بسم الله الرحمن الرحيم

Lectures' Notes
(Unrevised First Draft)

STAT – 324
Probability and Statistics for Engineers

Second Semester 1424/1425
Sections: 1190 and 5735

Teacher: Dr. Abdullah Al-Shiha
Office: 2B43 Building # 4
Office Hours: Saturday & Monday 10-11am
Department of Statistics and Operations Research
College of Science, King Saud University

Textbook:
Probability and Statistics for Engineers and Scientists
By: R. E. Walpole, R. H. Myers, and S. L. Myers
(Sixth Edition)

Chapter 1: Introduction to Statistics and Data Analysis:

1.1 Introduction:

* Populations and Samples:

Population	→	Sample = Observations
(Some Unknown Parameters)		**(We calculate Some Statistics)**
Example: KSU Students		Example: 20 Students from KSU
(Height Mean)		**(Sample Mean)**
N=Population Size		**n = Sample Size**

- Let $X_1, X_2, \ldots, X_N$ be the population values (in general, they are unknown)
- Let $x_1, x_2, \ldots, x_n$ be the sample values (these values are known)
- Statistics obtained from the sample are used to estimate (approximate) the parameters of the population.

* Scientific Data
* Statistical Inference

(1) Estimation:

$\rightarrow$ Point Estimation

$\rightarrow$ Interval Estimation (Confidence Interval)

(2) Hypotheses Testing

1.3 Measures of Location (Central Tendency):

- The data (observations) often tend to be concentrated around the center of the data.
- Some measures of location are: the mean, mode, and median.
- These measures are considered as representatives (or typical values) of the data. They are designed to give some quantitative measures of where the center of the data is in the sample.

The Sample mean of the observations ($\bar{x}$):

If $x_1, x_2, \ldots, x_n$ are the sample values, then the sample mean is

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \text{(unit)}$$

Unrevised First Draft

Example:

Suppose that the following sample represents the ages (in year) of a sample of 3 men:

$$x_1 = 30,\ x_2 = 35,\ x_3 = 27. \quad (n=3)$$

Then, the sample mean is:

$$\bar{x} = \frac{30+35+27}{3} = \frac{92}{3} = 30.67 \quad \text{(year)}$$

Note: $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$.

1.4 Measures of Variability (Dispersion or Variation):

- The variation or dispersion in a set of data refers to how spread out the observations are from each other.
- The variation is small when the observations are close together. There is no variation if the observations are the same.
- Some measures of dispersion are range, variance, and standard deviation
- These measures are designed to give some quantitative measures of the variability in the data.

The Sample Variance (S^2):

Let $x_1, x_2, \ldots, x_n$ be the observations of the sample. The sample variance is denoted by S^2 and is defined by:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1} \quad \text{(unit)}^2$$

where $\bar{x} = \sum_{i=1}^{n} x_i / n$ is the sample mean.

Note:

(n –1) is called the degrees of freedom (df) associated with the sample variance S^2.

The Standard Deviation (S):

The standard deviation is another measure of variation. It is the square root of the variance, i.e., it is:

$$S = \sqrt{S^2} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}} \quad \text{(unit)}$$

Example:

Compute the sample variance and standard deviation of the following observations (ages in year): 10, 21, 33, 53, 54.

Solution:

n=5

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{5} x_i}{5} = \frac{10+21+33+53+54}{5} = \frac{171}{5} = 34.2 \quad \text{(year)}$$

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1} = \frac{\sum_{i=1}^{5}(x_i - 34.2)^2}{5-1}$$

$$= \frac{(10-34.2)^2 + (21-34.2)^2 + (33-34.2)^2 + (53-34.2)^2 + (54-34.2)^2}{4}$$

$$= \frac{1506.8}{4} = 376.7 \quad \text{(year)}^2$$

The sample standard deviation is:

$$S = \sqrt{S^2} = \sqrt{376.7} = 19.41 \qquad \text{(year)}$$

* Another Formula for Calculating S^2:

$$S^2 = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1} \qquad \text{(It is simple and more accurate)}$$

For the previous Example,

x_i	10	21	33	53	54	$\sum x_i = 171$
x_i^2	100	441	1089	2809	2916	$\sum x_i^2 = 7355$

$$S^2 = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1} = \frac{7355 - (5)(34.2)^2}{5-1} = \frac{1506.8}{4} = 376.7 \ \text{(year)}^2$$

Chapter 2: Probability:

- **An Experiment:** is some procedure (or process) that we do and it results in an outcome.

2.1 The Sample Space:

Definition 2.1:

- The set of all possible outcomes of a statistical experiment is called the sample space and is denoted by S.
- Each outcome (element or member) of the sample space S is called a sample point.

2.2 Events:

Definition 2.2:

An event A is a subset of the sample space S. That is $A \subseteq S$.

- We say that an event A occurs if the outcome (the result) of the experiment is an element of A.
- $\phi \subseteq S$ is an event (ϕ is called the impossible event)
- $S \subseteq S$ is an event (S is called the sure event)

Example:

Experiment: Selecting a ball from a box containing 6 balls numbered 1,2,3,4,5 and 6. (or tossing a die)

- This experiment has 6 possible outcomes
  The sample space is $S=\{1,2,3,4,5,6\}$.
- Consider the following events:

E_1=getting an even number $=\{2,4,6\} \subseteq S$

E_2 =getting a number less than 4$=\{1,2,3\} \subseteq S$

E_3 =getting 1 or 3$=\{1,3\} \subseteq S$

E_4 =getting an odd number$=\{1,3,5\} \subseteq S$

E_5 =getting a negative number$=\{\ \}=\phi \subseteq S$

E_6 =getting a number less than 10 $= \{1,2,3,4,5,6\}=S \subseteq S$

Notation:

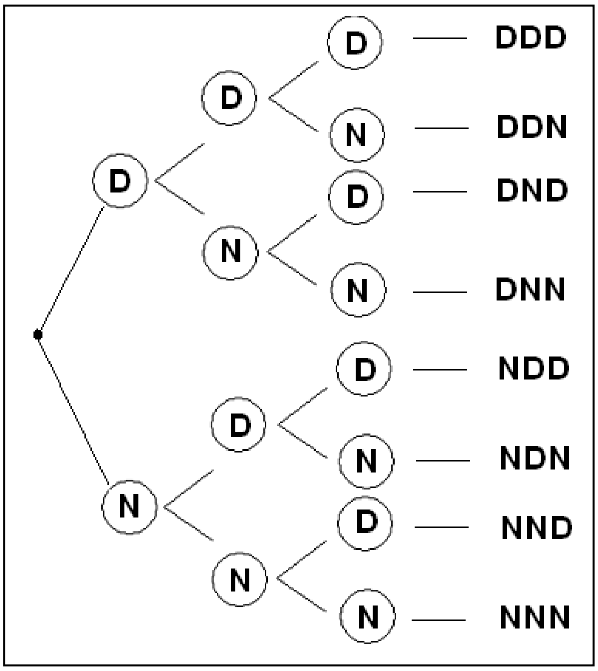
$n(S)$= no. of outcomes (elements) in S.

$n(E)$= no. of outcomes (elements) in the event E.

Example:

Experiment: Selecting 3 items from manufacturing process; each item is inspected and classified as defective (D) or non-defective (N).

- This experiment has 8 possible outcomes
  S={DDD,DDN,DND,DNN,NDD,NDN,NND,NNN}



- Consider the following events:

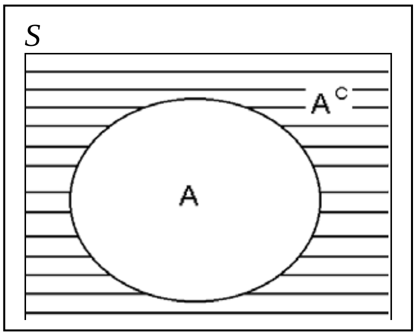
A={at least 2 defectives}= {DDD,DDN,DND,NDD}$\subseteq S$
B={at most one defective}={DNN,NDN,NND,NNN}$\subseteq S$
C={3 defectives}={DDD}$\subseteq S$

Some Operations on Events:

Let A and B be two events defined on the sample space S.

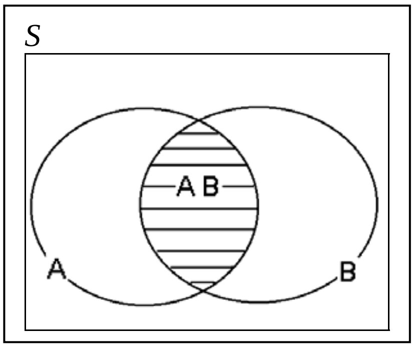
Definition 2.3: Complement of The Event A:

- A^c **or** A' **or** $\overline{A}$
- $A^c = \{x \in S:\ x \notin A\}$
- A^c consists of all points of S that are not in A.
- A^c occurs if A does not.

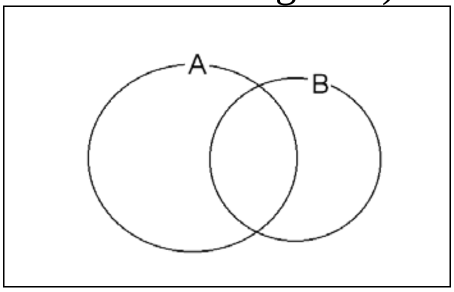


Definition 2.4: Intersection:

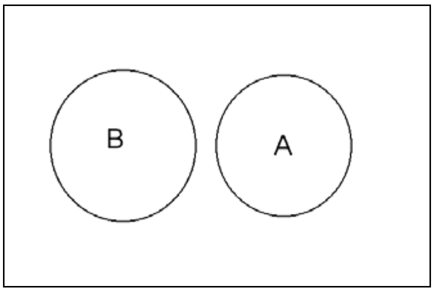
- $A \cap B = AB = \{x \in S:\ x \in A \text{ and } x \in B\}$
- $A \cap B$ Consists of all points in both A **and** B.
- $A \cap B$ Occurs if both A **and** B occur together.



Definition 2.5: Mutually Exclusive (Disjoint) Events:
Two events *A* and *B* are mutually exclusive (or disjoint) if and only if $A \cap B = \phi$; that is, *A* and *B* have no common elements (they do not occur together).



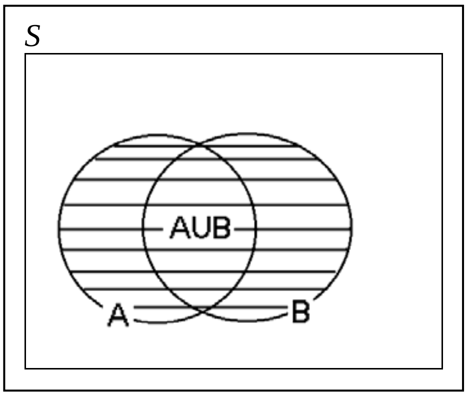
$A \cap B \neq \phi$
A and *B* are not mutually exclusive



$A \cap B = \phi$
A and *B* are mutually exclusive (disjoint)

Definition 2.6: Union:

- $A \cup B = \{x \in S: x \in A \text{ or } x \in B\}$
- $A \cup B$ Consists of all outcomes in *A* **or** in *B* **or** in both *A* and *B*.
- $A \cup B$ Occurs if *A* occurs, **or** *B* occurs, **or** both *A* and *B* occur. That is $A \cup B$ Occurs if at least one of *A* and *B* occurs.



<u>2.3 Counting Sample Points:</u>

- There are many counting techniques which can be used to count the number points in the sample space (or in some events) without listing each element.
- In many cases, we can compute the probability of an event by using the counting techniques.

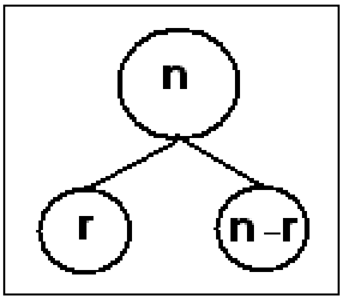
Combinations:
In many problems, we are interested in the number of ways of selecting *r* objects from *n* objects without regard to order. These selections are called combinations.

- Notation:
  n factorial is denoted by $n!$ and is defined by:
  $$n! = n \times (n-1) \times (n-2) \times \cdots \times (2) \times (1) \qquad for \quad n = 1, 2, \cdots$$
  $$0! = 1$$

Example: $5! = 5\times 4\times 3\times 2\times 1 = 120$

Theorem 2.8:

The number of combinations of *n* distinct objects taken *r* at a time is denoted by $\binom{n}{r}$ and is given by:



$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}; \qquad r = 0, 1, 2, \ldots, n$$

Notes:

- $\binom{n}{r}$ is read as " *n* " choose " *r* ".
- $\binom{n}{n} = 1$, $\binom{n}{0} = 1$, $\binom{n}{1} = n$, $\binom{n}{r} = \binom{n}{n-r}$
- $\binom{n}{r}$ = The number of different ways of selecting *r* objects from *n* distinct objects.
- $\binom{n}{r}$ = The number of different ways of dividing *n* distinct objects into two subsets; one subset contains *r* objects and the other contains the rest (*n*–*r*) bjects.

Example:

If we have 10 equal–priority operations and only 4 operating rooms are available, in how many ways can we choose the 4 patients to be operated on first?

Solution:

n = 10 *r* = 4

The number of different ways for selecting 4 patients from 10 patients is

$$\binom{10}{4} = \frac{10!}{4!\,(10-4)!} = \frac{10!}{4!\times 6!} = \frac{10\times 9\times 8\times 7\times 6\times 5\times 4\times 3\times 2\times 1}{(4\times 3\times 2\times 1)\ \times\ (6\times 5\times 4\times 3\times 2\times 1)}$$

$$= 210 \qquad (\textit{different ways})$$

2.4. Probability of an Event:

- To every point (outcome) in the sample space of an experiment *S*, we assign a weight (or probability), ranging from 0 to 1, such that the sum of all weights (probabilities) equals 1.

- The weight (or probability) of an outcome measures its likelihood (chance) of occurrence.
- To find the probability of an event *A*, we sum all probabilities of the sample points in *A*. This sum is called the probability of the event *A* and is denoted by P(*A*).

Definition 2.8:

The probability of an event *A* is the sum of the weights (probabilities) of all sample points in *A*. Therefore,

1. $0 \le P(A) \le 1$
2. $P(S) = 1$
3. $P(\phi) = 0$

Example 2.22:

A balanced coin is tossed twice. What is the probability that at least one head occurs?

Solution:

S = {HH, HT, TH, TT}

A = {at least one head occurs}= {HH, HT, TH}

Since the coin is balanced, the outcomes are equally likely; i.e., all outcomes have the same weight or probability.

Outcome	Weight (Probability)
HH	P(HH) = w
HT	P(HT) = w
TH	P(TH) = w
TT	P(TT) = w
sum	4w=1

4w =1 ⇔ w =1/4 = 0.25

P(HH)=P(HT)=P(TH)=P(TT)=0.25

The probability that at least one head occurs is:

P(*A*) = P({at least one head occurs})=P({HH, HT, TH})

= P(HH) + P(HT) + P(TH)

= 0.25+0.25+0.25

= 0.75

Theorem 2.9:

If an experiment has *n*(*S*)=*N* equally likely different outcomes, then the probability of the event *A* is:

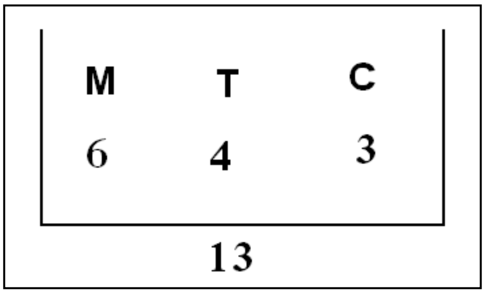
$$P(A) = \frac{n(A)}{n(S)} = \frac{n(A)}{N} = \frac{no.\ of\ outcomes\ \ in\ \ A}{no.\ of\ outcomes\ \ in\ \ S}$$

Example 2.25:

A mixture of candies consists of 6 mints, 4 toffees, and 3 chocolates. If a person makes a random selection of one of these candies, find the probability of getting:

(a) a mint

(b) a toffee or chocolate.



Solution:

Define the following events:

M = {getting a mint}

T = {getting a toffee}

C = {getting a chocolate}

Experiment: selecting a candy at random from 13 candies

$n(S)$ = no. of outcomes of the experiment of selecting a candy.

= no. of different ways of selecting a candy from 13 candies.

$= \binom{13}{1} = 13$

The outcomes of the experiment are equally likely because the selection is made at random.

(a) M = {getting a mint}

$n(M)$ = no. of different ways of selecting a mint candy from 6 mint candies

$= \binom{6}{1} = 6$

P(M)= P({getting a mint})$= \frac{n(M)}{n(S)} = \frac{6}{13}$

(b) $T \cup C$ = {getting a toffee or chocolate}

$n(T \cup C)$ = no. of different ways of selecting a toffee **or** a chocolate candy

= no. of different ways of selecting a toffee candy **+** no. of different ways of selecting a chocolate candy

$= \binom{4}{1} + \binom{3}{1} = 4 + 3 = 7$

= no. of different ways of selecting a candy from 7 candies

$$= \binom{7}{1} = 7$$

P($T \cup C$)= P({getting a toffee or chocolate})= $\frac{n(T \cup C)}{n(S)} = \frac{7}{13}$

Example 2.26:

In a poker hand consisting of 5 cards, find the probability of holding 2 aces and 3 jacks.

Solution:

Experiment: selecting 5 cards from 52 cards.

$n(S)$ = no. of outcomes of the experiment of selecting 5 cards from 52 cards.

$$= \binom{52}{5} = \frac{52!}{5! \times 47!} = 2598960$$

The outcomes of the experiment are equally likely because the selection is made at random.

Define the event A = {holding 2 aces **and** 3 jacks}

$n(A)$ = no. of ways of selecting 2 aces **and** 3 jacks

= (no. of ways of selecting 2 aces) × (no. of ways of selecting 3 jacks)

= (no. of ways of selecting 2 aces from 4 aces) × (no. of ways of selecting 3 jacks from 4 jacks)

$$= \binom{4}{2} \times \binom{4}{3}$$

$$= \frac{4!}{2! \times 2!} \times \frac{4!}{3! \times 1!} = 6 \times 4 = 24$$

P(A)= P({holding 2 aces **and** 3 jacks })

$$= \frac{n(A)}{n(S)} = \frac{24}{2598960} = 0.000009$$

2.5 Additive Rules:

Theorem 2.10:

If A and B are any two events, then:

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

Corollary 1:

If A and B are mutually exclusive (disjoint) events, then:

$$P(A \cup B) = P(A) + P(B)$$

Corollary 2:

If $A_1, A_2, \ldots, A_n$ are n mutually exclusive (disjoint) events, then:

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

Note: Two event Problems:

* In Venn diagrams, consider the probability of an event A as the area of the region corresponding to the event A.

* Total area= $P(S)=1$

* Examples:

$P(A) = P(A \cap B) + P(A \cap B^C)$

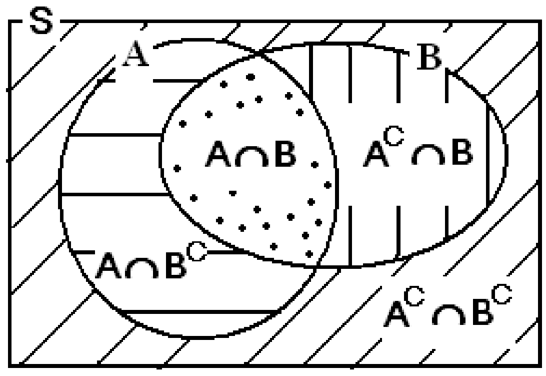
$P(A \cup B) = P(A) + P(A^C \cap B)$

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

$P(A \cap B^C) = P(A) - P(A \cap B)$

$P(A^C \cap B^C) = 1 - P(A \cup B)$

etc.,



Total area= $P(S)=1$

Example 2.27:

The probability that Paula passes Mathematics is 2/3, and the probability that she passes English is 4/9. If the probability that she passes both courses is 1/4, what is the probability that she will:

(a) pass at least one course?

(b) pass Mathematics and fail English?

(c) fail both courses?

Solution:

Define the events: M={Paula passes Mathematics}

E={Paula passes English}

We know that $P(M)=2/3$, $P(E)=4/9$, and $P(M \cap E)=1/4$.

(a) Probability of passing at least one course is:

$$P(M \cup E) = P(M) + P(E) - P(M \cap E)$$

$$= \frac{2}{3} + \frac{4}{9} - \frac{1}{4} = \frac{31}{36}$$

(b) Probability of passing Mathematics and failing English is:

$$P(M \cap E^C) = P(M) - P(M \cap E)$$

$$= \frac{2}{3} - \frac{1}{4} = \frac{5}{12}$$

(c) Probability of failing both courses is:

$$P(M^C \cap E^C) = 1 - P(M \cup E)$$

$$= 1 - \frac{31}{36} = \frac{5}{36}$$

Theorem 2.12:

If A and A^C are complementary events, then:

$$P(A) + P(A^C) = 1 \Leftrightarrow P(A^C) = 1 - P(A)$$

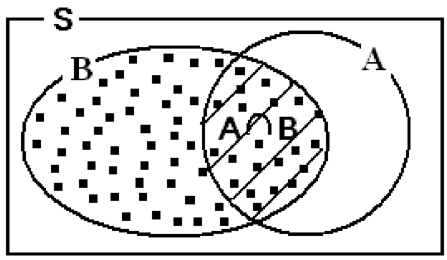
2.6 Conditional Probability:

The probability of occurring an event A when it is known that some event B has occurred is called the conditional probability of A given B and is denoted $P(A|B)$.

Definition 2.9:

The conditional probability of the event A given the event B is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad ; \; P(B) > 0$$



P(S)=Total area=1

Notes:

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

$$= \frac{n(A \cap B)/n(S)}{n(B)/n(S)} = \frac{n(A \cap B)}{n(B)}; \; \textit{for equally likely outcomes case}$$

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

3. $P(A \cap B) = P(A)\,P(B|A)$
$\quad = P(B)\,P(A|B)$ (Multiplicative Rule=Theorem 2.13)

Example:

339 physicians are classified as given in the table below. A physician is to be selected at random.

(1) Find the probability that:

(a) the selected physician is aged 40 – 49

(b) the selected physician smokes occasionally

(c) the selected physician is aged 40 – 49 and smokes occasionally

(2) Find the probability that the selected physician is aged 40 – 49 given that the physician smokes occasionally.

		Smoking Habit			
		Daily (B_1)	Occasionally (B_2)	Not at all (B_3)	Total
Age	20 - 29 (A_1)	31	9	7	47
	30 - 39 (A_2)	110	30	49	189
	40 - 49 (A_3)	29	21	29	79
	50+ (A_4)	6	0	18	24
	Total	176	60	103	339

Solution:

.$n(S) = 339$ equally likely outcomes.

Define the following events:

A_3 = the selected physician is aged 40 – 49

B_2 = the selected physician smokes occasionally

$A_3 \cap B_2$ = the selected physician is aged 40 – 49 and smokes occasionally

(1) (a) A_3 = the selected physician is aged 40 – 49

$$P(A_3) = \frac{n(A_3)}{n(S)} = \frac{79}{339} = 0.2330$$

(b) B_2 = the selected physician smokes occasionally

$$P(B_2) = \frac{n(B_2)}{n(S)} = \frac{60}{339} = 0.1770$$

(c) $A_3 \cap B_2$ = the selected physician is aged 40 – 49 and smokes occasionally.

$$P(A_3 \cap B_2) = \frac{n(A_3 \cap B_2)}{n(S)} = \frac{21}{339} = 0.06195$$

(2) $A_3|B_2$ = the selected physician is aged 40 – 49 given that the physician smokes occasionally

(i) $P(A_3 \mid B_2) = \dfrac{P(A_3 \cap B_2)}{P(B_2)} = \dfrac{0.06195}{0.1770} = 0.35$

(ii) $P(A_3 \mid B_2) = \dfrac{n(A_3 \cap B_2)}{n(B_2)} = \dfrac{21}{60} = 0.35$

(iii) We can use the restricted table directly: $P(A_3 \mid B_2) = \dfrac{21}{60} = 0.35$

Notice that $P(A_3|B_2)=0.35 > P(A_3)=0.233$.
The conditional probability does not equal unconditional probability; i.e., $P(A_3|B_2) \neq P(A_3)$! What does this mean?

Note:

- $P(A|B)=P(A)$ means that knowing B has no effect on the probability of occurrence of A. In this case A is independent of B.
- $P(A|B)>P(A)$ means that knowing B increases the probability of occurrence of A.
- $P(A|B)<P(A)$ means that knowing B decreases the probability of occurrence of A.

Independent Events:

Definition 2.10:

Two events A and B are independent if and only if $P(A|B)=P(A)$ and $P(B|A)=P(B)$. Otherwise A and B are dependent.

Example:

In the previous example, we found that $P(A_3|B_2) \neq P(A_3)$. Therefore, the events A_3 and B_2 are dependent, i.e., they are not independent. Also, we can verify that $P(B_2| A_3) \neq P(B_2)$.

2.7 Multiplicative Rule:

Theorem 2.13:

If $P(A) \neq 0$ and $P(B) \neq 0$, then:

$$P(A\cap B) = P(A)\,P(B|A) = P(B)\,P(A|B)$$

Example 2.32:

Suppose we have a fuse box containing 20 fuses of which 5 are defective (D) and 15 are non-defective (N). If 2 fuses are selected at random and removed from the box in succession without replacing the first, what is the probability that both fuses are defective?

Solution:

Define the following events:

A = {the first fuse is defective}

B = {the second fuse is defective}

$A\cap B$={the first fuse is defective and the second fuse is defective} = {both fuses are defective}

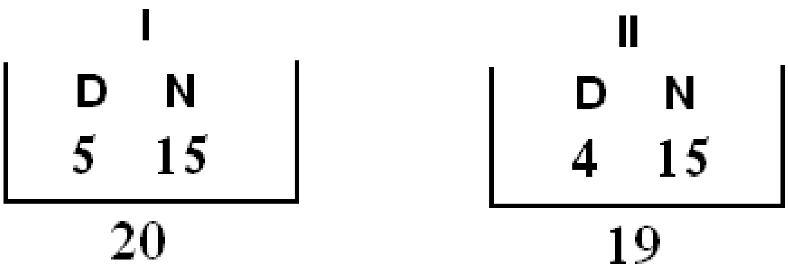
We need to calculate $P(A\cap B)$.

$P(A) = \frac{5}{20}$

$P(B|A) = \frac{4}{19}$

$P(A\cap B) = P(A)\, P(B|A)$

$= \frac{5}{20} \times \frac{4}{19} = 0.052632$



First Selection

Second Selection: given that the first is defective (D)

Theorem 2.14:

Two events A and B are independent if and only if

$$P(A\cap B)= P(A)\, P(B)$$

*(Multiplicative Rule for independent events)

Note:

Two events A and B are independent if one of the following conditions is satisfied:

(i) $P(A|B)=P(A)$

$\Leftrightarrow$ (ii) $P(B|A)=P(B)$

$\Leftrightarrow$ (iii) $P(A\cap B)= P(A)\, P(B)$

Theorem 2.15: (k=3)

- If A_1, A_2, A_3 are 3 events, then:

  $$P(A_1\cap A_2 \cap A_3)= P(A_1)\, P(A_2|A_1)\, P(A_3|A_1\cap A_2)$$

- If A_1, A_2, A_3 are 3 independent events, then:

  $$P(A_1\cap A_2 \cap A_3)= P(A_1)\, P(A_2)\, P(A_3)$$

Example 2.36:

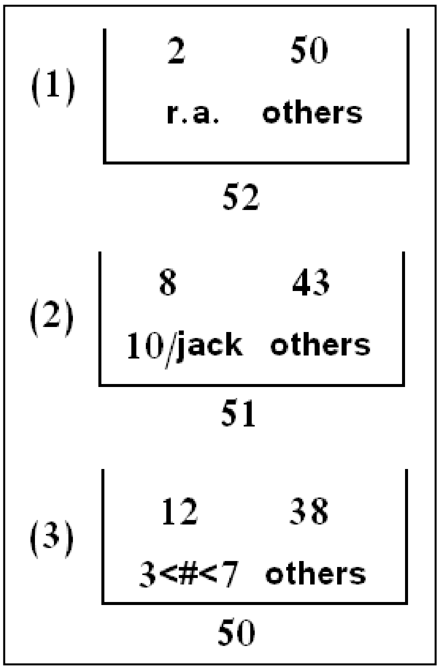
Three cards are drawn in succession, without replacement, from an ordinary deck of playing cards. Fined $P(A_1 \cap A_2 \cap A_3)$, where the events A_1, A_2 , and A_3 are defined as follows:

A_1 = {the 1-st card is a red ace}

A_2 = {the 2-nd card is a 10 or a jack}

A_3 = {the 3-rd card is a number greater than 3 but less than 7}

Solution:



$P(A_1) = 2/52$

$P(A_2 | A_1) = 8/51$

$P(A_3 | A_1 \cap A_2) = 12/50$

$P(A_1 \cap A_2 \cap A_3)$

$= P(A_1)\, P(A_2 | A_1)\, P(A_3 | A_1 \cap A_2)$

$= \frac{2}{52} \times \frac{8}{51} \times \frac{12}{50}$

$= \frac{192}{132600}$

$= 0.0014479$

2.8 Bayes' Rule:

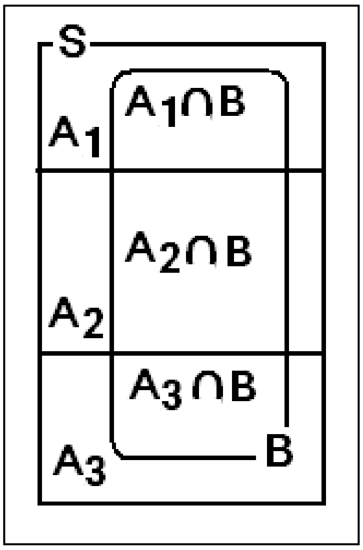
Definition:

The events $A_1, A_2, \ldots,$ and A_n constitute a partition of the sample space S if:

- $\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ \ldots \ \cup A_n = S$
- $A_i \cap A_j = \phi, \quad \forall\ i \neq j$

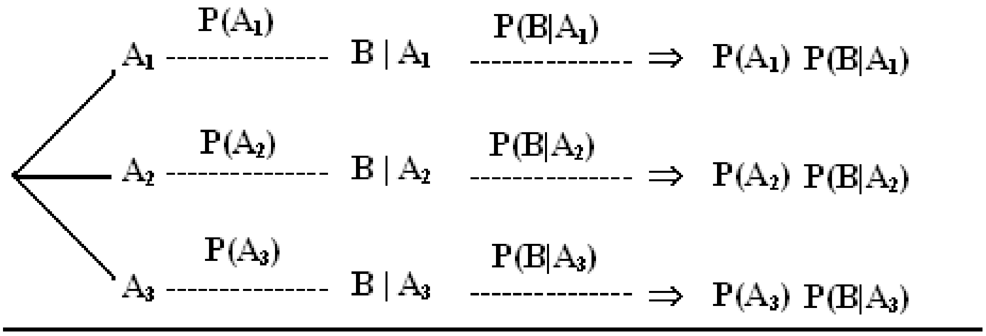
Theorem 2.16: (Total Probability)

If the events $A_1, A_2, \ldots,$ and A_n constitute a partition of the sample space S such that $P(A_k) \neq 0$ for $k = 1, 2, \ldots, n$, then for any event B:



$$P(B) = \sum_{k=1}^{n} P(A_k \cap B)$$

$$= \sum_{k=1}^{n} P(A_k) P(B | A_k)$$



المجموع $= P(B) = \sum_{k=1}^{n} P(A_k)\, P(B \mid A_k)$

Tree Diagram

Example 2.38:

Three machines A_1, A_2, and A_3 make 20%, 30%, and 50%, respectively, of the products. It is known that 1%, 4%, and 7% of the products made by each machine, respectively, are defective. If a finished product is randomly selected, what is the probability that it is defective?

Solution:

Define the following events:

B = {the selected product is defective}

A_1 = {the selected product is made by machine A_1}

A_2 = {the selected product is made by machine A_2}

A_3 = {the selected product is made by machine A_3}

$P(A_1) = \frac{20}{100} = 0.2$; $P(B|A_1) = \frac{1}{100} = 0.01$

$P(A_2) = \frac{30}{100} = 0.3$; $P(B|A_2) = \frac{4}{100} = 0.04$

$P(A_3) = \frac{50}{100} = 0.5$; $P(B|A_3) = \frac{7}{100} = 0.07$

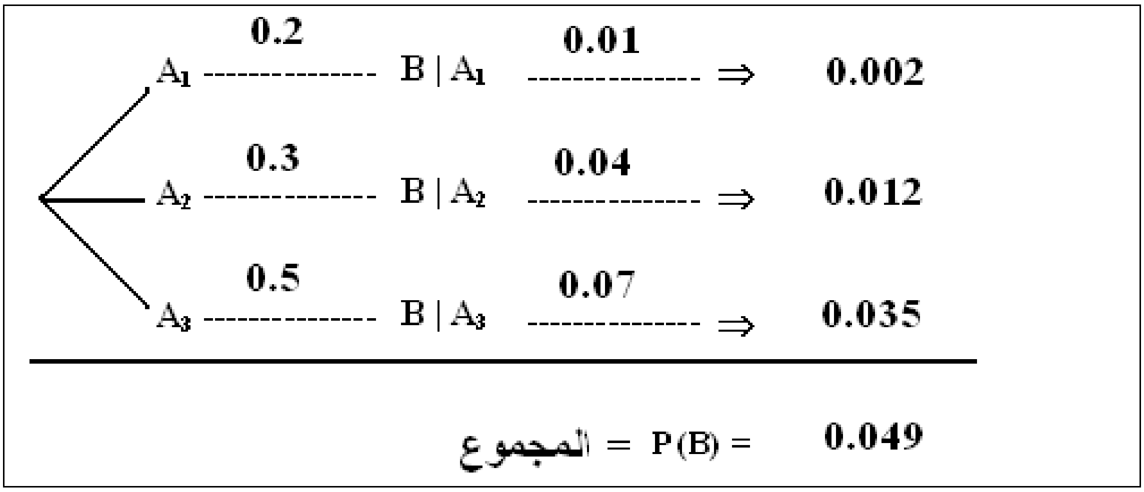
$$P(B) = \sum_{k=1}^{3} P(A_k) P(B \mid A_k)$$

$$= P(A_1)\,P(B|A_1) + P(A_2)\,P(B|A_2) + P(A_3)\,P(B|A_3)$$

$$= 0.2\times0.01 + 0.3\times0.04 + 0.5\times0.07$$

$$= 0.002 + 0.012 + 0.035$$

$$= 0.049$$



Question:

If it is known that the selected product is defective, what is the probability that it is made by machine A_1?

Answer:

$$P(A_1|B) = \frac{P(A_1 \cap B)}{P(B)} = \frac{P(A_1)P(B|A_1)}{P(B)} = \frac{0.2\times 0.01}{0.049} = \frac{0.002}{0.049} = 0.0408$$

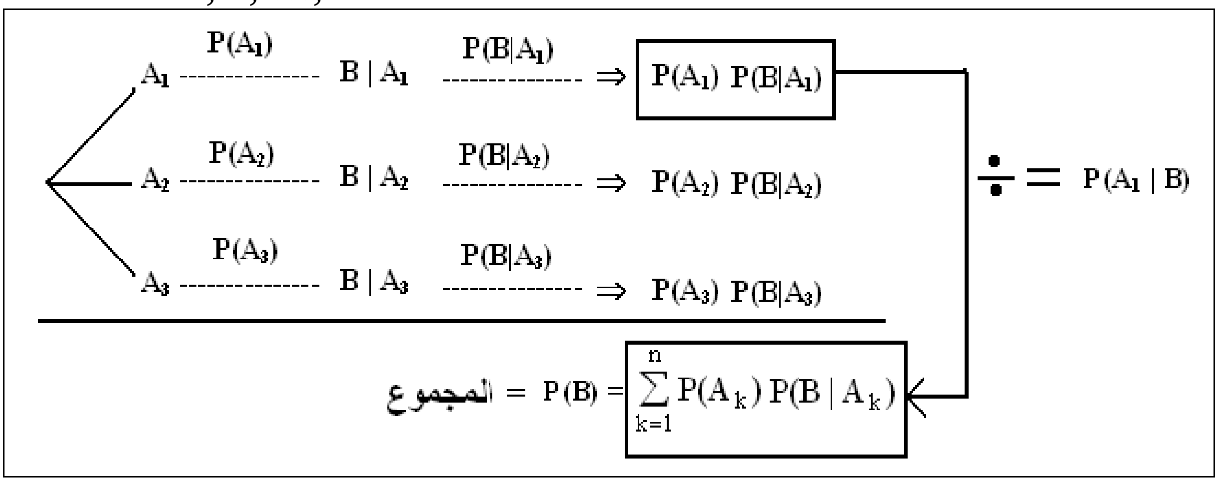
This rule is called Bayes' rule.

Theorem 2.17: (Bayes' rule)

If the events $A_1, A_2, \dots,$ and A_n constitute a partition of the sample space S such that $P(A_k) \neq 0$ for $k=1, 2, \dots, n$, then for any event B such that $P(B) \neq 0$:

$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{k=1}^{n} P(A_k)P(B|A_k)} = \frac{P(A_i)P(B|A_i)}{P(B)}$$

for i = 1, 2, …, n.



Example 2.39:

In Example 2.38, if it is known that the selected product is defective, what is the probability that it is made by:

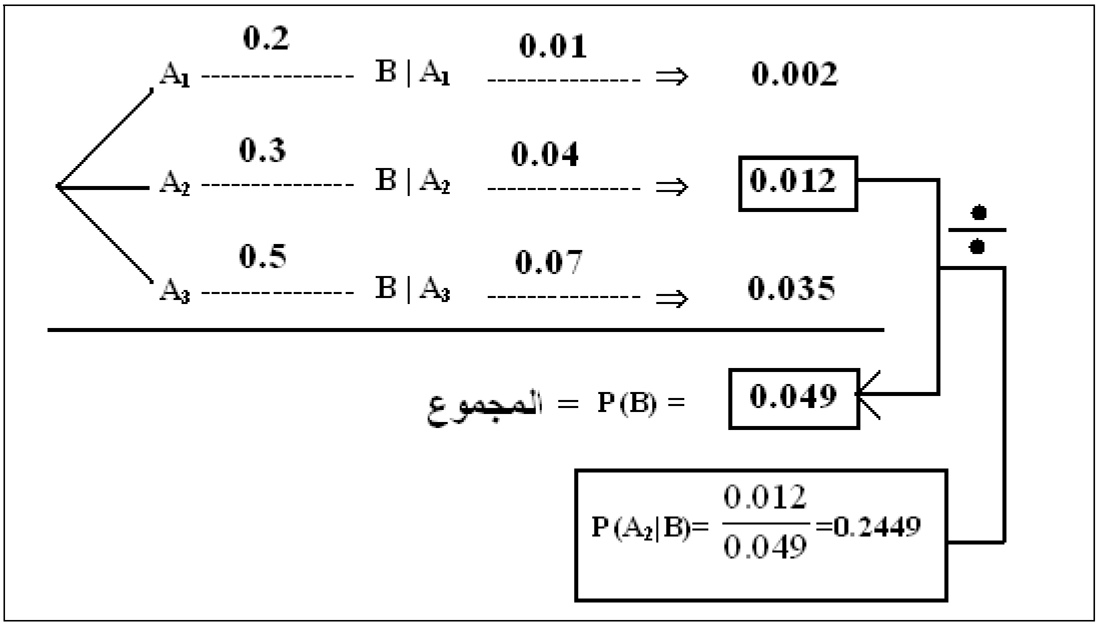
(a) machine A_2?

(b) machine A_3?

Solution:

(a) $P(A_2|B) = \dfrac{P(A_2)P(B|A_2)}{\sum_{k=1}^{n} P(A_k)P(B|A_k)} = \dfrac{P(A_2)P(B|A_2)}{P(B)}$

$$= \frac{0.3 \times 0.04}{0.049} = \frac{0.012}{0.049} = 0.2449$$



(b) $P(A_3|B) = \dfrac{P(A_3)P(B|A_3)}{\sum_{k=1}^{n} P(A_k)P(B|A_k)} = \dfrac{P(A_3)P(B|A_3)}{P(B)}$

$$= \frac{0.5 \times 0.07}{0.049} = \frac{0.035}{0.049} = 0.7142$$

Note:

$P(A_1|B) = 0.0408$, $P(A_2|B) = 0.2449$, $P(A_3|B) = 0.7142$

- $\sum_{k=1}^{3} P(A_k|B) = 1$
- If the selected product was found defective, we should check machine A_3 first, if it is ok, we should check machine A_2, if it is ok, we should check machine A_1.

Chapter 3: Random Variables and Probability Distributions:

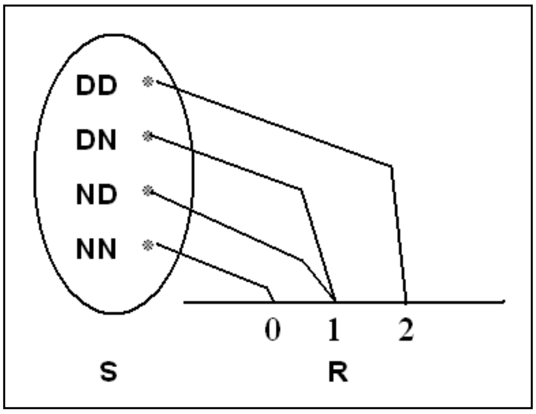
3.1 Concept of a Random Variable:

- In a statistical experiment, it is often very important to allocate numerical values to the outcomes.

Example:

- Experiment: testing two components. (D=defective, N=non-defective)
- Sample space: *S*={DD,DN,ND,NN}
- Let X = number of defective components when two components are tested.
- Assigned numerical values to the outcomes are:

Sample point (Outcome)	Assigned Numerical Value (x)
DD	2
DN	1
ND	1
NN	0



- Notice that, the set of all possible values of the random variable X is {0, 1, 2}.

Definition 3.1:

A random variable X is a function that associates each element in the sample space with a real number (i.e., $X : S \to \mathbf{R}$.)

Notation: " X " denotes the random variable .
" x " denotes a value of the random variable X.

Types of Random Variables:

- A random variable X is called a **discrete** random variable if its set of possible values is countable, i.e.,
  .$x \in \{x_1, x_2, \ldots, x_n\}$ or $x \in \{x_1, x_2, \ldots\}$
- A random variable X is called a **continuous** random variable if it can take values on a continuous scale, i.e.,
  .$x \in \{x: a < x < b; a, b \in R\}$

- In most practical problems:
    - A discrete random variable represents count data, such as the number of defectives in a sample of *k* items.
    - A continuous random variable represents measured data, such as height.

3.2 Discrete Probability Distributions

- A discrete random variable X assumes each of its values with a certain probability.

Example:

- Experiment: tossing a non-balance coin 2 times independently.
- H= head , T=tail
- Sample space: *S*={HH, HT, TH, TT}
- Suppose P(H)=½P(T) ⇔ P(H)=1/3 and P(T)=2/3
- Let X= number of heads

Sample point (Outcome)	Probability	Value of X (x)
HH	P(HH)=P(H) P(H)=1/3×1/3 = 1/9	2
HT	P(HT)=P(H) P(T)=1/3×2/3 = 2/9	1
TH	P(TH)=P(T) P(H)=2/3×1/3 = 2/9	1
TT	P(TT)=P(T) P(T)=2/3×2/3 = 4/9	0

- The possible values of X are: 0, 1, and 2.
- X is a discrete random variable.
- Define the following events:

Event (X=x)	Probability = P(X=x)
(X=0)={TT}	P(X=0) = P(TT)=4/9
(X=1)={HT,TH}	P(X=1) =P(HT)+P(TH)=2/9+2/9=4/9
(X=2)={HH}	P(X=2) = P(HH)= 1/9

- The possible values of X with their probabilities are:

x	0	1	2	Total
P(X=x)=f(x)	4/9	4/9	1/9	1.00

The function f(x)=P(X=x) is called the probability function (probability distribution) of the discrete random variable X.

Definition 3.4:

The function f(x) is a probability function of a discrete random variable X if, for each possible values x, we have:

1. $f(x) \geq 0$
2. $\sum_{all\ x} f(x) = 1$
3. $f(x) = P(X=x)$

Note:

- $P(X \in A) = \sum_{all\ x \in A} f(x) = \sum_{all\ x \in A} P(X = x)$

Example:

For the previous example, we have:

x	0	1	2	Total
f(x)= P(X=x)	4/9	4/9	1/9	$\sum_{x=0}^{2} f(x) = 1$

$P(X<1) = P(X=0)=4/9$

$P(X\leq 1) = P(X=0) + P(X=1) = 4/9+4/9 = 8/9$

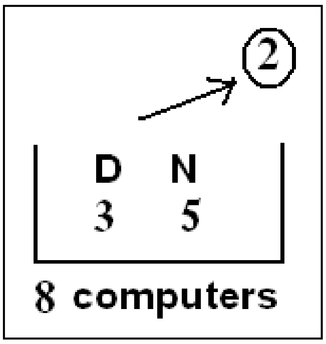
$P(X\geq 0.5) = P(X=1) + P(X=2) = 4/9+1/9 = 5/9$

$P(X>8) = P(\phi) = 0$

$P(X<10) = P(X=0) + P(X=1) + P(X=2) = P(S) = 1$

Example 3.3:

A shipment of 8 similar microcomputers to a retail outlet contains 3 that are defective and 5 are non-defective.
If a school makes a random purchase of 2 of these computers, find the probability distribution of the number of defectives.



Solution:

We need to find the probability distribution of the random variable: X = the number of defective computers purchased.

Experiment: selecting 2 computers at random out of 8

$$n(S) = \binom{8}{2} \text{ equally likely outcomes}$$

The possible values of X are: x=0, 1, 2.

Consider the events:

(X=0)={0D and 2N} $\Rightarrow$ n(X=0)= $\binom{3}{0}\times\binom{5}{2}$

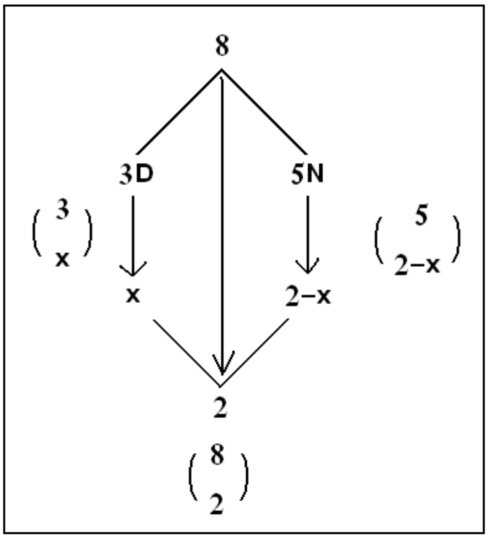
(X=1)={1D and 1N} $\Rightarrow$ n(X=1)= $\binom{3}{1}\times\binom{5}{1}$

(X=2)={2D and 0N} $\Rightarrow$ n(X=2)= $\binom{3}{2}\times\binom{5}{0}$

f(0)=P(X=0)= $\frac{n(X=0)}{n(S)}=\frac{\binom{3}{0}\times\binom{5}{2}}{\binom{8}{2}}=\frac{10}{28}$

f(1)=P(X=1)= $\frac{n(X=1)}{n(S)}=\frac{\binom{3}{1}\times\binom{5}{1}}{\binom{8}{2}}=\frac{15}{28}$

f(2)=P(X=2)= $\frac{n(X=2)}{n(S)}=\frac{\binom{3}{2}\times\binom{5}{0}}{\binom{8}{2}}=\frac{3}{28}$



In general, for x=0,1, 2, we have:

f(x)=P(X=x)= $\frac{n(X=x)}{n(S)}=\frac{\binom{3}{x}\times\binom{5}{2-x}}{\binom{8}{2}}$

The probability distribution of X is:

x	0	1	2	Total
f(x)= P(X=x)	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$	1.00

The probability distribution of X can be written as a formula as follows:

$$f(x)=P(X=x)=\begin{cases}\dfrac{\binom{3}{x}\times\binom{5}{2-x}}{\binom{8}{2}}; & x=0,1,2\\ 0; & otherwise\end{cases}$$

Hypergeometric Distribution

Definition 3.5:

The cumulative distribution function (CDF), F(x), of a discrete random variable X with the probability function f(x) is given by:

$$F(x) = P(X\le x) = \sum_{t\le x} f(t) = \sum_{t\le x} P(X=t) \ ; \ \text{for } -\infty<x<\infty$$

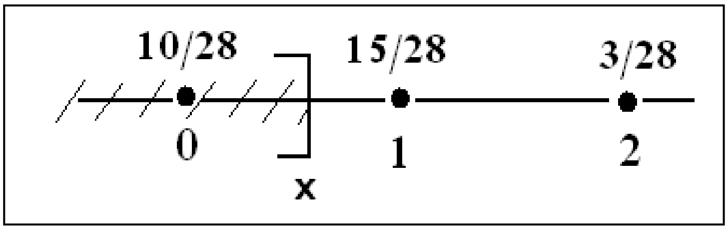
Example:

Find the CDF of the random variable X with the probability function:

x	0	1	2
f(x)	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

Solution:

F(x)=P(X≤x) for –∞<x<∞



For x<0: F(x)=0

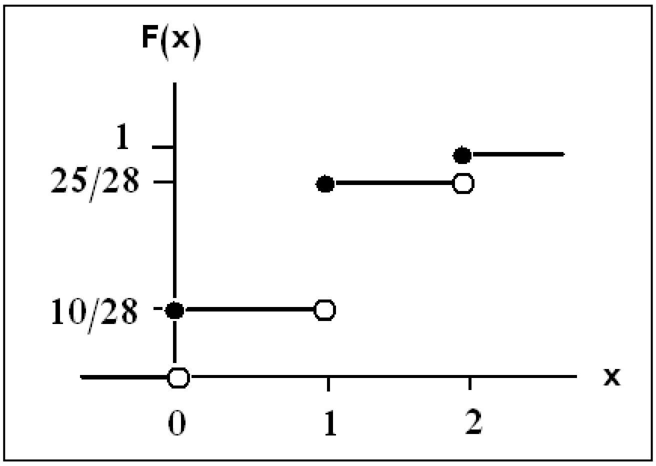
For 0≤x<1: $F(x)=P(X=0)=\frac{10}{28}$

For 1≤x<2: $F(x)=P(X=0)+P(X=1)=\frac{10}{28}+\frac{15}{28}=\frac{25}{28}$

For x≥2: $F(x)=P(X=0)+P(X=1)+P(X=2)=\frac{10}{28}+\frac{15}{28}+\frac{3}{28}=1$

The CDF of the random variable X is:

$$F(x)=P(X\le x)=\begin{cases} 0 \ ; & x<0 \\ \frac{10}{28} \ ; & 0\le x<1 \\ \frac{25}{28} \ ; & 1\le x<2 \\ 1 \ ; & x\ge 2 \end{cases}$$



Note:

F(–0.5) = P(X≤–0.5)=0

$F(1.5)=P(X\le 1.5)=F(1) = \frac{25}{28}$

F(3.8) =P(X≤3.8)=F(2)= 1

Result:

P(a < X ≤ b) = P(X ≤ b) – P(X ≤ a) = F(b) – F(a)

$P(a \leq X \leq b) = P(a < X \leq b) + P(X=a) = F(b) - F(a) + f(a)$

$P(a < X < b) = P(a < X \leq b) - P(X=b) = F(b) - F(a) - f(b)$

Result:

Suppose that the probability function of X is:

x	x_1	x_2	x_3	…	x_n
f(x)	$f(x_1)$	$f(x_2)$	$f(x_3)$	…	$f(x_n)$

Where $x_1 < x_2 < \ldots < x_n$. Then:

$F(x_i) = f(x_1) + f(x_2) + \ldots + f(x_i)$; $i=1, 2, \ldots, n$

$F(x_i) = F(x_{i-1}) + f(x_i)$; $i=2, \ldots, n$

$f(x_i) = F(x_i) - F(x_{i-1})$

Example:

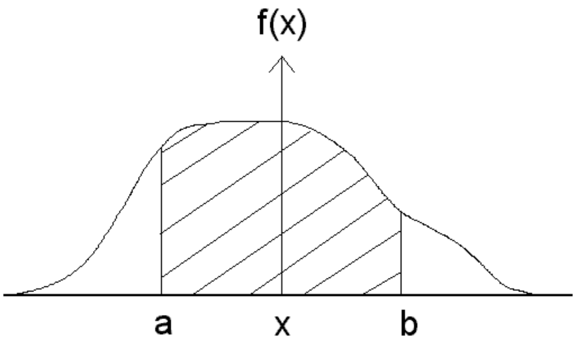
In the previous example,

$P(0.5 < X \leq 1.5) = F(1.5) - F(0.5) = \frac{25}{28} - \frac{10}{28} = \frac{15}{28}$

$P(1 < X \leq 2) = F(2) - F(1) = 1 - \frac{25}{28} = \frac{3}{28}$

<u>3.3. Continuous Probability Distributions</u>

For any continuous random variable, X, there exists a non-negative function f(x), called the probability density function (p.d.f) through which we can find probabilities of events expressed in term of X.



$f: R \to [0, \infty)$

$P(a < X < b) = \int_a^b f(x)dx$

= area under the curve of f(x) and over the interval (a,b)

$P(X \in A) = \int_A f(x)dx$

= area under the curve of f(x) and over the region A

Definition 3.6:

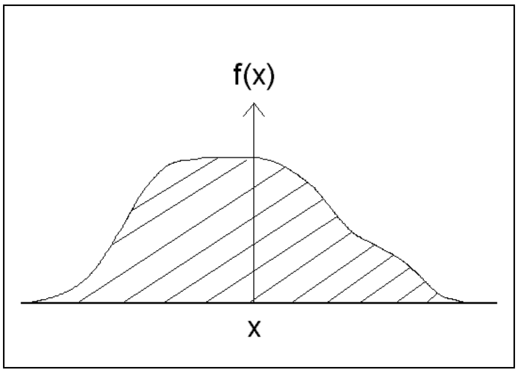
The function f(x) is a probability density function (pdf) for a continuous random variable X, defined on the set of real numbers, if:

1. $f(x) \geq 0 \quad \forall\, x \in R$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$
3. $P(a \leq X \leq b) = \int_{a}^{b} f(x)\,dx \quad \forall\, a, b \in R;\ a \leq b$

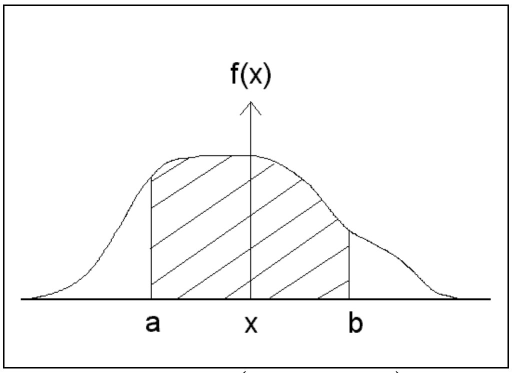
Note:

For a continuous random variable X, we have:

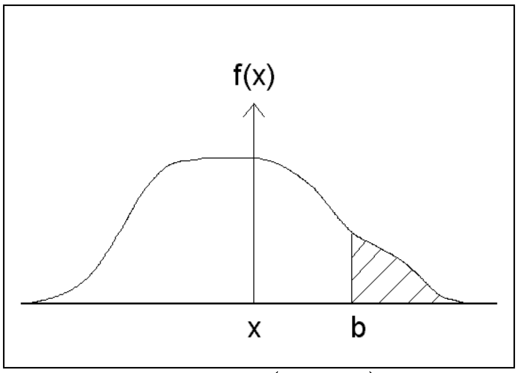
1. $f(x) \neq P(X=x)$ (in general)
2. $P(X=a) = 0$ for any $a \in R$
3. $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$
4. $P(X \in A) = \int_{A} f(x)\,dx$



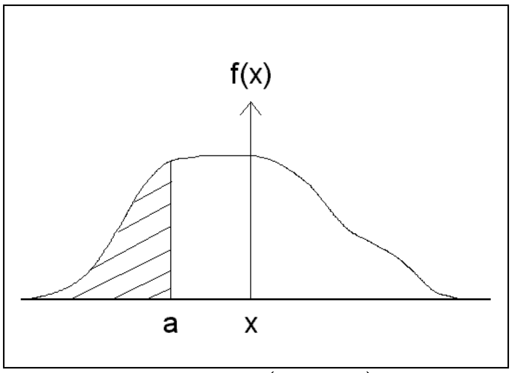
$Total\ area = \int_{-\infty}^{\infty} f(x)\,dx = 1$



$area = P(a \leq X \leq b) = \int_{a}^{b} f(x)\,dx$



$area = P(X \geq b) = \int_{b}^{\infty} f(x)\,dx$



$area = P(X \leq a) = \int_{-\infty}^{a} f(x)\,dx$

Example 3.6:

Suppose that the error in the reaction temperature, in ^oC, for a controlled laboratory experiment is a continuous random variable X having the following probability density function:

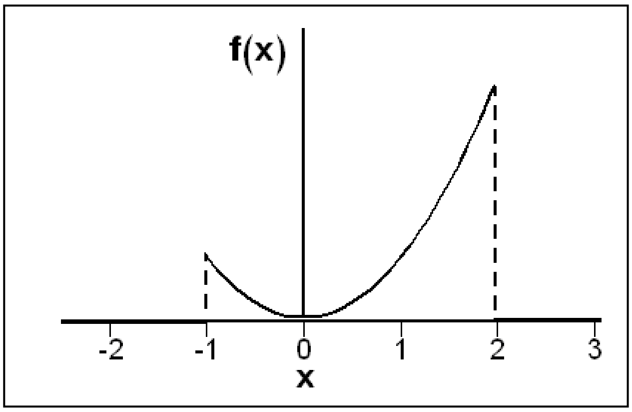
$$f(x) = \begin{cases} \frac{1}{3}x^2 \ ; -1 < x < 2 \\ 0 \ ; \ elsewhere \end{cases}$$

1. Verify that (a) f(x) ≥ 0 and (b) $\int_{-\infty}^{\infty} f(x)\,dx = 1$
2. Find P(0<X≤1)

Solution:

X = the error in the reaction temperature in ^oC.

X is continuous r. v.



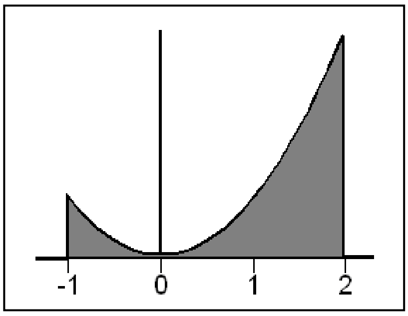
$$f(x) = \begin{cases} \frac{1}{3}x^2 \ ; -1 < x < 2 \\ 0 \ ; \ elsewhere \end{cases}$$

1. (a) f(x) ≥ 0 because f(x) is a quadratic function.

(b) $\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{-1} 0\,dx + \int_{-1}^{2} \frac{1}{3}x^2\,dx + \int_{2}^{\infty} 0\,dx$

$= \int_{-1}^{2} \frac{1}{3}x^2\,dx = \left[\frac{1}{9}x^3 \ \middle| \begin{matrix} x=2 \\ x=-1 \end{matrix} \right]$

$= \frac{1}{9}(8-(-1)) = 1$

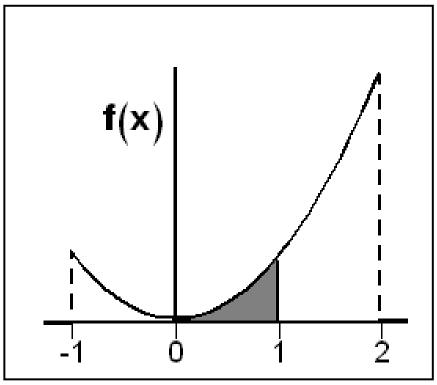


2. P(0<X≤1) = $\int_{0}^{1} f(x)\,dx = \int_{0}^{1} \frac{1}{3}x^2\,dx$

$= \left[\frac{1}{9}x^3 \ \middle| \begin{matrix} x=1 \\ x=0 \end{matrix} \right]$

$= \frac{1}{9}(1-(0))$

$= \frac{1}{9}$



Definition 3.7:
The cumulative distribution function (CDF), F(x), of a continuous random variable X with probability density function f(x) is given by:

$$F(x) = P(X\le x)= \int_{-\infty}^{x} f(t)\,dt \;;\; \text{for } -\infty<x<\infty$$

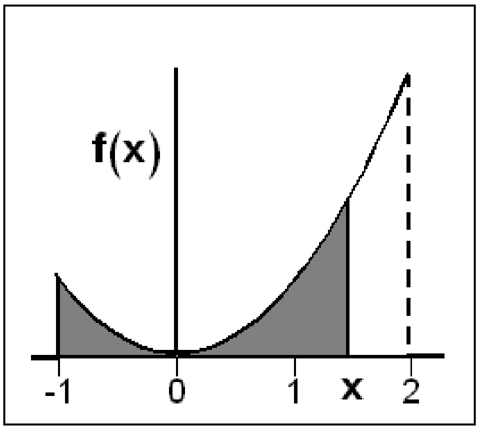
Result:
$P(a < X \le b) = P(X \le b) - P(X \le a) = F(b) - F(a)$

Example:
in Example 3.6,

1. Find the CDF
2. Using the CDF, find $P(0<X\le 1)$.

Solution:

$$f(x)=\begin{cases}\frac{1}{3}x^2 \;;\; -1<x<2\\ 0 \;;\; elsewhere\end{cases}$$



For $x< -1$:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0\,dt = 0$$

For $-1\le x<2$:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{-1} 0\,dt + \int_{-1}^{x} \frac{1}{3}t^2\,dt$$

$$= \int_{-1}^{x} \frac{1}{3}t^2\,dt$$

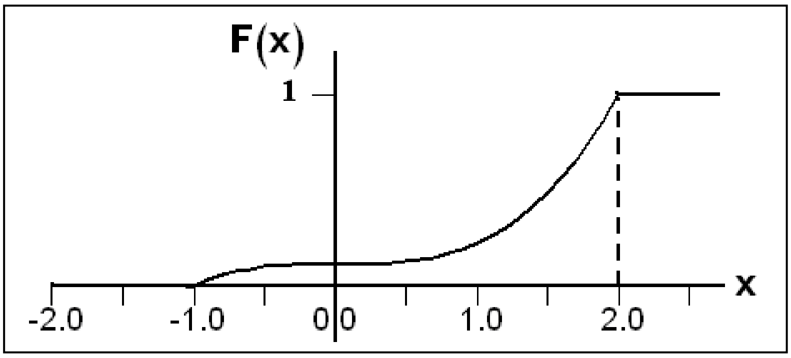
$$= \left[\frac{1}{9}t^3 \;\Big|_{t=-1}^{t=x}\right] = \frac{1}{9}(x^3-(-1)) = \frac{1}{9}(x^3+1)$$

For $x\ge 2$:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{-1} 0\,dt + \int_{-1}^{2} \frac{1}{3}t^2\,dt + \int_{2}^{x} 0\,dt = \int_{-1}^{2} \frac{1}{3}t^2\,dt = 1.$$

Therefore, the CDF is:

$$F(x)=P(X\le x)=\begin{cases}0 \;;\; x<-1\\ \frac{1}{9}(x^3+1) \;;\; -1\le x<2\\ 1 \;;\; x\ge 2\end{cases}$$



2. Using the CDF,

$P(0<X\leq 1) = F(1) - F(0) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$

Chapter 4: Mathematical Expectation:

4.1 Mean of a Random Variable:

Definition 4.1:

Let X be a random variable with a probability distribution f(x). The mean (or expected value) of X is denoted by μ_X (or E(X)) and is defined by:

$$E(X)=\mu_X=\begin{cases}\sum\limits_{all\ x} x\, f(x)\,; & if\ \ X\ \ is\ discrete\\ \int\limits_{-\infty}^{\infty} x\, f(x)\, dx\,; & if\ \ X\ \ is\ continuous\end{cases}$$

Example 4.1: (Reading Assignment)

Example: (Example 3.3)

A shipment of 8 similar microcomputers to a retail outlet contains 3 that are defective and 5 are non-defective. If a school makes a random purchase of 2 of these computers, find the expected number of defective computers purchased

Solution:

Let X = the number of defective computers purchased.

In Example 3.3, we found that the probability distribution of X is:

x	0	1	2
f(x)= P(X=x)	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

or:

$$f(x)=P(X=x)=\begin{cases}\dfrac{\binom{3}{x}\times\binom{5}{2-x}}{\binom{8}{2}}; & x=0,1,2\\ 0;\ otherwise\end{cases}$$

The expected value of the number of defective computers purchased is the mean (or the expected value) of X, which is:

$$E(X)=\mu_X=\sum_{x=0}^{2} x\, f(x)$$

$$= (0)\ f(0) + (1)\ f(1) + (2)\ f(2)$$

$$= (0)\frac{10}{28} + (1)\frac{15}{28} + (2)\frac{3}{28}$$

$$= \frac{15}{28} + \frac{6}{28} = \frac{21}{28} = 0.75 \text{ (computers)}$$

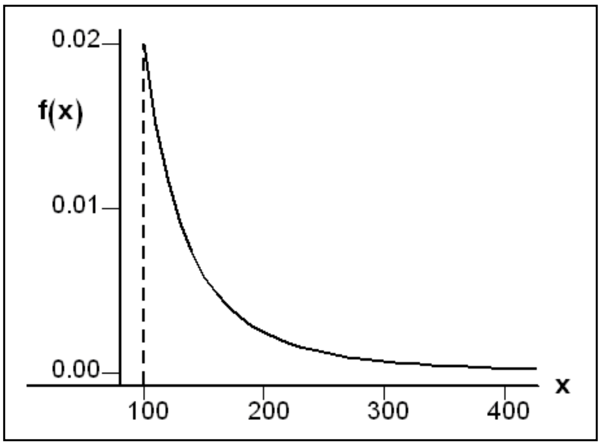
Example 4.3:

Let X be a continuous random variable that represents the life (in hours) of a certain electronic device. The pdf of X is given by:

$$f(x) = \begin{cases} \dfrac{20{,}000}{x^3} \; ; & x > 100 \\ 0 \; ; & elsewhere \end{cases}$$

Find the expected life of this type of devices.

Solution:



$$E(X) = \mu_X = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

$$= \int_{100}^{\infty} x \frac{20000}{x^3} dx$$

$$= 20000 \int_{100}^{\infty} \frac{1}{x^2} dx$$

$$= 20000 \left[-\frac{1}{x} \Big|_{x=100}^{x=\infty} \right]$$

$$= -20000 \left[0 - \frac{1}{100} \right] = 200 \text{ (hours)}$$

Therefore, we expect that this type of electronic devices to last, on average, 200 hours.

Theorem 4.1:

Let X be a random variable with a probability distribution f(x), and let g(X) be a function of the random variable X. The mean (or expected value) of the random variable g(X) is denoted by $\mu_{g(X)}$ (or E[g(X)]) and is defined by:

$$E[g(X)] = \mu_{g(X)} = \begin{cases} \sum_{all\ x} g(x)\, f(x) \; ; & if\ X\ is\ discrete \\ \int_{-\infty}^{\infty} g(x)\, f(x)\, dx \; ; & if\ X\ is\ continuous \end{cases}$$

Example:

Let X be a discrete random variable with the following probability distribution

x	0	1	2
f(x)	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

Find E[g(X)], where $g(X)=(X-1)^2$.

Solution:

$g(X)=(X-1)^2$

$$E[g(X)]=\mu_{g(X)}=\sum_{x=0}^{2} g(x)\,f(x)=\sum_{x=0}^{2}(x-1)^2\,f(x)$$

$$=(0-1)^2\,f(0)+(1-1)^2\,f(1)+(2-1)^2\,f(2)$$

$$=(-1)^2\,\frac{10}{28}+(0)^2\,\frac{15}{28}+(1)^2\,\frac{3}{28}$$

$$=\frac{10}{28}+0+\frac{3}{28}=\frac{13}{28}$$

Example:

In Example 4.3, find $E\left(\frac{1}{X}\right)$.

Solution:

$$f(x)=\begin{cases}\frac{20,000}{x^3}\ ; & x>100\\ 0\ ; & elsewhere\end{cases}$$

$$g(X)=\frac{1}{X}$$

$$E\left(\frac{1}{X}\right)=E[g(X)]=\mu_{g(X)}=\int_{-\infty}^{\infty} g(x)\,f(x)\,dx=\int_{-\infty}^{\infty}\frac{1}{x}\,f(x)\,dx$$

$$=\int_{100}^{\infty}\frac{1}{x}\frac{20000}{x^3}dx=20000\int_{100}^{\infty}\frac{1}{x^4}dx=\frac{20000}{-3}\left[\frac{1}{x^3}\Bigg|_{x=100}^{x=\infty}\right]$$

$$=\frac{-20000}{3}\left[0-\frac{1}{1000000}\right]=0.0067$$

<u>4.2 Variance (of a Random Variable):</u>

The most important measure of variability of a random variable X is called the variance of X and is denoted by Var(X) or σ_X^2.

Definition 4.3:

Let X be a random variable with a probability distribution f(x) and mean μ. The variance of X is defined by:

$$\text{Var}(X)=\sigma_X^2=E[(X-\mu)^2]=\begin{cases}\sum\limits_{all\ x}(x-\mu)^2\ f(x)\ ; & if\ X\ is\ discrete\\ \int\limits_{-\infty}^{\infty}(x-\mu)^2\ f(x)\,dx\ ; & if\ X\ is\ continuous\end{cases}$$

Definition:

The positive square root of the variance of X, $\sigma_X=\sqrt{\sigma_X^2}$, is called the standard deviation of X.

Note:

Var(X)=E[g(X)], where $g(X)=(X-\mu)^2$

Theorem 4.2:

The variance of the random variable X is given by:

$$\text{Var}(X)=\sigma_X^2=E(X^2)-\mu^2$$

where $E(X^2)=\begin{cases}\sum\limits_{all\ x}x^2\ f(x)\ ; & if\ X\ is\ discrete\\ \int\limits_{-\infty}^{\infty}x^2\ f(x)\,dx\ ; & if\ X\ is\ continuous\end{cases}$

Example 4.9:

Let X be a discrete random variable with the following probability distribution

x	0	1	2	3
f(x)	0.51	0.38	0.10	0.01

Find $\text{Var}(X)=\sigma_X^2$.

Solution:

$\mu=\sum\limits_{x=0}^{3}x\,f(x)$ = (0) f(0) + (1) f(1) +(2) f(2) + (3) f(3)

= (0) (0.51) + (1) (0.38) +(2) (0.10) + (3) (0.01)

= 0.61

1. First method:

$$\text{Var}(X)=\sigma_X^2=\sum_{x=0}^{3}(x-\mu)^2\ f(x)$$

$= \sum_{x=0}^{3}(x-0.61)^2 f(x)$

$=(0-0.61)^2 f(0)+(1-0.61)^2 f(1)+(2-0.61)^2 f(2)+(3-0.61)^2 f(3)$

$=(-0.61)^2 (0.51)+(0.39)^2 (0.38)+(1.39)^2 (0.10)+(2.39)^2 (0.01)$

$= 0.4979$

2. Second method:

$\text{Var}(X)=\sigma_X^2 = E(X^2)-\mu^2$

$E(X^2)= \sum_{x=0}^{3} x^2 f(x) = (0^2) f(0) + (1^2) f(1) + (2^2) f(2) + (3^2) f(3)$

$= (0)(0.51) + (1)(0.38) + (4)(0.10) + (9)(0.01)$

$= 0.87$

$\text{Var}(X)=\sigma_X^2 = E(X^2)-\mu^2 = 0.87-(0.61)^2 = 0.4979$

Example 4.10:

Let X be a continuous random variable with the following pdf:

$$f(x)=\begin{cases} 2(x-1) \; ; \; 1<x<2 \\ 0 \; ; \; elsewhere \end{cases}$$

Find the mean and the variance of X.

Solution:

$\mu = E(X)= \int_{-\infty}^{\infty} x\, f(x)\,dx = \int_{1}^{2} x[2(x-1)]\,dx = 2\int_{1}^{2} x(x-1)\,dx = 5/3$

$E(X^2)= \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_{1}^{2} x^2[2(x-1)]\,dx = 2\int_{1}^{2} x^2(x-1)\,dx = 17/6$

$\text{Var}(X)=\sigma_X^2 = E(X^2)-\mu^2 = 17/6-(5/3)^2 = 1/18$

4.3 Means and Variances of Linear Combinations of Random Variables:

If $X_1, X_2, \ldots, X_n$ are n random variables and $a_1, a_2, \ldots, a_n$ are constants, then the random variable :

$$Y=\sum_{i=1}^{n} a_i X_i = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$$

is called a linear combination of the random variables $X_1, X_2, \ldots, X_n$.

Theorem 4.5:

If X is a random variable with mean $\mu=E(X)$, and if a and b are constants, then:

$$E(aX \pm b) = a\,E(X) \pm b$$

$$\Leftrightarrow$$

$$\mu_{aX \pm b} = a\,\mu_X \pm b$$

Corollary 1: E(b) = b (a=0 in Theorem 4.5)

Corollary 2: E(aX) = a E(X) (b=0 in Theorem 4.5)

Example 4.16:

Let X be a random variable with the following probability density function:

$$f(x) = \begin{cases} \frac{1}{3}x^2 \;;\; -1 < x < 2 \\ 0 \;;\; elsewhere \end{cases}$$

Find E(4X+3).

Solution:

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{-1}^{2} x[\frac{1}{3}x^2]\, dx = \frac{1}{3}\int_{-1}^{2} x^3\, dx = \frac{1}{3}\left[\frac{1}{4}x^4 \Bigg|_{x=-1}^{x=2}\right] = 5/4$$

E(4X+3) = 4 E(X)+3 = 4(5/4) + 3=8

Another solution:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx \qquad ;\; g(X) = 4X+3$$

$$E(4X+3) = \int_{-\infty}^{\infty} (4x+3)\, f(x)\, dx = \int_{-1}^{2} (4x+3)\, [\frac{1}{3}x^2]\, dx = \cdots = 8$$

Theorem:

If $X_1, X_2, \ldots, X_n$ are n random variables and $a_1, a_2, \ldots, a_n$ are constants, then:

$$E(a_1X_1 + a_2X_2 + \ldots + a_nX_n) = a_1E(X_1) + a_2E(X_2) + \ldots + a_nE(X_n)$$

$$\Leftrightarrow$$

$$E(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i E(X_i)$$

Corollary:

If X, and Y are random variables, then:

$$E(X \pm Y) = E(X) \pm E(Y)$$

Theorem 4.9:

If X is a random variable with variance $Var(X) = \sigma_X^2$ and if a and b are constants, then:

$$Var(aX \pm b) = a^2 \, Var(X)$$

$$\Leftrightarrow$$

$$\sigma^2_{aX+b} = a^2 \sigma^2_X$$

Theorem:

If $X_1, X_2, \ldots, X_n$ are n <u>independent</u> random variables and $a_1, a_2, \ldots, a_n$ are constants, then:

$$Var(a_1X_1+a_2X_2+\ldots+a_nX_n)$$
$$= a_1^2 Var(X_1) + a_2^2 Var(X_2) + \ldots + a_n^2 Var(X_n)$$

$$\Leftrightarrow$$

$$Var(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i)$$

$$\Leftrightarrow$$

$$\sigma^2_{a_1X_1+a_2X_2+\ldots+a_nX_n} = a_1^2\sigma^2_{X_1} + a_2^2\sigma^2_{X_2} + \ldots + a_n^2\sigma^2_{X_n}$$

Corollary:

If X, and Y are <u>independent</u> random variables, then:

- $Var(aX+bY) = a^2 \, Var(X) + b^2 \, Var(Y)$
- $Var(aX-bY) = a^2 \, Var(X) + b^2 \, Var(Y)$
- $Var(X \pm Y) = Var(X) + Var(Y)$

Example:

Let X, and Y be two independent random variables such that E(X)=2, Var(X)=4, E(Y)=7, and Var(Y)=1. Find:

1. E(3X+7) and Var(3X+7)
2. E(5X+2Y−2) and Var(5X+2Y−2).

Solution:

1. $E(3X+7) = 3E(X)+7 = 3(2)+7 = 13$

   $Var(3X+7) = (3)^2 \, Var(X) = (3)^2 \, (4) = 36$

2. $E(5X+2Y-2) = 5E(X) + 2E(Y) - 2 = (5)(2) + (2)(7) - 2 = 22$

   $Var(5X+2Y-2) = Var(5X+2Y) = 5^2 \, Var(X) + 2^2 \, Var(Y)$

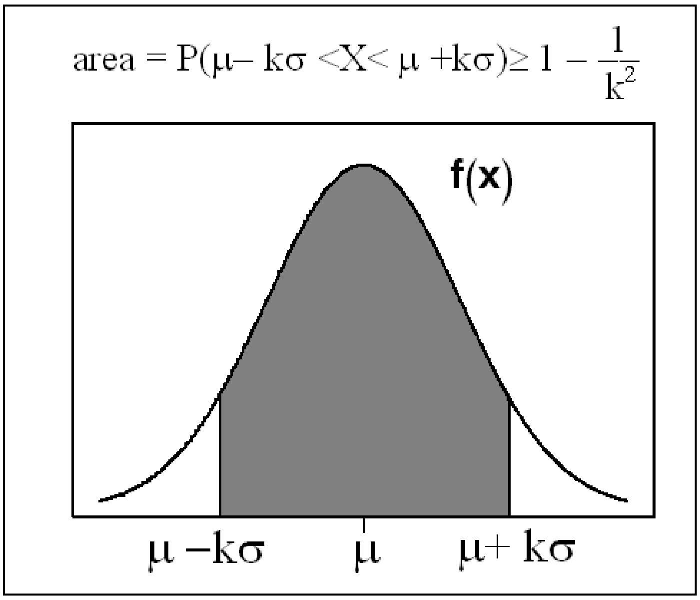
   $= (25)(4)+(4)(1) = 104$

4.4 Chebyshev's Theorem:

* Suppose that X is any random variable with mean $E(X)=\mu$ and variance $Var(X)=\sigma^2$ and standard deviation σ.

* Chebyshev's Theorem gives a conservative estimate of the probability that the random variable X assumes a value within k standard deviations ($k\sigma$) of its mean μ, which is $P(\mu - k\sigma < X < \mu + k\sigma)$.

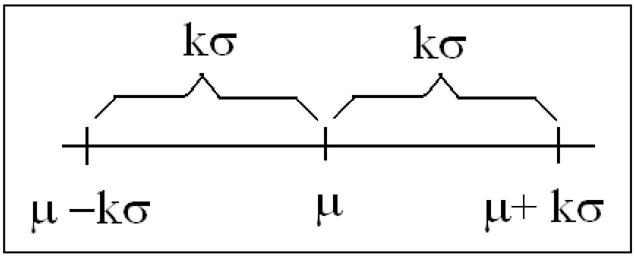
* $P(\mu - k\sigma < X < \mu + k\sigma) \approx 1 - \frac{1}{k^2}$



Theorem 4.11:(Chebyshev's Theorem)

Let X be a random variable with mean $E(X)=\mu$ and variance $Var(X)=\sigma^2$, then for k>1, we have:

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2} \quad \Leftrightarrow \quad P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$



Example 4.22:

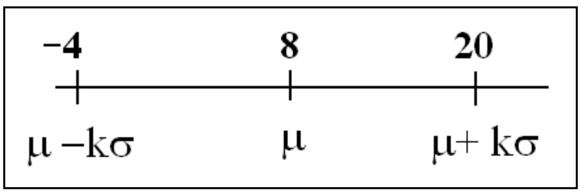
Let X be a random variable having an unknown distribution with mean $\mu=8$ and variance $\sigma^2=9$ (standard deviation $\sigma=3$). Find the following probability:

(a) $P(-4 < X < 20)$

(b) $P(|X-8| \geq 6)$

Solution:

(a) $P(-4 < X < 20) = ??$



$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

$(-4 < X < 20) = (\mu - k\sigma < X < \mu + k\sigma)$

$-4 = \mu - k\sigma \Leftrightarrow -4 = 8 - k(3)$ or $20 = \mu + k\sigma \Leftrightarrow 20 = 8 + k(3)$

$\Leftrightarrow -4 = 8 - 3k$ $\quad\quad \Leftrightarrow 20 = 8 + 3k$

$\Leftrightarrow 3k = 12$ $\quad\quad \Leftrightarrow 3k = 12$

$\Leftrightarrow$ **k=4** $\quad\quad \Leftrightarrow$ **k=4**

$$1 - \frac{1}{k^2} = 1 - \frac{1}{16} = \frac{15}{16}$$

Therefore, $P(-4 < X < 20) \geq \frac{15}{16}$, and hence, $P(-4 < X < 20) \approx \frac{15}{16}$ (approximately)

(b) $P(|X-8| \geq 6) = ??$

$P(|X-8| \geq 6) = 1 - P(|X-8| < 6)$

$P(|X-8| < 6) = ??$

$P(|X-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}$

$(|X-8| < 6) = (|X-\mu| < k\sigma)$

$6 = k\sigma \Leftrightarrow 6 = 3k \Leftrightarrow$ **k=2**

$$1 - \frac{1}{k^2} = 1 - \frac{1}{4} = \frac{3}{4}$$

$P(|X-8| < 6) \geq \frac{3}{4} \Leftrightarrow 1 - P(|X-8| < 6) \leq 1 - \frac{3}{4}$

$\Leftrightarrow 1 - P(|X-8| < 6) \leq \frac{1}{4}$

$\Leftrightarrow P(|X-8| \geq 6) \leq \frac{1}{4}$

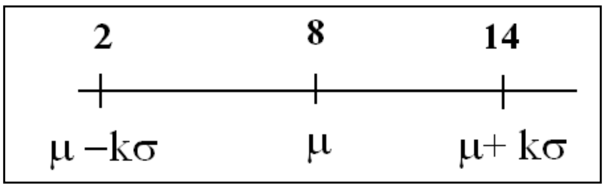
Therefore, $P(|X-8| \geq 6) \approx \frac{1}{4}$ (approximately)

Another solution for part (b):

$P(|X-8| < 6) = P(-6 < X-8 < 6)$

$= P(-6+8 < X < 6+8)$

$= P(2 < X < 14)$



$(2 < X < 14) = (\mu - k\sigma < X < \mu + k\sigma)$

$2 = \mu - k\sigma \Leftrightarrow 2 = 8 - k(3) \Leftrightarrow 2 = 8 - 3k \Leftrightarrow 3k = 6 \Leftrightarrow$ **k=2**

$$1-\frac{1}{k^2}=1-\frac{1}{4} = \frac{3}{4}$$

$$P(2<X < 14) \geq \frac{3}{4} \Leftrightarrow P(|X-8| < 6) \geq \frac{3}{4}$$

$$\Leftrightarrow 1 - P(|X-8| < 6) \leq 1 - \frac{3}{4}$$

$$\Leftrightarrow 1 - P(|X-8| < 6) \leq \frac{1}{4}$$

$$\Leftrightarrow P(|X-8| \geq 6) \leq \frac{1}{4}$$

Therefore, $P(|X-8| \geq 6) \approx \frac{1}{4}$ (approximately)

Chapter 5: Some Discrete Probability Distributions:

5.2: Discrete Uniform Distribution:

If the discrete random variable X assumes the values x_1, x_2, …, x_k with equal probabilities, then X has the discrete uniform distribution given by:

$$f(x) = P(X = x) = f(x;k) = \begin{cases} \frac{1}{k} \; ; \; x = x_1, x_2, \cdots, x_k \\ 0 \; ; \; elsewhere \end{cases}$$

Note:

- f(x)=f(x;*k*)=P(X=x)
- *k* is called the parameter of the distribution.

Example 5.2:

- Experiment: tossing a balanced die.
- Sample space: *S*={1,2,3,4,5,6}
- Each sample point of *S* occurs with the same probability 1/6.
- Let X= the number observed when tossing a balanced die.
- The probability distribution of X is:

$$f(x) = P(X = x) = f(x;6) = \begin{cases} \frac{1}{6} \; ; \; x = 1, 2, \cdots, 6 \\ 0 \; ; \; elsewhere \end{cases}$$

Theorem 5.1:

If the discrete random variable X has a discrete uniform distribution with parameter k, then the mean and the variance of X are:

$$E(X) = \mu = \frac{\sum_{i=1}^{k} x_i}{k}$$

$$Var(X) = \sigma^2 = \frac{\sum_{i=1}^{k} (x_i - \mu)^2}{k}$$

Example 5.3:

Find E(X) and Var(X) in Example 5.2.

Solution:

$$E(X) = \mu = \frac{\sum_{i=1}^{k} x_i}{k} = \frac{1+2+3+4+5+6}{6} = 3.5$$

$$Var(X) = \sigma^2 = \frac{\sum_{i=1}^{k} (x_i - \mu)^2}{k} = \frac{\sum_{i=1}^{k} (x_i - 3.5)^2}{6}$$

$$= \frac{(1-3.5)^2 + (2-3.5)^2 + \cdots + (6-3.5)^2}{6} = \frac{35}{12}$$

5.3 Binomial Distribution:

Bernoulli Trial:

- Bernoulli trial is an experiment with only two possible outcomes.
- The two possible outcomes are labeled:
  success (s) and failure (f)
- The probability of success is P(s)=p and the probability of failure is P(f)= q = 1−p.
- Examples:
  1. Tossing a coin (success=H, failure=T, and p=P(H))
  2. Inspecting an item (success=defective, failure=non-defective, and p=P(defective))

Bernoulli Process:

Bernoulli process is an experiment that must satisfy the following properties:

1. The experiment consists of n repeated Bernoulli trials.
2. The probability of success, P(s)=p, remains constant from trial to trial.
3. The repeated trials are independent; that is the outcome of one trial has no effect on the outcome of any other trial

Binomial Random Variable:

Consider the random variable :

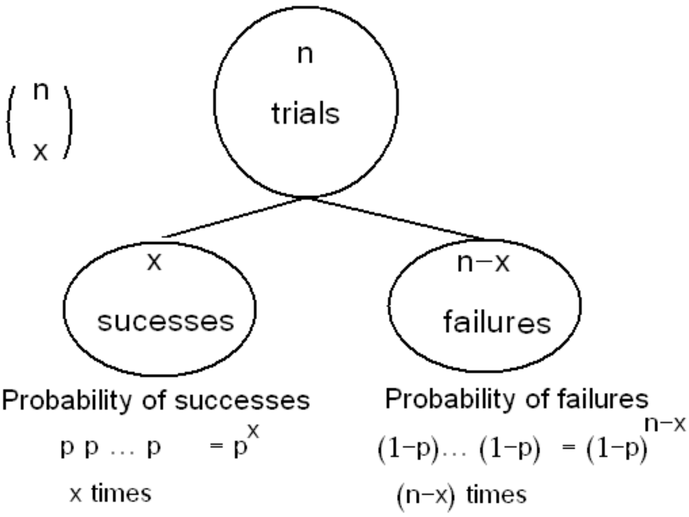
X = The number of successes in the n trials in a Bernoulli process

The random variable *X* has a binomial distribution with parameters *n* (number of trials) and *p* (probability of success), and we write:

$$X \sim \text{Binomial}(n,p) \quad \text{or} \quad X \sim b(x;n,p)$$

The probability distribution of *X* is given by:

$$f(x) = P(X=x) = b(x;n,p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} \; ; & x = 0, 1, 2, \dots, n \\ 0 \; ; & otherwise \end{cases}$$



We can write the probability distribution of *X* as a table as follows.

x	f(x)=P(X=x)=b(x;*n*,*p*)
0	$\binom{n}{0} p^0 (1-p)^{n-0} = (1-p)^n$
1	$\binom{n}{1} p^1 (1-p)^{n-1}$
2	$\binom{n}{2} p^2 (1-p)^{n-2}$
⋮	⋮
$n-1$	$\binom{n}{n-1} p^{n-1} (1-p)^1$
n	$\binom{n}{n} p^n (1-p)^0 = p^n$
Total	1.00

Example:

Suppose that 25% of the products of a manufacturing process are defective. Three items are selected at random, inspected, and classified as defective (D) or non-defective (N). Find the probability distribution of the number of defective items.

Solution:

- Experiment: selecting 3 items at random, inspected, and classified as (D) or (N).
- The sample space is
  S={DDD,DDN,DND,DNN,NDD,NDN,NND,NNN}
- Let X = the number of defective items in the sample
- We need to find the probability distribution of X.

(1) First Solution:

Outcome	Probability	x
NNN	$\frac{3}{4}\times\frac{3}{4}\times\frac{3}{4}=\frac{27}{64}$	0
NND	$\frac{3}{4}\times\frac{3}{4}\times\frac{1}{4}=\frac{9}{64}$	1
NDN	$\frac{3}{4}\times\frac{1}{4}\times\frac{3}{4}=\frac{9}{64}$	1
NDD	$\frac{3}{4}\times\frac{1}{4}\times\frac{1}{4}=\frac{3}{64}$	2
DNN	$\frac{1}{4}\times\frac{3}{4}\times\frac{3}{4}=\frac{9}{64}$	1
DND	$\frac{1}{4}\times\frac{3}{4}\times\frac{1}{4}=\frac{3}{64}$	2
DDN	$\frac{1}{4}\times\frac{1}{4}\times\frac{3}{4}=\frac{3}{64}$	2
DDD	$\frac{1}{4}\times\frac{1}{4}\times\frac{1}{4}=\frac{1}{64}$	3

The probability distribution .of X is

.x	.f(x)=P(X=x)
0	$\frac{27}{64}$
1	$\frac{9}{64}+\frac{9}{64}+\frac{9}{64}=\frac{27}{64}$
2	$\frac{3}{64}+\frac{3}{64}+\frac{3}{64}=\frac{9}{64}$
3	$\frac{1}{64}$

(2) Second Solution:

Bernoulli trial is the process of inspecting the item. The results are success=D or failure=N, with probability of success P(s)=25/100=1/4=0.25.

The experiments is a Bernoulli process with:

- number of trials: n=3
- Probability of success: p=1/4=0.25
- X ~ Binomial(n,p)=Binomial(3,1/4)

- The probability distribution of X is given by:

$$f(x)=P(X=x)=b(x;3,\frac{1}{4})=\begin{cases}\binom{3}{x}(\frac{1}{4})^x(\frac{3}{4})^{3-x}; & x=0,1,2,3\\ 0\ ; & otherwise\end{cases}$$

$$f(0)=P(X=0)=b(0;3,\frac{1}{4})=\binom{3}{0}(\frac{1}{4})^0(\frac{3}{4})^3=\frac{27}{64}$$

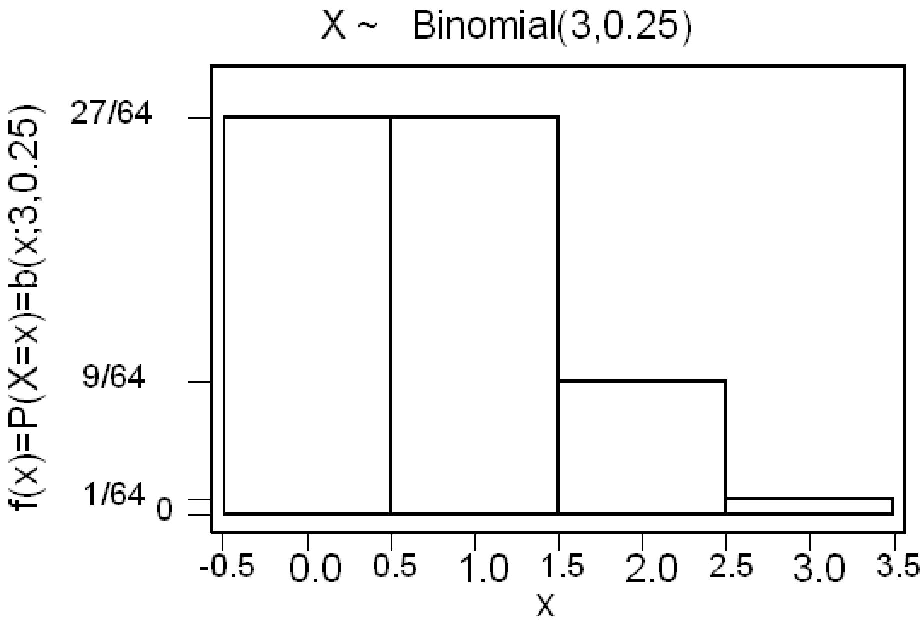
$$f(1)=P(X=1)=b(1;3,\frac{1}{4})=\binom{3}{1}(\frac{1}{4})^1(\frac{3}{4})^2=\frac{27}{64}$$

$$f(2)=P(X=2)=b(2;3,\frac{1}{4})=\binom{3}{2}(\frac{1}{4})^2(\frac{3}{4})^1=\frac{9}{64}$$

$$f(3)=P(X=3)=b(3;3,\frac{1}{4})=\binom{3}{3}(\frac{1}{4})^3(\frac{3}{4})^0=\frac{1}{64}$$

The probability distribution of X is

x	f(x)=P(X=x) =b(x;3,1/4)
0	27/64
1	27/64
2	9/64
3	1/64



Theorem 5.2:

The mean and the variance of the binomial distribution b(x;n,p) are:

$$\mu = n\,p$$
$$\sigma^2 = n\,p\,(1-p)$$

Example:

In the previous example, find the expected value (mean) and the variance of the number of defective items.

Solution:

- X = number of defective items
- We need to find E(X)=μ and Var(X)=σ^2
- We found that X ~ Binomial(n,p)=Binomial(3,1/4)
- .n=3 and p=1/4

The expected number of defective items is

$$E(X)=\mu = n\,p = (3)\,(1/4) = 3/4 = 0.75$$

The variance of the number of defective items is

$$Var(X)=\sigma^2 = n\,p\,(1-p) = (3)\,(1/4)\,(3/4) = 9/16 = 0.5625$$

Example:

In the previous example, find the following probabilities:

(1) The probability of getting at least two defective items.

(2) The probability of getting at most two defective items.

Solution:

X ~ Binomial(3,1/4)

$$f(x) = P(X = x) = b(x;3,\frac{1}{4}) = \begin{cases} \binom{3}{x} (\frac{1}{4})^x\, (\frac{3}{4})^{3-x} & for\;\; x = 0, 1, 2, 3 \\ 0 & otherwise \end{cases}$$

.x	.f(x)=P(X=x)=b(x;3,1/4)
0	27/64
1	27/64
2	9/64
3	1/64

(1) The probability of getting at least two defective items:

$$P(X\geq 2)=P(X=2)+P(X=3)= f(2)+f(3)=\frac{9}{64}+\frac{1}{64}=\frac{10}{64}$$

(2) The probability of getting at most two defective item:

$$P(X\leq 2) = P(X=0)+P(X=1)+P(X=2)$$

$$= f(0)+f(1)+f(2) = \frac{27}{64}+\frac{27}{64}+\frac{9}{64}=\frac{63}{64}$$

or

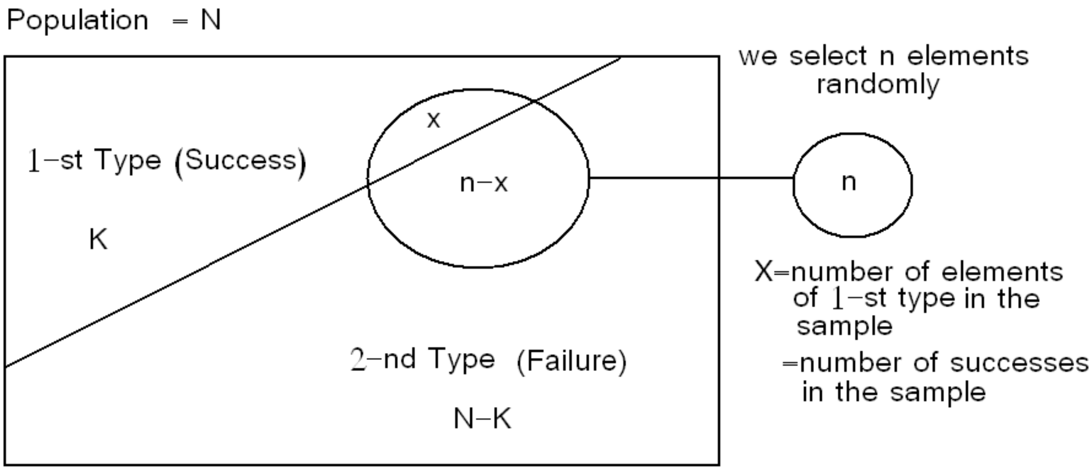
$$P(X\leq 2)= 1-P(X>2) = 1-P(X=3) = 1- f(3) = 1-\frac{1}{64}=\frac{63}{64}$$

Example 5.4: Reading assignment

Example 5.5: Reading assignment

Example 5.6: Reading assignment

5.4 Hypergeometric Distribution :



- Suppose there is a population with 2 types of elements:
  1-st Type = success
  2-nd Type = failure
- N= population size
- K= number of elements of the 1-st type
- $N - K$ = number of elements of the 2-nd type
- We select a sample of n elements at random from the population
- Let X = number of elements of 1-st type (number of successes) in the sample
- We need to find the probability distribution of X.

There are to two methods of selection:

1. selection with replacement
2. selection without replacement

(1) If we select the elements of the sample at random and with replacement, then

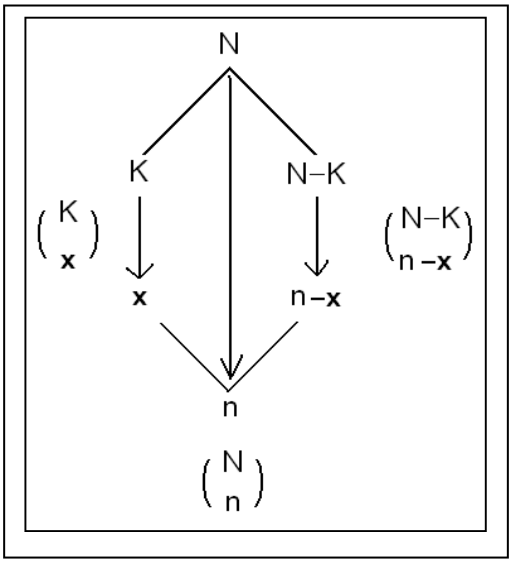
$$X \sim \text{Binomial}(n,p)\text{; where } p = \frac{K}{N}$$

(2) Now, suppose we select the elements of the sample at random and without replacement. When the selection is made without replacement, the random variable X has a hypergeometric distribution with parameters N, n, and K. and we write X~h(x;N,n,K).

The probability distribution of X is given by:

$$f(x) = P(X = x) = h(x; N, n, K)$$

$$= \begin{cases} \dfrac{\binom{K}{x} \times \binom{N-K}{n-x}}{\binom{N}{n}}; & x = 0, 1, 2, \cdots, n \\ 0; & otherwise \end{cases}$$



Note that the values of X must satisfy:

$0 \le x \le K$ and $0 \le n - x \le N - K$

$\Leftrightarrow$

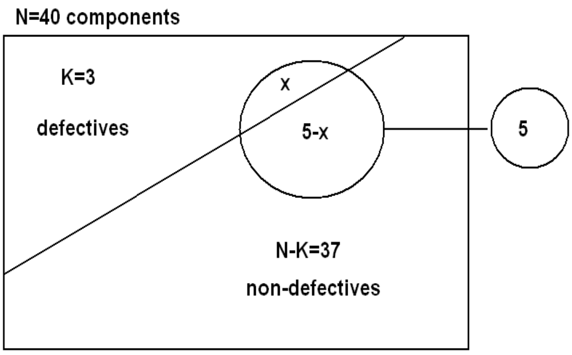
$0 \le x \le K$ and $n - N + K \le x \le n$

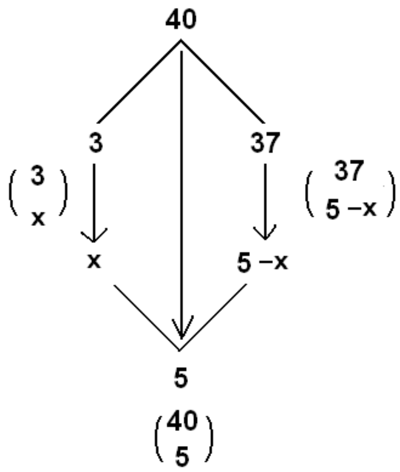
Example 5.8: Reading assignment

Example 5.9:

Lots of 40 components each are called acceptable if they contain no more than 3 defectives. The procedure for sampling the lot is to select 5 components at random (without replacement) and to reject the lot if a defective is found. What is the probability that exactly one defective is found in the sample if there are 3 defectives in the entire lot.

Solution:





- Let X= number of defectives in the sample
- N=40, K=3, and n=5
- X has a hypergeometric distribution with parameters N=40, n=5, and K=3.
- X~h(x;N,n,K)=h(x;40,5,3).
- The probability distribution of X is given by:

$$f(x) = P(X = x) = h(x;40,5,3) = \begin{cases} \dfrac{\binom{3}{x} \times \binom{37}{5-x}}{\binom{40}{5}}; & x = 0, 1, 2, \cdots, 5 \\ 0; & otherwise \end{cases}$$

But the values of X must satisfy:

$0 \le x \le K$ and $n - N + K \le x \le n \Leftrightarrow 0 \le x \le 3$ and $-42 \le x \le 5$

Therefore, the probability distribution of *X* is given by:

$$f(x) = P(X = x) = h(x;40,5,3) = \begin{cases} \dfrac{\binom{3}{x} \times \binom{37}{5-x}}{\binom{40}{5}}; & x = 0, 1, 2, 3 \\ 0; & otherwise \end{cases}$$

Now, the probability that exactly one defective is found in the sample is

$$.f(1)=P(X=1)=h(1;40,5,3)= \frac{\binom{3}{1} \times \binom{37}{5-1}}{\binom{40}{5}} = \frac{\binom{3}{1} \times \binom{37}{4}}{\binom{40}{5}} = 0.3011$$

Theorem 5.3:

The mean and the variance of the hypergeometric distribution h(x;*N*,*n*,*K*) are:

$$\mu = n\frac{K}{N}$$

$$\sigma^2 = n\frac{K}{N}\left(1 - \frac{K}{N}\right)\frac{N-n}{N-1}$$

Example 5.10:

In Example 5.9, find the expected value (mean) and the variance of the number of defectives in the sample.

Solution:

- *X* = number of defectives in the sample
- We need to find E(X)=μ and Var(X)=σ^2
- We found that *X* ~ h(x;40,5,3)
- *N*=40, *n*=5, and *K*=3

The expected number of defective items is

$$E(X)=\mu = n\frac{K}{N} = 5\times\frac{3}{40} = 0.375$$

The variance of the number of defective items is

$$Var(X)=\sigma^2 = n\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1} = 5\times\frac{3}{40}\left(1-\frac{3}{40}\right)\frac{40-5}{40-1} = 0.311298$$

Relationship to the binomial distribution:

* Binomial distribution: $b(x;n,p)=\binom{n}{x}p^x(1-p)^{n-x};\ x=0,1,\ldots,n$

* Hypergeometric distribution: $h(x;N,n,K)=\dfrac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}};x=0,1,\cdots,n$

If n is small compared to N and K, then the hypergeometric distribution h(x;N,n,K) can be approximated by the binomial distribution b(x;n,p), where $p=\frac{K}{N}$; i.e., for large N and K and small n, we have:

$$h(x;N,n,K) \approx b(x;n,\frac{K}{N})$$

$$\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \approx \binom{n}{x}\left(\frac{K}{N}\right)^x\left(1-\frac{K}{N}\right)^{n-x};x=0,1,\cdots,n$$

Note:
If n is small compared to N and K, then there will be almost no difference between selection without replacement and selection with replacement ($\frac{K}{N}\approx\frac{K-1}{N-1}\approx\cdots\approx\frac{K-n+1}{N-n+1}$).

Example 5.11: Reading assignment

5.6 Poisson Distribution:

- Poisson experiment is an experiment yielding numerical values of a random variable that count the number of outcomes occurring in a given time interval or a specified region denoted by *t*.

  X = The number of outcomes occurring in a given time interval or a specified region denoted by *t*.
- Example:
  1. X = number of field mice per acre (t= 1 acre)
  2. X= number of typing errors per page (t=1 page)
  3. X=number of telephone calls received every day (t=1 day)
  4. X=number of telephone calls received every 5 days (t=5 days)
- Let λ be the average (mean) number of outcomes per unit time or unit region (t=1).
- The average (mean) number of outcomes (mean of X) in the time interval or region *t* is:

$$\mu = \lambda\, t$$

- The random variable X is called a Poisson random variable with parameter μ ($\mu=\lambda t$), and we write X~Poisson(μ), if its probability distribution is given by:

$$f(x) = P(X = x) = p(x;\mu) = \begin{cases} \dfrac{e^{-\mu}\mu^{x}}{x!} & ; \quad x = 0, 1, 2, 3, \ldots \\ 0 & ; \quad otherwise \end{cases}$$

Theorem 5.5:

The mean and the variance of the Poisson distribution Poisson(x;μ) are:

$$\mu = \lambda t$$
$$\sigma^2 = \mu = \lambda t$$

Note:

- λ is the average (mean) of the distribution in the unit time (t=1).
- If X=The number of calls received in a month (unit time t=1 month) and X~Poisson(λ), then:

Unrevised First Draft

(i) Y = number of calls received in a year.
Y ~ Poisson (μ); μ=12λ (t=12)
(ii) W = number of calls received in a day.
W ~ Poisson (μ); μ=λ/30 (t=1/30)

Example 5.16: Reading Assignment

Example 5.17: Reading Assignment

Example:

Suppose that the number of typing errors per page has a Poisson distribution with average 6 typing errors.

(1) What is the probability that in a given page:

(i) The number of typing errors will be 7?

(ii) The number of typing errors will at least 2?

(2) What is the probability that in 2 pages there will be 10 typing errors?

(3) What is the probability that in a half page there will be no typing errors?

Solution:

(1) X = number of typing errors per page.

X ~ Poisson (6) (t=1, λ=6, μ=λt=6)

$$f(x) = P(X = x) = p(x;6) = \frac{e^{-6}6^x}{x!}\ ; \quad x = 0, 1, 2, \ldots$$

(i) $$f(7) = P(X = 7) = p(7;6) = \frac{e^{-6}6^7}{7!} = 0.13768$$

(ii) P(X≥2) = P(X=2)+ P(X=3)+ . . . = $\sum_{x=2}^{\infty} P(X = x)$

P(X≥2) = 1– P(X<2) = 1 – [P(X=0)+ P(X=1)]

$$=1 - [f(0) + f(1)] = 1 - [\frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!}]$$

= 1 – [0.00248+0.01487]

= 1 – 0.01735 = 0.982650

(2) X = number of typing errors in 2 pages

X ~ Poisson(12) (t=2, λ=6, μ=λt=12)

$$f(x) = P(X = x) = p(x;12) = \frac{e^{-12}12^x}{x!} : \quad x = 0\ , 1\ , 2\ldots$$

$$f(10) = P(X = 10) = \frac{e^{-12}12^{10}}{10} = 0.1048$$

(3) X = number of typing errors in a half page.

X ~ Poisson (3) (t=1/2, λ=6, $\mu=\lambda t$=6/2=3)

$$f(x) = P(X = x) = p(x;3) = \frac{e^{-3}\,3^x}{x!} : \quad x = 0, 1, 2\ldots$$

$$P(X = 0) = \frac{e^{-3}(3)^0}{0!} = 0.0497871$$

Theorem 5.6: (Poisson approximation for binomial distribution:

Let X be a binomial random variable with probability distribution b(x;n,p). If $n\to\infty$, $p\to 0$, and $\mu=np$ remains constant, then the binomial distribution b(x;n,p) can approximated by Poisson distribution p(x;μ).

- For large n and small p we have:

b(x;n,p) ≈ Poisson(μ) ($\mu=np$)

$$\binom{n}{x} p^x (1-p)^{n-x} \approx \frac{e^{-\mu}\mu^x}{x!}; x = 0,1,\cdots,n; \quad (\mu = np)$$

Example 5.18: Reading Assignment

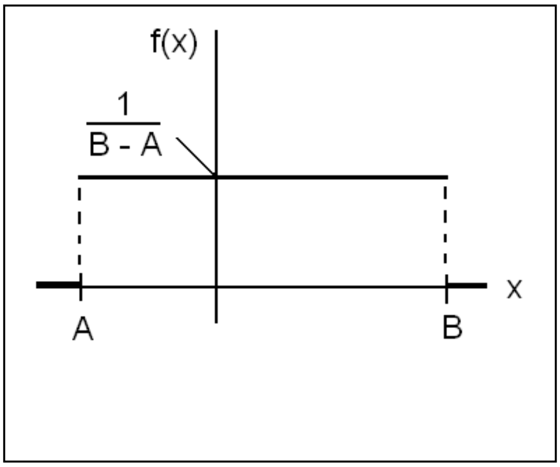
Chapter 6: Some Continuous Probability Distributions:

6.1 Continuous Uniform distribution:

(Rectangular Distribution)

The probability density function of the continuous uniform random variable X on the interval [*A*, *B*] is given by:

$$f(x) = f(x; A, B) = \begin{cases} \dfrac{1}{B-A} \; ; \; A \le x \le B \\ 0 \; ; \; elsewhere \end{cases}$$



We write X~Uniform(*A*,*B*).

Theorem 6.1:

The mean and the variance of the continuous uniform distribution on the interval [*A*, *B*] are:

$$\mu = \frac{A+B}{2}$$

$$\sigma^2 = \frac{(B-A)^2}{12}$$

Example 6.1:

Suppose that, for a certain company, the conference time, X, has a uniform distribution on the interval [0,4] (hours).

(a) What is the probability density function of X?

(b) What is the probability that any conference lasts at least 3 hours?

Solution:

(a) $f(x) = f(x;0,4) = \begin{cases} \dfrac{1}{4} \; ; \; 0 \le x \le 4 \\ 0 \; ; \; elsewhere \end{cases}$

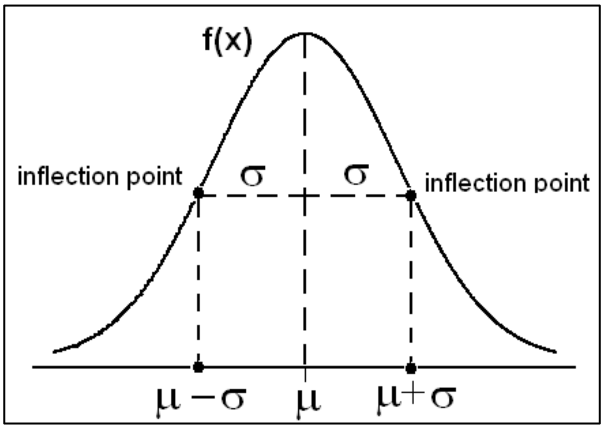
(b) $P(X \ge 3) = \int_3^4 f(x)\,dx = \int_3^4 \frac{1}{4}\,dx = \frac{1}{4}$

6.2 Normal Distribution:

■ The normal distribution is one of the most important continuous distributions.

■ Many measurable characteristics are normally or approximately normally distributed, such as, height and weight.

■ The graph of the probability density function (pdf) of a normal distribution, called the normal curve, is a bell-shaped curve.



■ The pdf of the normal distribution depends on two parameters: mean = E(X)= μ and variance =Var(X) = σ^2.

■ If the random variable X has a normal distribution with mean μ and variance σ^2, we write:

$$X \sim \text{Normal}(\mu,\sigma) \quad \text{or} \quad X \sim N(\mu,\sigma)$$

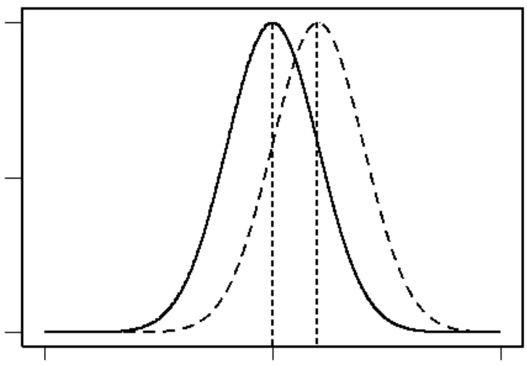
■ The pdf of $X \sim \text{Normal}(\mu,\sigma)$ is given by:

$$f(x) = n(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad ; \begin{cases} -\infty < x < \infty \\ -\infty < \mu < \infty \\ \sigma > 0 \end{cases}$$

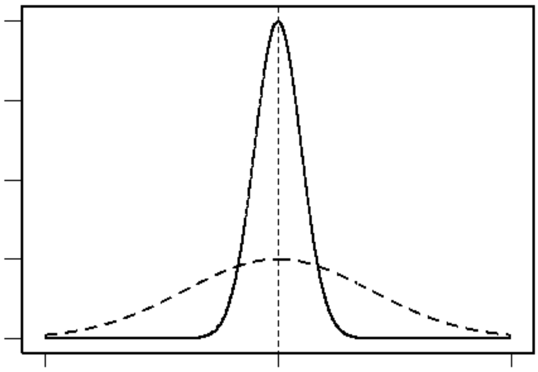
■ The location of the normal distribution depends on μ and its shape depends on σ.

Suppose we have two normal distributions:

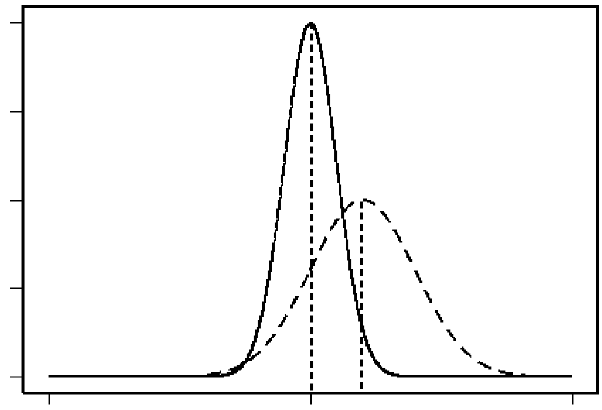
________ $N(\mu_1, \sigma_1)$
----------- $N(\mu_2, \sigma_2)$



$\mu_1 < \mu_2$, $\sigma_1 = \sigma_2$



$\mu_1 = \mu_2$, $\sigma_1 < \sigma_2$



$\mu_1 < \mu_2$, $\sigma_1 < \sigma_2$

■ Some properties of the normal curve $f(x)$ of N(μ,σ):

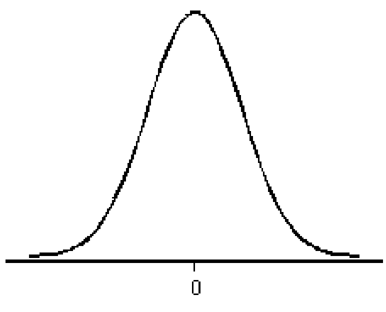
1. $f(x)$ is symmetric about the mean μ.
2. $f(x)$ has two points of inflection at x= $\mu \pm \sigma$.
3. The total area under the curve of $f(x)$ =1.
4. The highest point of the curve of $f(x)$ at the mean μ.

6.3 Areas Under the Normal Curve:

Definition 6.1:
The Standard Normal Distribution:

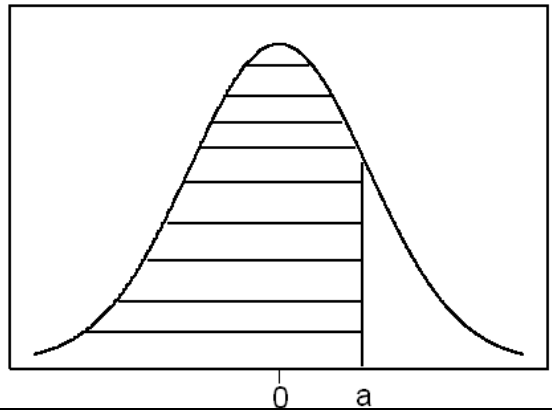
- The normal distribution with mean μ=0 and variance σ^2=1 is called the standard normal distribution and is denoted by Normal(0,1) or N(0,1). If the random variable Z has the standard normal distribution, we write Z~Normal(0,1) or Z~N(0,1).
- The pdf of Z~N(0,1) is given by:

$$f(z) = n(z;0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$



- The standard normal distribution, Z~N(0,1), is very important because probabilities of any normal distribution can be calculated from the probabilities of the standard normal distribution.

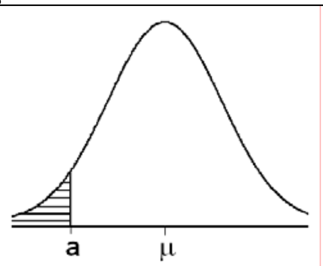
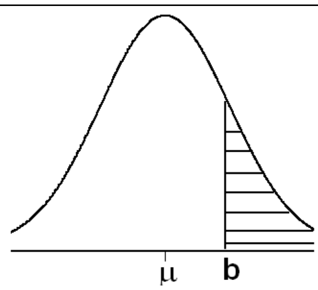
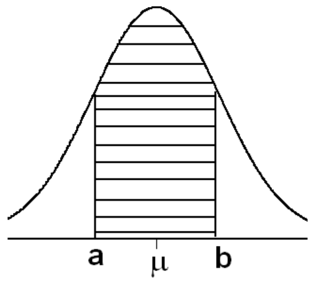
- Probabilities of the standard normal distribution Z~N(0,1) of the form P(Z≤a) are tabulated (Table A.3, p681).

	$P(Z\le a) = \int_{-\infty}^{a} f(z)\,dz$ $= \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\,dz$ = from the table

- We can transfer any normal distribution X~N(μ,σ) to the standard normal distribution, Z~N(0,1) by using the following result.
- Result: If X~N(μ,σ), then $Z = \frac{X-\mu}{\sigma}$~N(0,1).

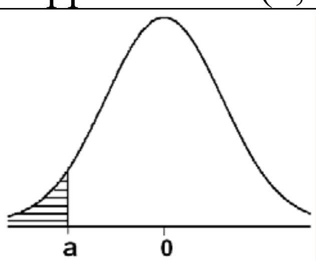
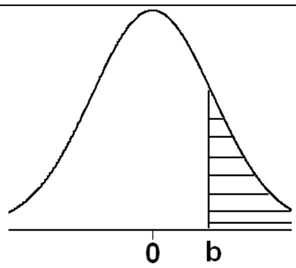
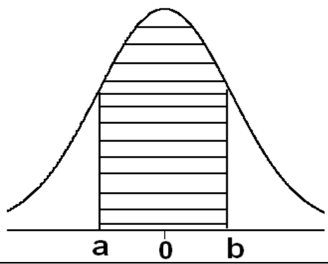
Areas Under the Normal Curve of X~N(μ,σ)

The probabilities of the normal distribution N(μ,σ) depends on μ and σ.

		
$P(X<a)= \int_{-\infty}^{a} f(x)\,dx$	$P(X>b)= \int_{b}^{\infty} f(x)\,dx$	$P(a<X<b)= \int_{a}^{b} f(x)\,dx$

Probabilities of Z~N(0,1):

Suppose Z ~ N(0,1).

		
P(Z≤a) =From Table (A.3)	P(Z≥b) = 1−P(Z≤b)	P(a≤Z≤b) = P(Z≤b)−P(Z≤a)

Note: P(Z=a)=0 for every a .

Example:
Suppose Z~N(0,1).

(1)
P(Z≤1.50)=0.9332

Z	0.00	0.01	…
:	⇓		
1.5 ⇒	0.9332		
:			

(2)
P(Z≥0.98)=1−P(Z≤0.98)
=1 – 0.8365
= 0.1635

Z	0.00	…	0.08
:	:	:	⇓
:	…	…	⇓
0.9⇒	⇒	⇒	0.8365

(3)
P(−1.33≤ Z≤2.42)
= P(Z≤2.42)
− P(Z≤−1.33)
= 0.9922 – 0.0918
= 0.9004

Z	…	0.02	0.03
:	:	⇓	⇓
−1.3	⇒		0.0918
:		⇓	
2.4	⇒	0.9922	

(4) P(Z≤0)=P(Z ≥0)=0.5

Example:
Suppose Z~N(0,1). Find the value of k such that P(Z≤k)= 0.0207.

Solution:
.k = −2.04

Z	…	0.04	
:	:	⇑ ⇑	
−2.0	⇐⇐	0.0207	
:			

Example 6.2: Reading assignment
Example 6.3: Reading assignment

Probabilities of X~N(μ,σ):

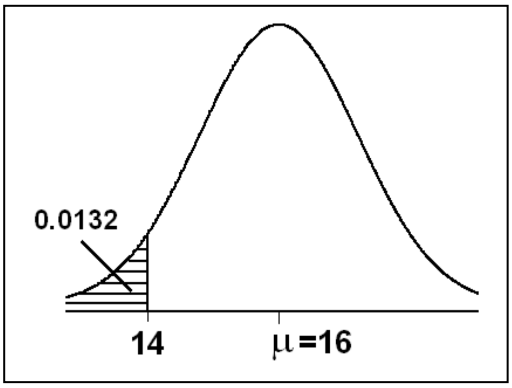
- Result: X ~N(μ,σ) ⇔ $Z = \frac{X - \mu}{\sigma}$ ~N(0,1)

- $X \le a \iff \dfrac{X-\mu}{\sigma} \le \dfrac{a-\mu}{\sigma} \iff Z \le \dfrac{a-\mu}{\sigma}$

1. $P(X \le a) = P\left(Z \le \dfrac{a-\mu}{\sigma}\right)$
2. $P(X \ge a) = 1 - P(X \le a) = 1 - P\left(Z \le \dfrac{a-\mu}{\sigma}\right)$
3. $P(a \le X \le b) = P(X \le b) - P(X \le a) = P\left(Z \le \dfrac{b-\mu}{\sigma}\right) - P\left(Z \le \dfrac{a-\mu}{\sigma}\right)$
4. P(X=a)=0 for every a.
5. $P(X \le \mu) = P(X \ge \mu) = 0.5$

Example 6.4: Reading assignment
Example 6.5: Reading assignment
Example 6.6: Reading assignment

Example:
Suppose that the hemoglobin level for healthy adults males has a normal distribution with mean $\mu=16$ and variance $\sigma^2=0.81$ (standard deviation $\sigma=0.9$).
(a) Find the probability that a randomly chosen healthy adult male has hemoglobin level less than 14.
(b) What is the percentage of healthy adult males who have hemoglobin level less than 14?
Solution:
Let X = the hemoglobin level for a healthy adult male
$X \sim N(\mu,\sigma) = N(16, 0.9)$.
(a) $P(X \le 14) = P\left(Z \le \dfrac{14-\mu}{\sigma}\right) = P\left(Z \le \dfrac{14-16}{0.9}\right)$
$= P(Z \le -2.22) = 0.0132$
(b) The percentage of healthy adult males who have hemoglobin level less than 14 is
$P(X \le 14) \times 100\% = 0.01320 \times 100\%$
$= 1.32\%$
Therefore, 1.32% of healthy adult males have hemoglobin level less than 14.



Unrevised First Draft

Example:

Suppose that the birth weight of Saudi babies has a normal distribution with mean μ=3.4 and standard deviation σ=0.35.

(a) Find the probability that a randomly chosen Saudi baby has a birth weight between 3.0 and 4.0 kg.

(b) What is the percentage of Saudi babies who have a birth weight between 3.0 and 4.0 kg?

Solution:

X = birth weight of a Saudi baby

$\mu = 3.4 \quad \sigma = 0.35 \ (\sigma^2 = 0.35^2 = 0.1225)$

$X \sim N(3.4, 0.35)$

(a) $P(3.0<X<4.0)=P(X<4.0)-P(X<3.0)$

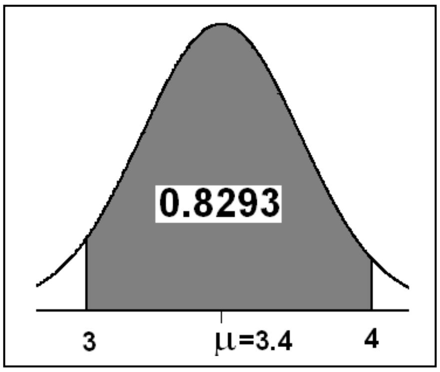
$$= P\left(Z \le \frac{4.0-\mu}{\sigma}\right) - P\left(Z \le \frac{3.0-\mu}{\sigma}\right)$$

$$= P\left(Z \le \frac{4.0-3.4}{0.35}\right) - P\left(Z \le \frac{3.0-3.4}{0.35}\right)$$

$= P(Z\le 1.71) - P(Z \le -1.14)$

$= 0.9564 - 0.1271$

$= 0.8293$



(b) The percentage of Saudi babies who have a birth weight between 3.0 and 4.0 kg is

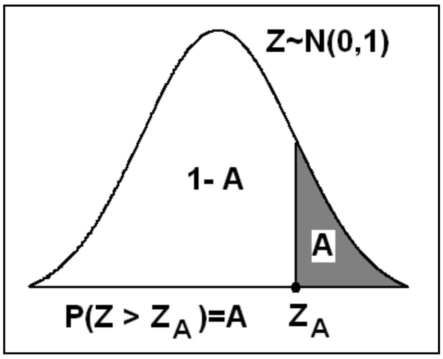
$P(3.0<X<4.0) \times 100\% = 0.8293 \times 100\% = 82.93\%$

Notation:

$P(Z \ge Z_A) = A$

Result:

$Z_A = -Z_{1-A}$



Example:

$Z \sim N(0,1)$

$P(Z \ge Z_{0.025}) = 0.025$

$P(Z \ge Z_{0.95}) = 0.95$

$P(Z \ge Z_{0.90}) = 0.90$

Example:

$Z \sim N(0,1)$

$Z_{0.025} = 1.96$

$Z_{0.95} = -1.645$

$Z_{0.90} = -1.285$

Z	…	0.06	
:	:	⇑ ⇑	
1.9	⇐⇐	0.975	

$P(Z \ge Z_{0.025}) = 0.025$

$Z_{0.025} = 1.96$

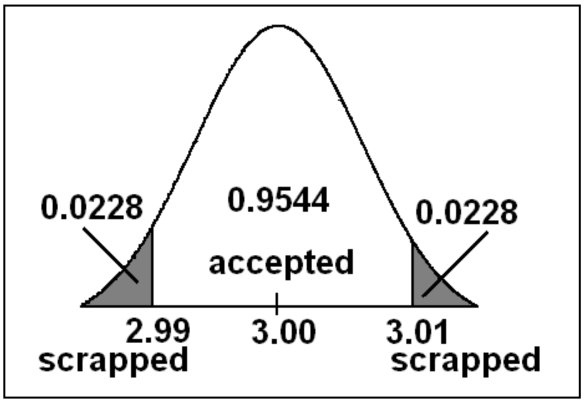
Unrevised First Draft

6.4 Application of the Normal Distribution:

Example 6.7: Reading assignment
Example 6.8: Reading assignment

Example 6.9:
In an industrial process, the diameter of a ball bearing is an important component part. The buyer sets specifications on the diameter to be 3.00±0.01 cm. The implication is that no part falling outside these specifications will be accepted. It is known that, in the process, the diameter of a ball bearing has a normal distribution with mean 3.00 cm and standard deviation 0.005 cm. On the average, how many manufactured ball bearings will be scrapped?

Solution:



$\mu=3.00$
$\sigma=0.005$
X=diameter
X~N(3.00, 0.005)
The specification limits are:
3.00±0.01
x_1=Lower limit=3.00–0.01=2.99
x_2=Upper limit=3.00+0.01=3.01

$P(x_1<X< x_2)=P(2.99<X<3.01)=P(X<3.01)-P(X<2.99)$

$$= P\left(Z \le \frac{3.01-\mu}{\sigma}\right) - P\left(Z \le \frac{2.99-\mu}{\sigma}\right)$$

$$= P\left(Z \le \frac{3.01-3.00}{0.005}\right) - P\left(Z \le \frac{2.99-3.00}{0.005}\right)$$

$$= P(Z\le 2.00) - P(Z \le -2.00)$$

$$= 0.9772 - 0.0228$$

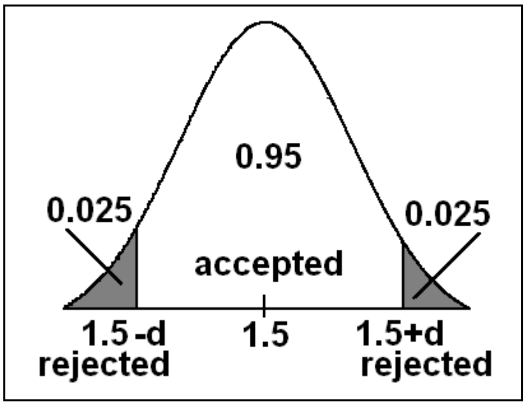
$$= 0.9544$$

Therefore, on the average, 95.44% of manufactured ball bearings will be accepted and 4.56% will be scrapped?

Example 6.10:
Gauges are use to reject all components where a certain dimension is not within the specifications 1.50±*d*. It is known

that this measurement is normally distributed with mean 1.50 and standard deviation 0.20. Determine the value *d* such that the specifications cover 95% of the measurements.

Solution:



$\mu=1.5$

$\sigma=0.20$

X= measurement

$X \sim N(1.5,\ 0.20)$

The specification limits are:

$1.5 \pm d$

x_1=Lower limit=$1.5-d$

x_2=Upper limit=$1.5+d$

$P(X > 1.5+d) = 0.025 \Leftrightarrow P(X < 1.5+d) = 0.975$

$P(X < 1.5-d) = 0.025$

$P(X < 1.5-d) = 0.025$

$$\Leftrightarrow P\left(\frac{X-\mu}{\sigma} \le \frac{(1.5-d)-\mu}{\sigma}\right)=0.025$$

$$\Leftrightarrow P\left(Z \le \frac{(1.5-d)-\mu}{\sigma}\right)=0.025$$

$$\Leftrightarrow P\left(Z \le \frac{(1.5-d)-1.5}{0.20}\right)=0.025$$

$$\Leftrightarrow P\left(Z \le \frac{-d}{0.20}\right)=0.025$$

$$\Leftrightarrow \frac{-d}{0.20} = -1.96$$

$$\Leftrightarrow -d = (0.20)(-1.96)$$

$$\Leftrightarrow d = 0.392$$

Z	…	0.06	
:	:	⇑ ⇑	
−1.9	⇐⇐	0.025	

$P(Z \le \frac{-d}{0.20}) = 0.025$

$\frac{-d}{0.20} = -1.96$

Note: $\frac{-d}{0.20} = Z_{0.025}$

The specification limits are:

x_1=Lower limit=$1.5-d = 1.5 - 0.392 = 1.108$

x_2=Upper limit=$1.5+d=1.5+0.392= 1.892$

Therefore, 95% of the measurements fall within the specifications (1.108, 1.892).

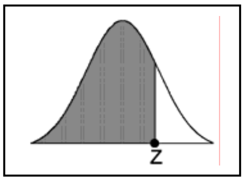
Example 6.11: Reading assignment

Example 6.12: Reading assignment

Example 6.13: Reading assignment

Example 6.14: Reading assignment

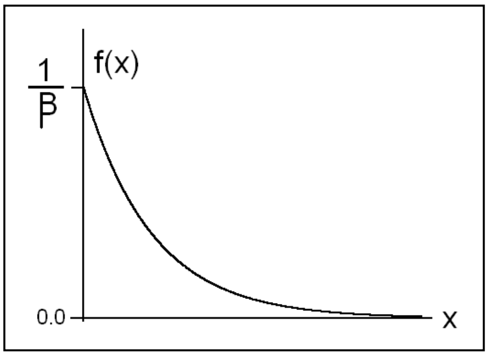
TABLE:
Areas Under The Standard Normal Curve
Z~Normal(0,1)



Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

6.6 Exponential Distribution:

Definition:
The continuous random variable X has an exponential distribution with parameter β, if its probability density function is given by:



$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} & ; x > 0 \\ 0 & ; elsewhere \end{cases}$$

and we write X~Exp(β)

Theorem:
If the random variable X has an exponential distribution with parameter β, i.e., X~Exp(β), then the mean and the variance of X are:

$$E(X) = \mu = \beta$$
$$Var(X) = \sigma^2 = \beta^2$$

Example 6.17:
Suppose that a system contains a certain type of component whose time in years to failure is given by *T*. The random variable *T* is modeled nicely by the exponential distribution with mean time to failure β=5. If 5 of these components are installed in different systems, what is the probability that at least 2 are still functioning at the end of 8 years?

Solution:
β=5
T~Exp(5)
The pdf of *T* is:

$$f(t) = \begin{cases} \frac{1}{5} e^{-t/5} & ; t > 0 \\ 0 & ; elsewhere \end{cases}$$

The probability that a given component is still functioning after 8 years is given by:

$$P(T>8) = \int_8^\infty f(t)\, dt = \int_8^\infty \frac{1}{5} e^{-t/5}\, dt = e^{-8/5} = 0.2$$

Now define the random variable:

X= number of components functioning after 8 years out of 5 components

X~ Binomial(5, 0.2) (n=5, p= P(T>8)= 0.2)

$$f(x) = P(X = x) = b(x;5,0.2) = \begin{cases} \binom{5}{x} 0.2^{x}\ 0.8^{5-x}\ ;\ x = 0, 1, \ldots, 5 \\ 0\ ; \qquad\qquad otherwise \end{cases}$$

The probability that at least 2 are still functioning at the end of 8 years is:

P(X≥2)=1 − P(X<2)=1 − [P(X=0)+P(X=1)]

$$= 1 - \left[\binom{5}{0} 0.2^{0}\ 0.8^{5-0} + \binom{5}{1} 0.2^{1}\ 0.8^{5-1}\right]$$

$$= 1 - [\ 0.8^{5} + 5 \times 0.2\ \times 0.8^{4}\ ]$$

$$= 1 - 0.7373$$

$$= 0.2627$$

Chapter 8: Fundamental Sampling Distributions and Data Descriptions:

8.1 Random Sampling:

Definition 8.1:

A population consists of the totality of the observations with which we are concerned. (Population=Probability Distribution)

Definition 8.2:

A sample is a subset of a population.

Note:

- Each observation in a population is a value of a random variable X having some probability distribution f(x).
- To eliminate bias in the sampling procedure, we select a random sample in the sense that the observations are made independently and at random.
- The random sample of size *n* is:

$$X_1, X_2, \ldots, X_n$$

It consists of *n* observations selected independently and randomly from the population.

8.2 Some Important Statistics:

Definition 8.4:

Any function of the random sample $X_1, X_2, \ldots, X_n$ is called a statistic.

Central Tendency in the Sample:

Definition 8.5:

If $X_1, X_2, \ldots, X_n$ represents a random sample of size *n*, then the sample mean is defined to be the statistic:

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n} \qquad \text{(unit)}$$

Note:

- $\overline{X}$ is a statistic because it is a function of the random sample $X_1, X_2, \ldots, X_n$.
- $\overline{X}$ has same unit of $X_1, X_2, \ldots, X_n$.
- $\overline{X}$ measures the central tendency in the sample (location).

Variability in the Sample:

Definition 8.9:

If $X_1, X_2, \ldots, X_n$ represents a random sample of size n, then the sample variance is defined to be the statistic:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1} = \frac{(X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + \cdots + (X_n - \overline{X})^2}{n-1} \quad \text{(unit)}^2$$

Theorem 8.1: (Computational Formulas for S^2)

$$S^2 = \frac{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}{n-1} = \frac{n\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2}{n(n-1)}$$

Note:

- S^2 is a statistic because it is a function of the random sample $X_1, X_2, \ldots, X_n$.
- S^2 measures the variability in the sample.

Definition 8.10:

The sample standard deviation is defined to be the statistic:

$$S = \sqrt{S^2} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}} \quad \text{(unit)}$$

Example 8.1: Reading Assignment

Example 8.8: Reading Assignment

Example 8.9: Reading Assignment

8.4 Sampling distribution:

Definition 8.13:

The probability distribution of a statistic is called a sampling distribution.

- Example: If $X_1, X_2, \ldots, X_n$ represents a random sample of size n, then the probability distribution of $\overline{X}$ is called the sampling distribution of the sample mean $\overline{X}$.

8.5 Sampling Distributions of Means:

Result:

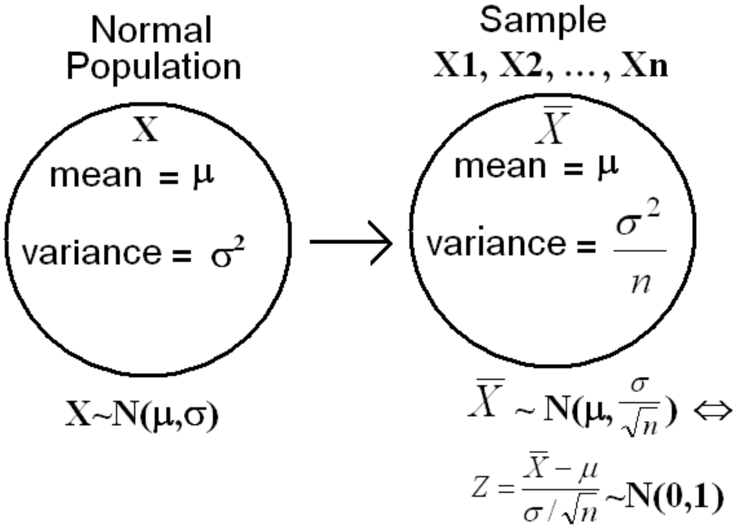
If $X_1, X_2, \ldots, X_n$ is a random sample of size n taken from a normal distribution with mean μ and variance σ^2, i.e. $N(\mu,\sigma)$, then the sample mean $\overline{X}$ has a normal distribution with mean

$$E(\overline{X}) = \mu_{\overline{X}} = \mu$$

and variance

$$Var(\overline{X}) = \sigma^2_{\overline{X}} = \frac{\sigma^2}{n}$$

- If X_1, X_2, …, X_n is a random sample of size n from N(μ,σ), then $\overline{X}$ ~N($\mu_{\overline{X}}, \sigma_{\overline{X}}$) or $\overline{X}$ ~N(μ, $\frac{\sigma}{\sqrt{n}}$).
- $\overline{X}$ ~ N(μ, $\frac{\sigma}{\sqrt{n}}$) $\Leftrightarrow$ $Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ ~N(0,1)



Theorem 8.2: (Central Limit Theorem)

If X_1, X_2, …, X_n is a random sample of size n from any distribution (population) with mean μ and finite variance σ^2, then, if the sample size n is large, the random variable

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

is approximately standard normal random variable, i.e., $Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ ~N(0,1) approximately.

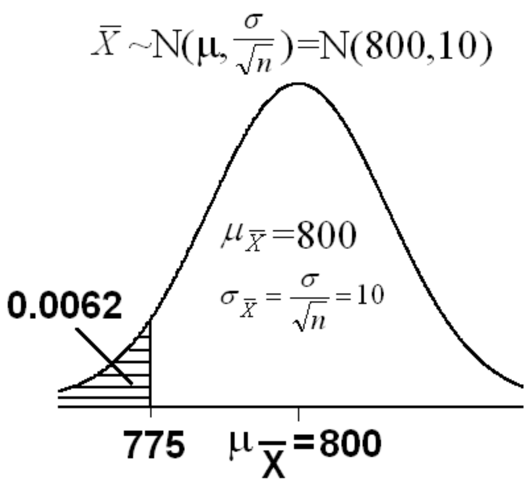
- $Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ ~N(0,1) $\Leftrightarrow$ $\overline{X}$ ~ N(μ, $\frac{\sigma}{\sqrt{n}}$)
- We consider n large when $n \geq 30$.

- For large sample size n, $\overline{X}$ has approximately a normal distribution with mean μ and variance $\frac{\sigma^2}{n}$, i.e., $\overline{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$ approximately.
- The sampling distribution of $\overline{X}$ is used for inferences about the population mean μ.

Example 8.13:

An electric firm manufactures light bulbs that have a length of life that is approximately normally distributed with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a random sample of 16 bulbs will have an average life of less than 775 hours.

Solution:



X= the length of life

μ=800 , σ=40

X~N(800, 40)

n=16

$\mu_{\overline{X}} = \mu = 800$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{40}{\sqrt{16}} = 10$$

$$\overline{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}}) = N(800,10)$$

$$\Leftrightarrow \quad Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = Z = \frac{\overline{X} - 800}{10} \sim N(0,1)$$

$$P(\overline{X} < 775) = P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < \frac{775 - \mu}{\sigma/\sqrt{n}}\right)$$

$$= P\left(\frac{\overline{X} - 800}{10} < \frac{775 - 800}{10}\right)$$

$$= P\left(Z < \frac{775 - 800}{10}\right)$$

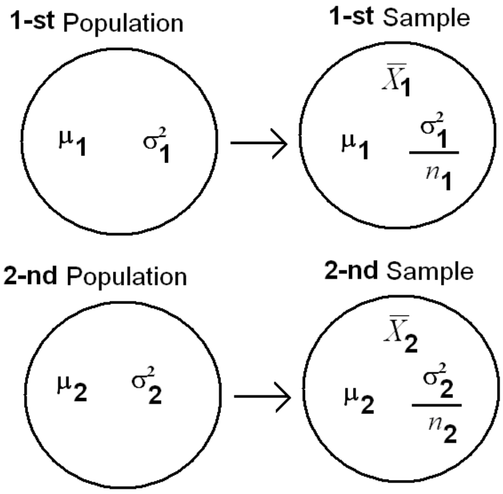
$$= P(Z < -2.50)$$

$$= 0.0062$$

Sampling Distribution of the Difference between Two Means:

Suppose that we have two populations:

- 1-st population with mean μ_1 and variance σ_1^2
- 2-nd population with mean μ_2 and variance σ_2^2
- We are interested in comparing μ_1 and μ_2, or equivalently, making inferences about $\mu_1 - \mu_2$.
- We independently select a random sample of size n_1 from the 1-st population and another random sample of size n_2 from the 2-nd population:
- Let $\bar{X}_1$ be the sample mean of the 1-st sample.
- Let $\bar{X}_2$ be the sample mean of the 2-nd sample.
- The sampling distribution of $\bar{X}_1 - \bar{X}_2$ is used to make inferences about $\mu_1 - \mu_2$.



Theorem 8.3:

If n_1 and n_2 are large, then the sampling distribution of $\bar{X}_1 - \bar{X}_2$ is approximately normal with mean

$$E(\bar{X}_1 - \bar{X}_2) = \mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$$

and variance

$$Var(\bar{X}_1 - \bar{X}_2) = \sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

that is:

$$\bar{X}_1 - \bar{X}_2 \sim \text{N}(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

$$\Leftrightarrow$$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim \text{N}(0,1)$$

Note:

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\sigma^2_{\bar{X}_1 - \bar{X}_2}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \neq \sqrt{\frac{\sigma_1^2}{n_1}} + \sqrt{\frac{\sigma_2^2}{n_2}} = \frac{\sigma_1}{\sqrt{n_1}} + \frac{\sigma_2}{\sqrt{n_2}}$$

Example 8.15: Reading Assignment

Example 8.16:
The television picture tubes of manufacturer *A* have a mean lifetime of 6.5 years and standard deviation of 0.9 year, while those of manufacturer *B* have a mean lifetime of 6 years and standard deviation of 0.8 year. What is the probability that a random sample of 36 tubes from manufacturer *A* will have a mean lifetime that is at least 1 year more than the mean lifetime of a random sample of 49 tubes from manufacturer *B*?

Solution:

Population A	Population B
μ_1=6.5	μ_2=6.0
σ_1=0.9	σ_2=0.8
n_1=36	n_2=49

- We need to find the probability that the mean lifetime of manufacturer *A* is at least 1 year more than the mean lifetime of manufacturer *B* which is P($\bar{X}_1 \geq \bar{X}_2 + 1$).
- The sampling distribution of $\bar{X}_1 - \bar{X}_2$ is

$$\bar{X}_1 - \bar{X}_2 \sim \text{N}(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

- $E(\bar{X}_1 - \bar{X}_2) = \mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$=6.5– 6.0=0.5
- $Var(\bar{X}_1 - \bar{X}_2) = \sigma^2_{\bar{X}_1 - \bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{(0.9)^2}{36} + \frac{(0.8)^2}{49}$=0.03556

- $\sigma_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} = \sqrt{0.03556}$ = 0.189
- $\overline{X}_1-\overline{X}_2$~N(0.5,0.189)
- Recall $Z = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$~N(0,1)

P($\overline{X}_1 \geq \overline{X}_2 + 1$) = P($\overline{X}_1 - \overline{X}_2 \geq 1$)

$$= P\left(\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \geq \frac{1-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\right)$$

$$= P\left(Z \geq \frac{1-0.5}{0.189}\right)$$

= P(Z≥2.65)
= 1 – P(Z<2.65)
= 1 – 0.9960
= 0.0040

8.7 t-Distribution:

- Recall that, if X_1, X_2, …, X_n is a random sample of size *n* from a normal distribution with mean μ and variance σ^2, i.e. N(μ,σ), then

$$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

- We can apply this result only when σ^2 is known!
- If σ^2 is unknown, we replace the population variance σ^2 with the sample variance $S^2 = \frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n-1}$ to have the following statistic

$$T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$$

Result:

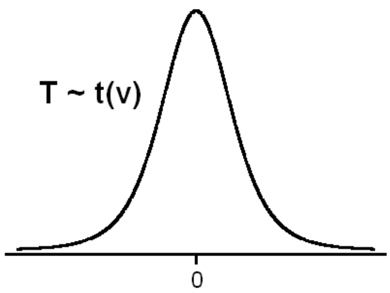
If X_1, X_2, …, X_n is a random sample of size *n* from a normal distribution with mean μ and variance σ^2, i.e. $N(\mu,\sigma)$, then the statistic

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

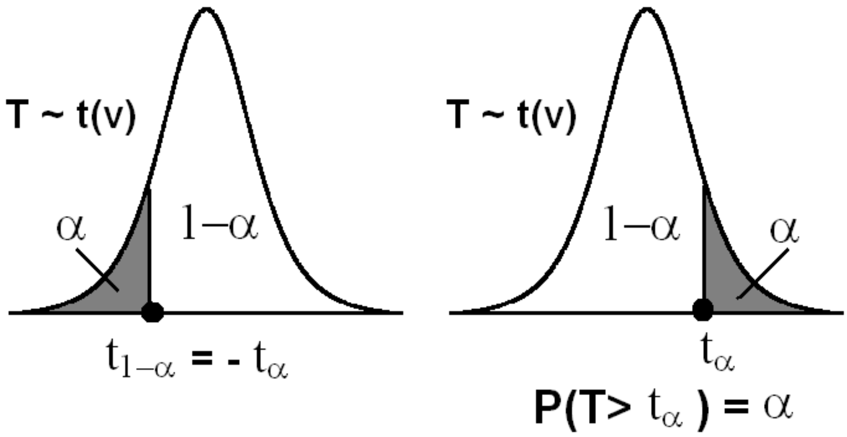
has a t-distribution with $\nu = n-1$degrees of freedom (df), and we write $T \sim t(\nu)$.

Note:

- t-distribution is a continuous distribution.
- The shape of t-distribution is similar to the shape of the standard normal distribution.



Notation:



- t_α = The t-value above which we find an area equal to α, that is $P(T > t_\alpha) = \alpha$
- Since the curve of the pdf of $T \sim t(\nu)$ is symmetric about 0, we have

$$t_{1-\alpha} = -t_\alpha$$

- Values of t_α are tabulated in Table A-4 (p.683).

Example:

Find the t-value with ν=14 (df) that leaves an area of:

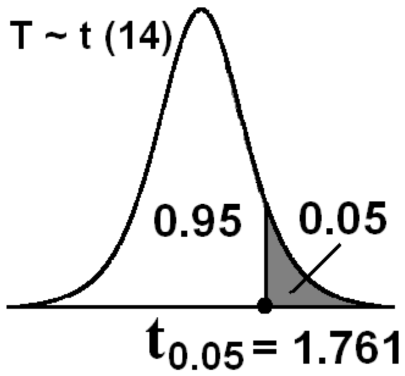
(a) 0.95 to the left.

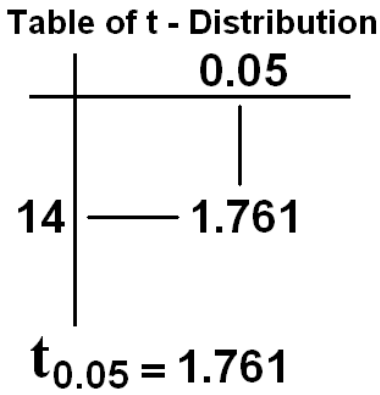
(b) 0.95 to the right.

Solution:

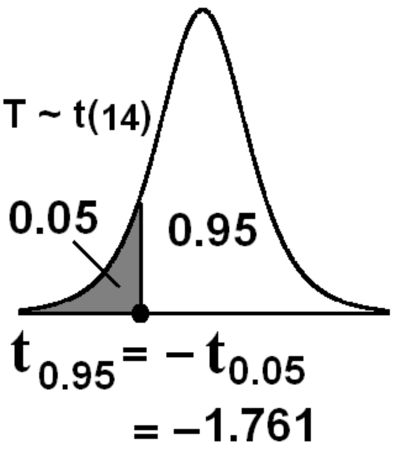
ν = 14 (df); $T \sim t(14)$

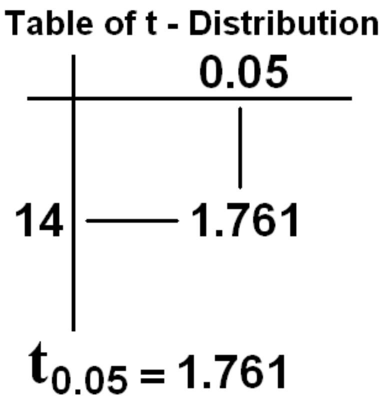
(a) The t-value that leaves an area of 0.95 to the left is
$t_{0.05} = 1.761$



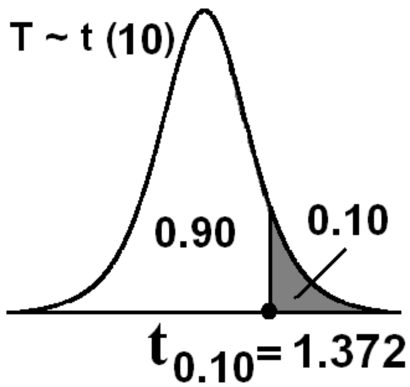


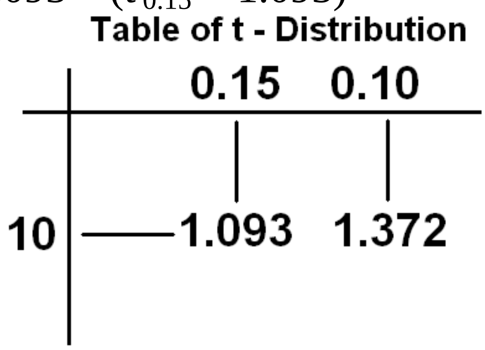
(b) The t-value that leaves an area of 0.95 to the right is
$t_{0.95} = -t_{1-0.95} = -t_{0.05} = -1.761$





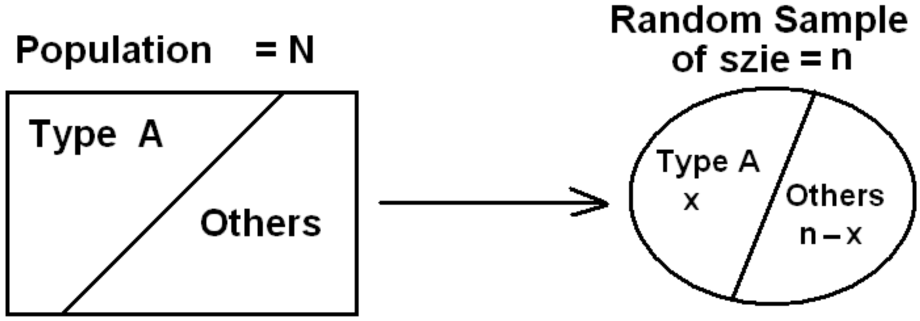
Example:
For $\nu = 10$ degrees of freedom (df), find $t_{0.10}$ and $t_{0.85}$.
Solution:
$t_{0.10} = 1.372$
$t_{0.85} = -t_{1-0.85} = -t_{0.15} = -1.093$ ($t_{0.15} = 1.093$)





Sampling Distribution of the Sample Proportion:

Suppose that the size of a population is N. Each element of the population can be classified as type A or non-type A. Let p be the proportion of elements of type A in the population. A random sample of size n is drawn from this population. Let $\hat{p}$ be the proportion of elements of type A in the sample.



Let X = no. of elements of type A in the sample

p =Population Proportion

$$= \frac{no.\ of\ \ elements\ of\ \ type\ \ A\ in\ the\ \ population}{N}$$

$\hat{p}$ = Sample Proportion

$$= \frac{no.\ of\ \ elements\ of\ \ type\ \ A\ in\ the\ sample}{n} = \frac{X}{n}$$

Result:

(1) X ~ Binomial (n, p)

(2) $E(\hat{p}) = E(\frac{X}{n}) = p$

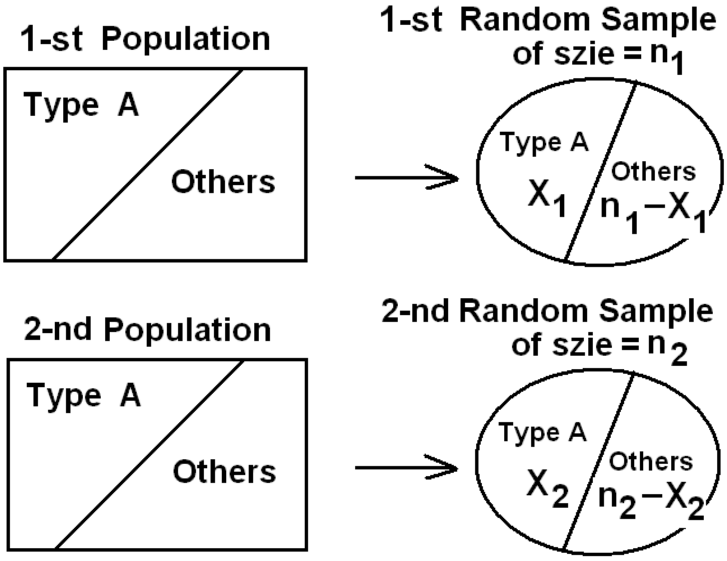
(3) $Var(\hat{p}) = Var(\frac{X}{n}) = \frac{pq}{n}$; $q = 1 - p$

(4) For large n, we have

$$\hat{p} \sim N(p, \sqrt{\frac{pq}{n}}) \quad \text{(Approximately)}$$

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \sim N(0,1) \quad \text{(Approximately)}$$

<u>Sampling Distribution of the Difference between Two Proportions:</u>



Suppose that we have two populations:

- p_1 = proportion of the 1-st population.
- P_2 = proportion of the 2-nd population.
- We are interested in comparing p_1 and p_2, or equivalently, making inferences about $p_1 - p_2$.
- We <u>independently</u> select a random sample of size n_1 from the 1-st population and another random sample of size n_2 from the 2-nd population:
- Let X_1 = no. of elements of type A in the 1-st sample.
- Let X_2 = no. of elements of type A in the 2-nd sample.
- $\hat{p}_1 = \frac{X_1}{n_1}$ = proportion of the 1-st sample
- $\hat{p}_2 = \frac{X_2}{n_2}$ = proportion of the 2-nd sample
- The sampling distribution of $\hat{p}_1 - \hat{p}_2$ is used to make inferences about $p_1 - p_2$.

Result:

(1) $E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$

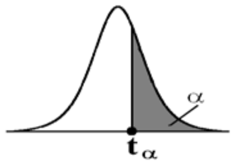
(2) $Var(\hat{p}_1 - \hat{p}_2) = \frac{p_1\, q_1}{n_1} + \frac{p_2\, q_2}{n_2} \quad ; q_1 = 1 - p_1,\ q_2 = 1 - p_2$

(3) For large n_1 and n_2, we have

$$\hat{p}_1 - \hat{p}_2 \sim N(\,p_1 - p_2\,, \sqrt{\frac{p_1\,q_1}{n_1} + \frac{p_2\,q_2}{n_2}}\,) \quad \text{(Approximately)}$$

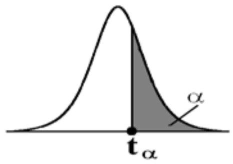
$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1\,q_1}{n_1} + \frac{p_2\,q_2}{n_2}}} \sim \mathrm{N}(0,1) \quad \text{(Approximately)}$$

Critical Values of the t-distribution (t_α)



ν	α						
	0.40	**0.30**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179
13	0.259	0.537	0.870	1.079	1.350	1.771	2.160
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960

Critical Values of the t-distribution (t_α)



ν	α						
	0.02	**0.015**	**0.01**	**0.0075**	**0.005**	**0.0025**	**0.0005**
1	15.895	21.205	31.821	42.434	63.657	127.322	636.590
2	4.849	5.643	6.965	8.073	9.925	14.089	31.598
3	3.482	3.896	4.541	5.047	5.841	7.453	12.924
4	2.999	3.298	3.747	4.088	4.604	5.598	8.610
5	2.757	3.003	3.365	3.634	4.032	4.773	6.869
6	2.612	2.829	3.143	3.372	3.707	4.317	5.959
7	2.517	2.715	2.998	3.203	3.499	4.029	5.408
8	2.449	2.634	2.896	3.085	3.355	3.833	5.041
9	2.398	2.574	2.821	2.998	3.250	3.690	4.781
10	2.359	2.527	2.764	2.932	3.169	3.581	4.587
11	2.328	2.491	2.718	2.879	3.106	3.497	4.437
12	2.303	2.461	2.681	2.836	3.055	3.428	4.318
13	2.282	2.436	2.650	2.801	3.012	3.372	4.221
14	2.264	2.415	2.624	2.771	2.977	3.326	4.140
15	2.249	2.397	2.602	2.746	2.947	3.286	4.073
16	2.235	2.382	2.583	2.724	2.921	3.252	4.015
17	2.224	2.368	2.567	2.706	2.898	3.222	3.965
18	2.214	2.356	2.552	2.689	2.878	3.197	3.922
19	2.205	2.346	2.539	2.674	2.861	3.174	3.883
20	2.197	2.336	2.528	2.661	2.845	3.153	3.850
21	2.189	2.328	2.518	2.649	2.831	3.135	3.819
22	2.183	2.320	2.508	2.639	2.819	3.119	3.792
23	2.177	2.313	2.500	2.629	2.807	3.104	3.768
24	2.172	2.307	2.492	2.620	2.797	3.091	3.745
25	2.167	2.301	2.485	2.612	2.787	3.078	3.725
26	2.162	2.296	2.479	2.605	2.779	3.067	3.707
27	2.158	2.291	2.473	2.598	2.771	3.057	3.690
28	2.154	2.286	2.467	2.592	2.763	3.047	3.674
29	2.150	2.282	2.462	2.586	2.756	3.038	3.659
30	2.147	2.278	2.457	2.581	2.750	3.030	3.646
40	2.125	2.250	2.423	2.542	2.704	2.971	3.551
60	2.099	2.223	2.390	2.504	2.660	2.915	3.460
120	2.076	2.196	2.358	2.468	2.617	2.860	3.373
∞	2.054	2.170	2.326	2.432	2.576	2.807	3.291

Chapter 9: One- and Two-Sample Estimation Problems:

9.1 Introduction:

- Suppose we have a population with some unknown parameter(s).
  Example: Normal(μ,σ)
  μ and σ are parameters.
- We need to draw conclusions (make inferences) about the unknown parameters.
- We select samples, compute some statistics, and make inferences about the unknown parameters based on the sampling distributions of the statistics.

* Statistical Inference
(1) Estimation of the parameters (Chapter 9)
→ Point Estimation
→ Interval Estimation (Confidence Interval)
(2) Tests of hypotheses about the parameters (Chapter 10)

9.3 Classical Methods of Estimation:

Point Estimation:
A point estimate of some population parameter θ is a single value $\hat{\theta}$ of a statistic $\hat{\Theta}$. For example, the value $\bar{x}$ of the statistic $\bar{X}$ computed from a sample of size n is a point estimate of the population mean μ.

Interval Estimation (Confidence Interval = C.I.):
An interval estimate of some population parameter θ is an interval of the form $(\hat{\theta}_L, \hat{\theta}_U)$, i.e, $\hat{\theta}_L<\theta<\hat{\theta}_U$. This interval contains the true value of θ "with probability 1−α", that is $P(\hat{\theta}_L<\theta<\hat{\theta}_U)=1-\alpha$.

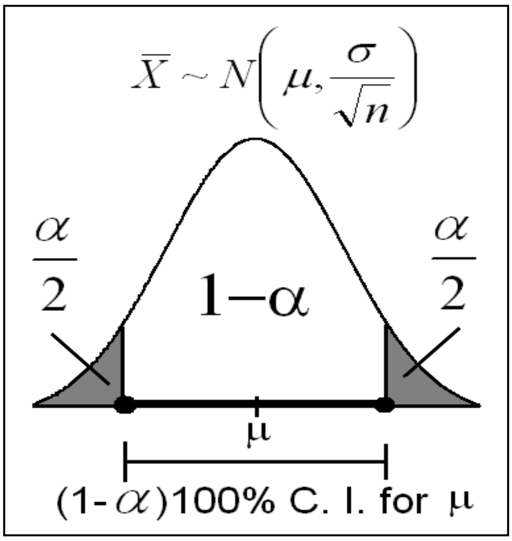
- $(\hat{\theta}_L, \hat{\theta}_U)=\hat{\theta}_L<\theta<\hat{\theta}_U$ is called a (1−α)100% confidence interval (C.I.) for θ.
- 1−α is called the confidence coefficient
- $\hat{\theta}_L$ = lower confidence limit

- $\hat{\theta}_U$ = upper confidence limit
- α=0.1, 0.05, 0.025, 0.01 ($0<\alpha<1$)

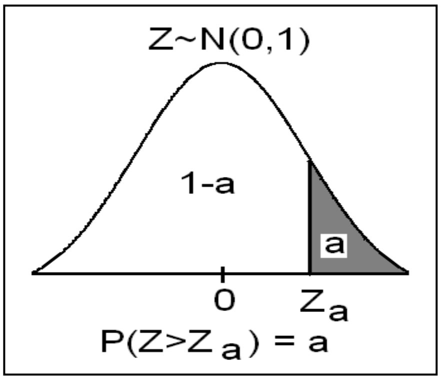
9.4 Single Sample: Estimation of the Mean (μ):

Recall:

- $E(\overline{X}) = \mu_{\overline{X}} = \mu$
- $Var(\overline{X}) = \sigma^2_{\overline{X}} = \dfrac{\sigma^2}{n}$
- $\overline{X} \sim N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$
- $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ ~N(0,1)
  (σ^2 is known)
- $T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}}$ ~ t(n–1)
  (σ^2 is unknown)
- We use the sampling distribution of $\overline{X}$ to make inferences about μ.



Notation:
Z_a is the Z-value leaving an area of a to the right; i.e., $P(Z>Z_a)=a$ or equivalently, $P(Z<Z_a)=1-a$



Point Estimation of the Mean (μ):

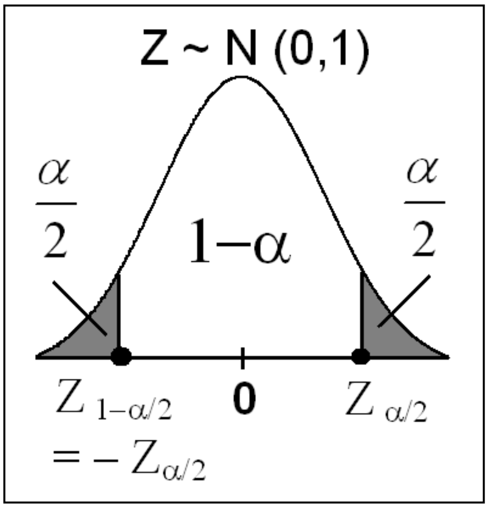
- The sample mean $\overline{X} = \sum_{i=1}^{n} X_i / n$ is a "good" point estimate for μ.

Interval Estimation (Confidence Interval) of the Mean (μ):

(i) First Case: σ^2 is known:

Result:

If $\overline{X} = \sum_{i=1}^{n} X_i / n$ is the sample mean of a random sample of size n from a population (distribution) with mean μ and known variance σ^2, then a $(1-\alpha)100\%$ confidence interval for μ is :



$$(\overline{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ,\ \overline{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$$

$$\Leftrightarrow\ \overline{X} \pm Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$$\Leftrightarrow\ \overline{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

where $Z_{\frac{\alpha}{2}}$ is the Z-value leaving an area of α/2 to the right; i.e., P(Z>$Z_{\frac{\alpha}{2}}$)=α/2, or equivalently, P(Z<$Z_{\frac{\alpha}{2}}$)=1−α/2.

Note:

We are (1−α)100% confident that $\mu \in (\overline{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ,\ \overline{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$

Example 9.2:

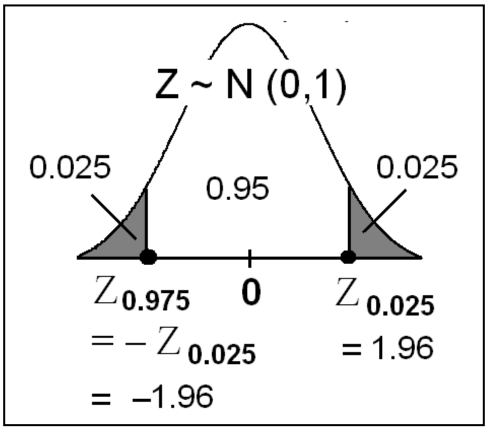
The average zinc concentration recorded from a sample of zinc measurements in 36 different locations is found to be 2.6 gram/milliliter. Find a 95% and 99% confidence interval (C.I.) for the mean zinc concentration in the river. Assume that the population standard deviation is 0.3.

Solution:

μ = the mean zinc concentration in the river.

Population	Sample
μ=??	n=36
σ=0.3	$\overline{X}$ =2.6

First, a point estimate for μ is $\overline{X}$ =2.6.



(a) We want to find 95% C.I. for μ.

α = ??

95% = (1−α)100%

⇔ 0. 95 = (1−α)

⇔ α=0.05

⇔ α/2 = 0.025

$Z_{\frac{\alpha}{2}} = Z_{0.025}$

$= 1.96$

A 95% C.I. for μ is

$$\overline{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$\Leftrightarrow \overline{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$\Leftrightarrow 2.6 - (1.96)\left(\frac{0.3}{\sqrt{36}}\right) < \mu < 2.6 + (1.96)\left(\frac{0.3}{\sqrt{36}}\right)$$

⇔ 2.6 – 0.098 < μ < 2.6 + 0.098

⇔ 2.502 < μ < 2.698

⇔ μ ∈(2.502 , 2.698)

We are 95% confident that μ ∈(2.502 , 2.698).

(b) Similarly, we can find that (Homework) A 99% C.I. for μ is

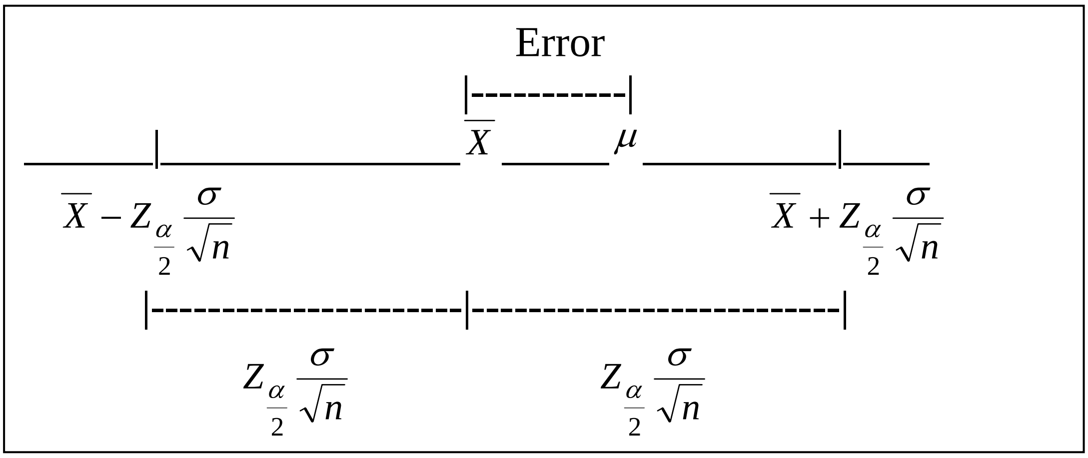
2.471 < μ < 2.729

⇔ μ ∈(2.471 , 2.729)

We are 99% confident that μ ∈(2.471 , 2.729)

Notice that a 99% C.I. is wider that a 95% C.I..

Note:



Theorem 9.1:

If $\overline{X}$ is used as an estimate of μ, we can then be (1–α)100% confident that the error (in estimation) will not exceed $Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$.

Example:

In Example 9.2, we are 95% confident that the sample mean $\overline{X} = 2.6$ differs from the true mean μ by an amount less than $Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = (1.96)\left(\frac{0.3}{\sqrt{36}}\right) = 0.098$.

Note:

Let e be the maximum amount of the error, that is $e = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$, then:

$$e = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \Leftrightarrow \sqrt{n} = Z_{\frac{\alpha}{2}} \frac{\sigma}{e} \Leftrightarrow n = \left(Z_{\frac{\alpha}{2}} \frac{\sigma}{e} \right)^2$$

Theorem 9.2:

If $\overline{X}$ is used as an estimate of μ, we can then be (1−α)100% confident that the error (in estimation) will not exceed a specified amount e when the sample size is $n = \left(Z_{\frac{\alpha}{2}} \frac{\sigma}{e} \right)^2$.

Note:

1. All fractional values of $n = (Z_{\frac{\alpha}{2}} \sigma / e)^2$ are rounded up to the next whole number.
2. If σ is unknown, we could take a preliminary sample of size n≥30 to provide an estimate of σ. Then using $S = \sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2 / (n-1)}$ as an approximation for σ in Theorem 9.2 we could determine approximately how many observations are needed to provide the desired degree of accuracy.

Example 9.3:

How large a sample is required in Example 9.2 if we want to be 95% confident that our estimate of μ is off by less than 0.05?

Solution:

We have σ= 0.3 , $Z_{\frac{\alpha}{2}} = 1.96$, e=0.05. Then by Theorem 9.2,

$$n = \left(Z_{\frac{\alpha}{2}} \frac{\sigma}{e} \right)^2 = \left(1.96 \times \frac{0.3}{0.05} \right)^2 = 138.3 \approx 139$$

Therefore, we can be 95% confident that a random sample of size *n*=139 will provide an estimate $\overline{X}$ differing from μ by an amount less than *e*=0.05.

<u>Interval Estimation (Confidence Interval) of the Mean (μ):</u>

<u>(ii) Second Case: σ^2 is unknown:</u>

Recall:

- $T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$ ~ t(*n*–1)

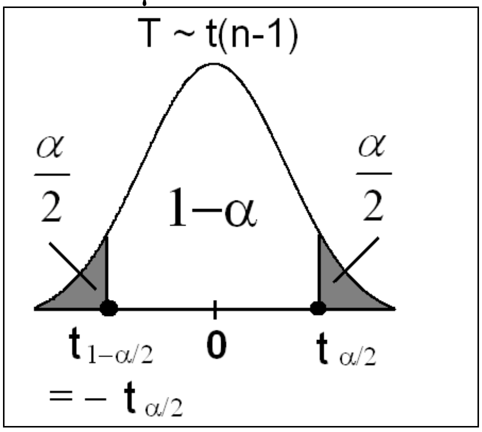
<u>Result:</u>

If $\overline{X} = \sum_{i=1}^{n} X_i / n$ and $S = \sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2 / (n-1)}$ are the sample mean and the sample standard deviation of a random sample of size *n* from <u>a normal</u> population (distribution) with unknown variance σ^2, then a (1–α)100% confidence interval for μ is :

$$(\overline{X} - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \, , \, \overline{X} + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}})$$

$$\Leftrightarrow \overline{X} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

$$\Leftrightarrow \overline{X} - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \overline{X} + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$



where $t_{\frac{\alpha}{2}}$ is the t-value with ν=*n*–1 degrees of freedom leaving an area of α/2 to the right; i.e., P(T>$t_{\frac{\alpha}{2}}$)=α/2, or equivalently, P(T<$t_{\frac{\alpha}{2}}$)=1–α/2.

Example 9.4:

The contents of 7 similar containers of sulfuric acid are 9.8, 10.2, 10.4, 9.8, 10.0, 10.2, and 9.6 liters. Find a 95% C.I. for the mean of all such containers, assuming an approximate normal distribution.

Solution:

.n=7 $\qquad \overline{X} = \sum_{i=1}^{n} X_i / n = 10.0 \qquad S = \sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2 / (n-1)} = 0.283$

First, a point estimate for μ is $\overline{X} = \sum_{i=1}^{n} X_i / n = 10.0$

Now, we need to find a confidence interval for μ.

α = ??

95%=(1–α)100% ⇔ 0. 95=(1–α) ⇔ α=0.05 ⇔ α/2=0.025

$t_{\frac{\alpha}{2}}$ = $t_{0.025}$ =2.447 (with ν=n–1=6 degrees of freedom)

	0.025
6	→ ↓ $t_{0.025}$=2.447

A 95% C.I. for μ is

$$\overline{X} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

$$\Leftrightarrow \overline{X} - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \overline{X} + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$

$$\Leftrightarrow 10.0 - (2.447)\left(\frac{0.283}{\sqrt{7}}\right) < \mu < 10.0 + (2.447)\left(\frac{0.283}{\sqrt{7}}\right)$$

⇔ 10.0 – 0.262< μ < 10.0 + 0.262

⇔ 9.74 < μ < 10.26

⇔ μ ∈(9.74 , 10.26)

We are 95% confident that μ ∈(9.74 , 10.26).

9.5 Standard Error of a Point Estimate:

- The standard error of an estimator is its standard deviation.
- We use $\overline{X} = \sum_{i=1}^{n} X_i / n$ as a point estimator of μ, and we used the sampling distribution of $\overline{X}$ to make a (1–α)100% C.I. for μ.
- The standard deviation of $\overline{X}$, which is $\sigma_{\overline{X}} = \sigma / \sqrt{n}$, is called the standard error of $\overline{X}$. We write $s.e.(\overline{X}) = \sigma / \sqrt{n}$.
- Note: a (1–α)100% C.I. for μ, when σ^2 is known, is

$$\overline{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} = \overline{X} \pm Z_{\frac{\alpha}{2}} s.e(\overline{X}).$$

- Note: a $(1-\alpha)100\%$ C.I. for μ, when σ^2 is unknown and the distribution is normal, is

$$\overline{X} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} = \overline{X} \pm t_{\frac{\alpha}{2}}\, \hat{s}.e(\overline{X})\,. \quad (\nu = n-1 \text{ df})$$

9.7 Two Samples: Estimating the Difference between Two Means ($\mu_1-\mu_2$):

Recall: For two independent samples:

- $\mu_{\overline{X}_1-\overline{X}_2} = \mu_1 - \mu_2 = E(\overline{X}_1 - \overline{X}_2)$
- $\sigma^2_{\overline{X}_1-\overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = Var(\overline{X}_1 - \overline{X}_2)$
- $\sigma_{\overline{X}_1-\overline{X}_2} = \sqrt{\sigma^2_{\overline{X}_1-\overline{X}_2}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
- $Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$

Point Estimation of $\mu_1-\mu_2$:

- $\overline{X}_1 - \overline{X}_2$ is a "good" point estimate for $\mu_1-\mu_2$.

Confidence Interval of $\mu_1-\mu_2$:

(i) First Case: σ_1^2 and σ_2^2 are known:

- $Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$
- Result:

a $(1-\alpha)100\%$ confidence interval for $\mu_1-\mu_2$ is :

$$(\overline{X}_1 - \overline{X}_2) - Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

or $(\overline{X}_1 - \overline{X}_2) \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

or $\left((\overline{X}_1 - \overline{X}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, (\overline{X}_1 - \overline{X}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right)$

(ii) Second Case: $\sigma_1^2 = \sigma_2^2 = \sigma^2$ is unknown:

- If σ_1^2 and σ_2^2 are unknown but $\sigma_1^2 = \sigma_2^2 = \sigma^2$, then the pooled estimate of σ^2 is

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

where S_1^2 is the variance of the 1-st sample and S_2^2 is the variance of the 2-nd sample. The degrees of freedom of S_p^2 is $\nu = n_1 + n_2 - 2$.

- Result:

a $(1-\alpha)100\%$ confidence interval for $\mu_1 - \mu_2$ is :

$$(\overline{X}_1 - \overline{X}_2) - t_{\frac{\alpha}{2}} \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + t_{\frac{\alpha}{2}} \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$$

or

$$(\overline{X}_1 - \overline{X}_2) - t_{\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + t_{\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

or $(\overline{X}_1 - \overline{X}_2) \pm t_{\frac{\alpha}{2}}\ S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

or $\left((\overline{X}_1 - \overline{X}_2) - t_{\frac{\alpha}{2}}\ S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, (\overline{X}_1 - \overline{X}_2) + t_{\frac{\alpha}{2}}\ S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$

where $t_{\frac{\alpha}{2}}$ is the t-value with $\nu = n_1 + n_2 - 2$ degrees of freedom.

Example 9.6: (1^{st} Case: σ_1^2 and σ_2^2 are known)

An experiment was conducted in which two types of engines, *A* and *B*, were compared. Gas mileage in miles per gallon was measured. 50 experiments were conducted using engine type *A* and 75 experiments were done for engine type *B*. The gasoline used and other conditions were held constant. The average gas mileage for engine *A* was 36 miles per gallon and the average for engine *B* was 42 miles per gallon. Find 96% confidence interval for $\mu_B - \mu_A$, where μ_A and μ_B are population mean gas mileage for engines *A* and *B*, respectively. Assume that the population standard deviations are 6 and 8 for engines *A* and *B*, respectively.

Solution:

Engine *A*	Engine *B*
$n_A=50$	$n_B=75$
$\overline{X}_A=36$	$\overline{X}_B=42$
$\sigma_A=6$	$\sigma_B=8$

A point estimate for $\mu_B-\mu_A$ is $\overline{X}_B-\overline{X}_A=42-36=6$.

α = ??

96% = $(1-\alpha)100\% \Leftrightarrow 0.96=(1-\alpha) \Leftrightarrow \alpha=0.04 \Leftrightarrow \alpha/2=0.02$

$Z_{\frac{\alpha}{2}} = Z_{0.02} = 2.05$

A 96% C.I. for $\mu_B-\mu_A$ is

$$(\overline{X}_B-\overline{X}_A)-Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_B^2}{n_B}+\frac{\sigma_A^2}{n_A}}<\mu_B-\mu_A<(\overline{X}_B-\overline{X}_A)+Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_B^2}{n_B}+\frac{\sigma_A^2}{n_A}}$$

$$(\overline{X}_B-\overline{X}_A)\pm Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_B^2}{n_B}+\frac{\sigma_A^2}{n_A}}$$

$$(42-36)\pm Z_{0.02}\sqrt{\frac{8^2}{75}+\frac{6^2}{50}}$$

$$6\pm(2.05)\sqrt{\frac{64}{75}+\frac{36}{50}}$$

$$6\pm 2.571$$

$3.43 < \mu_B-\mu_A < 8.57$

We are 96% confident that $\mu_B-\mu_A \in (3.43, 8.57)$.

Example 9.7: (2nd Case: $\sigma_1^2=\sigma_2^2$ unknown) Reading Assignment

Example: (2nd Case: $\sigma_1^2=\sigma_2^2$ unknown)

To compare the resistance of wire *A* with that of wire *B*, an experiment shows the following results based on two independent samples (original data multiplied by 1000):

Wire *A*: 140, 138, 143, 142, 144, 137

Wire *B*: 135, 140, 136, 142, 138, 140

Assuming equal variances, find 95% confidence interval for $\mu_A-\mu_B$, where μ_A (μ_B) is the mean resistance of wire *A* (*B*).

Solution:

Wire A	Wire B .
n_A=6	n_B=6
$\overline{X}_A$=140.67	$\overline{X}_B$=138.50
S^2_A=7.86690	S^2_B=7.10009

A point estimate for $\mu_A-\mu_B$ is $\overline{X}_A-\overline{X}_B$=140.67–138.50=2.17.

95% = (1–α)100% ⇔ 0. 95 = (1–α) ⇔ α=0.05 ⇔ α/2 = 0.025

ν= df = n_A+n_B-2= 10

$t_{\frac{\alpha}{2}} = t_{0.025} = 2.228$

$$S_p^2 = \frac{(n_A-1)S_A^2+(n_B-1)S_B^2}{n_A+n_B-2}$$

$$= \frac{(6-1)(7.86690)+(6-1)(7.10009)}{6+6-2} = 7.4835$$

$$S_p = \sqrt{S_p^2} = \sqrt{7.4835} = 2.7356$$

A 95% C.I. for $\mu_A-\mu_B$ is

$$(\overline{X}_A-\overline{X}_B)-t_{\frac{\alpha}{2}}S_p\sqrt{\frac{1}{n_A}+\frac{1}{n_B}} < \mu_A-\mu_B < (\overline{X}_A-\overline{X}_B)+t_{\frac{\alpha}{2}}S_p\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}$$

or $(\overline{X}_A-\overline{X}_B) \pm t_{\frac{\alpha}{2}}\ S_p\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}$

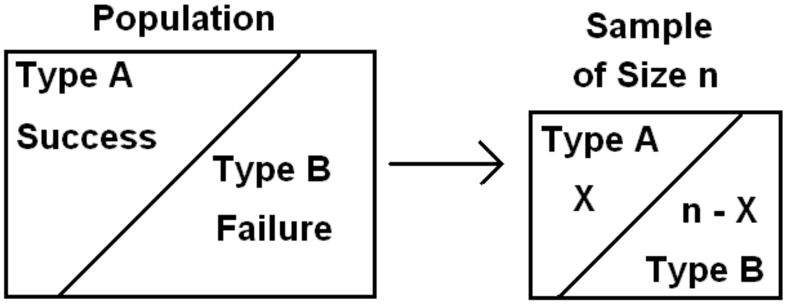
$$(140.67-138.50)\pm(2.228)(2.7356)\sqrt{\frac{1}{6}+\frac{1}{6}}$$

$$2.17 \pm 3.51890$$

$-1.35 < \mu_A-\mu_B < 5.69$

We are 95% confident that $\mu_A-\mu_B \in (-1.35, 5.69)$

9.9 Single Sample: Estimating of a Proportion:



.p = Population proportion of successes (elements of Type A) in the population

$$= \frac{A}{A+B} = \frac{no.\ of\ \ elements\ of\ \ type\ \ A}{Total\ \ no.\ of\ \ elements}$$

.n = sample size

X = no. of elements of type A in the sample of size n.

$\hat{p}$ = Sample proportion of successes (elements of Type A) in the sample

$$= \frac{X}{n}$$

Recall that:

(1) X ~ Binomial (n, p)

(2) E($\hat{p}$)= E($\frac{X}{n}$)= p

(3) Var($\hat{p}$) = Var($\frac{X}{n}$) = $\frac{pq}{n}$; $q = 1 - p$

(4) For large n, we have

$$\hat{p} \sim \text{N}(p, \sqrt{\frac{pq}{n}}) \quad \text{(Approximately)}$$

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \sim \text{N}(0,1) \quad \text{(Approximately)}$$

Point Estimation for *p*:

A good point estimator for the population proportion p is given by the statistic (sample proportion):

$$\hat{p} = \frac{X}{n}$$

Confidence Interval for *p*:

Result:

For large n, an approximate (1−α)100% confidence interval for p is :

$$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\,\hat{q}}{n}} \quad ; \quad \hat{q} = 1 - \hat{p}$$

or

$$\left(\hat{p} - Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\,\hat{q}}{n}} \;,\; \hat{p} + Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\,\hat{q}}{n}} \right)$$

or

$$\hat{p} - Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\,\hat{q}}{n}} < p < \hat{p} + Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\,\hat{q}}{n}}$$

Example 9.10:

In a random sample of *n*=500 families owing television sets in the city of Hamilton, Canada, it was found that *x*=340 subscribed to HBO. Find 95% confidence interval for the actual proportion of families in this city who subscribe to HBO.

Solution:

.*p* = proportion of families in this city who subscribe to HBO.

.*n* = sample size = 500

X = no. of families in the sample who subscribe to HBO=340.

$\hat{p}$ = proportion of families in the sample who subscribe to HBO.

$= \frac{X}{n} = \frac{340}{500} = 0.68$

$\hat{q} = 1 - \hat{p} = 1 - 0.68 = 0.32$

A point estimator for *p* is

$$\hat{p} = \frac{X}{n} = \frac{340}{500} = 0.68$$

Now,

95% = (1–α)100% ⇔ 0. 95 = (1–α) ⇔ α=0.05 ⇔ α/2 = 0.025

$Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96$

A 95% confidence interval for *p* is:

$$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\,\hat{q}}{n}} \quad ; \; \hat{q} = 1 - \hat{p}$$

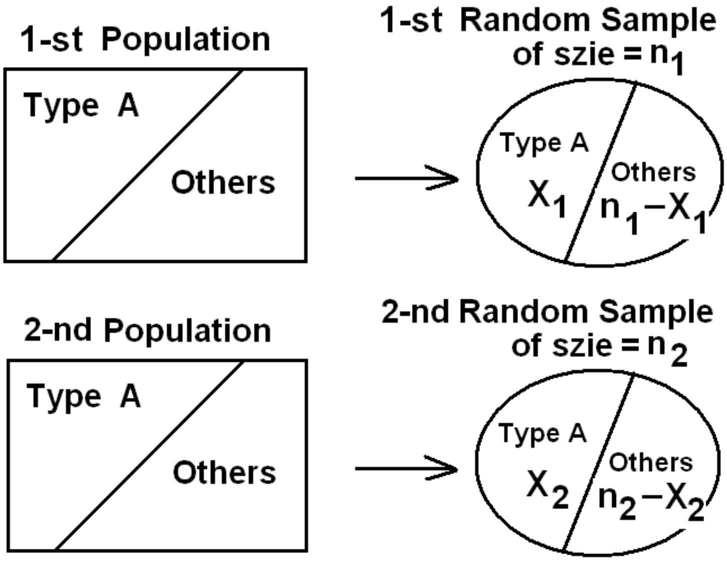
$$0.68 \pm 1.96 \sqrt{\frac{(0.68)(0.32)}{500}}$$

$$0.68 \pm 0.04$$

$$0.64 < p < 0.72$$

We are 95% confident that $p \in (0.64, 0.72)$.

9.10 Two Samples: Estimating the Difference between Two Proportions:



Suppose that we have two populations:

- .p_1 = proportion of the 1-st population.
- .p_2 = proportion of the 2-nd population.
- We are interested in comparing p_1 and p_2, or equivalently, making inferences about $p_1 - p_2$.
- We independently select a random sample of size n_1 from the 1-st population and another random sample of size n_2 from the 2-nd population:
- Let X_1 = no. of elements of type A in the 1-st sample.

$$X_1 \sim \text{Binomial}(n_1, p_1)$$
$$E(X_1) = n_1\, p_1$$
$$\text{Var}(X_1) = n_1\, p_1\, q_1 \quad (q_1 = 1 - p_1)$$

- Let X_2 = no. of elements of type A in the 2-nd sample.

$$X_2 \sim \text{Binomial}(n_2, p_2)$$
$$E(X_2) = n_2\, p_2$$
$$\text{Var}(X_2) = n_2\, p_2\, q_2 \quad (q_2 = 1 - p_2)$$

- $\hat{p}_1 = \frac{X_1}{n_1}$ = proportion of the 1-st sample
- $\hat{p}_2 = \frac{X_2}{n_2}$ = proportion of the 2-nd sample
- The sampling distribution of $\hat{p}_1 - \hat{p}_2$ is used to make inferences about $p_1 - p_2$.

Result:

(1) $E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$

(2) $Var(\hat{p}_1 - \hat{p}_2) = \frac{p_1\, q_1}{n_1} + \frac{p_2\, q_2}{n_2} \quad ; q_1 = 1 - p_1,\ q_2 = 1 - p_2$

(3) For large n_1 and n_2, we have

$$\hat{p}_1 - \hat{p}_2 \sim N(p_1 - p_2, \sqrt{\frac{p_1\, q_1}{n_1} + \frac{p_2\, q_2}{n_2}}) \quad \text{(Approximately)}$$

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1\, q_1}{n_1} + \frac{p_2\, q_2}{n_2}}} \sim N(0,1) \quad \text{(Approximately)}$$

Point Estimation for $p_1 - p_2$:

A good point estimator for the difference between the two proportions, $p_1 - p_2$, is given by the statistic:

$$\hat{p}_1 - \hat{p}_2 = \frac{X_1}{n_1} - \frac{X_2}{n_2}$$

Confidence Interval for $p_1 - p_2$:

Result:

For large n_1 and n_2, an approximate $(1-\alpha)100\%$ confidence interval for $p_1 - p_2$ is :

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

or

$$\left((\hat{p}_1 - \hat{p}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}},\ (\hat{p}_1 - \hat{p}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} \right)$$

or

$$(\hat{p}_1 - \hat{p}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Example 9.13:

A certain change in a process for manufacture of component parts is being considered. Samples are taken using both existing and the new procedure to determine if the new process results in an improvement. If 75 of 1500 items from the existing procedure were found to be defective and 80 of 2000 items from the new procedure were found to be defective, find 90% confidence interval for the true difference in the fraction of defectives between the existing and the new process.

Solution:

.p_1 = fraction (proportion) of defectives of the existing process
.p_2 = fraction (proportion) of defectives of the new process
$\hat{p}_1$= sample fraction of defectives of the existing process
$\hat{p}_2$= sample fraction of defectives of the new process

Existing Process	New Process
$n_1 = 1500$	$n_2 = 2000$
$X_1=75$	$X_2=80$
$\hat{p}_1 = \frac{X_1}{n_1} = \frac{75}{1500} = 0.05$	$\hat{p}_2 = \frac{X_2}{n_2} = \frac{80}{2000} = 0.04$
$\hat{q}_1 = 1 - 0.05 = 0.95$	$\hat{q}_2 = 1 - 0.04 = 0.96$

Point Estimation for $p_1 - p_2$:
A point estimator for the difference between the two proportions, $p_1 - p_2$, is:

$$\hat{p}_1 - \hat{p}_2 = 0.05 - 0.04 = 0.01$$

Confidence Interval for $p_1 - p_2$:
A 90% confidence interval for $p_1 - p_2$ is :

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{0.05} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$0.01 \pm 1.645 \sqrt{\frac{(0.05)(0.95)}{1500} + \frac{(0.04)(0.96)}{2000}}$$

$$0.01 \pm 0.01173$$

$$-0.0017 < p_1 - p_2 < 0.0217$$

We are 90% confident that $p_1 - p_2 \in (-0.0017, 0.0217)$.

Note:

Since $0 \in$ 90% confidence interval=(–0.0017, 0.0217), there is no reason to believe that the new procedure produced a significant decrease in the proportion of defectives over the existing method ($p_1 - p_2 \approx 0 \Leftrightarrow p_1 \approx p_2$).

Chapter 10: One- and Two-Sample Tests of Hypotheses:

Consider a population with some unknown parameter θ. We are interested in testing (confirming or denying) some conjectures about θ. For example, we might be interested in testing the conjecture that $\theta = \theta_o$, where θ_o is a given value.

10.1-10.3: Introduction +:

- A statistical hypothesis is a conjecture concerning (or a statement about) the population.
- For example, if θ is an unknown parameter of the population, we may be interested in testing the conjecture that $\theta > \theta_o$ for some specific value θ_o.
- We usually test the null hypothesis:

$H_o: \theta = \theta_o$ (Null Hypothesis)

Against one of the following alternative hypotheses:

$$H_1: \begin{cases} \theta \neq \theta_o \\ \theta > \theta_o \\ \theta < \theta_o \end{cases}$$ (Alternative Hypothesis or Research Hypothesis)

- Possible situations in testing a statistical hypothesis:

	H_o is true	H_o is false
Accepting H_o	Correct Decision	Type II error (β)
Rejecting H_o	Type I error (α)	Correct Decision

Type I error = Rejecting H_o when H_o is true

Type II error = Accepting H_o when H_o is false

P(Type I error) = P(Rejecting H_o | H_o is true) = α

P(Type II error) = P(Accepting H_o | H_o is false) = β

- The level of significance of the test:

α = P(Type I error) = P(Rejecting H_o | H_o is true)

- One-sided alternative hypothesis:

  $H_o: \theta = \theta_o$ or $H_o: \theta = \theta_o$

  $H_1: \theta > \theta_o$ $\quad$ $H_1: \theta < \theta_o$

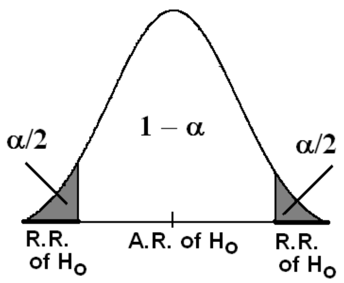
- Two-sided alternative hypothesis:

  $H_o: \theta = \theta_o$

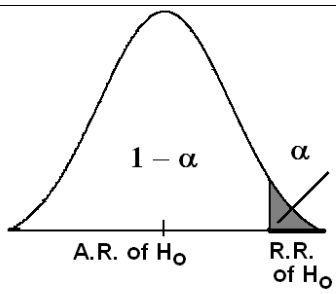
  $H_1: \theta \neq \theta_o$

- The test procedure for rejecting H_o (accepting H_1) or accepting H_o (rejecting H_1) involves the following steps:

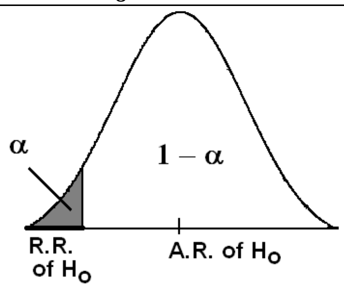
1. Determining a test statistic (T.S.)
2. Determining the significance level α
3. Determining the rejection region (R.R.) and the acceptance region (A.R.) of H_o .
   R.R. of H_o depends on H_1 and α
   - H_1 determines the direction of the R.R. of H_o
   - α determines the size of the R.R. of H_o



$H_1: \theta \neq \theta_o$
Two-sided alternative



$H_1: \theta > \theta_o$
One-sided alternative



$H_1: \theta < \theta_o$
One-sided alternative

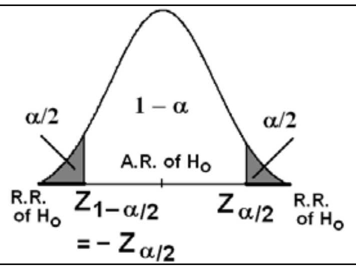
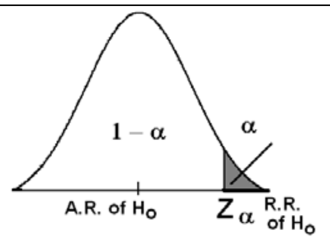
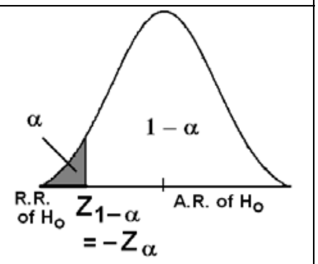
4. Decision:
   We reject H_o (accept H_1) if the value of the T.S. falls in the R.R. of H_o , and vice versa.

10.5: Single Sample: Tests Concerning a Single Mean (Variance Known):

Suppose that $X_1, X_2, \ldots, X_n$ is a random sample of size n from distribution with mean μ and (known) variance σ^2.

Recall:

- $E(\overline{X}) = \mu_{\overline{X}} = \mu$
- $Var(\overline{X}) = \sigma^2_{\overline{X}} = \dfrac{\sigma^2}{n}$
- $\overline{X} \sim N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right) \Leftrightarrow Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- Test Procedure:

Hypotheses	$H_o: \mu = \mu_o$ $H_1: \mu \neq \mu_o$	$H_o: \mu = \mu_o$ $H_1: \mu > \mu_o$	$H_o: \mu = \mu_o$ $H_1: \mu < \mu_o$
Test Statistic (T.S.)	$Z = \dfrac{\overline{X} - \mu_o}{\sigma/\sqrt{n}} \sim N(0,1)$		
R.R. and A.R. of H_o			
Decision:	Reject H_o (and accept H_1) at the significance level α if:		
	$Z > Z_{\alpha/2}$ or $Z < -Z_{\alpha/2}$ Two-Sided Test	$Z > Z_{\alpha}$ One-Sided Test	$Z < -Z_{\alpha}$ One-Sided Test

Example 10.3:

A random sample of 100 recorded deaths in the United States during the past year showed an average of 71.8 years. Assuming a population standard deviation of 8.9 year, does this seem to indicate that the mean life span today is greater than 70 years? Use a 0.05 level of significance.

Solution:

.n=100, $\overline{X}$=71.8, σ=8.9

μ=average (mean) life span

μ_o=70

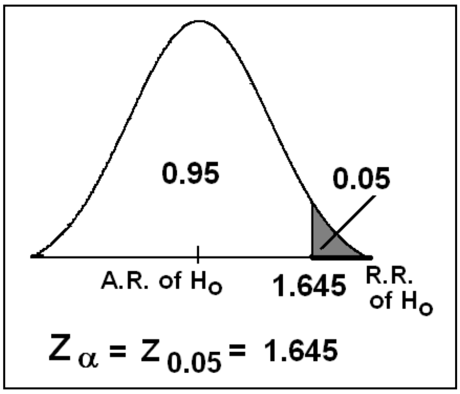
Hypotheses:

H_o: μ = 70

H_1: μ > 70

T.S. :

$$Z = \frac{\overline{X} - \mu_o}{\sigma / \sqrt{n}} = \frac{71.8 - 70}{8.9 / \sqrt{100}} = 2.02$$



Level of significance:

α=0.05

R.R.:

$Z_\alpha = Z_{0.05} = 1.645$

$Z > Z_\alpha = Z_{0.05} = 1.645$

Decision:

Since Z=2.02 ∈R.R., i.e., $Z=2.02>Z_{0.05}$, we reject H_o at α=0.05 and accept H_1: μ > 70. Therefore, we conclude that the mean life span today is greater than 70 years.

Example 10.4:

A manufacturer of sports equipment has developed a new synthetic fishing line that he claims has a mean breaking strength of 8 kilograms with a standard deviation of 0.5 kilograms. Test the hypothesis that μ=8 kg against the alternative that μ≠8 kg if a random sample of 50 lines is tested and found to have a mean breaking strength of 7.8 kg. Use a 0.01 level of significance.

Solution:

.n=50, $\overline{X}$ =7.8, σ=0.5,

α=0.01, α/2 = 0.005

μ= mean breaking strength

μ_o=8

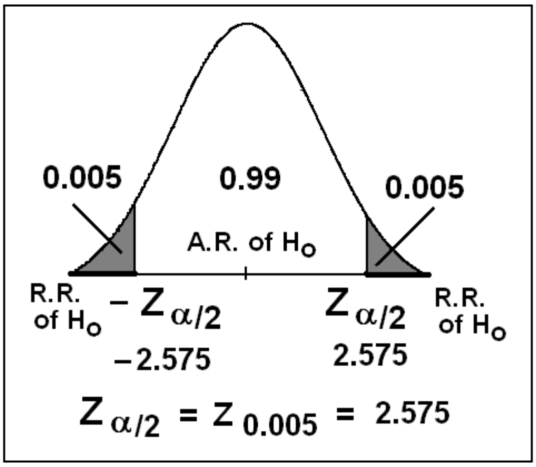
Hypotheses:

H_o: μ = 8

H_1: μ ≠ 8

T.S. :

$$Z = \frac{\overline{X} - \mu_o}{\sigma / \sqrt{n}} = \frac{7.8 - 8}{0.5 / \sqrt{50}} = -2.83$$



$Z_{\alpha/2} = Z_{0.005} = 2.575$ and $-Z_{\alpha/2} = -Z_{0.005} = -2.575$

Decision:

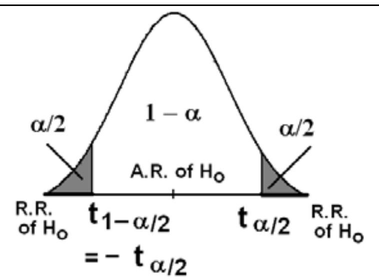
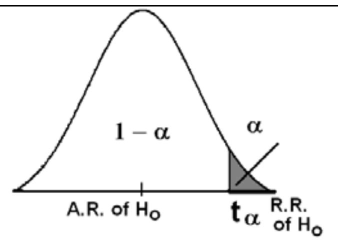
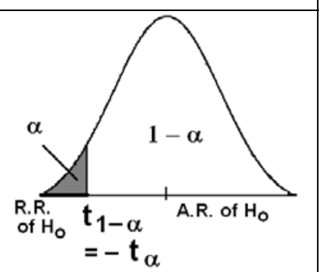
Since Z= −2.83 ∈R.R., i.e., Z= −2.83 < $-Z_{0.005}$, we reject H_o at α=0.01 and accept H_1: μ ≠ 8. Therefore, we conclude that the claim is not correct.

10.7: Single Sample: Tests on a Single Mean (Variance Unknown):

Suppose that $X_1, X_2, \ldots, X_n$ is a random sample of size n from normal distribution with mean μ and unknown variance σ^2.

Recall:

- $T = \dfrac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(\nu)$; $\nu = n-1$
- $S = \sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)}$
- Test Procedure:

Hypotheses	$H_o: \mu = \mu_o$ $H_1: \mu \neq \mu_o$	$H_o: \mu = \mu_o$ $H_1: \mu > \mu_o$	$H_o: \mu = \mu_o$ $H_1: \mu < \mu_o$
Test Statistic (T.S.)	$T = \dfrac{\bar{X} - \mu_o}{S/\sqrt{n}} \sim t(n-1)$		
R.R. and A.R. of H_o			
Decision:	Reject H_o (and accept H_1) at the significance level α if:		
	$T > t_{\alpha/2}$ or $T < -t_{\alpha/2}$ Two-Sided Test	$T > t_{\alpha}$ One-Sided Test	$T < -t_{\alpha}$ One-Sided Test

Example 10.5:

… If a random sample of 12 homes included in a planned study indicates that vacuum cleaners expend an average of 42 kilowatt-hours per year with a standard deviation of 11.9 kilowatt-hours, does this suggest at the 0.05 level of significance that the vacuum cleaners expend, on the average, less than 46 kilowatt-hours annually? Assume the population of kilowatt-hours to be normal.

Solution:

.n=12, $\overline{X}$ =42, S=11.9, α=0.05

μ=average (mean) kilowatt-hours annual expense of a vacuum cleaner

μ_o=46

Hypotheses:

H_o: μ = 46

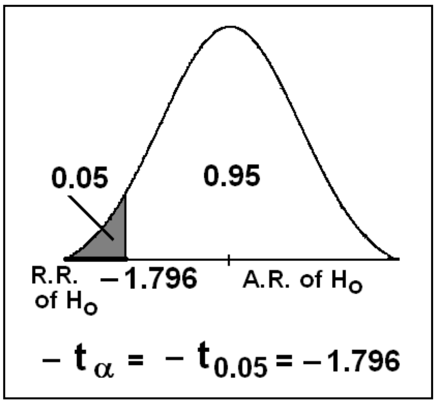
H_1: μ < 46

T.S. :

$$T = \frac{\overline{X} - \mu_o}{S/\sqrt{n}} = \frac{42-46}{11.9/\sqrt{12}} = -1.16 \;;$$

ν=df= n–1=11

$-\,t_{\alpha/2} = -\,t_{0.05} = -1.796$



Decision:

Since T=–1.16$\notin$R.R. (T=–1.16$\in$A.R.), we do not reject H_o at α=0.05 (i.e, accept H_o: μ = 46) and reject H_1: μ < 46. Therefore, we conclude that μ is not less than 46 kilowatt-hours.

<u>10.8: Two Samples: Tests on Two Means:</u>

Recall: For two independent samples:

- If σ_1^2 and σ_2^2 are known, then we have:

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

- If σ_1^2 and σ_2^2 are unknown but $\sigma_1^2 = \sigma_2^2 = \sigma^2$, then we have:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1+n_2-2)$$

Where the pooled estimate of σ^2 is

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

The degrees of freedom of S_p^2 is $\nu = n_1+n_2-2$.

Now, suppose we need to test the null hypothesis

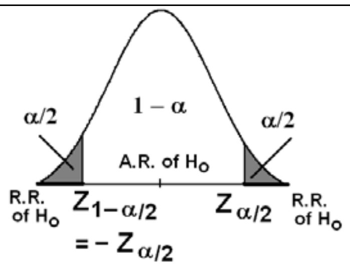
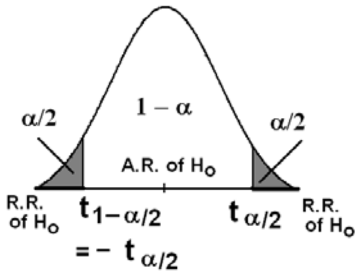
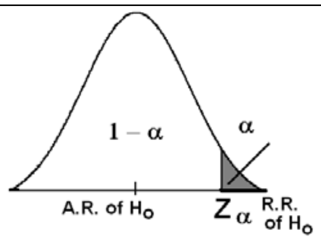
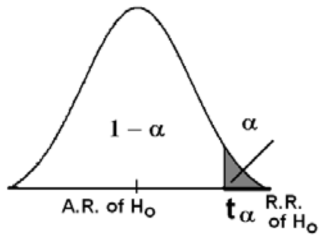
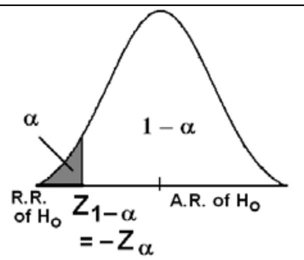
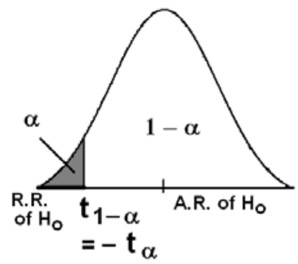
$$H_o: \mu_1 = \mu_2 \quad \Leftrightarrow \quad H_o: \mu_1 - \mu_2 = 0$$

Generally, suppose we need to test

$$H_o: \mu_1 - \mu_2 = d \qquad \text{(for some specific value } d\text{)}$$

Against one of the following alternative hypothesis

$$H_1: \begin{cases} \mu_1 - \mu_2 \neq d \\ \mu_1 - \mu_2 > d \\ \mu_1 - \mu_2 < d \end{cases}$$

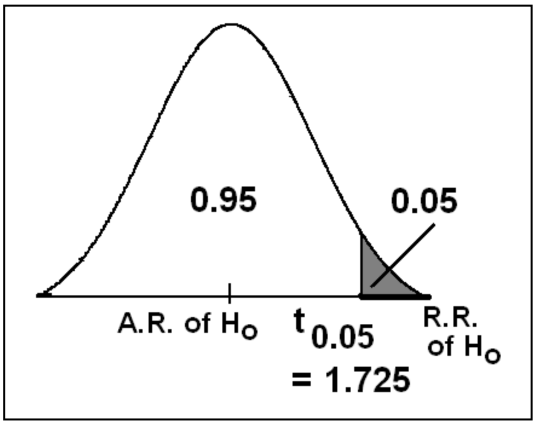
Hypotheses	$H_o: \mu_1 - \mu_2 = d$ $H_1: \mu_1 - \mu_2 \neq d$	$H_o: \mu_1 - \mu_2 = d$ $H_1: \mu_1 - \mu_2 > d$	$H_o: \mu_1 - \mu_2 = d$ $H_1: \mu_1 - \mu_2 < d$
Test Statistic (T.S.)	$Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - d}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ ~N(0,1) {if σ_1^2 and σ_2^2 are known} or $T = \dfrac{(\bar{X}_1 - \bar{X}_2) - d}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ ~t(n_1+n_2-2) {if $\sigma_1^2=\sigma_2^2=\sigma^2$ is unknown}		
R.R. and A.R. of H_o	 Or 	 Or 	 Or 
Decision:	Reject H_o (and accept H_1) at the significance level α if:		
	T.S. ∈ R.R. Two-Sided Test	T.S. ∈ R.R. One-Sided Test	T.S. ∈ R.R. One-Sided Test

Example10.6:

An experiment was performed to compare the abrasive wear of two different laminated materials. Twelve pieces of material 1 were tested by exposing each piece to a machine measuring wear. Teen pieces of material 2 were similarly tested. In each case, the depth of wear was observed. The samples of material 1 gave an average wear of 85 units with a sample standard deviation of 4, while the samples of materials 2 gave an average wear of 81 and a sample standard deviation of 5. Can we conclude at the 0.05 level of significance that the mean abrasive wear of material 1 exceeds that of material 2 by more than 2 units? Assume populations to be approximately normal with equal variances.

Solution:

Material 1	material 2
n_1=12	n_2=10
$\overline{X}_1$=85	$\overline{X}_2$=81
S_1=4	S_2=5



Hypotheses:

H_o: $\mu_1 = \mu_2 + 2$ (d=2)

H_1: $\mu_1 > \mu_2 + 2$

Or equivalently,

H_o: $\mu_1 - \mu_2 = 2$ (d=2)

H_1: $\mu_1 - \mu_2 > 2$

Calculation:

α=0.05

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{(12-1)(4)^2 + (10-1)(5)^2}{12+10-2} = 20.05$$

S_p=4.478

$\nu = n_1+n_2-2=12+10-2 = 20$

$t_{0.05}$ = 1.725

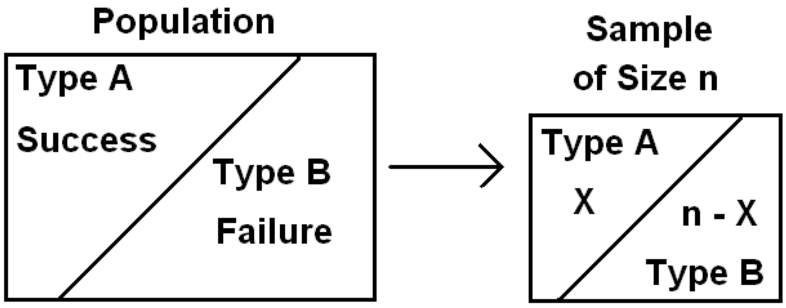
T.S.:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - d}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{(85-81)-2}{(4.478)\sqrt{\frac{1}{12}+\frac{1}{10}}} = 1.04$$

Decision:

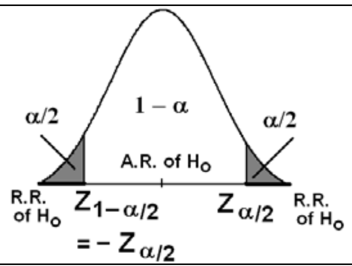
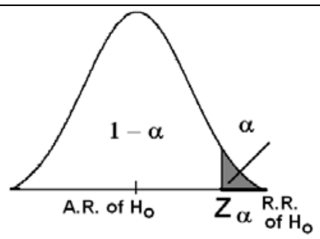
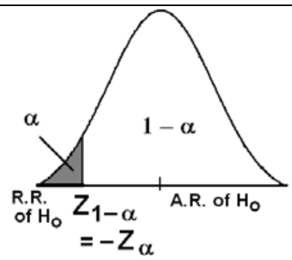
Since T=1.04 $\in$A.R. (T=1.04< $t_{0.05}$ = 1.725), we accept (do not reject) H_o and reject H_1: $\mu_1 - \mu_2 > 2$ at α=0.05.

10.11 One Sample: Tests on a Single Proportion:



Recall:

- .p = Population proportion of elements of Type A in the population
- $= \dfrac{A}{A+B} = \dfrac{no.\ of\ \ elements\ of\ \ type\ \ A}{Total\ \ no.\ of\ \ elements}$
- .n = sample size
- X = no. of elements of type A in the sample of size n.
- $\hat{p}$ = Sample proportion elements of Type A in the sample $= \dfrac{X}{n}$
- For large n, we have
  $Z = \dfrac{\hat{p}-p}{\sqrt{\dfrac{pq}{n}}} = \dfrac{X-np}{\sqrt{npq}} \sim N(0,1)$ (Approximately, $q=1-p$)

Hypotheses	$H_o: p = p_o$ $H_1: p \neq p_o$	$H_o: p = p_o$ $H_1: p > p_o$	$H_o: p = p_o$ $H_1: p < p_o$
Test Statistic (T.S.)	$Z = \dfrac{\hat{p}-p_o}{\sqrt{\dfrac{p_o q_o}{n}}} = \dfrac{X-np_o}{\sqrt{np_o q_o}} \sim N(0,1)$ ($q_o=1-p_o$)		
R.R. and A.R. of H_o			
Decision:	Reject H_o (and accept H_1) at the significance level α if:		
	$Z > Z_{\alpha/2}$ or $Z < -Z_{\alpha/2}$ Two-Sided Test	$Z > Z_{\alpha}$ One-Sided Test	$Z < -Z_{\alpha}$ One-Sided Test

Example 10.10:

A builder claims that heat pumps are installed in 70% of all homes being constructed today in the city of Richmond. Would you agree with this claim if a random survey of new homes in the city shows that 8 out of 15 homes had heat pumps installed? Use a 0.10 level of significance.

Solution:

.p = Proportion of homes with heat pumps installed in the city.

.n=15

X= no. of homes with heat pumps installed in the sample = 8

$\hat{p}$ = proportion of homes with heat pumps installed in the sample = $\frac{X}{n} = \frac{8}{15} = 0.5333$

Hypotheses:

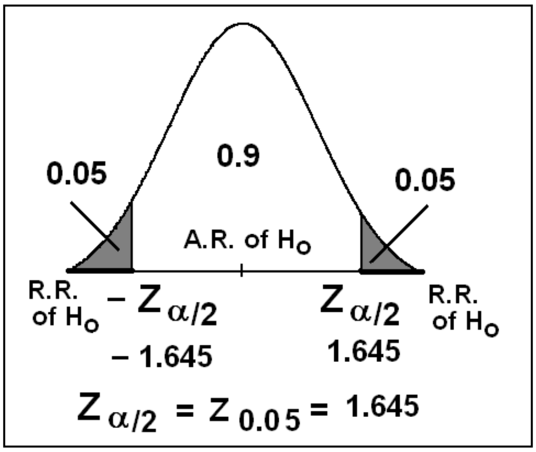
H_o: $p = 0.7$ (p_o=0.7)

H_1: $p \neq 0.7$

Level of significance:

α=0.10

T.S.:



$$Z = \frac{\hat{p} - p_o}{\sqrt{\frac{p_o q_o}{n}}} = \frac{0.5333 - 0.70}{\sqrt{\frac{(0.7)(0.3)}{15}}} = -1.41$$

or

$$Z = \frac{X - np_o}{\sqrt{np_o q_o}} = \frac{8 - (15)(0.7)}{\sqrt{(15)(0.7)(0.3)}} = -1.41$$

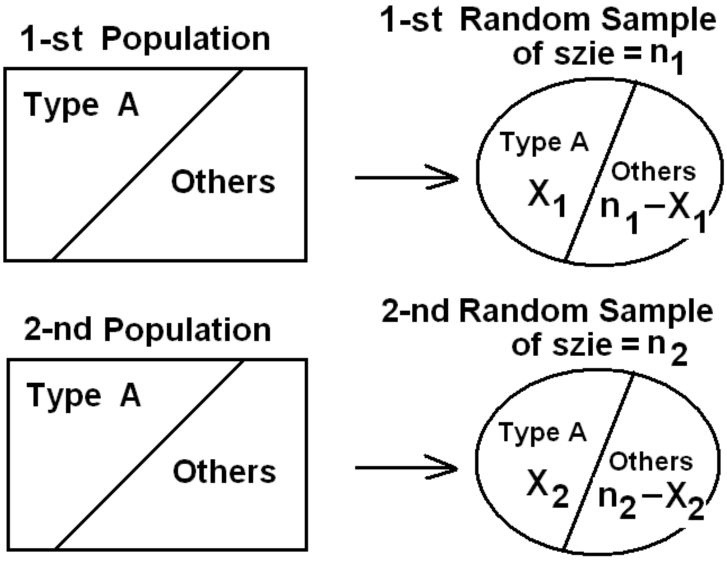
$Z_{\alpha/2}$= $Z_{0.05}$= 1.645

Decision:

Since Z= –1.41 $\in$A.R., we accept (do not reject) H_o: p=0.7 and reject H_1: $p \neq 0.7$ at α=0.1. Therefore, we agree with the claim.

Example 10.11: Reading Assignment

10.12 Two Samples: Tests on Two Proportions:



Suppose that we have two populations:

- .p_1 = proportion of the 1-st population.
- .p_2 = proportion of the 2-nd population.
- We are interested in comparing p_1 and p_2, or equivalently, making inferences about $p_1 - p_2$.
- We <u>independently</u> select a random sample of size n_1 from the 1-st population and another random sample of size n_2 from the 2-nd population:
- Let X_1 = no. of elements of type A in the 1-st sample.
- Let X_2 = no. of elements of type A in the 2-nd sample.
- $\hat{p}_1 = \frac{X_1}{n_1}$ = proportion of the 1-st sample
- $\hat{p}_2 = \frac{X_2}{n_2}$ = proportion of the 2-nd sample
- The sampling distribution of $\hat{p}_1 - \hat{p}_2$ is used to make inferences about $p_1 - p_2$.
- For large n_1 and n_2, we have

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1\, q_1}{n_1} + \frac{p_2\, q_2}{n_2}}} \sim \text{N}(0,1) \quad \text{(Approximately)}$$

Suppose we need to test:

H_o: $p_1 = p_2$

H_1: $\begin{cases} p_1 \neq p_2 \\ p_1 > p_2 \\ p_1 < p_2 \end{cases}$

Or, equivalently,

H_o: $p_1 - p_2 = 0$

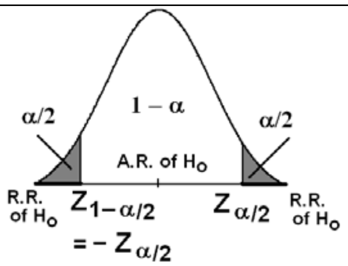
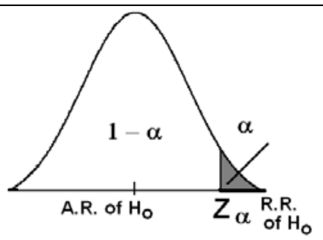
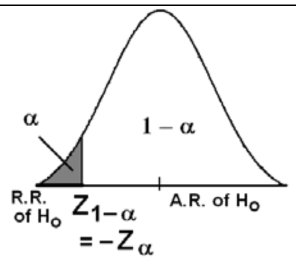
H_1: $\begin{cases} p_1 - p_2 \neq 0 \\ p_1 - p_2 > 0 \\ p_1 - p_2 < 0 \end{cases}$

Note, under H_o: $p_1 = p_2 = p$, the pooled estimate of the proportion p is:

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2} \qquad (\hat{q} = 1 - \hat{p})$$

The test statistic (T.S.) is

$$Z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

Hypotheses	H_o: $p_1 - p_2 = 0$ H_1: $p_1 - p_2 \neq 0$	H_o: $p_1 - p_2 = 0$ H_1: $p_1 - p_2 > 0$	H_o: $p_1 - p_2 = 0$ H_1: $p_1 - p_2 < 0$
Test Statistic (T.S.)	$Z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$		
R.R. and A.R. of H_o			
Decision:	Reject H_o (and accept H_1) at the significance level α if:		
	$Z > Z_{\alpha/2}$ or $Z < -Z_{\alpha/2}$ Two-Sided Test	$Z > Z_{\alpha}$ One-Sided Test	$Z < -Z_{\alpha}$ One-Sided Test

Unrevised First Draft

Example 10.12:

A vote is to be taken among the residents of a town and the surrounding county to determine whether a proposed chemical plant should be constructed. The construction site is within the town limits and for this reason many voters in the county feel that the proposal will pass because of the large proportion of town voters who favor the construction. To determine if there is a significant difference in the proportion of town voters and county voters favoring the proposal, a poll is taken. If 120 of 200 town voters favor the proposal and 240 of 500 county voters favor it, would you agree that the proportion of town voters favoring the proposal is higher than the proportion of county voters? Use a 0.025 level of significance.

Solution:

.p_1 = proportion of town voters favoring the proposal

.p_2 = proportion of county voters favoring the proposal

$\hat{p}_1$ = sample proportion of town voters favoring the proposal

$\hat{p}_2$ = sample proportion of county voters favoring the proposal

<u>Town</u>

$n_1 = 200$

$X_1 = 120$

$\hat{p}_1 = \frac{X_1}{n_1} = \frac{120}{200} = 0.60$

$\hat{q}_1 = 1 - 0.60 = 0.40$

<u>County</u>

$n_2 = 500$

$X_2 = 240$

$\hat{p}_2 = \frac{X_2}{n_2} = \frac{240}{500} = 0.48$

$\hat{q}_2 = 1 - 0.48 = 0.52$

The pooled estimate of the proportion p is:

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{120 + 240}{200 + 500} = 0.51$$

$\hat{q} = 1 - 0.51 = 0.49$

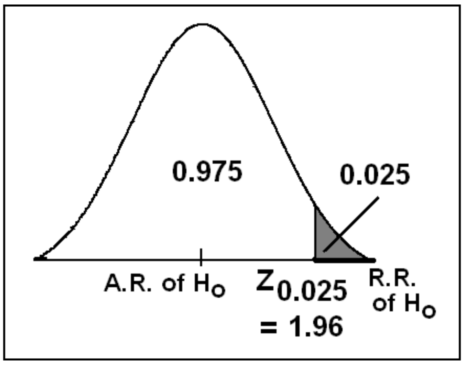
Hypotheses:

H_o: $p_1 = p_2$

H_1: $p_1 > p_2$

or

Unrevised First Draft



$H_o: p_1 - p_2 = 0$
$H_1: p_1 - p_2 > 0$

Level of significance:
$\alpha=0.025$
$Z_\alpha = Z_{0.025} = 1.96$
T.S.:

$$Z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{(0.60-0.48)}{\sqrt{(0.51)(0.49)\left(\frac{1}{200}+\frac{1}{500}\right)}} = 2.869$$

Decision:

Since Z= 2.869 $\in$R.R. ($Z= 2.869 > Z_\alpha = Z_{0.025} = 1.96$), we reject $H_o: p_1 = p_2$ and accept $H_1: p_1 > p_2$ at $\alpha=0.025$. Therefore, we agree that the proportion of town voters favoring the proposal is higher than the proportion of county voters.