THE
SIX SIGMA
GREEN BELT
PRIMER

THERE IS ALWAYS A 100% PROBABILITY THAT A PIECE OF TOAST WILL LAND BUTTERED SIDE DOWN ON NEW CARPET.

FROM “MURPHY'S LAWS”

Probability and Statistics

Probability is described in the following topic areas:

- **Probability and statistics**
  - **Drawing valid statistical conclusions**
  - **Central limit theorem**
  - **Basic probability concepts**
- **Probability distributions**

Basic Statistical Terms

Continuous Distributions	Distributions containing infinite (variable) data points that may be displayed on a continuous measurement scale. Examples: normal, uniform, exponential, and Weibull distributions.
Discrete Distributions	Distributions resulting from countable (attribute) data that has a finite number of possible values. Examples: binomial, Poisson, and hypergeometric distributions.
Decision Distributions	Distribution used to make decisions and construct confidence intervals. Examples: t, F, and chi-square distributions.
Parameter	The true numeric population value, often unknown, estimated by a statistic.
Population	All possible observations of similar items from which a sample is drawn.
Sample	A randomly selected set of units or items drawn from a population.
Statistic	A numerical data value taken from a sample that may be used to make an inference about a population.

(Omdahl, 1997)[6]

Drawing Valid Statistical Conclusions *

Analytical (Inferential) Studies

The objective of statistical inference is to draw conclusions about population characteristics based on the information contained in a sample. Statistical inference in a practical situation contains two elements: (1) the inference and (2) a measure of its validity. The steps involved in statistical inference are:

- Define the problem objective precisely
- Decide if the problem will be evaluated by a one tail or two tail test
- Formulate a null hypothesis and an alternate hypothesis
- Select a test distribution and a critical value of the test statistic reflecting the degree of uncertainty that can be tolerated (the alpha, α, risk)
- Calculate a test statistic value from the sample information
- Make an inference about the population by comparing the calculated value to the critical value. This step determines if the null hypothesis is to be rejected. If the null is rejected, the alternate must be accepted.
- Communicate the findings to interested parties

Everyday, in our personal and professional lives, we are faced with decisions between choice A or choice B. In most situations, relevant information is available; but it may be presented in a form that is difficult to digest. Quite often, the data seems inconsistent or contradictory. In these situations, an intuitive decision may be little more than an outright guess.

While most people feel their intuitive powers are quite good, the fact is that decisions made on gut-feeling are often wrong. The student should be aware that the subjects of null hypothesis and types of errors are previewed in Primer Section IX.

---

* A substantial portion of the material throughout this Section comes from the *CQE Primer* by Wortman (2005)[9].

Drawing Valid Statistical Conclusions (Continued)

Enumeration (Descriptive) Studies

Enumerative data is data that can be counted. For example: the classification of things, the classification of people into intervals of income, age or health. A census is an enumerative collection and study. Useful tools for tests of hypothesis conducted on enumerative data are the chi-square, binomial and Poisson distributions.

Deming, in 1975, defined a contrast between enumeration and analysis:

Enumerative study	A study in which action will be taken on the universe.
Analytic study	A study in which action will be taken on a process to improve performance in the future.

Descriptive Statistics

Numerical, descriptive measures create a mental picture of a set of data. These measures calculated from a sample are numerical, descriptive measures called statistics. When these measures describe a population, they are called parameters.

Measures	Statistics	Parameters
Mean	$\bar{X}$	μ
Standard Deviation	s	σ

Table 7.1 Statistics and Parameters

Table 7.1 shows examples of statistics and parameters for the mean and standard deviation. These two important measures are called central tendency and dispersion.

Summary of Analytical and Enumerative Studies

Analytical studies start with the hypothesis statement made about population parameters. A sample statistic is then used to test the hypothesis and either reject, or fail to reject, the null hypothesis. At a stated level of confidence, one should then be able to make inferences about the population.

Central Limit Theorem

If a random variable X has mean μ and finite variance σ^2, as n increases, $\bar{X}$ approaches a normal distribution with mean μ and variance $\sigma^2_{\bar{X}}$. Where, $\sigma^2_{\bar{X}} = \frac{\sigma^2_X}{n}$ and n is the number of observations on which each mean is based.

Sampling Distribution of the Mean

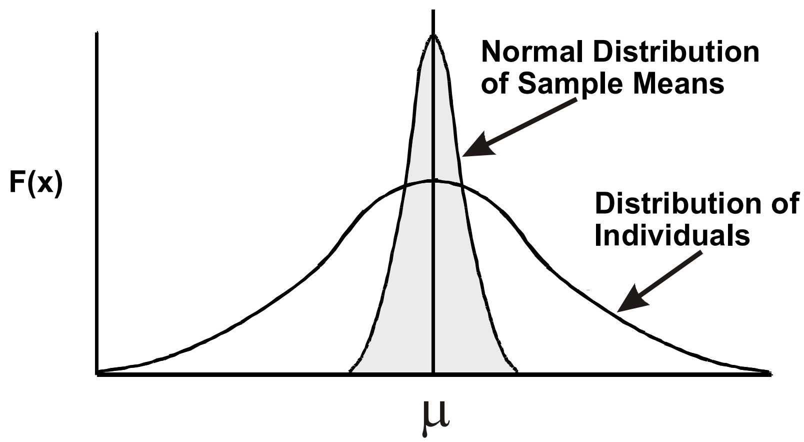


Figure 7.2 Distributions of Individuals Versus Means

The Central Limit Theorem States:

- **The sample means ($\bar{X}$s) will be more normally distributed around μ than individual readings (Xs). The distribution of sample means approaches normal, regardless of the shape of the parent population. This is why $\bar{X}$ - R control charts work!**

- **The spread in sample means ($\bar{X}$s) is less than Xs with the standard deviation of $\bar{X}$s equal to the standard deviation of the population (individuals) divided by the square root of the sample size. $\sigma_{\bar{X}}$ is referred to as the standard error of the mean:**

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}} \quad \text{Which is estimated by } s_{\bar{X}} = \frac{s_X}{\sqrt{n}}$$

Example 7.1: Assume the following are weight variation results: $\bar{X}$ = 20 grams and s = 0.124 grams. Estimate $\sigma_{\bar{X}}$ for a sample size of 4:

Solution: $s_{\bar{X}} = \frac{s_X}{\sqrt{n}} = \frac{0.124}{\sqrt{4}} = 0.062$ **grams**

Central Limit Theorem (Continued)

The significance of the central limit theorem on control charts is that the distribution of sample means approaches a normal distribution. Refer to Figure 7.3 below:

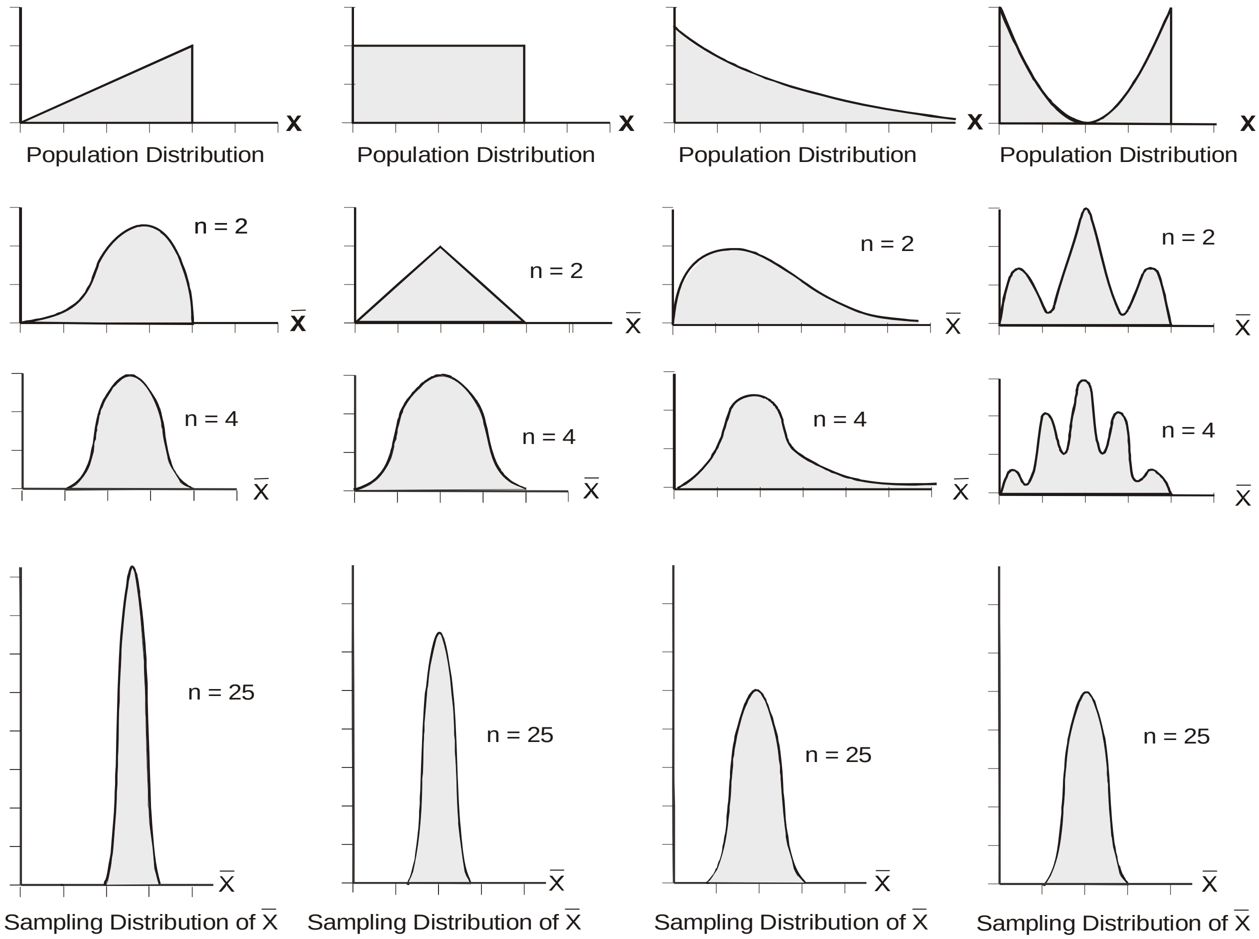


Figure 7.3 Illustration of Central Tendency

(Lapin, 1982)[5]

In Figure 7.3, a variety of population distributions approach normality for the sampling distribution of $\bar{X}$ as n increases. For most distributions (but not all), a near normal sampling distribution is attained with a sample size of 4 or 5.

Probability

Most quality theories use statistics to make inferences about a population based on information contained in samples. The mechanism one uses to make these inferences is probability. For a more expansive treatment of probability, see Triola (1994)[8] referenced at the end of this Section.

The probability of any event (E) lies between 0 and 1. The sum of the probabilities of all possible events (E) in a sample space (S) = 1. The ratio of the chances favoring an event to the total number of chances for and against the event. Probability (P) is always a ratio.

$$P = \frac{\text{Chances Favoring}}{\text{Chances Favoring Plus Chances Not Favoring}}$$

Simple Events

An event that cannot be decomposed is a simple event (E). The set of all sample points for an experiment is called the sample space (S).

If an experiment is repeated a large number of times, (N), and the event (E) is observed n_E times, the probability of E is approximately:

$$PE \approx \frac{n_E}{N}$$

Example 7.2: The probability of observing 3 on the toss of a single die is:

$$PE_3 = \frac{1}{6}$$

Example 7.3: What is the probability of getting 1, 2, 3, 4, 5, or 6 by throwing a die?

$$PE_T = P(E_1) + P(E_2) + P(E_3) + P(E_4) + P(E_5) + P(E_6)$$

$$PE_T = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$$

Compound Events

Compound events are formed by a composition of two or more events. They consist of more than one point in the sample space. For example, if two dice are tossed, what is the probability of getting an 8? A die and a coin are tossed. What is the probability of getting a 4 and tail? The two most important probability theorems are the additive and multiplicative (covered later in this Section). For the following discussion, E_A = A and E_B = B.

I. **Composition.** Consists of two possibilities -- a union or intersection.

A. Union of A and B.

If A and B are two events in a sample space (S), the union of A and B ($A \cup B$) contains all sample points in event A or B or both.

Example 7.4: In the die toss of Example 7.3, consider the following:

If A = E_1, E_2 and E_3 (numbers less than 4)
and B = E_1, E_3 and E_5 (odd numbers), then $A \cup B$ = E_1, E_2, E_3 and E_5.

B. Intersection of A and B.
If A and B are two events in a sample space (S), the intersection of A and B ($A \cap B$) is composed of all sample points that are in both A and B.

Example 7.5: Refer to Example 7.4. $A \cap B$ = E_1 and E_3

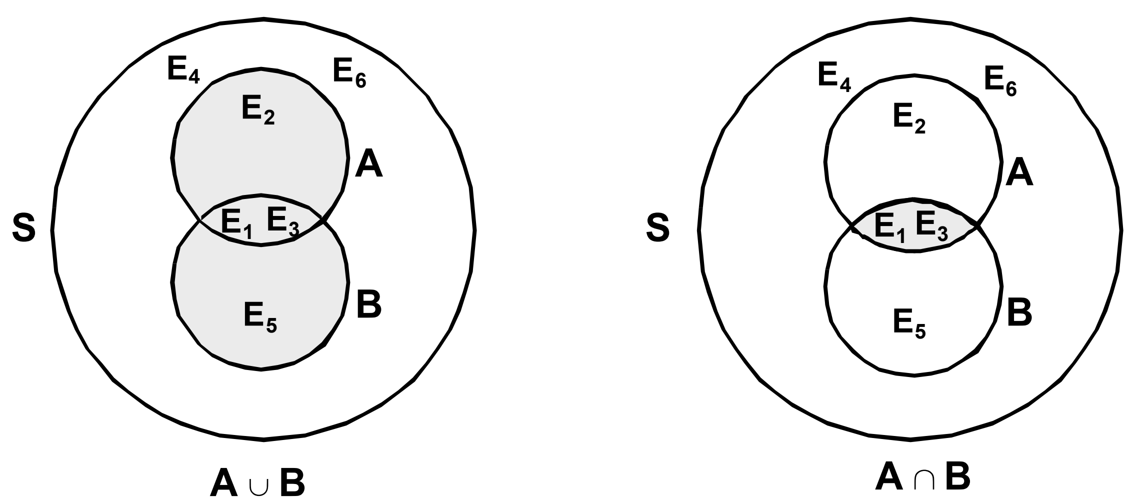


Figure 7.4 Venn Diagrams Illustrating Union and Intersection

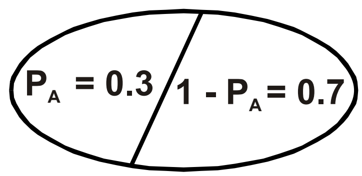
Compound Events (Continued)

II. Event Relationships. There are three relationships in finding the probability of an event: complementary, conditional and mutually exclusive.

A. Complement of an Event.

The complement of an event A is all sample points in the sample space (S), but not in A. The complement of A is $1\text{-}P_A$.

Example 7.6: If P_A (cloudy days) is 0.3, the complement of A would be $1 - P_A = 0.7$ (clear).



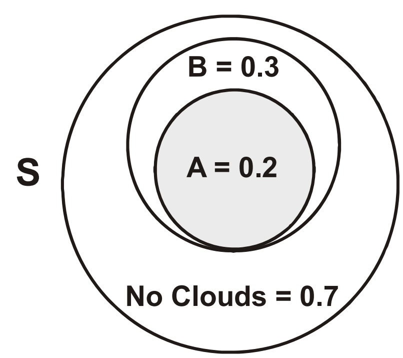
B. Conditional Probabilities.

The conditional probability of event A, given that B has occurred, is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{if } P(B) \neq 0$$

Example 7.7: If event A (rain) = 0.2, and event B (cloudiness) = 0.3, what is the probability of rain on a cloudy day? (Note, it will not rain without clouds.)

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.3} = 0.67$$



Event A and B are said to be independent if either:

$P(A|B) = P(A)$ or $P(B|A) = P(B)$ However,
$P(A|B) = 0.67$ and $P(A) = 0.2$ = no equality, and
$P(B|A) = 1.00$ and $P(B) = 0.3$ = no equality

Therefore, the events are said to be dependent.

Compound Events (Continued)

C. Mutually Exclusive Events.

If event A contains no sample points in common with event B, then they are said to be mutually exclusive.

Example 7.8: Obtaining a 3 and a 2 on the toss of a single die is a mutually exclusive event. The probability of observing both events simultaneously is zero. The probability of obtaining either a 3 or a 2 is:

$$PE_2 + PE_3 = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

D. Testing for Event Relationships.

Example 7.9: Refer to Example 7.4.

Event A: E_1, E_2, E_3
Event B: E_1, E_3, E_5

Are A and B mutually exclusive, complementary, independent, or dependent? A and B contain two sample points in common so they are not mutually exclusive. They are not complementary because B does not contain all points in S that are not in A.

To determine if they are independent requires a check.

Does $P(A|B) = P(A)$?

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2/6}{1/2} = \frac{2}{3} \qquad P(A) = \frac{1}{2}$$

Therefore $P(A|B) \neq P(A)$

By definition, events A and B are dependent.

The Additive Law

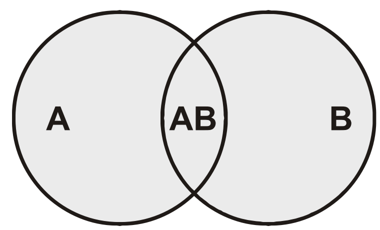
If the two events are not mutually exclusive:

1. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Note that $P(A \cup B)$ is shown in many texts as $P(A + B)$ and is read as the probability of A or B.

Example 7.10: If one owns two cars and the probability of each car starting on a cold morning is 0.7, what is the probability of getting to work?

$$P(A \cup B) = 0.7 + 0.7 - (0.7 \times 0.7)$$
$$= 1.4 - 0.49$$
$$= 0.91 = 91\ \%$$

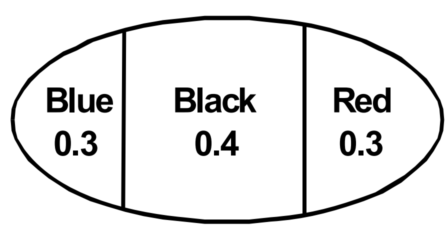


If the two events are mutually exclusive, the law reduces to:

2. $P(A \cup B) = P(A) + P(B)$ also $P(A + B) = P(A) + P(B)$

Example 7.11: If the probability of finding a black sock in a dark room is 0.4 and the probability of finding a blue sock is 0.3, what is the chance of finding a blue or black sock?

$$P(A \cup B) = 0.4 + 0.3 = 0.7 = 70\ \%$$



Note: The problem statements center around the word "or." Will car A <u>or</u> B start? Will one get a black sock <u>or</u> blue sock?

The Multiplicative Law

If events A and B are dependent, the probability of A influences the probability of B. This is known as conditional probability and the sample space is reduced.

For any two events A and B such that P(B) ≠ 0,

1. $P(A|B) = \frac{P(A \cap B)}{P(B)}$ and $P(A \cap B) = P(A|B)P(B)$

Note, in some texts P(A ∩ B) is shown as P(A • B) and is read as the probability of A and B. P(B|A) is read as the probability of B given that A has occurred.

Example 7.12: If a shipment of 100 T.V. sets contains 30 defective units and two samples are obtained, what is probability of finding both defective? (Event A is the first sample and the sample space is reduced, and event B is the second sample.)

$$P(A \cap B) = \frac{30}{100} \text{ x } \frac{29}{99} = \frac{870}{9900} = 0.088$$
$$P(A \cap B) = 8.8\ \%$$

If events A and B are independent:

2. $P\ (A \cap B) = P(A) \text{ X } P(B)$

Example 7.13: One relay in an electric circuit has a probability of working equal to 0.9. Another relay in series with the first has a chance of 0.8. What's the probability that the circuit will work?

$$P\ (A \cap B) = 0.9 \text{ X } 0.8 = 0.72$$
$$P\ (A \cap B) = 72\%$$

Note: The problems center around the word "and." Will T.V. A <u>and</u> B work? Will relay A <u>and</u> B operate?

