Applied Statistics I
(IMT224β/AMT224β)

Department of Mathematics
University of Ruhuna

A.W.L. Pubudu Thilan

Statistical Applications with Probability Models

Statistical experiment

A statistical experiment is any process by which an observation or a measurement is made.

Statistical experiment
Examples

1. Measure the weight in kg of students in the university.
2. Measure the daily rainfall in inches.
3. Count the number of defective items in daily production of a company.
4. Measure the height of students in Science Faculty.

Statistical experiment
Features of a statistical experiment

- Each experiment is capable of being repeated indefinitely under essentially unchanged conditions.
- Although we are in general not able to state what a particular outcome will be, we are able to describe the set of all possible outcomes of the experiment.
- As the experiment is performed repeatedly, the individual outcomes seem to occur in a haphazard manner. However as the experiment is repeated a large number of times, a definite pattern or regularity appears.

Random variable

- A **variable** is a symbol (P, Q, x, y, etc.) that can take on any of a specified set of values.
- When the value of a variable is the outcome of a statistical experiment, that variable is a **random variable**.
- As opposed to other variables, a random variable conceptually does not have a single, fixed value.

Random variable
Cont....

- A random variable is a variable whose value is subject to variations due to chance.
- A random variable, X, represents a quantity being measured.
- It can take on a set of possible different values, each with an associated probability.

Random variable
Examples

1. X=the weight in kg of students in the university.
2. X=the amount of rain in each day in inches.
3. X=the number of defective items in daily production of a company.
4. X=the height of students in Science Faculty.

Discrete and continuous random variables

- When a random variable X can take on only countable values (such as 0, 1, 2, 3, . . .), then X is said to be a discrete random variable.
- When a random variable X can take on any value in an interval, then X is said to be a continuous random variable.

Which of the following random variables are discrete and which are continuous?

1. The time it takes a student to register for second semester in MIS.
2. The number of students in Applied Statistics lecture.
3. The height of students in Faculty of Science.
4. The air pressure in an automobile tire.
5. The number of eggs in a nest.

Probability distribution

- A probability distribution is an assignment of probabilities to specific values of a random variable (discrete) or to a range of values of a random variable (continuous).
- A probability distribution is a table or an equation that links each outcome of a statistical experiment with its probability of occurence.

Usually we use a capital letter to represent a random variable and a lower-case letter, to represent one of its values. For example,

- X represents the random variable X.
- $P(X)$ represents the probability of X.
- $P(X = x)$ refers to the probability that the random variable X is equal to a particular value, denoted by x.
- As an example, $P(X = 1)$ refers to the probability that the random variable X is equal to 1.

Probability distribution
Example

Consider a die rolling experiment. The table below, which associates each outcome with its probability, is an example of a probability distribution.

Outcome (X)	Probability $P(X)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Why do we need probability distribution models?

- We need to study probability so that we can calculate the chance that our sample leads us to the wrong conclusion about the population.
- To do that we need to model the process of taking the sample from the population.
- Since there are many different types of data and many different ways we might collect a sample of data we need lots of different probability models.

Probability distribution models

A **probability distribution model** is a mathematical representation of a random phenomenon. In here we mainly focus on following probability distribution models.

1. Binomial probability distribution model
2. Bernoulli probability distribution model
3. Poisson probability distribution model
4. Normal probability distribution model

Binomial probability distribution model

Binomial experiment

A **binomial experiment** (also known as a **Bernoulli trial**) is a statistical experiment that has the following properties:

1. The experiment consists of n repeated trials.
2. Each trial can result in just two possible outcomes. We call one of these outcomes a **success** and the other, a **failure**.
3. The probability of success, denoted by p, is the same on every trial.
4. The trials are independent; that is, the outcome on one trial does not affect the outcome on other trials.

An example for binomial experiment

Suppose you flip a coin 2 times and count the number of times the coin lands on heads. This is a binomial experiment because:

- The experiment consists of repeated trials. We flip a coin 2 times.
- Each trial can result in just two possible outcomes, heads or tails.
- The probability of success is constant (0.5) on every trial.
- The trials are independent; that is, getting heads on one trial does not affect whether we get heads on other trials.

Notations used in binomial probability distribution model

The following notations are used, when we talk about binomial probability.

- k: The number of successes.
- n: The number of trials.
- p: The probability of success on an individual trial.
- q: The probability of failure on an individual trial.
- $bin(k; n, p)$: Binomial probability, the probability that an n-trial binomial experiment results in exactly k successes, when the probability of success on an individual trial is p.

What is meant by binomial probability?

- A **binomial random variable** is the number of successes k in n repeated trials of a binomial experiment.
- The probability distribution of a binomial random variable is called a **binomial distribution**.
- The **binomial probability** refers to the probability that a binomial experiment results in exactly k successes.

Formula for binomial probability

Suppose a binomial experiment consists of n trials and results in k successes. If the probability of success on an individual trial is p, then the binomial probability is:

$$Pr(X = k) = bin(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}; \quad k = 0, 1, 2, ..., n.$$

Mean and variance of a binomial random variable

If $X \sim bin(n, p)$ (that is, X is a binomially distributed random variable), then the expected value of X is:

$$E[X] = np,$$

and the variance is

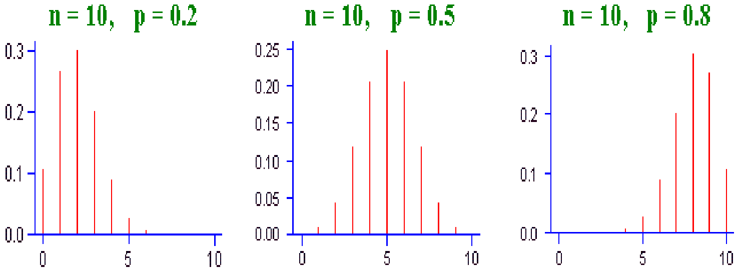
$$Var[X] = np(1 - p).$$

Shapes of binomial distributions

The skewness of a binomial distribution will depend upon the values of n and p. In general,

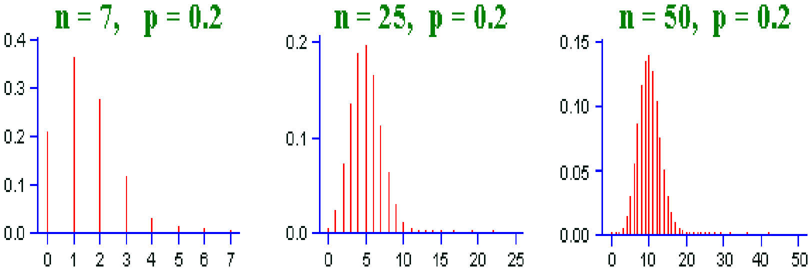
- If $p < 0.5$ the distribution will exhibit positive skew.
- If $p = 0.5$ the distribution will be symmetric.
- If $p > 0.5$ the distribution will exhibit negative skew.

Shapes of binomial distributions
With same n and different p values



Shapes of binomial distributions

With same p and different n values



Example 1

Suppose a die is tossed 5 times. What is the probability of getting outcome four exactly two times?

This is a binomial experiment in which the number of trials (n) is equal to 5, the number of successes (k) is equal to 2, and the probability of success (p) on a single trial is 1/6 or about 0.167. Therefore, the binomial probability is:

$$\begin{aligned} Pr(X=k) = bin(k;n,p) &= \binom{n}{k} p^k (1-p)^{n-k} \\ Pr(X=2) = bin(2;5,0.167) &= \binom{5}{2} (0.167)^2 (1-0.167)^{5-2} \\ &= \binom{5}{2} (0.167)^2 (0.833)^3 \\ &= 0.161 \end{aligned}$$

Example 2

The probability that a student is accepted to a prestigeous college is 0.3. If 5 students from the same school apply, what is the probability that at most 2 are accepted?

Example 2
Solution

$$
\begin{aligned}
\mathrm{bin}(k \leq 2; 5, 0.3) &= \mathrm{bin}(k = 0; 5, 0.3) + \mathrm{bin}(k = 1; 5, 0.3) \\
&+ \mathrm{bin}(k = 2; 5, 0.3) \\
&= \binom{5}{0}(0.3)^0(1-0.3)^{5-0} \\
&+ \binom{5}{1}(0.3)^1(1-0.3)^{5-1} \\
&+ \binom{5}{2}(0.3)^2(1-0.3)^{5-2} \\
&= 0.8369
\end{aligned}
$$

Example 3

A manufacturer of metal pistons finds that on the average, 12% of his pistons are rejected because they are either oversize or undersize. What is the probability that a batch of 10 pistons will contain,

(a) no more than 2 rejects?

(b) at least 2 rejects?

Example 3
Solution

(a) Let X = number of rejected pistons

In this case, ”success” means rejection!

Here, $n = 10, p = 0.12, q = 0.88$.

The probability of getting no more than 2 rejects is:

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

Example 3
Solution⇒Cont...

$$
\begin{aligned}
Pr(X=0) &= \binom{10}{0}(0.12)^0(0.88)^{10-0} = 0.2785 \\
Pr(X=1) &= \binom{10}{1}(0.12)^1(0.88)^{10-1} = 0.37977 \\
Pr(X=2) &= \binom{10}{2}(0.12)^2(0.88)^{10-2} = 0.23304
\end{aligned}
$$

So, the probability of getting no more than 2 rejects is:

$$
\begin{aligned}
Pr(X \leq 2) &= 0.2785 + 0.37977 + 0.23304 \\
&= 0.89131
\end{aligned}
$$

(b) We could work out all the cases for $X = 2, 3, 4, ..., 10$, but it is much easier to proceed as follows:

$$\begin{aligned}\text{Probability of}& \\ \text{at least 2 rejects} &= 1 - Pr(X \leq 1) \\ &= 1 - (Pr(X = 0) + Pr(X = 1)) \\ &= 1 - (0.2785 + 0.37977) \\ &= 0.34173\end{aligned}$$

Example 4

A certain type of missiles can attack a target with probability $p = 0.2$. Suppose we fire n missiles. If the probability of attacking the target at least by one of the missiles is at least 90%. Find n.

Example 4
Solution

X=Number of missiles hit the target

$$
\begin{aligned}
Pr(X=0) &= \binom{n}{0}(0.2)^0(0.8)^n \\
&= (0.8)^n
\end{aligned}
$$

The probability of attacking the target by at least one missile is

$$
\begin{aligned}
&= 1-(0.8)^n \\
1-(0.8)^n &\geq 0.9 \\
0.1 &\geq (0.8)^n \\
\log(0.1) &\leq n\log(0.8) \\
n &\geq \frac{\log 0.1}{\log 0.8} \\
n &\geq 10.31 \\
n &= 11
\end{aligned}
$$

Excercise

(a) Bits are sent over a communications channel in packets of 12. If the probability of a bit being corrupted over this channel is 0.1 and such errors are independent, what is the probability that no more than 2 bits in a packet are corrupted?

(b) If 6 packets are sent over the channel, what is the probability that at least one packet will contain 3 or more corrupted bits?

(c) Let X denote the number of packets containing 3 or more corrupted bits. What is the probability that X will exceed its mean by more than 2 standard deviations?

Distribution of summation of binomial random variables

$$
\begin{aligned}
\text{Let } X_i &\sim \text{bin}(n_i, p) \\
\text{Let } Y &= \sum_{i=1}^{k} X_i; \quad X_i\text{'s are independent} \\
\text{Then } Y &\sim \text{bin}\left(\sum_{i=1}^{k} n_i, p\right) \\
E(Y) &= \sum_{i=1}^{k} n_i p \\
Var(Y) &= \sum_{i=1}^{k} n_i p q
\end{aligned}
$$

Example

Consider the two pair of coins. Let

X=Number of heads in 100 tosses of the first coin

Y=Number of heads in 50 tosses of the second coin

Find the probability that $X + Y \leq 140$.

Example
Solution

$$
\begin{aligned}
X &\sim \text{bin}\left(100, \frac{1}{2}\right) \\
Y &\sim \text{bin}\left(50, \frac{1}{2}\right) \\
X + Y &\sim \text{bin}\left(150, \frac{1}{2}\right) \\
\Pr(X + Y \leq 140) &= \sum_{k=0}^{140} \binom{150}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{150-k} \quad \text{or} \\
\Pr(X + Y \leq 140) &= 1 - \Pr(X + Y > 140) \\
&= 1 - \sum_{k=141}^{150} \binom{150}{k} \left(\frac{1}{2}\right)^k \left(1 - \frac{1}{2}\right)^{150-k}
\end{aligned}
$$

Bernoulli probability distribution model

Bernoulli experiment

- A single experiment which can have one of two possible outcomes is called **Bernoulli experiment**.
- The two possible outcomes are "success" and "failure".
- The binomial experiment is an n times repeated Bernoulli trial.
- When a binomial trial occurs once we get a Bernoulli trial.

Examples for Bernoulli experiments

- Flipping a coin.
- Rolling a die, where a six is "success" and everything else a "failure".
- Either you pass an exam or you do not pass an exam.
- Either you get the job you applied for or you do not get the job.

Formula for Bernoulli probability

For a Bernoulli random variable probability distribution is given by:

$$Pr(X = k) = p^k(1-p)^{1-k} \quad k = 0, 1.$$

- Where $k = 0$ and $k = 1$ represent two outcomes.
- When $k = 0$ we get a failure with probability $1 - p = q$.
- When $k = 1$ we get a success with probability p.

Mean and variance of a Bernoulli random variable

If X is a random variable which has a Bernoulli distribution with success probability p, then it is denoted by,

$$X \sim \mathrm{Ber}(p).$$

Then mean of X is p and variance of X is pq. That is,

$$\begin{aligned} E(X) &= p \\ Var(X) &= pq. \end{aligned}$$

Example

Suppose the probability of passing Applied Statistics course unit is 0.75. Let the random variable X represents someone sitting the examination for the course unit. What is the probability of failing that student, what is the mean and variance of X?

Example
Solution

$$
\begin{aligned}
Pr(X = k) &= p^k(1-p)^{1-k} \quad k = 0, 1. \\
Pr(X = 0) &= 0.75^0(1-0.75)^{1-0} \\
Pr(X = 0) &= 0.25 \\
E(X) &= p = 0.75 \\
Var(X) &= pq = 0.75 \times 0.25 = 0.1875.
\end{aligned}
$$

Distribution of summation of Bernoulli random variables

Suppose $X_i \sim \mathrm{Ber}(\mathsf{p})$. Then if we consider the distribution of $\sum_{i=1}^{n} X_i$, it would be a binomial distribution as given below.

$$\sum_{i=1}^{n} X_i \quad \sim \quad \mathrm{bin}(n, p)$$

Example

Suppose that in each week a person buys a lottery ticket which gives him a chance of $1/100$ of a win.

(a) What is the chance of no wins in a week?

(b) What is the chance of 3 wins in the year (approximately)?

(a)

$$\begin{aligned} X &= \text{number of wins in a week, X=0, 1} \\ X &\sim \text{Ber}(0.01) \\ \Pr(X=0) &= (0.01)^0(1-0.01)^1 \\ &= 0.99 \end{aligned}$$

(b)

$$
\begin{aligned}
X &\sim \mathrm{Ber}(0.01) \\
1 \text{ year} &\Rightarrow 52 \text{ weeks} \\
\sum_{i=1}^{52} X_i &\sim \mathrm{bin}(52, 0.01) \\
\text{Let } Y &= \sum_{i=1}^{52} X_i \\
\Pr(Y = 3) &= \binom{52}{3}(0.01)^3(1-0.01)^{49} \\
&= 0.013
\end{aligned}
$$

Poisson probability distribution model

Poisson experiment

- The Poisson distribution was first derived in 1837 by the French mathematician *Simeon Denis Poisson*.
- A Poisson experiment examines the number of times an event occurs during a specified interval. The interval could be anything-a unit of time, length, volume, etc.
- The number of successes in two disjoint time intervals is independent.

Examples for Poisson experiments

- Failure of a machine in one month
- The number of flaws in a fibre optic cable
- Number of typing errors on a page
- The number of bacteria on a plate

Properties of a Poisson experiment

- The experiment results in outcomes that can be classified as successes or failures.
- The average number of successes (μ) that occurs in a specified region is known.
- The probability that a success will occur is proportional to the size of the region.
- The probability that a success will occur in an extremely small region is virtually zero.

Formula for Poisson probability

The probability distribution of a Poisson random variable, X representing the number of successes occurring in a given time interval or a specified region of space is given by the formula:

$$Pr(X = k) = \frac{e^{-\mu}\mu^k}{k!},$$

where

$$\begin{aligned} k &= 0, 1, 2, 3... \\ e &= 2.71828 \\ \mu &= \text{mean number of successes in the given time} \\ & \quad \text{interval or region of space.} \end{aligned}$$

Mean and variance of a Poisson random variable

- If μ is the average number of successes occurring in a given time interval or region in the Poisson distribution, then the mean and the variance of the Poisson distribution are both equal to μ.

$$\begin{aligned} X &\sim Pois(\mu) \\ E(X) &= \mu \\ Var(X) &= \mu. \end{aligned}$$

- In a Poisson distribution, only one parameter, μ is needed to determine the probability of an event.

Example 1

Suppose web site of the Department of Mathematics is visited by 12 people daily. What is the probability, that the web site will have 20 visitors a day?

Example 1
Solution

$$
\begin{aligned}
Pr(X = k) &= \frac{e^{-\mu}\mu^k}{k!} \\
k &= 20 \\
e &= 2.71828 \\
\mu &= 12 \\
Pr(X = 20) &= \frac{e^{-12}12^{20}}{20!} \\
&= 0.0097
\end{aligned}
$$

Example 2

A company makes electric motors. The probability an electric motor is defective is 0.01. What is the probability that a sample of 300 electric motors will contain exactly 5 defective motors?

Example 2
Solution

The average number of defectives in 300 motors is $\mu = np = 0.01 \times 300 = 3$.

The probability of getting 5 defectives is:

$$
\begin{aligned}
Pr(X = k) &= \frac{e^{-\mu}\mu^{k}}{k!} \\
k &= 5 \\
e &= 2.71828 \\
\mu &= 3 \\
Pr(X = 5) &= \frac{e^{-3}3^{5}}{5!} \\
&= 0.10082
\end{aligned}
$$

This problem looks similar to a binomial distribution problem, that we met in the last section.
If we do it using binomial distribution, with
$n = 300, k = 5, p = 0.01$ and $q = 0.99$, we get:

$$\begin{aligned}\Pr(X = 5) &= \binom{300}{5}(0.01)^5(0.99)^{295} \\ &= 0.10099\end{aligned}$$

We see that the result is very similar. We can use binomial distribution to approximate Poisson distribution (and vice-versa) under certain circumstances.

Example 3

If electricity power failures occur according to a Poisson distribution with an average of 3 failures every twenty weeks, calculate the probability that there will not be more than one failure during a particular week.

Example 3
Solution

Not more than one failure means we need to include the probabilities for zero failures plus one failure.

$$\begin{aligned}\Pr(X=0)+\Pr(X=1) &= \frac{e^{-0.15}(0.15)^0}{0!}+\frac{e^{-0.15}(0.15)^1}{1!}\\ &= 0.98981\end{aligned}$$

Distribution of summation of Poisson random variables

$$\begin{aligned} \text{let } X_i &\sim \text{Pois}(\mu_i) \\ \text{Let } Y &= \sum_{i=1}^{n} X_i \ ; \quad X_i\text{'s are independent} \\ \text{Then } Y &\sim \text{Pois}\left(\sum_{i=1}^{n} \mu_i\right). \end{aligned}$$

Example 1

A book contains 300 pages. The number of mistakes on a page is distributed with a Poisson distribution with average 2.

(a) Find the probability that number of mistakes in the book is more than 630.

(b) Find the probability that there is no mistakes in the book.

(c) Find the expected number of pages with no mistakes.

(a)

$$\begin{aligned}
X_i &= \text{number of mistakes on } i^{th} \text{ page} \\
X_i &= \text{Pois}(2) \\
\text{Let } Y &= \text{number of mistakes in the book} \\
Y &= \sum_{i=1}^{300} X_i \\
Y &\sim \text{Pois}\left(\sum_{i=1}^{300} \mu_i\right) \\
Y &\sim \text{Pois}(300 \times 2) \\
Y &\sim \text{Pois}(600) \\
\Pr(Y > 630) &= 1 - \Pr(Y \leq 630) = 1 - \sum_{k=0}^{630} \frac{e^{-600}(600)^k}{k!}
\end{aligned}$$

(b)

$$\Pr(Y=0) = \frac{e^{-600}600^0}{0!} = e^{-600} \simeq 0$$

(c)

$$\Pr(X=0) = \frac{e^{-2}2^0}{0!} = e^{-2} \simeq = 0.135$$

the expected number of pages with no mistakes

$$\begin{aligned} &= np \\ &= 300 \times 0.135 \\ &= 40.5 \\ &\simeq 41 \end{aligned}$$

Example 2

Let X and Y be two independent Poisson random variables with parameters 1 and 2 respectively.

(a) Find the probability that $X + Y > 2$.

(b) Find the probability that $X = 1$ and $Y = 2$.

Example 2
Solution

(a)

$$
\begin{aligned}
X &\sim \mathrm{Poiss}(1) \\
Y &\sim \mathrm{Poiss}(2) \\
X + Y &\sim \mathrm{Poiss}(3) \\
\Pr(X + Y > 2) &= 1 - \Pr(X + Y \leq 2) \\
&= 1 - \Pr(X + Y = 0) - \Pr(X + Y = 1) - \\
&\quad \Pr(X + Y = 2) \\
&= 1 - \frac{e^{-3}3^0}{0!} - \frac{e^{-3}3^1}{1!} - \frac{e^{-3}3^2}{2!} \\
&= 1 - e^{-3}\left(1 + 3 + \frac{3^2}{2!}\right)
\end{aligned}
$$

(b)

$$
\begin{aligned}
\Pr(X = 1 \text{ and } Y = 2) &= \Pr(X = 1) \,.\, \Pr(Y = 2) \\
&= \frac{e^{-1}1^1}{1!} \cdot \frac{e^{-2}2^2}{2!} \\
&= 0.0993
\end{aligned}
$$

Exercise 1

An average of 0.1 customers per minutes will arrive at a check point. Find the probability that less than two customers arrives at the check point within 4 consecutive minutes.

Exercise 2

A bank receives an average of 6 bad checks per day. Find the probability that it will receive 4 bad checks per day.

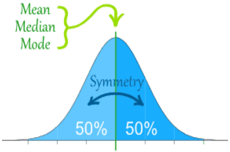
Normal probability distribution model

Introduction

- The normal distribution is the most important and most widely used distribution in statistics.
- It is also called the "Gaussian distribution" after the mathematician Karl Friedrich Gauss.
- The normal distribution is a continuous probability distribution, defined on the entire real line.

Properties of normal distribution

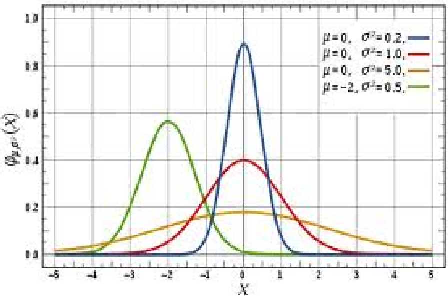
- All normal distributions are symmetric with relatively more values at the center of the distribution and relatively few in the tails.
- Normal distributions are symmetric around their mean.
- The mean, median, and mode of a normal distribution are equal.



Properties of normal distribution
Cont...

- The graph of the normal distribution depends on two factors, the mean (μ) and the standard deviation (σ).
- The mean of the distribution determines the location of the center of the graph, and the standard deviation determines the height and width of the graph.
- When the standard deviation is large, the curve is short and wide.
- When the standard deviation is small, the curve is tall and narrow.

Properties of normal distribution
Cont...



Normal probability density function

- Probability density function is a function that describes the relative likelihood for this random variable to take on a given value.
- The probability for the random variable to fall within a particular region is given by the integral of this variables density over the region.
- The probability density function is nonnegative everywhere, and its integral over the entire space is equal to one.

Normal probability density function

Cont...

The random variable X is said to be normally distributed if its probability density function $f_X(x)$ is

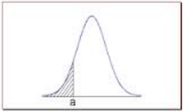
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2};$$

where μ and σ^2 are real and $\sigma > 0$.

Normal probability density function
Properties

The normal distribution is a continuous probability distribution and it has several implications for probability.

- The total area under the normal curve is equal to 1.
- The probability that a normal random variable X equals any particular value is 0.
- The probability that X is less than **a** equals the area under the normal curve bounded by **a** and minus infinity.
- The probability that X is greater than **a** equals the area under the normal curve bounded by **a** and plus infinity.



Mean and variance of a normal random variable

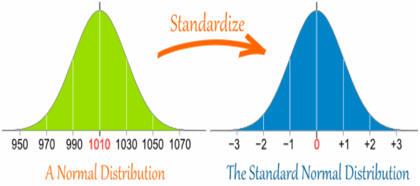
If X is a normally distributed random variable with mean μ and standard deviation σ, then it is denoted by,

$$\begin{aligned} X &\sim N(\mu, \sigma^2) \text{ and where} \\ E(X) &= \mu \\ Var(X) &= \sigma^2. \end{aligned}$$

Standard normal distribution

- The standard normal distribution is a normal distribution with a mean of 0 and a standard deviation of 1.
- Normal distributions can be transformed to standard normal distributions by the formula:

$$z = \frac{X - \mu}{\sigma}.$$



Mean and variance of a standard normal random variable

The standard normal distribution is sometimes called the z distribution.

$$\begin{aligned} z &\sim N(0,1), \\ E(z) &= 0, \\ Var(z) &= 1. \end{aligned}$$

The importance of standard normal probability table

- There are an infinite variety of normal distributions.
- So, it is not possible to print probability tables for every normal distribution.
- As a matter of fact, it is common practice to convert a normal to a standard normal and then use the standard normal table to find probabilities.

How to find probabilities using standard normal table

Find $P(Z < a)$

- The table shows the $P(Z < a)$.
- $P(Z < 1.13) = 0.8708$.
- $P(Z < 0) = 0.5000$.
- $P(Z < -1.20) = 0.1151$.

How to find probabilities using standard normal table
Find $P(Z > a)$

- The table shows the $P(Z < a)$.
- Then $P(Z > a) = 1 - P(Z < a)$.
- $P(Z > 3.00) = 1 - P(Z < 3.00) = 1 - 0.9987 = 0.0013$.

How to find probabilities using standard normal table

Find $P(a < Z < b)$.

- $P(a < Z < b) = P(Z < b) - P(Z < a)$.
- $P(-1.40 < Z < -1.20) = P(Z < -1.20) - P(Z < -1.40) = 0.1151 - 0.0808 = 0.0343$.

Example 1

Find the following probabilities:

(a) $P(Z > 1.06)$.

(b) $P(Z < -2.15)$.

(c) $P(1.06 < Z < 4.00)$.

(d) $P(-1.06 < Z < 4.00)$.

Example 2

It was found that the mean length of 100 parts produced by a lathe was 20.05 mm with a standard deviation of 0.02 mm. Find the probability that a part selected at random would have a length,

(a) between 20.03 mm and 20.08 mm.

(b) between 20.06 mm and 20.07 mm.

(c) less than 20.01 mm.

(d) greater than 20.09 mm.

(a) X=length of a part produced by the lathe

$$
\begin{aligned}
X &\sim N(20.05, 0.02^2) \\
&= Pr(20.03 < X < 20.08) \\
&= Pr\left(\frac{20.03-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{20.08-\mu}{\sigma}\right) \\
&= Pr\left(\frac{20.03-20.05}{0.02} < \frac{X-20.05}{0.02} < \frac{20.08-20.05}{0.02}\right) \\
&= Pr(-1 < z < 1.5) \\
&= Pr(z < 1.5) - Pr(z < -1) \\
&= Pr(z < 1.5) - Pr(z > 1) \\
&= Pr(z < 1.5) - [1 - Pr(z < 1)] \\
&= 0.9332 - 1 + 0.8413 \\
&= 0.7745
\end{aligned}
$$

(b)

$$
\begin{aligned}
&= Pr(20.06 < X < 20.07) \\
&= Pr\left(\frac{20.06-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{20.07-\mu}{\sigma}\right) \\
&= Pr\left(\frac{20.06-20.05}{0.02} < \frac{X-20.05}{0.02} < \frac{20.07-20.05}{0.02}\right) \\
&= Pr(0.5 < z < 1) \\
&= Pr(z < 1) - Pr(z < 0.5) \\
&= 0.1498
\end{aligned}
$$

(c)

$$\begin{aligned} &= Pr(X < 20.01) \\ &= Pr(\frac{X-\mu}{\sigma} < \frac{20.01-\mu}{\sigma}) \\ &= Pr(\frac{X-20.05}{0.02} < \frac{20.01-20.05}{0.02}) \\ &= Pr(z < -2) \\ &= 0.0228 \end{aligned}$$

(d) 20.09 is 2 s.d. above the mean, so the answer will be the same as (c). So, $Pr(X > 20.09) = 0.0228$.

Example 3

Suppose that the amount of cosmic radiation to which a person is expected while flying by a jet is random variable having a normal distribution with mean 4.35 (units) and standard deviation 0.59 (units). What is the probability that a person will be expected to have more than 5 (units) of cosmic radiation on such a flight. If the number of travelers in a flight is 100, what is the expected number of travelers exposed to cosmic radiation more than 5 (units).

Example 3
Solution

X=amount of cosmic radiation on a flight

$$\begin{aligned}
X &\sim N(4.35, 0.59^2) \\
Pr(X > 5) &= Pr\left(\frac{X-\mu}{\sigma} > \frac{5-\mu}{\sigma}\right) \\
&= Pr\left(\frac{X-4.35}{0.59} > \frac{5.00-4.35}{0.59}\right) \\
&= Pr(z > 1.102) \\
&= 1 - 0.8643 \\
&= 0.1357
\end{aligned}$$

Expected number of travelers exposed to cosmic radiation more than 5 (units) $= 100 \times 0.1357$.

Example 4

In school term test, a candidate fails if he obtains less than 40 marks out of the 100 marks and he has to obtain at least 76 marks in order to pass examination with a distinction. The top 10% of the students have received passes with distinction and the bottom 30% of the students have failed the examination. Assuming that the distribution of marks is normal, find the mean and the standard deviation of the marks.

Example 4
Solution

Let X be the marks of any student
Let

$$
\begin{aligned}
X &\sim N(\mu, \sigma^2) \\
Pr(X < 40) &= 30\% = 0.3 \\
Pr\left(\frac{X-\mu}{\sigma} < \frac{40-\mu}{\sigma}\right) &= 0.3 \\
Pr\left(z < \frac{40-\mu}{\sigma}\right) &= 0.3 \\
Pr(z < -0.52) &= 0.3 \\
\frac{40-\mu}{\sigma} &= -0.52 \qquad (1)
\end{aligned}
$$

Example 4

Solution⇒Cont...

$$\begin{aligned}
Pr(x \geq 76) &= 10\% = 0.1 \\
Pr\left(\frac{X-\mu}{\sigma} \geq \frac{76-\mu}{\sigma}\right) &= 0.1 \\
Pr\left(z \geq \frac{76-\mu}{\sigma}\right) &= 0.1 \\
\frac{76-\mu}{\sigma} &= 1.28 \qquad (2) \\
\text{From (1) and (2)} & \\
\frac{40-\mu}{76-\mu} &= \frac{-0.52}{1.28} \\
\mu &= 50.4 \\
\sigma &= 20
\end{aligned}$$

Distribution of summation of normal random variables

$$\begin{aligned} \text{Let } X_i &\sim N(\mu_i, \sigma_i^2);\ X_i'\text{s are independent.} \\ \text{Let } Y &= \sum_{i=1}^{n} X_i \\ \text{Then } \ Y &\sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right). \end{aligned}$$

Distribution of summation of normal random variables
Special case I

$$\begin{aligned} \text{Let } X_i &\sim N(\mu, \sigma^2);\ X_i'\text{s are independent.} \\ \text{Let } Y &= \sum_{i=1}^{n} X_i \\ \text{Then } Y &\sim N\left(n\mu, n\sigma^2\right). \end{aligned}$$

Distribution of summation of normal random variables

Special case II

$$\begin{aligned} \text{Let } X_i &\sim N(\mu, \sigma^2);\ X_i'\text{s are independent.} \\ \text{Let } \overline{X} &= \frac{\sum_{i=1}^{n} X_i}{n};\ \text{for independent } X_i \\ \text{Then } \overline{X} &\sim N\left(\mu, \frac{\sigma^2}{n}\right). \end{aligned}$$

Distribution of summation of normal random variables

Special case III

If X_1 and X_2 are independent normally distributed random variables with distribution,

$$\begin{aligned} X_1 &\sim N(\mu_1, \sigma_1^2) \\ X_2 &\sim N(\mu_2, \sigma_2^2), \\ \text{Then } X_1 - X_2 &\sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2). \end{aligned}$$

Example 1

A foot bridge of negligible weight is capable of carrying out a maximum weight of $400Kg$ at once. If 9 persons go by it at the same time, find the probability that the foot bridge is out of dangerous. You may assume that the weight of a person has a normal distribution with the mean of $50Kg$ and the standard deviation of $6Kg$.

Example 1
Solution

$$\begin{aligned}
\text{Let } X &= \text{weight of a person} \\
X &\sim N(50, 6^2) \\
\text{Let } Y &= \sum_{i=1}^{9} X_i \\
\text{Then } Y &\sim N(9 \times 50, 9 \times 6^2) \\
Y &\sim N(450, 9 \times 6^2)
\end{aligned}$$

Example 1
Solution⇒Cont...

$$
\begin{aligned}
Pr(Y < 400) &= Pr\left(\frac{Y - 450}{\sqrt{9 \times 6^2}} < \frac{400 - 450}{\sqrt{9 \times 6^2}}\right) \\
&= Pr\left(z < -\frac{50}{18}\right) \\
&= Pr(z < -2.778) \\
&= 1 - 0.9973 \\
&= 0.0027
\end{aligned}
$$

Example 2

The nicotine content of a certain brand of cigarettes is normally distributed with mean equal to $25mg$ and standard deviation equal to $4mg$. A random sample of 25 cigarettes is selected. What is the probability that the average nicotine content in the sample is at least $24mg$.

Example 2
Solution

$$\begin{aligned}
\text{Let } X &= \text{the nicotine content of the cigarette} \\
X &\sim N(25, 4^2) \\
\text{Let } \overline{X} &\sim \text{average/mean of nicotine in the sample} \\
\overline{X} &\sim N\left(\mu, \frac{\sigma^2}{n}\right) \\
\mu &= 25 \\
\frac{\sigma^2}{n} &= \frac{4^2}{25} = \frac{16}{25} \\
\overline{X} &\sim N\left(25, \frac{16}{25}\right)
\end{aligned}$$

$$
\begin{aligned}
Pr(\overline{X} \geq 24) &= Pr\left(\frac{\overline{X}-25}{4/5} > \frac{24-25}{4/5}\right) \\
&= Pr\left(z > -\frac{5}{4}\right) \\
&= Pr(z > -1.25) \\
&= 0.8944
\end{aligned}
$$

Example 3

Let X_1 and X_2 represent the life time of two types of bulbs A and B.

$$X_1 \sim N(1010, 5^2) \quad \text{and } X_2 \sim N(1020, 10^2)$$

What is the probability that A has higher life time that of B.

Example 3
Solution

$$
\begin{aligned}
Pr(X_1 > X_2) &= Pr(X_1 - X_2 > 0) \\
X_1 - X_2 &\sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2) \\
\text{Since } X_1 &\sim N(1010, 5^2) \quad \text{and } X_2 \sim N(1020, 10^2) \\
X_1 - X_2 &\sim N(-10, 125) \\
Pr(X_1 - X_2 > 0) &= Pr\left(\frac{(X_1 - X_2) - (-10)}{\sqrt{125}} > \frac{0 - (-10)}{\sqrt{125}}\right) \\
&= Pr\left(z > \frac{10}{\sqrt{125}}\right) \\
&= Pr(z > 0.89) \\
&= 1 - 0.8133 \\
&= 0.1867
\end{aligned}
$$

Probability approximations

Some possible probability approximations

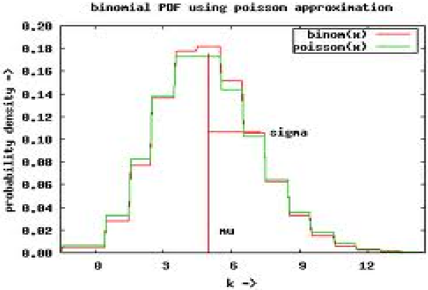
1. The Poisson approximation to binomial probabilities
2. Normal approximation to the binomial probabilities
3. Normal approximation to the Poisson probabilities

[1] The Poisson approximation to binomial probabilities

- The binomial distribution converges towards the Poisson distribution as the number of trials goes to infinity while the product np remains fixed.
- Therefore the Poisson distribution with parameter $\mu = np$ can be used as an approximation to $bin(n, p)$ of the binomial distribution if n is sufficiently large and p is sufficiently small.
- This approximation is good if $n \geq 20$ and $p \leq 0.05$ or if $n \geq 100$ and $np \leq 10$.

[1] The Poisson approximation to binomial probabilities

When $n = 50$; $p = 0.1$



Example

A certain disease occurs in 1.2% of the population. 100 people are selected at random from the population.

(a) Find the probability that no one has the disease.

(b) Find the probability that exactly 2 have the disease.

- This is a binomial experiment.
- The number of trials $n = 100$.
- $p = 0.012$.
- This satisfies $n \geq 20$ and $p \leq 0.05$.
- Therefore we can approximate binomial probabilities using Poisson distribution.
- For this approximation Poisson parameter, $\mu = np = 100 \times 0.012 = 1.2$.

Example
Solution⇒Cont...

(a)

$$\begin{aligned} Pr(X = k) &= \frac{e^{-\mu}\mu^k}{k!} \\ Pr(X = 0) &= \frac{e^{-1.2}(1.2)^0}{0!} \\ &\simeq 0.301 \end{aligned}$$

(b)

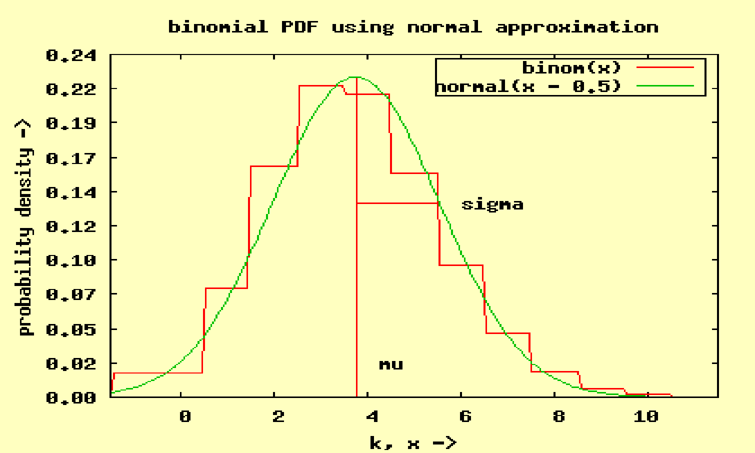
$$\begin{aligned} Pr(X = k) &= \frac{e^{-\mu}\mu^k}{k!} \\ Pr(X = 2) &= \frac{e^{-1.2}(1.2)^2}{2!} \\ &\simeq 0.217 \end{aligned}$$

[2] Normal approximation to the binomial distribution
continuity adjustment

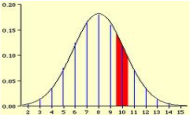
- A discrete random variable can take on only specified values while a continuous random variable can take on any values within a continuum or interval around those specified values.
- Hence, when using the normal distribution to approximate the binomial distributions, a correction for continuity adjustment is needed.

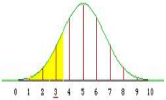
[2] Normal approximation to the binomial distribution

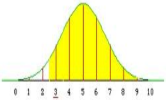
When $n = 25$; $p = 0.15$



- A continuous distribution (such as the normal), the probability of obtaining a particular value of a random variable is zero.
- When the normal distribution is used to approximate a discrete distribution, a correction for continuity adjustment can be employed so that the probability of a specific value of the discrete distribution can be approximated.







[2] Normal approximation to the binomial probabilities

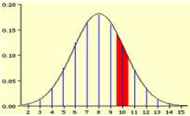
Why this type of approximation is needed?

- The larger the number of observations in the sample, it is difficult to compute the exact probabilities of success by use of binomial formula.
- Fortunately, though, whenever the sample size is large, the normal distribution can be used to approximate the exact probabilities of success.

- The binomial distribution is symmetric (like the normal distribution) whenever $p = 0.5$.
- When $p \neq 0.5$ the binomial distribution will not be symmetric.
- However, the closer p is to 0.5 and the larger the number of sample observations n, the more symmetric the distribution becomes.

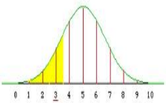
- The normal distribution provides a good approximation to the binomial when n is large and p is close to 0.5.
- However when n is not fairly large and p differs from 1/2 (but p is not close to zero or one) the normal approximation can be used when both np and $n(1-p)$ are greater than 5.

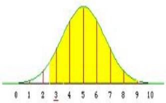
The binomial probability $Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ can be approximated by the normal probability $Pr(k - 0.5 \leq X \leq k + 0.5)$ considering the continuity correction.



[2] Normal approximation to the binomial probabilities
For one sided binomial probabilities

- $Pr(X \leq k)_{\text{bin}} \simeq Pr(X \leq k + 0.5)_{\text{normal}}$
- $Pr(X \geq k)_{\text{bin}} \simeq Pr(X \geq k - 0.5)_{\text{normal}}$
- $Pr(X < k)_{\text{bin}} = Pr(X \leq k - 1)_{\text{bin}} \simeq Pr(X \leq k - 0.5)_{\text{normal}}$
- $Pr(X > k)_{\text{bin}} = Pr(X \geq k + 1)_{\text{bin}} \simeq Pr(X \geq k + 0.5)_{\text{normal}}$





[2] Normal approximation to the binomial probabilities

For two sided binomial probabilities

- $Pr(a < X < b)_{\text{bin}} \simeq Pr(a + 0.5 \leq X \leq b - 0.5)_{\text{normal}}$
- $Pr(a \leq X \leq b)_{\text{bin}} \simeq Pr(a - 0.5 \leq X \leq b + 0.5)_{\text{normal}}$

Example

In a production process 10% units produced are defective. Let X be the number of defectives found in a sample of 100 units selected at random.

(a) Find $Pr(X \leq 13)$.

(b) Find $Pr(X > 7)$.

(c) Find $Pr(5 \leq X < 11)$.

(a)

$$\begin{aligned} X &- \text{number of defectives} \\ X &\sim \text{bin}(100, 0.1) \\ Pr(X \leq 13) &= \sum_{k=0}^{13} \binom{100}{k} (0.1)^k (0.9)^{100-k} \end{aligned}$$

It is very hard to find the value of the right hand expression. But in here:

$$\begin{aligned} np &= 100 \times 0.1 = 10 > 5 \\ n(1-p) &= 100 \times 0.9 = 90 > 5 \end{aligned}$$

Since np and $n(1-p)$ both greater than 5, we can use normal approximation to the binomial distribution.

Example

Solution⇒Cont...

$$
\begin{aligned}
Pr(X \leq 13)_{\text{bin}} &\simeq Pr(X \leq 13.5)_{\text{normal}} \\
\mu &= np = 10 \\
\sigma &= \sqrt{npq} \\
&= 3 \\
Pr(X \leq 13.5)_{\text{normal}} &= Pr\left(\frac{\overline{x} - 10}{3} \leq \frac{13.5 - 10}{3}\right) \\
&= Pr\left(z \leq \frac{3.5}{3}\right) \\
&= Pr(z \leq 1.167) \\
&= 0.8790
\end{aligned}
$$

Example
Solution⇒Cont...

(b)

$$\begin{aligned} Pr(X > 7)_{\text{bin}} = Pr(X \geq 8)_{\text{bin}} &= Pr(X \geq 7.5)_{\text{normal}} \\ &= Pr\left(\frac{\overline{x} - 10}{3} \geq \frac{7.5 - 10}{3}\right) \\ &= Pr(z \geq -0.833) \\ &= 0.7967 \end{aligned}$$

(c)

$$Pr(5 \leq X < 11)_{\text{bin}} = \sum_{k=5}^{10} \binom{100}{k} (0.1)^k (0.9)^{100-k}$$

Using the normal approximation to the binomial distribution

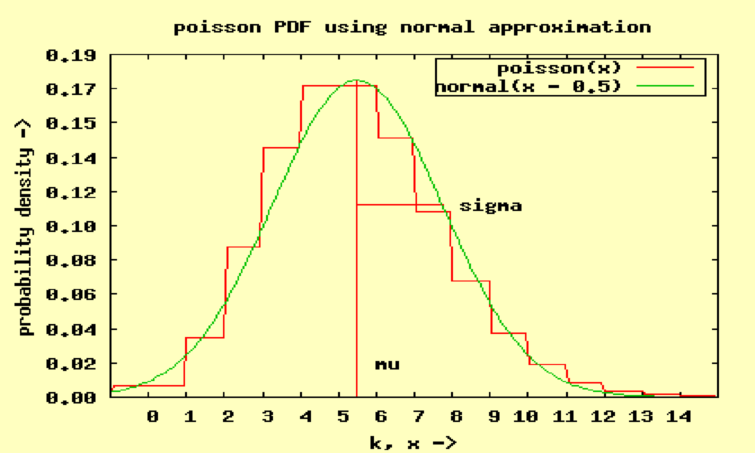
$$\begin{aligned}
Pr(5 \leq X < 11)_{\text{bin}} &= Pr(5 \leq X \leq 10)_{\text{bin}} \\
&= Pr(4.5 \leq X \leq 10.5)_{\text{normal}} \\
&= Pr\left(\frac{4.5-10}{3} \leq \frac{\overline{x}-10}{3} \leq \frac{10.5-10}{3}\right) \\
&= Pr(-1.83 \leq z \leq 0.167) \\
&= 0.5675 - (1 - 0.9664) \\
&= 0.5339
\end{aligned}$$

[3] Normal approximation to the Poisson probabilities

The normal distribution can also be used to approximate the Poisson distribution whenever the parameter μ, the expected number of successes, equals or exceeds 5.

[3] Normal approximation to the Poisson probabilities

Normal approximation when $\mu = 5$



[3] Normal approximation to the Poisson probabilities

Transformation equation

Since the value of the mean and the variance of a Poisson distribution are the same,

$$\begin{aligned} X &\sim Pois(\mu) \\ Var(X) &= E(X) \\ \sigma^2 &= \mu \\ \sigma &= \sqrt{\mu}. \end{aligned}$$

Substituting into the transformation equation,

$$z = \frac{X - \mu}{\sigma}$$
$$z = \frac{X - \mu}{\sqrt{\mu}}$$

So that, for large enough μ, the random variable z is approximately normally distributed.

Example

Suppose that at a certain automobile plant the average number of work stoppages per day due to equipment problems during the production process is 12.0. What is the approximate probability of having 15 or fewer work stoppages due to equipment problems on any given day?

In here $\mu = 12$ and it satisfies the condition that $\mu \geq 5$. So we can approximate this Poisson probability using the normal distribution.

$$\begin{aligned}
Pr(X \leq 15)_{\text{Pois}} &\simeq Pr(X \leq 15.5)_{\text{normal}} \\
&= Pr\left(\frac{X-\mu}{\sqrt{\mu}} \leq \frac{15.5-\mu}{\sqrt{\mu}}\right) \\
&= Pr\left(z \leq \frac{15.5-12}{\sqrt{12}}\right) \\
&= Pr(z \leq 1.01) \\
&= 0.8438
\end{aligned}$$

Example
Solution⇒Cont...

Therefore, the approximate probability of having 15 or fewer work stoppages due to equipment problems on any given day is 0.8438. This approximation compares quite favorably to the exact Poisson probability, 0.8445.

Pass paper 2011 (4)

(a) A certain type of storage battery lasts, on average, 3.0 years, with standard deviation of 0.5 year. Assuming that the battery lives are normally distributed, find the probability that a given battery will last less than 2.3 years.

(b) The LDL cholesterol level is normally distributed with mean 4.8 and standard deviation 0.6. A person has moderate risk if, the difference between his cholesterol level and the mean is more than 1 standard deviations but less than 2 standard deviations. What proportion of the population has moderate risk according to this criterion?

(c) In an industrial process the diameter of a ball bearing is an important component part. The buyer sets specifications on the diameter to be 3.0 ± 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 3.0 and standard deviation $\sigma = 0.005$. On the average, how many manufactured ball bearings will be scrapped?

Pass paper 2011 (4)
Solution

(a) X-Life time of a battery in years

$$
\begin{aligned}
X &\sim N(3, 0.5^2) \\
&= P\left(X < 2.3\right) \\
&= P\left(\frac{X-\mu}{\sigma} < \frac{2.3-\mu}{\sigma}\right) \\
&= P\left(\frac{X-3}{0.5} < \frac{2.3-3}{0.5}\right) \\
&= P\left(z < \frac{2.3-3}{0.5}\right) \\
&= P\left(z < -1.4\right)
\end{aligned}
$$

The probability that a given battery will last less than 2.3 is 0.0808.

(b) X- LDL cholesterol level

$$\begin{aligned} X &\sim N(4.8, 0.6^2) \\ &= P\left(\sigma < X - \mu < 2\sigma\right) \\ &= P\left(1 < \frac{X-\mu}{\sigma} < 2\right) \\ &= P\left(1 < z < 2\right) \\ &= p(z < 2) - p(z < 1) \\ &= 0.9772 - 0.8413 \\ &= 0.1359 \end{aligned}$$

Thus, 13.6% of the population have moderate risk.

(c) X-The diameter of a ball bearing

$$\begin{aligned}
X &\sim N(3.0, 0.005^2) \\
&= P(2.99 < X < 3.01) \\
&= P\left(\frac{2.99 - 3.00}{0.005} < \frac{X - 3.00}{0.005} < \frac{3.01 - 3.00}{0.005}\right) \\
&= P(-2.0 < z < 2.0) \\
&= P(z < 2.0) - P(z < -2.0) \\
&= 0.9544
\end{aligned}$$

As a result, it is anticipated that on the average, 4.56% of manufactured ball bearing will be scrapped.

(a) The probability of a flight arriving on time is determined to be 0.9. On four different occasions a person is taking flights. Find the probability that

(i) the person arrive on time on all four flights.

(ii) the person arrive on time on at least one occasion.

Pass paper 2011 (5)
Cont...

(b) It is known that 3% of the transistors from a production line are defective. If a random sample of 100 transistors is taken from this production line, use the Poisson approximation to estimate the probability that the sample contains,

   (i) exactly 2 defective transistors.
   (ii) at least 2 defective transistors.

(c) The probability that a patient recovers from a rare blood disease is 0.4. If 150 people are known to have infected this disease, what is the probability that less than 50 survive?

Pass paper 2011 (5)
Solution

(a) Let X be the number of times a flight is on time.

$$X \sim bin(4, 0.9)$$
$$Pr(X=k) = bin(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k};\ k = 0, 1, 2, ..., n.$$

(i)

$$\begin{aligned} Pr(X=4) = bin(4; 4, 0.9) &= \binom{4}{4} (0.9)^4 (1-0.9)^{4-4} \\ &= 0.656 \end{aligned}$$

Pass paper 2011 (5)
Solution

(ii)

$$
\begin{aligned}
Pr(X \geq 1) &= 1 - P(X = 0) \\
&= 1 - \binom{4}{0}(0.9)^0(1-0.9)^{4-0} \\
&= 1 - 0.1^4 \\
&= 0.9999
\end{aligned}
$$

(b) This is a binomial problem. Let X be the number of defective transistors.

$$\begin{aligned} p &= 0.03 \\ n &= 100 \end{aligned}$$

This satisfies $n \geq 20$ and $p \leq 0.05$. Therefore we can approximate binomial probabilities using Poisson distribution.

$$\begin{aligned} \mu &= np = 3 \\ X &\sim Pois(3) \\ P(X = k) &= \frac{e^{-\mu}(\mu)^k}{k!} \end{aligned}$$

Pass paper 2011 (5)
Solution

(i)

$$\begin{aligned} P(X=2) &= \frac{e^{-3}(3)^2}{2!} \\ &= \frac{(0.3679)^3 3^2}{2} \\ &= 0.224 \end{aligned}$$

(ii)

$$\begin{aligned} P(X \geq 2) &= 1 - [P(X=0) + P(X=1)] \\ &= 1 - \left[\frac{e^{-3}(3)^0}{0!} + \frac{e^{-3}(3)^1}{1!}\right] \\ &= 1 - e^{-3}[1+3] \\ &= 1 - 0.1992 \\ &= 0.8008 \end{aligned}$$

(c) Let X be the number of patient recovers from the rare blood disease.

$$\begin{aligned} X &\sim bin(150, 0.4) \\ np &= 150 \times 0.4 = 60 \\ n(1-p) &= 150 \times 0.6 = 90 \end{aligned}$$

Since both np and $n(1-p)$ greater than 5, we can use normal approximation to the binomial distribution.

$$\begin{aligned}
\mu &= np = 60 \\
\sigma &= \sqrt{npq} = \sqrt{150 \times 0.4 \times 0.6} = \sqrt{36} = 6 \\
P(X < 50)_{bin} &= P(X \leq 49.5)_{normal} \\
&= P\left(\frac{X - 60}{6} \leq \frac{49.5 - 60}{6}\right) \\
&= P(z \leq -1.75) \\
&= P(z \geq 1.75) \\
&= 1 - P(z < 1.75) \\
&= 1 - 0.9599 \\
&= 0.0401
\end{aligned}$$

(a) The number of flaws in a fibre optic cable follows a Poisson distribution. The average number of flaws in 50m of cable is 1.2. What is the probability of

(i) exactly three flaws in 150m of cable?
(ii) at least two flaws in 100m of cable?
(iii) exactly one flaw in the first 50m of cable and exactly one flaw in the second 50m of cable?

(b) The number of calls per hour received by an office switchboard follows a Poisson distribution with parameter 36. Using normal approximation to the Poisson distribution, find the probability that, in one hour,

(i) there are more than 33 calls.
(ii) there are between 25 and 28 calls.

(a) (i) Mean number of flaws in 150m of cable is 3.6. Let X be the number of flaws in 150m of cable. So the probability of exactly three flaws in 150m of cable:

$$
\begin{aligned}
X &\sim Pois(3.6) \\
P(X=k) &= \frac{e^{-\mu}(\mu)^k}{k!} \\
P(X=3) &= \frac{e^{-3.6}(3.6)^3}{3!} \\
&= \frac{(e^{-1.2})^3(3.6)^3}{3!} \\
&= \frac{(0.3012)^3(3.6)^3}{3!} \\
&= 0.212
\end{aligned}
$$

(ii) Mean number of flaws in 100m of cable is 2.4. Let Y be the number of flaws in 100m of cable.

$$
\begin{aligned}
Y &\sim Pois(2.4) \\
P(Y = k) &= \frac{e^{-\mu}(\mu)^k}{k!} \\
P(Y \geq 2) &= 1 - [P(Y = 0) + P(Y = 1)] \\
&= 1 - \left[\frac{e^{-2.4}(2.4)^0}{0!} + \frac{e^{-2.4}(2.4)^1}{1!}\right] \\
&= 1 - (0.091 + 0.218) = 0.691
\end{aligned}
$$

(iii) Now let L denote the number of flaws in a 50m section of cable. Then we know that:

$$\begin{aligned} L &\sim Pois(1.2) \\ P(L=k) &= \frac{e^{-\mu}(\mu)^k}{k!} \\ P(L=1) &= \frac{e^{-1.2}(1.2)^1}{1!} \\ &= 0.361 \end{aligned}$$

As L follows a Poisson distribution, the occurrence of flaws in the first and second 50m of cable are independent. Thus the probability of exactly one flaw in the first 50m and exactly one flaw in the second 50m is $(0.361)(0.361) = 0.13$.

(b) Let X be the number of calls per hour received. In here $\mu = 36$ and it satisfies the condition that $\mu \geq 5$. So we can approximate these probabilities using normal distribution.

(i)

$$
\begin{aligned}
X &\sim Pois(36) \\
P(X > 33)_{Pois} &\simeq P(X \geq 33.5)_{normal} \\
&= P\left(\frac{X-\mu}{\sqrt{\mu}} \geq \frac{33.5-\mu}{\sqrt{\mu}}\right) \\
&= P\left(\frac{X-36}{\sqrt{36}} \geq \frac{33.5-36}{\sqrt{36}}\right) \\
&= P(z \geq -0.417) \\
&= P(z \leq 0.417) \\
&= 0.6628
\end{aligned}
$$

(ii)

$$
\begin{aligned}
X &\sim Pois(36) \\
P(25 \leq X \leq 28)_{Pois} &\simeq P(24.5 \leq X \leq 28.5)_{normal} \\
&= P\left(\frac{24.5-\mu}{\sqrt{\mu}} \leq \frac{X-\mu}{\sqrt{\mu}} \leq \frac{28.5-\mu}{\sqrt{\mu}}\right) \\
&= P\left(\frac{24.5-36}{\sqrt{36}} \leq \frac{X-36}{\sqrt{36}} \leq \frac{28.5-36}{\sqrt{36}}\right) \\
&= P(-1.92 \leq z \leq -1.25) \\
&= P(1.25 \leq z \leq 1.92) \\
&= 0.9726 - 0.8944 \\
&= 0.0782
\end{aligned}
$$

Thank You