0.1 Dice Questions

Suppose a pair of fair dice is thrown.

(a) What is the probability of getting a sum of 9 from two throws of a dice

Find the probability that the sum is 10 or greater if

(b) a 5 appears on the first die,

(c) a 5 appears on at least one of the dice.

Tickets numbered 1 to 20 are mixed up and then a ticket is drawn at random. What is the probability that the ticket drawn has a number which is a multiple of 3 or 5

Chapter 1

Mathematical Fundamentals

The factorial function

The factorial function (symbol: !) just means to multiply a series of descending natural numbers. Examples:

- $4! = 4 \times 3 \times 2 \times 1 = 24$
- $7! = 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 5,040$
- $1! = 1$
- $0! = 1$

Importantly

$$n! = n \times (n-1)! = n \times (n-1) \times (n-2)!$$

For Example

$$6! = 6 \times 5! = 6 \times 5 \times 4!$$

- $\binom{6}{2} = 15$
- $\binom{5}{2} = 10$
- $\binom{4}{0} = 1$
- $\binom{4}{3} = 4$
- factorials

$$n! = (n) \times (n-1) \times (n-2) \times \ldots \times 1$$

  - $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$
  - $3! = 3 \times 2 \times 1$

- Zero factorial

$$0! = 1$$

9B.1 Permutation

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}$$

$$\binom{6}{3} = \frac{6!}{(6-3)!3!} = \frac{6!}{3! \times 3!}$$

$$\frac{6!}{3! \times 3!} = \frac{6 \times 5 \times 4 \times 3!}{3! \times 3!} = \frac{120}{6} = 120$$

- $\binom{6}{2} = 15$
- $\binom{5}{2} = 10$
- $\binom{4}{0} = 1$
- $\binom{4}{3} = 4$

1.1 Counting Problems

- Permutations where repetition is allowed:

$$n!$$

- Permutations where repetition is not allowed

$$\frac{n!}{(n-k)!}$$

A factorial is a positive whole number, based on a number n , and which is written as "$n!$". The factorial $n!$ is defined as follows:

$$n! = n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1$$

Remark $n! = n \times (n-1)!$

Example:

- $3! = 3 \times 2 \times 1 = 6$
- $4! = 4 \times 3! = 4 \times 3 \times 2 \times 1 = 24$

Remark $0! = 1$ not 0.

Often we are concerned with computing the number of ways of selecting and arranging groups of items.

- A ***combination*** describes the selection of items from a larger group of items.
- A ***permutation*** is a combination that is arranged in a particular way.

- Suppose we have items A,B,C and D to choose two items from.
- AB is one possible selection, BD is another. AB and BD are both combinations.
- More importantly, AB is one combination, for which there are two distinct permutations: AB and BA.

Combinations: The number of ways of selecting k objects from n unique objects is:

$${}^{n}C_{k} = \frac{n!}{k! \times (n-k)!}$$

In some texts, the notation for finding the number of possible combination is written

$${}^{n}C_{k} = \binom{n}{k}$$

How many ways are there of selecting two items from possible 5?

$$^5C_2 \left(\text{ also } \binom{5}{2} \right) = \frac{5!}{2! \times 3!} = \frac{5 \times 4 \times 3!}{2 \times 1 \times 3!} = 10$$

Discuss how combinations can be used to compute the number of rugby matches for each group in the Rugby World Cup.

The number of different permutations of r items from n unique items is written as ^nP_k

$$^nP_k = \frac{n!}{(n-k)!}$$

Example: How many ways are there of arranging 3 different jobs, between 5 workers, where each worker can only do one job?

$$^5P_3 = \frac{5!}{(5-3)!} = \frac{5!}{2!} = 60$$

A committee of 4 must be chosen from 3 females and 4 males.

- In how many ways can the committee be chosen.
- In how many cans 2 males and 2 females be chosen.
- Compute the probability of a committee of 2 males and 2 females are chosen.
- Compute the probability of at least two females.

Part 1

We need to choose 4 people from 7:

This can be done in

$$^{7}C_4 = \frac{7!}{4! \times 3!} = \frac{7 \times 6 \times 5 \times 4!}{4! \times 3!} = 35 \text{ ways.}$$

Part 2

With 4 men to choose from, 2 men can be selected in

$$^{4}C_2 = \frac{4!}{2! \times 2!} = \frac{4 \times 3 \times 2!}{2! \times 2!} = 6 \text{ ways.}$$

Similarly 2 women can be selected from 3 in

$$^{3}C_2 = \frac{3!}{2! \times 1!} = \frac{3 \times 2!}{2! \times 1!} = 3 \text{ ways.}$$

When implementing combination calculations in `R`, we use the `choose()` function.

```
> choose(5,0)
[1] 1
> choose(5,1)
[1] 5
> choose(5,2)
[1] 10
> choose(5,3)
[1] 10
> choose(5,4)
[1] 5
```

```
> choose(5,5)
[1] 1
```

Part 2

Thus a committee of 2 men and 2 women can be selected in $6 \times 3 = 18$ ways.

Part 3

The probability of two men and two women on a committee is

$$\frac{\text{Number of ways of selecting 2 men and 2 women}}{\text{Number of ways of selecting 4 from 7}} = \frac{18}{35}$$

Part 4

- The probability of at least two females is the probability of 2 females or 3 females being selected.
- We can use the addition rule, noting that these are two mutually exclusive events.
- From before we know that probability of 2 females being selected is 18/35.

Part 4

- We have to compute the number of ways of selecting 1 male from 4 (4 ways) and the number of ways of selecting three females from 2 (only 1 way)

- The probability of selecting three females is therefore $\frac{4\times 1}{35} = 4/35$
- So using the addition rule

$$Pr(\text{ at least 2 females }) = Pr(\text{ 2 females }) + Pr(\text{ 3 females })$$

$$Pr(\text{ at least 2 females }) = 18/35 + 4/35 = 22/35$$

1.2 Combinations and Permutations

A factorial is a positive whole number, based on a number n , and which is written as "$n!$". The factorial $n!$ is defined as follows:

$$n! = n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1$$

Remark $n! = n \times (n-1)!$

Example:

- $3! = 3 \times 2 \times 1 = 6$
- $4! = 4 \times 3! = 4 \times 3 \times 2 \times 1 = 24$

Remark $0! = 1$ not 0.

Often we are concerned with computing the number of ways of selecting and arranging groups of items.

- A ***combination*** describes the selection of items from a larger group of items.

- A ***permutation*** is a combination that is arranged in a particular way.

- Suppose we have items A,B,C and D to choose two items from.
- AB is one possible selection, BD is another. AB and BD are both combinations.
- More importantly, AB is one combination, for which there are two distinct permutations: AB and BA.

Combinations: The number of ways of selecting k objects from n unique objects is:

$$^{n}C_k = \frac{n!}{k! \times (n-k)!}$$

In some texts, the notation for finding the number of possible combination is written

$$^{n}C_k = \binom{n}{k}$$

How many ways are there of selecting two items from possible 5?

$$^{5}C_2 \left(\text{ also } \binom{5}{2} \right) = \frac{5!}{2! \times 3!} = \frac{5 \times 4 \times 3!}{2 \times 1 \times 3!} = 10$$

Discuss how combinations can be used to compute the number of rugby matches for each group in the Rugby World Cup.

The number of different permutations of r items from n unique items is written as ${}^{n}P_{k}$

$$ {}^{n}P_{k} = \frac{n!}{(n-k)!} $$

Example: How many ways are there of arranging 3 different jobs, between 5 workers, where each worker can only do one job?

$$ {}^{5}P_{3} = \frac{5!}{(5-3)!} = \frac{5!}{2!} = 60 $$

A committee of 4 must be chosen from 3 females and 4 males.

- In how many ways can the committee be chosen.
- In how many cans 2 males and 2 females be chosen.
- Compute the probability of a committee of 2 males and 2 females are chosen.
- Compute the probability of at least two females.

Part 1
We need to choose 4 people from 7:
This can be done in

$$ {}^{7}C_{4} = \frac{7!}{4! \times 3!} = \frac{7 \times 6 \times 5 \times 4!}{4! \times 3!} = 35 \text{ ways.} $$

Part 2

With 4 men to choose from, 2 men can be selected in

$$^4C_2 = \frac{4!}{2! \times 2!} = \frac{4 \times 3 \times 2!}{2! \times 2!} = 6 \text{ ways.}$$

Similarly 2 women can be selected from 3 in

$$^3C_2 = \frac{3!}{2! \times 1!} = \frac{3 \times 2!}{2! \times 1!} = 3 \text{ ways.}$$

When implementing combination calculations in R, we use the `choose()` function.

```
> choose(5,0)
[1] 1
> choose(5,1)
[1] 5
> choose(5,2)
[1] 10
> choose(5,3)
[1] 10
> choose(5,4)
[1] 5
> choose(5,5)
[1] 1
```

Part 2

Thus a committee of 2 men and 2 women can be selected in $6 \times 3 = 18$ ways.

Part 3

The probability of two men and two women on a committee is

$$\frac{\text{Number of ways of selecting 2 men and 2 women}}{\text{Number of ways of selecting 4 from 7}} = \frac{18}{35}$$

Part 4

- The probability of at least two females is the probability of 2 females or 3 females being selected.
- We can use the addition rule, noting that these are two mutually exclusive events.
- From before we know that probability of 2 females being selected is 18/35.
- We have to compute the number of ways of selecting 1 male from 4 (4 ways) and the number of ways of selecting three females from 2 (only 1 way)
- The probability of selecting three females is therefore $\frac{4\times1}{35} = 4/35$
- So using the addition rule

$$Pr(\text{ at least 2 females }) = Pr(\text{ 2 females }) + Pr(\text{ 3 females })$$

$$Pr(\text{ at least 2 females }) = 18/35 + 4/35 = 22/35$$

Combinations and Permutations

Combinations

In mathematical terms, a combination is an subset of items from a larger set such that the order of the items does not matter.

1.2.1 Permutations

- The notion of permutation relates to the act of permuting (rearranging) objects or values.
- Informally, a permutation of a set of objects is an arrangement of those objects into a particular order.
- For example, there are six permutations of the set $\{1, 2, 3\}$, namely (1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), and (3,2,1).
- As another example, an anagram of a word is a permutation of its letters.

If the probability of C is 70% then the probability of C' is 30%

Combinations

Formula

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots 1},$$

which can be written using factorials as whenever $k \leq n$

Example 1

$$\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5.4.3!}{2!.3!} = \frac{5.4}{2.1} = 10$$

Example 2

$$\binom{5}{0} = \frac{5!}{0!(5-0)!} = \frac{5!}{0!.5!} = \frac{5!}{2!} = 1$$

Recall $0! = 1$

1.3 Urn Questions

Urn Questions Suppose an urn contains seven white, four black and three red beads. Three beads are picked at random without replacement. Find the probability that all three beads are the different in colour. at least two beads are the same colour.

- A bag contains 2 red, 3 green and 2 blue balls. Two balls are drawn at random. What is the probability that none of the balls drawn is blue?

  Independent Events

- Competitors A and B fire at their respective targets. The probability that A hits a target is 1/3 and the probability that B hits a target is 1/5. Find the probability that:

i. (2 marks) A does not hit the target,

ii. (2 marks) both hit their respective targets,

iii. (2 marks) only one of them hits a target,

iv. (2 marks) neither A nor B hit their targets.

- Four Students work independently on a mathematical problem. The probability that the four students have of solving the problem are as follows:

- An IT consultant is responsible for three software engineering projects X, Y and Z. He knows that the probability of completing project X in time is 0.99, for project Y this probability is 0.95 and for project Z it is 0.80.

  a What assumption do you need to make in order to calculate the probability of completing all three projects in time, from the information given?

  b Calculate the probability of completing all three projects in time.

  c Calculate the probability that only projects X and Y will be completed on time.

- A doctor treating a patient issues a prescription for antibiotics and provides for two repeat prescriptions. The probability that the infection will be cleared by the first prescription is p_1 =0.6. The probability that successive treatments are successful, given that previous prescriptions were not successful are $p_2 = 0.5$, $p_3 = 0.4$. Calculate the probability that:

a a patient will require the third prescription,

b the patient is still infected after the third prescription,

c the patient is cured by the second prescription, given that the patient is eventually cured.

- Two people look at the letters in the word discovery. Independently of each other, each person writes down two of the letters from the word discovery. What is the probability that

  (i) One person writes down two vowels and the other person

  (ii)

- Three cards are drawn, one after the other, without replacement, from a pack of 52 playing cards. Find the probability that the

- On completion of a programming project, three programmers from a team submit a collection of subroutines to an acceptance group.

  The following table shows the percentage of subroutines each programmer submitted and the probability that a subroutine submitted by each programmer will pass the certification test based on historical data.

Programmer	A	B	C
Proportion of subroutines submitted	0.40	0.35	0.25
Probability of acceptance	0.75	0.95	0.85

i. (3 marks) What is the proportion of subroutines that pass the acceptance test?

ii. (3 marks) After the acceptance tests are completed, one of the subroutines is selected at random and found to have passed the test. What is the probability that it was written by Programmer A?

1.4 Permutations

- Permutations where repetition is allowed:

$$n!$$

- Permutations where repetition is not allowed

$$\frac{n!}{(n-k)!}$$

Choose Operator

1 Choose Operator

$$\binom{n}{k} = \frac{n!}{k! \times (n-k)!}$$

Evaluate the following:

1 $\binom{5}{2}$	3 $\binom{6}{3}$	5 $\binom{10}{1}$
2 $\binom{5}{0}$	4 $\binom{6}{6}$	6 $\binom{10}{9}$

2 In how many ways can a group of four people be selected from three men and four women? In how many of these groups are there more women than men?

3 In how many ways can a group of five be selected from ten people How many groups can be selected if two particular people from the ten can not be selected in the same group?

Counting Sets using Venn Diagrams

4 The Venn Diagram shows the number of elements in each subset of set S. If $P(A) = 3/10$ and $P(B) = 1/2$, find the values of x and y

5 How many different four digit numners greater than 5000 can be formed from the digits **2,4,5,8,9** if each digit can only be used once in any given number. How many of these numbers are odd?

Type of Permutations

There are two types of permutation:

1. Repetition is Allowed: such as the lock above. It could be "333".
2. No Repetition: for example the first three people in a running race. You can't be first and second.

Summary

- If the order doesn't matter, it is a Combination.
- If the order does matter it is a Permutation.

1.4.1 Permutations

How many anagrams (permutations of the letters) are there of the following words

1. ANSWER

2. PERMUTE
3. ANAGRAM
4. LITTLE

Part 1 : ANSWER

Examples:

ASNWRE, SANERW, REWSAN, ...

Since ANSWER has 6 distinct letters, the number of permutations (anagrams) is

$$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = \mathbf{720}$$

Part 2 : PERMUTE

- The word PERMUTE has 7 letters, but only 6 different letters.
- There are 7! ways to arrange 7 letters.
- However, interchanging the two Es does not result in a new permutation. There would be two identical anagrams.

PERMUTE, MUTEPER, PETEMUR, ..
PERMUTE, MUTEPER, PETEMUR, ..

Part 2 : PERMUTE

- The number of permutations (anagrams) is half of 7! .

$$\frac{7!}{2} = \frac{5040}{2} = \mathbf{2520}$$

Part 3 : ANAGRAM

- The word ANAGRAM has 7 letters, but there are three As.
- From before, there are 7! ways to arrange 7 letters.
- How many new permutations are found by re-arranging the As?
- ANAGRAM
- ANAGRAM
- ANAGRAM

- ANAGRAM
- ANAGRAM

- ANAGRAM

Part 3 : ANAGRAM

- We divide 7! by 3! to account for the identical anagrams.

$$\frac{7!}{3!} = \frac{5040}{6} = \mathbf{840}$$

Part 2 : PERMUTE

1.4.2 Permutations

- We re-express the answer from part 2 as follows:

$$\frac{7!}{2!} = \frac{5040}{2} = \mathbf{2520}$$

Part 4 : LITTLE

- The word LITTLE has 6 letters, but there are two Ls and two Ts.
- From before, there are 6! ways to arrange 6 letters.
- Again, interchanging the two Ls and Ts does not result in a new permutation.

$$\frac{6!}{2! \times 2!} = \frac{720}{4} = \mathbf{180}$$

1.4.3 Permutations

- In how many permutations are there of counting a subset of k elements, when there are n elements in total.
- The number of permutations of a set of n elements is denoted n! (pronounced n factorial.)

1.4.4 Permutation Formula

A formula for the number of possible permutations of k objects from a set of n. This is usually written ${}^{n}P_{k}$.

Formula:

$$ {}^{n}P_{k} = \frac{n!}{(n-k)!} = n.(n-1).(n-2)\dots.(n-k+1) $$

Example:
How many ways can 4 students from a group of 15 be lined up for a photograph?
Answer:

There are ${}^{15}P_{4}$ possible permutations of 4 students from a group of 15.

$$ {}^{15}P_{4} = \frac{15!}{11!} = 15 \times 14 \times 13 \times 12 = 32760 $$

There are 32760 different lineups.

1.5 Counting

Given S is the set of all 5 digit binary strings, E is the set of a 5 digit binary strings beginning with a 1 and F is the set of all 5 digit binary strings ending with two zeroes.

(a) Find the cardinality of S, E and F.

(b) Draw a Venn diagram to show the relationship between the sets S, E and F. Show the relevant number of elements in each region of your diagram.

Probability Distributions (Question 2 for End Of Year Exam)

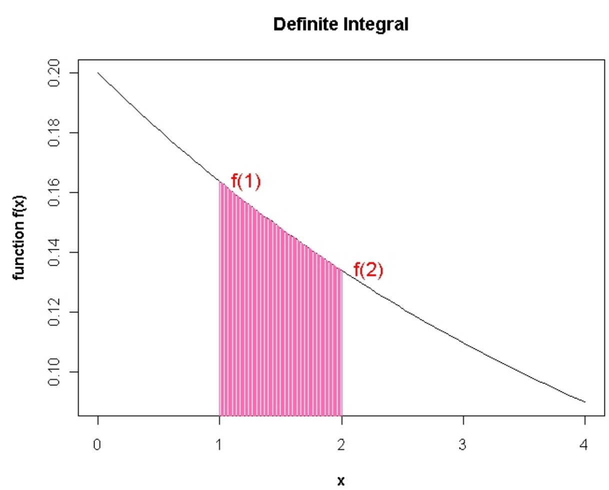
- Discrete Probability Distributions
  - Binomial Probability Distribution (Week 3)
  - Geometric Probability Distribution (Week 3)
  - Poisson Probability Distribution (Week 3/4)
- Continuous Probability Distributions
  - Exponential Probability Distribution (Week 4)
  - Uniform Probability Distribution (Week 4)
  - Normal Probability Distribution (Week 4/5)

- Previously we have been studying discrete random variables, such as the Binomial and the Poisson random variables.
- Now we turn our attention to continuous random variables.
- Recall that a continuous random variable is one which takes an infinite number of possible values, rather than just a countable number of distinct values.
- Continuous random variables are usually measurements.
- Examples include height, weight, the amount of sugar in an orange, the time required to run a mile.

Remarks: This is for continuous distributions only.

- The probability that a continuous random variable will take an exact value is infinitely small. We will usually treat it as if it was zero.
- When we write probabilities for continuous random variables in mathematical notation, we often retain the equality component (i.e. the "...or equal to..").
  For example, we would write expressions $P(X \leq 2)$ or $P(X \geq 5)$.
- Because the probability of an exact value is almost zero, these two expression are equivalent to $P(X < 2)$ or $P(X > 5)$.
- Also, the complement of $P(X \geq k)$ can be written as $P(X \leq k)$.
- Integration is not part of the syllabus, and it is assumed that students are not familiar with how to compute definite integrals.
- However, it is useful to know what the purpose of definite integrals are, because we will be using the results derived from definite integrals.
- It is assumed that students are familiar with functions.

Some function $f(x)$ evaluated at $x = 1$.



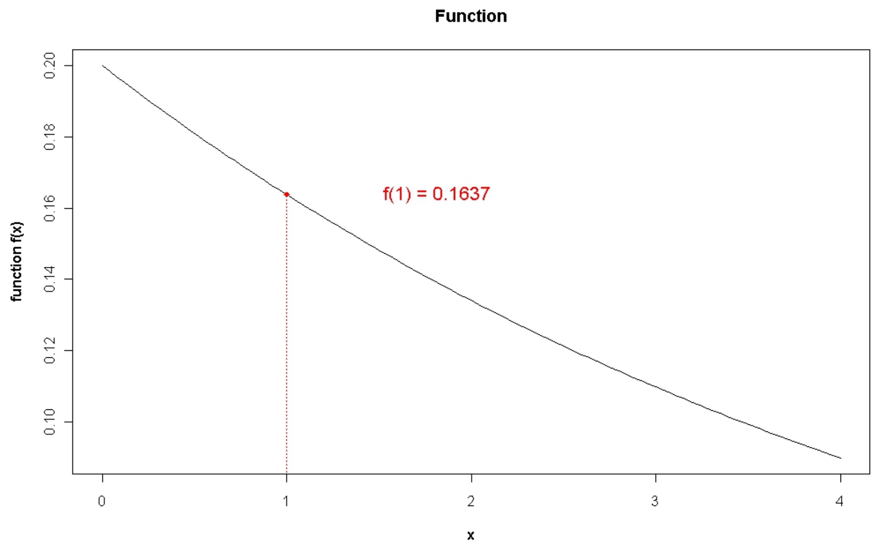


Figure 1.1:

Definite integral of function is area under curve between X=1 and X=2.

- Definite integrals are used to compute the "***area under curves***".
- Definite integrals are defined by a lower and upper limit.
- The area under the curve between X=1 and X=2 is depicted in the previous slide.
- By computing the definite integral, we are able to determine a value for this area.
- Probability can be represented as an area under a curve.

- In probability theory, a ***probability density function*** (PDF) (or "density" for short) of a continuous random variable is a function that describes the relative likelihood for this random variable to occur at a given point.

- The PDF for a continuous random variable X is often denoted $f(x)$.
- The probability density function can be integrated to obtain the probability that the random variable takes a value in a given interval.
- The probability for the random variable to fall within a particular interval is given by the integral of this variable's density over the region.
- The probability density function is non-negative everywhere, and its integral over the entire space is equal to one.
- A plot of the PDF is referred to as a '***density curve***'.
- A density curve that is always on or above the horizontal axis and has total area underneath equal to one.
- Area under the curve in a range of values indicates the proportion of values in that range.
- Density curves come in a variety of shapes, but the normal distribution's bell-shaped densities are perhaps the most commonly encountered.

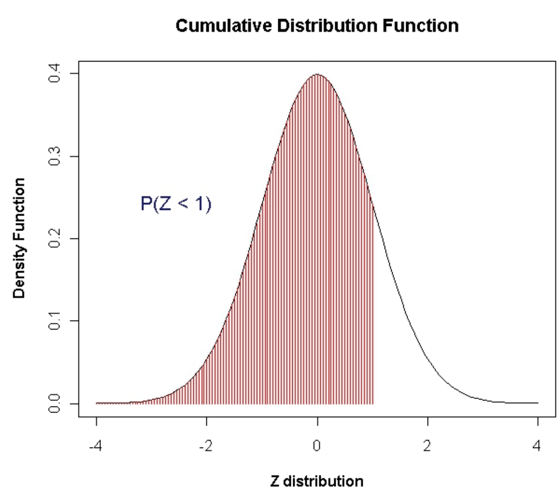
- Remember the density is only an approximation, but it simplifies analysis and is generally accurate enough for practical use.

Recall:

- The ***cumulative distribution function*** (CDF), (or just distribution function), describes the probability that a continuous random variable X with a given probability distribution will be found at a value less than or equal to x.

$$F(x) = P(X \leq x)$$

- Intuitively, it is the "area so far" function of the probability distribution.



Cumulative Distribution Function $P(Z \leq 1)$
Here the random variable is called Z (we will see why later)

- Probability Density Function
- Cumulative Density Function

If X is a continuous random variable then we can say that the probability of obtaining a **precise** value x is infinitely small, i.e. close to zero.

$$P(X = x) \approx 0$$

Consequently, for continuous random variables (only), $P(X \leq x)$ and $P(X < x)$ can be used interchangeably.

$$P(X \leq x) \approx P(X < x)$$

A random variable X is called a continuous uniform random variable over the interval (a, b) if it's probability density function is given by

$$f_X(x) = \frac{1}{b-a} \qquad \text{when } a \leq x \leq b$$

The corresponding cumulative density function is

$$F_x(x) = \frac{x-a}{b-a} \qquad \text{when } a \leq x \leq b$$

The mean of the continuous uniform distribution is

$$E(X) = \frac{a+b}{2}$$

$$V(X) = \frac{(b-a)^2}{12}$$

The most interesting property of the exponential distribution is the ***memoryless*** property. By this , we mean that if the lifetime of a component is exponentially distributed, then an item which has been in use for some time is a good as a brand new item with regards to the likelihood of failure.

The exponential distribution is the only distribution that has this property.

Session 09: Probability

9A.1 Counting Methods

9A.2 Counting using Sets

9A.3 Probability

9A.4 Independent Events

9B.1 Permutation

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}$$

$$\binom{6}{3} = \frac{6!}{(6-3)!3!} = \frac{6!}{3! \times 3!}$$

$$\frac{6!}{3! \times 3!} = \frac{6 \times 5 \times 4 \times 3!}{3! \times 3!} = \frac{120}{6} = 120$$

- $\binom{6}{2} = 15$
- $\binom{5}{2} = 10$
- $\binom{4}{0} = 1$
- $\binom{4}{3} = 4$
- pairwise disjoint sets
- The addition principle

Theorem

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Probability

9B.2 The sample space of an experiment (S)

9B.3 The size of a sample space

9B.4 Indepedent Evcents (9.3.1)

Session 9 Probability

- The outcome of an experiment need not be a number, for example, the outcome when a coin is tossed can be 'heads' or 'tails'.
- However, we often want to represent outcomes as numbers.
- A ***random variable*** is a function that associates a unique numerical value with every outcome of an experiment.
- The value of the random variable will vary from trial to trial as the experiment is repeated.
- Numeric values can be assigned to outcomes that are not usually considered numeric.
- For example, we could assign a 'head' a value of 0, and a 'tail' a value of 1, or vice versa.

There are two types of random variable - discrete and continuous. The distinction between both types will be important later on in the course.

Examples

- A coin is tossed ten times. The random variable X is the number of tails that are noted. X can only take the values $\{0, 1, ..., 10\}$, so X is a discrete random variable.

- A light bulb is burned until it burns out. The random variable Y is its lifetime in hours. Y can take any positive real value, so Y is a continuous random variable.
- A discrete random variable is one which may take on only a countable number of distinct values such as $\{0, 1, 2, 3, 4, ...\}$.
- Discrete random variables are usually (but not necessarily) counts.
- If a random variable can take only a finite number of distinct values, then it must be discrete.
- Examples of discrete random variables include the number of children in a family, the Friday night attendance at a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.
- A continuous random variable is one which takes an infinite number of possible values.
- Continuous random variables are usually measurements.
- Examples include height, weight, the amount of sugar in an orange, the time required to run a computer simulation.

A pair of dice is thrown. Let X denote the minimum of the two numbers which occur. Find the distributions and expected value of X. A fair coin is tossed four times. Let X denote the longest string of heads. Find the distribution and expectation of X.

A fair coin is tossed until a head or five tails occurs. Find the expected number E of tosses of the coin.

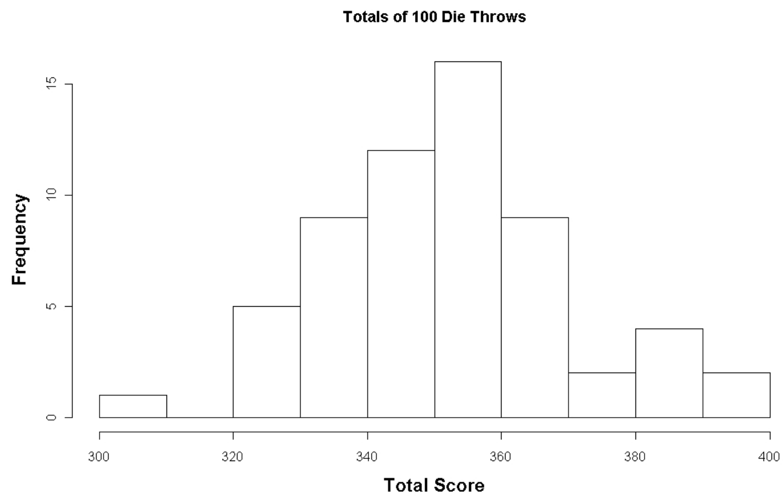
The coin is tossed three times. Let X denote the number of heads that appear.

- (a) Find the distribution f of X.
- (b) Find the expectation E(X).
- Bar-plots
- Histograms
- Boxplots
- Consider an experiment in which each student in a class of 60 rolls a die 100 times.
- Each score is recorded, and a total score is calculated.
- As the expected value of rolled die is 3.5, the expected total is 350 for each student.
- At the end of the experiment the students reported their totals.
- The totals were put into ascending order, and tabulated as follows (next slide).

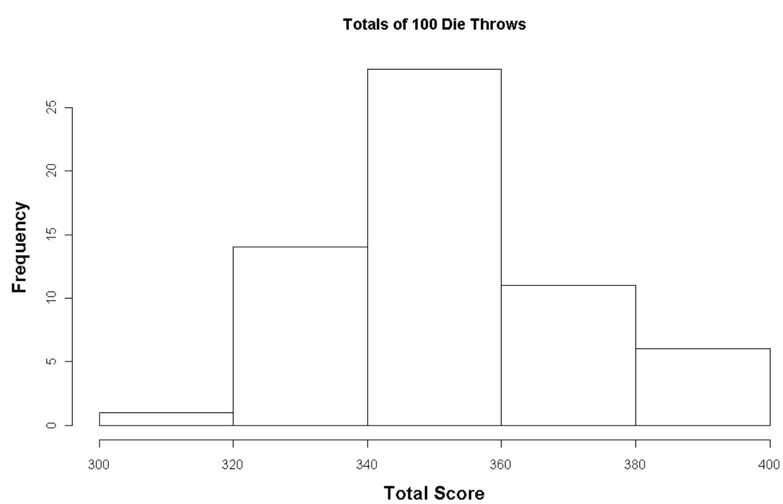
307	321	324	328	329	330	334	335	336	337
337	337	338	339	339	342	343	343	344	344
346	346	347	348	348	348	350	351	352	352
353	353	353	354	354	356	356	357	357	358
358	360	360	361	362	363	365	365	369	369
370	370	374	378	381	384	385	386	392	398

- What proportion of outcomes are less than or equal to 330?
  (Answer: 10%)
- What proportion of outcomes are greater than or equal to 370?
  (Answer: 16.66%)

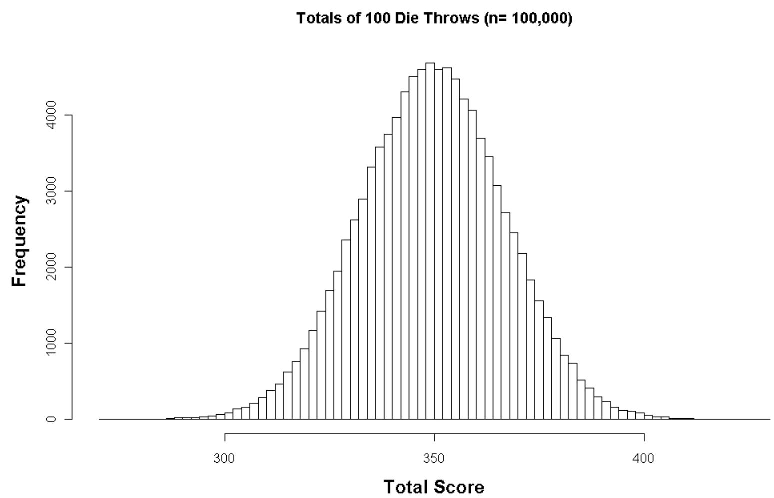
TEXT HERE
For the die-throw experiment;



- Compute an appropriate number of class intervals.
- As a rule of thumb, the number of class intervals is usually approximately the square root of the number of observations.
- As there are 60 observations, we would normally use 7 or 8 class intervals.
- To save time, we will just use 5 class intervals.



- Suppose that the experiment of throwing a die 100 times and recording the total was repeated 100,000 times.
- (If implemented on a computer, we would call this a simulation study)
- The histogram of data (with a class interval width of 2) is shown on the next slide.
- How should the shape of the histogram be described?
- "Bell-shaped" would be a suitable description.



A couple of remarks about the simulation study, some of which will be relevant later on.

- Approximately 68.7% of the values in the simulation study are between 332 and 367.
- Approximately 95% of the values are between 316 and 383.
- 2.5% of the values output are less than 316.
- 2.5% of the values study output are greater than 383.
- 175 values are greater than or equal to 400, whereas 198 values are less than or equal to 300.
- Results such as these are unusual, but they are not impossible.

A pair of dice is thrown. Let X denote the minimum of the two numbers which occur. Find the distributions and expected value of X.

A fair coin is tossed four times. Let X denote the longest string of heads. Find the distribution and expectation of X.

A fair coin is tossed until a head or five tails occurs. Find the expected number E of tosses of the coin.

The coin is tossed three times. Let X denote the number of heads that appear.

- (a) Find the distribution f of X.
- (b) Find the expectation E(X).
- Now consider an experiment with only two outcomes. Independent repeated trials of such an experiment are called Bernoulli trials, named after the Swiss mathematician Jacob Bernoulli (16541705).
- The term ***independent trials*** means that the outcome of any trial does not depend on the previous outcomes (such as tossing a coin).
- We will call one of the outcomes the “success” and the other outcome the “failure”.
- Let p denote the probability of success in a Bernoulli trial, and so $q = 1 - p$ is the probability of failure. A binomial experiment consists of a fixed number of Bernoulli trials.
- A binomial experiment with n trials and probability p of success will be denoted by

$$B(n, p)$$

Binomial Coefficients

- factorials

$$n! = (n) \times (n-1) \times (n-2) \times \ldots \times 1$$

  - $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$
  - $3! = 3 \times 2 \times 1$

- Zero factorial

$$0! = 1$$

The complement rule in Probability
$P(C') = 1 - P(C)$

- Over the next set of lectures, we are now going to look at two important discrete probability distributions
- The first is the ***binomial*** probability distribution.
- The second is the Poisson probability distribution.
- In `R`, calculations are performed using the `binom` family of functions and `pois` family of functions respectively.
- Now consider an experiment with only two outcomes. Independent repeated trials of such an experiment are called Bernoulli trials, named after the Swiss mathematician Jacob Bernoulli (16541705).
- The term ***independent trials*** means that the outcome of any trial does not depend on the previous outcomes (such as tossing a coin).
- We will call one of the outcomes the "success" and the other outcome the "failure".
- Let p denote the probability of success in a Bernoulli trial, and so $q = 1 - p$ is the probability of failure. A binomial experiment consists of a fixed number of Bernoulli trials.
- A binomial experiment with n trials and probability p of success will be denoted by

$$B(n, p)$$

- a probability mass function (pmf) is a function that gives the probability that a discrete random variable is exactly equal to some value.
- The probability mass function is often the primary means of defining a discrete probability distribution

Thirty-eight students took the test. The X-axis shows various intervals of scores (the interval labeled 35 includes any score from 32.5 to 37.5). The Y-axis shows the number of students scoring in the interval or below the interval.

cumulative frequency distribution A can show either the actual frequencies at or below each interval (as shown here) or the percentage of the scores at or below each interval. The plot can be a histogram as shown here or a polygon.

Session 09: Probability

9B.1 Permutation

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}$$

$$\binom{6}{3} = \frac{6!}{(6-3)!3!} = \frac{6!}{3! \times 3!}$$

$$\frac{6!}{3! \times 3!} = \frac{6 \times 5 \times 4 \times 3!}{3! \times 3!} = \frac{120}{6} = 120$$

- $\binom{6}{2} = 15$
- $\binom{5}{2} = 10$
- $\binom{4}{0} = 1$

- $\binom{4}{3} = 4$
- pairwise disjoint sets
- The addition principle

Theorem

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Probability

9B.2 The sample space of an experiment (S)

9B.3 The size of a sample space

9B.4 Indepedent Evcents (9.3.1)

Suppose an electronics assembly subcontractor receives resistors from two suppliers: A and B

- Supplier A supplies 80% of the resistors

- Supplier B supplies 20% of the resistors

Suppose an electronics assembly subcontractor recieves resistors from two suppliers A and B

- Supplier A supplies 80% of the resistors

- *Probability that a randomly chosen resistor comes from A is 80 %*
- *P(A) = 0.80*
- Supplier B supplies 20% of the resistors
- *Probability that a randomly chosen resistor comes from B is therefore 20%*
- *P(B) = 0.20*
- We are giving information about the rate of faulty components from each supplier.
  (Faulty : resistor fails some quality test)
- 1% of the resistors supplied by A are faulty
- 3% of the resistors supplied by B are faulty
- We are giving information about the rate of faulty components from each supplier.
  (Faulty : resistor fails some quality test)
- *P(F) probability that randomly selecting compoent is faulty*
- 1% of the resistors supplied by A are faulty.

- *We write this as* $P(F|A) = 0.01$
- 3% of the resistors supplied by B are faulty
- *We write this as* $P(F|B) = 0.03$

Question 1:

- What is the probability that a randomly selected resistor fails the final test?
- In mathematical terms, compute P(F)

Law of Total Probability:

- Faulty Resistors are either from Supplier A or Supplier B.
- *Resistors MUST come from one of the two suppliers.*
- *A and B are mutually exclusive.*

$$P(F) = P(F \text{ and } A) + P(F \text{ and } B)$$

Conditional Probability

$$P(X|Y) = \frac{P(X \text{ and } Y)}{P(Y)}$$

Re-arranging

$$P(X \text{ and } Y) = P(X|Y) \times P(Y)$$

Therefore we can say

$$P(F \text{ and } A) = P(F|A) \times P(A)$$

$$P(F \text{ and } B) \equiv P(F|B) \times P(B)$$

$$P(F \text{ and } A) \equiv P(F|A) \times P(A)$$

$$P(F \text{ and } B) = P(F|B) \times P(B)$$

$$P(F \text{ and } A) = P(F|A) \times P(A) = 0.80 \times 0.01$$

$$P(F \text{ and } A) = 0.008$$

$$P(F \text{ and } B) = P(F|B) \times P(B) = 0.20 \times 0.03$$

$$P(F \text{ and } B) = 0.006$$

Recall:

$$P(F) = P(F \text{ and } A) + P(F \text{ and } B)$$

Session 9 Probability

The complement rule in Probability

$P(C') = 1 - P(C)$

If the probability of C is 70% then the probability of C' is 30%

- Permutations where repetition is allowed:

$$n!$$

- Permutations where repetition is not allowed

$$\frac{n!}{(n-k)!}$$

Expected Values of Random Variables

Example Exercise

www.MathsResource.com

If the random variable Z has a distribution which is standard normal, show that the expected value of e^{sZ} is given as follows:

$$E(e^{sZ}) = e^{\frac{s^2}{2}}$$

- In general, the expected value is computed using this formula

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

- The expected value of a ***transformed*** random variable is computed using this formula

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

- The probability density function for the standard normal distribution is

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

- The probability density function for the standard normal distribution is

$$f(z) = f(x, \mu = 0, \sigma = 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}^2}$$

$$E(e^{sZ}) = \int_{-\infty}^{\infty} e^{sx} \; \times \; \frac{1}{\sqrt{2\pi}} \; e^{-\frac{x}{2}^2}$$

$$E(e^{sZ}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\left[sx - \frac{(x)^2}{2}\right]} dx$$

$$E(e^{sZ}) = e^{\frac{s^2}{2}} \; \times \; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\left[-\frac{(x-s)^2}{2}\right]} dx$$

Mathematical Identity

- Proven in a separate video

$$\int_{-\infty}^{\infty} e^{-\frac{y}{2}^2} dy = \sqrt{2\pi}$$

$$E(e^{sZ}) = e^{\frac{s^2}{2}} \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\left[-\frac{(x-s)^2}{2}\right]} dx$$

$$E(e^{sZ}) = e^{\frac{s^2}{2}} \times \frac{1}{\sqrt{2\pi}} \left[\sqrt{2\pi}\right]$$

Suppose an electronics assembly subcontractor recieves resistors from two suppliers A and B

Supplier A supplies 80

P(A) = 0.80 probability that a randomly chosen resistor comes from A

Supplier B supplies 20

P(B) = 0.20 probability that a randomly chosen resistor comes from B

- 1
- 3

Qeustion: What is the probability that a randomly selected resistor fails the final test?

Compute P(F)

P(F) = P(FandA) + P(FandB)

A probability distribution is a mathematical approach to quantifying uncertainty.

There are two main classes of probability distributions: Discrete and continuous.

Discrete distributions describe variables that take on discrete values only (typically the positive integers), while continuous distributions describe variables that can take on arbitrary values in a continuum (typically the real numbers).

Binomial Distribution

A manufacturer of hospital equipment knows from experience that 5

Number of independent trials n

Probability of a "success" p

A basic introduction to the concept

Example

Certain events happen at unpredictable intervals. But for some reason, no matter how recent or long ago last event was, the probability that another event will occur within the next hour is exactly the same (say, 10

Then the number of events per day is Poisson distributed.

Formal definition

Let X be a stochastic variable taking non-negative integer values with probability density function

$$P(X = k) = f(k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Then X follows the Poisson distribution with parameter λ.

Characteristics of the Poisson distribution

If X is a Poisson distribution stochastic variable with parameter λ, then

- The expected value $E[X] = \lambda$
- The variance $Var[X] = \lambda$

The Normal Distribution

Symmetric Intervals

Symmetric Intervals

$$P(-z \leq Z \leq z)$$

The Normal Distribution

The Symmetry Rule

The Normal Distribution

The Symmetry Rule

From statistical tables, we could determine the following:

- $P(Z \leq 1.5)$
- $P(Z \geq 1.5)$

Consider the normally distributed random variable X

$$X \sim \mathcal{N}(\mu = 1000, \sigma^2 = 2500)$$

Paramters:

- $\mu = 1000$
- $\sigma = 50$

Questions

- $P(X \leq 925)$
- $P(X \geq 925)$

Z-score

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

$$X \sim \mathcal{N}(\mu = 1000, \sigma^2 = 2500)$$

- Mean $\mu = 1000$
- Standard Deviation $\sigma = 50$

$$P(X \leq 925) = P(Z \leq -1.5)$$

Applying the Symmetry Rule

$$P(Z \leq -1.5) = P(Z \geq 1.5) = 0.0668$$

Theerefore we can say

$$P(X \leq 925) = 0.0668$$

As a consequence of Property 1, it is possible to relate all normal random variables to the standard normal.

If $X \sim N(\mu, \sigma^2)$, then

$$Z = \frac{X - \mu}{\sigma}$$

is a standard normal random variable: $Z \sim N(0, 1)$. An important consequence is that the cdf of a general normal distribution is therefore

$$\Pr(X \leq x) = \Phi\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x - \mu}{\sigma\sqrt{2}}\right)\right)$$

.

Conversely, if Z is a standard normal distribution, $Z \sim N(0, 1)$, then

$$X = \sigma Z + \mu$$

is a normal random variable with mean μ and variance σ^2.

The standard normal distribution has been tabulated (usually in the form of value of the cumulative distribution function F), and the other normal distributions are the simple transformations, as described above, of the standard one. Therefore, one can use tabulated values of the cdf of the standard normal distribution to find values of the cdf of a general normal distribution.

1.5.1 Statistics

1. Sample mean
$$\bar{x} = \frac{\sum x_i}{n}.$$
2. Sample standard deviation
$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}.$$
3. Conditional probability:
$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}.$$

Complement Rule

The complement rule in Probability

$P(C') = 1 - P(C)$

If the probability of C is 70% then the probability of C' is 30%

- The outcome of an experiment need not be a number, for example, the outcome when a coin is tossed can be ‘heads’ or ‘tails’.

- However, we often want to represent outcomes as numbers.
- A ***random variable*** is a function that associates a unique numerical value with every outcome of an experiment.
- The value of the random variable will vary from trial to trial as the experiment is repeated.
- Numeric values can be assigned to outcomes that are not usually considered numeric.
- For example, we could assign a 'head' a value of 0, and a 'tail' a value of 1, or vice versa.

There are two types of random variable - discrete and continuous. The distinction between both types will be important later on in the course.

Examples

- A coin is tossed ten times. The random variable X is the number of tails that are noted. X can only take the values $\{0, 1, ..., 10\}$, so X is a discrete random variable.

- A light bulb is burned until it burns out. The random variable Y is its lifetime in hours. Y can take any positive real value, so Y is a continuous random variable.

- A discrete random variable is one which may take on only a countable number of distinct values such as $\{0, 1, 2, 3, 4, .$

- Discrete random variables are usually (but not necessarily) counts.

- If a random variable can take only a finite number of distinct values, then it must be discrete.

- Examples of discrete random variables include the number of children in a family, the Friday night attendance at a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.

- A continuous random variable is one which takes an infinite number of possible values.

- Continuous random variables are usually measurements.

- Examples include height, weight, the amount of sugar in an orange, the time required to run a computer simulation.

A pair of dice is thrown. Let X denote the minimum of the two numbers which occur. Find the distributions and expected value of X.

A fair coin is tossed four times. Let X denote the longest string of heads. Find the distribution and expectation of X.

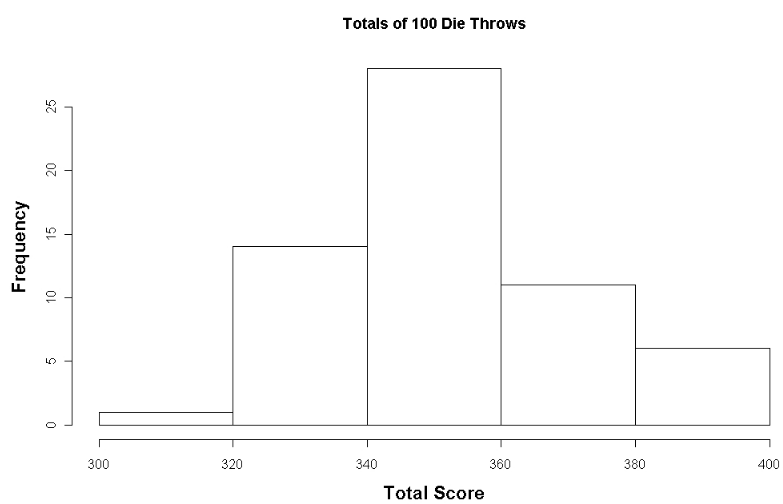
A fair coin is tossed until a head or five tails occurs. Find the expected number E of tosses of the coin.

The coin is tossed three times. Let X denote the number of heads that appear.

- (a) Find the distribution f of X.
- (b) Find the expectation E(X).
- Bar-plots
- Histograms
- Boxplots
- Consider an experiment in which each student in a class of 60 rolls a die 100 times.
- Each score is recorded, and a total score is calculated.
- As the expected value of rolled die is 3.5, the expected total is 350 for each student.
- At the end of the experiment the students reported their totals.
- The totals were put into ascending order, and tabulated as follows (next slide).

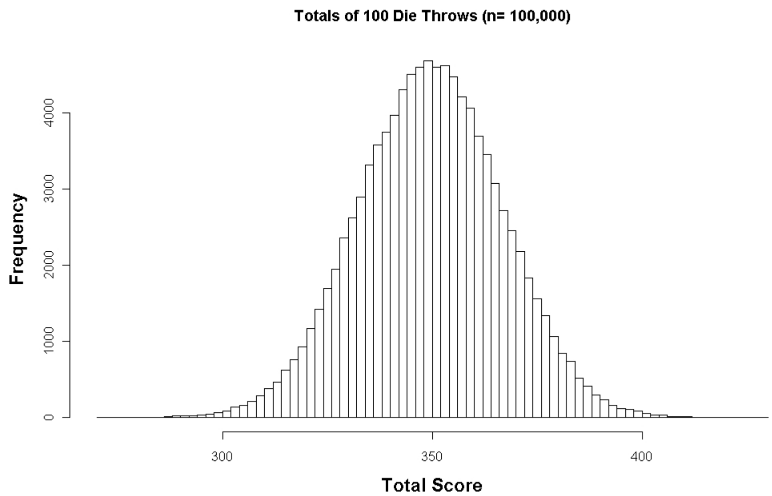
307	321	324	328	329	330	334	335	336	337
337	337	338	339	339	342	343	343	344	344
346	346	347	348	348	348	350	351	352	352
353	353	353	354	354	356	356	357	357	358
358	360	360	361	362	363	365	365	369	369
370	370	374	378	381	384	385	386	392	398

- What proportion of outcomes are less than or equal to 330?
  (Answer: 10%)
- What proportion of outcomes are greater than or equal to 370?
  (Answer: 16.66%)
- Compute an appropriate number of class intervals.
- As a rule of thumb, the number of class intervals is usually approximately the square root of the number of observations.
- As there are 60 observations, we would normally use 7 or 8 class intervals.
- To save time, we will just use 5 class intervals.



- Suppose that the experiment of throwing a die 100 times and recording the total was repeated 100,000 times.
- (If implemented on a computer, we would call this a simulation study)

- The histogram of data (with a class interval width of 2) is shown on the next slide.
- How should the shape of the histogram be described?
- "Bell-shaped" would be a suitable description.



A couple of remarks about the simulation study, some of which will be relevant later on.

- Approximately 68.7% of the values in the simulation study are between 332 and 367.
- Approximately 95% of the values are between 316 and 383.
- 2.5% of the values output are less than 316.
- 2.5% of the values study output are greater than 383.
- 175 values are greater than or equal to 400, whereas 198 values are less than or equal to 300.
- Results such as these are unusual, but they are not impossible.

A pair of dice is thrown. Let X denote the minimum of the two numbers which occur. Find the distributions and expected value of X.

A fair coin is tossed four times. Let X denote the longest string of heads. Find the distribution and expectation of X.

A fair coin is tossed until a head or five tails occurs. Find the expected number E of tosses of the coin.

The coin is tossed three times. Let X denote the number of heads that appear.

- (a) Find the distribution f of X.
- (b) Find the expectation E(X).
- Now consider an experiment with only two outcomes. Independent repeated trials of such an experiment are called Bernoulli trials, named after the Swiss mathematician Jacob Bernoulli (16541705).
- The term ***independent trials*** means that the outcome of any trial does not depend on the previous outcomes (such as tossing a coin).
- We will call one of the outcomes the "success" and the other outcome the "failure".
- Let p denote the probability of success in a Bernoulli trial, and so $q = 1 - p$ is the probability of failure. A binomial experiment consists of a fixed number of Bernoulli trials.
- A binomial experiment with n trials and probability p of success will be denoted by

$$B(n, p)$$

Chapter 2

Discrete Probability Distributions

2.1 Probability Formulae

- Conditional probability:

$$P(B|A) = \frac{P\,(A \text{ and } B)}{P\,(A)}.$$

- Bayes' Theorem:

$$P(B|A) = \frac{P\,(A|B) \times P(B)}{P\,(A)}.$$

Section 3 : Probability

How to Compute Probability: Equally Likely Outcomes Sometimes, a statistical experiment can have n possible outcomes, each of which is equally likely. Suppose a subset of r outcomes are classified as "successful" outcomes.

The probability that the experiment results in a successful outcome (S) is:

P(S) = (Number of successful outcomes) / (Total number of equally likely outcomes) = r / n

Consider the following experiment. An urn has 10 marbles. Two marbles are red, three are green, and five are blue. If an experimenter randomly selects 1 marble from the urn, what is the probability that it will be green?

In this experiment, there are 10 equally likely outcomes, three of which are green marbles. Therefore, the probability of choosing a green marble is 3/10 or 0.30.

- Conditional probability

- Independent events
- Repeated independent events

2.2 Mutually Exclusive Events

Mutually exclusive events are events that cannot happen at the same time.

$$P(AandB) = P(A) + P(B)$$

2.3 Quantiles

The quantile (this term was first used by Kendall, 1940) of a distribution of values is a number xp such that a proportion p of the population values are less than or equal to xp. For example, the .25 quantile (also referred to as the 25th percentile or lower quartile) of a variable is a value (xp) such that 25% (p) of the values of the variable fall below that value.

Similarly, the 0.75 quantile (also referred to as the 75th percentile or upper quartile) is a value such that 75% of the values of the variable fall below that value and is calculated accordingly.

See

2.4 Probability Distribution

A statistical function that describes all the possible values and likelihoods that a random variable can take within a given range. This range will be between the minimum and maximum statistically possible values, but where the possible value is likely to be plotted on the probability distribution depends on a number of factors, including the distributions mean, standard deviation, skewness and kurtosis.

2.5 The binomial distribution

The binomial distribution is a discrete probability distribution that is applicable as a model for decisionmaking situations in which a sampling process can be assumed to conform to a Bernoulli process. A Bernoulli process is a sampling process in which

(1) Only two mutually exclusive possible outcomes are possible in each trial, or observation. For convenience these are called success and failure.

(2) The outcomes in the series of trials, or observations, constitute independent events.

(3) The probability of success in each trial, denoted by p, remains constant from trial to trial. That is, the process is stationary.

The binomial distribution can be used to determine the probability of obtaining a designated number of successes in a Bernoulli process. Three values are required: the designated number of successes (X); the number of trials, or observations (n); and the probability of success in each trial (p). Where $q = (1 - p)$, the formula for determining the probability of a specific number of successes X for a binomial distribution is

Formula

Chapter 3

Normal Probability Distribution

3.1 Normal Distribution: Worked Examples

MA4413 Computer Maths 3 January 2007

Q5. a) Assume that the amount of wine poured into a bottle has a normal distribution with a mean of 750ml and a variance of 144ml^2.

(i) Calculate the probability that a bottle contains more than 765ml. (2 marks)

(ii) Calculate the probability that a bottle contains between 744ml and 759ml. (3 marks)

A machine fills bags with animal feed. The nominal weight of a bag is 50kg. Because random variations the weight of a filled bag is normally distributed $N(\mu, \sigma^2)$. The variance (σ^2) is known to be 0.01kg^2 and μ is set by the operator to a particular value.

(i) If ? = 50kg calculate the probability of a bag containing less than 49.95kg? (ii) Calculate the value of "?" such that only 2% of the output are under the nominal weight?

MA4004 Engineering Statistics SPRING 2008

a) The amount of beer in a bottle has a normal distribution with mean 500ml and variance 25ml2. i) Calculate the probability that the amount of beer in the bottle is between 498ml and 504ml. ii) What volume is exceeded by 20% of the bottles? (6 marks)

Question 1 : Probability Distribution

Introduction

Consider playing a game in which you are winning when a ***fair die*** is showing 'six' and losing otherwise.

Part 1

If you play three such games in a row, find the probability mass function (pmf) of the number X of times you have won.

- Firstly: what type of probability distribution is this?
- Is this the distribution ***discrete*** or ***continuous***?
- The outcomes are whole numbers - so the answer is discrete.
- So which type of discrete distribution? (We have two to choose from. See first page of formulae)
- **Binomial:** characterizing the number of ***successes*** in a series of n ***independent trials***, with the ***probability of a success*** in each trial being p.
- **Poisson:** characterizing the ***number of occurrences*** in a ***unit space*** (i.e. a unit length, unit area or unit volume, or a unit period in time), where λ is the the number of occurrences per unit space.

3.1.1 The Standard Normal Distribution

3.1.2 Standardisation Formula

$$Z = (X - \mu)/\sigma \tag{3.1}$$

Uniform Distribution: Exercise 24

Use the uniform distribution to simulate 100 throws of two dice. The outcome is the combined values of both dice. Use the appropriate R command to discretize values.

- What is the mean and standard deviation of the outcomes?
- Make a stem-and-leaf plot of the outcomes.
- Make a histogram of the outcomes. (hint: use breaks =seq(1.5,12.5))

3.1.3 Example 2

A machine produces components whose thicknesses are normally distributed with a mean of 0.40 cm and a standard deviation of 0.02 cm. Components are rejected if they have a thickness outside the range 0.38 cm to 0.41 cm. (i) What is the probability that a component will have a thickness exceeding 0.41 cm? (4 marks) (ii) What is the probability that a component will have a thickness between 0.38 cm and 0.41 cm? (4 marks) (iii) What is the thickness below which 25% of the components will be? (4 marks)

3.1.4 Example 3

A charity believes that when it puts out an appeal for charitable donations the donations it receives will normally distributed with a mean 50 and standard deviation 6, and it is assumed that donations will be independent of each other.

- Find the probability that the first donation it receives will be greater than 40.
- Find the probability that it will be between 55 and 60.
- Find the value x such that 5% of donations are more than x.

9B.1 Permutation

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}$$

$$\binom{6}{3} = \frac{6!}{(6-3)!3!} = \frac{6!}{3! \times 3!}$$

$$\frac{6!}{3! \times 3!} = \frac{6 \times 5 \times 4 \times 3!}{3! \times 3!} = \frac{120}{6} = 120$$

- pairwise disjoint sets
- The addition principle

Theorem

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Probability

9B.2 The sample space of an experiment (S)

9B.3 The size of a sample space

9B.4 Indepedent Evcents (9.3.1)

Session 09: Probability

- pairwise disjoint sets
- The addition principle

Theorem

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Probability

9B.2 The sample space of an experiment (S)

9B.3 The size of a sample space

9B.4 Indepedent Evcents (9.3.1)

Session 9 Probability

3.2 Discrete Random Variables

1. The probability distribution of discrete random variable X is tabulated below. There are 6 possible outcome of X, i.e. 0, 1, 2, 4 ,8 and 10.

x_i	0	1	2	4	8	10
$P(x_i)$	0.25	0.15	0.25	0.15	k	0.10

   i. (1 marks) Compute the value for k.
   ii. (3 marks) Determine the expected value $E(X)$.
   iii. (2 marks) Evaluate $E(X^2)$.
   iv. (3 marks) Compute the variance of random variable X.

2. Suppose X is a random variable with

   - $E(X^2) = 3.6$
   - $P(X = 2) = 0.6$
   - $P(X = 3) = 0.1$

   (a) The random variable takes just one other value besides 2 and 3. This value is greater than 0. What is this value?
   (b) What is the variance of X?

3. Consider the random variables X and Y. Both X and Y take the values 0, 1 and 2. The joint probabilities for each pair are given by the following table.

	$X = 0$	$X = 1$	$X = 2$
$Y = 0$	0.1	0.15	0.1
$Y = 1$	0.1	0.1	0.1
$Y = 2$	0.2	0.05	0.1

   Compute the $E(U)$ expected value of U, where $U = X - Y$.

- Suppose we have a set of **n** items.
- From that set, we create a subset of **k** items.
- The **order** in which items are selected is recorded. (The ordering of selected items is very important.)
- The total number of **ordered subsets** of **k** items chosen from a set of **n** items is

$$\frac{n!}{n-k!}$$

An ordered sequence of four digits is formed by choosing digits without repetition from the set $\{1, 2, 3, 4, 5, 6, 7\}$.

(i) the total number of such sequences; (780)

(ii) the number of sequences which begin with an odd number; (480) N(A)

(iii) the number of sequences which end with an odd number; (480) (NB)

(iv) the number of sequences which begin and end with an odd number;(240)

(v) the number of sequneces which begin with an odd number or end with an odd number or both; (720)

(vi) the number of sequences which begin with an odd number or end with an odd number but not both. (480)

A college teaches a range of courses including maths, physics and IT. Students choose a range of courses from these three subject areas. Currently 600 students are enrolled of whom 300 study maths courses, 120 study IT and 380 study physics courses.

- 40 students study courses from all three subject areas.
- 200 maths students study physics as well. 60 physics students also study IT and 70 IT students also study maths. 20 students study physics and IT, but not maths.

- How many students study none of these courses at all? (90)
- How many students study maths but not physics or IT? (70)
- How many students study both maths and physics but not IT? (160)
- How many students study courses from precisely two of these subject areas? (210)