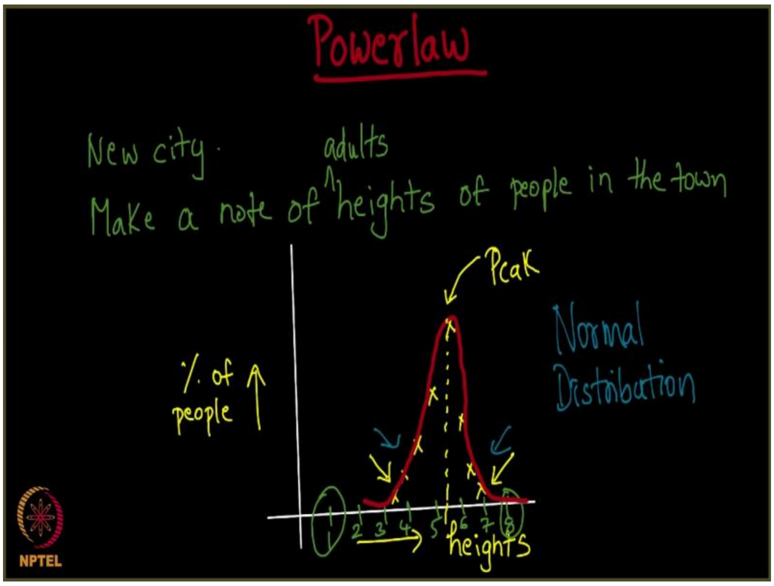
**Social Networks**
**Prof. S. R. S. Iyengar**
**Department of Computer Science**
**Indian Institute of Technology, Ropar**
**Rich Get Richer Phenomenon**

**Lecture – 115**
**Introduction to Power law**

(Refer Slide Time: 00:08)




So, we are now starting a brand-new chapter called the Power Law. So, it is difficult to; difficult on your minds if I start talking about power law directly. So, as and always it has been customary, that we start with an example right with a nice story rather ok. So, what I will do is I will now ask you all this question. Assume if I go to a new city and I make a note of everybody's heights of everyone in the town. I make a note of every ones height of people in the town. I come back and then I look at this data. And I try plotting the following very simple graph. The graph will be this all my x axis, I will have heights; on my y axis I will have percentage of people with that height.

Now, let me illustrate this with a numeric value. So, what I do is let us say this is 1, this is 2, this is 3; 1, 2, 3, 4, 5, 6, 7. So, I forgot to mention that I am making note of only height of adults in the town ok. As you would have guessed they there is no reason why we should take one foot 1 foot or let us say 8 and above, we do not know of anyone whose of this height adults especially.

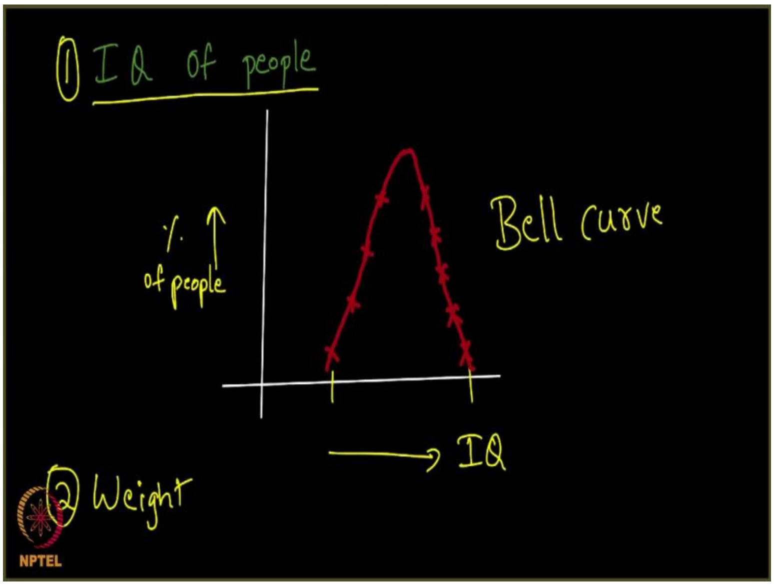
So, what I do is this plot mainly goes like this. I look at the percentage of people who are 5 feet tall, percentage of people who are 6 feet tall. And most of the people as you would have guessed will be between 5 and 6. So, around 5.3, 5.4 will be the peak. There will be very less people with 6 and over very very few people who will be 7 feet tall, very unlikely in many countries although you will find tall people; it is very country specific. You might for find all people in different countries, but most the countries will not have very tall people above 7 feet.

So, they are going to be very very less. And people just above four are going to be less; people who are 4 feet are very very less. And this is how your plot will look like as you would have guessed ok. It is it goes up and comes down like this. So, what do we observe? We observe that there are very few people with very less height. And very few people with who are very very tall and most the people are between this ok.

Let us say between 5 and 6 feet correct. This is varied peaks; such a distribution is called a normal distribution; this is called a normal distribution. What do you? What do you mean by normal distribution? The distribution where it is less on the left and right, and it peaks somewhere in the middle right ok.

So, we can see many things in life which has normal distribution things.

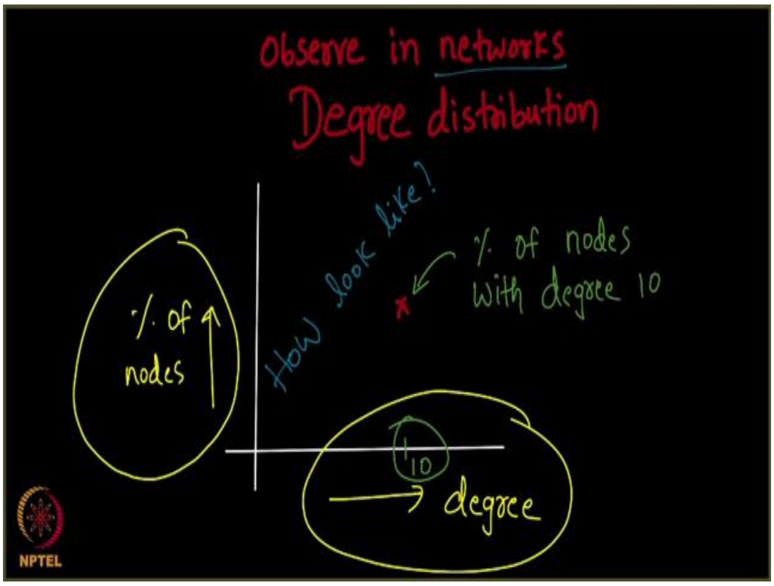
(Refer Slide Time: 03:57)



For example, let us say IQ of people; IQ of people in a classroom; you will see you will see if I were to plot on the x and y axis. As and always as I told you this is going to be my in this example IQ of people. And y axis will be percentage of people with that IQ right. So, let us say this is slightly on the lower side; this is slightly on the higher side; you will find very few people with very very low IQ; meaning below average very few people with extremely high IQ. And most of them will sort of lie in the middle, and if you actually try plotting values for everything; you will indeed get a curve like this. And this curve in statistics is called the Bell curve. It resembles the shape of the bell that's why it is called the bell curve. In fact, it is easy to capture it mathematically as well although we will not discuss it right now.

So, IQ of people forms a bell curve you can think of many such examples right; weight of people in a town, weight is another such example, IQ being one such example, height being another example right. So, let us look at an example which is very specific to our work to our subject and that is going to be next.

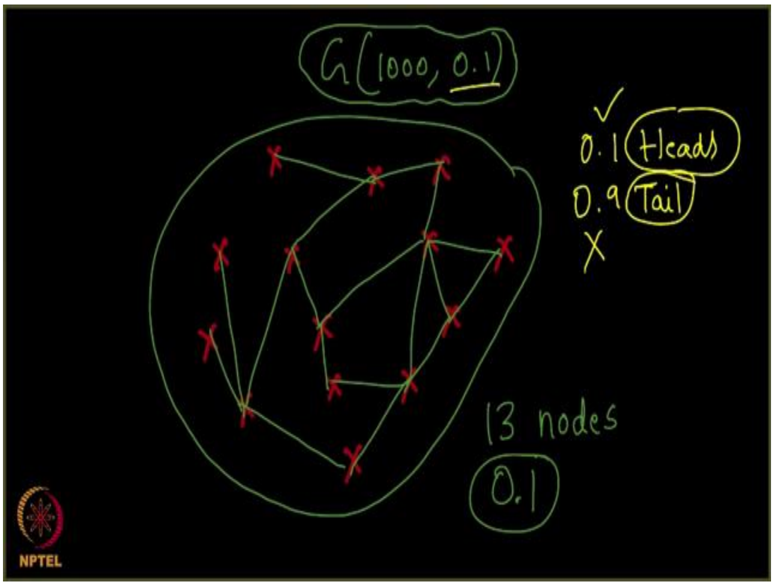
(Refer Slide Time: 05:35)




So, if I what do we observe, what do we observe? If we were to take a network and try to ask the degree distribution in the network; degree distribution what do you even mean by this? By degree distribution I mean again x and y axis ok. This is my x axis and my x axis. How long my x axis will be degree of nodes. And along my y axis will be percentage of nodes with that degree; with that I mean if you have something like 10 here; just this point will denote the let me just write the downs its seems into your minds. This denotes the percentage of nodes with degree 10, degree 10.

And big question right now is, how does this plot look like? How does this plot look like? For that we should first take a network, now here is a good the chance for you all to write a piece of code and observe the following.

(Refer Slide Time: 07:03)



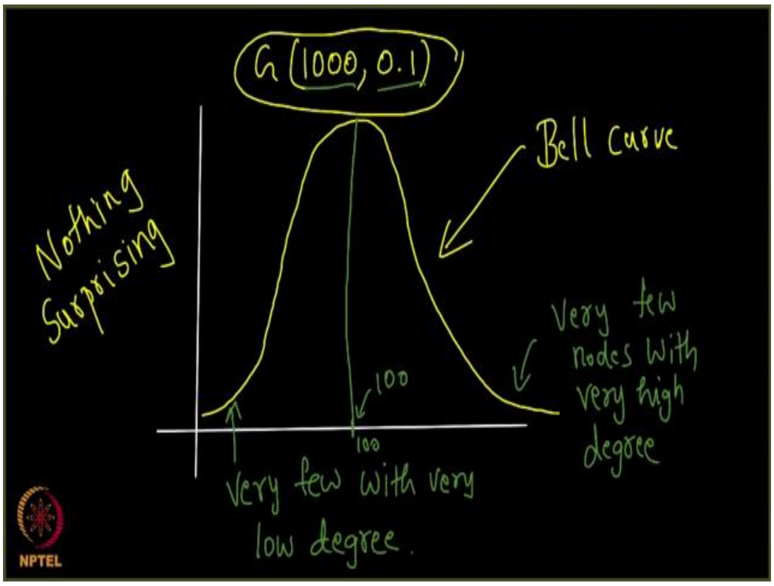

Let me take you all should probably consider a graph with 1000 nodes. And with probability let us say point 1 you put edges between them. What do you mean by this? By this you mean let us say we had 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 nodes something like this.

What you do is you will put edges by considering the probability 0.1. What do what do I mean by probability 0.1? We have discussed the set length before, but I will just recalculate. You toss a coin with head and tail probability 0.1; heads probability and 0.9 will be your tail probability right.

If you get a head, you put an edge; if you get a tail, you do not put an edge that is how I am continuing this right. Assume I got a network something like this right. This will be my final network. And this is my graph G with probability, how many how many nodes are here? 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 nodes, this is graph 13 nodes and edge probability is 0.1 which means I toss a coin with probability of it being heads is 0.1 and tail being 0.9, I put an edge if I get a head I do not put an edge; if it is a tail.

I repeat this process for every possible edge on this 13 node graph ok. I showed you a simple example, but what you should be doing is do it for a 1000 node graph right.

(Refer Slide Time: 08:56)      .




Do it for a 1000 node graph. Once you finish programming, you get 1000 node graph. And put edges with probability 0.1. This is called a G (1000, 0.1). And what should you do? As you would have guessed compute it is degree distribution. We discussed in our, previous slide; what is the degree distribution. You put degree on the x axis and you put the percentage of nodes on the y axis, and see how this plots looks like.

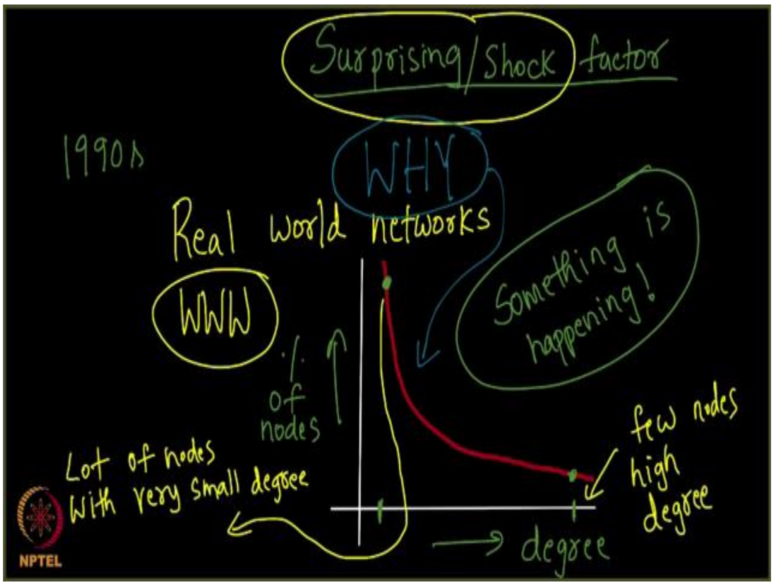
You will be started to see, that this plot will resemble a bell curve once again with very few nodes, very few nodes with very high degree with very high degree. At the same time, you will see very few nodes, very few nodes with very low degree. And again, it will peak somewhere in the middle.

If you can guess, it peaks around 100 for x being 100. It peaks you can probably tell me why it is the product of 1000 and 0.1. It is lift as a it is left as a quiz for you all to say why does it peak at 100. Anyways, the moral the story is simple. When we draw out a graph with 1000 nodes and 0.1 edge probability, we will see again a bell curve. You probably you would tell me yeah, we saw it in the previous cases, we saw many examples right it is not at all surprising for us right.

So, IQ of people you saw, you saw, that the degree of nodes follow. What is called the normal distribution? And everything else looks similar right. If you look at weights of people, height of people, IQ of people all of them seem to follow this bell curve only nothing surprising about it right. So, let me write that down there is nothing surprising it nothing surprising.

Then in the late 1990s scientist observed something that is actually very surprising.

(Refer Slide Time: 11:31)




So, what do you mean by the very word surprise? Surprise is something that happens unexpectedly. You think that this is how it should be, but nature has something else in store. You see something else happening where what you thought should have happened is not what is happening there. And that is what creates a surprising or what is called a shock factor. In the late 1990's network scientist observed a very nice phenomenon. They took at the, they looked at several real-world networks. So, until 1990's people did not have big networks. That was because we did not have computational ability back then. It is only in the 1880's and 1990's that we started collecting a lot of data and we had data sets we had computers which would crunch the data sets.

So, let us say they if you consider the network of the World Wide Web. It is a few millions of nodes. People ask the same question. What will be the distribution of degree on these networks? And they all expected it will be a bell curve. And as you would have guessed, there was a surprise stroke shock factor there and the fact was the following. They observed that it was no way close to the bell curve. The curve looked something like this.

There was a drop; it was not a bell like this. It was not like this; it was a drop like this. So, once again what is the x and y axis? x axis is the degree and y axis is the percentage of nodes that have so much degree. So, people observed that there is a drop here. Why is it happening right? So, if let us say you try drinking milk tomorrow morning and milk taste salty let us say; you will actually start wondering what went wrong, why is it why is my milk tasting, salty of all things right; something should I gone wrong right, something should be there is some background information that you are missing.

So, similarly when scientist observed that drop of the curve like this which is nowhere close to the bell curve, people thought something must be happening here that we do not seem to understand. So, let we write that now something is happening here that we do not understand. So, what could be happening why exactly is this observed. What is this even mean?

You see there are a lot of nodes with very low degree and very few nodes right; very few nodes with very high degree that is what it means right. Few nodes with high degree and then let me write this here lot of nodes with very small degree. And this startle the community of network scientist. They started wondering why this a property is being seen, what is making this property emerge.

Now, is the right time for you all to pause the video and think about the question. Why is this happening the big why? Why is it that we are not able to see a normal distribution here? Why are we seeing a drop in the curve like this?