Probability & Statistics

Def: A random experiment is one whose outcome is not predictable in advance. We can determine the range of possible outcomes but not which specific outcome will occur.

An outcome is the result of conducting a particular outcome.

For eg. Roll a six-sided die. A 3 is obtained which is the outcome of that random experiment.

Def: The sample space for an experiment S, is the set of all possible outcomes for that experiment. For a given experiment the sample space is always the same.

Eg. Toss a coin. The sample space $S = \{H, T\}$ ie heads or tails. $\# S = 2$.

Eg. Toss two coins. The sample space $S = \{HH, HT, TH, TT\}$. $\# S = 4$.

Eg. Roll a die. $S = \{1, 2, 3, 4, 5, 6\}$, $\# S = 6$.

Eg. Roll two dice. $S = \{11, 12, 13, 14, 15, 16, 21, 22, \cdots, 66\}$.

$\# S = 36$.

Eg. An urn contains 1 red, 1 green and 1 blue ball. A ball is drawn at random. $S = \{R, G, B\}$, $\#S = 3$.

If three balls are drawn at random (successively) without replacement then

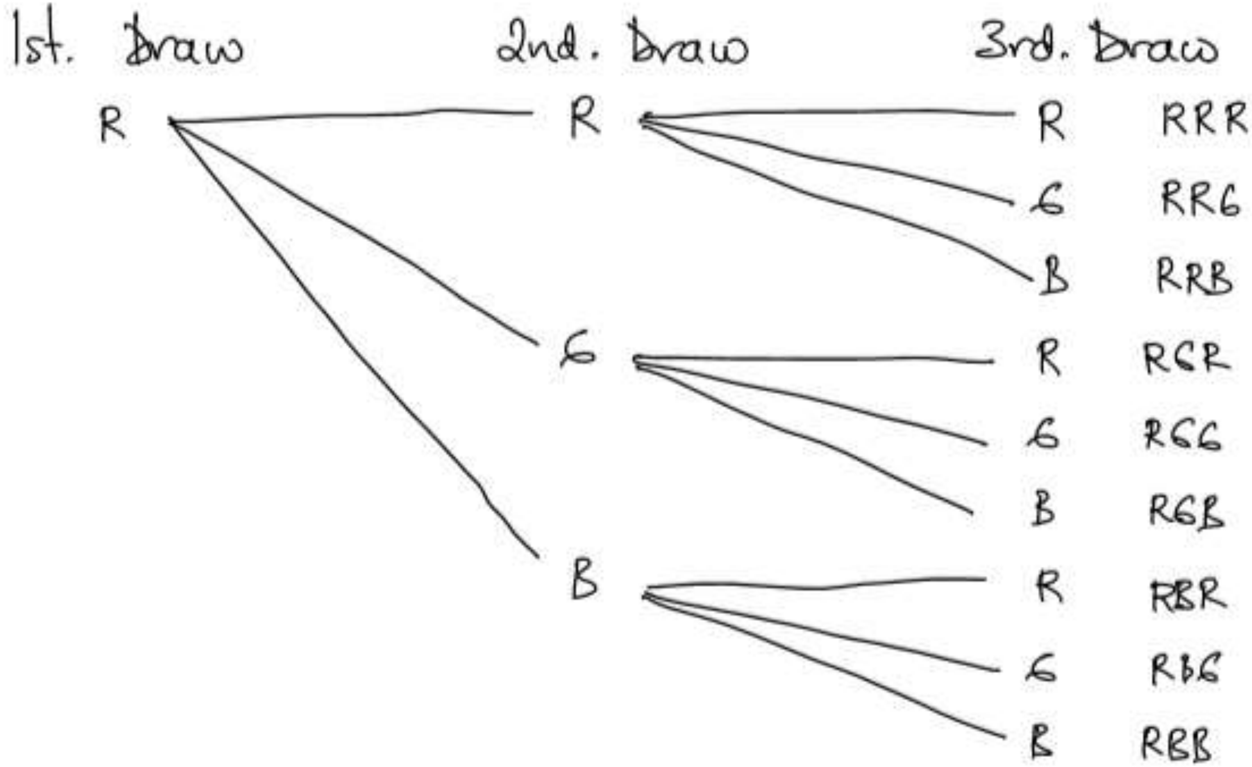
$$S = \{RGB, RBG, GRB, GBR, BRG, BGR\}$$

$\#S = 6$.

If instead we draw with replacement

$$S = \{RRR, RRG, RRB, RGR, \cdots\}$$

$\#S = 27$. We can use a tree to generate all the possible outcomes as follows.


1st. Draw
2nd. Draw
3rd. Draw
R
R
G
B
R
G
B
R
G
B
R
G
B
RRR
RRG
RRB
RGR
RGG
RGB
RBR
RBG
RBB

etc.

Combinatorics (counting):

Let A be a set containing n elements. A permutation of A is an arrangement of the elements of A.

Eg. $A = \{1, 2, 3\}$ then $3, 2, 1$ is a permutation of A.

There are $n!$ permutations of a set with n elements. ie

$$n! = n(n-1)(n-2)\cdots 3(2)(1)$$

How many ways can r elements be chosen from n if the order matters. The ans. is

$$n(n-1)(n-2)\cdots(n-r+1) =$$

$$\frac{n(n-1)(n-2)\cdots(n-r+1)(n-r)(n-r-1)\cdots 3(2)(1)}{(n-r)(n-r-1)\cdots(3)(2)(1)} =$$

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

This is sometimes denoted by ${}^nP_r$. (No. of permutations of r things from n things).

How many ways can r elements be chosen from n if the order doesn't matter? For a particular choice of r elements there are $r!$ permutations all of which we consider the same. We call such a choice a <u>combination</u>. The no. of combinations of r things chosen from n things is denoted ${}^nC_r$ and is

$$ {}^nC_r = \frac{{}^nP_r}{r!} = \frac{n!}{r!\,(n-r)!} $$

Eg. How many ways can 4 people be arranged in a row? Ans: $4! = 24$.

Eg. How many ways can a deck of 52 cards be arranged? Ans: $52!$ (68 digit no.).

Infinite Sets: There two kinds of infinite set – countable and uncountable. A set is countable if it can be placed in 1–1 (one-to-one) correspondence with the positive integers. That is, there is a func. which assigns a unique pos. int. to every element of the set. A set which is not countable is said to be uncountable.

Sets which are countable include $\mathbb{Z}^+$, $\mathbb{Z}$, $\mathbb{Q}$ (ie the rational nos.). Uncountable sets include $\mathbb{Q}'$ (the irrational nos.) and $\mathbb{R}$.

<u>Events</u>: An event E is a subset of the sample space S for a random experiment, to which a probability can be assigned. If a sample space is finite or countable then every subset is an event. If the sample space is uncountable then certain subsets may not be events, ie they cannot be assigned a probability. We only consider <u>discrete</u> sample spaces (ie finite or countable).

Eg. Roll a fair die, then $S = \{1, 2, 3, 4, 5, 6\}$ and we can define some events as

$$A = \{2, 4, 6\}$$

this event is the event that the outcome of rolling the die (our rand. exp.) is an even no.

$$B = \{1, 5\}$$
$$C = \emptyset$$
$$D = S$$

Eg. Toss a fair coin twice. $S = \{HH, HT, TH, TT\}$, $\#S = 4$.
An event could be that we get at least one head, ie

$$\{HT, TH\}$$

Eg. Roll two fair 6-sided dice. We can calc. the sample space as follows:

2nd die

1st. die	1	2	3	4	5	6
1	11	12	13	14	15	16
2	21	22	23	24	25	26
3	31	32	33	34	35	36
4	41	42	43	44	45	46
5	51	52	53	54	55	56
6	61	62	63	64	65	66

$$S = \{11, 12, \cdots, 16,\ 21, 22, \cdots, 26,\ \vdots\ 61, 62, \cdots, 66\}$$

$$\#S = 36$$

Consider the events:

A = event both nos. are even

B = event both nos. are the same

C = event the sum is less than 4

D = event the sum is 7

$A = \{22, 24, 26, 42, 44, 46, 62, 64, 66\}$ $\quad \# A = 9$

$B = \{11, 22, 33, 44, 55, 66\}$ $\quad \# B = 6$

$C = \{11, 12, 21\}$ $\quad \# C = 3$

$D = \{16, 25, 34, 43, 52, 61\}$ $\quad \# D = 6$

An event can be the null set (ie no outcomes occur)

the entire sample space or just a single outcome.

We often need combinations of event. For eg. suppose in the prev. eg. the interest is that the numbers rolled are both even and their sum is greater than 6. This is a combination of two events

E = event both nos. are even

F = event the sum is greater than 6

In this case the overall event K will only occur if both E and F occur. That is, the outcome of our experiment must be in both the set E and F ie

$$K = E \cap F$$

$E = \{22, 24, 26, 42, 44, 46, 62, 64, 66\}$ $\quad \# E = 9$

$F = \{16, 25, 26, 34, 35, 36, 43, 44, 45, 46, 52, 53, 54, 55, 56, 61, 62, 63, 64, 65, 66\}$ $\quad \# F = 21$

$K = \{26, 44, 46, 62, 64, 66\}$ $\quad \# K = 6.$

Suppose D is the event that the outcome consists of two numbers, one of which is even and one of which is odd <u>or</u> whose sum is 7. The combined event occurs if either the 1st. event occurs or the second event occurs, and this corresponds to the union of the two events.

$$A = \text{one even, one odd} = \{12, 14, 16, 21, 23, 25, 32, 34, 36, 41, 43, 45, 52, 54, 56, 61, 65, 65\}$$

$$B = \text{sum is } 7 = \{16, 61, 25, 52, 34, 43\}$$

$$D = \text{one even, one odd } \underline{\text{or}} \text{ sum is } 7 = A \cup B$$
$$= \{12, 14, 16, 21, 23, 25, 32, 34, 36, 41, 43, 45, 52, 54, 56, 61, 63, 65\}.$$

Sometimes we're interested in an event <u>not</u> occurring. For eg. suppose F is the event that the outcome isn't 11 (ie snake-eyes). This event is the complement of the event G that the outcome is snake-eyes. That is

$$F = G^c$$

where G is the event the outcome is 11.

If A is a subset of S then A^c is everything in S that isn't in A.

Eg. $S = \{1, 2, 3, 4, 5, 6\}$. $A = \{1, 3\}$, $A^c = \{2, 4, 5, 6\}$.

A <u>probability space</u> consists of a sample space S, a collection of subsets $\mathcal{E}$ of S called events and a probability function $p : \mathcal{E} \longrightarrow \mathbb{R}$ with the following properties:

(i) $S \in \mathcal{E}$

(ii) $\phi \in \mathcal{E}$

(iii) if $A_1, A_2, A_3, \cdots$ is a countable collection of events, ie $A_i \in \mathcal{E}\ \forall i$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{E}$.

(iv) if $A \in \mathcal{E}$ then $A^c \in \mathcal{E}$.

(v) $p(A) \geq 0$, $\forall A \in \mathcal{E}$

(vi) $p(S) = 1$

(vii) If $A_1, A_2, A_3, \cdots$ are mutually exclusive events then $p(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} p(A_i)$ (countable additivity).

Note: $\bigcup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup A_3 \cup \cdots$

$$\sum_{i=1}^{\infty} A_i = A_1 + A_2 + A_3 + \cdots$$

Note: Two events are mutually exclusive if they cannot occur simultaneously, that is, they have no outcomes in common. Thus events A and B are mutually exclusive if $A \cap B = \emptyset$. A collection of events is pairwise mutually exclusive if any two events have no elements in common. Countable additivity requires pairwise mutual exclusivity. A collection of events can have element in common but not be pairwise mutually exclusive.

If A is any event then $p(A) = \dfrac{\#A}{\#S}$

ie $\quad p(A) = \dfrac{\text{no. of outcomes in } A}{\text{no. of outcomes in } S}$

where S is the sample space.

This defines a valid probability space with Σ being the power set of S ie the collection of all subsets of S.

Eg. Roll a die. What is the prob. the outcome is

(i) a 1?

(ii) a 1 or 2?

(iii) odd and less than 4?

(iv) not 4 or 5?

(v) even or less than 3?

Assuming the die is fair the random exp. is equi-probable with $S = \{1,2,3,4,5,6\}$.

(i) A = event the outcome is 1 $= \{1\}$

$\#A = 1$

$\#S = 6$

so $p(A) = p(\{1\}) = \frac{\#A}{\#S} = \frac{1}{6}$

(ii) $B = \{1,2\}$ is the event the outcome is a 1 or 2.

$\#B = 2$ so $p(B) = \frac{\#B}{\#S} = \frac{2}{6} = \frac{1}{3}$

(iii) $C = \{1,3,5\}$ is the event the outcome is odd.

$\#C = 3$ so $p(C) = \frac{\#C}{\#S} = \frac{3}{6} = \frac{1}{2}$.

$D = \{1,2,3\}$ is the event the outcome is less than 4.

$E = C \cap D$ is the event the outcome is odd and less than 4.

$E = \{1,3\}$, $\#E = 2$

$$p(E) = \frac{2}{6} = \frac{1}{3}$$

(iv) $F = \{1, 2, 3, 6\}$ is the event the outcome isn't 4 or 5.

$\#F = 4$ so $p(F) = \frac{4}{6} = \frac{2}{3}$

Sometimes it easier to calc. the prob. of the complementary event ie F^c in this case, and then use the fact that

$$p(F) = 1 - p(F^c)$$

[Follows since $F \cup F^c = S$ and $\#F + \#F^c = \#S$].

$F^c = \{4, 5\}$, $\#F^c = 2$, $p(F^c) = \frac{2}{6} = \frac{1}{3}$

and $p(F) = 1 - p(F^c) = 1 - \frac{1}{3} = \frac{2}{3}$.

(v) $H = \{2, 4, 6\}$ (outcome is even)

$G = \{1, 2\}$ (outcome is < 3)

$K = \{1, 2, 4, 6\} = H \cup G$ (or means union)

$\#K = 4$ so $p(K) = \frac{4}{6} = \frac{2}{3}$.

Eg. Roll two fair dice and add the numbers rolled.
What is the prob. of obtaining

(i) 2

(ii) 3

(iii) a num. less than 5

(iv) an even no.

(v) an even no. less than 6

(vi) a number less than or equal to 10.

At 1st. sight this doesn't seem equiprobable as there are more ways to obtain a total of 5 say than a total of 2. The sum of the numbers isn't the sample space here even though we want their probs.

The sample space is

$$S = \{11, 12, 13, \cdots, 16, \cdots, 66\}$$

ie

	1	2	3	4	5	6
1	11	12	13	14	15	16
2	21	22	23	24	25	26
3	31	32	33	34	35	36
4	41	42	43	44	45	46
5	51	52	53	54	55	56
6	61	62	63	64	65	66

$\#S = 36$

These outcomes are equiprobable.

(i) $A = \{(1,1)\}$ "total is a 2"

Note (i,j) indicates a roll of i on the 1st. die and j on the second.

$$p(A) = \frac{\#A}{\#S} = \frac{1}{36}$$

(ii) $B = \{(1,2), (2,1)\}$ "total is 3"

$$p(B) = \frac{2}{36} = \frac{1}{18}$$

(iii) $C = \{(1,1), (1,2), (2,1), (1,3), (2,2), (3,1)\}$ "total <5"

$$p(C) = \frac{6}{36} = \frac{1}{6}$$

(iv) To get an even total both nos. rolled must even or both odd.

There are 3 ways to get an even no. on the 1st. die times 3 ways to get an even no. on the 2nd. die. In total there are 9 ways to get two even nos. Similarly there are 9 ways to get two odd nos. There are thus 18 ways to get an even total for a prob. of $\frac{18}{36} = \frac{1}{2}$.

(v) $D = \{(1,1), (1,3), (2,2), (3,1)\}$ "total is even and < 6"

$$p(D) = \frac{4}{36} = \frac{1}{9}$$

(vi) This is easiest to do using complements.

E = "total is less than or equal to 10"

E^c = "total is greater than 10"

$= \{(5,6), (6,5), (6,6)\}$

$p(E^c) = 3 = \underline{1}$

$$p(E) = 1 - \frac{1}{12} = \frac{11}{12}$$

Properties of Probability Spaces:

(i) If $A \cap B = \emptyset$ (ie A and B are mutually exclusive) then

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

(ii) $P(A^c) = 1 - P(A)$ for all events A.

(iii) $A \subseteq B \Rightarrow P(A) \leq P(B)$

The converse isn't true ie $P(A) \leq P(B)$ doesn't imply $A \subseteq B$.

(iv) The inclusion-exclusion principle:

For all events A, B

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

Eg. Roll two fair dice. Define events:

A: sum is 8

B: both nos. are even.

What is the prob. the outcome is an 8 or even?

A, B are not mut. exc. (mutually exclusive).

$\#S = 36$

$\#A = 5$ since $A = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$

$\#B = 9$ since $B = \{(2,2), (2,4), (2,6), \cdots, (6,6)\}$

$$p(A) = \frac{5}{36}, \quad p(B) = \frac{9}{36} = \frac{1}{4}$$

$p(A \cup B) = ?$

$A \cap B = \{(2,6), (4,4), (6,2)\}$ $\quad \#A \cap B = 3$ so $p(A \cap B) = \frac{3}{36} = \frac{1}{12}$

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$
$$= \frac{5}{36} + \frac{9}{36} - \frac{3}{36} = \frac{11}{36}.$$

De Morgan's Laws:

For any two sets A, B:

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

Eg. A card is drawn at random from a shuffled 52 card deck. What is the probability the card is a 10 or a spade.

Define events:

T = card is a 10

S = card is a spade

Not mutually exc. so $P(T \cup S) = p(T) + p(S) - p(T \cap S)$

$p(T) = 4/52 = 1/13$

$p(S) = 13/52 = 1/4$

$T \cap S$ = ten of spades $\Rightarrow P(T \cap S) = 1/52$

$$P(T \cup S) = \frac{1}{13} + \frac{1}{4} - \frac{1}{52} = \frac{4}{13}$$

What is the prob. the card is neither a four nor a club?

Define events:

$F =$ card is a 4 $\Rightarrow F^c =$ card is not a 4

$C =$ card is a club $\Rightarrow C^c =$ card is not a club

Neither a 4 nor a club $= F^c \cap C^c$.

Using DeMorgan

$$F^c \cap C^c = (F \cup C)^c$$

and therefore $P(F^c \cap C^c) = P((F \cup C)^c)$

$$= 1 - P(F \cup C)$$

F and C are not mutually exclusive so

$$P(F \cup C) = P(F) + P(C) - P(F \cap C)$$

$P(F) = 4/52 = 1/13$

$P(C) = 13/52 = 1/4$

$P(F \cap C) = 1/52$

so

$$P(F \cup C) = 1/13 + 1/4 - 1/52 = 4/13$$

Finally

$$P(F^c \cap C^c) = 1 - P(F \cup C)$$

$= 1 - 4/13$

$= 9/13$

Note: If A and B are mut. exc. events the inclusion-exclusion principle gives the correct result since $A \cap B = \emptyset$ and $p(\emptyset) = 0$, so

$$p(A \cup B) = p(A) + p(B) + 0$$

Conditional Probability:

Sometimes we'd like to know the prob. of an event B occurring given that we know that some event A has already occurred. This is called a conditional prob. denoted by

$$P(B|A)$$

read as "prob. of B given A".

If A and B are mutually exc. then $A \cap B = \emptyset$ and if A has occurred then B cannot and $p(B|A) = 0$. If $A = B$ then if A has occurred so has B and $p(B|A) = 1$. If $B \subset A$ then $p(B|A) = 1$ since if A has occurred so has B.

If $A \not\subset B$ and $B \not\subset A$ what is $P(B|A)$? Since A has occurred any outcomes in A^c are irrelevant as they cannot now occur and effectively as far as $P(B|A)$ is concerned we have a reduced sample space consisting only of the outcomes in A.
Note: Although we know A has occurred we do not know which outcome in A has occurred.

Eg. A bag contains 5 balls, 3 red and 2 blue. Two balls are drawn at random without replacement. What is the prob. the second ball is red given that the 1st. ball is red.

Since the 1st. ball is red there are now 2 red and 2 blue left in the bag so the prob. that the second is red is just $2/4$ ie $1/2$.

Eg. Roll two dice. Given that the no. on the 1st. die is even what is the probability the sum of the dice is ≤ 5.

Let A = the event the 1st. die is even.

$$A = \{ (2,1), (2,2), \cdots (2,6), \\ (4,1), (4,2), \cdots (4,6), \\ (6,1), (6,2), \cdots (6,6) \}$$

$\#A = 18$.

Since we know A occurred our sample space is now effectively just A. Let B = event the sum is ≤ 5. There are only 4 outcomes in A that lead to that result, ie $(2,1), (2,2), (2,3)$ and $(4,1)$. Thus

$$P(B|A) = \frac{4}{18} = \frac{2}{9}$$

Notice that the outcomes in A that give rise to B are just the outcomes in $A \cap B$ ie

$$p(B|A) = \frac{\#(A \cap B)}{\#A}$$

$$= \frac{\#(A \cap B)/\#S}{\#A/\#S}$$

$$= \frac{p(A \cap B)}{p(A)}$$

You can show this is always true, ie

$$p(B|A) = \frac{p(A \cap B)}{p(A)}$$

Eg. An urn contains 5 red and 3 blue balls. Two balls are drawn without replacement. What is the prob the second ball is red if the 1st. ball is red? What is the prob. if the 1st. was blue?

$S = \{(b_1, b_2) : b_1$ is 1st. drawn, b_2 is the 2nd.$\}$

For e.g. (B, R) indicates a blue ball 1st. and a red 2nd.

$\# S = 8(7) = 56$ (8 choices for 1st. and 7 for the second).

Let F = event 1st. ball is red.

S = " 2nd. ball is red.

$F = \{(R, b_2)\}$, $\#F = 5(7) = 35$

$S = \{(b_1, R)\}$, $\#S = 7(5) = 35$

$F \cap S$ = event both balls drawn are red. There are 5 choices for the 1st. red times 4 choices for the second.

$\#(F \cap S) = 5(4) = 20$

$$P(S|F) = \frac{P(S \cap F)}{P(F)} = \frac{20/56}{35/56} = \frac{20}{35} = \frac{4}{7}$$

Exercise: Show the prob. of a red second if a blue is 1st. is $5/7$.

We define the conditional probability

$$p(B|A) := \frac{P(A \cap B)}{P(A)}$$

provided $p(A) \neq 0$.

Eg. In a class, 25% failed maths, 15% failed chemistry and 10% failed both. A student is selected at random from the class. Find the prob.

(i) he faild maths if he failed chemistry

(ii) he failed chemistry if he failed maths

(iii) he failed maths or chemistry

(iv) he failed neither maths nor chemistry

Define events: M = failed maths

C = failed chemistry

$P(M) = 0.25$

$P(C) = 0.15$

$P(M \cap C) = 0.10$

(i) $P(M|C) = \dfrac{P(M \cap C)}{P(C)} = \dfrac{0.10}{0.15} = \dfrac{2}{3}$

(ii) $P(C|M) = \dfrac{p(C \cap M)}{p(M)} = \dfrac{0.10}{0.25} = \dfrac{2}{5}$

(iii) Failed maths or chemistry $= M \cup C$ and

$$p(M \cup C) = p(M) + p(C) - p(M \cap C)$$
$$= 0.25 + 0.15 - 0.10 = 0.30$$

(iv) Failed neither maths nor chemistry $= M^c \cap C^c$

Using De Morgan

$$p(M^c \cap C^c) = p((M \cup C)^c)$$
$$= 1 - p(M \cup C)$$
$$= 1 - 0.3$$
$$= 0.7$$

Multiplication Theorem: Rewriting the rule for conditional prob. gives

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

Eg. An urn contains 5 red and three blue balls. Two balls are drawn at random without replacement. What is the prob. both are red.

A = event the 1st. ball is red.
B = " " 2nd. " " "
C = " both are red = $A \cap B$

$p(C) = p(A \cap B) = P(A)P(B|A)$
$p(A) = 5/8$
$p(B|A) = 4/7$
then $p(C) = \frac{5}{8} \cdot \frac{4}{7} = \frac{5}{14}$

[Can also use $p(C) = \frac{{}^5C_2}{{}^8C_2} = \frac{5}{14}$]

Independence

<u>Def:</u> Two events A, B are <u>independent</u> if the occurrance or non-occurrance of one doesn't affect the prob. of the occurrance of the other.

Eg. Flip a coin twice. The results of the two flips are independent because what is obtained on the second flip doesn't depend on the outcome of the 1st. or vice-versa.

B is independent of A if p(B) and p(B|A) are the same. That is

$$p(B) = p(B|A) = \frac{p(A \cap B)}{p(A)}$$

ie if ind. $\quad p(A \cap B) = p(A)p(B)$

We now adopt this as our def. of independance.

Def. Two events A, B are independent if and only if

$$p(A \cap B) = p(A)\,p(B).$$

Note: We often stipulate that certain events are ind., in which case you can assume $p(A \cap B) = p(A)p(B)$.

Example: Flip a coin 3 times. Define events:

A = 1st. flip is a head

B = 2nd. " " " "

C = two heads in a row = $\{HHT, THH\}$.

$p(A) = 1/2$ $\{HHH, HHT, HTH, HTT\}$

$p(B) = 1/2$ $\{HHH, HHT, THH, THT\}$

$p(C) = 2/8 = 1/4$

$p(A \cap B) = 2/8 = 1/4$

$p(A \cap C) = 1/8$

$p(B \cap C) = 2/8 = 1/4$

$$p(A)p(B) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4} = P(A \cap B) \Rightarrow A \text{ and } B \text{ are ind.}$$

$$p(A)p(C) = \left(\frac{1}{2}\right)\left(\frac{1}{4}\right) = \frac{1}{8} = P(A \cap C) \Rightarrow A \text{ and } C \text{ are ind.}$$

$$p(B)p(C) = \left(\frac{1}{2}\right)\left(\frac{1}{4}\right) = \frac{1}{8} \neq P(B \cap C) \Rightarrow B \text{ and } C \text{ are } \underline{\text{not}} \text{ ind.}$$

Events which are not ind. are <u>dependent</u>.

Eg. A study relating smoking and cancer found the following probs.

	Gets Cancer	Doesn't get Cancer
Smokers	0.5	0.2
Non-smokers	0.1	0.2

A = an individual selected at random is a smoker

B = an individual gets cancer.

Are A and B independent?

$p(A) = 0.5 + 0.2 = 0.7$

$p(B) = 0.5 + 0.1 = 0.6$

$p(A \cap B) = 0.5$

$p(A \cap B) \neq P(A)P(B) = (0.7)(0.6) = 0.42$

Thus A and B are not independent ie whether the individual is a smoker or not affects the prob. they get cancer.

Random Variable: This is a misnomer since a random variable (RV) is neither random nor a variable.

Def. A random var. X on a sample space S is a func. from S to $\mathbb{R}$ ie

$$X : S \longrightarrow \mathbb{R}.$$

It maps outcomes to real nos.

A RV is discrete if it's range of values in $\mathbb{R}$ is finite or countable.

The Binomial RV or Distribution:

Def. A Bernoulli trial is a random experiment with two outcomes, success and failure. If the prob. of success is p then prob. of failure is $1-p$.

Eg. An archer takes 3 shots at a target. Define the random var. X to be the no. of hits. Then X is a random variable with possible values 0, 1, 2, or 3.

For a sequence of n Bernoulli trials we define a RV X which is the number of successes. Such a RV is said to be Binomial. If the prob. of a success is p then the prob. of exactly r successes in n Bernoulli trials is

$$P(X=r) = {}^nC_r\, p^r (1-p)^{n-r}$$

where ${}^nC_r$ is the number of ways of choosing r items from n items (ignoring order).

Eg. A photocopier broke down 17 times over 100 days. The copier is then checked daily for 4 days. Find

(a) The prob. it is broken on one of the days

(b) " " " " " " two days.

(c) " " " " " on at least one day.

(d) " " " " " " at most 2 days.

In this case a 'trial' is whether it's working on a given day. Assuming independence the no. of days on which it's not working is a binomial RV. We're conducting 4 Bernoulli trials so $n=4$. "Success" = not working on a given day and $p = 17/100 = 0.17$. Let X = no. of days when it's not working, then X takes values 0, 1, 2, 3 or 4.

(a) $P(X=1) = {}^4C_1 (0.17)^1 (0.83)^{4-1} = 4(0.17)(0.83)^3 = 0.389$

(note: $p = 0.17$, $1-p = 1-0.17 = 0.83$)

(b) $P(X=2) = {}^4C_2 (0.17)^2 (0.83)^{4-2}$

$= 6(0.17)(0.83)^2$

$= 0.119$

(c) The prob. it's broken on at least one of the days is 1 minus the prob. it's not broken on any day. Thus

$$P(X \geq 1) = 1 - P(X=0)$$

$$P(X=0) = {}^4C_0 p^0 (1-p)^{n-0}$$

$$= (0.83)^4$$

$$= 0.475$$

$$P(X \geq 1) = 1 - 0.475$$

$$= 0.525$$

(d) Want $P(X \leq 2)$. This corresponds to either $X=0$ or $X=1$ or $X=2$ which are mutually exclusive so the prob. is the sum of these

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

$$= 0.475 + 0.389 + 0.119$$

$$= 0.983$$

Eg. A biased coin has a 60% chance of coming up heads. Suppose the coin is flipped twice and let X be the RV consisting of the number of heads. So $X = 0, 1$ or 2.

Let S = success = a head, F = failure = a tail and assume each flip is an independent trial. Then X is a binomial random variable with parameters $n = 2$ and $p = 3/5$.

Calculate $P(X=0)$, $P(X=1)$, $P(X=2)$. What is the prob. of.

(i) At least one head

(ii) Zero or one head.

$$p(X=0) = {}^2C_0 (3/5)^0 (2/5)^2 = 4/25$$
$$p(X=1) = {}^2C_1 (3/5)^1 (2/5) = 12/25$$
$$p(X=2) = {}^2C_2 (3/5)^2 (2/5)^0 = 9/25$$

Note that $P(X=0)+P(X=1)+P(X=2) = 1$.

(i) $P(X \geq 1) = P(X=1) + P(X=2)$
$= 12/25 + 9/25 = 21/25$

(ii) $P(X \leq 1) = P(X=0) + P(X=1)$
$= 4/25 + 12/25 = 16/25$

Can also use $P(X \leq 1) = 1 - P(X=2) = 1 - 9/25 = 16/25$.