The basics of probability – Part I

Introduction

In assessing a reported measurement on a sample of material or product against compliance to specification or regulation limits, we need to consider two important concepts: "Measurement uncertainty" and "Measurement decision risk".

For such decision risk, we have to prescribe the acceptance or rejection of the product or test parameter based on the measurement result, its uncertainty and the specification limit or limits, taking into account of the acceptable level of the probability of making a wrong decision. Indeed, there are two types of risk involved in making a decision, i.e. false accept risk and false reject risk.

A *false accept risk* (sometimes called "customer's risk") is the probability of an out-of-specification item or parameter being unknowingly accepted by test or calibration, whilst a *false reject risk* (also known as "measurement risk") is the probability of an in-specification item or parameter being unknowingly rejected by testing and calibration.

Therefore, measurement decision risk uses probability expressions.

We need to revise our basic knowledge of probability, and the concepts of joint and conditional probability before we link them in risk analysis.

What is a probability?

A probability is interpreted as a *relative frequency* or occurrence in an infinitely large number of events (experiments).

Hence, probability theory tries to describe the variation or randomness within order so that underlying order may be better understood. Once understood, we can more effectively formulate strategies and evaluate their risks involved.

Fundamentally we enumerate all the possible events or outcomes which are possible within a given context. For example, by flipping a coin, the possible outcomes are "heads" and "tails". If we roll a die, we do not get the same number every time as there are six possible outcomes, all equally probable.

In the simplest situation, we can register whether or not a random event occurs, for example, if a single die shows an even pip number, denoted as E and let N_E be the number of occurrence of E out of N rolls, then N_E/N is the relative frequency of the event E.

Statistically, as there are 3 possible even numbers of pips in a single die, we estimate the probability of obtaining an even number to be:

$$P(E) = N_E / N = 3/6 = 0.5$$

The relative frequency stabilizes as N increases. This probability estimate becomes the limit of the actual probability of occurrence as the number N approaches infinity:

$$P(E) = \lim_{N \to \infty} N_E / N$$

Such relation is called the *law of large numbers*.

Outcomes, events and probabilities

We denote U as the set of all possible outcomes from an experiment, called the *sample space*. A possible outcome, u, is an element in the sample space, so we say $u \in U$. The possible outcomes when we throw a die are 1, 2, … , 6, so the sample space is the set $U = \{1, 2, \ldots, 6\}$.

An event, A, is a subset of the sample space, and we write $A \subseteq U$. We say that the event A has occurred if the outcome of the experiment belongs to the set A.

For example, if we are interested in the event of rolling an odd number of pips on a single die, then the event A is the set {1,3,5}.

The empty set ∅ is an event that never occurs, whilst the entire sample space, U, is event that always occur.

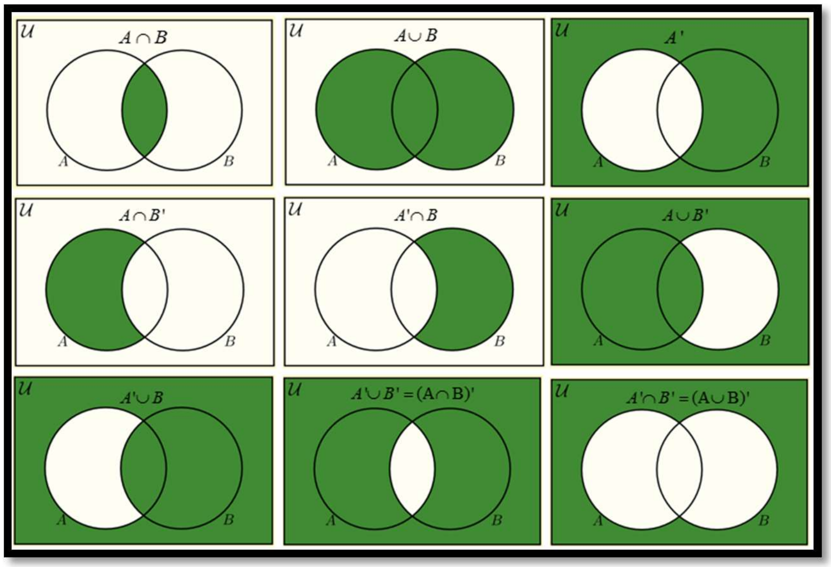
Let's use set theory to combine two or more sets. If A and B are events, then we can have the following situations:

- Intersection: $A \cap B$ is the event that both A and B occur
- Union: $A \cup B$ is the event that either A or B or both occur
- Complement: A^c is the event that A does not occur

A is said to imply B if A is a subset of B ($A \subseteq B$) because B must occur if A occurs.

A and B are said to be disjoint if $A \cap B = \emptyset$. That means that they cannot both occur at once.

The figure below shows the various relationships between two events, A and B.



Example 1:

If we throw a regular die, the sample space is U = {1,2,3,4,5,6}. Let the two events "odd number of pips" and "at least 5 pips" be:

$$A = \{1,3,5\}, \text{ and } B = \{5,6\}$$

For these two events, we have that:

$A \cap B = \{5\}$; $A \cup B = \{,1,3,5,6\}$

Set difference $A \setminus B = \{1,3\}$; $A^c = \{2,4,6\}$

If N_A and N_B are the frequencies of events A and B each occurring out of N experiments, then we have:

- $0 \leq N_A/N \leq 1$, because $0 \leq N_A \leq N$;
- $N_U/N = 1$, because $N_U = N$
- $N_{A \cup B}/N = N_A/N + N_B/N$ if A and B are disjoint (independent) (i.e., $A \cap B = \emptyset$), since $A \cup B$ occurs $N_A + N_B$ times when A and B cannot both occur at the same time.