

Section 3.5a - Applying the Normal Distribution

MDM4U

Jensen

Part 1: Normal Distribution Video

While watching the video, answer the following questions

1. What is another name for the Empirical rule?

The 68-95-99.7% Rule.

2. How tall must a woman be to join the Boston Beanstalks Club?

At least 5 feet 10 inches tall.

3. How do you calculate a z-score?

Suppose x is an observation from a normal distribution with mean μ and standard deviation σ . To calculate the z-score, subtract μ from x and then divide the result by σ .

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

4. Based on z-scores, are eligibility requirements to join the Boston Beanstalks more difficult to meet for men or for women?

The eligibility z-score for women (1.48) is higher than for men (0.98). So, in order to join the Beanstalks, women's heights must be at least 1.48 standard deviations above the mean height for women while men's heights need only be at least 0.98 standard deviations above the mean height for men.

Part 2: The Standard Normal Distribution

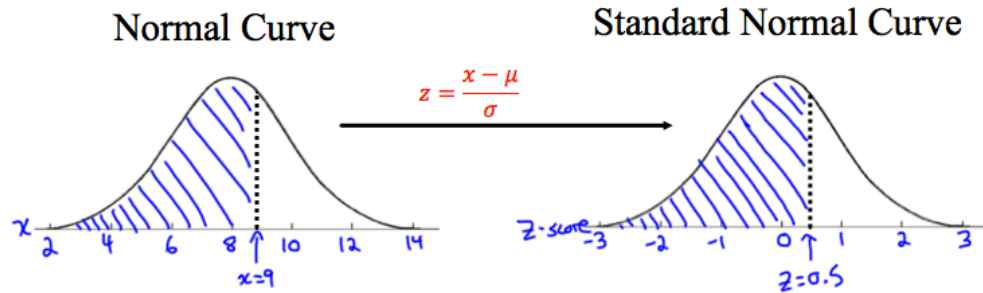
All normal distributions are the same when we measure how many standard deviations an observation x lies away from the mean, which we calculate as follows:

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

The standardized value for x , calculated by the formula above, is called its z-score. Observations from different normal distributions are best compared by comparing their standardized values, or z-scores. The z-score states how many standard deviations the original observation falls away from the mean and in which direction. Observations larger than the mean have positive z-scores, while observations smaller than the mean have negative z-scores. Converting to standardized values allows us to find proportions that we can't get from the Empirical Rule.

Example 1: Suppose we want to know the percentage of data from a normal distribution with mean $\mu = 8$ and standard deviation $\sigma = 2$ that falls below $x = 9$. We can use the Empirical Rule to learn that the percentage is between 50% and 84%, but that is not a very accurate estimate. Instead, we convert $x = 9$ into a z-score:

$$z_9 = \frac{9 - 8}{2} = 0.5$$



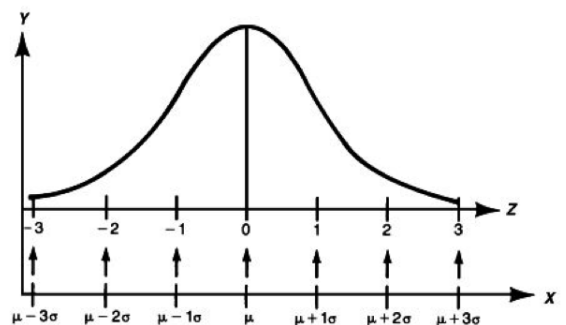
The standard normal graph shows how the value $x = 9$ is 0.5 standard deviations above the mean. This value is in the same position on this graph as it was for the normal distribution graph.

Now, we find the proportion of standard normal data that falls below 0.5. To find this proportion, we use a standard normal table, which gives the proportion of data that falls below any value for z . Using the portion of a z-table, we look down the z-column to locate 0.5 and then move to the right under the .00 column. Our answer is 0.6915, or around 69.15%.

	0.00	0.01	0.02	0.03	0.04
0.0	0.5000	0.5040	0.5080	0.5120	0.5160
0.1	0.5398	0.5438	0.5478	0.5517	0.5557
0.2	0.5793	0.5832	0.5871	0.5910	0.5948
0.3	0.6179	0.6217	0.6255	0.6293	0.6331
0.4	0.6554	0.6591	0.6628	0.6664	0.6700
0.5	0.6915	0.6950	0.6985	0.7019	0.7054
0.6	0.7257	0.7291	0.7324	0.7357	0.7389

Standardizing a data value by finding the z-score places it on the standard normal distribution. This is a special normal distribution with a mean of 0 and a standard deviation of 1.

$$X \sim N(0, 1^2)$$



Part 3: Practice Finding and Interpreting Z-Scores

Example 2: Caley scored 84% in her Data Management course, while Lauren, who attends a different Data Management class, scored 83%. If Caley's class average is 74% with a standard deviation of 8, and Lauren's class average is 70% with a standard deviation of 9.8, use z-scores to determine who has the better mark.

Note: z-scores are used to standardize the data so that they can be accurately compared.

Caley

$$z_{84} = \frac{84 - 74}{8}$$
$$= 1.25$$

Lauren

$$z_{84} = \frac{83 - 70}{9.8}$$
$$= 1.327$$

Lauren's result is 1.327 standard deviations above the mean while Caley's is 1.25 standard deviations above the mean. Lauren's result is slightly better.

Part 4: Area to Left (Percentile)

The proportion (percentage) of data with a **lower** z-score is equal to the **percentile** the value can be categorized by. The areas given on the z-score table are considered percentiles because the z-score table always gives areas to the left of the given z-score.

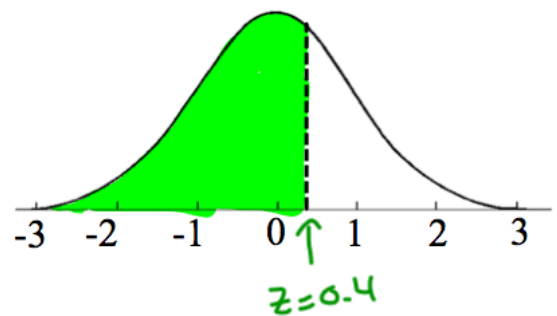
Percentile: the kth percentile is the least data value that is greater than k% of the population.

For example; If your test result is in the 85th percentile, that means that 85% of tests had a lower score than yours.

Example 3: Perch in a lake have a mean length of 20cm and a standard deviation of 5cm. Find the percent of the population that is less than 22 cm.

$$z_{22} = \frac{22 - 20}{5}$$
$$= 0.4$$

	0.00	0.01	0.02
0.0	0.5000	0.5040	0.5080
0.1	0.5398	0.5438	0.5478
0.2	0.5793	0.5832	0.5871
0.3	0.6179	0.6217	0.6255
0.4	0.6554	0.6591	0.6628
0.5	0.6915	0.6950	0.6985



From table:

Area to the left is 0.6554. This means that of the fish in the lake, 65.54% are 22cm long or less. This fish is in the 66th percentile.

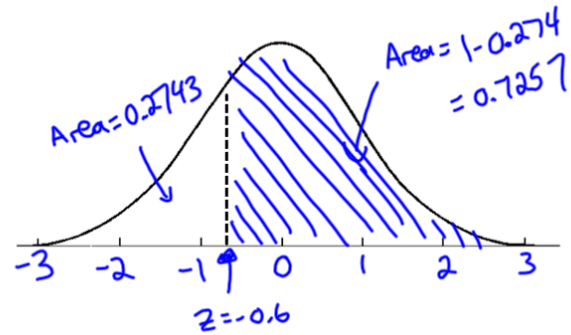
Part 5: Area to the Right

Example 4: Perch in a lake have a mean length of 20cm and a standard deviation of 5cm. Find the percent of the population that is greater than to 17 cm.

$$z_{17} = \frac{17 - 20}{5}$$

$$= -0.6$$

-0.8	0.2119	0.2090	0.2061	0.2033
-0.7	0.2420	0.2389	0.2358	0.2327
-0.6	0.2743	0.2709	0.2676	0.2643
-0.5	0.3085	0.3050	0.3015	0.2981
-0.4	0.3446	0.3409	0.3372	0.3336
-0.3	0.3821	0.3783	0.3745	0.3707



From table:

Area to the left is 0.2743.

To find the area to the right, we must use the fact that the total area under the curve is 1. Therefore, the area to the right = $1 - 0.2743 = 0.7257$. This means that about 72.57% of fish in the lake have a length greater than 17 cm.

Part 6: Area Between Two Values

Example 5: Using the normal distribution $X \sim N(7, 2.2^2)$, find the percent of data that is within the interval $3 < x < 6$

$$z_3 = \frac{3-7}{2.2}$$

$$= -1.82$$

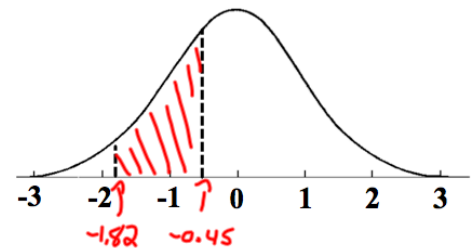
$$z_6 = \frac{6-7}{2.2}$$

$$= -0.45$$

From table:

Area to the left is 3.44%

Area to the left is 32.64%



To find the area between the two z-scores, subtract the two percentages obtained from the z-score table.

Therefore, the percent of data between 3 and 6 = $32.64\% - 3.44\% = 29.2\%$

Part 7: Percentile to Scores: z in Reverse

Finding areas from z-scores is the simplest way to work with the normal distribution. But sometimes we start with areas and are asked to work backward to find the corresponding z-score or even the original data value.

Example 6: SAT test scores have a mean of 500 and a standard deviation of 100. Suppose a college says it only admits students who have scores in at least the 90th percentile. How high a score does it take to be eligible?

From the z-score table, the z-score with an area to left as close to 0.9 as possible is 1.28.

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

$$1.28 = \frac{x - 500}{100}$$

$$x = 628$$

Therefore, it takes a score of at least 628 to be eligible to be admitted to the college.

	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

Using the Ti-84

Using ti-83 for a Normal Distribution

Access normalcdf function: 2nd → VARS (DISTR) → NORMALCDF

For a normal distribution:

normalcdf(left boundary, right boundary, mean, standard deviation)

For a standard normal distribution:

normalcdf(left boundary z-score, right boundary z-score)

Note: Use 1E99 for positive infinity and -1E99 for negative infinity

Example 3: $X \sim N(20, 5^2)$. Find percent less than 22.

```
NORMAL FLOAT AUTO REAL RADIAN MP
normalcdf
lower: -1E99
upper: 22
μ: 20
σ: 5
Paste
normalcdf(-1E99, 22, 20, 5)
.....
.6554216971
```

Example 4: $X \sim N(20, 5^2)$. Find percent greater than 17.

```
NORMAL FLOAT AUTO REAL RADIAN MP
normalcdf
lower: 17
upper: E99
μ: 20
σ: 5
Paste
normalcdf(17, E99, 20, 5)
.....
.7257469354
```

Example 5: $X \sim N(7, 2.2^2)$. Find percent between 3 and 6.

```
NORMAL FLOAT AUTO REAL RADIAN MP
normalcdf
lower: 3
upper: 6
μ: 7
σ: 2.2
Paste
normalcdf(3, 6, 7, 2.2)
.....
.2902000428
```

Example 6: $X \sim N(500, 100^2)$. Score in 90th percentile.

```
NORMAL FLOAT AUTO REAL RADIAN MP
invNorm
area: 0.9
μ: 500
σ: 100
Paste
invNorm(0.9, 500, 100)
.....
628.1551567
```