VII. MEASURE - PROBABILITY

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. Examples: normal, uniform, exponential, and Weibull distributions.**
Discrete Distributions	**Distributions resulting from countable (attribute) data that has a finite number of 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.**

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 it will be evaluated by a one or two tail test
- Formulate a null 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 from the sample
- Compare the calculated value to the critical value and determine if the null hypothesis is to be rejected. If the null is rejected, the alternate must be accepted.

Drawing Valid Conclusions (Continued)

Enumeration (Descriptive) Studies

Enumerative data is data that can be counted. Useful tools for tests of hypothesis conducted on enumerative data are the chi-square, binomial and Poisson distributions.

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

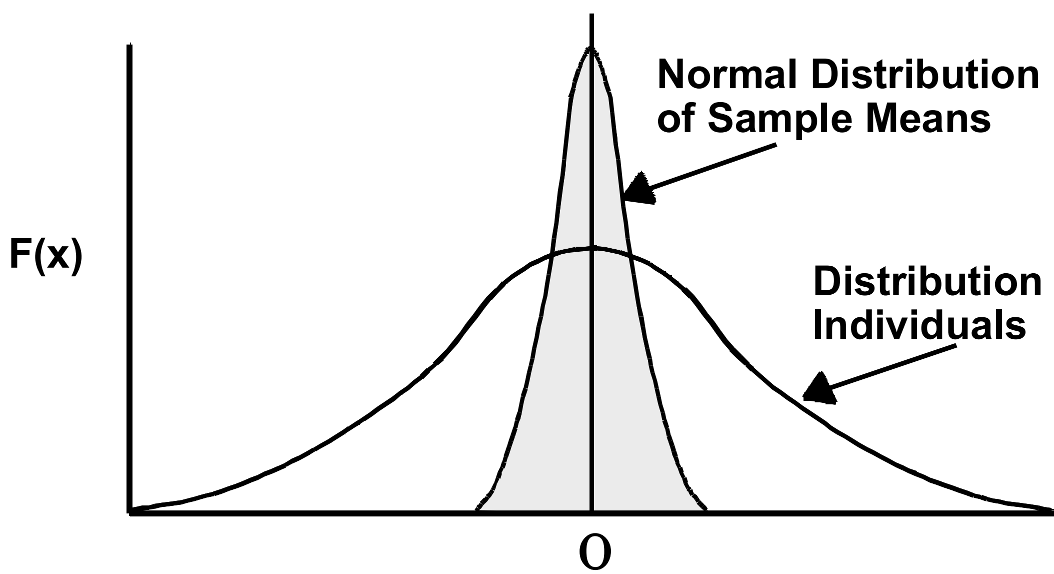
Numerical, descriptive measures calculated from a sample are called statistics. When these measures describe a population, they are called parameters.

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

Central Limit Theorem

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

Sampling Distribution of the Mean



Distributions of Individuals Versus Means

Central Limit Theorem

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. $S_{\bar{X}}$ is referred to as the standard error of the mean:

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}$$

Which is estimated by:

$$s_{\bar{X}} = \frac{s_X}{\sqrt{n}}$$

Central Limit Theorem (Continued)

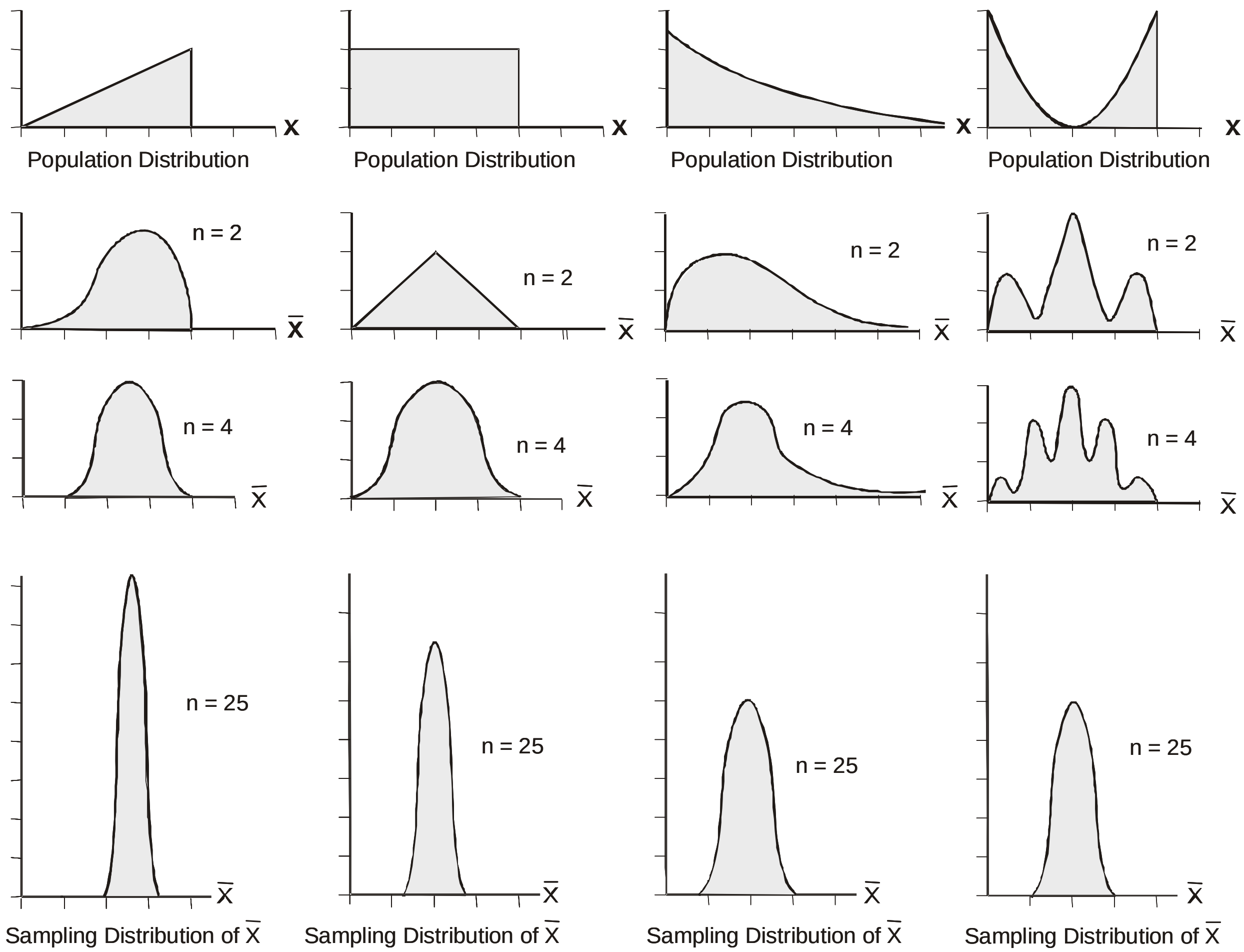


Illustration of Central Tendency

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

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

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:

$$P(E) \approx \frac{n_E}{N}$$

Compound Events

Compound events are formed by a composition of two or more events. The two most important probability theorems are the additive and multiplicative. For the following discussion, E_A = A and E_B = B.

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

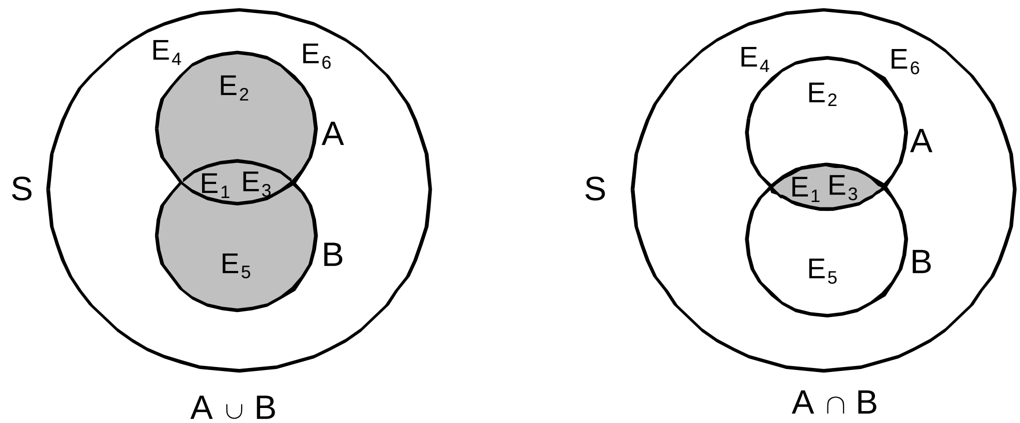
A. Union of A and B.

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.

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.

Compound Events (Continued)



Venn Diagrams Illustrating Union and Intersection

II. Event Relationships.

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-P_A.

B. Conditional Probabilities.

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

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

Compound Events (Continued)

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

$$P(A|B) = P(A) \quad \text{or} \quad P(B|A) = P(B)$$

C. Mutually Exclusive Events.

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

D. Testing for Event Relationships.

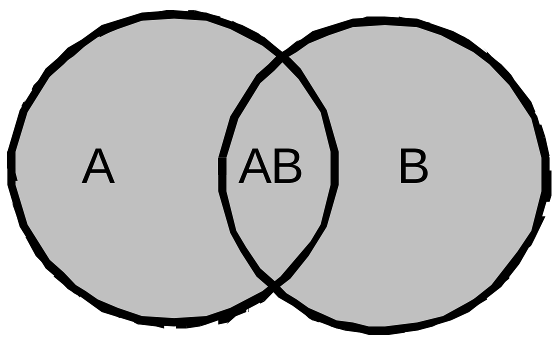
Are A and B mutually exclusive, complementary, independent, or dependent? If A and B contain one or more sample points in common, they are not mutually exclusive. If event B does not contain all points in S that are not in A, then they are not complementary.

The Additive Law

If the two events are not mutually exclusive:

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

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



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

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

The problem statements center around the word "or."

The Multiplicative Law

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

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

P(A ∩ B) is shown as P(A • B) in some texts 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.

If events A and B are independent:

2. $P(A \cap B) = P(A) \times P(B)$

The problem statements center around the word "and."