Topic 6

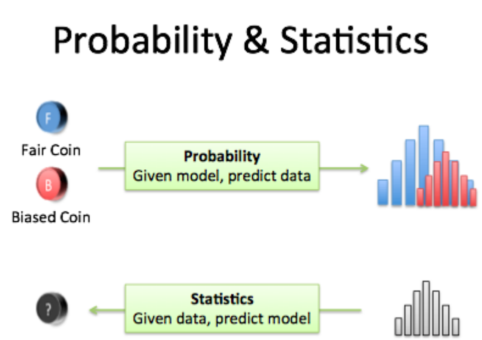
Probability Distributions

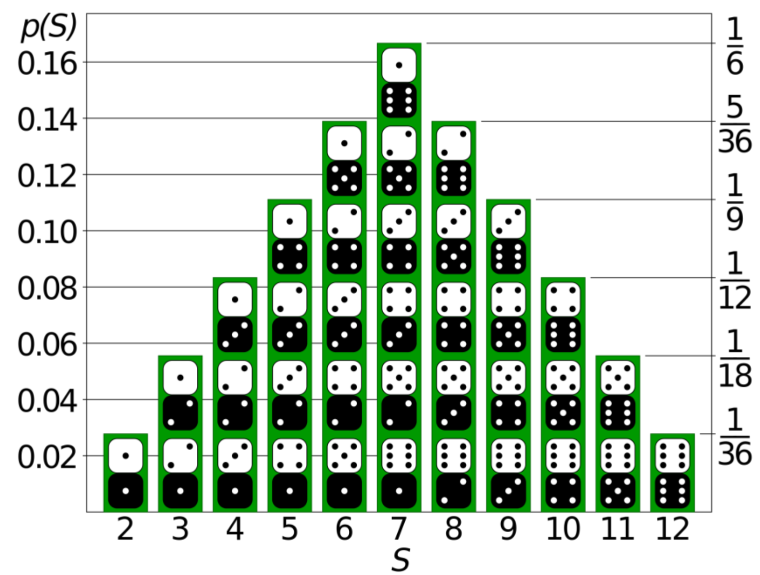
Contents:

6.1 Binomial Distribution

6.2 Poisson Distribution

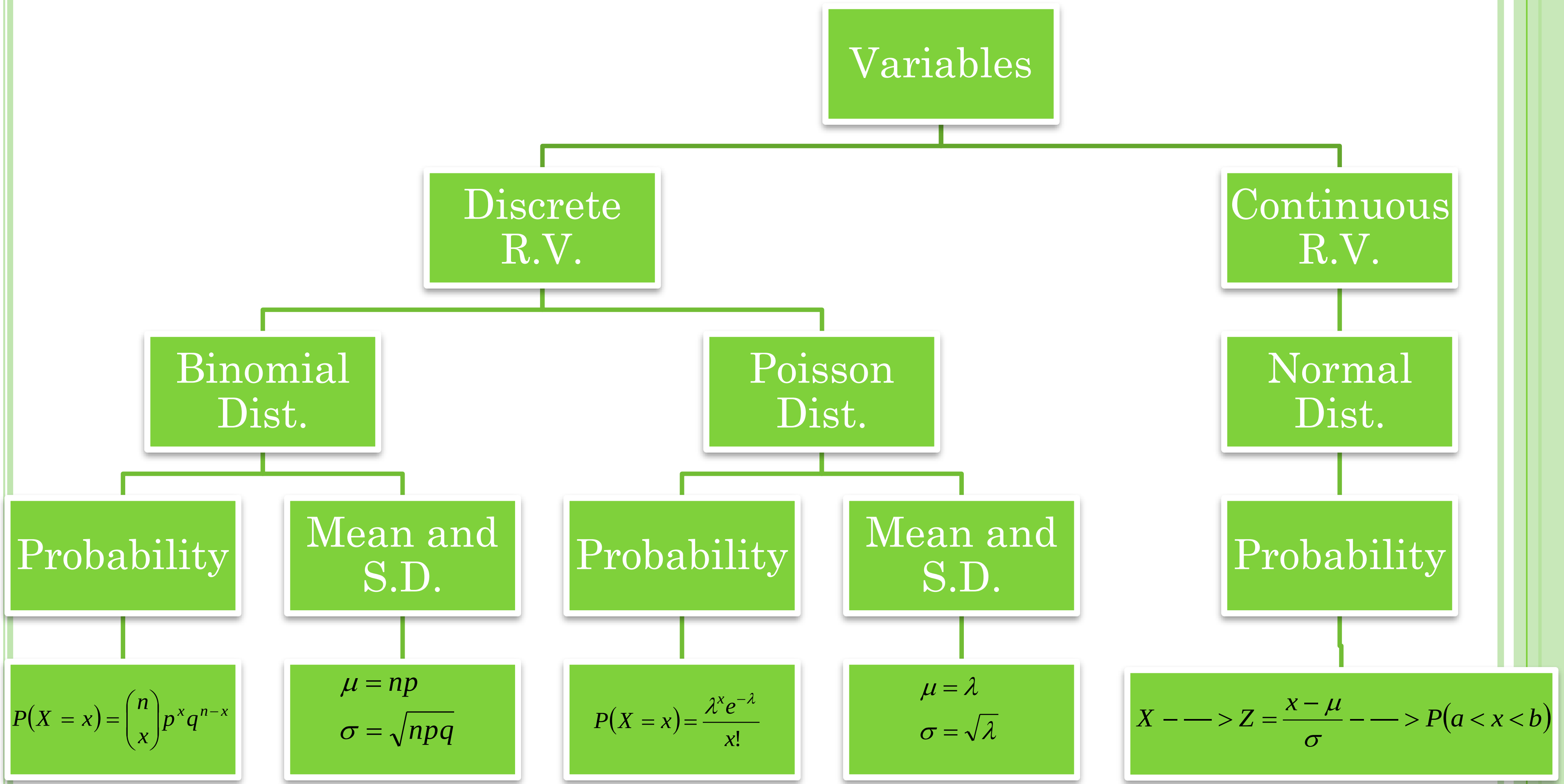
6.3 Normal Distribution





Subtopics:

- 6.1.1 The Bernoulli Process
- 6.1.2 Binomial Formula
- 6.1.3 Mean and Standard Deviation of Binomial Distribution
- 6.2.1 Poisson Formula
- 6.2.2 Mean and Standard Deviation of Poisson Distribution
- 6.3.1 Properties of a normal curve


Variables
Discrete R.V.
Continuous R.V.
Binomial Dist.
Poisson Dist.
Normal Dist.
Probability
Mean and S.D.
Probability
Mean and S.D.
Probability
$P(X=x)=\binom{n}{x}p^x q^{n-x}$
$\mu = np$
$\sigma = \sqrt{npq}$
$P(X=x)=\frac{\lambda^x e^{-\lambda}}{x!}$
$\mu = \lambda$
$\sigma = \sqrt{\lambda}$
$X ---> Z = \frac{x-\mu}{\sigma} ---> P(a<x<b)$

6.1 Binomial Distribution

6.1.1 The Bernoulli Process

- The binomial distribution is applied to find the probability that an outcomes will occur x times in n times in n performances of an experiment which satisfies the **Bernoulli Process**.

Four Conditions of Bernoulli Process

1. The experiment consists of ***n* repeated trials**.
2. Each trial results in two outcomes that may be classified as a success (p) or a failure (q).
3. The probability of success (p) and failure (q) remains constant from trial to trial.
4. The repeated trials are independent.

EXAMPLE

Consider the experiment consisting of 10 tosses of a coin. Determine whether or not it is a binomial experiment.

Solution:

- There are a total of 10 trials (tosses), and they are all identical. All 10 tosses are performed under identical condition.
- Each trial (toss) has only two possible outcomes: a head (success) and a tail (failure).
- $P(Head) = \frac{1}{2}$ and $P(Tail) = \frac{1}{2}$ remain the same for each toss. The sum of these two probabilities is 1.0.
- The trials (tosses) are independent. The outcome of one trial does not affect the outcome of another trial.

$\Rightarrow$ **Hence, it is a Binomial experiment**

EXAMPLE

The number of people with blood type O-negative based upon a simple random sample of size 10 is recorded. According to the Information Please Almanac, 6% of the human population is blood type O-negative. Determine whether or not it is a binomial experiment.

Answer: This is a binomial experiment

Note:

Four Conditions of Bernoulli Process

1) The experiment consists of ***n* repeated trials**.
2) Each trial results in **two outcomes** that may be classified as a ***success (p) or a failure (q).***
3) The probability of success (p) and failure (q) **remains constant** from trial to trial.
4) The repeated trials are **independent.**

EXAMPLE

A probability experiment in which three balls are drawn from a basket contains seven red balls and five blue balls without replacement and the number of red balls is recorded. Determine whether or not it is a binomial experiment.

Answer: No, it's not a binomial experiment

Note:

Four Conditions of Bernoulli Process

1) The experiment consists of ***n* repeated trials**.
2) Each trial results in **two outcomes** that may be classified as a ***success (p) or a failure (q).***
3) The probability of success (p) and failure (q) **remains constant** from trial to trial.
4) The repeated trials are **independent**.

6.1 Binomial Distribution

6.1.2 Binomial Formula

- For a binomial experiment, the probability of exactly x successes in n trials is given by the **binomial formula**

$$P(x) = \binom{n}{x} p^x q^{n-x}$$

where n = total number of trials

p = probability of success

$q = 1-p$ = probability of failure

x = number of successes in n trials

$n-x$ = number of failures in n trials

Example

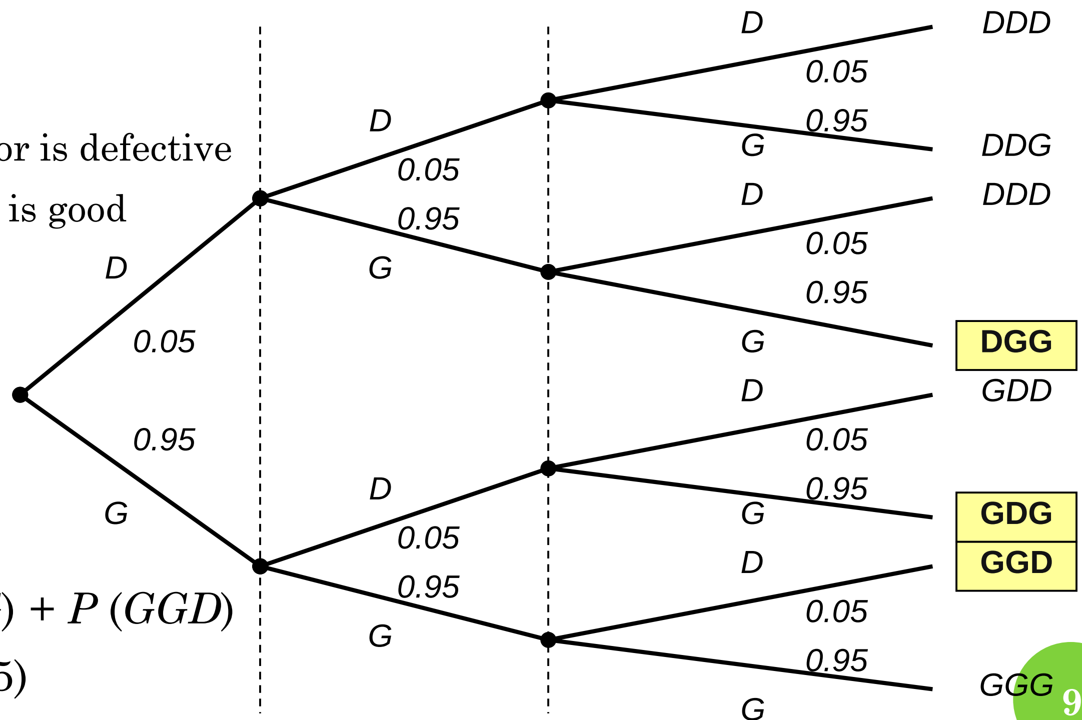
Five percent of all calculators manufactured by a large electronics company are defective. A quality control inspector randomly selects three calculators from the production line. What is the probability that exactly one of these calculators is defective?

Solution:

1st Method: Tree Diagram

Let D = a selected calculator is defective

G = a selected calculator is good



$P(D) = 0.05$ and $P(G) = 0.95$

$P(x = 1) = P(DGG) + P(GDG) + P(GGD)$

$= 3(0.05 \text{ x } 0.95 \text{ x } 0.95)$

$= 0.1354$

EXAMPLE

Five percent of all calculators manufactured by a large electronics company are defective. A quality control inspector randomly selects three calculators from the production line. What is the probability that exactly one of these calculators is defective?

Solution:

2nd Method: Binomial formula $P(x) = B(n, p) = \binom{n}{x} p^x q^{n-x}$

Let n = total number of trials = 3

p = probability of success = 0.05

q = 1 – p = 0.95

x = number of successes in n trials = 1

$$P(X = 1) = \binom{3}{1}(0.05^1)(0.95^{3-1}) = 0.1354$$

EXAMPLE

The probability that John hits the target is $\frac{1}{4}$. He fires 6 times.
Find the probability that he hits the target:

a) exactly 2 times

b) more than 4 times

c) at least once

Answer:

a) 0.2967

b) 0.0046

c) 0.8220

Example

One prominent physician claims that 70% of those with lung cancers are chain smokers. If his assertion is correct. Find the probability that:

a) of 10 such patients recently admitted to a hospital, fewer than half are chain smokers.
b) of 20 such patients recently admitted to a hospital, fewer than half are chain smokers.

Answer:

a) 0.0473
b) 0.0171

6.1. Binomial Distribution

6.1.3 Mean and Standard Deviation of Binomial Distribution

- Mean, $\mu = np$
- Standard deviation, $\sigma = \sqrt{npq}$

Example:
According to a study, 25% of CEO said that their favorite luxury car is Mercedes. Assume that this result holds true for the current population of all Malaysia CEOs. A sample of 40 CEOs is selected. Let x denotes the number of CEOs in this sample who hold this view. Find the mean and standard deviation of the probability distribution of x.

Answer:
Mean = 10
Standard deviation = 2.739

Example

One prominent physician claims that 70% of those with lung cancers are chain smokers. If his assertion is correct. Find the expected value and standard deviation of the probability distribution of x of which 15 such patients recently admitted to a hospital are chain smokers.

Answer:
Mean = 10.5
Standard deviation = 1.775

6.2 Poisson Distribution

Three Conditions of a Poisson Probability Distribution

1	• x is a discrete random variable.
2	• The events occur at random (do not follow any pattern).
3	• The events occur independently (consider with respect to an interval or region).
4	• Average rate does not change over the period of interest.

<u>Example of Poisson condition</u>

1) The number of accidents that occur on a given highway during a one-week period
2) The number of customers entering a grocery store during a one hour interval
3) The number of television sets sold at a department store during a given week.

6.2 Poisson Distribution

6.2.1 Poisson Formula

- The probability distribution of the Poisson random variable X, representing the number of events occurring in a **given time interval** or **specified region** is

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{where} \quad x = 0, 1, 2, \ldots$$

where λ = average/mean number of occurrences

x = actual number of occurrences

e ≅ 2.71828

Example

The mean number of bacteria per cm^3 of liquid is known to be 3. Assuming that the number of bacteria follows a Poisson distribution, find the probability that there will be

a) no bacteria in 1 cm^3 of liquid

b) 4 bacteria in 1ml of liquid

c) less than two bacteria in $2cm^3$ of liquid

d) in $\frac{1}{2}$ ml of liquid there will be more than 2 bacteria

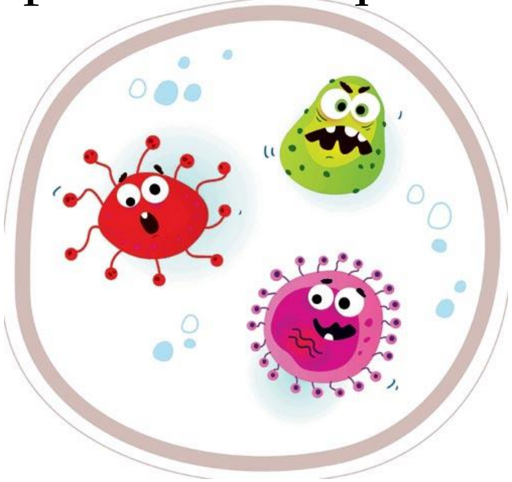
Solution:

Let's X = number of bacteria and $\lambda = 3$ bacteria per cm^3 of liquid

$\mu = \lambda = 3$ bacteria per cm^3 of liquid

a)

$$P(X = 0) = \frac{3^0 e^{-3}}{0!} = 0.0498$$



Example

The mean number of bacteria per cm^3 of liquid is known to be 3. Assuming that the number of bacteria follows a Poisson distribution, find the probability that there will be

a) no bacteria in 1 cm^3 of liquid

b) 4 bacteria in 1ml of liquid

c) less than two bacteria in 2cm^3 of liquid

d) in $\frac{1}{2}$ ml of liquid there will be more than 2 bacteria

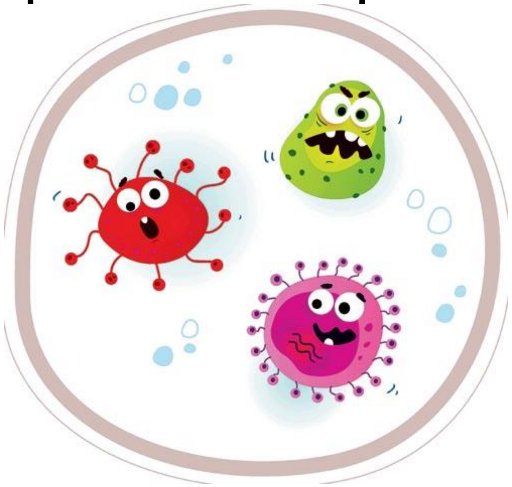
Solution:

Let's $X =$ number of bacteria and $\lambda = 3$ bacteria per cm^3 of liquid

$\mu = \lambda = 3$ bacteria per cm^3 of liquid

b)

$$P(X = 4) = \frac{3^4 e^{-3}}{4!} = 0.1680$$



EXAMPLE

The mean number of bacteria per cm^3 of liquid is known to be 3. Assuming that the number of bacteria follows a Poisson distribution, find the probability that there will be

a) no bacteria in 1 cm^3 of liquid

b) 4 bacteria in 1ml of liquid

c) less than two bacteria in 2cm^3 of liquid

d) in $\frac{1}{2}$ ml of liquid there will be more than 2 bacteria

Solution:

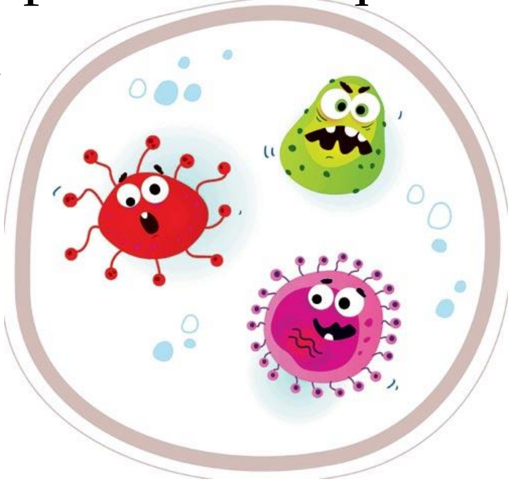
Let's X = number of bacteria and $\lambda = 3$ bacteria per cm^3 of liquid

$\mu = \lambda = 3 \text{ x } 2 = 6$ bacteria per cm^3 of liquid

c) $P(X < 2)$

$$= P(X = 0) + P(X = 1)$$

$$= \frac{6^0 e^{-6}}{0!} + \frac{6^1 e^{-6}}{1!} = 0.0174$$



EXAMPLE

The mean number of bacteria per cm^3 of liquid is known to be 3. Assuming that the number of bacteria follows a Poisson distribution, find the probability that there will be

d) in $\frac{1}{2}$ ml of liquid there will be more than 2 bacteria

Solution:

Let's $X =$ number of bacteria and $\lambda = 3$ bacteria per cm^3 of liquid

$$\mu = \lambda = 3 \times \frac{1}{2} = 1.5 \ per \quad \frac{1}{2} cm^3 \quad of \quad liquid$$

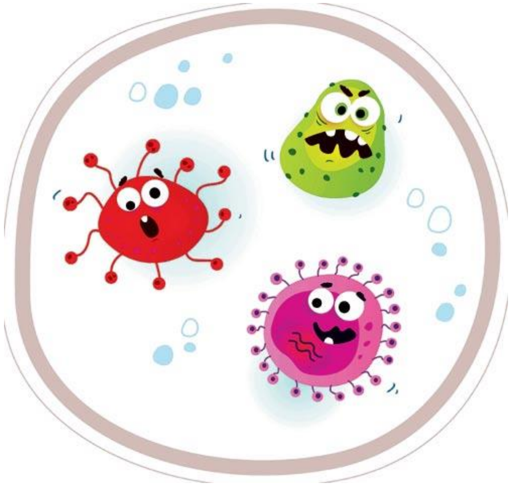
d) $P(X > 2)$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$= 1 - [\frac{1.5^0 e^{-1.5}}{0!} + \frac{1.5^1 e^{-1.5}}{1!} + \frac{1.5^2 e^{-1.5}}{2!}]$$

$$= 1 - 0.8088$$

$$= 0.1912$$



Try This

On average, two new accounts are opened per day at a bank. Using the table of Poisson distribution function, find the probability that on a given day, the number of new accounts opened at this bank will be

a) exactly 6

b) at most 3

c) at least 7

TRY THIS

A washing machine in a Laundromat breaks down an average of three times per month. Find the probability that during the next month this machine will have

a) exactly two breakdown

b) at most one breakdown

6.2 Poisson Distribution

6.2.2 Mean and Standard Deviation of Poisson Distribution

- Mean, $\mu = \lambda$
- Standard deviation, $\sigma = \sqrt{\lambda}$

Example:

The number of absentees per week in a probability class follows a Poisson distribution with standard deviation 2. Find the probability that

a) No absentees in the third week.
b) More than four absentees in first three weeks.
c) Less than 3 absentees in a particular week.

Solution:

a) 0.0183
b) 0.9924
c) 0.2381

Try This

The number of customers arriving per hour at a certain automobile service facility is assumed to follow a Poisson distribution with mean, $\lambda = 7$

a) What is the mean number of arrivals during a 2-hour period?

b) Compute the probability that more than 10 customers will arrive in a 2 hour period.

Try This

On average a certain intersection results in 3 traffic accidents per month. What is the probability that for any given month at this intersection

a) exactly five accidents will occur.

b) less than three accidents will occur.

c) at least two accidents will occur.

Try This

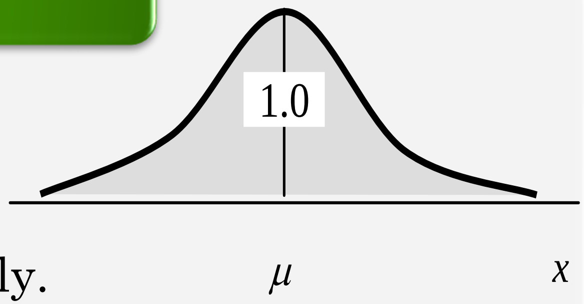
In the inspection of tin plate produced by a continuous electrolytic process, 0.2 imperfections are spotted on the average per minute. Find the probability of spotting:

a) one imperfection in three minutes.
b) at least two imperfections in five minutes.
c) at most one imperfection in 15 minutes.

6.3 Normal Distribution

6.3.1 Properties of a normal curve

- The total area under the curve is 1.0.
- The curve is symmetric about the mean μ.
- The two tails of the curve extend indefinitely.



- A continuous random variable X having the bell-shaped distribution is called a normal random variable.
- A continuous random variable x is said to have a normal distribution with parameters μ and σ, where $-\infty < \mu < \infty$ and $\sigma > 0$, if the probability distribution function of x is

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

AREA UNDER THE NORMAL DISTRIBUTION CURVE

- If x is continuous, then

$$P(a \le x \le b) = P(a < x < b) = P(a \le x < b) = P(a < x \le b)$$

- Therefore,

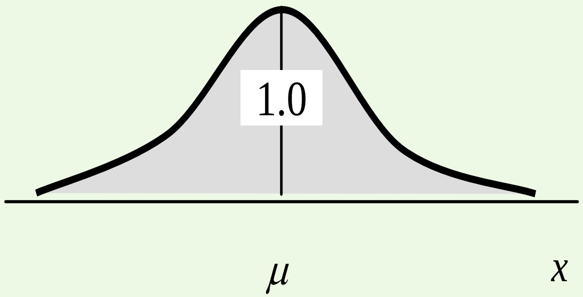
$$P(x_1 < X < x_2)$$

$$= \int_{x_1}^{x_2} f(x)dx$$

$$= \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Normal Distribution,

$X \sim N(\mu, \sigma)$



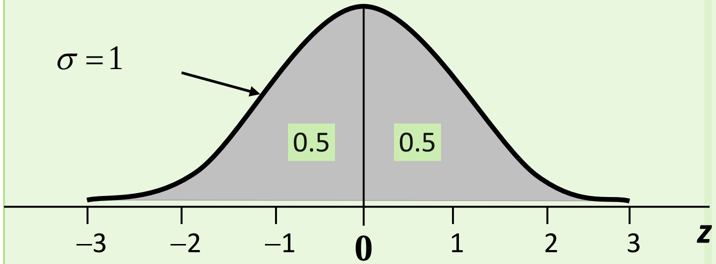
$$P(x_1 < X < x_2) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{x_1}^{x_2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Transformation

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

Standard Normal Distribution,

$Z \sim N(0, 1)$ where μ=0 and σ =1



Refer to statistical table

Methods for Obtaining area under a Standard Normal Curve

1) Find the area to the left of $Z = z_0$
2) Find the area to the right of $Z = z_0$
3) Find the area between $Z = z_0$ and $Z = z_1$

Example

Find the area under the standard normal curve:

a) between $z = 0$ and $z = 1.95$

b) area to the right of $z = 2.32$

c) area to the right of $z = -0.75$

d) area to the left of $z = -1.54$

e) between $z = -2.17$ and $z = 0$

f) between $z = -1.56$ and $z = 2.31$

Standard Normal Variable

For a normal random variable x, a particular value of x can be converted to its corresponding z value by using the formula

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

, where μ and σ are the mean and standard deviation of the normal distribution of x, respectively.

Step for Finding the Area under any Normal Curve / Finding the Probability

1. Convert the value of X to Z-score
2. Draw a standard normal curve with putting the z-values
3. Shade the desired area on the curve
4. Find the area under the standard normal curve.

EXAMPLE

Let x be a continuous random variable that is normally distribution with a mean of 25 and a standard deviation of 4. Find the are

a) between $x = 25$ and $x = 32$

b) between $x = 18$ and $x = 34$

FINDING THE *Z* AND *X* VALUES WHEN AN AREA UNDER THE NORMAL DISTRIBUTION CURVE IS KNOWN

Three important cautions:

1) Don't confuse z scores and areas
2) Choose the correct side of the graph
3) A z score must be negative whenever it is located to the left of the center line of zero.

<u>Example</u>

Find the value of h of a standard normal distribution such that

a) $P(Z > h) = 0.3015$
b) $P(h < Z < -0.18) = 0.4197$

EXAMPLE

1) Given a standard normal distribution, find the value of *k* such that

a) $P(Z > k) = 0.2946$

b) $P(Z < k) = 0.0427$

c) $P(-0.93 < Z < k) = 0.7235$

2) Given that X has a normal distribution with a mean of 20 and a variance of 4, find the value of k.

a) $P(X > k) = 0.2119$

b) $P(X < k) = 0.9678$

c) $P(x \leq k) = 0.00347$

EXAMPLE

A company pays its employees an average wage of RM15.90 an hour with a standard deviation of RM1.50. If the wages are approximately normal distributed,

a) What is the probability that the workers receive wages between RM13.75 and RM16.22 an hour inclusive?
b) What is the probability that the workers receive wages within RM3.90 from the average?
c) What is the probability that the workers receive wages less than the average wage by RM2.50 or more?

Example

A research scientist reports that mice will live an average of 40 months when their diets are sharply restricted and enriched with vitamins and proteins. Assuming that the lifetimes of such mice are normally distributed with a standard deviation of 6.3 months, find the probability that the given mice will live

a) more than 32 months

b) less than 28 months

c) between 37 and 49 months