

Lesson Plan

- Discuss Quizzes/Answer Questions
- Areas Under Normal Curves
- Statistical Graphics (Maps)

4. Areas Under Normal Curves

In the last lecture, we learned how to find areas under a standard normal distribution (mean $\mu = 0$, standard deviation $sd = 1$). This required use of the table on page A-105.

A region under a normal curve corresponds to a proportion of the population. This is because a normal curve can be viewed as the limit of a series of histograms, in which the sample gets large while the bin-size goes to zero.

For example, 50% of the area under the standard normal lies to left of 0. Thus if this represents temperatures in Nome, then about half the time the temperature is below 0.

We now show how to convert a question about an arbitrary normal distribution into an equivalent question about the standard normal, and vice-versa. Thus we can use the table on A-105 to answer questions about all normal distributions, not just the standard normal.

Let X be a random value from a population with mean μ and standard deviation σ . We write this as $X \sim N(\mu, \sigma)$.

Define a new random variable $Z = \frac{X - \mu}{\sigma}$. Then one can prove that $Z \sim N(0, 1)$.

This is called the z -transformation. To go the other way, we convert the standard normal value to an arbitrary normal distribution by solving for X . So $X = \mu + Z\sigma$.

4.1 Using the z-Transformation

Reggie Jackson has an IQ of 140. What percentage of people are smarter?

Assume that IQs are normally distributed with mean 100 and standard deviation 16.

We want the area under the normal distribution for IQ that lies to the right of 140. By the z-transformation, this is equivalent to the area under the standard normal distribution that lies to the right of

$$z = \frac{X - \mu}{\sigma} = \frac{140 - 100}{16} = 2.5.$$

From the table on A-105, the area between ± 2.5 is 98.76%. Thus the area above 2.5 is $\frac{1}{2}(100 - 98.76) = .62\%$.

Now we go the other way. We find the X value that corresponds to a given percentage.

To join Mensa one must be in the top 2% of the IQ distribution. What score do you need?

On the A-105 table, look up 96%. That gives the z -value of approximately 2.05. We know that 2% of the area under the standard normal is above 2.05, and 2% is below -2.05.

Now we use the inverse z -transformation. So

$$X = \mu + Z\sigma = 100 + (2.05)(16) = 132.8.$$

One needs a score of at least 132.8 to join.

Always draw a picture. Assume heights are normally distributed with mean 5'8" and standard deviation 4". What proportion of people are shorter than 5'5"?

Assume heights are normally distributed with mean 5'8" and standard deviation 4". Then 30% of people are shorter than what?

.

4.2 Statistical Graphics (on Maps)

John Snow helped to end the 1854 cholera outbreak through use of a statistical graphic based on a city map of London. The map shows the pattern of the disease outbreak, and illustrates the importance of exception analysis.

Snow was Queen Victoria's physician and a protégé of Florence Nightingale.

He also found a smart way to estimate the literacy rate. Guess how he did it?

The second graphic shows the incidence of stomach cancer separately for men and women.

- Is there a gender difference?
- What is going on in Nevada?
- What is going on in New Mexico?
- What is going on in Wisconsin, Minnesota, and North Dakota?
- What about Pittsburgh?
- What about Maine?

How do we interpret single-county hotspots?

The third graphic shows the pedestrian fatality rates by state. Florida is the worst, and has the top five cities in the country. What might explain this (consider also New Mexico and Arizona).

How do you interpret the north-south pattern?

The fourth graphic shows the numbers of poets in the Tang and Sung dynasties. You can see that the capital shifted from Xian to Hangzhou.

The Tang dynasty was the golden age, with Li Po, Wang Wei, and Tu Fu. The Sung was the silver age—the most famous are the Su-Ma Chien family.

The fifth graphic is by Charles-Joseph Minard, an inspector of bridges and tunnels in France. It shows the wine exports in 1864.

The sixth is also by Minard, and has been hailed as the best statistical graphic ever. It shows the size of Napoleon's army in 1812-1813, as he attacks Czar Alexander III in Moscow and then retreats.

The graphic includes information on:

- location (two dimensions)
- time
- temperature
- size of the army