The Z-Tables

Suppose you need to find **probability** of occurrence of a specific mineral grade or grade interval, and you know its mean and standard deviation, what to do?

1. Construct a CDF
2. Go to Z-Tables

What is the Z variable?

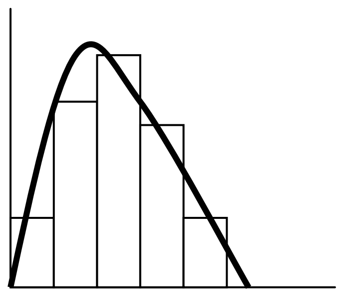
It is a **standardized normally distributed** variable with m = 0 and V = 1 (i.e. **N(0,1)**)and it has no units.

$$Z = \frac{x - m}{S}$$

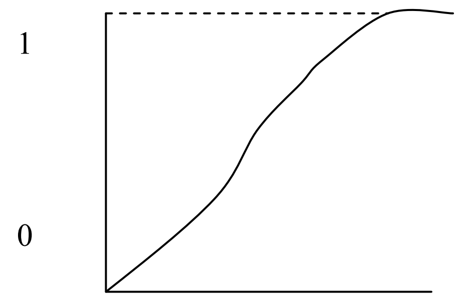
Why is it used?

We standardize values of variable x so that they come from the standard normal distribution because it is not possible to produce different probability tables for different variables and values.

What is the area under the PDF curve?



PDF

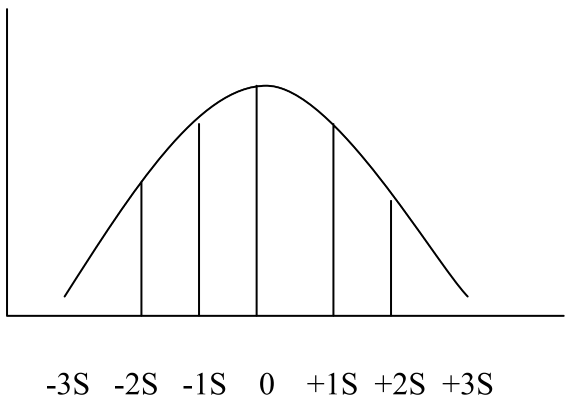


CDF

$$CDF = \int PDF$$

The area under the curve is always equal to **ONE**.

In a **normal distribution**:



68.5% of data lie within + - 1S
95.5% of data lie within + - 2S
99.7% of data lie within + - 3S

How to calculate probability using Z-Tables?

Example: Suppose the mean concentration of lead (Pb) = 367 ppm and its V = 32400 ppm^2. Calculate the probability of having Pb <= 450 ppm

Solution: m = 367 ppm, S = 180 ppm, x = 450 ppm
Therefore,

$$Z = \frac{450 - 367}{180} = 0.46$$

P (x <= 450 ppm) = P (Z <= 0.46) = 0.68 = 68%

P (x > 450 ppm) = 1 - P (Z <= 0.46) = 1 – 0.68 = 0.32 = 32 %

Point and Interval Estimates

An **estimate** of a population is expressed by a **single value (i.e. point value)** like mean of a mineral grade, porosity, mean of quartz in granite, mean of sulfur content in a volcanic gas … etc.

However a **single value** is statistically **meaningless**. To be more meaningful, an interval with upper and lower limits should be associated with it. This interval is called **Confidence Interval (C.I.)** and its limits are called **Confidence Limits (C.L.).**

To calculate **C.I.**, either **Z-Distribution** or **t-Distribution** should be used depending on the available number of samples **(n)**

Student's t-Distribution

t-Distribution was discovered by W. S. Gossett who published his work under the name "student".

It is used to define **Confidence Intervals (CI)** to the estimated mean of a population.

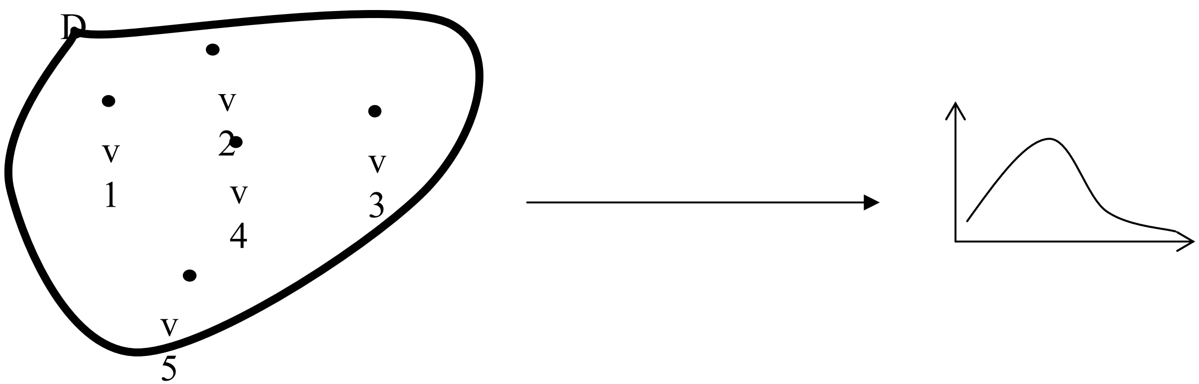
t-Distribution is used if the number of available data samples is less than 30 (i.e. **$n<30$**). However **Z-distribution** is used to assign CI to the estimated mean if **$n>=30$**.

Both distributions follow the normal distribution assumption.

Both distribution, in case of defining the C.I. satisfy the requirements of the **Central Limit Theory**

Central Limit Theory

If a random sample size n is taken from a **normal** distribution with mean **m** and standard deviation **s**, the distribution of sample means will also be **normal** with mean **m** and standard deviation **s/sqrt (n).**



<u>If n >= 30 samples</u>

$$m = \left(\overline{x}\right) \pm Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

<u>If n < 30 samples</u>

$$m = \left(\overline{x}\right) \pm t_{(\frac{\alpha}{2},\nu)} \frac{s}{\sqrt{n}}$$