	T, 0.08 given	Y, 0.16 given	M, 0.45 given	S, 0.31 given
At least one collision	$P(C\|T)$ $=0.15$ given	$P(C\|Y)$ $=0.08$ given	$P(C\|M)$ $=0.04$ given	$P(C\|S)$ $=0.05$ given
	$P(C\cap T)$ $=0.15\times 0.08$ $=0.012$	$P(C\cap Y)$ $=0.08\times 0.16$ $=0.0128$	$P(C\cap M)$ $=0.04\times 0.45$ $=0.018$	$P(C\cap S)$ $=0.05\times 0.31$ $=0.0155$

$P(\text{at least one Collision}) = P(C) = P(C\cap T)+P(C\cap Y)+P(C\cap M)+P(C\cap S)$
$= .012 + .0128 + .018 + .0155 = .0583$.

$P(\text{young adult}|\text{at least one collision}) = P(Y|C) = \frac{P(Y\cap C)}{P(C)} = \frac{0.0128}{0.0583} = 0.2196.$ Answer: D

24. R_1, R_2 and R_3 denote the events that the 1st, 2nd and 3rd ball chosen is red, respectively.
$P(R_3\cap R_2\cap R_1) = P(R_3|R_2\cap R_1)\times P(R_2\cap R_1)$
$= P(R_3|R_2\cap R_1)\times P(R_2|R_1)\times P(R_1) = 1\times\frac{5}{11}\times\frac{5}{10} = \frac{5}{22}$. Answer: D

25. From the given information, 400 of those surveyed are both hockey and lacrosse fans, 200 are lacrosse fans and not hockey fans, and 400 are hockey fans an not lacrosse fans. This is true because there are 1000 fans in the survey, but a combined total of $800+600=1400$ sports preferences, so that 400 must be fans of both. Of the 600 lacrosse fans, 400 are also hockey fans, so 200 are not hockey fans. The probability that a Canadian sports fans is not a hockey fan given that she/he is a lacrosse fan is $\frac{200}{600}=\frac{1}{3}$. Answer: C

26. Suppose there are B blue balls in urn II.
$P[\text{both balls are same color}] = P[\text{both blue} \cup \text{both red}] = P[\text{both blue}] + P[\text{both red}]$
(the last equality is true since the events "both blue" and "both red" are disjoint).

$P[\text{both blue}] = P[\text{blue from urn I} \cap \text{blue from urn II}]$
$= P[\text{blue from urn I}]\times P[\text{blue from urn II}]$ (choices from the two urns are independent)
$= \frac{6}{10}\times\frac{B}{16+B}$,

$P[\text{both red}] = P[\text{rd from urn I} \cap \text{red from urn II}]$
$= P[\text{red from urn I}]\times P[\text{red from urn II}] = \frac{4}{10}\times\frac{16}{16+B}$,

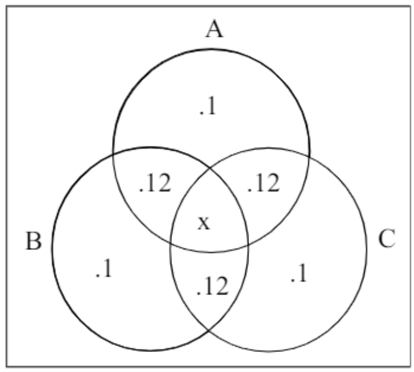
We are given $\frac{6}{10}\times\frac{B}{16+B}+\frac{4}{10}\times\frac{16}{16+B} = 0.44 \rightarrow \frac{6B+64}{10(16+B)} = 0.44 \rightarrow B=4$. Answer: A

27. We are given $P[A \cap B' \cap C'] = P[A' \cap B \cap C'] = P[A' \cap B' \cap C] = 0.1$ (having exactly one risk factor means not having either of the other two). We are also given

$$P[A \cap B \cap C'] = P[A \cap B' \cap C] = P[A' \cap B \cap C] = 0.12.$$

And we are given $P[A \cap B \cap C | A \cap B] = \frac{1}{3}$. We are asked to find $P[A' \cap B' \cap C' | A']$.
From $P[A \cap B \cap C | A \cap B] = \frac{1}{3}$ we get $\frac{P[A \cap B \cap C]}{P[A \cap B]} = \frac{1}{3}$, and then
$P[A \cap B \cap C] = \frac{1}{3} \times P[A \cap B]$.

The following Venn diagram illustrates the situation:



We see that $P[A \cap B \cap C] = x$ and $P[A \cap B] = x + 0.12$, so that

$$x = \frac{1}{3} \times (x + 0.12) \rightarrow x = P[A \cap B \cap C] = 0.06 .$$

Alternatively, we can use the rule $P[D] = P[D \cap E] + P[D \cap E']$ to get

$$P[A \cap B] = P[A \cap B \cap C] + P[A \cap B \cap C'] = P[A \cap B \cap C] + 0.12.$$

Then, $P[A \cap B] = P[A \cap B \cap C] + 0.12 = \frac{1}{3} \times P[A \cap B] + 0.12 \rightarrow P[A \cap B] = 0.18$
and $P[A \cap B \cap C] = \frac{1}{3} \times 0.18 = 0.06$.

We can also see from the diagram that $P[A \cap B'] = 0.1 + 0.12 = 0.22$.

Alternatively, we can use the rule above again to get

$$P[A \cap B'] = P[A \cap B' \cap C] + P[A \cap B' \cap C'] = 0.12 + 0.1 = 0.22.$$

Then, $P[A] = P[A \cap B] + P[A \cap B'] = 0.18 + 0.22 = 0.4$, and $P[A'] = 1 - P[A] = 0.6$.

We are asked to find $P[A' \cap B' \cap C' | A'] = \frac{P[A' \cap B' \cap C']}{P[A']} = \frac{P[A' \cap B' \cap C']}{0.6}$, so we must find $P[A' \cap B' \cap C']$. From the Venn diagram, we see that

$$P[A' \cap B' \cap C'] = 1 - (0.1 + 0.1 + 0.1 + 0.12 + 0.12 + 0.12 + 0.06) = 0.28.$$

Finally, $P[A' \cap B' \cap C' | A'] = \frac{P[A' \cap B' \cap C']}{P[A']} = \frac{P[A' \cap B' \cap C']}{0.6} = \frac{0.28}{0.6} = 0.467$. Answer: C

For trial use only.
Content may not
represent the most
current version

SECTION 3 - COMBINATORIAL PRINCIPLES, PERMUTATIONS AND COMBINATIONS

Factorial notation:

$n!$ denotes the quantity $n(n-1)(n-2)\cdots 2\cdot 1$;

$0!$ is defined to be equal to 1.

Permutations:

(a) Given n distinct objects, the number of different ways in which the objects may be **ordered** (or **permuted**) is $n!$. For example, the set of 3 letters $\{a, b, c\}$ can be ordered in the following $3! = 6$ ways:

$$abc\,,\ acb\,,\ bac\,,\ bca\,,\ cab\,,\ cba\,.$$

We say that we are choosing an ordered subset of size k **without replacement** from a collection of n objects if after the first object is chosen, the next object is chosen from the remaining $n-1$, the next after that from the remaining $n-2$, etc. The number of ways of doing this is

$$\frac{n!}{(n-k)!} = n \times (n-1) \times \cdots \times (n-k+1)$$

and is denoted

$${}_nP_k \ \text{ or } \ P_{n,k} \ \text{ or } \ P(n,k)$$

Using the set $\{a, b, c\}$ again, the number of ways of choosing an ordered subset of size 2 is

$$\frac{3!}{(3-2)!} = \frac{6}{1} = 6 \text{ - } ab\,,\ ac\,,\ ba\,,\ bc\,,\ ca\,,\ cb$$

(b) Given n objects, of which n_1 are of Type 1, n_2 are of Type 2, . . . , and n_t are of Type t ($t \geq 1$ is an integer), and $n = n_1 + n_2 + \cdots + n_t$, the number of ways of ordering all n objects (where objects of the same Type are indistinguishable) is

$$\frac{n!}{n_1!\cdot n_2!\cdots n_t!}$$

which is sometimes denoted

$$\binom{n}{n_1\ n_2\ \cdots\ n_t}$$

For example, the set $\{a, a, b, b, c\}$ has 5 objects, 2 are a's (Type 1), 2 are b's (Type 2) and 1 is c (Type 3). According to the formula above, there should be $\frac{5!}{2!\cdot 2!\cdot 1!} = 30$ distinct ways of ordering the 5 objects. These are

$aabbc$, $aabcb$, $aacbb$, $ababc$, $abacb$, $abbac$, $abbca$, $abcab$, $abcba$, $acabb$, $acbab$, $acbba$,

$bbaac$, $bbaca$, $bbcaa$, $babac$, $babca$, $baabc$, $baacb$, $bacba$, $bacab$, $bcbaa$, $bcaba$, $bcaab$,

$caabb$, $cabab$, $cabba$, $cbaab$, $cbaba$, $cbbaa$

Combinations:

Given n distinct objects, the number of ways of choosing a subset of size $k \le n$ without replacement and without regard to the order in which the objects are chosen is

$\frac{n!}{k!\cdot(n-k)!}$, which is usually denoted $\binom{n}{k}$ (or ${}_nC_k$, $C_{n,k}$ or $C(n,k)$) and is read "n choose k". $\binom{n}{k}$ is also called a **binomial coefficient** (and can be defined for any real number n and non-negative integer k). Note that if n is an integer and k is a non-negative integer, then

$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}$, and $\binom{n}{0} = \binom{n}{n} = 1$, and

$\binom{n}{1} = \binom{n}{n-1} = n$, and $\binom{n}{k} = \binom{n}{n-k}$.

Using the set $\{a,b,c\}$ again, the number of ways of choosing a subset of size 2 without replacement is $\binom{3}{2} = \frac{3!}{2!\cdot(3-2)!} = 3$; the subsets are $\{a,b\}, \{a,c\}, \{b,c\}$. When considering combinations, the order of the elements in the set is irrelevant, so $\{a,b\}$ is considered the same combination as $\{b,a\}$. When considering permutations, the order is important, so $\{a,b\}$ is a different permutation from $\{b,a\}$.

The name "binomial coefficient" arises from the fact that these factors appear as coefficients in a "binomial expansion". For instance,

$$\begin{aligned}(x+y)^4 &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4 \\ &= \tbinom{4}{0}x^4y^0 + \tbinom{4}{1}x^{4-1}y^1 + \tbinom{4}{2}x^{4-2}y^2 + \tbinom{4}{3}x^{4-3}y^3 + \tbinom{4}{4}x^{4-4}y^4\end{aligned}$$

A general form of this expansion is found in the binomial theorem.

Binomial Theorem:

In the power series expansion of $(1+t)^N$, the coefficient of t^k is

$\binom{N}{k}$, so that $(1+t)^N = \sum_{k=0}^{\infty} \binom{N}{k} \times t^k = 1 + Nt + \frac{N(N-1)}{2}t^2 + \frac{N(N-1)(N-2)}{6}t^3 + \cdots$

If N is an integer, then the summation stops at $k = N$ and the series is valid for any real number t, but if N is not an integer, then the series is valid if $|t| < 1$.

Multinomial Theorem:

Given n objects, of which n_1 are of Type 1, n_2 are of Type 2, . . . , and n_t are of Type t ($t \ge 1$ is an integer), and $n = n_1 + n_2 + \cdots + n_t$, the number of ways of choosing a subset of size $k \le n$ (without replacement) with k_1 objects of Type 1, k_2 objects of Type 2,..., and k_t objects of Type t, where $k = k_1 + k_2 + \cdots + k_t$ is $\binom{n_1}{k_1} \times \binom{n_2}{k_2} \times \cdots \times \binom{n_t}{k_t}$. A general form of the relationship is found in the multinomial theorem

In the power series expansion of $(t_1 + t_2 + \cdots + t_s)^N$ where N is a positive integer, the coefficient of $t_1^{k_1} \times t_2^{k_2} \times \cdots \times t_s^{k_s}$ (where $k_1 + k_2 + \cdots + k_s = N$)

is $\binom{N}{k_1\ k_2\ \cdots\ k_s} = \frac{N!}{k_1!\cdot k_2!\cdots k_s!}$. For example, in the expansion of $(1+x+y)^4$, the coefficient of xy^2 is the coefficient of $1^1x^1y^2$, which is $\binom{4}{1\ 1\ 2} = \frac{4!}{1!\times 1!\times 2!} = 12$.

Important Note

In questions involving coin flips or dice tossing, it is understood, unless indicated otherwise successive flips or tosses are independent of one another. Also, in making a random selection of an object from a collection of n objects, it is understood, unless otherwise indicated, that each object has the same chance of being chosen, which is $\frac{1}{n}$. In questions that arise involving choosing k objects at random from a total of n objects, or in constructing a random permutation of a collection of objects, it is understood that each of the possible choices or permutations is equally likely to occur. For instance, if a purse contains one quarter, one dime, one nickel and one penny, and two coins are chosen, there are $\binom{4}{2} = 6$ possible ways of choosing two coins without regard to order of choosing; these are

Q-D , Q-N , Q-P , D-N , D-P , N-P

(the choice Q-D is regarded as the same as D-Q, etc.). It would be understood that each of the 6 possible ways are equally likely, and each has (uniform) probability of $\frac{1}{6}$ of occurring; the probability space would consist of the 6 possible pairs of coins, and each sample point would have probability $\frac{1}{6}$. Then, the probability of a particular event occurring would be $\frac{j}{6}$, where j is the number of sample points in the event. If A is the event "one of the coins is either a quarter or a dime", then $P[A] = \frac{5}{6}$, since event A consists of the 5 of the sample points

{Q-D , Q-N , Q-P , D-N , D-P }

Example 3-1:

An ordinary die and a die whose faces have 2, 3, 4, 6, 7, 9, dots are tossed independently of one another, and the total number of dots on the two dice is recorded as N. Find the probability that $N \geq 10$.

Solution:

It is assumed that for each die, each face has a $\frac{1}{6}$ probability of turning up. If the number of dots turning up on die 1 and die 2 are d_1 and d_2, respectively, then the tosses that result in $N = d_1 + d_2 \geq 10$ are

$$(1,9),\ (2,9),\ (3,7),\ (3,9),\ (4,6),\ (4,7), (4,9),\ (5,6),\ (5,7),\ (5,9),\ (6,4),\ (6,6),\ (6,7),\ (6,9)\ ,$$

14 combinations out of a total of $6 \times 6 = 36$ combinations that can possibly occur. Since each of the 36 (d_1, d_2) combinations is equally likely, the probability is $\frac{14}{36}$. □

Example 3-2:

Three nickels, one dime and two quarters are in a purse. In picking three coins at one time (without replacement), what is the probability of getting a total of at least 35 cents?

Solution:

In order to get at least 35 cents, at least one quarter must be chosen. The possible choices are $1Q +$ any 2 of the non-quarters, or $2Q +$ any 1 of the non-quarters. The total number of ways of choosing three coins from the six coins is $\binom{6}{3} = 20$.

If we label the two quarters as Q_1 and Q_2 , then the number of ways of choosing the three coins so that only Q_1 (and not Q_2) is in the choice is $\binom{4}{2} = 6$ (this is the number of ways of choosing the other two coins from the three nickels and one dime), and therefore, the number of choices that contain only Q_2 (and not Q_1) is also 6.

The number of ways of choosing the three coins so that both Q_1 and Q_2 are in the choice is 4 (this is the number of ways of choosing the other coin from the three nickels and one dime). Thus, the total number of choices for which at least one of the three coins chosen is a quarter is 16. The probability in question is $\frac{16}{20}$.

An alternative approach is to find the number of three coin choices that do not contain any quarters is $\binom{4}{3} = 4$ (the number of ways of choosing the three coins from the 4 non-quarters), so that number of choices that contain at least one quarter is $20 - 4 = 16$. □

Example 3-3:

A and B draw coins in turn without replacement from a bag containing 3 dimes and 4 nickels. A draws first. It is known that A drew the first dime. Find the probability that A drew it on the first draw.

Solution:

$P[A \text{ draws dime on first draw} | A \text{ draws first dime}] = \frac{P[A \text{ draws dime on first draw}]}{P[A \text{ draws first dime}]}$

$P[A \text{ draws dime on first draw}] = \frac{3}{7}$. Since there only 3 dimes, in order for A to draw the first dime, this must happen on A's first, second or third draw. Thus,

$P[A \text{ draws first dime}] = P[A \text{ draws dime on first draw}]$

$\quad + P[A \text{ draws first dime on second draw}] + P[A \text{ draws first dime on third draw}]$

$P[A \text{ draws dime on second draw}] = \frac{4}{7} \times \frac{3}{6} \times \frac{3}{5} = \frac{6}{35}$, since A's first draw is one of the four non-dimes, and B's first draw is one of the three remaining non-dimes after A's draw, and A's second draw is one of the three dimes of the five remaining coins. In a similar way,

$P[A \text{ draws first dime on third draw}] = \frac{4}{7} \times \frac{3}{6} \times \frac{2}{5} \times \frac{1}{4} \times 1 = \frac{1}{35}$.

Then, $P[A \text{ draws first dime}] = \frac{3}{7} + \frac{6}{35} + \frac{1}{35} = \frac{22}{35}$, and

$P[A \text{ draws dime on first draw} | A \text{ draws first dime}] = \frac{3/7}{22/35} = \frac{15}{22}$. □

Example 3-4:

Three people, X, Y and Z, in order, roll an ordinary die. The first one to roll an even number wins. The game continues until someone rolls an even number. Find the probability that X will win.

Solution:

Since X rolls first, fourth, seventh, etc. until the game ends, the probability that X will win is the probability that in throwing a die, the first even number will occur on the 1st, or 4th, or 7th, or . . . throw. The probability that the first even number occurs on the n-th throw is $(\frac{1}{2})^{n-1}(\frac{1}{2}) = \frac{1}{2^n}$. This is true since it requires $n-1$ odd throws followed by an even throw. Assuming independence of successive throws, with $A_i =$ "throw i is even", the probability that the first even throw occurs on throw n is

$$P[A_1' \cap A_2' \cap \cdots \cap A_{n-1}' \cap A_n] = P[A_1'] \times P[A_2'] \times \cdots \times P[A_{n-1}'] \times P[A_n] \;=\; (\tfrac{1}{2})^{n-1}(\tfrac{1}{2}) = \tfrac{1}{2^n}$$

Thus,

P[first even throw is on 1st, or 4th, or 7th, or ...] $= \frac{1}{2} + \frac{1}{2^4} + \frac{1}{2^7} + \cdots = \frac{1}{2}(1 + \frac{1}{8} + \frac{1}{8^2} + \cdots) = \frac{4}{7}$ □

Example 3-5:

Urn I contains 7 red and 3 black balls, and Urn II contains 4 red and 5 black balls. After a randomly selected ball is transferred from Urn I to Urn II, 2 balls are randomly drawn from Urn II without replacement. Find the probability that both balls drawn from Urn II are red.

Solution:

Define the following events:

R_1: the ball transferred from Urn I to Urn II is red

B_1: the ball transferred from Urn I to Urn II is black

R_2: two red balls are selected from Urn II after the transfer from Urn I to Urn II

Since R_1 and B_1 are mutually exclusive,

$$P[R_2] = P[R_2 \cap (R_1 \cup B_1)] = P[R_2 \cap R_1] + P[R_2 \cap B_1]$$
$$= P[R_2|R_1] \times P[R_1] + P[R_2|B_1] \times P[B_1] = \frac{\binom{5}{2}}{\binom{10}{2}} \cdot \frac{7}{10} + \frac{\binom{4}{2}}{\binom{10}{2}} \times \frac{3}{10} = \frac{44}{225}$$ □

Example 3-6:

A calculator has a random number generator button which, when pressed, displays a random digit $0, 1, ..., 9$. The button is pressed four times. Assuming that the numbers generated are independent of one another, find the probability of obtaining one "0", one "5", and two "9"'s in any order.

Solution:

There are $10^4 = 10,000$ four-digit orderings that can arise, from 0-0-0-0 to 9-9-9-9. From the notes above on permutations, if we have four digits, with one "0", one "5" and two "9"'s, the number of orderings is $\frac{4!}{1! \times 1! \times 2!} = 12$. The probability in question is then $\frac{12}{10,000}$. □

Example 3-7:

In Canada's national 6-49 lottery, a ticket has 6 numbers each from 1 to 49, with no repeats. Find the probability of matching all 6 numbers if the numbers are all randomly chosen. The ticket cost is \$2. If you match exactly 3 of the 6 numbers chosen, you win \$10. Find the probability of winning \$10.

Solution:

There are $\binom{49}{6} = \frac{49!}{6! \times 43!} = \frac{49 \times 48 \times 47 \times 46 \times 45 \times 44}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = 13,983,816$ possible combinations of 6 numbers from 1 to 49 (we are choosing 6 numbers from 1 to 49 without replacement), so the probability of matching all 6 numbers is $\frac{1}{13{,}983{,}816} = 0.0000000715112$

(about 1 in 14 million).

Suppose you have bought a lottery ticket. There are $\binom{6}{3} = 20$ ways of picking 3 numbers from the 6 numbers on your ticket. Suppose we look at one of those subsets of 3 numbers from your ticket. In order for the winning ticket number to match exactly those 3 of your 6 numbers, the other 3 winning ticket numbers must come from the 43 numbers between 1 and 49 that are not numbers on your ticket. There are $\binom{43}{3} = \frac{43 \times 42 \times 41}{3 \times 2 \times 1} = 12,341$ ways of doing that, and since there are 20 subsets of 3 numbers on your ticket, there are $20 \times 12,341 = 246,820$ ways in which the winning ticket numbers match exactly 3 of your ticket numbers. Since there are a total of 13,983,816 ways of picking 6 out of 49 numbers, your chance of matching exactly 3 of the winning numbers is $\frac{246{,}820}{13{,}983{,}816} = 0.01765$ (about $\frac{1}{57}$). So you have about a one in 57 chance of turning \$2 into \$10. □

Example 3-8:

In a poker hand of 5 cards from an ordinary deck of 52 cards, a "full house" is a hand that consist of 3 of one rank and 2 of another rank (such as 3 kings and 2 5's). If 5 cards are dealt at random from an ordinary deck, find the probability of getting a full house.

Solution:

There are $\binom{52}{5} = 2,598,960$ possible hands that can be dealt from the 52 cards. There are 13 ranks from deuce (2) to ace, and there are $\binom{13}{2} = 78$ pairs of ranks. For each pair of ranks, there are $\binom{4}{3} \times \binom{4}{2} = 24$ combinations consisting of 3 cards of the first rank and 2 cards of the second rank, and there are 24 combinations consisting of 2 cards of the first rank and 3 cards of the second rank, for a total of 48 possible full house hands based on those two ranks. Since there are 78 pairs of ranks, there are $78 \times 48 = 3744$ distinct poker hands that are a full house. The probability of being dealt a full house is $\frac{3744}{2{,}598{,}960} = 0.00144058$ (a little better chance than 1 in 700). □

PROBLEM SET 3
Combinatorial Principles

1. A class contains 8 boys and 7 girls. The teacher selects 3 of the children at random and without replacement. Calculate the probability that number of boys selected exceeds the number of girls selected.
A) $\frac{512}{3375}$ B) $\frac{28}{65}$ C) $\frac{8}{15}$ D) $\frac{1856}{3375}$ E) $\frac{36}{65}$

2. There are 97 men and 3 women in an organization. A committee of 5 people is chosen at random, and one of these 5 is randomly designated as chairperson. What is the probability that the committee includes all 3 women and has one of the women as chairperson?
A) $\frac{3(4!\,97!)}{2(100!)}$ B) $\frac{5!\,97!}{2(100!)}$ C) $\frac{3(5!\,97!)}{2(100!)}$ D) $\frac{3!\,5!\,97!}{100!}$ E) $\frac{3^3\,97^2}{100^5}$

3. A box contains 4 red balls and 6 white balls. A sample of size 3 is drawn without replacement from the box. What is the probability of obtaining 1 red ball and 2 white balls, given that at least 2 of the balls in the sample are white?
A) $\frac{1}{2}$ B) $\frac{2}{3}$ C) $\frac{3}{4}$ D) $\frac{9}{11}$ E) $\frac{54}{55}$

4. When sent a questionnaire, the probability is 0.5 that any particular individual to whom it is sent will respond immediately to that questionnaire. For an individual who did not respond immediately, there is a probability of 0.4 that the individual will respond when sent a follow-up letter. If the questionnaire is sent to 4 persons and a follow-up letter is sent to any of the 4 who do not respond immediately, what is the probability that at least 3 never respond?
A) $(0.3)^4 + 4(0.3)^3(0.7)$ B) $4(0.3)^3(0.7)$ C) $(0.1)^4 + 4(0.1)^3(0.9)$ D) $4(0.3)(0.7)^3 + (0.7)^4$
E) $(0.9)^4 + 4(0.9)^3(0.1)$

5. A box contains 35 gems, of which 10 are real diamonds and 25 are fake diamonds. Gems are randomly taken out of the box, one at a time without replacement. What is the probability that exactly 2 fakes are selected before the second real diamond is selected?
A) $\frac{225}{5236}$ B) $\frac{675}{5236}$ C) $\frac{\binom{25}{2}\binom{10}{2}}{\binom{35}{4}}$ D) $\binom{3}{2}(\frac{10}{35})^2(\frac{25}{35})^2$ E) $\binom{4}{2}(\frac{10}{35})^2(\frac{25}{35})^2$

6. (SOA) An insurance company determines that N the number of claims received in a week, is a random variable with $P[N=n] = \frac{1}{2^{n+1}}$, where $n \geq 0$. The company also determines that the number of claims received in a given week is independent of the number of claims received in any other week. Determine the probability that exactly seven claims will be received during a given two-week period.
A) $\frac{1}{256}$ B) $\frac{1}{128}$ C) $\frac{7}{512}$ D) $\frac{1}{64}$ E) $\frac{1}{32}$

7. Three boxes are numbered 1, 2 and 3. For $k = 1, 2, 3$, box k contains k blue marbles and $5 - k$ red marbles. In a two-step experiment, a box is selected and 2 marbles are drawn from it without replacement. If the probability of selecting box k is proportional to k, what is the probability that the two marbles drawn have different colors?

A) $\frac{17}{60}$ B) $\frac{34}{75}$ C) $\frac{1}{2}$ D) $\frac{8}{15}$ E) $\frac{17}{30}$

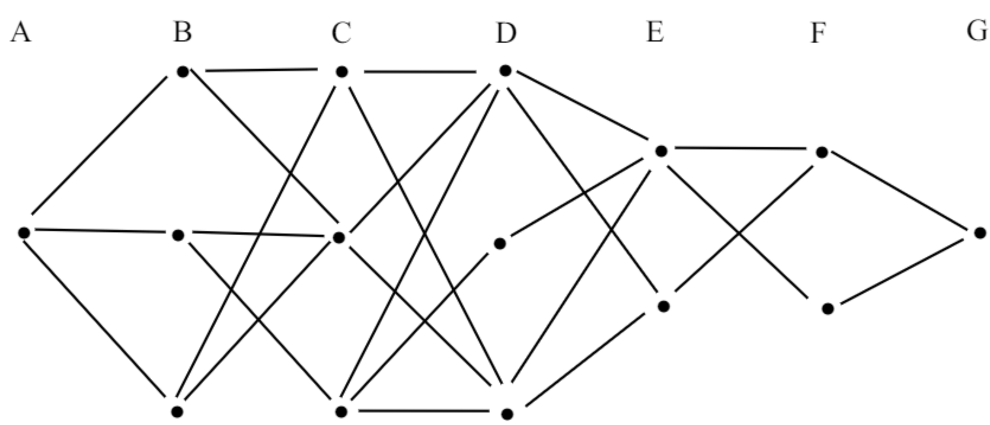
8. In Canada's national 6-49 lottery, a ticket has 6 numbers each from 1 to 49, with no repeats. Find the probability of matching exactly 4 of the 6 winning numbers if the winning numbers are all randomly chosen.

A) 0.00095 B) 0.00097 C) 0.00099 D) 0.00101 E) 0.00103

9. A number X is chosen at random from the series 2,5,8, ... and another number Y is chosen at random from the series 3,7,11, ... Each series has 100 terms. Find $P[X = Y]$.

A) 0.0025 B) 0.0023 C) 0.0030 D) 0.0021 E) 0.0033

10. In the following diagram, A, B,... refer to successive states through which a traveler must pass in order to get from A to G, moving from left to right. A path consists of a sequence of line segments from one state to the next. A path must always move to the next state until reaching state G. Determine the number of possible paths from A to G.



A) 30 B) 32 C) 34 D) 36 E) 38

11. (SOA) A store has 80 modems in its inventory, 30 coming from Source A and the remainder from Source B. Of the modems from Source A, 20% are defective. Of the modems from Source B, 8% are defective. Calculate the probability that exactly two out of a random sample of five modems from the store's inventory are defective.

A) 0.010 B) 0.078 C) 0.102 D) 0.105 E) 0.125

12. (SOA) Thirty items are arranged in a 6-by-5 array as shown.

A_1	A_2	A_3	A_4	A_5
A_6	A_7	A_8	A_9	A_{10}
A_{11}	A_{12}	A_{13}	A_{14}	A_{15}
A_{16}	A_{17}	A_{18}	A_{19}	A_{20}
A_{21}	A_{22}	A_{23}	A_{24}	A_{25}
A_{26}	A_{27}	A_{28}	A_{29}	A_{30}

Calculate the number of ways to form a set of three distinct items such that no two of the selected items are in the same row or column.

A) 200 B) 760 C) 1200 D) 4560 E) 7200

13. (SOA) From 27 pieces of luggage, an airline luggage handler damages a random sample of four. The probability that exactly one of the damaged pieces of luggage is insured is twice the probability that none of the damaged pieces are insured.

Calculate the probability that exactly two of the four damaged pieces are insured.

A) 0.06 B) 0.13 C) 0.27 D) 0.30 E) 0.31

PROBLEM SET 3 SOLUTIONS

1. There are $\binom{15}{3} = \frac{15!}{3!\times 12!} = \frac{15\times 14\times 13}{3\times 2\times 1} = 455$ ways of selecting 3 children from a group of 15 without replacement. The number of boys selected exceeds the number of girls selected if either
(i) 3 boys and 0 girls are selected, or (ii) 2 boys and 1 girl are selected.
There are $\binom{8}{3} \times \binom{7}{0} = \frac{8!}{3!\times 5!} \cdot \frac{7!}{0!\times 7!} = 56$ ways in which selection (i) can occur , and there are $\binom{8}{2} \times \binom{7}{1} = \frac{8!}{2!\times 6!} \times \frac{7!}{1!\times 6!} = 196$ ways in which selection (ii) can occur.
The probability of either (i) or (ii) occurring is $\frac{56+196}{455} = \frac{36}{65}$. Answer: E

2. Let A be the event that the committee has a woman as chairperson, and let B be the event that the committee includes all 3 women. Then, $P[A\cap B] = P[A|B] \times P[B]$.
The conditional probability $P[A|B]$ is equal to $\frac{3}{5}$ since the chairperson is chosen at random from the 5 committee members, and, given B, 3 of the committee members are women. There are $\binom{100}{5}$ ways of choosing a 5-member committee from the group of 100. Out of all 5-member committees, there are $\binom{97}{2}$ committees that include all 3 women (i.e., 2 men are chosen from the 97 men). Thus,
$$P[B] = \frac{\binom{97}{2}}{\binom{100}{5}} = \left(\frac{97!}{2!\,95!}\right)/\left(\frac{100!}{5!\,95!}\right) = \frac{5!\,97!}{2!\,100!}$$
and
$$P[A\cap B] = \frac{5!\,97!}{2!\,100!} \times \frac{3}{5} = \frac{3\cdot 4!\,97!}{2!\,100!}$$ Answer: A

3. $P[R, 2W|\text{at least } 2W] = \frac{P[R,2W]}{P[\text{at least } 2W]} = \frac{\binom{4}{1}\binom{6}{2}}{\binom{4}{1}\binom{6}{2}+\binom{4}{0}\binom{6}{3}} = \frac{3}{4}$. Answer: C

4. The probability that an individual will not respond to either the questionnaire or the follow-up letter is $0.5 \times 0.6 = 0.3$. The probability that all 4 will not respond to either the questionnaire or the follow-up letter is $(0.3)^4$.
$$\begin{aligned} P[\text{3 don't respond}] &= P[\text{1 response on 1st round, no additional responses on 2nd round}] \\ &\quad + P[\text{no responses on 1st round, 1 response on 2nd round}] \\ &= 4[(0.5)^4(0.6)^3] + 4[(0.5)^4(0.6)^3(0.4)] = 4(0.3)^3(0.7). \end{aligned}$$

Then, $P[\text{at least 3 don't respond}] = (0.3)^4 + 4(0.3)^3(0.7)$. Answer: A

5. Exactly 2 fakes must be picked in the first 3 picks and the second real diamond must occur on the 4th pick.

The possible ways in which this may occur are (F-fake, R-real)

$FFRR$ (prob. $\frac{25\times24\times10\times9}{35\times34\times33\times32}$) , $FRFR$ (prob. $\frac{25\times10\times24\times9}{35\times34\times33\times32}$) , $RFFR$ (prob. $\frac{10\times25\times24\times9}{35\times34\times33\times32}$)

The overall probability is $3\times\frac{25\times24\times10\times9}{35\times34\times33\times32}=\frac{675}{5236}$. Answer: B

6. The following combinations result in a total of 7 claims in a 2 week period:

Week 1 , Prob.	Week 2 - Prob.	Combined Probability
$0,\ \frac{1}{2}$	$7,\ \frac{1}{2^8}$	$\frac{1}{2}\times\frac{1}{2^8}=\frac{1}{2^9}$
$1,\ \frac{1}{2^2}$	$6,\ \frac{1}{2^7}$	$\frac{1}{2^2}\times\frac{1}{2^7}=\frac{1}{2^9}$
	$\vdots$	
$7,\ \frac{1}{2^8}$	$1,\ \frac{1}{2}$	$\frac{1}{2^8}\times\frac{1}{2}=\frac{1}{2^9}$

The total probability of exactly 7 claims in a two week period is $8\times\frac{1}{2^9}=\frac{1}{64}$. Answer: D

7. If the probability of selecting box 1 is p, then $p+2p+3p=1\to p=\frac{1}{6}$.

Then the probability in question is

$$\begin{aligned} &P[\text{2 different colors}|\text{box 1 selected}]\times P[\text{box 1 selected}]\\ &\quad+P[\text{2 different colors}|\text{box 2 selected}]\times P[\text{box 2 selected}]\\ &\quad+P[\text{2 different colors}|\text{box 3 selected}]\times P[\text{box 3 selected}]\\ &=\frac{1\cdot4}{\binom{5}{2}}\times\frac{1}{6}+\frac{2\cdot3}{\binom{5}{2}}\times\frac{2}{6}+\frac{3\cdot2}{\binom{5}{2}}\times\frac{3}{6}=\frac{34}{10\times6}=\frac{17}{30}.\end{aligned}$$

Answer: E

8. Suppose you have bought a lottery ticket. There are $\binom{6}{4}=15$ ways of picking 4 numbers from the 6 numbers on your ticket.

Suppose we look at one of those subsets of 4 numbers from your ticket. In order for the winning ticket number to match exactly those 4 of your 6 numbers, the other 2 winning ticket numbers must come from the 43 numbers between 1 and 49 that are not numbers on your ticket.

There are $\binom{43}{2}=\frac{43\times42}{2\times1}=903$ ways of doing that, and since there are 15 subsets of 4 numbers on your ticket, there are $15\times903=13,545$ ways in which the winning ticket numbers match exactly 4 of your ticket numbers. Since there are a total of 13,983,816 ways of picking 6 out of 49 numbers, your chance of matching exactly 4 of the winning numbers is $\frac{13,545}{13,983,816}=0.00096862$. Answer: B

9. There are $100^2=10,000$ equally likely possible choices of (X,Y).

Of these choices, the pairs that equal X and Y are

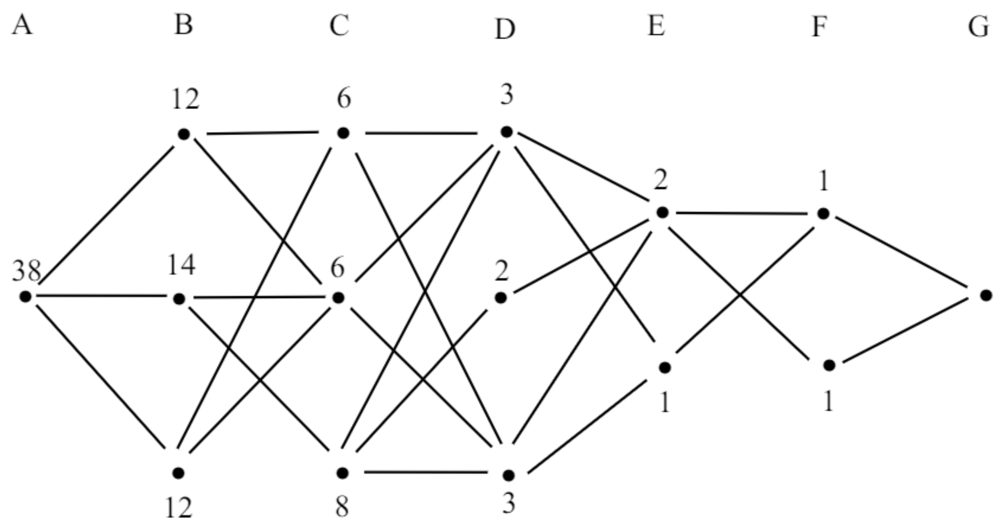
$$(11,11)\text{-}1\,,\ (23,23)\text{-}2\,,\ (35,35)\text{-}3\,,\ \ldots,\ (299,299)\text{-}25$$

(they are of the form $(12k-1,12k-1)$. The probability is $\frac{25}{10,000}$. Answer: A

10. This problem can be solved by a "backward induction" on the diagram. At each node we find the number of paths from that node to state G. We first apply backward induction to the two nodes in state F. At the upper node there is 1 path to G and at the lower node there is 1 path to G.

Then we look at the notes in state E and look at the next segments that can be taken. We see that there are $1+1=2$ possible paths from the upper node at F to G and 1 possible path from the lower node.

We continue in this way at state D. From the top node of state D there are $2+1=3$ paths to state G, from the middle node of state D there are paths, and from the lower node there are 3 paths. Continuing in this way back to state A, there will be a total of 38 paths from state A. The diagram below indicates the number of paths to state G from each node.



11. The probability is $\frac{\text{number of ways of choosing 2 defective and 3 non-defective}}{\text{number of ways of choosing 5 modems}}$.

There are a total of $0.2\times 30+0.08\times 50=10$ defective modems in total.

The number of ways of choosing 5 modems at random from the 80 modems is $\binom{80}{5}$.

The number of ways of choosing 2 defective and 3 non-defective is $\binom{10}{2}\times\binom{70}{3}$, since there are 10 defective and 70 non-defective. The probability is $\frac{\binom{10}{2}\times\binom{70}{3}}{\binom{80}{5}}=0.102$.

Answer: C

12. There are $\binom{5}{3} = 10$ ways to choose three of the five columns. Suppose that the columns chosen are labeled C_1, C_2 and C_3. There are 6 ways of choosing an item from C_1, then there are 5 ways of choosing an item from C_2 that is not in the same row as that chosen from C_1, and then there are 4 ways of choosing an item from C_3 that is not in either of the rows of the items chosen from C_1 and C_2. This results in $10 \times 6 \times 5 \times 4 = 1200$ was of choosing the three items.
We could have picked 3 of the 6 rows first. There are $\binom{6}{3} = 20$ ways of doing this. Then there would be $5 \times 4 \times 3 = 60$ ways of choosing an item from the first row chosen, another item from the second row chosen that is not in the same row as the first item, and an item from the third row chosen that is not in the same row as the first two. Total number of ways of choosing the three items is $20 \times 5 \times 4 \times 3 = 1200$, same as before. Answer: C

13. Suppose that n of the original pieces of luggage are insured. Then the probability that exactly one of the damaged pieces is insured is $\frac{\binom{n}{1} \times \binom{27-n}{3}}{\binom{27}{4}}$ and the probability that none of the damaged pieces are insured is $\frac{\binom{27-n}{4}}{\binom{27}{4}}$. We are given that $\frac{\binom{n}{1} \times \binom{27-n}{3}}{\binom{27}{4}} = 2 \times \frac{\binom{27-n}{4}}{\binom{27}{4}}$,
or equivalently, $\frac{n \times (27-n) \times (26-n) \times (25-n)}{6} = 2 \times \frac{(27-n) \times (26-n) \times (25-n) \times (24-n)}{24}$.
This simplifies to $\frac{n}{6} = 2 \times \frac{24-n}{24}$, from which we get $n = 8$.
The probability that exactly two of the damaged pieces are insured is
$\frac{\binom{8}{2} \times \binom{19}{2}}{\binom{27}{4}} = \frac{\frac{8 \times 7}{2} \times \frac{19 \times 18}{2}}{\frac{27 \times 26 \times 25 \times 24}{24}} = 0.273$. Answer: C

SECTION 4 - RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

Random variable X:

The formal definition of a random variable is that it is a function on a probability space S. This function assigns a real number $X(s)$ to each sample point $s \in S$. The less formal, but more typical way to describe a random variable is to describe the possible values that can occur and the probabilities of those values occurring. It is usually implicitly understood that there is some underlying random experiment whose outcome determines the value of X. For example, suppose that a gamble based on the outcome of the toss of a die pays \$10 if an even number is tossed, and pays \$20 if an odd number is tossed. If the die is a fair die, then there is probability of $\frac{1}{2}$ of tossing an even number and the same probability of $\frac{1}{2}$ of tossing an odd number. If the gamble had been described in terms of the flip of a fair coin with a payoff of \$10 if a head is flipped and a payoff of \$20 if a tail is flipped, then the probabilities of \$10 and \$20 are still each $\frac{1}{2}$. The crucial components of the description of this random variable are the possible outcomes (\$10 and \$20) and their probabilities (both $\frac{1}{2}$), and the actual experiment (even-or-odd die toss, or head-or-tail coin flip) leading to the outcome is not particularly significant, except that it tells us the probabilities of the possible outcomes. It would be possible to define this random variable without any reference to die toss or coin flip. We would say that the random variable X takes on either the value 10 or the value 20 and the probability is $\frac{1}{2}$ for each of these outcomes. That completely describes the random variable.

Discrete random variable:

The random variable X is discrete and is said to have a **discrete distribution** if it can take on values only from a finite or countable infinite sequence (usually the integers or some subset of the integers). As an example, consider the following two random variables related to successive tosses of a coin:

$X = 1$ if first head occurs on an even-numbered toss, $X = 0$ if first head occurs on an odd-numbered toss;

$Y = n$, where n is the number of the toss on which the first head occurs.

Both X and Y are discrete random variables, where X can take on only the values 0 or 1, and Y can take on any positive integer value. The "probability space" or set of possible outcomes for X is $\{0, 1\}$, and the probability space for Y is $\{1, 2, 3, 4, \ldots\}$.

Probability function of a discrete random variable:

The probability function (pf) of a discrete random variable is usually denoted $p(x)$, $f(x)$, $f_X(x)$ or p_x, and is equal to the probability that the value x occurs. This probability is sometimes denoted $P[X = x]$.

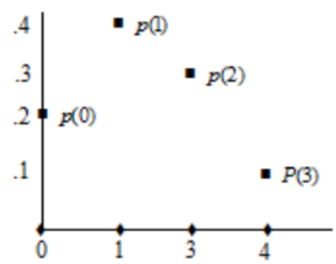
The probability function must satisfy

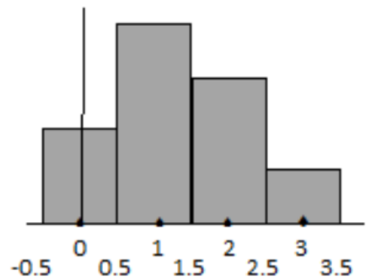
(i) $\boldsymbol{0 \leq p(x) \leq 1}$ for all x, and

(ii) $\boldsymbol{\sum_x p(x) = 1}$

Given a set A of real numbers (possible outcomes of X), the probability that X is one of the values in A is $\boldsymbol{P[X \in A] = \sum_{x \in A} p(x) = P[A]}$.

Probability plot and histogram: The probability function of a discrete random variable can be described in a probability plot or in a histogram. Suppose that X has the probability function $p(0) = 0.2$, $p(1) = 0.4$, $p(2) = 0.3$ and $p(3) = 0.1$ (note that the required conditions (i) and (ii) listed above are satisfied for this random variable X). The graph below on the left is the probability plot, and the graph at the right is the histogram for this distribution. For an integer valued random variable, a histogram is a bar graph. For each integer k, the base of the bar is from $k - \frac{1}{2}$ to $k + \frac{1}{2}$, and the height of the bar is the probability $p(k)$ at the point $X = k$. Histograms are also used to graph distributions that are described in interval form.





We can find various probabilities for this random variable X. For example

$P[X \text{ is odd}] = P[X = 1, 3] = P[X = 1] + P[X = 3] = 0.4 + 0.1 = 0.5$ and

$P[X \leq 2] = P[X = 0, 1, 2] = P[X = 0] + P[X = 1] + P[X = 2] = 0.9$.

We can find conditional probabilities also. For example,

$$P[X \geq 1 | X \leq 2] = \frac{P[(X\geq 1)\cap(X\leq 2)]}{P[X\leq 2]} = \frac{P[X=1,2]}{P[X=0,1,2]} = \frac{0.7}{0.9} = \frac{7}{9}.$$

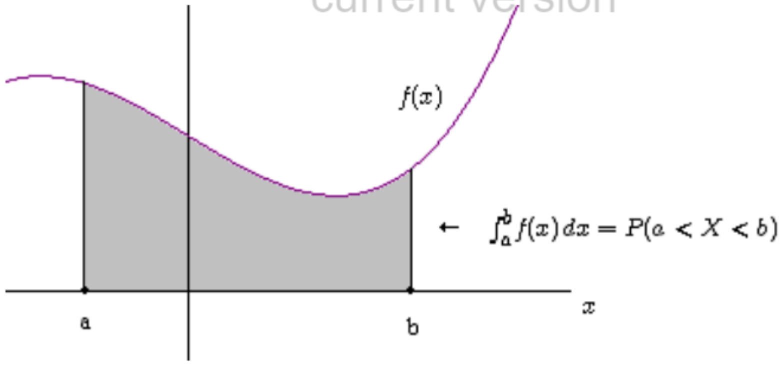
The probability at a point of a discrete random variable is sometimes called a **probability mass**. X above has a probability mass of 0.2 at $X = 0$, etc.

Continuous random variable:

A continuous random variable usually can assume numerical values from an interval of real numbers, or perhaps the whole set of real numbers. The probability space for the random variable is this interval. As an example, the length of time between successive streetcar arrivals at a particular (in service) streetcar stop could be regarded as a continuous random variable (assuming that time measurement can be made perfectly accurate).

Probability density function:

A continuous random variable X has a **probability density function (pdf)** usually denoted $f(x)$ or $f_X(x)$, which is a continuous function except possibly at a finite number of points. Probabilities related to X are found by integrating the density function over an interval. The probability that X is in the interval (a, b) is $\boldsymbol{P[X \in (a,b)] = P[a < X < b]}$**, which is defined to be equal to** $\boldsymbol{\int_a^b f(x)\,dx}$ (probability on an interval for a continuous random variable is the area under the density curve on that interval). Note that for a continuous random variable $P[X = c] = 0$ for any individual point c, since $P[X = c] = \int_c^c f(x)\,dx = 0$. For a continuous random variable there can only be probability over an interval, not at a single point.



Note that for a continuous random variable X, the following are all equal:

$$P(a < X < b)\ ,\ P(a < X \leq b)\ ,\ P(a \leq X < b)\ ,\ P(a \leq X \leq b)$$

This is true since the probability at a single point is 0, so it doesn't matter whether or not we include the endpoints a and b or not.

For a discrete random variable, probabilities are calculated as the sum of probabilities at individual points, so it does matter whether not an endpoint of an interval is included. For instance, for a fair die toss for which X denotes the outcome of the toss, $P(X \leq 3) = \frac{3}{6}$, but $P(X < 3) = \frac{2}{6}$.

The pdf $f(x)$ must satisfy (i) $f(x) \geq 0$ for all x, and (ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

Condition (ii) can be restated by saying that the integral of $f(x)$ over the probability space must be 1. Often, the region of non-zero density (the probability space of X) is a finite interval, and $f(x) = 0$ outside that interval. If $f(x)$ is continuous except at a finite number of points, then probabilities are defined and calculated as if $f(x)$ was continuous everywhere (the discontinuities are ignored).

Example 4-1:

Suppose that X has density function $f(x) = \begin{cases} 2x \text{ for } 0<x<1 \\ 0, \text{ elsewhere} \end{cases}$.

(i) Show that f satisfies the requirements for being a density function.

(ii) Find $P[0.2 < X < 0.5]$.

(iii) Find $P[0.2 < X < 0.5|X > 0.25]$.

Solution:

(i) f satisfies the requirements for a density function, since $f(x) \geq 0$ for all x and $\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 2x\,dx = 1$

(ii) $P[.2 < X < .5] = \int_{0.2}^{0.5} 2x\,dx = x^2\Big|_{0.2}^{0.5} = 0.21$. Note that this is equal to $P(0.2 \leq X \leq 0.5)$.

(iii) $P[0.2 < X < 0.5|X > 0.25] = \frac{P[(0.2<X<05)\cap(X>0.25)]}{P[X>0.25]} = \frac{P[0.25<X<0.5]}{P[X>0.25]} = \frac{\int_{0.25}^{0.5} 2x\,dx}{\int_{0.25}^{1} 2x\,dx}$

$= \frac{0.1875}{0.9375} = 0.2$ □

Example 4-2:

Y has the pdf $f(y) = \frac{20{,}000}{(100+y)^3}$ for $y > 0$.

(i) Show that f satisfies the requirements for being a density function.

(ii) Find $P(Y > t)$ if $t > 0$.

(iii) Find $P(Y > t + y | Y > t)$ if $t > 0$.

Solution:

(i) $\int_0^\infty \frac{20{,}000}{(100+y)^3}\, dy = \frac{20{,}000(100+y)^{-2}}{-2}\Big|_{y=0}^{y=\infty} = -0 + \frac{20{,}000}{2(100^2)} = 1$

and $f(y) \geq 0$ for all y

(ii) $P(Y > t) = \int_t^\infty \frac{20{,}000}{(100+y)^3}\, dy = \frac{20{,}000(100+y)^{-2}}{-2}\Big|_{y=t}^{y=\infty} = -0 + \frac{20{,}000}{2(100+t)^2} = (\frac{100}{100+t})^2$

(iii) $P(Y > t + y | Y > t) = \frac{P(Y>t+y)}{P(Y>t)} = \frac{10{,}000}{(100+t+y)^2} \Big/ \frac{10{,}000}{(100+t)^2} = (\frac{100+t}{100+t+y})^2$ $\square$

Mixed distribution: A random variable may have some points with non-zero probability mass combined with a continuous pdf on one or more intervals. Such a random variable is said to have a **mixed distribution**. The probability space is a combination of the set of discrete points of probability for the discrete part of the random variable along with the intervals of density for the continuous part. The sum of the probabilities at the discrete points of probability plus the integral of the density function on the continuous region for X is the total probability for X, and this must be 1. For example, suppose that X has probability of .5 at $X = 0$, and X is a continuous random variable on the interval $(0, 1)$ with density function $f(x) = x$ for $0 < x < 1$, and X has no density or probability elsewhere. This satisfies the requirements for a random variable since the total probability over the probability space is

$$P[X = 0] + \int_0^1 f(x)\, dx = 0.5 + \int_0^1 x\, dx = 0.5 + 0.5 = 1$$

Then,

$$P[0 < X < 0.5] = \int_0^{0.5} x\, dx = 0.125$$

and

$$P[0 \leq X < 0.5] = P[X = 0] + P[0 < X < 0.5] = 0.5 + 0.125 = 0.625$$

(since $X = 0$ is a discrete point of probability, we must include that probability in any interval that includes $X = 0$).

Cumulative distribution function (and survival function): Given a random variable X, the cumulative distribution function of X (also called the **distribution function**, or **cdf**) is $\boldsymbol{F(x) = P[X \leq x]}$ (also denoted $F_X(x)$). $F(x)$ is the cumulative probability to the left of (and including) the point x. The **survival function** is the complement of the distribution function, $\boldsymbol{S(x) = 1 - F(x) = P[X > x]}$. The event $X > x$ is referred to as a "tail" (or right tail) of the distribution.

For a discrete random variable with probability function $p(x)$, $\boldsymbol{F(x) = \sum_{w \leq x} p(w)}$, and in this case $F(x)$ is a "step function", it has a jump (or step increase) at each point with non-zero probability, while remaining constant until the next jump.