Inferential Statistics

- Inferential Statistics involves making inferences about a population based on a sample of that population

Populations and Samples

- Population: everything that qualifies as a potential case
  - If you are studying US college students, population is ALL US college students
- Sample: those cases you collect data on
  - 50 UCI student participants
  - Use samples to make inferences about the population

Research Design

- When designing new research projects, defining your population, and constructing a valid sampling design, is crucial

Populations and Samples

- We use different symbols to distinguish population statistics from sample statistics
- Mean:
    - Sample: $\bar{x}$
    - Population: μ
- Standard Deviation
    - Sample: s
    - Population: σ

Populations and Samples

- The equation of sample and population means are the same

Population

$$\mu = \frac{\sum x_i}{N}$$

Sample

$$\bar{x} = \frac{\sum x_i}{n}$$

- But the equations for standard deviation are slightly different

Population

$$\sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}}$$

Sample

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

Populations and Samples

- In general, we call population statistics parameters.
- Sample statistics are simply called statistics
- Parameters have one true value, but we rarely know what it is (can’t collect data)
- Statistics will have different values depending on the sample you take

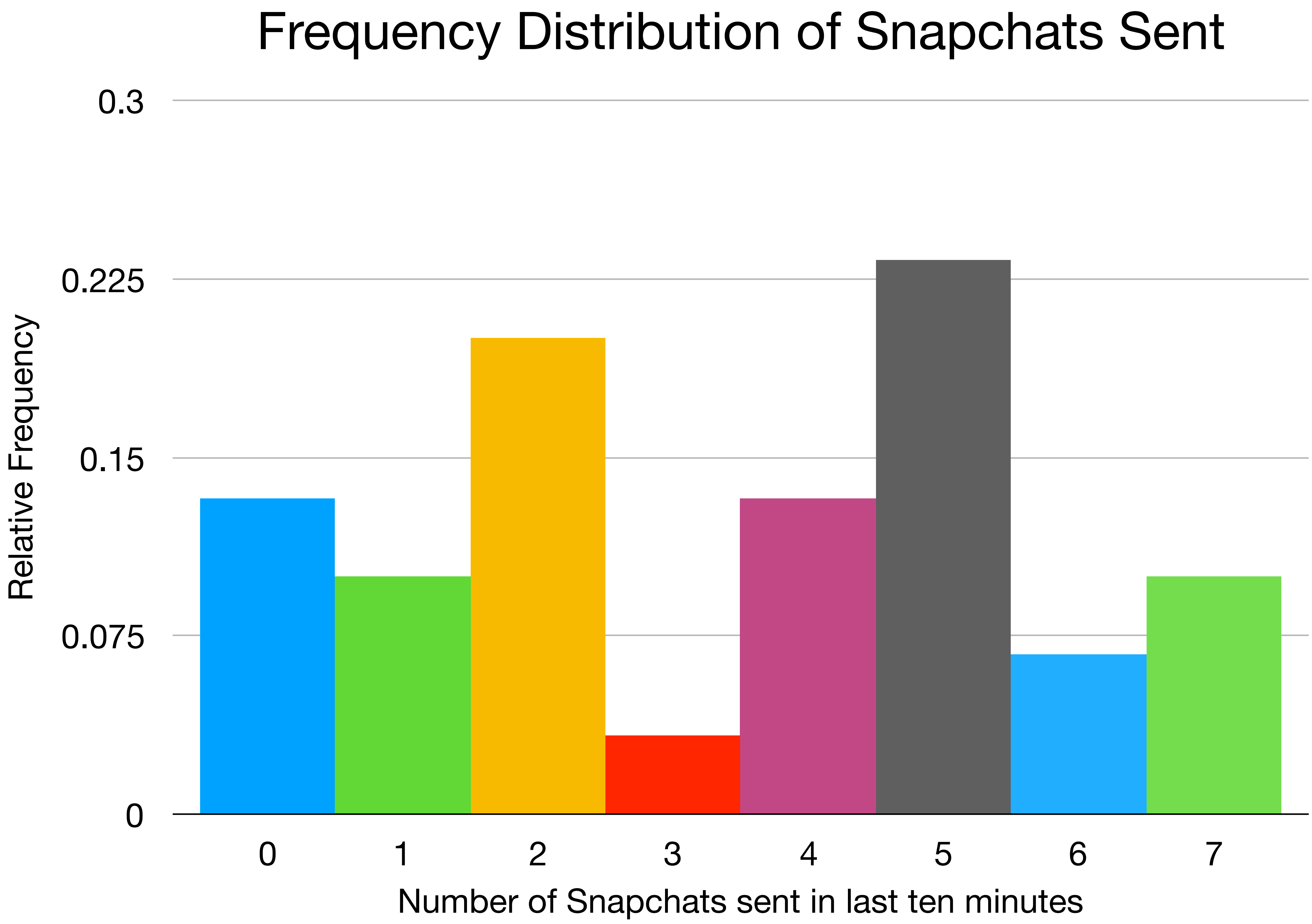
Distributions

- When doing inferential statistics, we can consider three things:
  - How individual cases are distributed within a population (don’t usually know this)
  - How individual cases are distributed within our sample (not as useful)
  - How multiple samples of a population are distributed within that population (most useful)

Distributions

Snapchats Sent	Frequency	Relative Frequency	Cumulative Frequency	Cumulative Relative Frequency
0	4	0.133	4	0.133
1	3	0.100	7	0.233
2	6	0.200	13	0.433
3	1	0.033	14	0.467
4	4	0.133	22	0.600
5	7	0.233	26	0.833
6	2	0.067	28	0.900
7	3	0.100	30	1.000
	30	1.000		

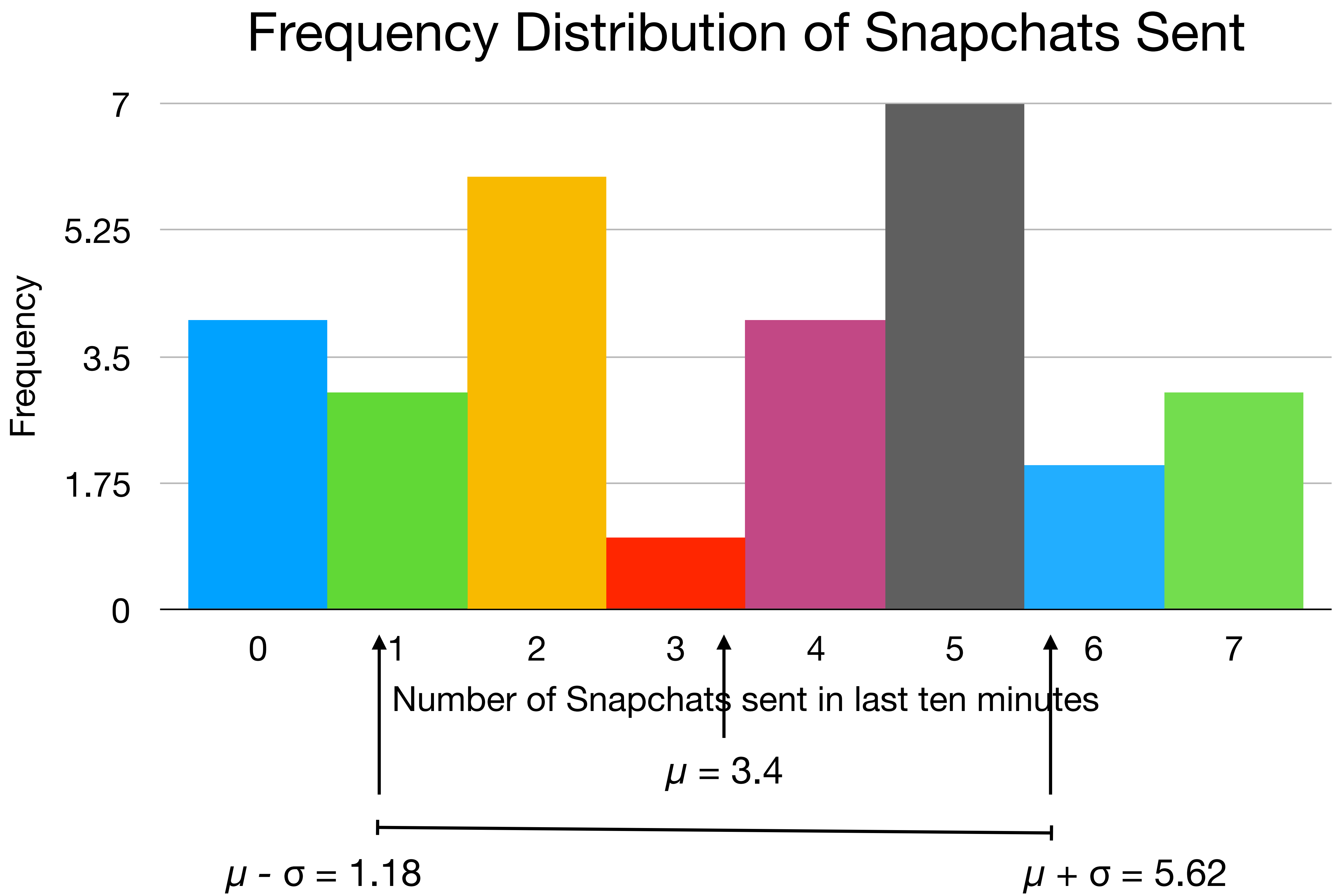
Population Parameters



Population Parameters

Snapchats Sent (x)	Frequency	Freq*x	x-μ	(x-μ)^2	Freq*(x-μ)^2
0	4	0	-3.4	11.56	46.24
1	3	3	-2.4	5.76	17.28
2	6	12	-1.4	1.96	11.76
3	1	3	-0.4	0.16	0.16
4	4	16	0.6	0.36	1.44
5	7	35	1.6	2.56	17.92
6	2	12	2.6	6.76	13.52
7	3	21	3.6	12.96	38.88
N →	30	102		sum of squares →	147.20
	$\mu = \frac{\sum x_i \times \text{freq}(x_i)}{N}$ →	3.4		$\sigma^2 = \frac{\sum (x_i - \mu)^2 \times \text{freq}(x_i)}{N}$ →	4.91
				$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2 \times \text{freq}(x_i)}{N}}$ →	2.22

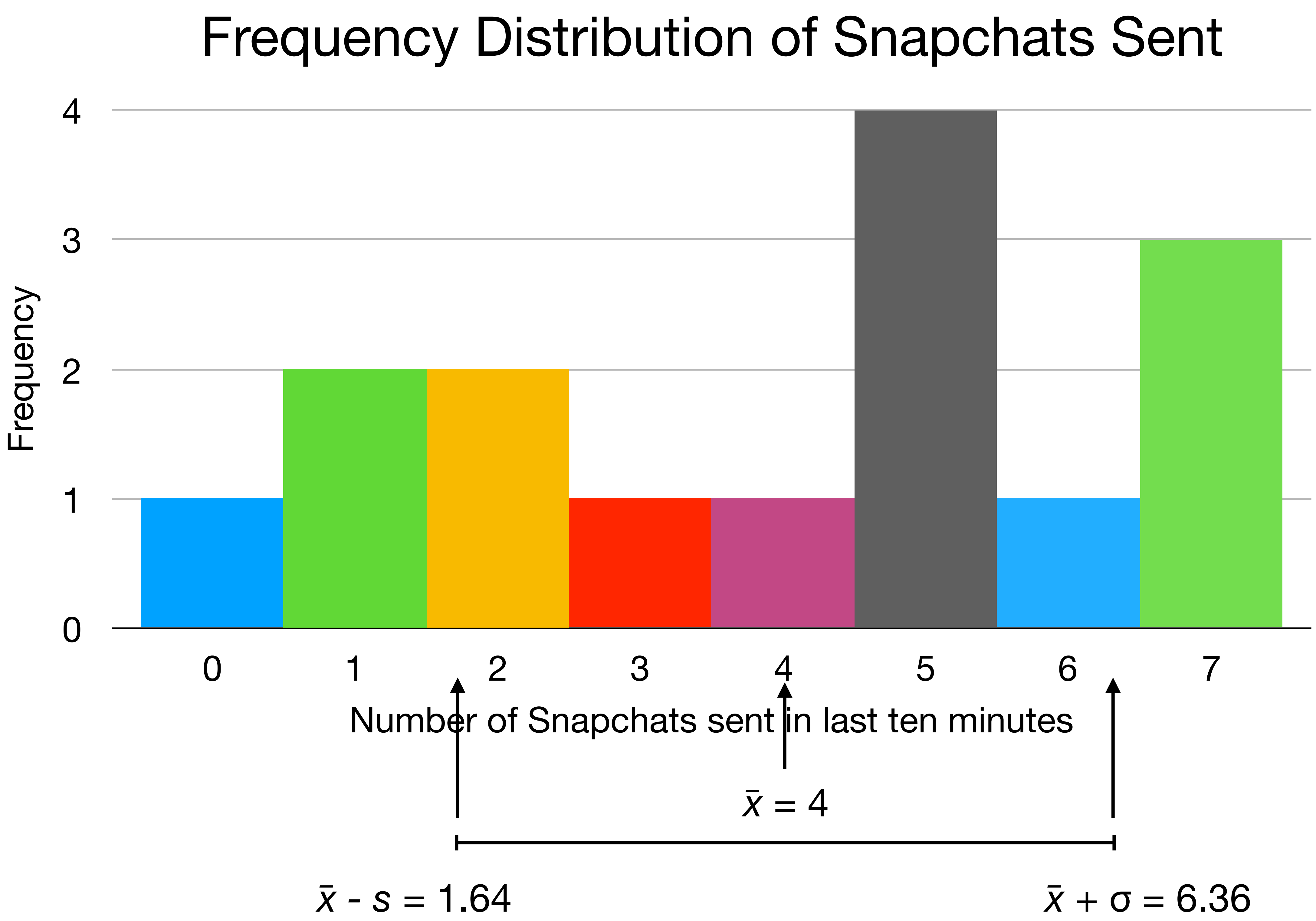
Population Parameters



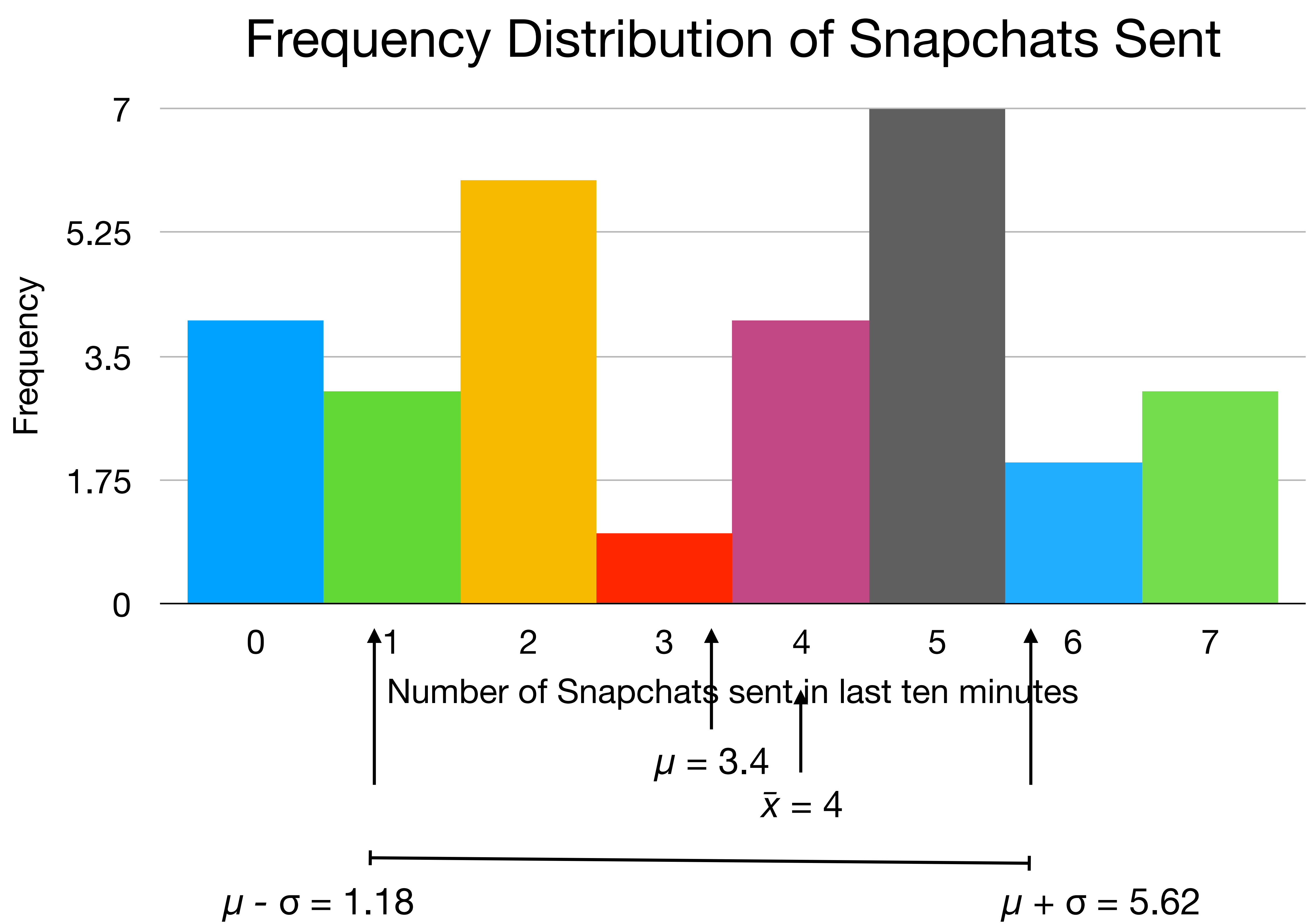
Sample Statistics

Snapchats Sent (x)	Frequency	Freq*x	x-$\bar{x}$	(x-$\bar{x}$)^2	Freq*(x-$\bar{x}$)^2
0	1	0	-4.0	16.00	16.00
1	2	2	-3.0	9.00	18.00
2	2	4	-2.0	4.00	8.00
3	1	3	-1.0	1.00	1.00
4	1	4	0.0	0.00	0.00
5	4	20	1.0	1.00	4.00
6	1	6	2.0	4.00	4.00
7	3	21	3.0	9.00	27.00
n →	15	60		sum of squares →	78.00
$\bar{x} = \frac{\sum x_i \times \text{freq}(x_i)}{n}$ →		4		$s^2 = \frac{\sum (x_i - \bar{x})^2 \times \text{freq}(x_i)}{n-1}$ →	5.57
				$s = \sqrt{\frac{\sum (x_i - \bar{x})^2 \times \text{freq}(x_i)}{n-1}}$ →	2.36

Sample Statistics



Population Parameters



Distributions

- What happens when we take several samples from a population?

Sampling Distribution

- Distribution of means from multiple samples within a single population
- We can apply the same math/logic for calculating mean and standard deviation that we did before; we are just doing it for multiple samples rather than multiple cases

Sampling Distribution

- Let's say you are interested in a research question in which the population is all undergrads at UCI (~22,000 students)
- Instead of interviewing all 22,000, you interview a sample of 100 students and ask about GPA

Sampling Distribution

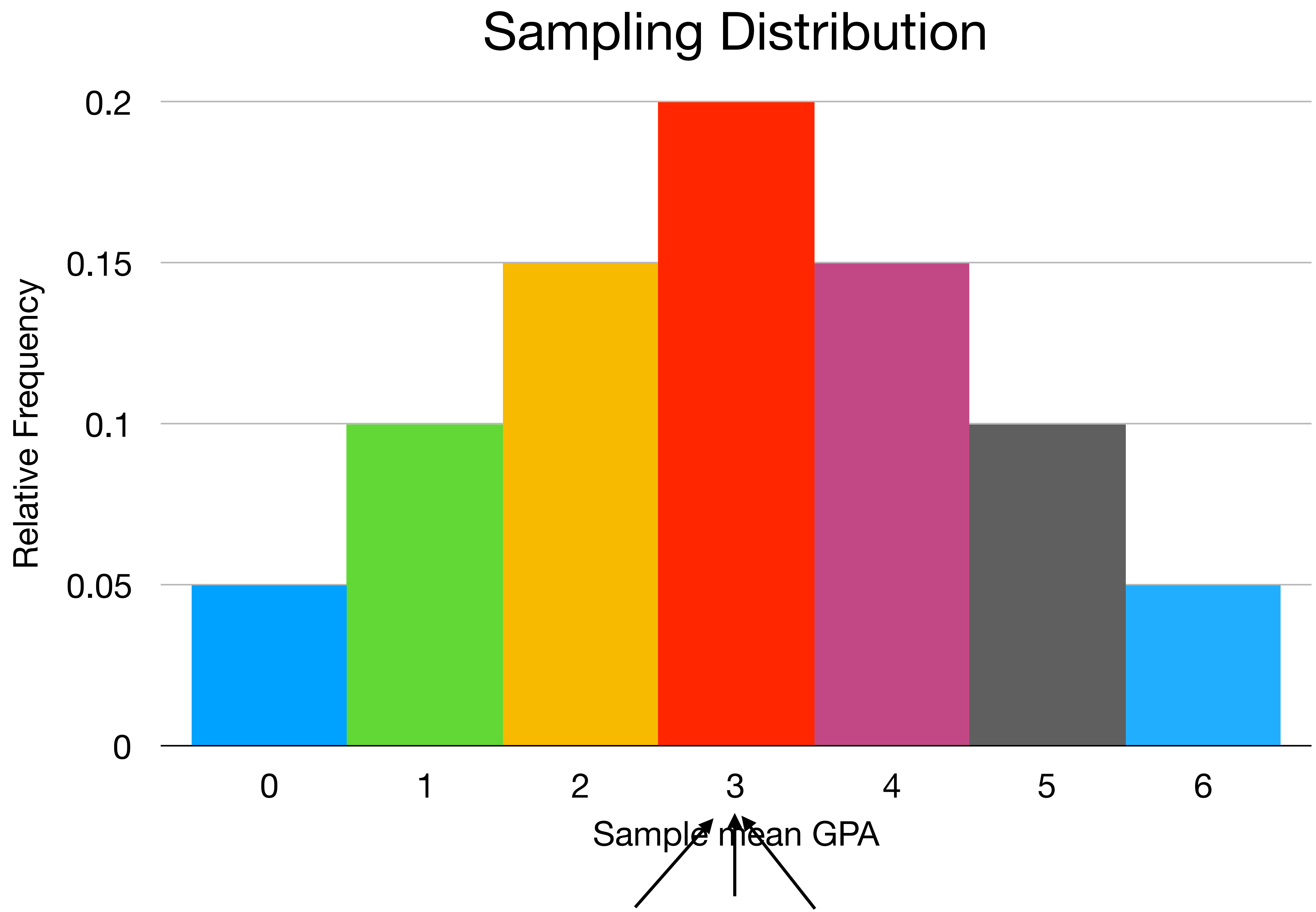
- Say that according to UCI’s website, the population mean GPA in Fall 2016 was 3.00
- You ask 100 students one day and your sample mean GPA was 2.93
- You ask 100 students the next day and your sample mean GPA was 3.02

Sampling Distribution

- The sample mean and standard deviation of your sample will depend on who you interviewed
- Samples have natural variability, and the sample statistics will reflect that variability.

Sampling Distribution

- There are lots and lots (1.5×10^{276}) of possible samples of 100 students from 22,000 undergrads
- The sample statistics from all these samples produce a distribution of sample statistics
- This is the Sampling Distribution



$\bar{x}_1 = 2.93 \quad \mu = 3.00 \quad \bar{x}_2 = 3.02$

Sampling Distribution

- The mean of the sampling distribution is calculated the same way as our previous means: by adding up all the sample means and dividing by the number of samples

the sample means

mean of the sampling distribution

$$\mu_{\bar{x}} = \frac{\sum \bar{x}_i}{N}$$

the number of samples

Standard Error

- of the mean of the sampling distribution

standard error

sample
standard deviation

$$\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}$$

sample size

Sampling Distribution

- In order to perform the statistics we are familiar with, our sampling distributions must be (or are assumed to be) normal
- Two theorems tell us that large sampling distributions are normal
- Later, we’ll talk about sampling distributions that aren’t normal (the t-distribution)

Law of Large Numbers

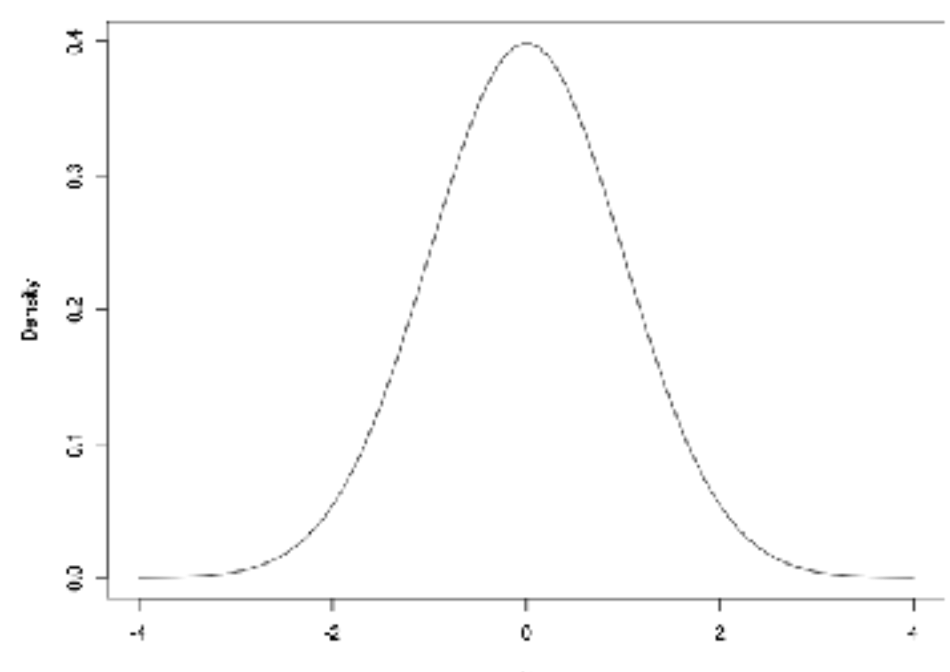
- As *n* (the sample size) increases, the sample mean will approach the population mean

$$\mu_{\bar{X}} = \mu$$

Central Limit Theorem

- If you take samples of size *n* from a population and compute their means, these means will be normally distributed for large *n*

- This distribution will be centered around mean μ with standard error $\sigma/\sqrt{n}$



Sampling Distribution

- Population: all UCI undergrads
- Sample: 30 undergrads chosen randomly when walking across campus
- Sampling distribution: mean GPAs for each sample of 30 that we took

Normal Distribution

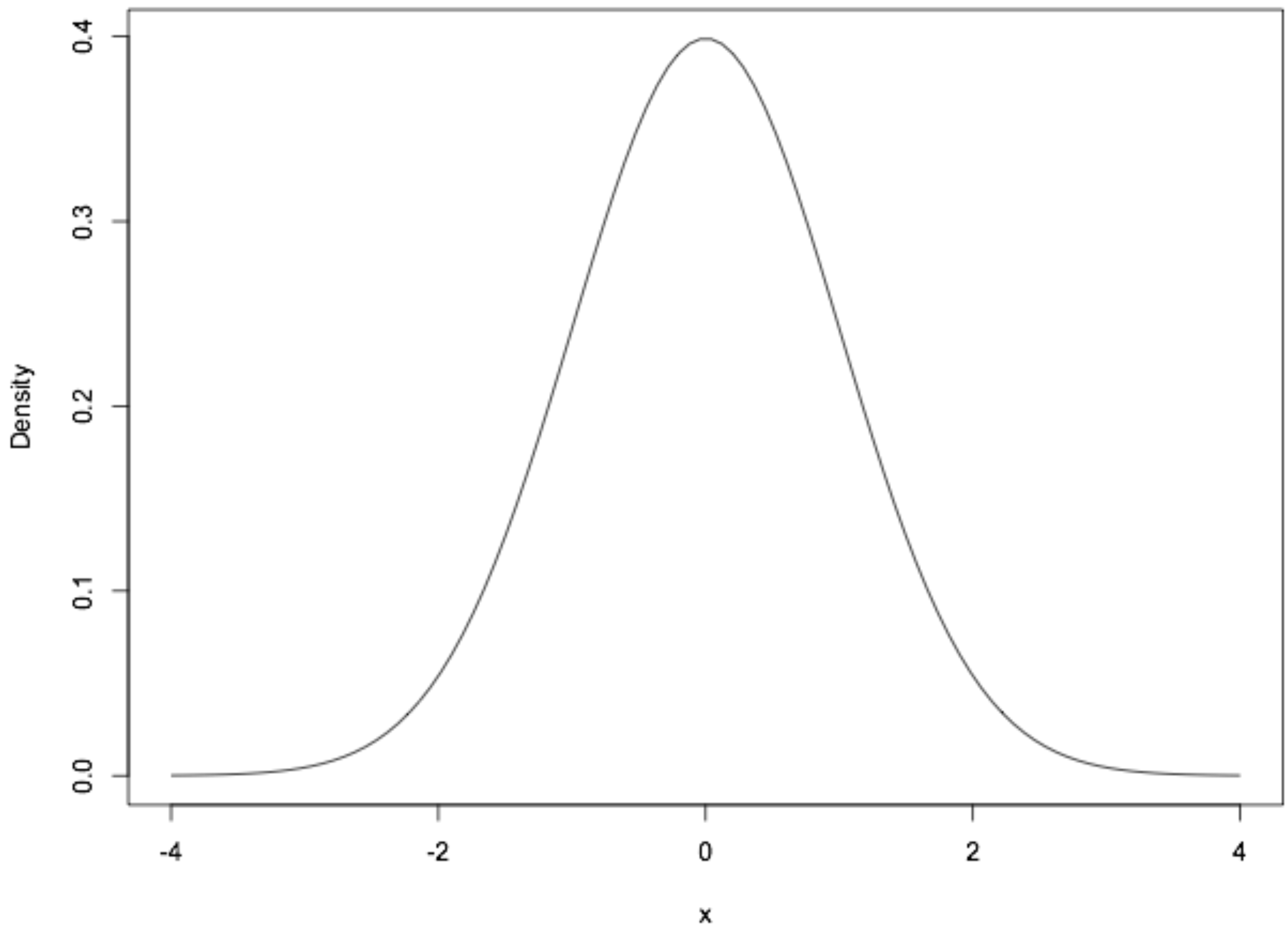
- Given its utility, the normal (bell-shaped) distribution is the most common distribution in social science

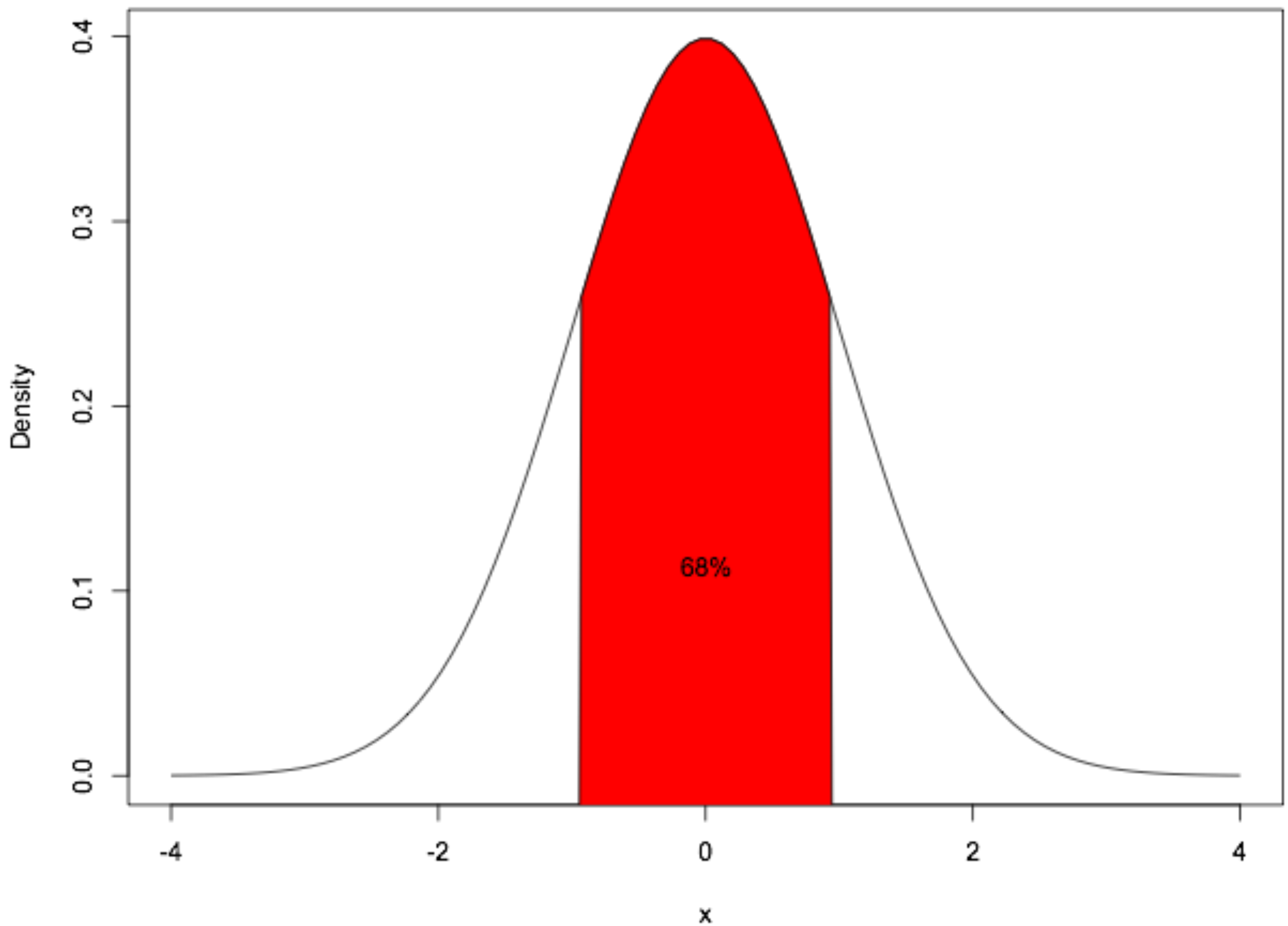
- Standard deviation (or error) tracks how far individual cases (or samples) are from the center of a distribution

- For a normal distribution, a set proportion of observations fall within certain known deviations from the mean

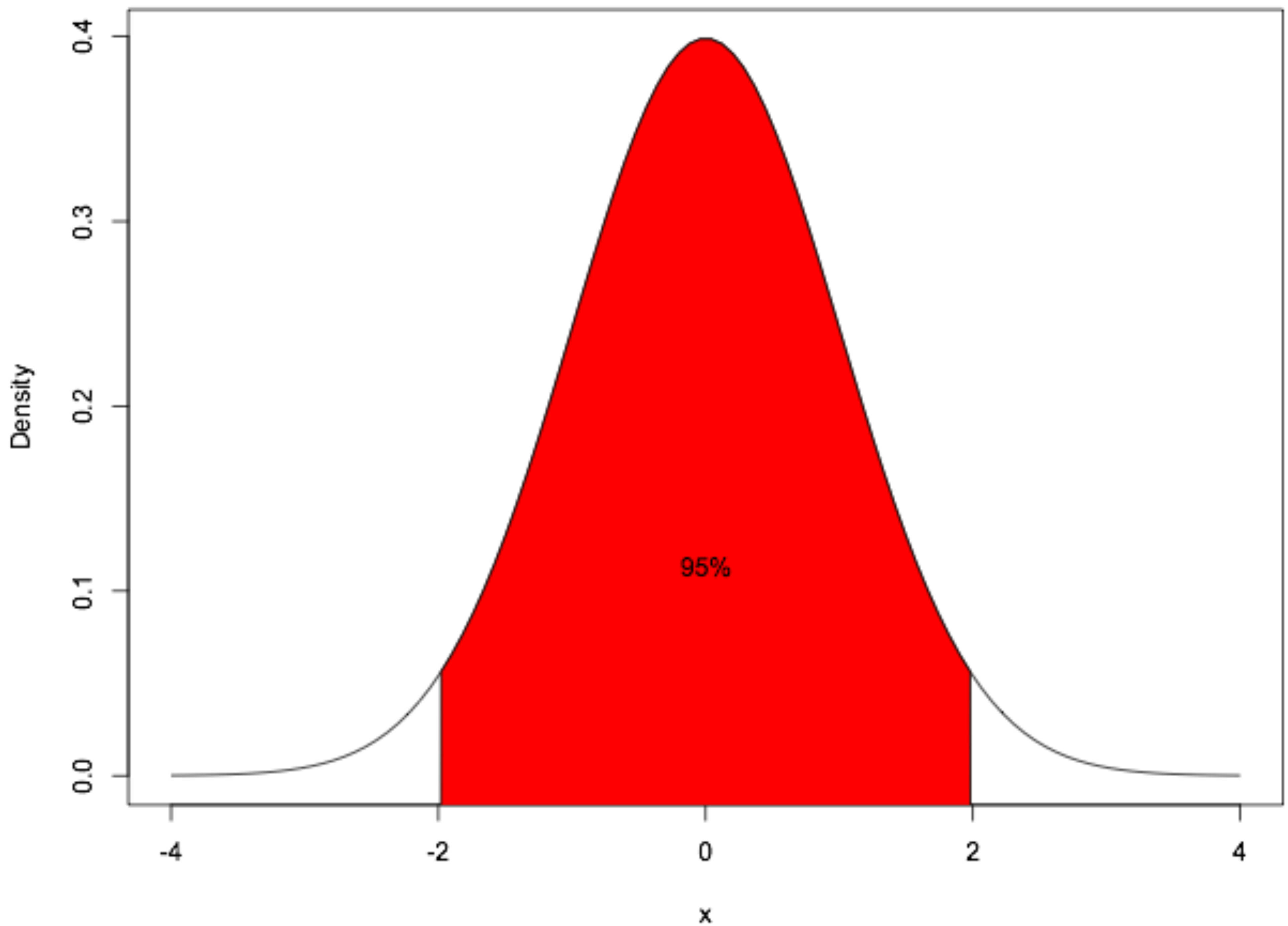
Normal Distribution

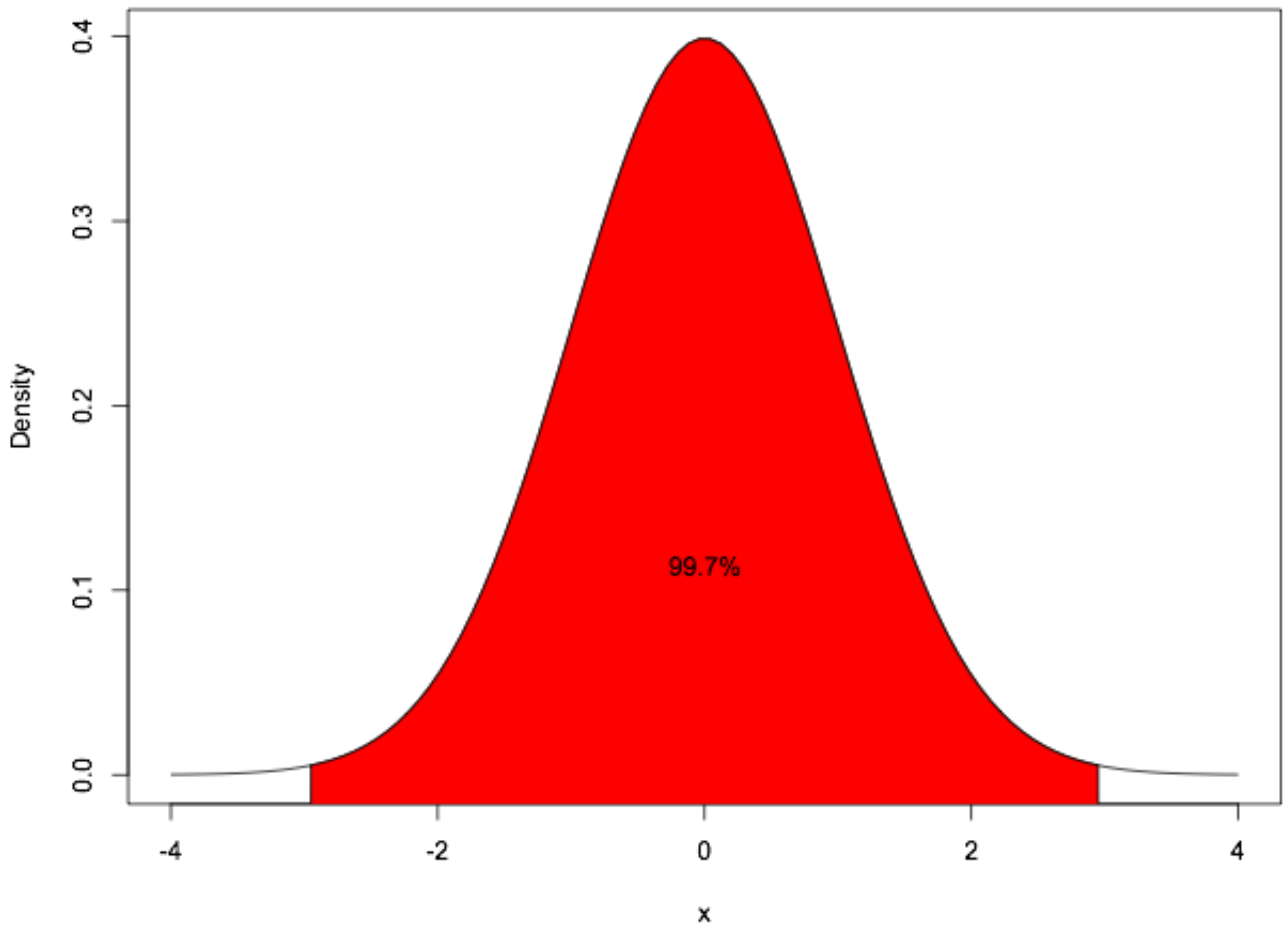
- Properties:
  - Mean and median are the same (the distribution is symmetrical)
  - 68.2% of observations are within 1 standard deviation of the mean
  - 95.4% of observations are within 2 standard deviations of the mean
  - 99.7% of observations are within 3 standard deviations of the mean


0.0
0.1
0.2
0.3
0.4
Density
-4
-2
0
2
4
x



Density
0.0
0.1
0.2
0.3
0.4
68%
-4
-2
0
2
4
x


0.4
0.3
0.2
0.1
0.0
Density
95%
-4
-2
0
2
4
x



99.7%
Density
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
4
x

Normal Distribution

- This information is useful for determining where a single individual (or sample) falls within its distribution

- Since the normal distribution will be used for a wide range of variables with different values, it is also useful to transform these values into a standard metric

1 in = 2.54 cm $\longrightarrow$ **7 in = 17.78 cm**

Z scores

Z scores

- Universal system of comparison
- Represent standardized values along a normal curve
- The ratio of an observation’s deviance score to the average deviance score (standard deviation)
- A normal distribution with a mean of 0 and standard deviation of 1 is called a standard normal distribution

Z scores

$$Z = \frac{\text{individual value} - \text{mean value}}{\text{standard deviation}}$$

(unitless)

Z scores

individual observation

sample mean

$$Z = \frac{X_i - \bar{X}}{s}$$

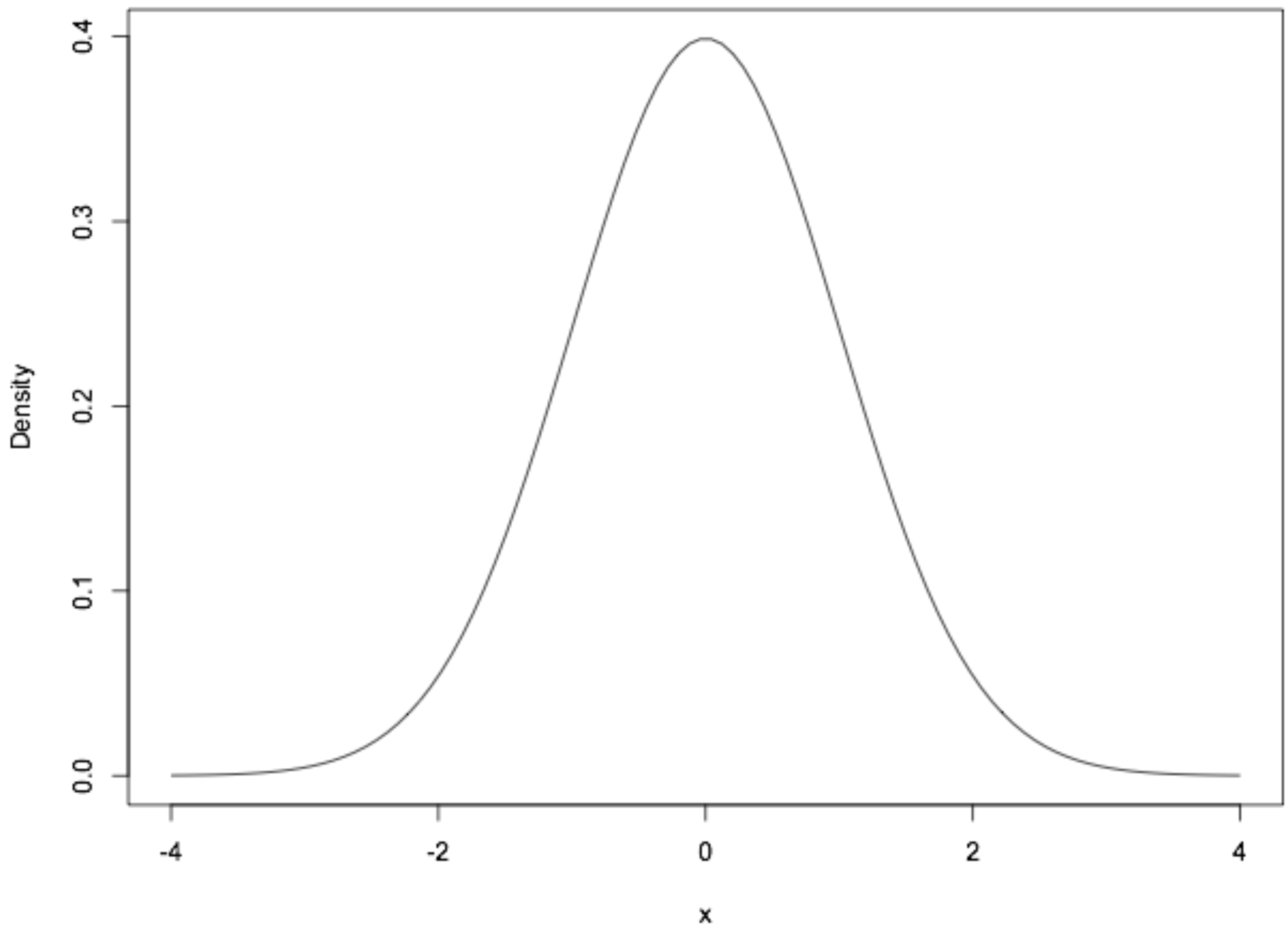
sample standard deviation

Practice!

Given a normally distributed set of observations with mean 4.6 and standard deviation 2.1, what is the Z score for a single observation with value 3? (How far away is it from the mean?)

$$\bar{X} = 4.6 \qquad s = 2.1 \qquad X_i = 3$$

$$Z = \frac{X_i - \bar{X}}{s} = \frac{3 - 4.6}{2.1} = \frac{-1.6}{2.1} = -0.76$$

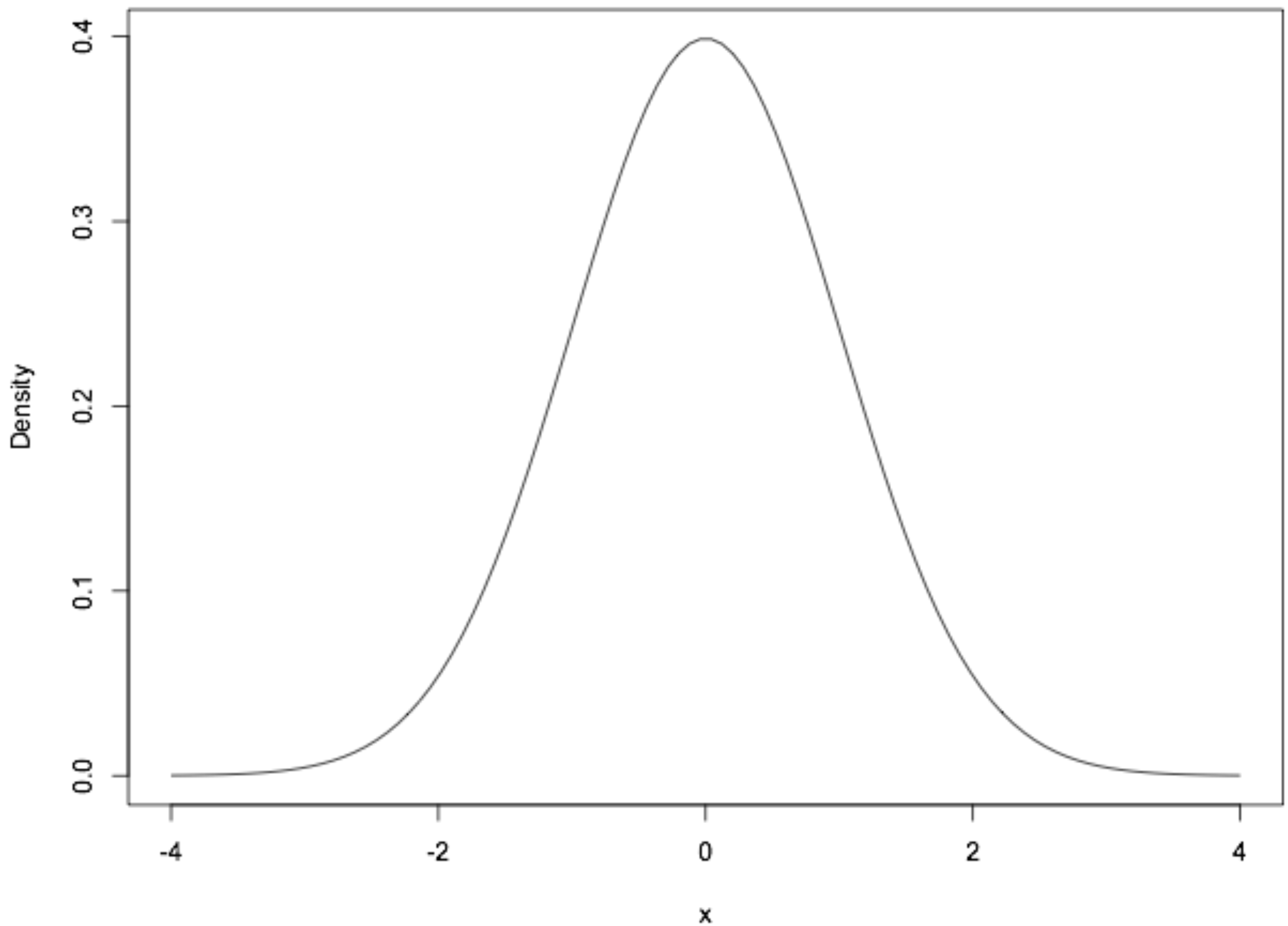

0.4
0.3
0.2
0.1
0.0
Density
-4
-2
0
2
4
x

Practice!

Given a normally distributed set of observations with mean 45 and standard deviation 10, what is the z score for a single observation of 62? (How far away is this observation from the mean?)

$$\bar{X} = 45 \qquad s = 10 \qquad X_i = 62$$

$$Z = \frac{X_i - \bar{X}}{s} = \frac{62 - 45}{10} = \frac{17}{10} = 1.7$$



0.0
0.1
0.2
0.3
0.4
Density
-4
-2
0
2
4
x

Z scores and Probability

- From our data and standardization (calculating the Z score) we know a little more about where a given sample mean falls within a distribution of sample means

- The question social scientists then ask is, how likely was this sample mean given what we know about the population?

- From answers to these questions we can make more concrete, precise statements about our data

Probability

- Probabilities represent the chances of picking a certain outcome (event) given what we know about all the possible outcomes (event space)
- I have 10 M&Ms. 2 are blue, 3 are red, 5 are yellow. What are the chances of me randomly picking a blue to eat?

Probability

- In the case of social statistics, we want the chance of picking a certain outcome (our sample) given what we know about all the possible outcomes (the population)

Probability

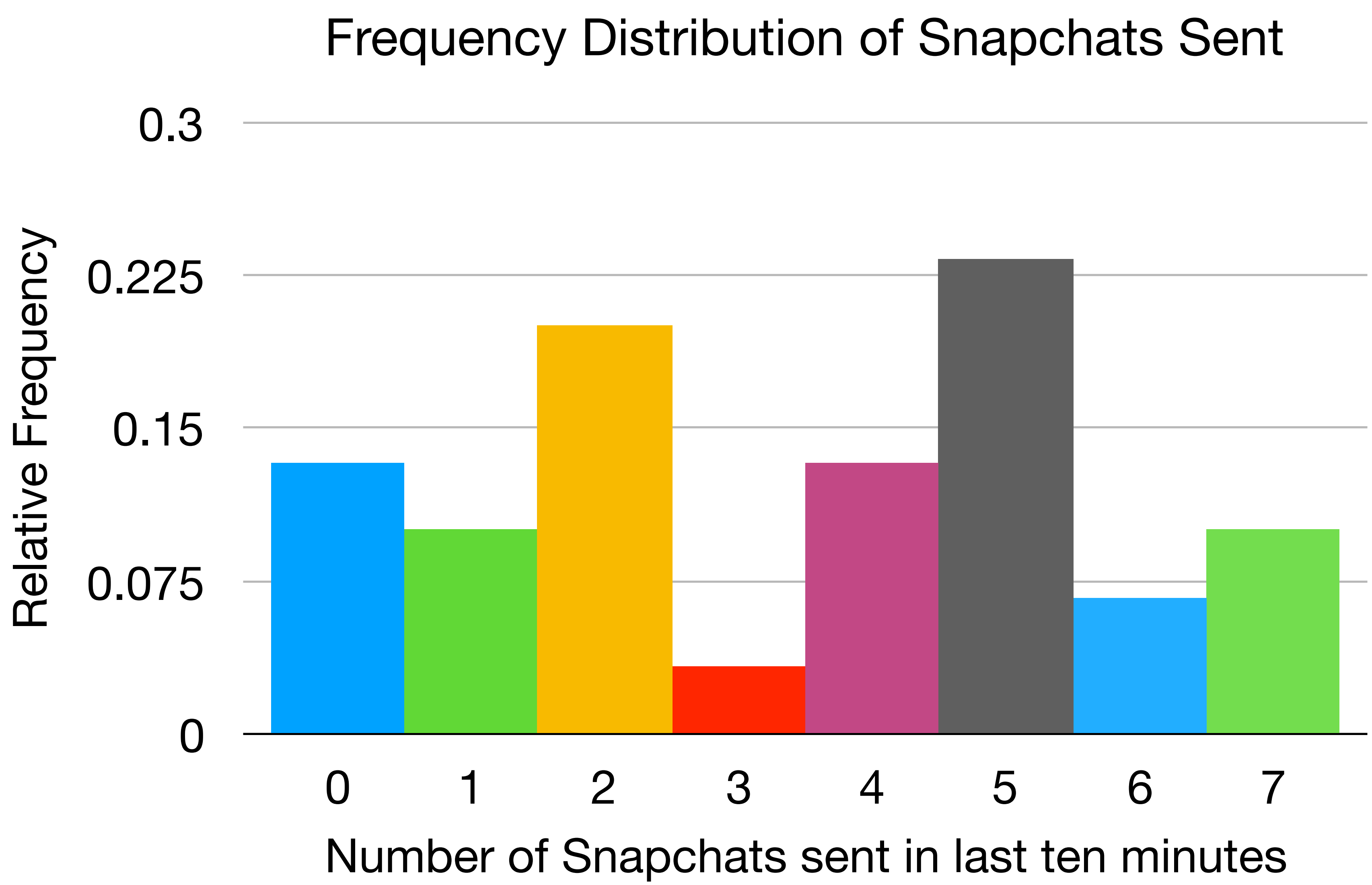
- The mean number of cats owned by my sample of graduate students is 5. If previous data suggests that graduate students own an average of 1 cat each, what does this suggest about my sample?
    - 5 is larger than 1, but only by a little bit, so the previous data is probably correct
    - What? No. 5 cats take up much more space than 1 cat. The previous data needs to be updated.

Probability

- What are the chances that my sample just randomly had a lot of cat lovers in it?
- AKA What is the probability that this sample is correct given what we know about the population?
- To answer this question, we need to think about how we quantify our total population

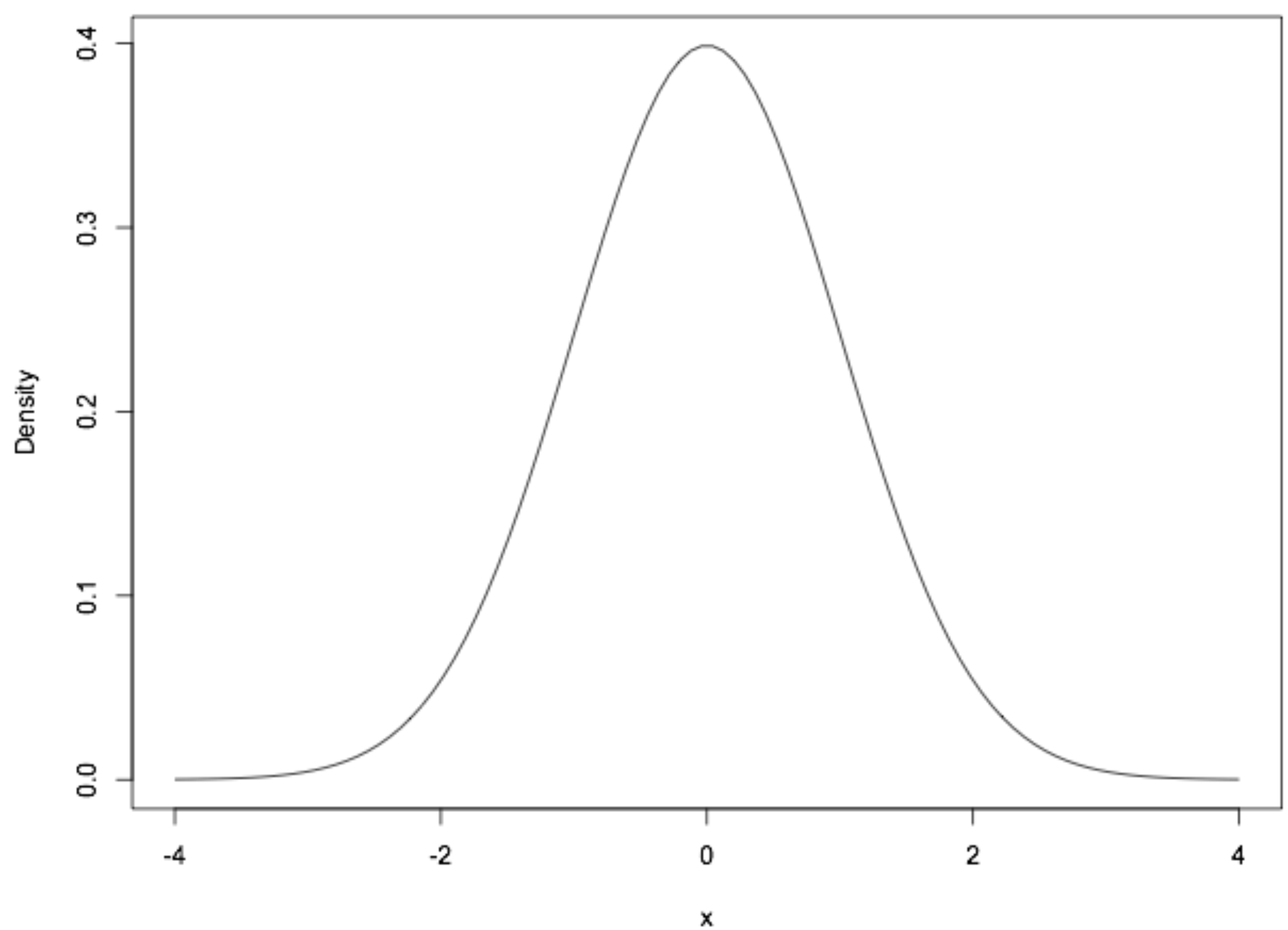
Z scores and Probability

- Area under the distribution curve gives us the "total" out of which we are pulling our sample



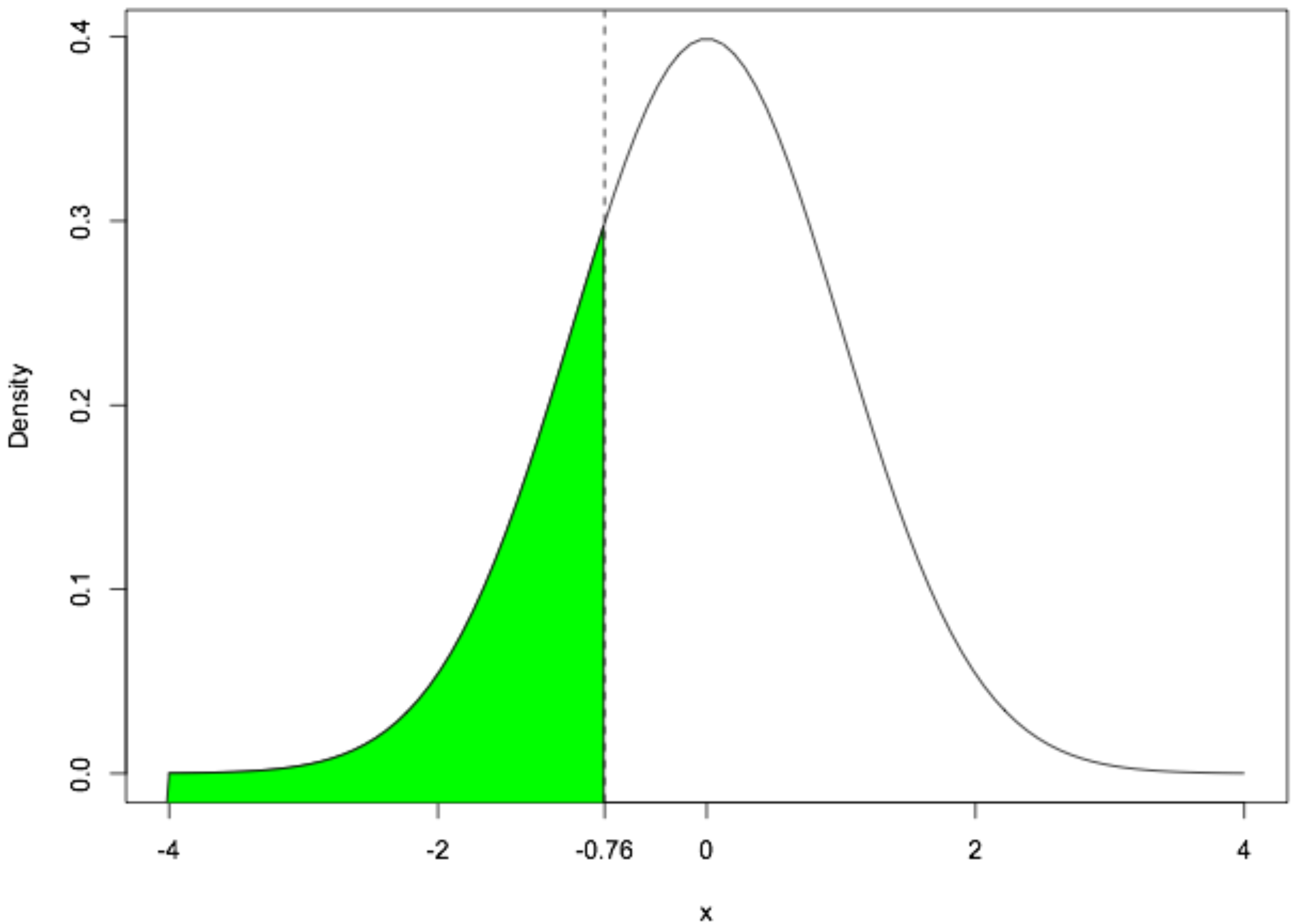
Z scores and Probability

- Thanks to calculus and the central limit theorem, we have ways of calculating this total population



Z scores and Probability

- We can figure out what the probability would be of selecting a sample at random and having it fall in a certain part of the distribution
- This is the ratio of the number of observations that fall above or below our sample to total number of observations
- Using Z scores (means and standard deviations) and their related proportions makes this (relatively) easier



Density
0.0
0.1
0.2
0.3
0.4
-4
-2
-0.76
0
2
4
x

Use a Z Table

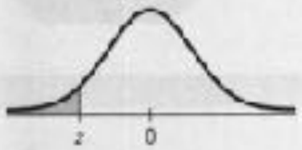
- Can't calculate the area directly (unless you use calculus)

- Use Z tables, printed in the back of text books and available online

$$\Phi(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z} e^{-\frac{t^2}{2}} dt$$

Table Z

Areas under the standard Normal curve

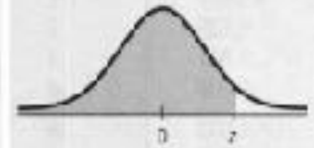


Second decimal place in z										
0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00	z
									0.0000†	−3.9
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	−3.8
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	−3.7
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	−3.6
0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	−3.5
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	−3.4
0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005	−3.3
0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	−3.2
0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010	−3.1
0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	−3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019	−2.9
0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026	−2.8
0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035	−2.7
0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047	−2.6
0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062	−2.5
0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082	−2.4
0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107	−2.3
0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139	−2.2
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179	−2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228	−2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287	−1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359	−1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446	−1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548	−1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668	−1.5
0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808	−1.4
0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968	−1.3
0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151	−1.2
0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357	−1.1
0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587	−1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841	−0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119	−0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420	−0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743	−0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085	−0.5
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821	−0.3
0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207	−0.2
0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602	−0.1
0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000	−0.0

†For $z \le -3.90$, the areas are 0.0000 to four decimal places.

Table Z (cont.)

Areas under the standard Normal curve

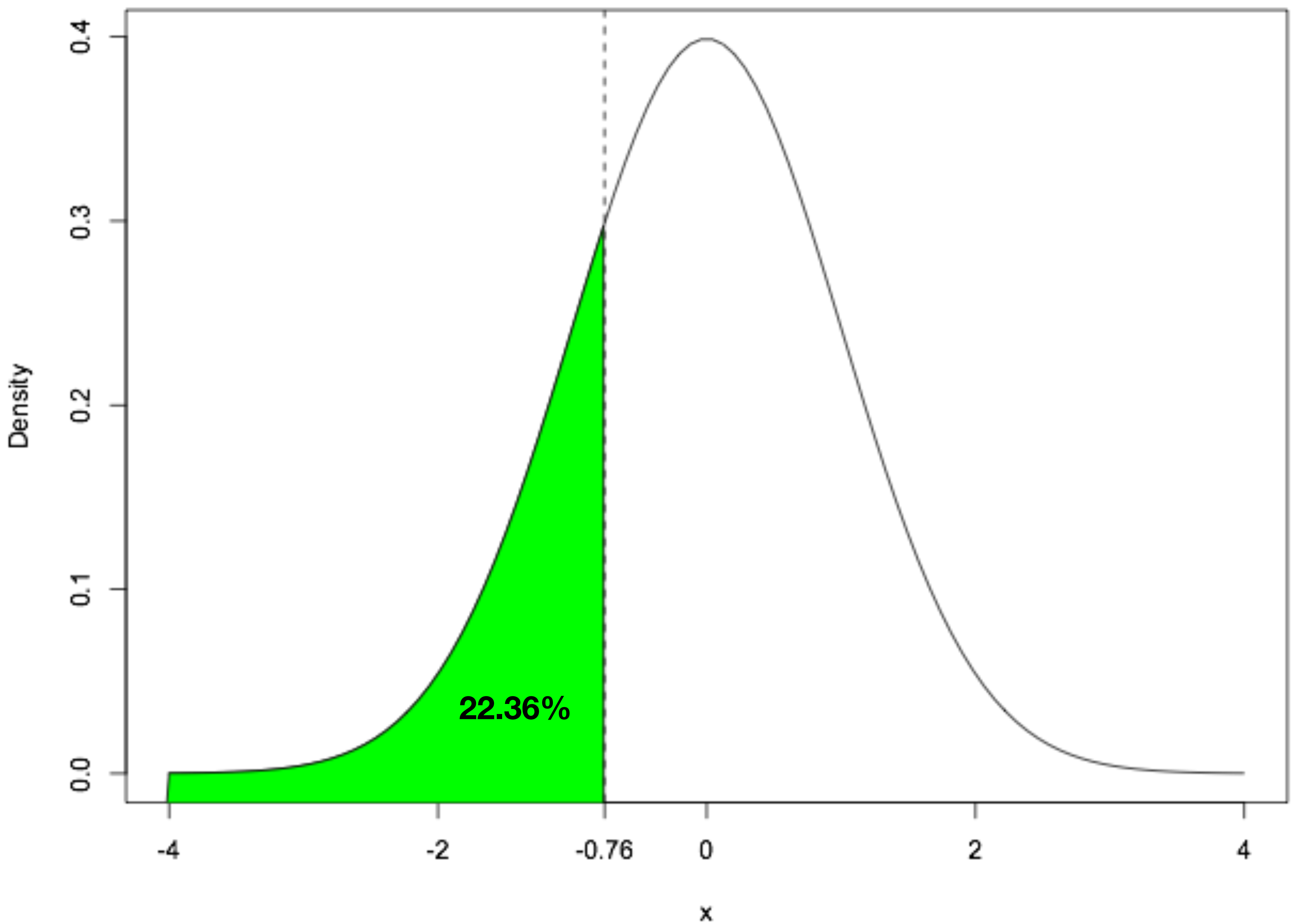


	Second decimal place in z									
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000†									

†For $z \ge 3.90$, the areas are 1.0000 to four decimal places.

Second decimal place in z										
0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00	z
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841	−0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119	−0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420	−0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743	−0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085	−0.5

$$P(Z \leq -0.76) = 0.2236$$



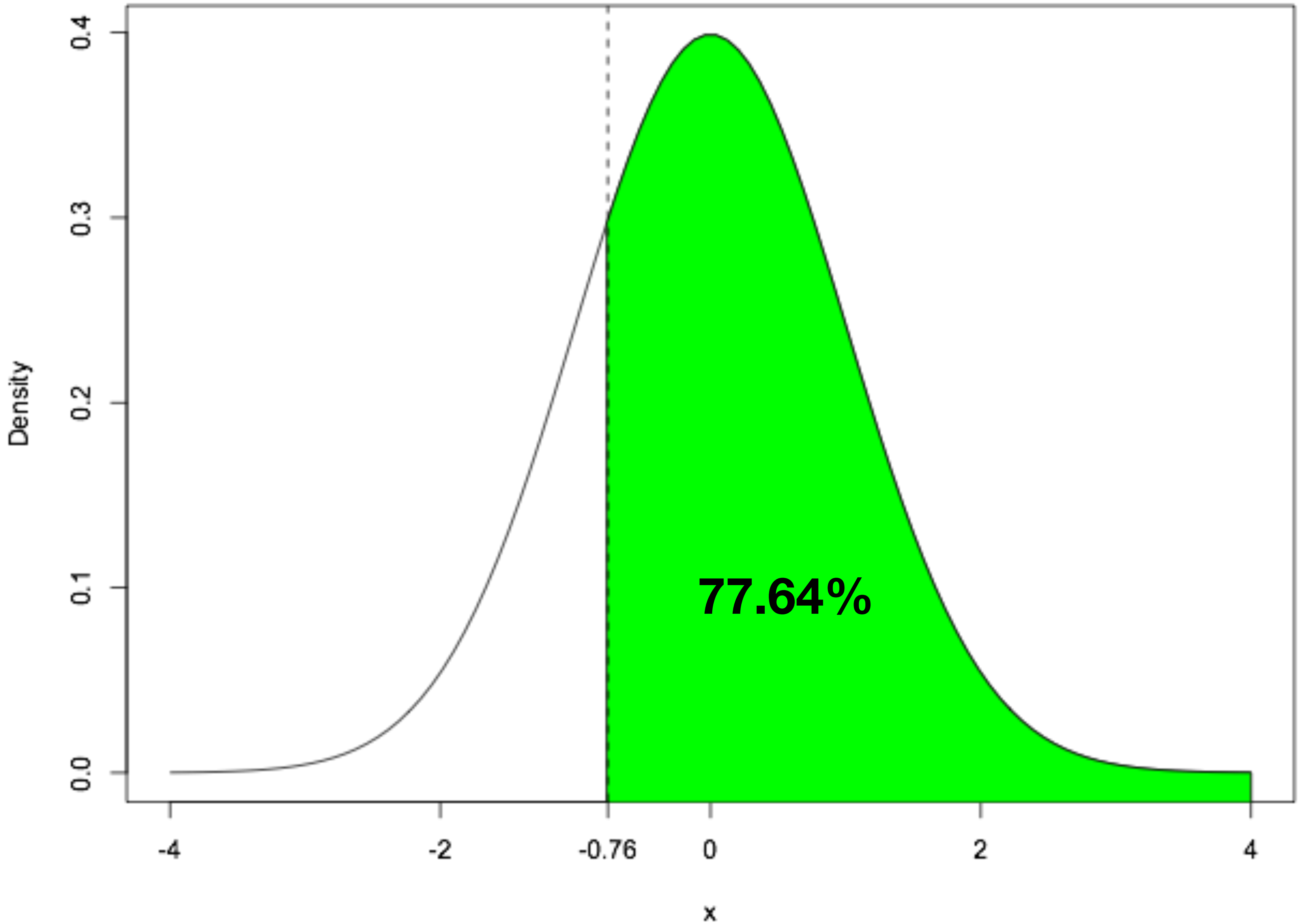
0.4
0.3
0.2
0.1
0.0
Density
22.36%
-4
-2
-0.76
0
2
4
x

Z scores

- Because the area under the curve by definition makes up 100% of the samples, knowing the proportion of one part of the curve gives us the other

- To find the proportion above a number, we can subtract the proportion below from 1

$$P(Z \geq -0.76) = 1 - 0.2236 = 0.7764$$



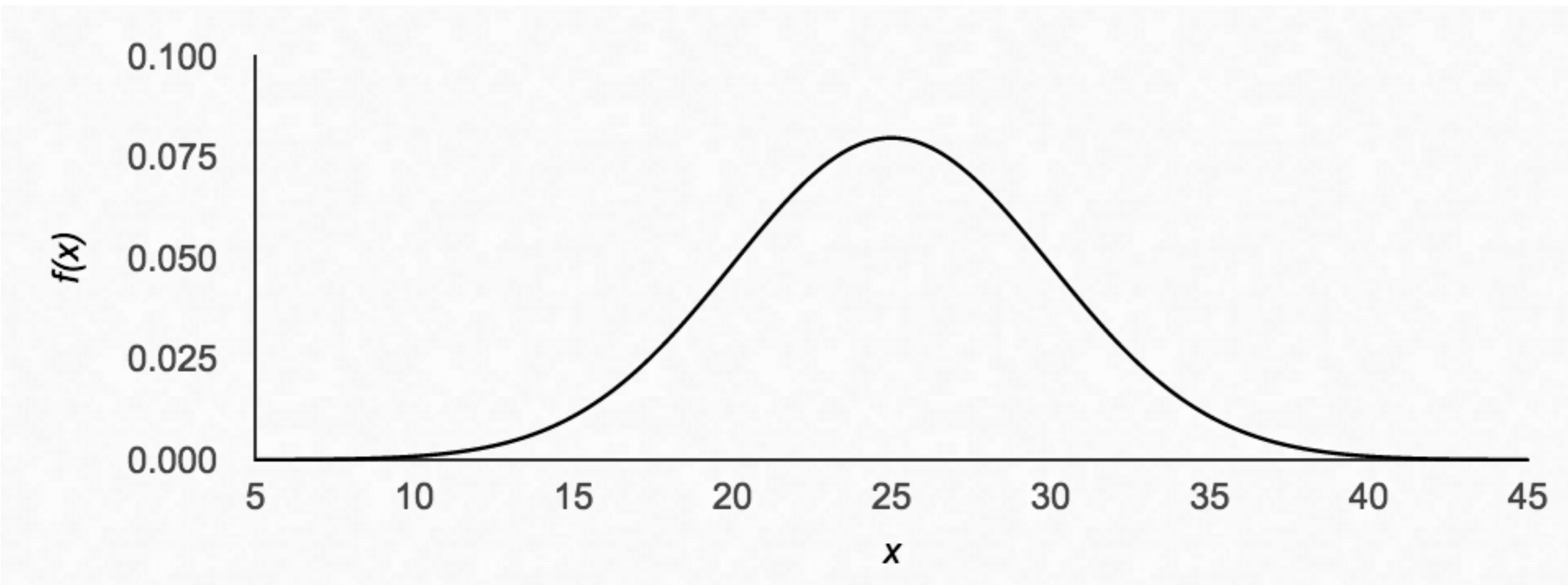
0.4
0.3
0.2
0.1
0.0
Density
77.64%
-4
-2
-0.76
0
2
4
x

Sample Question

- I’ve collected data on the number of hours spent playing video games in the past week. My sample is normally distributed, has a mean of 25 and a standard deviation of 5. I randomly pick a single respondent and they’ve only spent 10 hours playing. What were the chances that I chose someone with this response?

$$\bar{X} = 25 \qquad s = 5 \qquad X_i = 10$$

P(X≤10) = ?



$$\bar{X} = 25 \qquad s = 5 \qquad X_i = 10$$

$$Z = \frac{10 - 25}{5} = \frac{-15}{5} = -3$$

0.00	z
0.0000†	−3.9
0.0001	−3.8
0.0001	−3.7
0.0002	−3.6
0.0002	−3.5
0.0003	−3.4
0.0005	−3.3
0.0007	−3.2
0.0010	−3.1
$\boxed{0.0013}$	−3.0

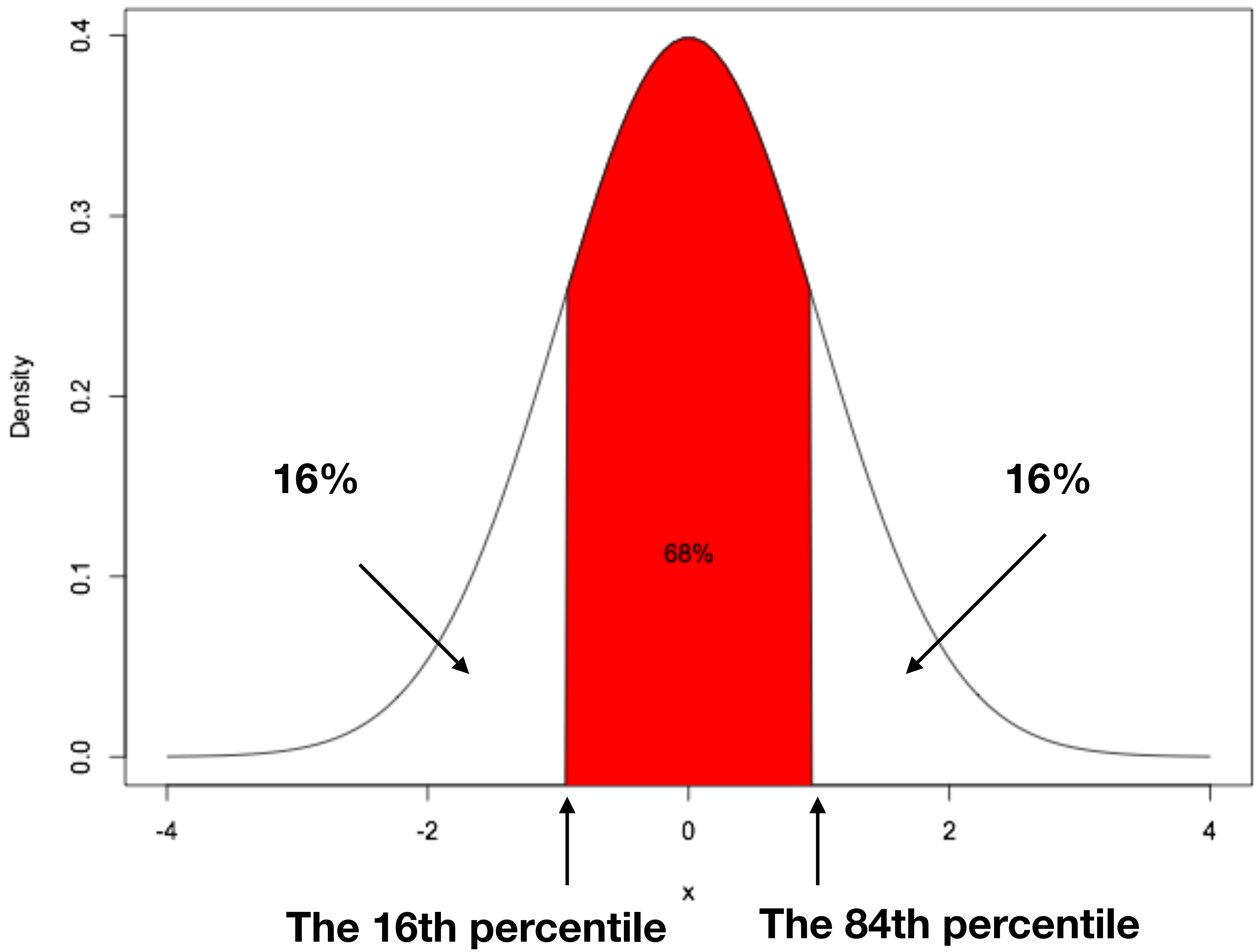
$$P(X \leq 10) = P(Z \leq -3) = 0.0013$$

The probability of randomly selecting a number less than 10 is 0.0013 or 0.13%

- http://www.intmath.com/counting-probability/normal-distribution-graph-interactive.php

Z Scores

- Z-scores are also useful when trying to find specific percentiles
- We’ve talked about the probability of selecting values above and below a certain threshold
- We can also think of these thresholds as percentiles, as they represent a certain percentage of respondents that fall under a specific section of the curve


0.4
0.3
0.2
0.1
0.0
Density
16%
16%
68%
-4
-2
0
2
4
x
The 16th percentile
The 84th percentile

Z Scores

- One common percentile we've already talked about is the median a.k.a. the 50th percentile

- The 25th and 75th percentiles are also called quartiles

- We can calculate the percentiles of a given sample with important pieces of information, like mean and standard deviation

Practice!

- Assume weekly hours worked by employees are normally distributed with a mean of 36 and a standard deviation of 4.

- (a) Calculate the 16th percentile of hours worked

$P(X \le .16) \longrightarrow Z = -1.00 \longrightarrow X = 36+4(-1)=32$

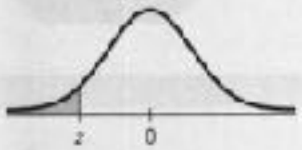
- (b) Calculate the 97.5th percentile of hours worked

$P(X \le .975) \longrightarrow Z = 1.96 \longrightarrow X = 36+4(1.96)=43.84$

Worksheet & Break

Table Z

Areas under the standard Normal curve

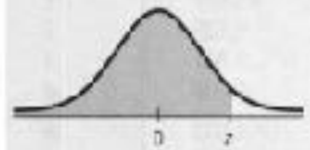


Second decimal place in z										
0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00	z
									0.0000†	−3.9
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	−3.8
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	−3.7
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	−3.6
0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	−3.5
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	−3.4
0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005	−3.3
0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	−3.2
0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010	−3.1
0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	−3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019	−2.9
0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026	−2.8
0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035	−2.7
0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047	−2.6
0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062	−2.5
0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082	−2.4
0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107	−2.3
0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139	−2.2
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179	−2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228	−2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287	−1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359	−1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446	−1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548	−1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668	−1.5
0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808	−1.4
0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968	−1.3
0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151	−1.2
0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357	−1.1
0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587	−1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841	−0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119	−0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420	−0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743	−0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085	−0.5
0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446	−0.4
0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821	−0.3
0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207	−0.2
0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602	−0.1
0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000	−0.0

†For $z \le -3.90$, the areas are 0.0000 to four decimal places.

Table Z (cont.)

Areas under the standard Normal curve



	Second decimal place in z									
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000†									

†For $z \ge 3.90$, the areas are 1.0000 to four decimal places.