THE NORMAL DISTRIBUTION METHOD ARTICLE NO.: 10080

By Ken Standfield, Director Research & Development, KNOWCORP
http://www.knowcorp.com *Email:* ks@knowcorp.com

INTRODUCTION

The following methods have been developed over the years that I lectured statistics. What follows is not a new discovery, in itself, but rather a strict ordering in what has remained a largely unstructured area. It is a method designed to provide a framework through which practically every normal distribution problem can be analyzed. I would like to extend my thanks to my friend and colleague Dr Roger Wallace at Deakin University.

I have found that students find difficulty understanding the concept of 'areas under the standard normal curve'. This problem can be solved by reading article 10150: *Understanding Areas Under Curves - ND & CLT*. Statistical formulae and statistical symbols have also caused students considerable confusion, articles 10060: *Understanding Statistical Formulae* and Article 10050: *Explaining Statistical Symbols* should clear up any problems.

THE NORMAL DISTRIBUTION METHOD

We solve normal distribution problems by translating the normal distribution into a standard normal distribution through use of the following formula:

$$z_x = \frac{x - m_x}{s_x}$$

This formula is a methodological formulae (see Article 10060: *Understanding Statistical Formulae*), and hence it has a definite method of solution

TYPES OF PROBLEMS

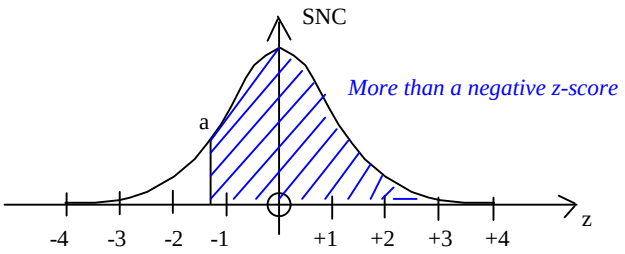
There are two basic normal distribution problems: (1) *more* or *less* than problems and (2) *between* problems. We will investigate each in turn.

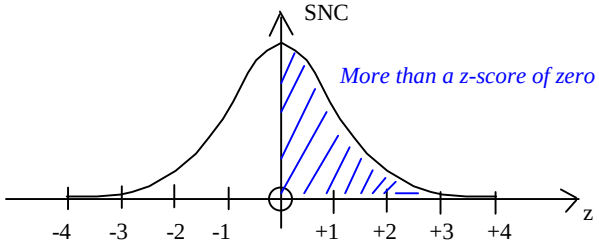
MORE THAN OR LESS THAN PROBLEMS

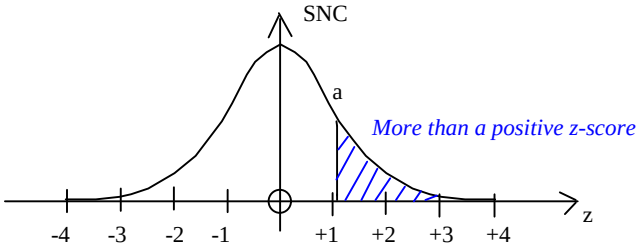
It is very easy to identify more or less than problems. For *more than* problems, look for the words "more than" or "increases" or "exceeds" or other words, in the question, that have the same meaning as *more than*. For *less than* problems, look for the words "less than" or "decreases" or other words, in the question, that have the same meaning as *less than*. Before we look at the specific method, we will investigate the two categories (more and less than) broadly.

1.1 "MORE THAN" PROBLEM TYPES

There are three categories that *more than* problems can fall into:

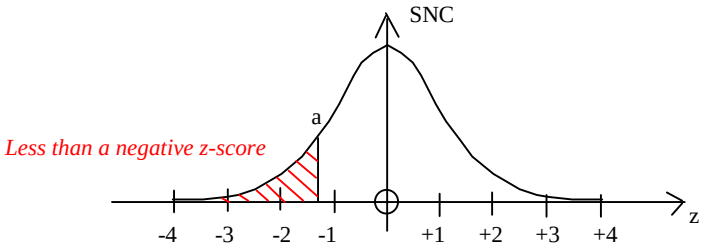


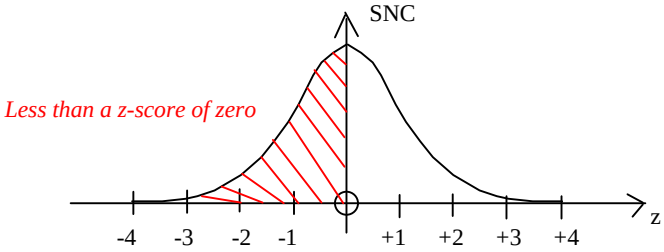


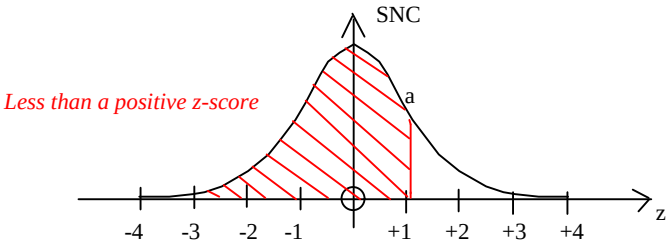


2.1 "LESS THAN" PROBLEM TYPES

Less than problems can fall into three categories:







3.1 THE MORE THAN METHOD

The following represents the more than method. The method is divided into four sections:

(1) Basic Mathematics
(2) Drawing the Correct Curve
(3) Work out the Appropriate Area
(4) Conclusion

3.1.1 SECTION #1 – BASIC MATHEMATICS

$$P_N = (x > a_x) \quad = \quad P_{SN}(z_x > \frac{x - m_x}{s_x}) \quad = \quad P_{SN}(z_x > \frac{a_x - m_x}{s_x})$$

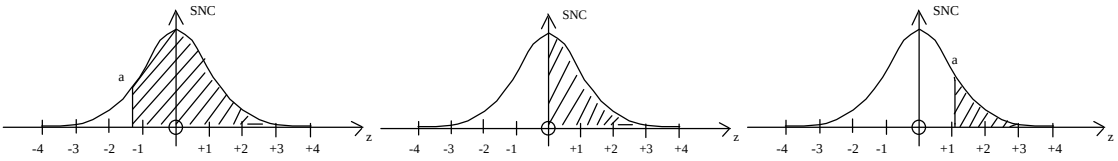
$$= \quad P_{SN}(z_x > \frac{Simp}{s_x})$$

$$= \quad P_{SN}(z_x > a_z)$$

Let's translate the above section of the method: (1) determine the x value, (2) proceed to translate the x-value into a standard normal value by applying the standard normal curve formula; substitute the value for the value of x; simplify (Simp) the numerator of the standard normal formula; simplify the whole fraction to arrive at the z-value of x. If terminology like numerator and denominator confuses you, refer to Article 10050: *Explaining Statistical Formulae*, it will clear up any problems.

3.1.2 SECTION #2 – DRAWING THE CORRECT CURVE

This is a three-stage process. Step 1 is to analyze the value of z (positive, zero, or negative). Step 2 involves marking in the value of z and step 3 involves shading in the correct direction and shading the correct distance, for example:



Shading from a negative z-score	Shading from a zero z-score	Shading from positive z-score

The appropriately shaded diagram is integrated into stage #1 - you will see how later.

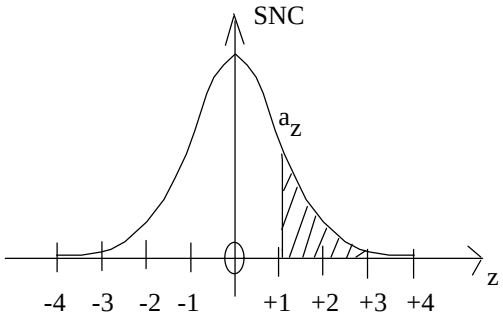
3.1.3 SECTION #3 – WORK OUT THE APPROPRIATE AREA

Immediately under the appropriate diagram (determined from Stage #2), we start determining the appropriate value. We do this by drawing our conclusions together in the following format:

$$P_N = (x > a_x) \;=\; P_{SN}(z_x > \frac{x - m_x}{s_x}) \;=\; P_{SN}(z_x > \frac{a_x - m_x}{s_x})$$

$$= \; P_{SN}(z_x > \frac{Simp}{s_x})$$

$$= \; P_{SN}(z_x > a_z)$$



Hence:

$$P_N = (x > a_x) \;=\; P_{SN}(z_x > a_z) \;=\; 0.5000 \; - f(z_{a_z})$$

$$= \; 0.5000 \; - \text{det} \; / \text{d}$$

$$= \; \text{det} \; / \text{d}$$

Article 10150: *Understanding Areas under Curves - ND & CLT*, shows how to determine the area under any curve - which is adding or subtracting 0.5000 from z-values (etc). It also explains phi-notation - ϕ. The *det/d* part of the solution above indicates that this is a value that has been *determined*.

3.1.4 SECTION #4 – CONCLUSION

You have determined you answer. To ensure that you maintain a completely logical approach to solving problems, you need to draw together your findings using words and numbers, not just numbers.

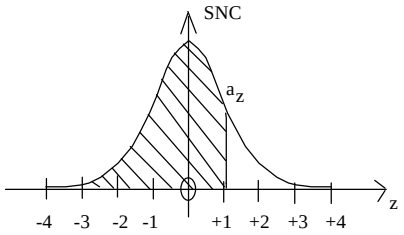
4.1 THE LESS THAN METHOD

We have investigated the *more than method*, now we shall attention to the *less than method.* The *less than method* is identical to the *more than method*, we only change a few symbols around.

$$P_N = (x < a_x) \;=\; P_{SN}(z_x < \frac{x - m_x}{s_x}) \;=\; P_{SN}(z_x < \frac{a_x - m_x}{s_x})$$

$$= \; P_{SN}(z_x < \frac{Simp}{s_x})$$

$$= \; P_{SN}(z_x < a_z)$$



Hence,

$$P_N = (x < a_x) \;=\; P_{SN}(z_x < a_z) \;=\; 0.5000 + f(z_{a_z})$$

$$= \; 0.5000 + \text{det} / \text{d}$$

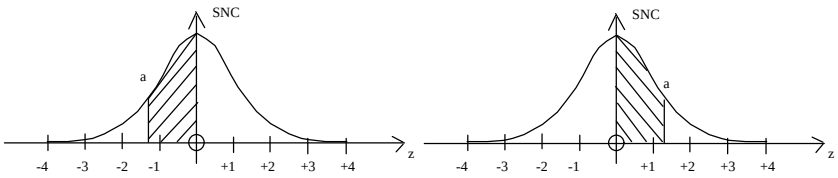
$$= \; \text{det} / \text{d}$$

Conclusion: Put the mathematics into words *and* answer the question.

5 ANALYZING BETWEEN PROBLEMS

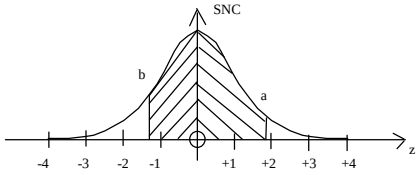
The other major type of problem involves "between" problems. These problems are composed of a more than problem and a less than problem put together to form a between problem. Hence between problems are recognized easily - two z-values must be determined.

5.1 BETWEEN A NUMBER AND ZERO.



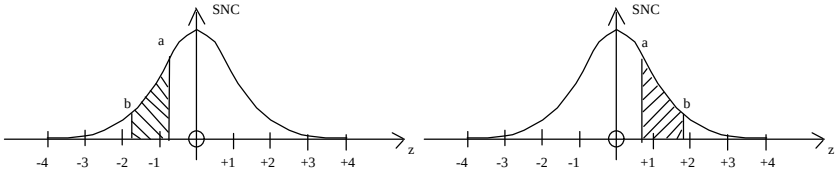
These problems involve problems that seek us to determine the area (probability) between the mean and some other number.

5.2 BETWEEN TWO NUMBERS ON DIFFERENT SIDES



5.3 BETWEEN TWO NUMBERS ON THE SAME SIDE

When we have two numbers on the same side, we must decompose the problem



Note that Article 10150: *Understanding Areas under Curves - ND & CLT*, shows the solutions to a variety of shading problems.

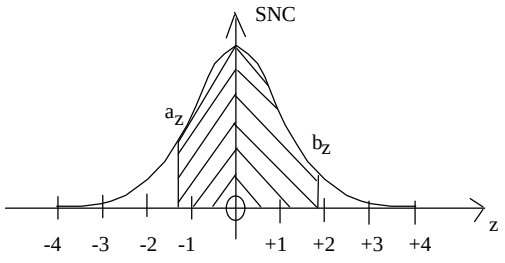
THE BETWEEN METHOD

The following is the general format used when solving between problems:

$$P_N = (a_x < x < b_x) = P_{SN}\left(\frac{x_1 - m_x}{s_x} < z_x < \frac{x_2 - m_x}{s_x}\right) = P_{SN}\left(\frac{a_x - m_x}{s_x} < z_x < \frac{b_x - m_x}{s_x}\right)$$

$$= P_{SN}\left(\frac{Simp}{s_x} < z_x < \frac{Simp}{s_x}\right)$$

$$= P_{SN}\left(a_z < z_x < b_z\right)$$



Hence,

$$P_N = (a_x < x < b_x) = P_{SN}(a_z < z_x < b_z) = f(z_{a_z}) + f(z_{b_z})$$

$$= \text{det} / \text{d} + \text{det} / \text{d}$$

$$= \text{det} / \text{d}$$

CONCLUSION

Put the mathematics into words *and* answer the question. If between terminology (a < x < b or a < z < b) causes you confusion, look at Article 10050: *Explaining Statistical Symbols*, it should clear up the confusion.

ABOUT THE AUTHOR

See http://www.knowcorp.com/developer.html for more information.