8-7 Fitting to a Normal Distribution

Going Deeper

Essential question: *How do you find percents of data and probabilities of events associated with normal distributions?*

COMMON CORE Standards for Mathematical Content

CC.9-12.S.ID.4 Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, ... and tables to estimate areas under the normal curve.*

Normal Curves All normal curves have the following properties, sometimes collectively called the *68-95-99.7 rule*:

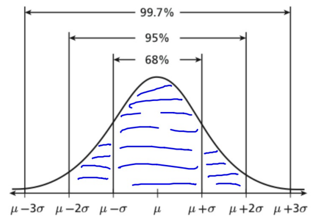
- 68% of the data fall within 1 standard deviation of the mean.
- 95% of the data fall within 2 standard deviations of the mean.
- 99.7% of the data fall within 3 standard deviations of the mean.

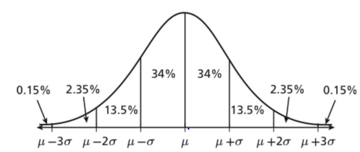
The figure at the left below illustrates the 68-95-99.7 rule.

A normal curve's symmetry allows you to separate the area under the curve into eight parts and know the percent of the data in each part, as shown at the right below.

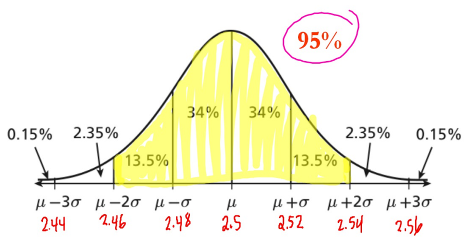
$99.7 - 95 = 4.7$, $\frac{4.7}{2} = 2.35$

$95 - 68 = 27$, $\frac{27}{2} = 13.5$





Suppose the masses (in grams) of pennies minted in the US after 1982 are normally distributed with a mean of 2.5 g and a std dev of 0.02g. Find the percent of pennies that have a mass between 2.46g and 2.54g.



RECALL: if it's more than 3 std dev away; it is an outlier. 99.7% of the data must be within 3 std dev away from the mean.

RECALL

Z - test

$n \geq 30$

μ **population mean**

$\bar{x}$ **sample mean**

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- If $|z| > 1.96$, then you can reject the null hypothesis with 95% certainty.
- If $|z| < 1.96$, then you do not have enough evidence to reject the null hypothesis.

This test is used to test a hypothesis
Works for ANY distribution

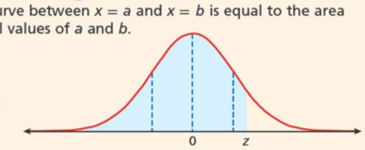
Standard Normal Values

If a random variable x is normally distributed with mean μ and standard deviation σ, then the **standard normal value** of x is given by the following formula:

$$z = \frac{x - \mu}{\sigma}$$

The area under the normal curve between $x = a$ and $x = b$ is equal to the area between the standard normal values of a and b.

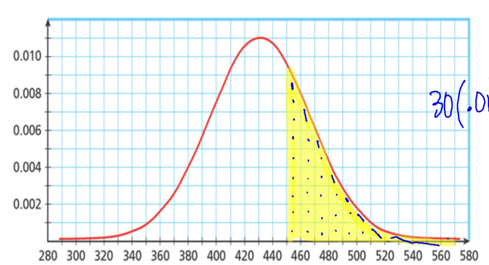
The table shows the approximate area under the standard normal curve for all values less than z for selected values of z.



z	−2.5	−2	−1.5	−1	−0.5	0	0.5	1	1.5	2	2.5
Area	0.01	0.02	0.07	0.16	0.31	0.5	0.69	0.84	0.93	0.98	0.99

This is only for NORMAL distribution.
Use to find area under the curve (probability)

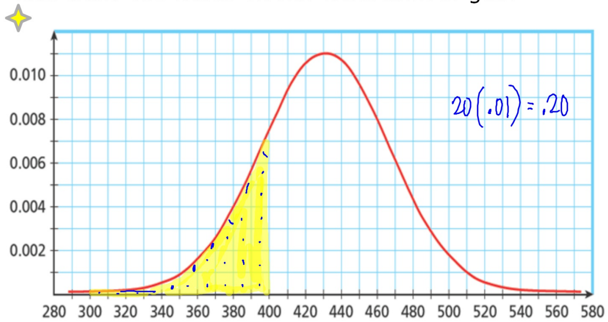
Jamie can drive her car an average of 432 miles per tank of gas, with a standard deviation of 36 miles. Use the graph to estimate the probability that Jamie will be able to drive more than 450 miles on her next tank of gas.



$30(.01) = .30$

The area under the normal curve is always equal to 1. Each square on the grid has an area of 10(0.001) = 0.01. Count the number of grid squares under the curve for values of *x greater than 450. There are about 31 squares under* the graph, so the probability is about 31(0.01) = 0.31 that she will be able to drive more than 450 miles on her next tank of gas.

Estimate the probability that Jamie will be able to drive less than 400 miles on her next tank of gas?



$20(.01) = .20$

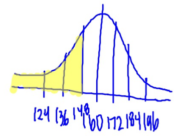
Scores on a test are normally distributed with a mean of 160 and a standard deviation of 12.

. Estimate the probability that a randomly selected student scored less than 148.

z	−2.5	−2	−1.5	−1	−0.5	0	0.5	1	1.5	2	2.5
Area	0.01	0.02	0.07	0.16	0.31	0.5	0.69	0.84	0.93	0.98	0.99

$\boxed{z = \frac{x-\mu}{\sigma}}$

$z = \frac{148-160}{12} = \frac{-12}{12} = -1$



16%

B. Estimate the probability that a randomly selected student scored between 154 and 184.

z	−2.5	−2	−1.5	−1	−0.5	0	0.5	1	1.5	2	2.5
Area	0.01	0.02	0.07	0.16	0.31	0.5	0.69	0.84	0.93	0.98	0.99

$z_{184} = \frac{184-160}{12} = \frac{24}{12} = 2 \Rightarrow 98\%$

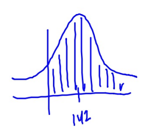
$z_{154} = \frac{154-160}{12} = \frac{-6}{12} = -0.5 \Rightarrow 31\%$

$98 - 31$

67%

Scores on a test are normally distributed with a mean of 142 and a standard deviation of 18. Estimate the probability of scoring above 106.

z	−2.5	−2	−1.5	−1	−0.5	0	0.5	1	1.5	2	2.5
Area	0.01	0.02	0.07	0.16	0.31	0.5	0.69	0.84	0.93	0.98	0.99



$z_{106} = \frac{106-142}{18} = \frac{-36}{18} = -2 \Rightarrow 2\%$

$1 - .02 = .98$

98%

Suppose the heights (in inches) of adult females in the US are normally distributed with a mean of 63.8 in and a std dev of 2.8 in.

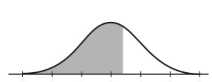
Find the percent of women who are no more than 65 in tall.

Standard Normal Table

z	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
−3	.0013	.0010	.0007	.0005	.0003	.0002	.0002	.0001	.0001	.0000+
−2	.0228	.0179	.0139	.0107	.0082	.0062	.0047	.0035	.0026	.0019
−1	.1587	.1357	.1151	.0968	.0808	.0668	.0548	.0446	.0359	.0287
−0	.5000	.4602	.4207	.3821	.3446	.3085	.2743	.2420	.2119	.1841
0	.5000	.5398	.5793	.6179	.6554	.6915	.7257	.7580	.7881	.8159
1	.8413	.8643	.8849	.9032	.9192	.9332	.9452	.9554	.9641	.9713
2	.9772	.9821	.9861	.9893	.9918	.9938	.9953	.9965	.9974	.9981
3	.9987	.9990	.9993	.9995	.9997	.9998	.9998	.9999	.9999	1.000−

$z_{65} = \frac{65-63.8}{2.8} \approx 0.4$

- Convert 65 to a z-score: $z_{65} = \frac{65-\mu}{\sigma} = \frac{65 - 63.8}{2.8} \approx 0.4$
- Recognize that the phrase "no more than 65 inches" means that $z \leq z_{65}$. Read the decimal from the appropriate row and column of the standard normal table: 0.6554



- Write the decimal as a percent, rounding to the nearest whole percent: about 66%

The probability that a randomly chosen woman is between 60 inches and 63 inches tall

Standard Normal Table

z	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
−3	.0013	.0010	.0007	.0005	.0003	.0002	.0002	.0001	.0001	.0000+
−2	.0228	.0179	.0139	.0107	.0082	.0062	.0047	.0035	.0026	.0019
−1	.1587	.1357	.1151	.0968	.0808	.0668	.0548	.0446	.0359	.0287
−0	.5000	.4602	.4207	.3821	.3446	.3085	.2743	.2420	.2119	.1841
0	.5000	.5398	.5793	.6179	.6554	.6915	.7257	.7580	.7881	.8159
1	.8413	.8643	.8849	.9032	.9192	.9332	.9452	.9554	.9641	.9713
2	.9772	.9821	.9861	.9893	.9918	.9938	.9953	.9965	.9974	.9981
3	.9987	.9990	.9993	.9995	.9997	.9998	.9998	.9999	.9999	1.000−

$z_{63} = \frac{63-63.8}{2.8} \approx -0.3 \Rightarrow 38.21\%$

$z_{60} = \frac{60-63.8}{2.8} \approx -1.4 \Rightarrow 8.08\%$

30.13%

- below a z-value. Use the area given in the chart for that value.
- between two z-values. Subtract the areas given in the chart.
- above a z-value. Subtract the area from 1.