Understanding the Standard Normal Distribution And Z-Scores

A **Z-SCORE (z)** is the number ofstandard deviations a given value (x) falls from the mean value(μ).

On a normal distribution curve graph, Z-scores are the values on the horizontal axis.



The entire area under the curve is 100%. As a decimal, this is ‘1’. The area between any to z-values on the graph is the **proportion**, or **probability** that the population will fall between those two values.

Since the normal curve is standardized to the population, the area under the curve is also standard. The Normal Curve, illustrated below, shows that:

* Approximately **68%** of the population always lies between a **z-score of -1 to +1**
* Approximately **95%** of the population always lies between a **z-score of -2 to +2**
* Approximately **99.7%** of the population always lies between a **z-score of -3 to +3**



# An Example

An example of a standardized normal distribution is the Intelligence Quotient or ‘IQ” score.

The mean (**μ**) IQ score for the entire population is 100. The standard deviation (**σ**) is 15 points.

If you have an IQ score of 115 (100 + 15), then you are 1 standard deviation above the mean (z = +1).

If you have an IQ score of 85 (100 – 15), then you are 1 standard deviation below the mean (z = -1).

To put this in perspective, we need to know the area under the curve between the IQ scores of 85 and 115. This is the area under the standard normal curve between z = -1 and z = +1. Therefore, we know that 68% of the population have and IQ score between 85 and 115.

Similarly, if we look at the scores that are 2 deviations away from the mean (z = +/-2), we would be looking at people with IQ scores between 70 and 130. We know that 95% of the population lies between these z scores which tells us that 95% of the people in the world have an IQ score between 70 and 130.

Any score within 2 standard deviations of the mean is considered “normal”. Outside of these values, statisticians consider the scores to be “extreme” or different from the normal population. Only 5% of the population will have IQ scores above 130 or below 70. In statistics, these scores are considered outside of the normal range.



Any IQ score above 145 or below 55 (z ≥ +/-3) is even more extreme with less than 1% of the population having those scores.

# Not All Z-Scores Are 1, 2 Or 3

What if a person has an IQ score of 110? How many standard deviations are they from the mean?

To calculate this, we use the z-score formula:



Z = $\frac{110-100}{15}=0.67$

An IQ score of 110 is 0.67 standard deviations away from the mean.

The probability of having an IQ score of 110 or **less** is the percentage of the area under the normal curve up to a z-score of 0.67 (shaded).



In order to find this probability, we need to use a z-score chart which provides the area to the left of any z-score.



The z-score chart indicates that P(X=110) = .7486 (the area under the curve up to the z-score of .67 is 74.86%).

The probability of having an IQ score of 110 or **higher** is the percentage of the area under the normal curve up to a z-score of 0.67 (shown in blue). To find this probability we subtract 0.7486 from 1 (the compliment). 1 - .7486 = 0.2514. Therefore P[X≥110] = .2514



# How Does ‘R’ Help Us Make These Calculations Easier?

The program “R” will calculate z-scores and probabilities for us without the use of a z-score chart. The command “**pnorm**” will calculate the probability for any raw value, x.

## Example 1

For example, if we want the probability of someone having an IQ score of 110 or less (just like in the last example), we would open the console in “R” and type;

**pnorm(110, 100, 15, lower.tail=TRUE)**

In general terms, **pnorm(x, μ, σ, lower.tail=TRUE)** will always calculate the area to the **left** of any z-score.



**> pnorm(110, 100, 15,lower.tail=TRUE)**

**[1] 0.7475075**

Note: this probability is slightly lower than the one we calculated from the chart because “R” is able to work with more decimal places.

## Example 2

If we want to find the probability of someone having an IQ score of 110 or higher, we would type;

**pnorm(110, 100, 15, lower.tail=FALSE)**

In general terms, **pnorm(x, μ, σ, left.tail=FALSE)** will always calculate the area to the **right** of any z-score.



**> pnorm(110, 100, 15,lower.tail=FALSE)**

**[1] 0.2524925**

## Example 3

If we want to find the probability of someone having an IQ score between two values (for example between 105 and 120), we need to preform two “pnorm” calculations and subtract the results. We would type;

**pnorm(120, 100, 15, lower.tail=TRUE)** - **pnorm(105, 100, 15, lower.tail=TRUE)**



**>pnorm(120, 100, 15, lower.tail=TRUE)-pnorm(105,100,15,lower.tail=TRUE)**

**[1] 0.2782301**

**qnorm(.8, 100, 15)** will calculate the score (x) that a person would have to get to be in the top 80% of the population. For someone having an IQ score of 110 or greater, In general terms, this is **qnorm(percentile, μ, σ).** This will always calculate the raw score if you know the probability.

# Finding a Z-Score When You Are Given a Probability

Use **qnorm(p)** to find a z-score when given a specific probability. For example, if P(X) = .5645, the z-score is calculated using **qnorm(.5645)**

**> qnorm(.5645)**

**[1] 0.1623884**

**Therefore, z = 0.16**

# Finding a Probability When You Know a Z-Score

Use **pnorm(z)** to find the probability when you know just the z-score. For example, for a z-score of 1.5, P[x≤1.5], the probability is

**> pnorm(1.5,lower.tail=TRUE)**

**[1] 0.9331928**

**Therefore, p = 0.93 (or 93%)**