

Some Basic Probability Concept and Probability Distribution



Odisha Social Science
Research & Consultancy

CONTENT

- Some basic probability concept
- **Screening test, sensitivity, specificity, predictive value positive and negative**
- Probability distribution
 - Introduction
 - Probability distribution of discrete variable
 - The Binomial Distribution
 - The Poisson Distribution
- Continuous Probability distribution
 - The Normal Distribution

Some Basic Probability Concept

- The theory of probability provides the foundation for statistical inference.
- Some Common probability parlance in medical:
 - A patient has a 50-50 chance of surviving a certain operation.
 - Another physician may say that she is 95% certain that a patient has a particular disease.
 - A public health nurse may say that nine times out of ten a certain client will break an appointment.
- We measure the probability of occurrence of some event by a number between zero and one.

Some Basic Probability Concept

- The more likely the event, the closer the prob. is to one and more unlikely it is the closer the prob. to '0'.
- An event that can't occur has a probability zero, and an event that is certain to occur has a probability of one.
- Health sciences researchers continually ask themselves if the result of their efforts could have occurred by chance alone or if some other force was operating to produce the observed effect.
- For example, suppose six out of ten patients suffering from some disease are cured after receiving a certain treatment.
- Is such a cure rate likely to have occurred if the patients had not received the treatment, or is it evidence of a true curative effect on the part of the treatment?
- Question such as these can be answered through the application of the concepts and laws of probability.

Some Basic Probability Concept Contd..

- **TWO VIEWS OF PROBABILITY- OBJECTIVE AND SUBJECTIVE**
- **Objective Probability**
 - Classical or a priori probability
 - The relative frequency , or a posteriori,
- **Classical Probability**
- Definition- If an event can occur in N mutually exclusive and equally likely ways, and if 'm' of these possess a trait, E, the probability of the occurrence of E is equal to m/N.
- Symbolically-
 - P(E) – Prob. of E

$$P(E) = \frac{m}{N}$$

Some Basic Probability Concept Contd..

■ Relative frequency probability

- If the process is repeated a large number of times, n , and if some resulting event with the characteristic E occurs ' m ' times, the relative frequency of occurrence of E , m/n , will be approximately equal to the probability of E . **Symbolically $P(E) = m/n$**

■ Subjective Probability (In the early 1950s, L.J. Savage)

- Probability measures the confidence that a particular individual has in the truth of a particular proposition. This concept doesn't rely on the repeatability of any process. In fact, by applying this concept of probability, one may evaluate the probability of an event that can only happen once, for example, the probability that a cure for cancer will be discovered within the next 10 years.

Some Basic Probability Concept Contd..

ELEMENTARY PROPERTIES OF PROBABILITY (Axiomatic approach to probability--Russian Mathematician A. N. Kolmogorov in 1933)

The basic of this approach is embodied in three properties:

Axiom of non-negativity: Given some process (or experiment) with n mutually exclusive outcomes (called events), $E_1, E_2, \dots, E_n$, the probability of any event E_i is assigned a nonnegative number. That is $P(E_i) \geq 0$

Some Basic Probability Concept Contd..

Axiom of exhaustiveness: The sum of the probabilities of the all mutually exclusive and exhaustive outcomes is equal to 1. i.e. $P(E_1) + P(E_2) + \dots + P(E_n) = 1$

This is the property of exhaustiveness.

Axiom of additive-ness: Consider any two mutually exclusive event, E_i and E_j . The probability of the occurrence of either E_i or E_j is equal to the sum of their individual probabilities.

$$P(E_i \text{ or } E_j) = P(E_i) + P(E_j)$$

Suppose the two events were not mutually exclusive; in that case the probability:

$$P(E_i \text{ or } E_j) = P(E_i) + P(E_j) - P(E_i \cap E_j)$$

CALCULATING THE PROBABILITY OF AN EVENT

Example:

The subjects in the study consisted of a sample of 75 men and 36 women. The subjects are a fairly representative sample of “typical” adult users who were neither in treatment nor in jail. Table 1 shows the life time frequency of cocaine use and the gender of these subjects. Suppose we pick a person at random from this sample. What is the probability that this person will be a male?

Life frequency of cocaine use	Male(M)	Female(F)	total
1-19 times(A)	32	7	39
20-99 times(B)	18	20	38
100+ times (C)	25	9	34
Total	75	36	111

- We define the desired probability as the number of subjects with the characteristic of interest (male) divided by the total number of subjects. Symbolically
 - $P(M) = \text{number of males} / \text{total number of subjects} = 75/111 = .6757$

- **Conditional Probability**

- When probabilities are calculated with a subset of the total group as the denominator, the result is a conditional probability.

- **Example:-**

- Suppose we pick a subject at random from the 111 subjects and find that he is a male (M). What is the probability that this male will be one who has used cocaine 100 times or more during his lifetime?

- **Solution:**

- This is a conditional probability and is written as $P(C|M)$. The 75 males become the denominator of this conditional probability, and 25 the number of males who have used cocaine 100 times or more during their lifetime, becomes the numerator. Our desired probability, then is $P(C|M) = 25/75 = 0.33$

- **Joint Probability**

- The probability that a subject picked at a random from a group of subjects possesses two characteristics at the same time.

- **Example:**Ref. (Table 1) what is the probability that a person picked at a random from the 111 subjects will be a male (M) and be person who has used cocaine 100 times or more during his lifetime (C)?

- **Solution:**The probability is written as $P(M \cap C)$

- $(M \cap C)$ indicates the joint occurrence of conditions M and C.

- No. of subjects satisfying both of the desired conditions is found in Table-1 at the intersection of the column labeled M and the row labeled C i.e. 25 since the selection will be made from the total set of subjects, the denominator is 111.

- $$P(M \cap C) = 25/111 = 0.2252$$

- **Multiplication Rule**

- A probability may be computed from other probability. For example, a joint probability may be computed as the product of an appropriate marginal probability and an appropriate conditional probability. This relationship is known as the multiplication rule of probability.

- We may state the multiplication rule in general terms as follows for any two events A & B:

- $P(A \cap B) = P(B) P(A|B)$, if $P(B) \neq 0$ or
- $P(A \cap B) = P(A) P(B|A)$, if $P(A) \neq 0$

- **Conditional Probability**

- The conditional probability of A given B is equal to the probability of $A \cap B$ divided by the probability of B, provided the probability of B is not zero.

- That is $\Rightarrow P(A|B) = P(A \cap B) / P(B)$, $P(B) \neq 0$ and same way $P(B|A) = P(A \cap B) / P(A)$, $P(A) \neq 0$

- **Example:**

- Using the equation and the data of Table-1 to find the conditional probability, $P(C|M)$.

- **Solution:**

- According to the equation

- $P(C|M) = P(C \cap M) / P(M)$

- Earlier we found $P(C \cap M) = P(M \cap C) = 25/111 = 0.2252$. We have also determined that $P(M) = 75/111 = 0.6757$.

- Using these results we are able to compute

- $P(C|M) = 0.2252 / 0.6757 = 0.3333$

- **Addition Rule:**

- Given two events A and B, the probability that event A, or event B, or both occur is equal to the probability that event A occurs, plus the probability that event B occurs, minus the probability that the events occur simultaneously.

- The addition rule may be written $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- **Example:**

- If we select a person at random from the 111 subjects represented in Table-1 what is the probability that this person will be a male (M) or will have used cocaine 100 times or more during his lifetime (C) or both?

- **Solution:**

- The probability is $P(M \cup C)$. By the addition rule as expressed this probability may be written as $P(M \cup C) = P(M) + P(C) - P(M \cap C)$
- We have already found that $P(M) = 75/111 = 0.6757$, and $P(M \cap C) = 25/111 = 0.2252$
- $P(C) = 34/111 = 0.3063$
- Substituting these results in to the equation for $P(M \cup C) = 0.6757 + 0.3063 - 0.2252 = 0.7568$.

- **Independence of Events**
- Suppose that the probability of event A is the same regardless of whether or not B occurs. In this situation, $P(A|B) = P(A)$. In such cases we say that A and B are independent events. The multiplication rule for two independent events, then may be written as:
 - $P(A \cap B) = P(B)P(A); P(A) \neq 0, P(B) \neq 0$
 - Thus we see that if two events are independent, the probability of their joint occurrence is equal to the product of the probabilities of their individual occurrences.
 - When two events with non zero probabilities are independent each of the following statements is true:
 - $P(A|B) = P(A), P(B|A) = P(B), P(A \cap B) = P(A) P(B)$

- **Complementary Events:**
- Earlier using the data in Table 1 we computed the probabilities:
- $P(M) = 75/111 = .6757$. $P(F) = 36/111 = .3243$. The sum of these two probabilities = 1. This is because the events being male and being female are complementary events.
- The probability of an event A is equal to 1 minus the probability of its complement, which is written as A' , and
- $$P(A') = 1 - P(A)$$
- This follows from the third property of probability since the event, A and its complement A' , are mutually exclusive and exhaustive
- **Example:**
- Suppose that of 1200 Admissions to a general hospital during a certain period of time, 750 are private admissions. If we designate these as set A' , then A is equal to 1200 minus 750, or 450. we may compute
- $P(A) = 750/1200 = 0.625$ and $P(A') = 450/1200 = 0.375$ and see that $P(A) + P(A') = 1$

- **Marginal Probability**

- Given some variable that can be broken down in to 'm' categories as $A_1, A_2, \dots, A_i, \dots, A_m$ and another jointly occurring variable into 'n' categories as $B_1, B_2, \dots, B_j, \dots, B_n$ the marginal probability of A_i , $P(A_i)$, is equal to the sum of the joint probabilities of A_i with all the categories of B .

- That is $P(A_i) = \sum P(A_i \cap B_j)$, for all values of j .

- **Example:**

- We use the equation and the data in Table-1 to compute the marginal probability $P(M)$.

- **Solution:**

- The variable gender is broken down in to two categories male and female and the variable 'frequency of cocaine use' in to 3 categories, 1-19 times (A), 20-99 times (B), 100+times (C). The category male occurs jointly with all 3 categories of the variable 'frequency of cocaine use'. The three joint probabilities are:

- $P(M \cap A) = 32/111 = 0.2883$, $P(M \cap B) = 18/111 = 0.1662$,

- $P(M \cap C) = 25/111 = 0.2252$

- We obtain the marginal probability: $P(M) = P(M \cap A) + P(M \cap B) + P(M \cap C) = 0.2883 + 0.1622 + 0.2252 = 0.6757$

Screening Test, Sensitivity, Specificity, Predictive Value Positive And Negative

- Screening tests are not always infallible. That is a testing procedure may yield a false positive or a false negative.
- A false positive results when a test indicates a positive status when the true status is negative.
- A false negative results when a test indicates a negative status when the true status is positive.

- The following questions must be answered in order to evaluate the usefulness of test results.
 - Given that a subject has the disease, what is the probability of a positive test result (or the presence of a symptom)?
 - Given that a subject doesn't have the disease, what is the probability of a negative test result (or the absence of a symptom)?
 - Given a positive screening test (or the presence of a symptom), what is the probability that the subject has the disease?
 - Given a negative screening test (or the absence of a symptom), what is the probability that the subject doesn't have the disease?

Table-2 Sample of n subjects (where n is large) cross-classified according to disease status and screening test result.

Test result	Present (D)	Absent (D')	Total
Positive(T)	a	b	a+b
Negetive (T')	c	d	c+d
Total	a+c	b+d	n

Sensitivity of a Test (Probability of True positive)

It is the probability of a positive test result (or presence of the system) given the presence of the disease.

Symbolically this is $P(T|D) = a/(a+c)$

Specificity of a Test (Probability of True Negative)

It is the probability of a negative test result (absence of the system) given the absence of the disease.

Symbolically estimate of Specificity is given by the conditional probability $P(T'|D') = d/(b+d)$

The Predictive Value of a Positive Test

The predictive value positive of a screening test (or symptom) is the probability that a subject has the disease given that the subject has a positive screening result (or has the symptom).

Symbolically, this is given by the conditional probability
 $P(D|T) = a/(a+b)$

Predictive Value of a Negative test

The predictive value negative of a screening test (or symptom) is the probability that a subject doesn't have the disease given that the subject has a negative screening test result (or doesn't have the symptom).

Symbolically, this is given by the conditional probability
 $P(D'|T') = d/(c+d)$

- **Probability of false negative:** It is the probability that the screening test is negative given that the subject has disease.
 - $P(T'|D) = c / (a + c)$.
- **Probability of false positive:** It is the probability that the screening test is positive given that the subject does not have disease.
 - $P(T|D') = b / (b + d)$.

Some Theoretical Probability Distribution

- ❑ The relationship between the values of a random variable and the probability of their occurrence summarized by means of a device called a probability distribution.
- ❑ It may be expressed in the form of a table, a graph or a formula.
- ❑ Knowledge of probability distribution of a random variable provides the researchers with a powerful tool for summarizing and describing a set of data and for reaching conclusion about a population on the basis of sample.

Definition

- The probability distribution of a discrete random variable is a table, graph, formula, or other device used to specify all possible values of a discrete random variable along with their respective probabilities.

Example: Prevalence of prescription and nonprescription drug use in pregnancy among women delivered at a large eastern hospital. N=4185

No of drugs	0	1	2	3	4	5	6	7	8	9	10	12
f	1425	1351	793	348	156	58	28	15	6	3	1	1

Construct a probability distribution table.

What is the probability that a randomly selected women who use 3 drugs?

What is the probability that a woman picked at random who used 2 or fewer drugs?

What is the probability that a woman use 5 or more drugs?

The Binomial Distribution

- ❑ Most widely encountered probability distribution.
- ❑ Derived from a process known as a Bernoulli trial.
- ❑ When a random process or experiment, called a trial can result only one of two mutually exclusive outcomes, such as dead or alive, sick or well, male or female, the trial is called a Bernoulli trial.

Characteristics of a binomial random variable

- The experiment consists of n identical trials
- There are only 2 possible outcomes on each trial. We will denote one outcome by S (for Success) and the other by F (for Failure).
- The probability of S remains the same from trial to trial. This probability will be denoted by p , and the probability of F will be denoted by q ($q = 1-p$).
- The trials are independent.
- The binomial random variable x is the number of S ' in n trials.

The probability distribution:

$$P_x = {}^nC_x p^x q^{n-x} \quad (x = 0, 1, 2, \dots, n),$$

Where,

p = probability of a success on a single trial, $q=1-p$

n = number of trials, x = number of successes in n trials

$${}^nC_x = \frac{n!}{x!(n-x)!} = \text{combination of } x \text{ from } n$$

Binomial Table

- The calculation of probability using the probability mass function is tedious if the sample size is large. There are binomial probability tables available where probabilities for different values of n , p and x have been tabulated.

Binomial Parameters

- The binomial distribution has two parameters n and p which are sufficient to specify a binomial distribution.
- The mean of the binomial distribution is $\mu=np$
- The variance of the binomial distribution $\sigma^2 = np(1-p)$
- Strictly speaking the binomial distribution is applicable where the sampling is from an infinite population or finite population with replacement. But in practice the sampling is mostly without replacement. Hence question arises on the appropriateness of this distribution under these circumstances.

Example:

In a certain population 10% of the population is color blind. If a random sample of 25 people is drawn from this population. Find the probability that:

a) Five or fewer will be color blind

$$P(X \leq 5) = .9666$$

b) Six or more will be color blind

The probability that six or more are color blind is the complement of the probability that five or fewer are not color blind.

$$P(X \geq 6) = 1 - P(X \leq 5) = 1 - .9666 = .0334$$

c) Between six and nine inclusive will be color blind

$$P(6 \leq X \leq 9) = P(X \leq 9) - P(X \leq 5) = .9999 - .9666 = .0333$$

d) Two, three or four will be color blind

$$P(2 \leq X \leq 4) = P(X \leq 4) - P(X \leq 1) = .9020 - .2712 = .6308$$

The Poisson Distribution

- ❖ This distribution has been used extensively as a probability model in biology and medicine.
If x is the number of occurrences of some random events in an interval of time or space, the probability that x will occur is:

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad x=0,1,2,\dots, \lambda>0$$

λ = parameter (average number of occurrence of the random event in the interval)
 e = constant = 2.7183

Characteristics defining a Poisson random variable

- Poisson distribution occurs when there are events which don't occur as outcome of a definite numbers of trials of an experiment but which occur at random point of time and space and our interest lies in the occurrence of the event only.
- The experiment consists of counting the number x of times a particular event occurs during a given unit of time
- The probability that an event occurs in a given unit of time is the same for all units.
- The number of events that occur in one unit of time is independent of the number that occur in other units of time.

- **Theoretically an infinite number of occurrence of the event must be possible in the interval.**
- The poisson event is an infrequently occurring events with very small probability of occurrence.
- Some situations where the poisson process can be employed successfully are
 - **No. of deaths from a disease such as heart attack or cancer or due to snake bite**
 - **No. of suicide reported in a particular city**

Example:

In a study of suicide it was found that monthly distribution of suicide in Cook country, between 1977 and 1987 closely followed a poisson distribution with parameter $\lambda = 2.75$. Find the probability that a randomly selected month will be one in which three adolescent suicide occurs.

Solution

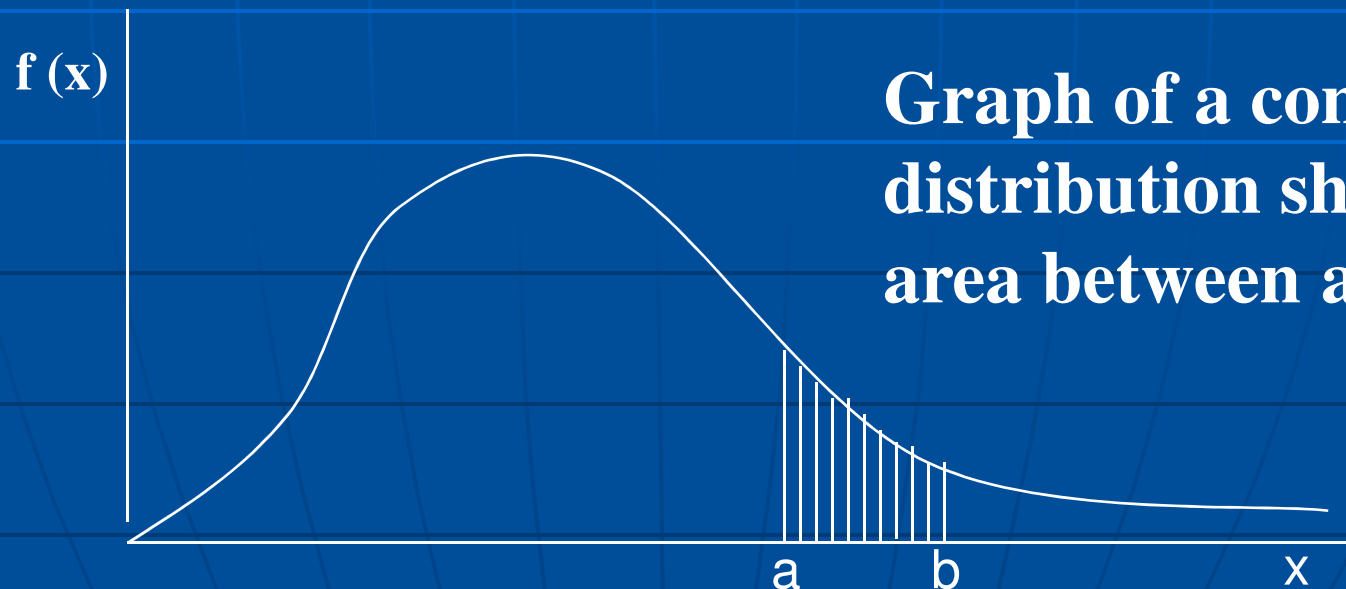
$$\begin{aligned} P(X=3) &= \frac{e^{-2.75} 2.75^3}{3!} \\ &= \frac{(.063928) (20.796875)}{6} \\ &= .221584 \end{aligned}$$

Continuous Probability Distribution

Continuous variable:

- Assumes any value within a specified interval/range.
- Consequently any two values within a specified interval, there exists an infinite number of values.
- As the number of observation, n , approaches infinite and the width of the class interval approaches zero, the frequency polygon approaches smooth curve.
- Such smooth curves are used to represent graphically the distribution of continuous random variable

- This has some important consequences when we deal with probability distributions.
- The total area under the curve is equal to 1 as in the case of the histogram.
- The relative frequency of occurrence of values between any two points on the x-axis is equal to the total area bounded by the curve, the x-axis and the perpendicular lines erected at the two points.



Graph of a continuous distribution showing area between a and b

Definition of probability distribution:

- ❑ A density function is a formula used to represent the probability distribution of a continuous random variable.
- ❑ This is a nonnegative function $f(x)$ of the continuous r.v, x if the total area bounded by its curve and the x -axis is equal to 1 and if the sub area under the curve bounded by the curve, the x -axis and perpendiculars erected at any two points a and b gives the probability that x is between the point a and b .

Normal Distribution(C.F.Gauss, 1777-1855)

- ❑ The distribution is frequently called the Gaussian distribution.
- ❑ It is a relative frequency distribution of errors, such errors of measurement. This curve provides an adequate model for the relative frequency distributions of data collected from many different scientific areas.
- ❑ The density function for a normal random variable

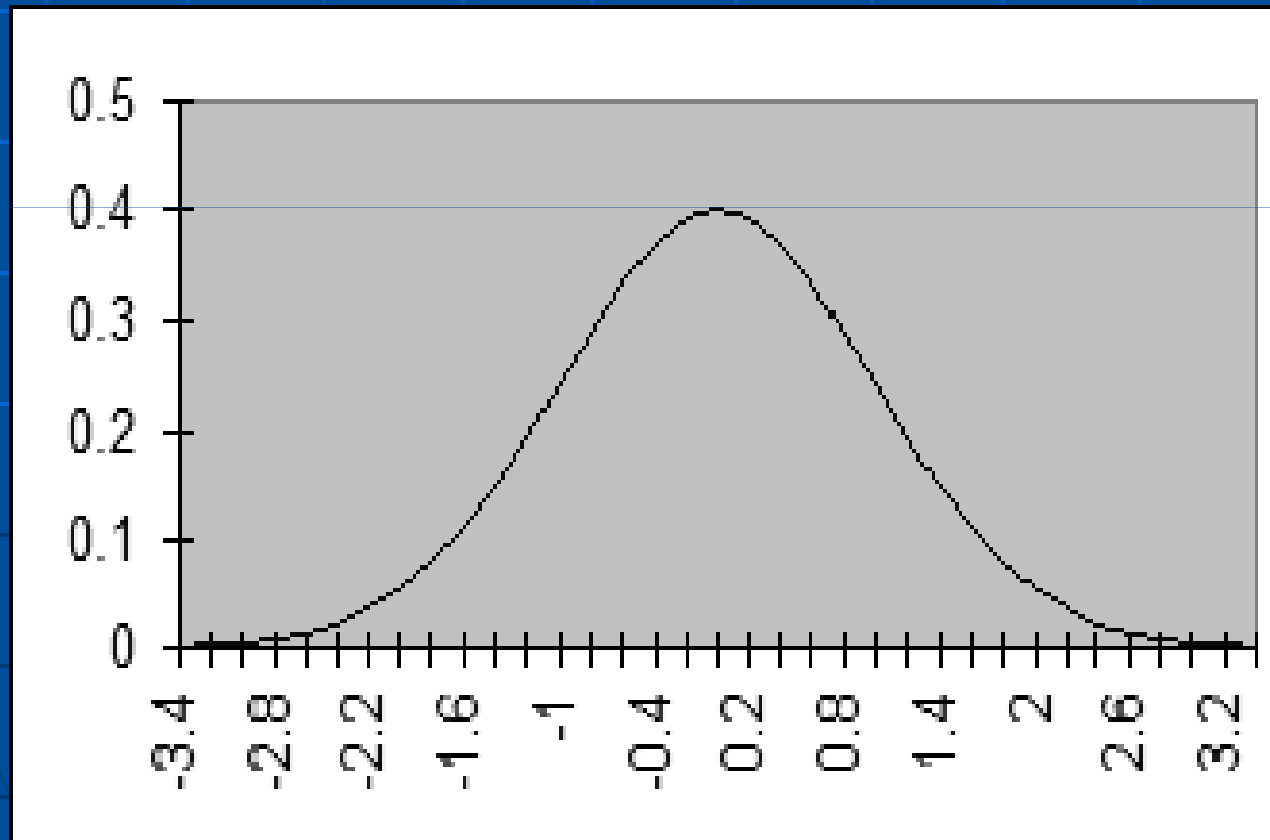
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

- ❑ The parameters μ and σ^2 are the mean and the variance, respectively, of the normal random variable

Characteristics Of The Normal Distribution

- It is symmetrical about its mean, μ .
- The mean, the median, and the mode are all equal.
- The total area under the curve above the x-axis is one square unit.
- This characteristic follows that the normal distribution is a probability distribution.
- Because of the symmetry already mentioned, 50% of the area is to the right of a perpendicular erected at the mean, and 50% is to the left.

➤ If $\mu = 0$ and $\sigma = 1$ then . The distribution with this density function is called the standardized normal distribution. The graph of the standardized normal density distribution is shown in Figure



- If 'x' is a normal random variable with the mean μ and variance σ then

1) the variable

$$z = \frac{x - \mu}{\sigma}$$

is the standardized normal random variable.

The equation of pdf for standard normal distribution

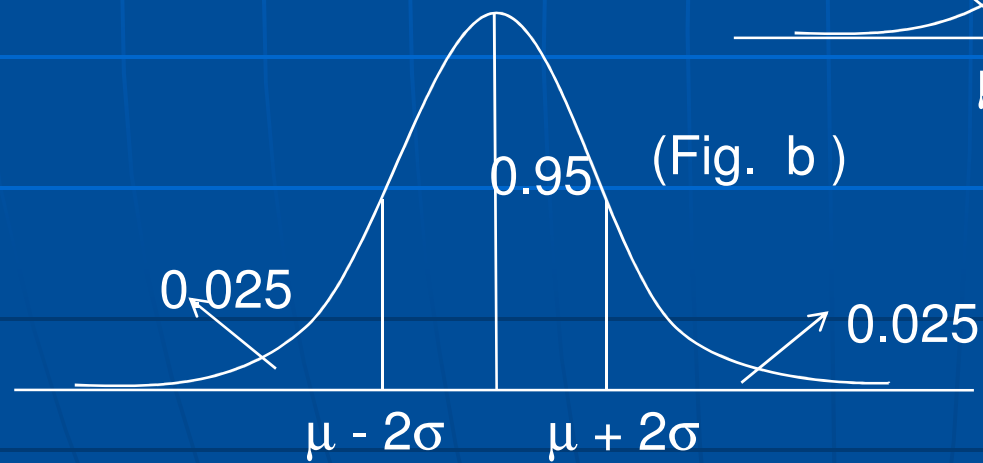
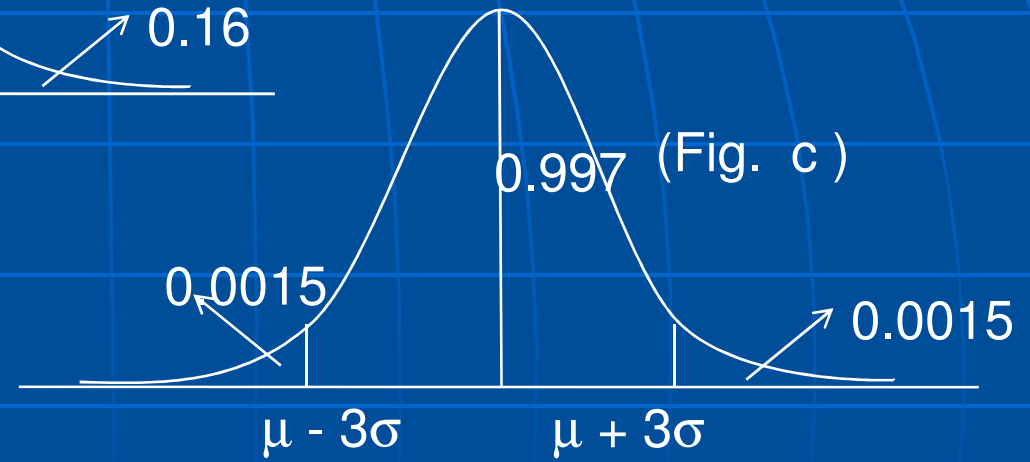
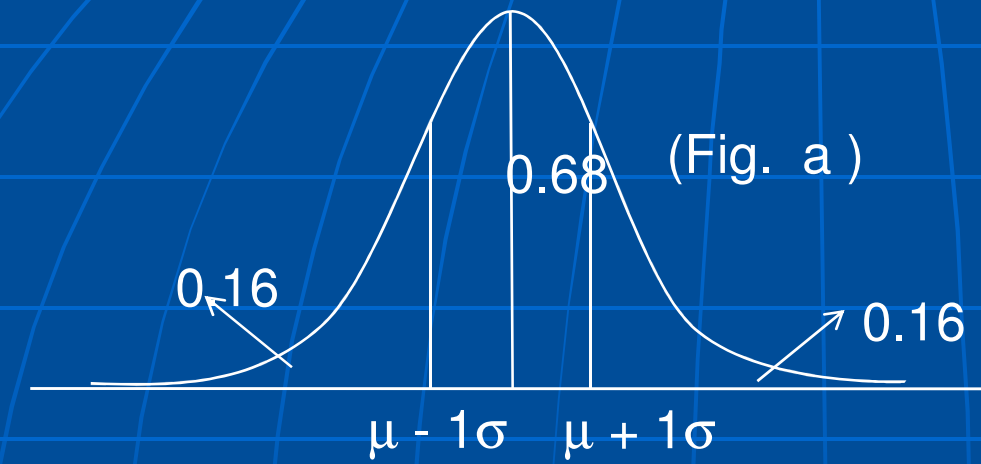
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, -\infty < z < \infty$$

- Area properties of normal distribution

$$P(|x - \mu| \leq 1\sigma) = 0.680$$

$$P(|x - \mu| \leq 2\sigma) = 0.9544$$

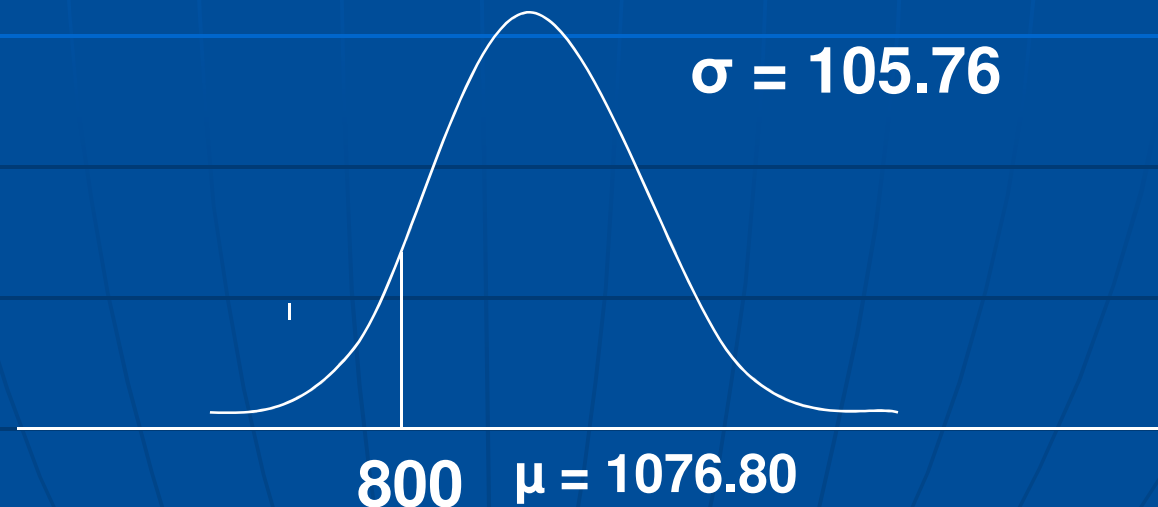
$$P(|x - \mu| \leq 3\sigma) = 0.9973$$



- Namely, if a population of measurements has approximately a normal distribution the probability that a random selected observation falls within the intervals $(\mu - \sigma, \mu + \sigma)$, $(\mu - 2\sigma, \mu + 2\sigma)$, and $(\mu - 3\sigma, \mu + 3\sigma)$, is approximately 0.6826, 0.9544 and 0.9973, respectively.

□ Normal Distribution Application

Example:1 As a part of a study of Alzheimer's disease, reported data that are compatible with the hypothesis that brain weights of victims of the disease are normally distributed. From the reported data, we may compute a mean of 1076.80 grams and a standard deviation of 105.76 grams. If we assume that these results are applicable to all victims of Alzheimer's disease, find the probability that a randomly selected victim of the disease will have a brain that weighs less than 800 grams.



Solution:

R.V x 'Brain weights' follows a Normal distribution with $\mu=1076.80$ and $\sigma = 105.76$)

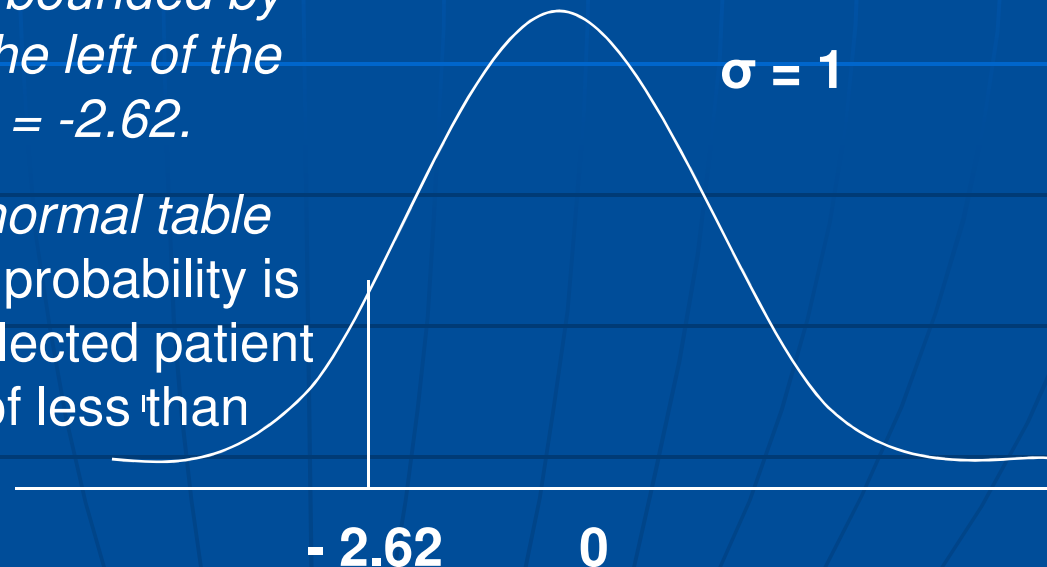
The Corresponding Standard Normal Variate

$$= z = \frac{x - 1076.80}{105.76}$$

$$z = \frac{x - \mu}{\sigma}$$

We have to find out $P(x < 800)$ i.e $P(z < -2.62)$. This is the area bounded by the curve, x axis and to the left of the perpendicular drawn at $z = -2.62$.

Thus from the standard normal table this prob., $p = .0044$. The probability is .0044 that a randomly selected patient will have a brain weight of less than 800 grams.



Example: 2

Suppose it is known that the heights of a certain population of individuals are approximately normally distributed with a mean of 70 inches and a standard deviation of 3 inches. What is the probability that a person picked at random from this group will be between 65 and 74 inches tall.

Solution:

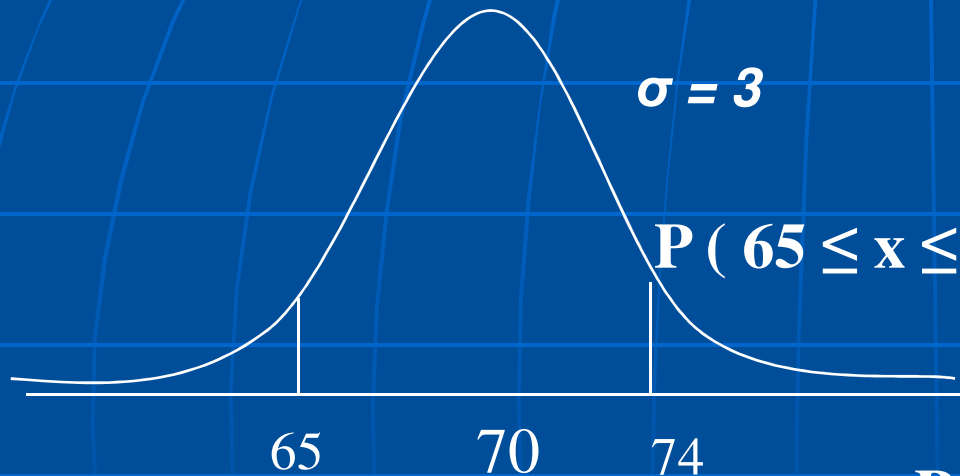
X (height) follows a Normal distribution with mean $(\mu) = 70$ inches and σ 3 inches.

For $x = 65$, Z we have

$$z = \frac{65-70}{3} = -1.76$$

For $x = 74$, Z we have

$$z = \frac{74-70}{3} = 1.33$$



$\sigma = 3$

$P(65 \leq x \leq 74)$

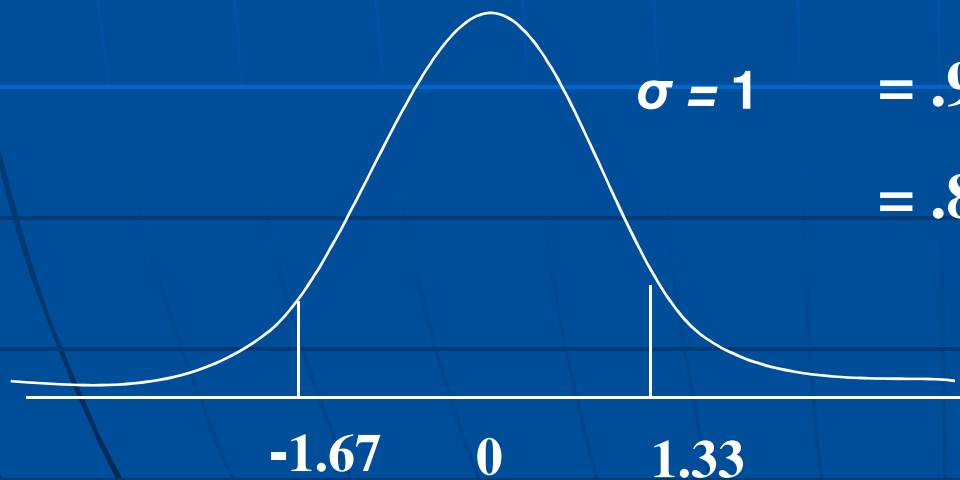
$$= P\left(\frac{65-70}{3} \leq z \leq \frac{74-70}{3}\right)$$

$$= P(-1.76 \leq z \leq 1.33)$$

$$= P(-\infty \leq z \leq 1.33) - P(-\infty \leq z \leq -1.67)$$

$$= .9082 - .0475$$

$$= .8607$$



$\sigma = 1$

-1.67

0

1.33

The required probability asked for in our original question, is 0.8607.

Example: 3

In a population of 10,000 of the people described in previous example how many would you expect to be 6 feet 5 inches tall or taller?

Solution:

we first find the probability that one person selected at random from the population would be 6 feet 5 inches tall or taller. That is,

$$\begin{aligned} P(x \geq 77) &= P\left(z \geq \frac{77-70}{3}\right) \\ &= P(z \geq 2.33) = 1 - .9901 = .0099 \end{aligned}$$

Out of 10,000 people we would expect $10,000 (.0099) = 99$ to be 6 feet 5 inches (77 inches tall or taller).

Exercise:

1. Given the standard normal distribution, find the area under the curve, above the z-axis between $z = -\infty$ and $z = 2$.
2. What is the probability that a z picked at random from the population of z's will have a value between -2.55 and + 2.55?
3. What proportion of z values are between -2.74 and 1.53?
4. Given the standard normal distribution, find $P(z \geq 2.71)$
5. Given the standard normal distribution, find $P(.84 \leq z \leq 2.45)$.

The background is a solid blue color with a subtle grid pattern of thin, lighter blue lines. The grid consists of both horizontal and vertical lines, creating a series of small squares across the entire page.

THANK YOU