A Problem:

Arpan and Nikhil are equally good table tennis players. A match is set up between them with 9 games to be played and the winner will receive 800 Rs. So far 5 games are played: A won 3 and N won 2. Unfortunately the match has to be stopped for reasons beyond our control. How should the money be shared?

1. A says that he is the winner at this stage and so he should receive Rs 800. Thus A:800 and N:0.

2. N says that the match is stopped for no fault of his. So it should be regarded as draw and they should each gat Rs 400. Thus A:400 and N:400.

3. Then A says: since I won three games the money should be shared in the ratio 3:2. Thus A: 480; N:320.

4. Fermat and Pascal then think; like one of you did; and say the following: Why did this problem arise at all: Because of uncertainty regarding the remaining unplayed 4 games. So the solution should be got by trying to understand this uncertainty. Let us see. There are 16 possible scenarios in the remaining four games. Of these 11 scenarios make A winner while 5 scenarios make N winner. Thus if the unplayed games result in any one of the 11 scenarios then A wins and similarly for N. Therefore it is reasonable to award the money in this proportion 11:5.

Thus A:550 and N:250.

Thus entered the idea of understanding uncertainty. Now-a-days you will reformulate the last argument as: the chances of A winning (from now on) is 11/16 and the chances of N winning (from now on) is 5/16 and the money is shared in accordance to their chances of winning.

uncertainty:

Suppose that an experiment has total number of scenarios n. Of these, you are interested in an event B having k scenarios. Then chances of the event B is k/n — where we assumed that the scenarios are equally likely. Most of the initial uses of this concept was in gambling. Playing dice was also recorded in Mahabharata. So gambling is (at least) as old as that.

Later, this concept of chance found interesting applications in biology. Would the pea plant be tall or short; what are the chances? Would the baby be boy or girl; what are the chances? And so on.

One realized that life is full of uncertainty and thus chance analysis can be applied to understand many practical problems. For instance, will it rain tomorrow? uncertain, but need to know in order to advise formers and fishermen. You can try to predict. Yes, it is possible that sometimes we go wrong in predictions, no problem, good enough to be correct most of the time. You get data about the necessary parameters to understand this uncertainty — like wind speed, cloud concentration, pressure, temperature etc; make a model and predict.

Is it wise to release a particular medicine in the market? You need to conduct experiments; collect data regarding its effectiveness, side effects etc and then make a model and then take a decision.

Understanding chance phenomena was of supreme importance in physics too. It enters in several ways. Imagine a large container with water and also a pollen particle. Assume that water is in equilibrium. Actually it is never in equilibrium. What we mean is that there are no external forces that make pollen particle move — like water currents, air bubbles, evaporation etc. In practice you see the pollen particle still continues to perform motion. why does it move and how does it move. This is what you call Brownian Motion because it was first observed by Robert Brown.

Here is a lesson for us. Keep your eyes open and observe things minutely. Brown, not only observed this motion, but also persisted that reasons must be found for this motion!

Many many years later Einstein explained the reason. It is all due to chance phenomenon: the surrounding water molecules keep hitting the pollen particle and bump it this way, that way continuously and this causes motion. Actually you do not consider *one* pollen particle. Throw some pollen particles. Ask: At time $t > 0$, what proportion of particles are at a distance x from their initial position. But let us not complicate life now. You should know that this probabilistic analysis was the basis for the first calculation of Avogadro number.

Several years later Norbert Wiener proved that such motions can be mathematically modelled. It took some more years for people to take note of this work of Wiener and appreciate. Once this happened, Brownian motion became indispensable for modelling in several contexts.

When you learn Statistical Mechanics, you will see more probability. When you learn quantum mechanics you will see much more of probability.

Now a days we use probability, in an essential way, in computer science, (randomized algorithms, simulation $\cdots$) ; modelling of stock market prices and so on.

Modelling uncertainty:

So how do you model chance experiments. The first thing is the following: you must know what happens when you do the experiment, that is, the possible scenarios. They are called **outcomes**. For example when we toss a coin twice, one after other, the scenarios are HH, HT, TH, TT with obvious understanding of the symbols used. Each one of these is called an outcome of the experiment.

The set of all outcomes is called **sample space** of the experiment. Usually sample space is denoted by Ω. Any subset of the sample space is called an **event**.

Experiment: Toss coin twice
Sample space Ω: $\{HH, HT, TH, TT\}$.

Experiment: Roll a six faced die once.
Sample space Ω: $\{1, 2, 3, 4, 5, 6\}$

Experiment: Toss a coin till you get heads and then stop.
Sample space Ω: $\{H;\ TH;\ TTH;\ TTTH;\ \cdots\cdots\}$.

Experiment: Observe the bulb glowing above, do not switch off. Note down the life time of this bulb. That is the duration till it ceases to work; fix some units of measurement.
Sample space Ω: $[0, \infty)$

Experiment: Pick a point from the interval $(0, 1)$.
Sample space Ω: $(0, 1)$

So on and on.

In the first two examples there are finitely many outcomes. In the third example there are too many outcomes, but countable. In the last two example the number of outcomes are far tooooo many, uncountable.

We want to model chances of outcomes. Any such activity starts with understanding and modelling simple experiments. If they do not fit reality, then use the understanding thus far gained and model complicated experiments; And so on. Basic philosophy: understand simple things first.

equally likely outcomes:

We first start understanding experiments which have finitely many outcomes.

We further assume that the outcomes are equally likely; we have only intuitive idea of what this means and must make a mathematical model. For example in tossing a coin once, there are two outcomes H, T and we assume that they are equally likely, so each must have chance $1/2$.

Suppose that the sample space of the experiment is Ω, a finite set. For each $\omega \in \Omega$ we associate a number chance of that outcome, $p(\omega)$. This number will tell you the chances of observing the outcome ω. Since we believe chances of some thing happening should be non-negative, we should have $p(\omega) \geq 0$.

Suppose we have two (different) outcomes ω_1, ω_2. It stands to reason to believe: the chances that one of the outcomes ω_1 or ω_2 appears is $p(\omega_1) + p(\omega_2)$. More generally if A is any set of outcomes then the chances that one outcome in A is observed should equal $\sum\limits_{\omega \in A} p(\omega)$.

In particular, the chances that an outcome in Ω is observed, equals $\sum\limits_{\omega \in \Omega} p(\omega)$. But Ω being the set of all possible scenarios, we are sure to observe something from Ω. In other words this sum should equal one.

Thus a good assignment of probability for points of a sample space is a map that associates with each $\omega \in \Omega$ a non-negative number $p(\omega)$ which all add to one. Then Probability of an event A equals $\sum\limits_{\omega \in A} p(\omega)$.

The number $p(\omega)$ is the chances of observing the outcome ω.

After this thought process, let us return to the special case of equally likely outcomes. Let Ω be the finite sample space of an experiment. If outcomes of Ω are equally likely, there is a number c such that $p(\omega) = c$ for all ω and since they should add to one; we conclude $p(\omega) = 1/|\Omega|$. This immediately leads to the conclusion that chances of any event equals the fraction of outcomes in that event.

In an experiment with finitely many outcomes, the probability of an event A equals the ratio

$$\frac{\textit{number of outcomes in } A}{\textit{Total number of outcomes}}$$

Thus calculating probabilities reduces to counting number of outcomes in sets of interest. This is one reason for considering finite sample spaces and equally likely outcomes. You can use your expertise in counting learnt in high school.

You must understand two things. Most of the time we consider tossing coin or throwing balls into boxes. These are only symbolic and apply to many situations. For example any analysis involving coin tossing can be applied to experiments where there are just two outcomes; H/T or Success/Failure or On/Off or $+1/-1$ or Good/Defective etc. Balls could be particles and boxes could be energy levels.

Secondly, equally likely is only a first step. For example if you are manager of a company manufacturing bolts; each bolt could be good or defective.

It is suicidal for the company to assume that they are equally likely.

Example: I have 30 letters addressed to different people and their different addresses on 30 envelopes. I put the letters in the envelopes (one in each) and all possible arrangements are equally likely. What is the probability that at least one letter is put in its envelope.

Here the experiment consists of putting the letters in the envelopes and you can see that there are 30! possible ways of doing this. These are the out comes. For example one outcome is: letter i is put in envelope i for all i. Another outcome is: letter i is put in envelope $30 - i + 1$. Another outcome is: letter i is put in envelope $2i$ for $1 \leq i \leq 15$ and letter $2i$ is put in envelope i for $1 \leq i \leq 15$.

The reason I am giving these examples is that many times without having any idea or feeling for the outcomes, you start counting and realize any error only afterwards. It is not necessary to take such a route. Understand, think and do.

Let A be the set of all outcomes in which at least one letter is put in its envelope. It is not easy to count $|A|$, number of elements in A. So we split this event into simpler events and use 'inclusion-exclusion' formula. Let, for $1 \leq i \leq 30$, A_i be the set of all outcomes in which letter i is put in its envelope. Clearly

$$A = \cup A_i.$$

■ Theorem:

$$|\overset{n}{\underset{1}{\cup}} A_i| = S_1 - S_2 + S_3 - S_4 + \cdots$$

where

$$S_1 = \sum_{1}^{n} |A_i|; \qquad S_2 = \sum_{i<j} |A_i \cap A_j|;$$

$$S_3 = \sum_{i<j<k} |A_i \cap A_j \cap A_k|; \quad \cdots . \qquad ■$$

I leave for you to check, in our example,

$$|A_i| = 29!; \qquad |A_i \cap A_j| = 28! \quad (i \neq j)$$

$$|A_i \cap A_j \cap A_k| = 27! \quad (i < j < k) \qquad \cdots$$

Thus

$$S_1 = 30 \times 29! = 30|$$

$$S_2 = \binom{30}{2} \times 28! = \frac{30!}{2!}$$
$$S_3 = \binom{30}{3} \times 27! = \frac{30!}{3!}$$

Thus
$$P(A) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots$$
Thus the chances that no letter goes to its envelope is given by

$$1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots \pm \frac{1}{30!}$$

Instead of 30 letters and envelopes, if we have n letters and n envelopes, then the chances that there is no match is given by

$$1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots \pm \frac{1}{n!}$$

This is pretty close to $1/e$ if n is pretty large.

So how do you prove the above theorem. Let us see if we can do for $n = 2$. To show
$$|A \cup B| = |A| \ + \ |B| \ - \ |A \cap B|$$
We know
$$A \cup B = (A) \ \cup \ (B - A \cap B)$$
Recall $B - C$ stands for set of all points which belong to B and which are not in C. The two sets A and $B - A \cap B$ are disjoint sets and thus the number of elements adds up. Thus

$$|A \cup B| = |A| \ + \ |B - A \cap B| \qquad (\bullet)$$

By same argument $B - A \cap B$ and $A \cap B$ being disjoint with union equal to B we get
$$|B| = |B - A \cap B| \ + \ |A \cap B|$$
or
$$|B - A \cap B| = |B| \ - \ |A \cap B|$$
Substituting this in $(\bullet)$ we get

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

as stated.

We shall prove the theorem by induction on n. Assume we have proved for any $n \leq m$. In particular if you take any m sets the formula is correct. We shall prove it for $m+1$ sets.

So take $m+1$ sets $A_1, A_2, \cdots, A_{m+1}$. Denote

$$A = A_1 \cup \cdots \cup A_m; \qquad B = A_{m+1}$$

thus

$$\bigcup_1^{m+1} A_i = A \cup B$$

So from what was proved above

$$|\bigcup_1^{m+1} A_i| = |A| \ + \ |B| \ - \ |A \cap B| \qquad (\spadesuit)$$

Let us denote

$$T_1 = \sum_1^m |A_i|; \qquad T_2 = \sum_{1 \leq i < j \leq m} |A_i \cap A_j|;$$

$$T_3 = \sum_{1 \leq i < j < k \leq m} |A_i \cap A_j \cap A_k|; \quad \cdots .$$

so that by induction hypothesis applied for the m sets $A_1, \cdots, A_m$ we get

$$|A| \ = \ |\bigcup_1^m A_i| \ = \ T_1 \ - \ T_2 \ + \ T_3 \ - \ \cdots \qquad (\star)$$

Let us observe

$$A \cap B \ = \ (\bigcup_1^m A_i) \cap A_{m+1} \ = \ \bigcup_1^m (A_i \cap A_{m+1})$$

Let us denote

$$U_1 = \sum_1^m |A_i \cap A_{m+1}|; \qquad U_2 = \sum_{1 \leq i < j \leq m} |A_i \cap A_j \cap A_{m+1}|;$$

$$U_3 = \sum_{1 \leq i < j < k \leq m} |A_i \cap A_j \cap A_k \cap A_{m+1}|; \quad \cdots .$$

Applying the induction hypothesis for the m sets

$$A_1 \cap A_{m+1}, \quad A_2 \cap A_{m+1}, \cdots, \quad A_m \cap A_{m+1}$$

we see

$$|A \cap B| \; = \; U_1 \; - \; U_2 \; + \; U_3 \; - \; \cdots \qquad (\star\star)$$

Using $(\star)$ and $(\star\star)$ in $(\spadesuit)$ and remembering that $B = A_{m+1}$ we get

$$|\bigcup_1^{m+1} A_i| =$$

$$\begin{aligned} &[T_1 - T_2 + T_3 - \cdots] \; + \; |A_{m+1}| \; - \; [U_1 - U_2 + U_3 - \cdots] \\ &= [T_1 + |A_{m+1}|] \; - \; [T_2 + U_1] \; + \; [T_3 + U_2] \; - \; \cdots \\ &= S_1 - S_2 + S_3 - \cdots \end{aligned}$$

as required. You should think about the last equality and be convinced.

This complete proof of the theorem.

There are several other things you can do with this experiment. But let us discuss other experiments.

Example: Let $A = \{1, 2, \cdots, N\}$. Experiment consists of picking one after another, twenty times, points from this set A. Remember always selected from A, we are not excluding the first selected element when selecting later point.

Sample space Ω consists of all sequences $(x_1, x_2, \cdots, x_{20})$ where for each i, $x_i \in A$.

Check $|\Omega| = N^{20}$.
Each outcome has probability $1/N^{20}$.

This experiment is called **sampling with replacement, of size 20**. With replacement because when we take second item the first item is not removed from the set A; the second selection is made as if you noted down the first item and put it back before second selection. Each outcome of this experiment is called sample (with replacement) of size 20.

What are the chances that first observation (first point of the sample) is 1? Answer: there are N^{19} outcomes for which first item is 1. So the required probability is $N^{19}/N^{20} = 1/N$

What are the chances that second item is 1? Again there are N^{19} outcomes where second item is 1. So required probability is again $1/N$.

What are the chances that first and second item are 1? There are N^{18} such outcomes and so the probability is $N^{18}/N^{20} = 1/N^2$.

What are the chances that 1 is included in the sample, that is some x_i in the sample is 1? There are $(N-1)^{19}$ outcomes where $x_i \neq 1$ for all i and hence the required probability equals

$$\frac{N^{20} - (N-1)^{20}}{N^{20}} = 1 - (1 - \frac{1}{N})^{20}$$

Example: Again let $A = \{1, 2, \cdots, N\}$. Experiment consists of picking one after another, twenty times, points from this set A, each time NOT replacing the points selected earlier.

Sample space Ω consists of all sequences $(x_1, x_2, \cdots, x_{20})$ where for each i, $x_i \in A$ and they are distinct. That is $x_i \neq x_j$ for $i \neq j$.

Check $|\Omega| = N(N-1)\cdots(N-19)$.
We assume that $N \geq 20$. Otherwise the sample space is empty set.
Each outcome has probability

$$\frac{1}{N(N-1)\cdots(N-19)}.$$

This experiment is called **sampling without replacement, of size 20**. Without replacement because when we take second item the first item is removed from the set A; similarly at each stage selected item is removed and then next selection is made. Each outcome of this experiment is called a sample (without replacement) of size 20.

What are the chances that first observation (first point of the sample) is 1? Answer: there are $(N-1)(N-2)\cdots(N-19)$ outcomes for which first item is 1. So the required probability is $1/N$

What are the chances that second item is 1? Again there are $(N-1)(N-2)\cdots(N-19)$ outcomes where second item is 1. So required probability is again $1/N$.

What are the chances that first and second item are 1? Zero.

What are the chances that 1 is included in the sample, that is some x_i in the sample is 1? There are $(N-1)(N-2)\cdots(N-20)$ outcomes where $x_i \neq 1$ for all i and hence the required probability equals

$$1 - \frac{N-20}{N} = \frac{20}{N}$$

Example: Again let $A = \{1, 2, \cdots, N\}$. Experiment consists of picking a subset of A of size 20.

Remember, we are not picking elements one by one. We grab a subset consisting of 20 elements.

Sample space Ω consists of all subsets of A which have 20 elements. That is all $\omega \subset A$ with $|\omega| = 20$

Check $|\Omega| = \binom{N}{20}$

We assume that $N \geq 20$. Otherwise the sample space is empty set.

Each outcome has probability

$$\frac{1}{\binom{N}{20}}.$$

This experiment is called **selecting a subset of size 20**.

What are the chances that first observation (first point of the sample) is 1? this question does not make sense. There is no first or second element. We have a subset consisting of 20 points, that is all.

What are the chances that second item is 1? Again meaningless question.

What are the chances that 1 is included in the selected set? that is $1 \in \omega$? There are

$$\binom{N-1}{19}$$

such outcomes and hence the probability equals

$$\frac{\binom{N-1}{19}}{\binom{N}{20}} = \frac{(N-1)!}{19!(N-20)!}\frac{20!(N-20)!}{N!} = \frac{20}{N}$$

same as in the case of sampling without replacement.

Suppose that in the $N = 50$ items there are 40 good items and 10 defective items. What are the chances that if I take set of 10 items from the lot it contains exactly 3 defective items? The chances are

$$\frac{\binom{40}{7}\binom{10}{3}}{\binom{50}{10}}$$

You can ask and answer several questions. You see the importance.

For instance you ask for supply of 1000 items for your company, confirming to some specifications. The suppliers delivers. Would you accept without making sure that there are not too many defectives? How do you find out? You estimate. You take some items and check how many defectives you got. Base your decision on this. This you will learn when (and if) you learn statistical methods.

Sometimes there may be only one kind of items and you *create* two kinds! Consider estimating number of fish in a pond. You do not remove water and count, right? what would you do? You catch some fish, mark them and leave them again in the pond. Now there are two kinds: marked and unmarked. wait for a day. Then 'recapture' some fish and count how many marked fish are there in this recaptured bunch. make your estimate based on this data.

Example: I have 30 boxes numbered: $1, 2, \cdots, 30$. I have 20 balls numbered: $1, 2, \cdots, 20$. Experiment is to throw the balls into the boxes.

$$|\Omega| = 30^{20}$$

all these outcomes are equally likely.

What are the chances that the first box is empty?

$$\frac{29^{20}}{30^{20}}$$

What are the chances that boxes 1,5,7 contain respectively 10,7,3 balls?

$$\frac{\binom{20}{10}\binom{10}{7}}{30^{20}}$$

This experiment is called **Maxwell-Boltzman** experiment. Boxes are energy levels and balls are elementary particles. The interesting point is: apparently, no known particles obey this rule!

Example: I have 30 boxes numbered: $1, 2, \cdots, 30$. I have 20 balls all looking alike. There is no way to distinguish one ball from the other. Experiment is to throw the balls into the boxes. How many different arrangements can our eye perceive?

$$|\Omega| = \binom{49}{29}$$

This can be seen in several ways. Put 49 star marks in a row; select 29 of these and convert them into vertical lines. You see a picture of 20 balls (the remaining star marks) put in 30 boxes (made by the vertical lines). And every method of putting balls into the boxes is achieved this way. remember, that now it makes no sense to say: where did ball one go? all balls look alike. so two arrangements are different only when the 'occupancy numbers' are different, that is, the vector $(n_1, n_2, \cdots, n_{30})$ where n_i is the number of balls in box i; determines the arrangement. Two different vectors give two different arrangements.

Bose-Einstein Rule: These arrangements are equally likely.

This experiment is called **Bose-Einstein** experiment. Again boxes are energy levels and balls are elementary particles.

What are the chances that box one is empty?

$$\frac{\binom{48}{28}}{\binom{49}{29}}$$

You can simplify and see.

What are the chances that boxes 1,5,7 contain respectively 10,7,3 balls?

$$\frac{1}{\binom{49}{29}}$$

This is because there is only one outcome satisfying the given condition.

Photons are known to obey this rule.

To understand how outrageous is this rule consider the following. I have a box with 1 green ball and 100 red balls numbered 1,2,$\cdots$, 100. I pick a ball at random. What are the chances it is green:

$$\frac{1}{101}$$

Suppose now I tell you that all red balls look alike, there are no numbers on them. How many outcomes your eye can perceive? only two: Red, green; Can I say that these are equally likely and conclude that chance of green ball is 1/2? Ridiculous. The B-S rule *does not* apply here.

Why photons obey B-E rule is unclear, but they do.

Example: I have 30 boxes numbered: $1, 2, \cdots, 30$. I have 20 balls all looking alike. There is no way to distinguish one ball from the other. Experiment is to throw the balls into the boxes subject to the condition: No more than one ball in a box.

$$|\Omega| = \binom{30}{20}$$

These are equally likely.

What are the chances that box one is empty?

$$\frac{\binom{29}{20}}{\binom{30}{20}}$$

What are the chances that boxes 1,5,7 contain respectively 10,7,3 balls? Zero.

This is called **Fermi-Dirac** experiment. Protons obey this rule.

Example: Kishlaya and Sandeep belong to a club with 100 members. They stand for election, say for presidentship of the club. K got 60 votes and S got 40 votes and thus K is the winner.

What are the chances that through out vote-counting, K is leading?

We discuss this problem because it teaches a new counting technique and the method has lots of applications. Further the problem itself has a neat answer:

$$\frac{60 - 40}{60 + 40} = \frac{20}{100}$$

What is vote counting? Let us count a vote for K as $+1$ and a vote for S as -1. Thus vote counting means all possible sequences of ± 1 of length 100 which have 60 ones and 40 minus ones.. This is the sample space. Thus

$$|\Omega| = \binom{100}{60}$$

We assume that all these outcomes are equally likely. Thus the problem boils down to finding $|A|$ where A consists of sequences which have at every stage more $+1$ than -1. Denote this set by A.

If we think of outcomes as simply sequences of ± 1 then to see if an outcome is in A or not we need to add and check at every stage. So let us think of Ω in a different and convenient manner.

A path is a sequence

$$(0, s_0), (1, s_1), \cdots (k, s_k)$$

where each s_i is an integer and $s_i - s_{i-1} = \pm 1$ for all $i \geq 1$. The path is said to start at s_0 and end at s_k and is of length k.

We claim Ω can be thought of as the set of paths of length 100 starting at 0 and ending at 20. obviously any such path defines a vote counting; simply

$$(s_i - s_{i-1} : i = 1, 2, \cdots, 100)$$

Verify that this has exactly 60 ones and 40 minus ones (because path ends at 20 and starts at zero).

Conversely, given a vote counting

$$(\epsilon_i : i = 1, 2, \cdots, 100)$$

you can define a path by taking

$$s_0 = 0, \qquad s_m = \sum_1^m \epsilon_i \quad 1 \leq m \leq 100.$$

Verify that this path starts at zero and ends at 20.

You can draw the axes and join successive points by straight line and visualize a path. such a picturesque visualization suggests ideas as we see now.

Digression:

1. A set is a well-defined collection of objects. This is not a very good definition but good enough for us. Well-defined means given an object you

should know whether it belongs to your set or not. If there is an object about which you can not decide whether it is in your collection or not, then conclude that your collection is not a set. Thus a set can be considered as a bag of objects.

If A is a set and an object x is in the set, we write $x \in A$.

2. Two sets are same if they have the same objects. Thus two sets A and B are same, written $A = B$, if: $x \in A$ iff $x \in B$. When you list elements of a set and write it (for *your* understanding) you need not repeat elements. This is simply because the set $A = \{1, 2, 1\}$ is same as $B = \{1, 2\}$. Indeed, you can show $x \in A$ iff $x \in B$. Also B is same as $C = \{2, 1\}$ for the same reason. Thus there is no order of appearance either among elements of the set. When you list a set you list the distinct elements of the set.

Thus listing elements of a set is only for your convenience and for your understanding; a set is described just by a property possessed by its elements. An object having the property is in the set and an object not having the property is not in the set. The set does not depend on the listing or on the person making the list or on the order the elements are being listed etc.

$A \subset B$ means $[x \in A$ implies $x \in B]$. This is also written as $B \supset A$. read A is a subset of B or B is a superset of A.

Prove: $A = B$ iff $[A \subset B \& B \subset A]$.

3. Empty set is the set which has no elements, denoted $\emptyset$.

4. Let $n \in \{0, 1, 2, \cdots\}$ the set of non-negative integers.

We say that empty set has zero elements, $|\emptyset| = 0$ or $\#\emptyset = 0$.

We say A has n elements; write, $|A| = n$ or $\#A = n$ if there is a bijection

$$f : \{1, 2, \cdots, n\} \to A.$$

Sometimes one equivalently says that there is a bijection

$$f : \{0, 1, 2 \cdots, n-1\} \to A.$$

Bijection $f : U \to V$ means a function which is one-to-one and onto. Here one-to-one means

$$x \neq y, x \in U, y \in U \quad \to \quad f(x) \neq f(y).$$

And onto means

$$\forall z \in V \ \ \exists x \in U; \ \ f(x) = z.$$

5. A set A is finite if it is empty set or if for some n there exists a bijection between $\{0, 1, 2, \cdots, n-1\}$ and A. Otherwise the set is infinite.

6. A set A is countably infinite if there is a bijection

$$f : \{1, 2, \cdots\cdots\} \to A.$$

7. A set is countable if it is either finite or countably infinite. Roughly speaking, you can count elements of a countable set. If the set is finite then the counting ends. If the set is infinite then counting continues for ever. You might wonder whether it can be called counting at all, in case it never ends. Yes. we *are* prescribing an algorithm to count: $f(1)$, then $f(2)$, then $f(3)$, etc etc. With this algorithm, every element of A is 'listed' sooner or later and no element is listed twice.

8. A set which is not countable is called uncountable. For example set of real numbers is an uncountable set.

9. For sets A, B we define $A \cup B$ (union) as the set of all objects which belong to at least one of the sets A, B. We define $A \cap B$ (intersection) as all objects which belong to both A, B. Similarly you define union and intersection of any (yes, *any*) family of sets.

10. For sets A and B, we define $A - B$ as set of all points which are in A but not in B. This is also same as $A - (A \cap B)$.

If we have fixed one grand set Ω during a discussion and $A \subset \Omega$ then we denote $\Omega - A = A^c$. Remember A^c does not make sense unless we know what is Ω.

11. There are several properties concerning set operations and you should be able to prove them when used. For example:

$$(A_1 \cup A_2) \cap B = (A_1 \cap B) \cup (A_2 \cap B)$$

∎ Proof: Let $x \in (A_1 \cup A_2) \cap B$ Then $x \in A_1 \cup A_2$ and $x \in B$. But then either $x \in A_1$ or $x \in A_2$. In case $x \in A_1$ then $x \in A_1 \cap B$ and in case $x \in A_2$ then $x \in A_2 \cap B$. In either case $x \in (A_1 \cap B) \cup (A_2 \cap B)$.

Conversely, let $x \in (A_1 \cap B) \ \cup \ (A_2 \cap B)$. Either $x \in (A_1 \cap B)$ or $x \in (A_2 \cap B)$. In the first case $x \in A_1$ and $x \in B$. In the second case $x \in A_2$

and $x \in B$. Thus in either case $x \in B$. And also in either case $x \in A_1 \cup A_2$. Thus $x \in (A_1 \cup A_2) \cap B$. ∎

HA1 EX7:

Here is the thought process and I felt that it should be easy to convert the thought into maths avoiding outrageous statements and unnecessary words like independence. I see it involves some work.

Let $p > 1$ be prime. The chances that a positive integer selected at random is divisible by p is $1/p$ — because a proportion $1/p$ of positive integers are divisible by p.

The chances that two positive integers selected at random independently are both divisible by p is

$$\frac{1}{p} \cdot \frac{1}{p} = \frac{1}{p^2}$$

The chances that two positive integers selected at random independently are not simultaneously divisible by p (that is at least one of them is not divisible by p) is

$$1 - \frac{1}{p^2}$$

Let the above event be A_p. Let $p_1, p_2, \cdots$ be enumeration of all primes larger than one. Note that

$$\bigcap_1^\infty A_{p_i}$$

is precisely the event that the selected pair does not have a common prime factor (and hence coprime).

While passing, note that divisibility by different primes p_1, p_2 are independent events because proportion of integers divisible by both equals product of the individual proportions. Thus

$$P(\bigcap_1^k A_{p_i}) = \prod_1^k (1 - \frac{1}{p_i^2})$$

So required probability equals

$$P(\bigcap_1^\infty A_{p_i}) = \prod_1^\infty (1 - \frac{1}{p_i^2}) \qquad (\boxplus)$$

Note

$$\prod_1^\infty (1 - \frac{1}{p_i^2})^{-1} = \prod_1^\infty \left[1 + \frac{1}{p_i^2} + \frac{1}{p_i^4} + \frac{1}{p_i^6} + \cdots\right]$$

$$= \sum_1^\infty \frac{1}{n^2} \qquad (\boxtimes)$$

Clearly ($\boxtimes$) and ($\boxplus$) lead to required solution.

This is to be converted to math: selecting an integer at random does not make sense; use of independence is to be avoided etc. Need to think.

Books:

You can consult any books on probability you see in the Library. here are some suggestions:

William Feller   Volume 1;

Paul Hoel , Sidney Port, Charles Stone (three authors)   Volume 1;

Sheldon Ross   Probability

Most of HomeAssignment problems should be in some book or other.