ECON 3818

Chapter 3

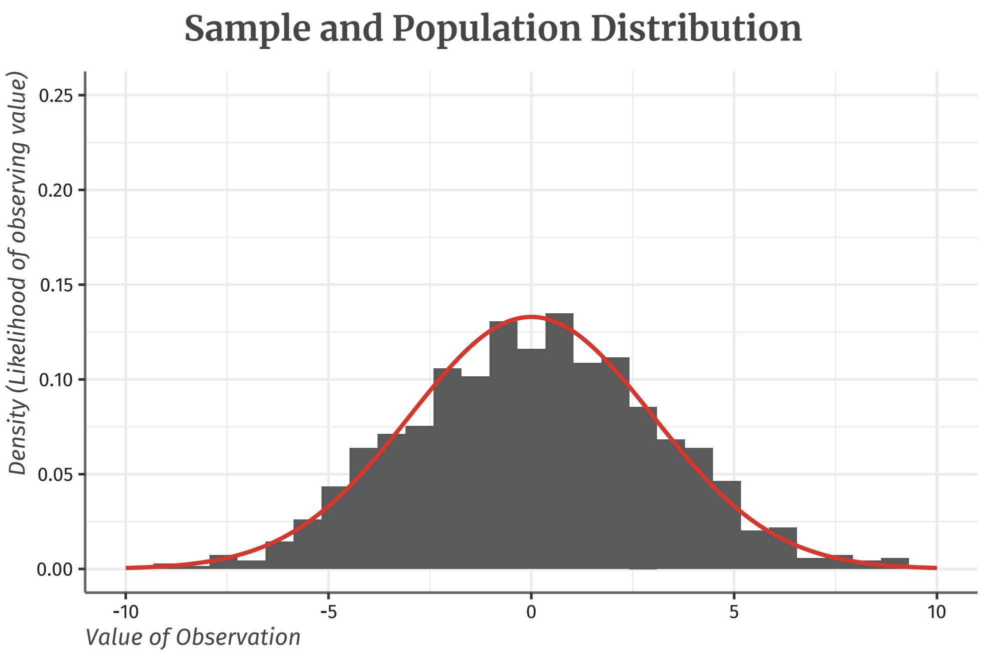
Kyle Butts

10 August 2021

Chapter 3: The Normal Distribution

Normal Distribution

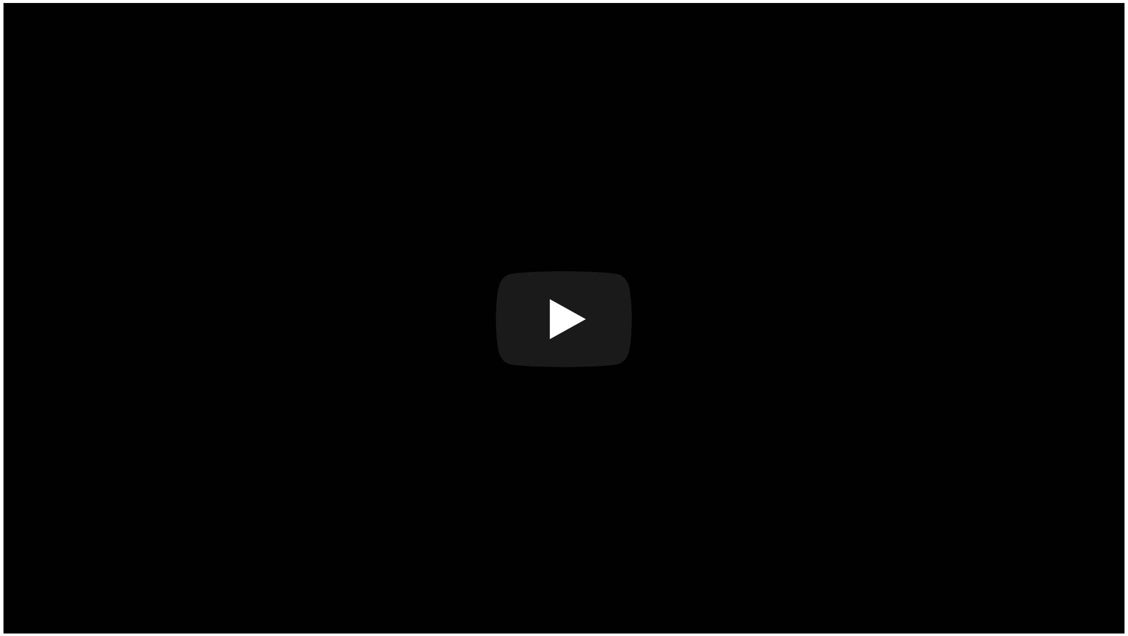
Normal curve is symmetric about the mean and bell-shaped:



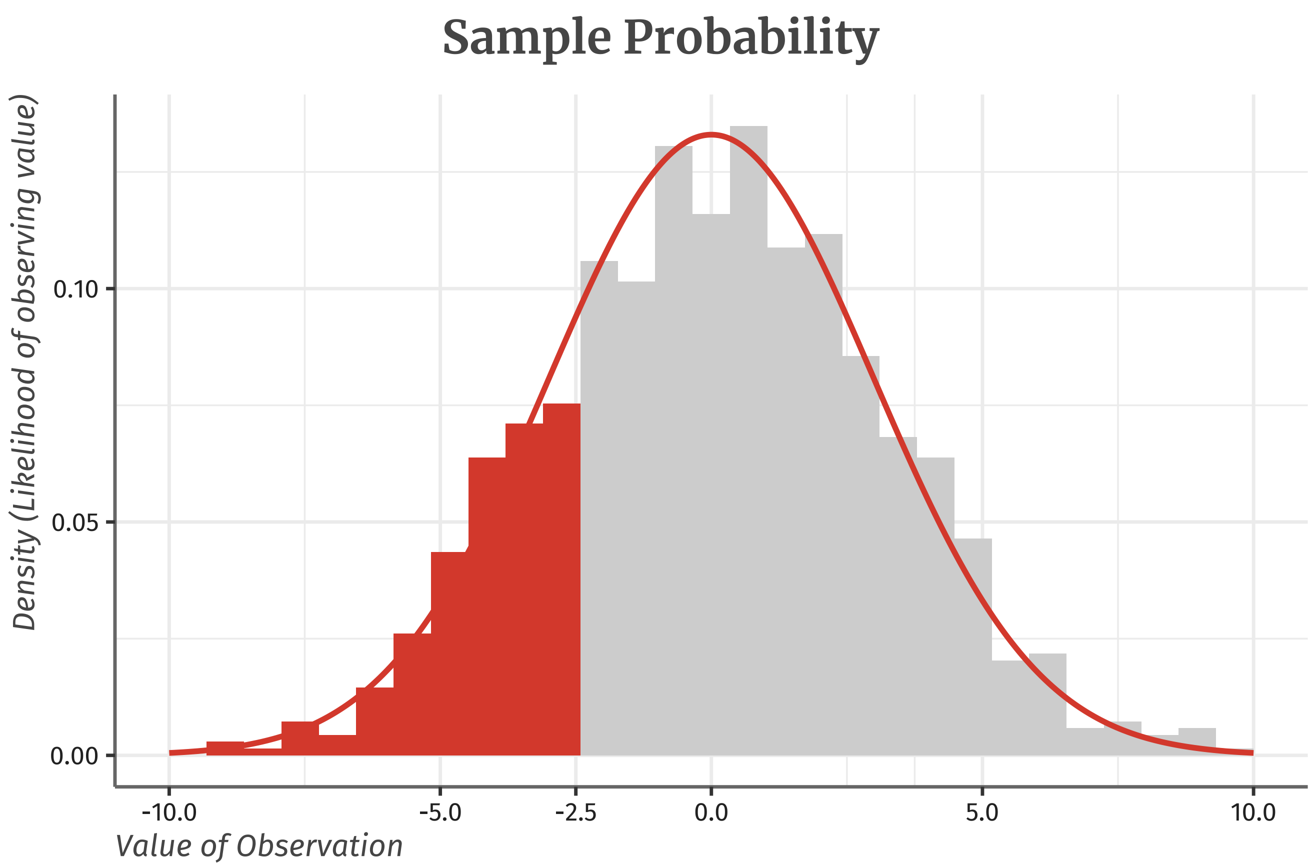
Lots of data naturally follow this distribution

- heights of people, blood pressure, grades on a test

Galton Board

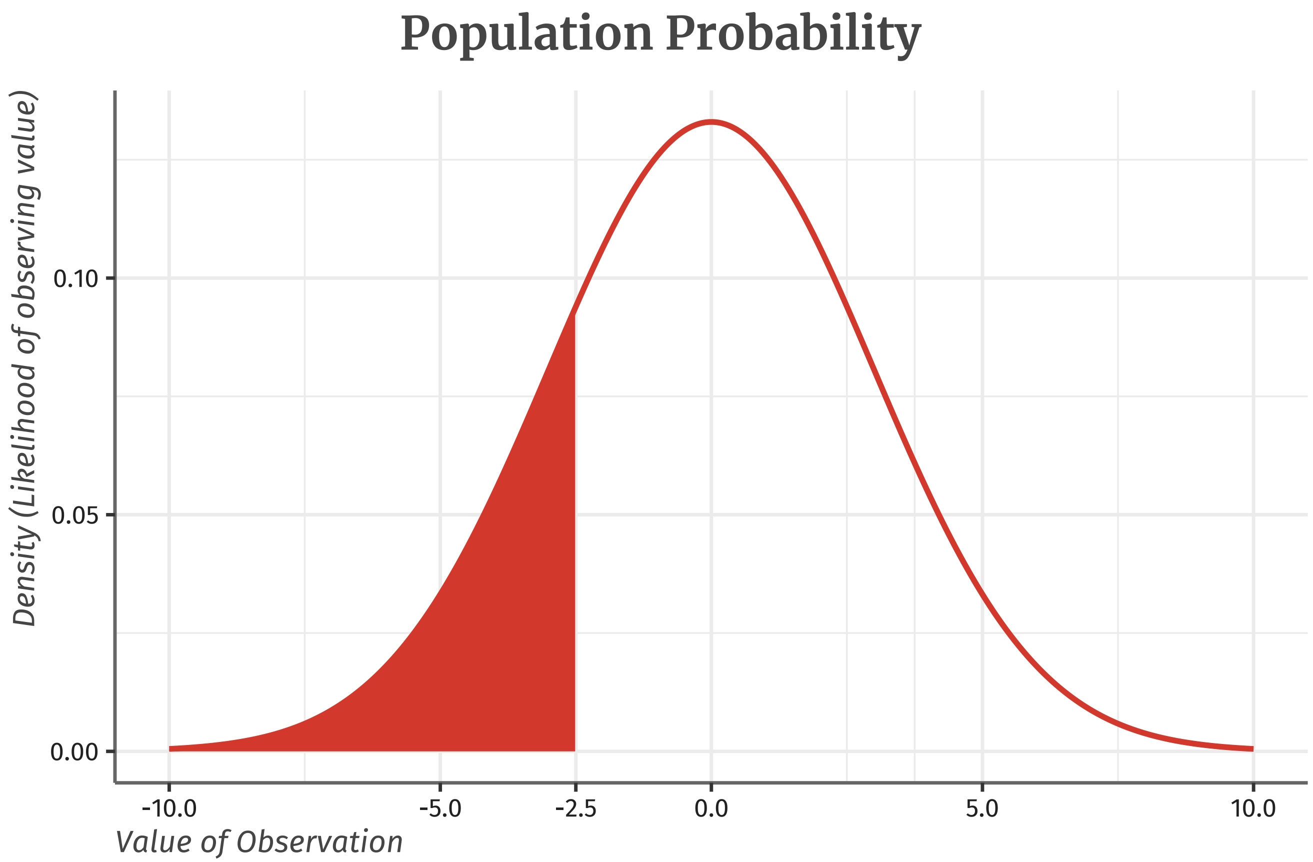


Sample Probability



- What's the probability that the x value is less than -2.5 **in our sample**?

Population Probability



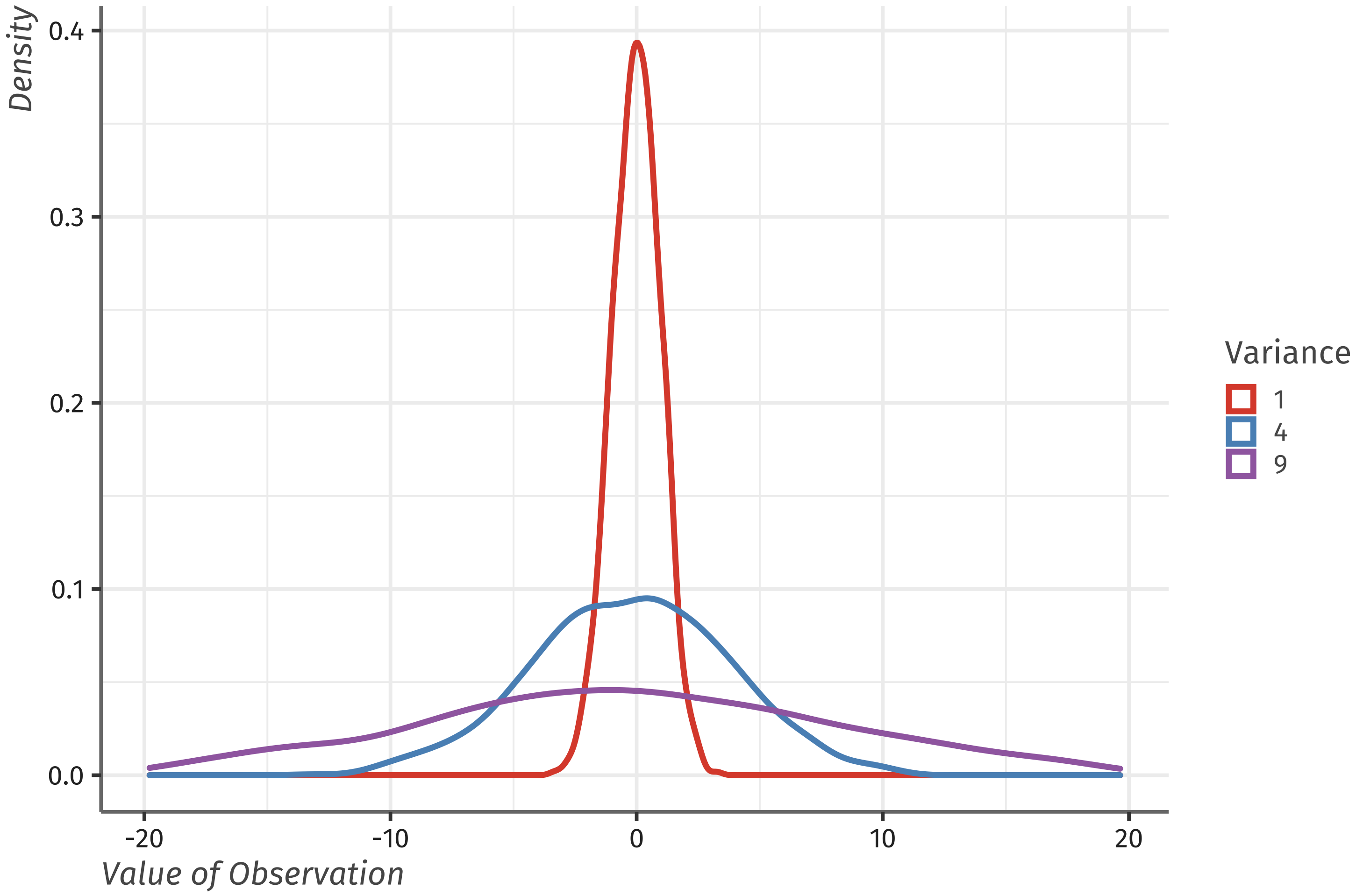
- What's the probability that the x value is less than -2.5 **in our population distribution**?

Parameters of Normal Distribution

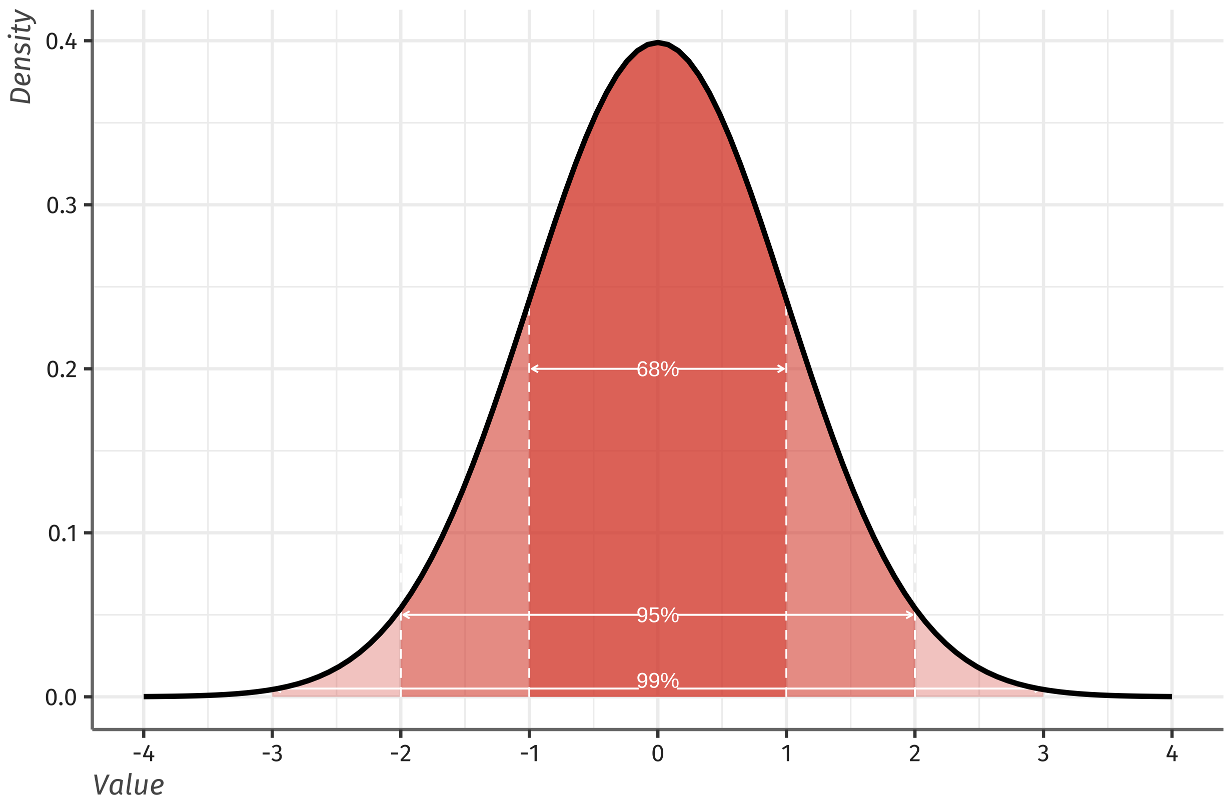
Normal distribution is described ***completely*** by two parameters-- its mean μ and variance σ^2

- The mean is located at the center of the symmetric curve
  - It is the same as the median
- Changing μ (without changing σ^2), moves the curve along the horizontal axis
- The variance describes the variability of the curve
- Higher variance means a flatter and wider distribution

Different Variances



68-95-99 Rule



- 68.2% of data is within ± 1 standard deviation of the mean
- 95.4% of data is within ± 2 standard deviation of the mean
- 99.6% of data is within ± 3 standard deviation of the mean

Clicker Question

Suppose that the mean birthweight in the sample is 113 oz. with a standard deviation of $\sqrt{484} \approx 22$ oz. Assuming babies' birthweight is normally distributed, how heavy are the middle 95% of babies?

a. 47 to 179 oz

b. 69 to 157 oz

c. 91 and 135 oz

d. 111 to 120 oz

Normal Distribution Notation

If X is distributed normally, we denote it the following way:

$$X \sim N(\mu, \sigma^2)$$

- This notation tells us everything we need to know about the normal distribution
- The distribution has mean μ
- The distribution has variance σ^2

Standard Normal Distribution

Standard normal distribution is a specific type of normal distribution

If a variable X follows a normal distribution with $\mu = 0$ and $\sigma^2 = 1$, we say that X follows the **standard normal distribution**

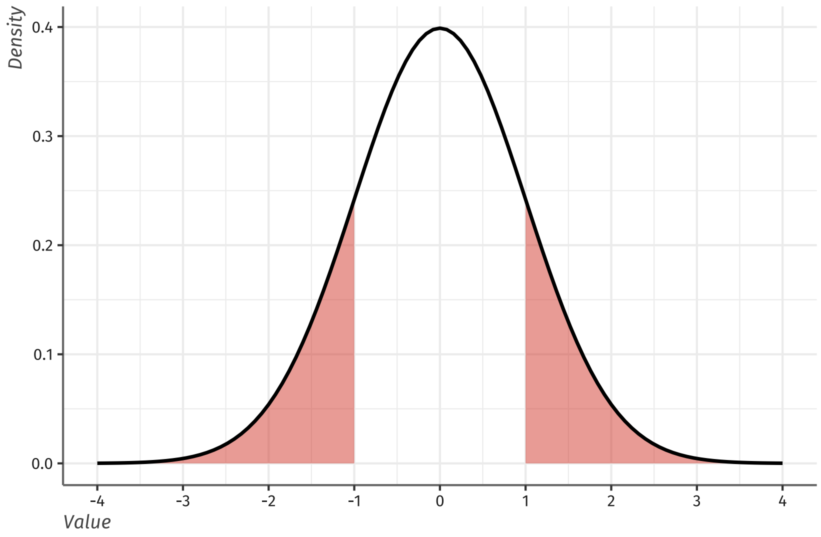
- Since it is so common, it is denoted as $Z \sim N(0, 1)$
- It is easier to find out probabilities about normal distributions if they are in the standard form
- Therefore we often will **standardize** any general normal distribution to be a standard normal

Properties of Standard Normal

Graph of the standard normal distribution has two important properties

Symmetric

$$P(Z < -1) = P(Z > 1)$$



Area under the curve sums to one

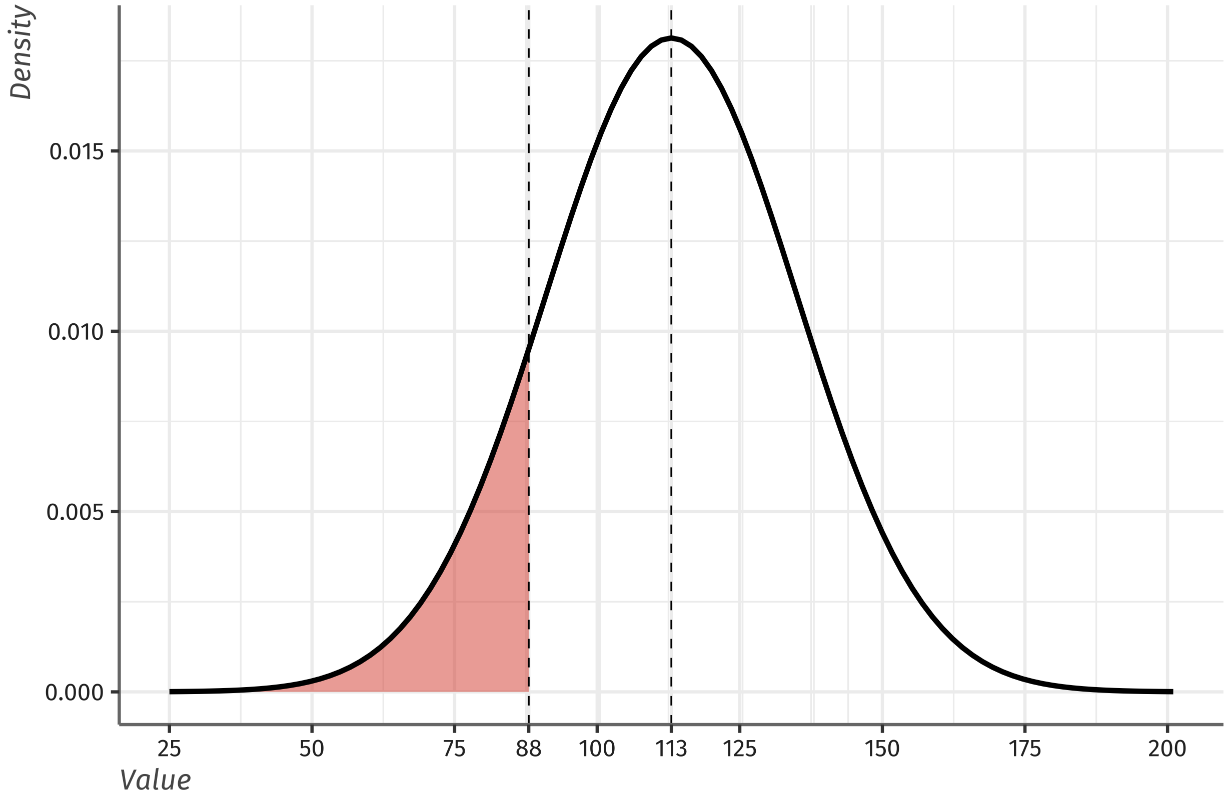
$$P(Z < 1) + P(Z > 1) = 1$$

Probabilities: Cumulative Proportions

Suppose we want to know the likelihood of a baby being born underweight (less than 88 oz).

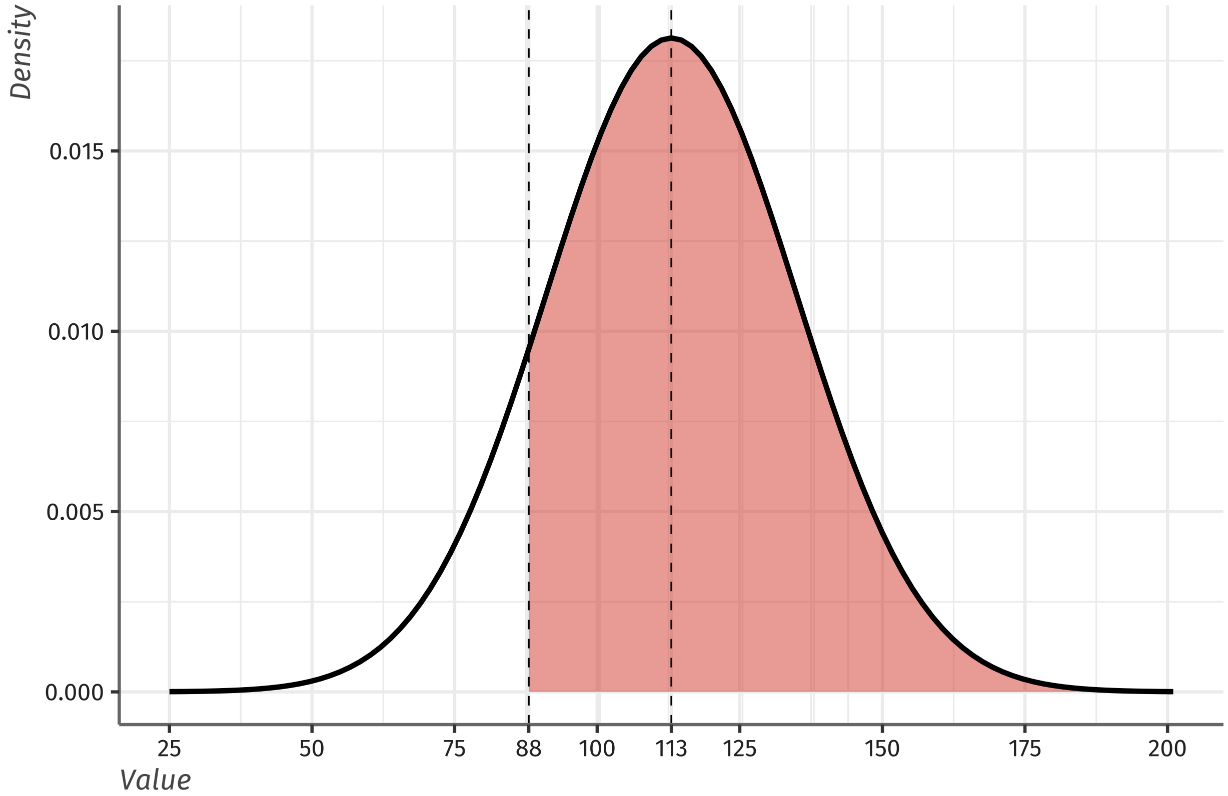
The data suggests $BW \sim N(113, 22^2)$, that is a mean of 113 and a standard deviation of 22. The probability of a baby being underweight is equal to $P(BW \leq 88)$. Graphically:

Left-tail probability



This probability is called the **left-tail probability** as it's every value to the left.

Right-tail Probability



If you want the **right-tail probability**, $P(BW > 88)$, you can use

$$P(BW > 88) = 1 - P(BW < 88)$$

Standardization

If a variable X has any normal distribution, $X \sim N(\mu, \sigma^2)$, then the standardized variable:

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

We call the standardized value the **Z-score**.

The Z-score is equivalent to the number of standard deviations that X is away from the mean.

Standardization

Since Z-scores are measured in number of standard deviations, we can compare across samples without having to worry about units.

For example:

- SAT scores are $X \sim N(1500, 250^2)$
- ACT scores are $Y \sim N(20.8, 2.8^2)$

You scored an 1860 on the SAT, your neighbor scored a 29 on the ACT. Who did better? Just compare Z-scores!

Calculating Probabilities

Coming back to the birth-weight example, let's do a little standardizing.

How do we actually calculate $P(BW \leq 88)$ (when $BW \sim N(113, 22^2)$)

First, standardize the distribution

$$P(BW \leq 88) = P(\frac{BW - \mu}{\sigma} \leq \frac{88 - 113}{22}) = P(Z \leq -1.14)$$

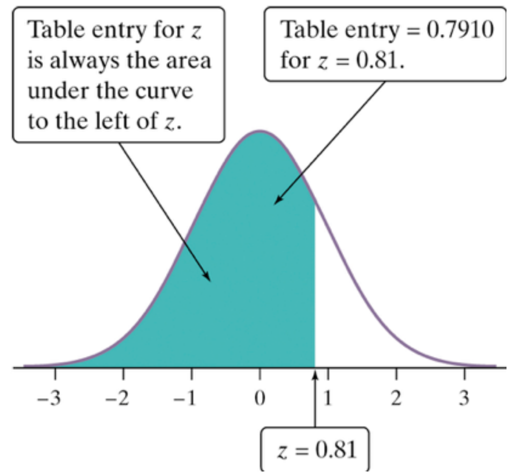
Then, we actually have a big table of left-tail probabilities for the *standard* normal distribution

- Table is either left-tail or right-tail (Z table on exam is left-tailed)

Standard Normal Tables

$$P(z < 0.81) = 0.7910$$

z	.00	.01	.02
0.7	.7580	.7611	.7642
0.8	.7881	.7910	.7939
0.9	.8159	.8186	.8212



Standard normal tables show the cumulative probability of different z-scores

- A table like this shows the **left-tail** probabilities. (One is available on the course site.)
- **Be Careful!** Some (but not many) tables display **right-tail** probabilities.

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.6	0.0006	0.0006	0.0007	0.0007	0.0007	0.0007	0.0008	0.0008	0.0008	0.0008
-3.5	0.0009	0.0009	0.0009	0.0010	0.0010	0.0010	0.0011	0.0011	0.0012	0.0012
-3.4	0.0012	0.0013	0.0013	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017
-3.3	0.0017	0.0018	0.0018	0.0019	0.0020	0.0020	0.0021	0.0022	0.0022	0.0023
-3.2	0.0024	0.0025	0.0025	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032
-3.1	0.0033	0.0034	0.0035	0.0036	0.0037	0.0038	0.0039	0.0040	0.0042	0.0043
-3.0	0.0044	0.0046	0.0047	0.0048	0.0050	0.0051	0.0053	0.0055	0.0056	0.0058
-2.9	0.0060	0.0061	0.0063	0.0065	0.0067	0.0069	0.0071	0.0073	0.0075	0.0077
-2.8	0.0079	0.0081	0.0084	0.0086	0.0088	0.0091	0.0093	0.0096	0.0099	0.0101
-2.7	0.0104	0.0107	0.0110	0.0113	0.0116	0.0119	0.0122	0.0126	0.0129	0.0132
-2.6	0.0136	0.0139	0.0143	0.0147	0.0151	0.0154	0.0158	0.0163	0.0167	0.0171
-2.5	0.0175	0.0180	0.0184	0.0189	0.0194	0.0198	0.0203	0.0208	0.0213	0.0219
-2.4	0.0224	0.0229	0.0235	0.0241	0.0246	0.0252	0.0258	0.0264	0.0270	0.0277
-2.3	0.0283	0.0290	0.0297	0.0303	0.0310	0.0317	0.0325	0.0332	0.0339	0.0347
-2.2	0.0355	0.0363	0.0371	0.0379	0.0387	0.0396	0.0404	0.0413	0.0422	0.0431
-2.1	0.0440	0.0449	0.0459	0.0468	0.0478	0.0488	0.0498	0.0508	0.0519	0.0529
-2.0	0.0540	0.0551	0.0562	0.0573	0.0584	0.0596	0.0608	0.0620	0.0632	0.0644
-1.9	0.0656	0.0669	0.0681	0.0694	0.0707	0.0721	0.0734	0.0748	0.0761	0.0775
-1.8	0.0790	0.0804	0.0818	0.0833	0.0848	0.0863	0.0878	0.0893	0.0909	0.0925
-1.7	0.0940	0.0957	0.0973	0.0989	0.1006	0.1023	0.1040	0.1057	0.1074	0.1092
-1.6	0.1109	0.1127	0.1145	0.1163	0.1182	0.1200	0.1219	0.1238	0.1257	0.1276
-1.5	0.1295	0.1315	0.1334	0.1354	0.1374	0.1394	0.1415	0.1435	0.1456	0.1476
-1.4	0.1497	0.1518	0.1539	0.1561	0.1582	0.1604	0.1626	0.1647	0.1669	0.1691
-1.3	0.1714	0.1736	0.1758	0.1781	0.1804	0.1826	0.1849	0.1872	0.1895	0.1919
-1.2	0.1942	0.1965	0.1989	0.2012	0.2036	0.2059	0.2083	0.2107	0.2131	0.2155
-1.1	0.2179	0.2203	0.2227	0.2251	0.2275	0.2299	0.2323	0.2347	0.2371	0.2396
-1.0	0.2420	0.2444	0.2468	0.2492	0.2516	0.2541	0.2565	0.2589	0.2613	0.2637
-0.9	0.2661	0.2685	0.2709	0.2732	0.2756	0.2780	0.2803	0.2827	0.2850	0.2874
-0.8	0.2897	0.2920	0.2943	0.2966	0.2989	0.3011	0.3034	0.3056	0.3079	0.3101
-0.7	0.3123	0.3144	0.3166	0.3187	0.3209	0.3230	0.3251	0.3271	0.3292	0.3312
-0.6	0.3332	0.3352	0.3372	0.3391	0.3410	0.3429	0.3448	0.3467	0.3485	0.3503
-0.5	0.3521	0.3538	0.3555	0.3572	0.3589	0.3605	0.3621	0.3637	0.3653	0.3668
-0.4	0.3683	0.3697	0.3712	0.3725	0.3739	0.3752	0.3765	0.3778	0.3790	0.3802
-0.3	0.3814	0.3825	0.3836	0.3847	0.3857	0.3867	0.3876	0.3885	0.3894	0.3902
-0.2	0.3910	0.3918	0.3925	0.3932	0.3939	0.3945	0.3951	0.3956	0.3961	0.3965
-0.1	0.3970	0.3973	0.3977	0.3980	0.3982	0.3984	0.3986	0.3988	0.3989	0.3989
-0.0	0.3989	0.3989	0.3989	0.3988	0.3986	0.3984	0.3982	0.3980	0.3977	0.3973

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.3989	0.3989	0.3989	0.3988	0.3986	0.3984	0.3982	0.3980	0.3977	0.3973
0.1	0.3970	0.3965	0.3961	0.3956	0.3951	0.3945	0.3939	0.3932	0.3925	0.3918
0.2	0.3910	0.3902	0.3894	0.3885	0.3876	0.3867	0.3857	0.3847	0.3836	0.3825
0.3	0.3814	0.3802	0.3790	0.3778	0.3765	0.3752	0.3739	0.3725	0.3712	0.3697
0.4	0.3683	0.3668	0.3653	0.3637	0.3621	0.3605	0.3589	0.3572	0.3555	0.3538
0.5	0.3521	0.3503	0.3485	0.3467	0.3448	0.3429	0.3410	0.3391	0.3372	0.3352
0.6	0.3332	0.3312	0.3292	0.3271	0.3251	0.3230	0.3209	0.3187	0.3166	0.3144
0.7	0.3123	0.3101	0.3079	0.3056	0.3034	0.3011	0.2989	0.2966	0.2943	0.2920
0.8	0.2897	0.2874	0.2850	0.2827	0.2803	0.2780	0.2756	0.2732	0.2709	0.2685
0.9	0.2661	0.2637	0.2613	0.2589	0.2565	0.2541	0.2516	0.2492	0.2468	0.2444
1.0	0.2420	0.2396	0.2371	0.2347	0.2323	0.2299	0.2275	0.2251	0.2227	0.2203
1.1	0.2179	0.2155	0.2131	0.2107	0.2083	0.2059	0.2036	0.2012	0.1989	0.1965
1.2	0.1942	0.1919	0.1895	0.1872	0.1849	0.1826	0.1804	0.1781	0.1758	0.1736
1.3	0.1714	0.1691	0.1669	0.1647	0.1626	0.1604	0.1582	0.1561	0.1539	0.1518
1.4	0.1497	0.1476	0.1456	0.1435	0.1415	0.1394	0.1374	0.1354	0.1334	0.1315
1.5	0.1295	0.1276	0.1257	0.1238	0.1219	0.1200	0.1182	0.1163	0.1145	0.1127
1.6	0.1109	0.1092	0.1074	0.1057	0.1040	0.1023	0.1006	0.0989	0.0973	0.0957
1.7	0.0940	0.0925	0.0909	0.0893	0.0878	0.0863	0.0848	0.0833	0.0818	0.0804
1.8	0.0790	0.0775	0.0761	0.0748	0.0734	0.0721	0.0707	0.0694	0.0681	0.0669
1.9	0.0656	0.0644	0.0632	0.0620	0.0608	0.0596	0.0584	0.0573	0.0562	0.0551
2.0	0.0540	0.0529	0.0519	0.0508	0.0498	0.0488	0.0478	0.0468	0.0459	0.0449
2.1	0.0440	0.0431	0.0422	0.0413	0.0404	0.0396	0.0387	0.0379	0.0371	0.0363
2.2	0.0355	0.0347	0.0339	0.0332	0.0325	0.0317	0.0310	0.0303	0.0297	0.0290
2.3	0.0283	0.0277	0.0270	0.0264	0.0258	0.0252	0.0246	0.0241	0.0235	0.0229
2.4	0.0224	0.0219	0.0213	0.0208	0.0203	0.0198	0.0194	0.0189	0.0184	0.0180
2.5	0.0175	0.0171	0.0167	0.0163	0.0158	0.0154	0.0151	0.0147	0.0143	0.0139
2.6	0.0136	0.0132	0.0129	0.0126	0.0122	0.0119	0.0116	0.0113	0.0110	0.0107
2.7	0.0104	0.0101	0.0099	0.0096	0.0093	0.0091	0.0088	0.0086	0.0084	0.0081
2.8	0.0079	0.0077	0.0075	0.0073	0.0071	0.0069	0.0067	0.0065	0.0063	0.0061
2.9	0.0060	0.0058	0.0056	0.0055	0.0053	0.0051	0.0050	0.0048	0.0047	0.0046
3.0	0.0044	0.0043	0.0042	0.0040	0.0039	0.0038	0.0037	0.0036	0.0035	0.0034
3.1	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026	0.0025	0.0025
3.2	0.0024	0.0023	0.0022	0.0022	0.0021	0.0020	0.0020	0.0019	0.0018	0.0018
3.3	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014	0.0013	0.0013
3.4	0.0012	0.0012	0.0012	0.0011	0.0011	0.0010	0.0010	0.0010	0.0009	0.0009
3.5	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007	0.0007	0.0007	0.0006
3.6	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0004

Using a Z-Table

Back to our birth-weight example. We want

$$P(Z < -1.14)$$

Method 1: if you have negative values in Z-Table:

- Look up Z = -1.14 in z-table

Method 2: If you have only positive values in Z-Table:

$$\begin{aligned} P(Z \leq -1.14) &= P(Z \geq 1.14) \\ &= 1 - P(Z \leq 1.14) \\ &= 1 - .8729 \\ &= 0.1271 \end{aligned}$$

Normal Example

A company chooses its new entry-level employees from a pool of recent college graduates. The cumulative GPA of the candidates is used as a tie-breaker. GPAs for the successful interviewees are normally distributed, with a mean of 3.3 and a standard deviation of 0.4. What proportion of candidates have a GPA under 3.0?

a. 2.3% b. 22.7% c. 55.1% d. 77.3%

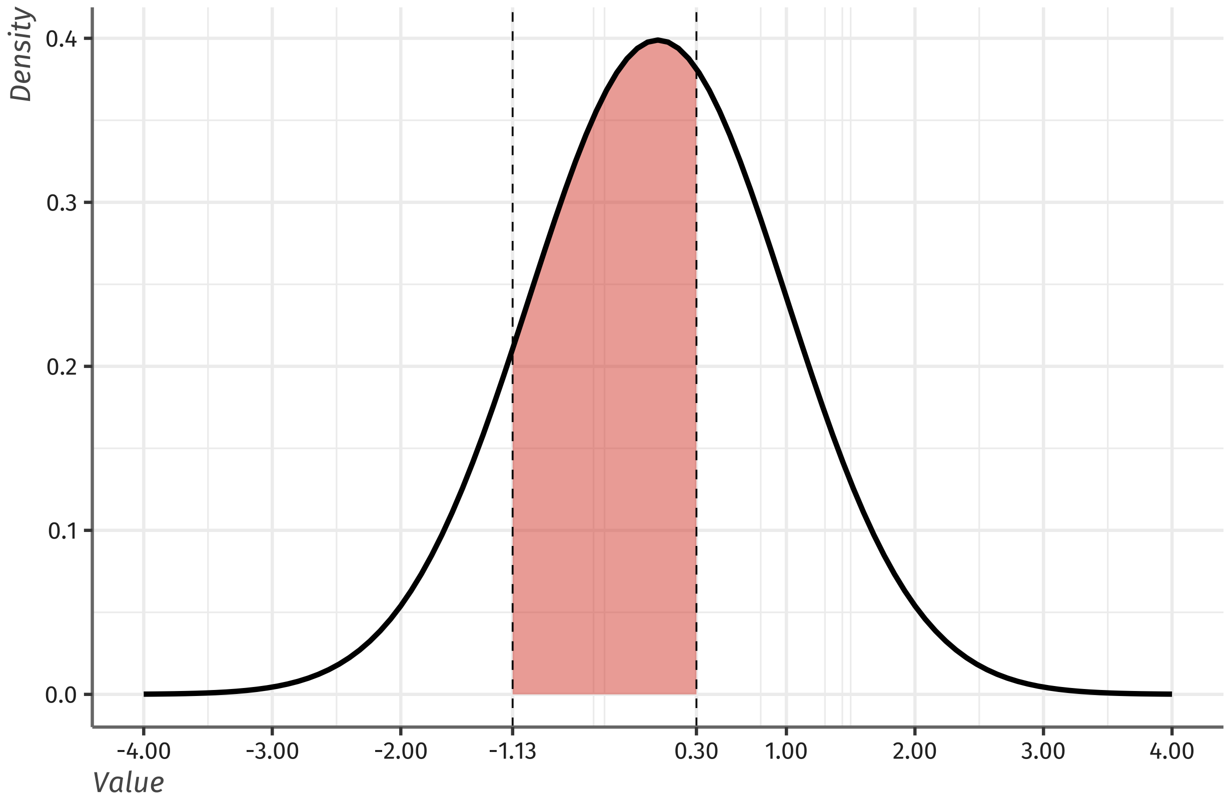
Clicker Question

Consider the scenario on the previous slide, where $GPA \sim N(3.3, 0.4^2)$. What percent of candidates have a GPA above 3.9?

a. 2.3% b. 6.7% c. 93.3% d. 97.7%

Area In Between Z-Scores

Suppose we want to calculate $P(-1.13 \leq Z \leq 0.3)$ Graphically we want to calculate the following shaded area:



Area In Between Z-Scores

In order to calculate the area between two z-scores,

$$P(-1.13 \leq Z \leq 0.3) = P(Z \leq 0.3) - P(Z \leq -1.13)$$

So we calculate:

- $P(Z \leq 0.3) = 0.6179$
- $P(Z \leq -1.13) = 0.1292$

$$\begin{aligned} P(-1.13 \leq Z \leq 0.3) &= P(Z \leq 0.3) - P(Z \leq -1.13) \\ &= 0.6179 - 0.1292 \\ &= 0.4887 \end{aligned}$$

Clicker Question

A typical college freshman spends an average of $\mu = 150$ minutes per day with a standard deviation of $\sigma = 50$ minutes, on social media. The distribution of time on social media is known be Normal. What is the probability a college freshman spends between 2 and 3 hours on social media?

a. 72.57% b. 27.43% c. 45.14%

Using probability to calculate Z-score

So far, we've used z-scores to calculate probabilities (values inside the table)

In some cases, we will use probabilities to calculate z-scores (values outside the table)

Example

Scores on the SAT verbal test follow approximately the $N(515, 109^2)$ distribution. How high must a student score in order to place in top 5% of all students taking the SAT?

Example

Back to the example discussing the distribution of GPAs, where GPA $\sim N(3.3, 0.4^2)$. If the company is interviewing 163 people, but only 121 can be hired, then what cut-off GPA should the company use?

Clicker Question

Suppose that $P(Z \leq z^*) = 0.025$. Using a standard normal table, find z^*

a. 0.5478 b. -0.5478 c. 1.96 d. -1.96

Review of Normal Distribution

Consider men's height to be distributed normally with a mean of 5.9 feet and a standard deviation or 0.4 feet. Calculate the following:

- $P(X > 6.5)$
- $P(X > 5)$
- What is the top 10% of men's height?
- What is the bottom 20% of men's height?