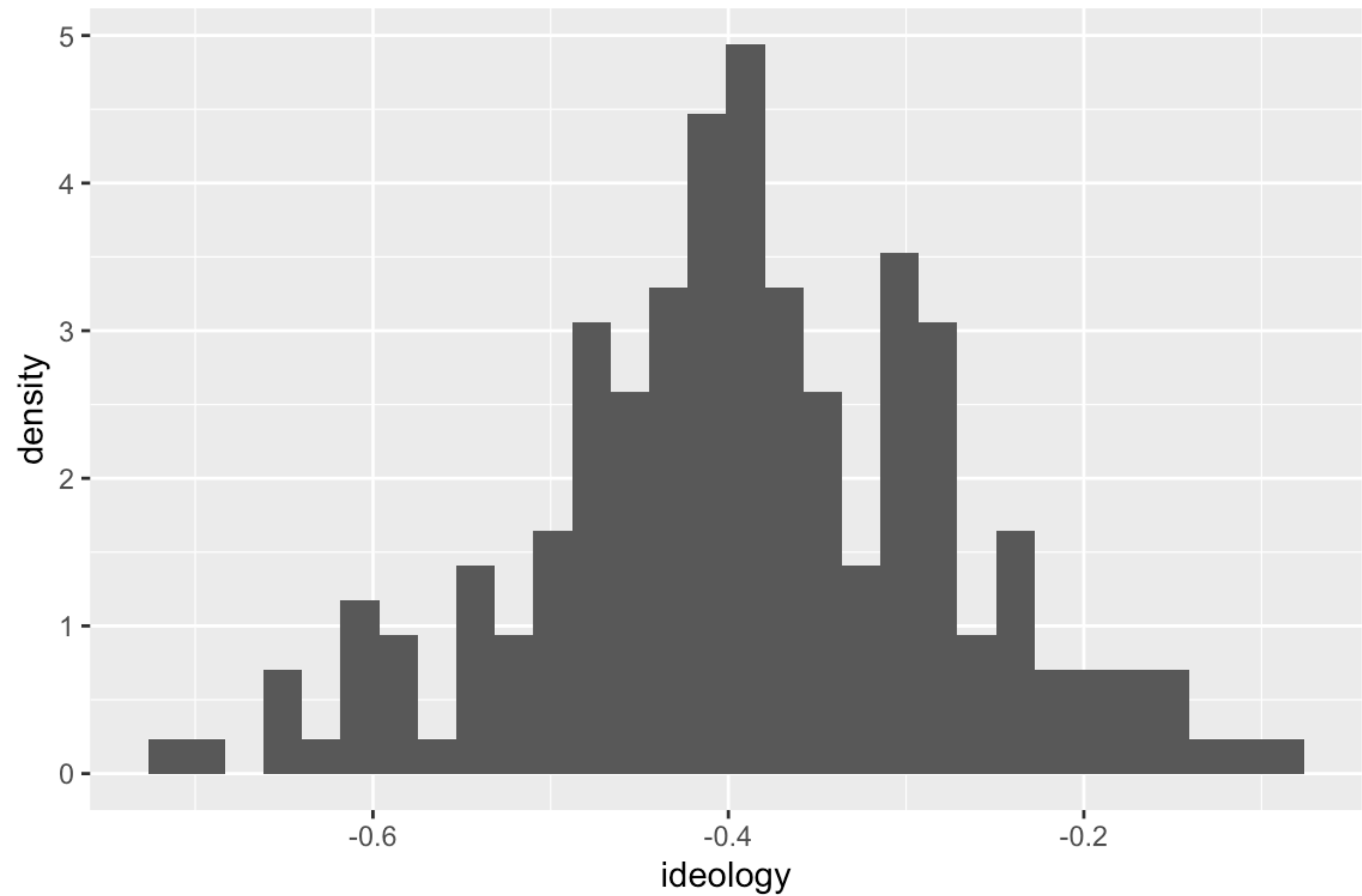
“The values are about _____, give or take _____ or so.”

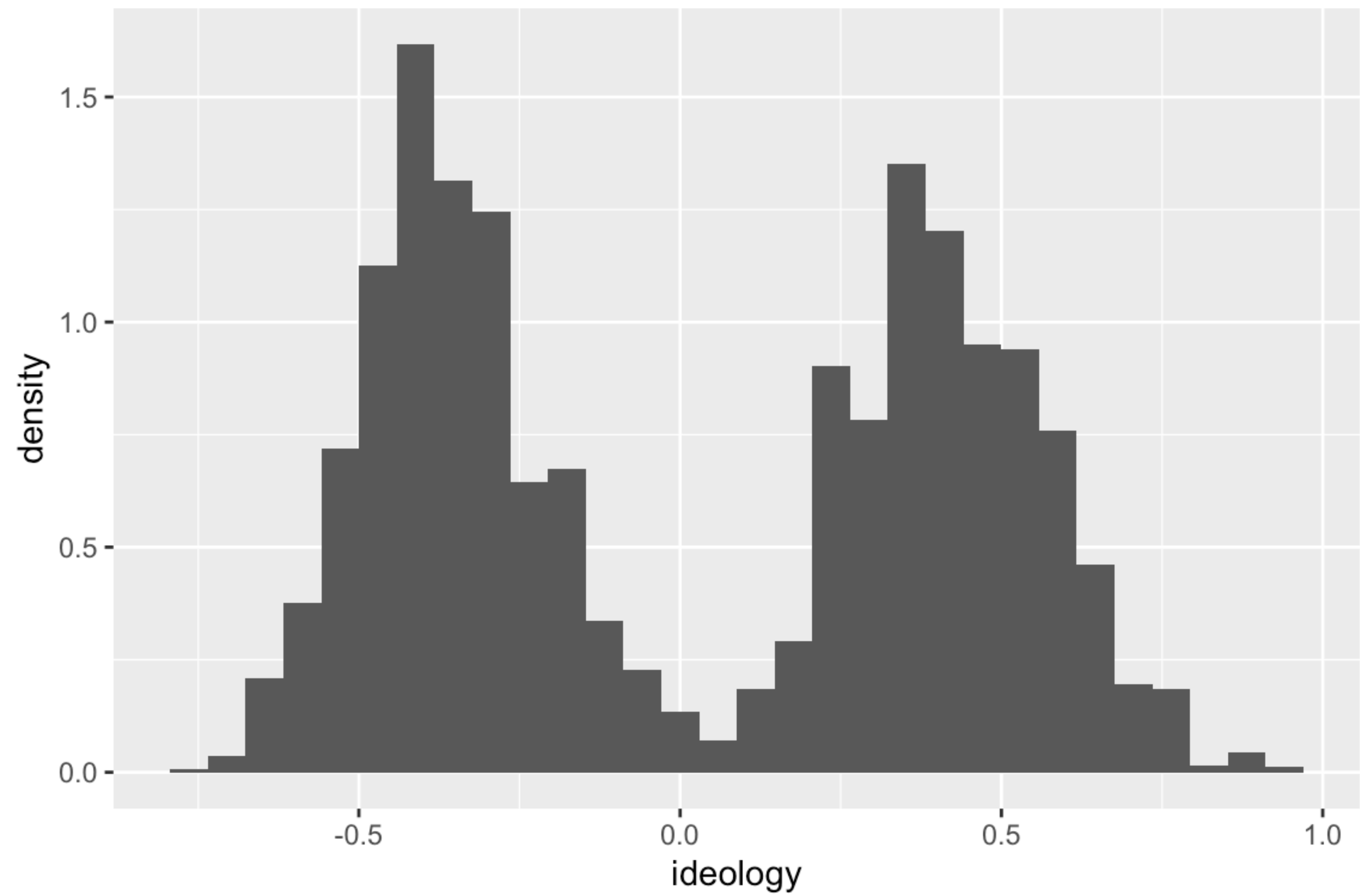
ch. 5

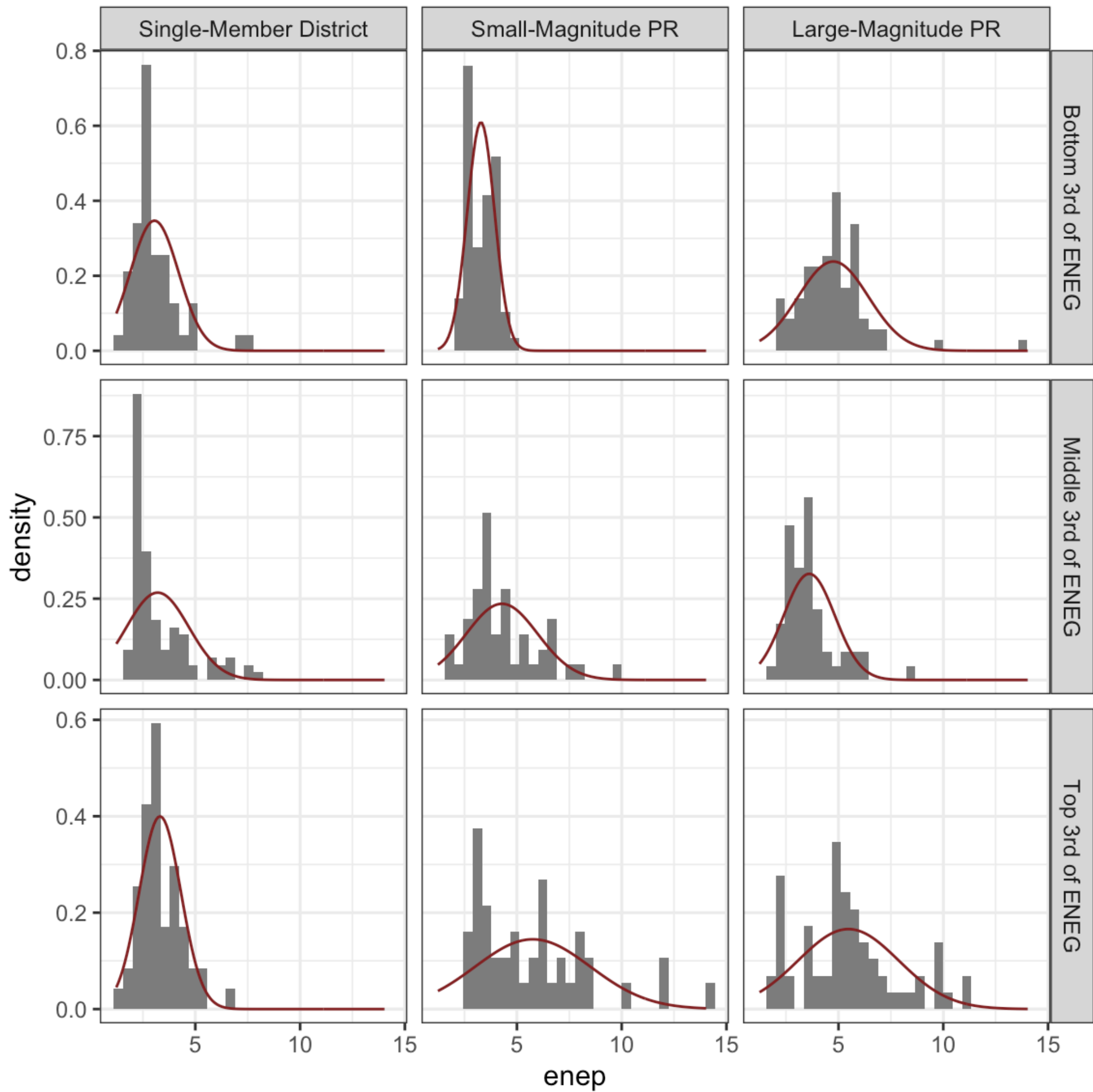
“This histogram follows the normal curve.”

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







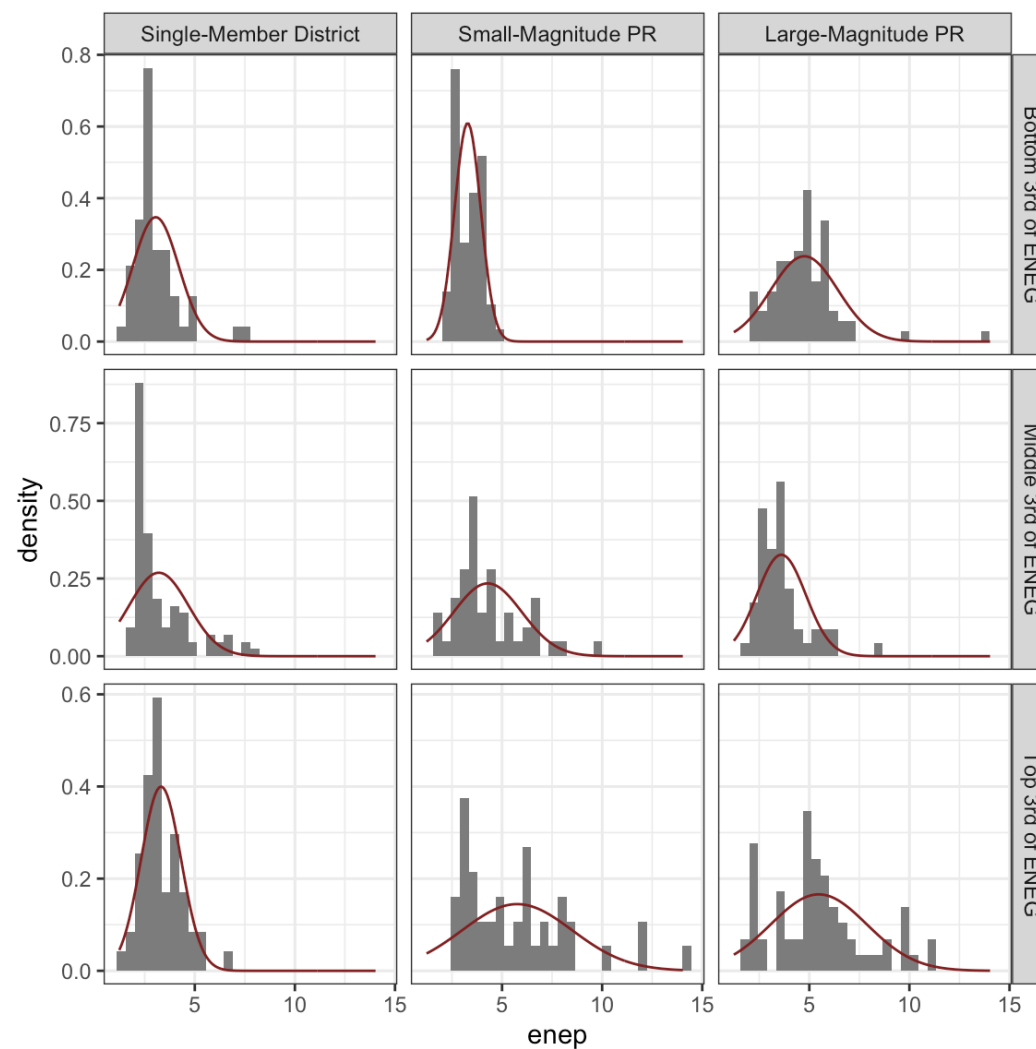


The Empirical Rule

If the variable roughly follows the normal curve (or is “bell-shaped”), then we have the following two rules:

1. About **68%** of the data (i.e., “most”) fall within **1 SD** of the average.
2. About **95%** of the data (i.e., “almost all”) fall within **2 SDs** of the average.

Electoral System	Social Heterogeneity	within 1 SD	within 2 SDs
Single-Member District	Bottom 3rd of ENEG	87%	96%
Single-Member District	Middle 3rd of ENEG	87%	93%
Single-Member District	Top 3rd of ENEG	74%	98%
Small-Magnitude PR	Bottom 3rd of ENEG	68%	97%
Small-Magnitude PR	Middle 3rd of ENEG	73%	96%
Small-Magnitude PR	Top 3rd of ENEG	76%	93%
Large-Magnitude PR	Bottom 3rd of ENEG	80%	98%
Large-Magnitude PR	Middle 3rd of ENEG	77%	96%
Large-Magnitude PR	Top 3rd of ENEG	65%	97%



The Normal Approximation

1. Draw a picture.
2. Convert to standard units.
3. Use the rules.

Using a Normal Approximation

To do a normal approximation, follow these steps:

1. **Draw a picture. This is really important. If you can draw the correct picture, it's really easy to find the correct approximation.**
 - A. Draw the normal curve.
 - B. Label the points of interest. I find it helpful to label the average as well.
 - C. Shade the area of interest.
2. **Convert the points of interest to z-scores or “standard units.” I like to add the z-scores in parentheses underneath the points of interest.**

$$z\text{-score} = \frac{\text{value} - \text{average}}{\text{SD}}$$

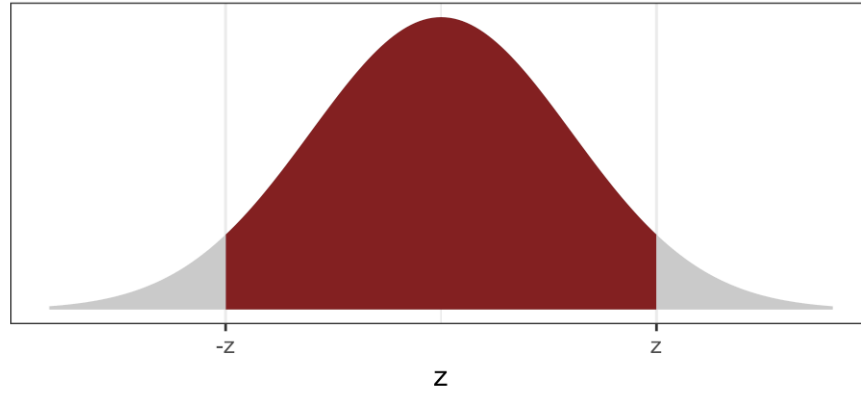
3. **Use the Rules of the Normal Curve.**

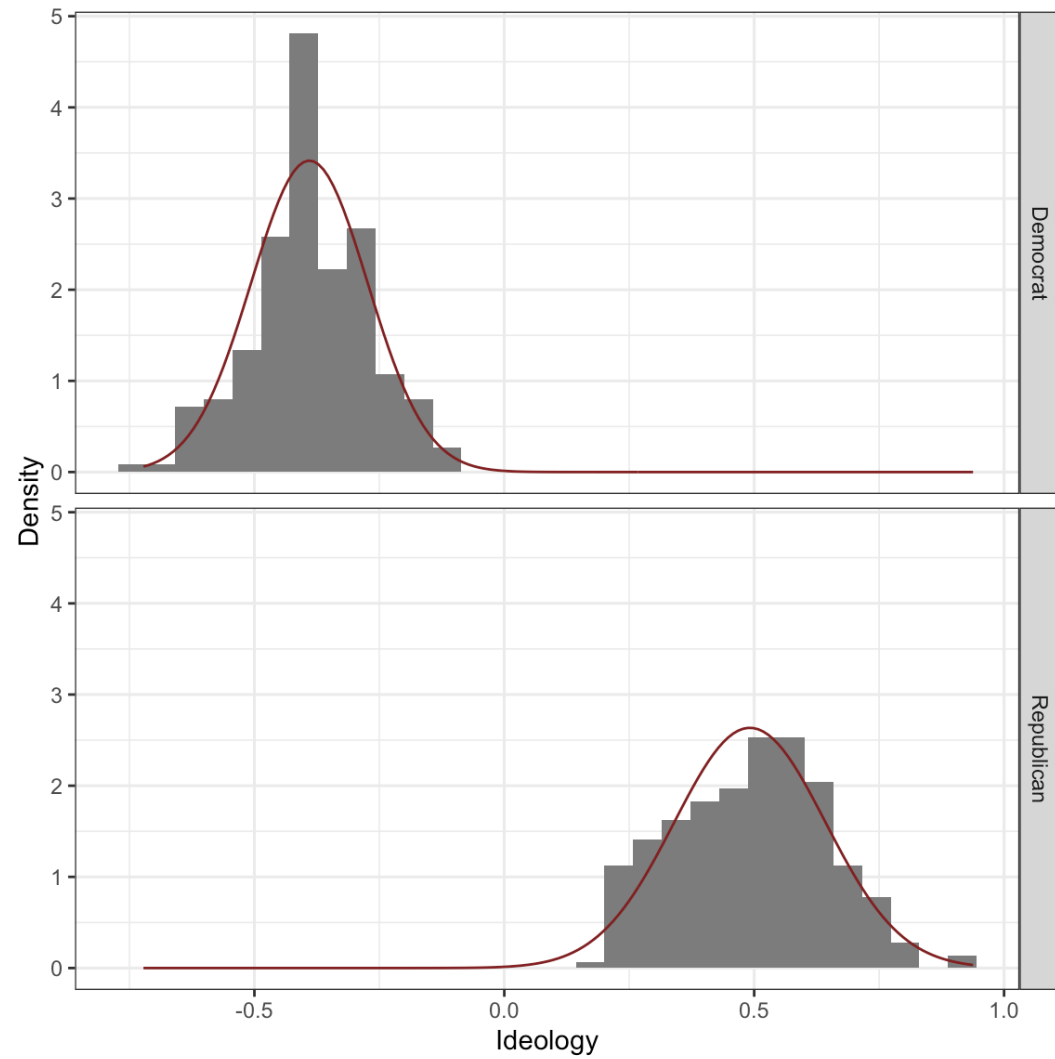
- A. The normal table gives the area between $-z$ and z . See FPP A-104 or the appendix in these notes. Usually start here.
- B. The area under the entire curve is 100%.
- C. The curve is symmetric, so that the area above a particular value z equals the area below the value $-z$.

A NORMAL TABLE

<i>z</i>	<i>Height</i>	<i>Area</i>	<i>z</i>	<i>Height</i>	<i>Area</i>	<i>z</i>	<i>Height</i>	<i>Area</i>
0.00	39.89	0	1.50	12.95	86.64	3.00	0.443	99.730
0.05	39.84	3.99	1.55	12.00	87.89	3.05	0.381	99.771
0.10	39.69	7.97	1.60	11.09	89.04	3.10	0.327	99.806
0.15	39.45	11.92	1.65	10.23	90.11	3.15	0.279	99.837
0.20	39.10	15.85	1.70	9.40	91.09	3.20	0.238	99.863
0.25	38.67	19.74	1.75	8.63	91.99	3.25	0.203	99.885
0.30	38.14	23.58	1.80	7.90	92.81	3.30	0.172	99.903
0.35	37.52	27.37	1.85	7.21	93.57	3.35	0.146	99.919
0.40	36.83	31.08	1.90	6.56	94.26	3.40	0.123	99.933
0.45	36.05	34.73	1.95	5.96	94.88	3.45	0.104	99.944
0.50	35.21	38.29	2.00	5.40	95.45	3.50	0.087	99.953
0.55	34.29	41.77	2.05	4.88	95.96	3.55	0.073	99.961
0.60	33.32	45.15	2.10	4.40	96.43	3.60	0.061	99.968
0.65	32.30	48.43	2.15	3.96	96.84	3.65	0.051	99.974
0.70	31.23	51.61	2.20	3.55	97.22	3.70	0.042	99.978
0.75	30.11	54.67	2.25	3.17	97.56	3.75	0.035	99.982
0.80	28.97	57.63	2.30	2.83	97.86	3.80	0.029	99.986
0.85	27.80	60.47	2.35	2.52	98.12	3.85	0.024	99.988
0.90	26.61	63.19	2.40	2.24	98.36	3.90	0.020	99.990
0.95	25.41	65.79	2.45	1.98	98.57	3.95	0.016	99.992
1.00	24.20	68.27	2.50	1.75	98.76	4.00	0.013	99.9937
1.05	22.99	70.63	2.55	1.54	98.92	4.05	0.011	99.9949
1.10	21.79	72.87	2.60	1.36	99.07	4.10	0.009	99.9959
1.15	20.59	74.99	2.65	1.19	99.20	4.15	0.007	99.9967
1.20	19.42	76.99	2.70	1.04	99.31	4.20	0.006	99.9973
1.25	18.26	78.87	2.75	0.91	99.40	4.25	0.005	99.9979
1.30	17.14	80.64	2.80	0.79	99.49	4.30	0.004	99.9983
1.35	16.04	82.30	2.85	0.69	99.56	4.35	0.003	99.9986
1.40	14.97	83.85	2.90	0.60	99.63	4.40	0.002	99.9989
1.45	13.94	85.29	2.95	0.51	99.68	4.45	0.002	99.9991

Area Between a Particular Value and Its Opposite



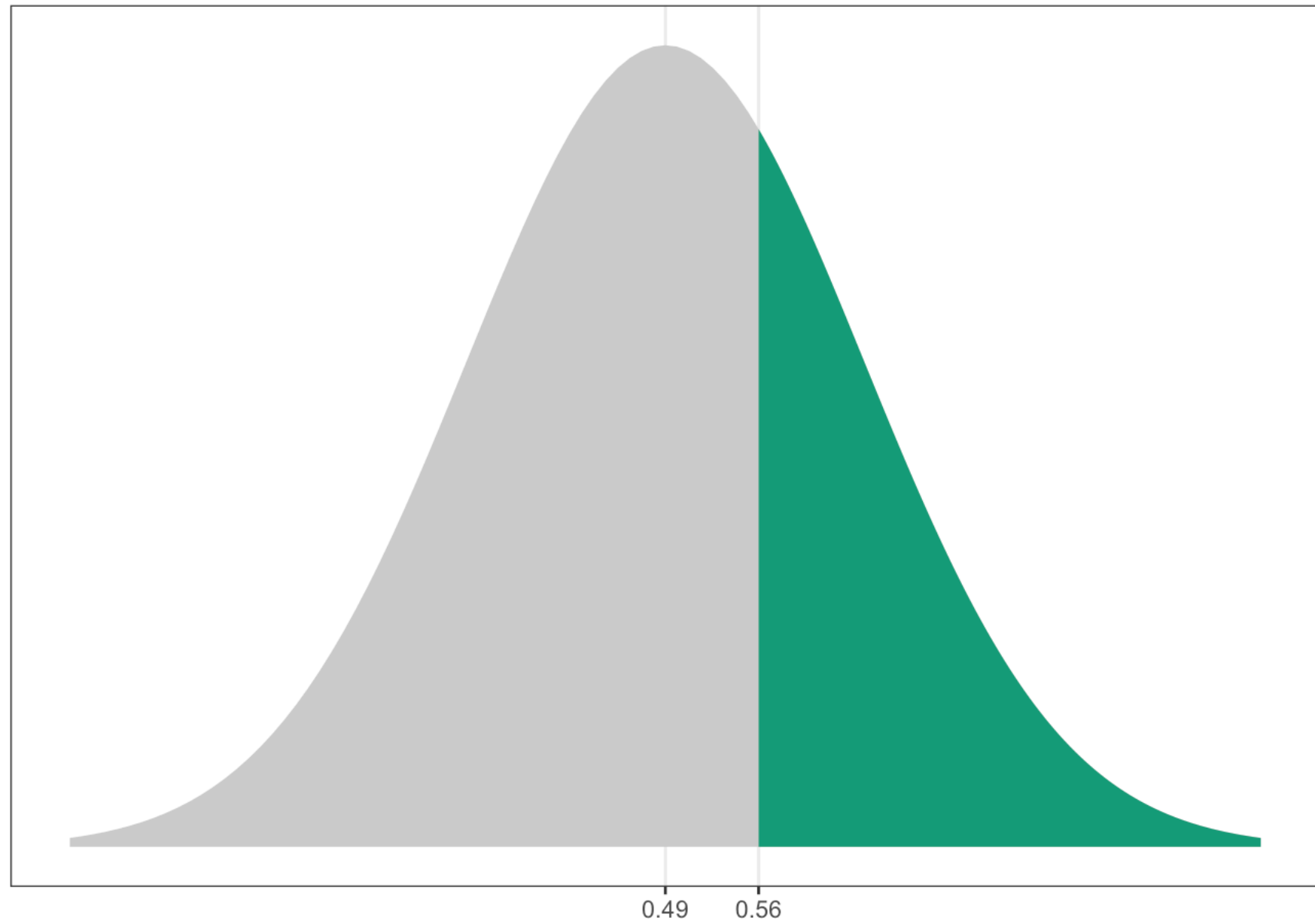


Name	Party	Position	Ideology Score	Inspect Histogram	Normal Approximation	Actual
RYAN, Paul D.	Republican	Speaker of the House	0.56			
MCCARTHY, Kevin	Republican	Majority Leader	0.46			
SCALISE, Steve	Republican	Majority Whip	0.56			
McMORRIS RODGERS, Cathy	Republican	Conference Chair	0.43			
PELOSI, Nancy	Democrat	Minority Leader	-0.49			
HOYER, Steny Hamilton	Democrat	Minority Whip	-0.38			
CLYBURN, James Enos	Democrat	Assistant Democratic Leader	-0.46			
LEWIS, John R.	Democrat	Senior Chief Deputy Minority Whip	-0.59			

Party	Average	SD
Democrat	-0.39	0.12
Republican	0.49	0.15

Step 1

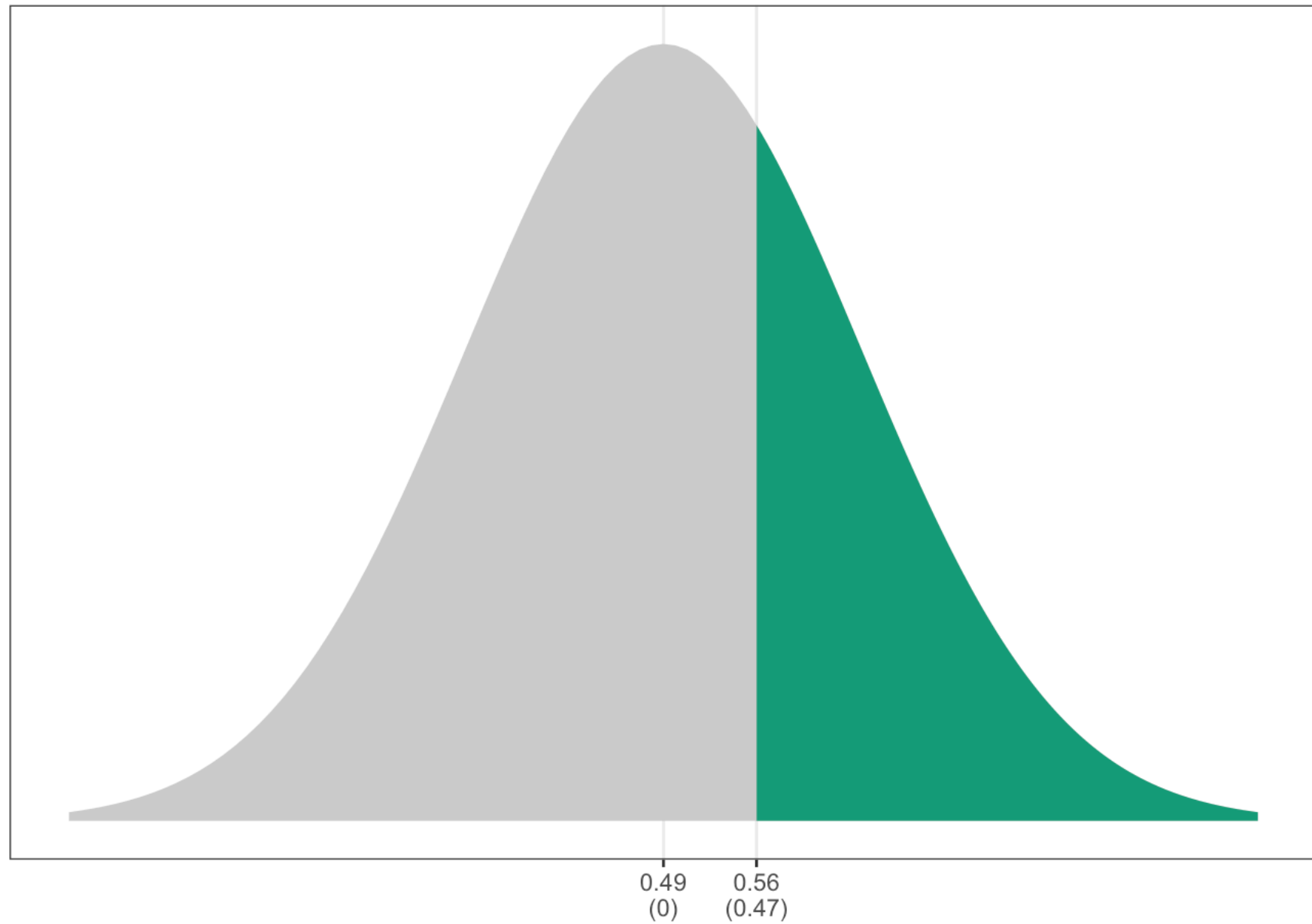
Draw a Picture



Step 2

Convert to SUs

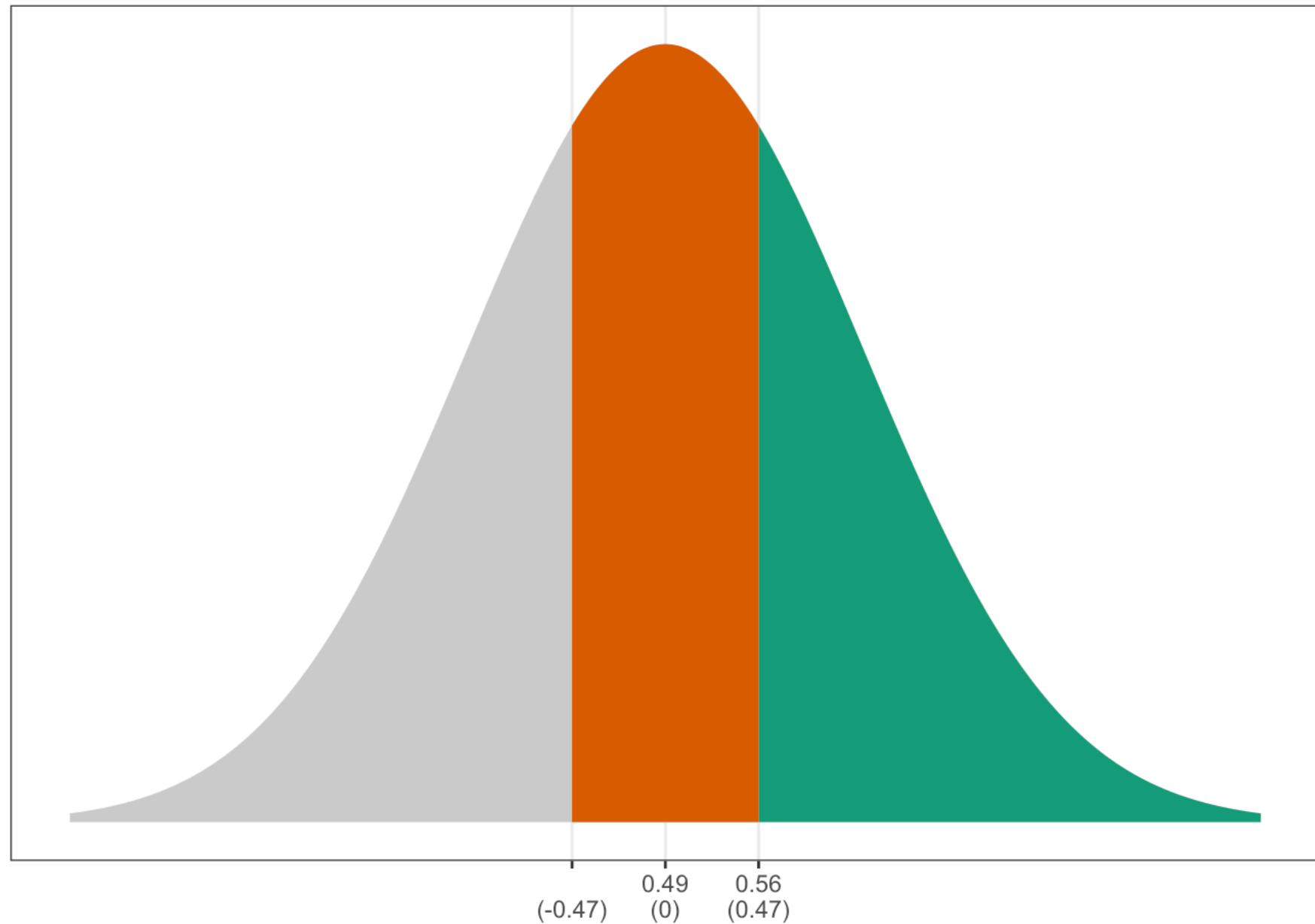
$$z\text{-score} = \frac{\text{value} - \text{average}}{\text{SD}}$$



Hint: $(0.56 - 0.49)/0.15 = 0.47$

Step 3

Use Rules!



There's 35% in the middle (**orange**).

Then there's $100\% - 35\% = 65\%$ left in both tails together (**green** plus **grey**).

Then there's $65\%/2 = 33\%$ in the right tail (**green**). That's the area we wanted.

Let's Try Nancy Pelosi