Dealing with Materials Data: Collection, Analysis and Interpretation

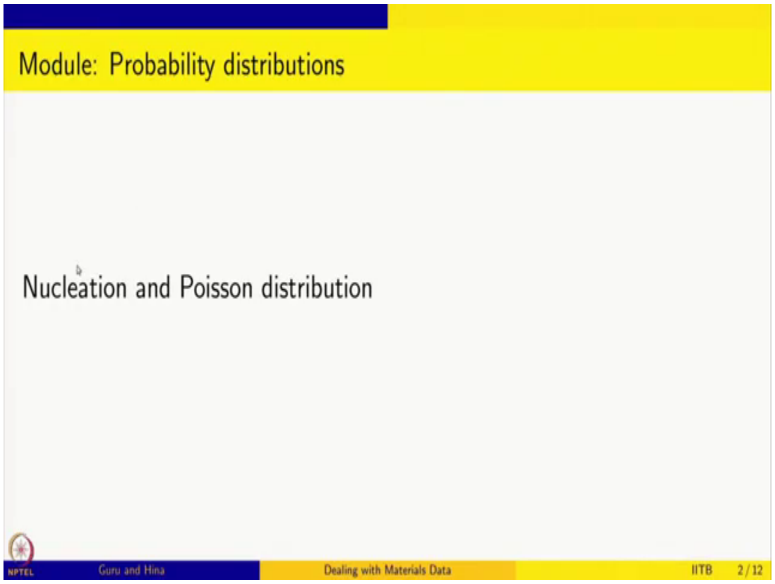
Professor M P Gururanjan
Professor Hina A Gokhale
Department of Metallurgical Engineering and Materials Science
Indian Institute of Technology, Bombay

Lecture 43

Nucleation and Poisson distribution.

Welcome to Dealing with Materials Data, this is a course on collection analysis and interpretation of data from material science and engineering. So, we are looking at the module of probability distributions.

(Refer Slide Time: 0:31)

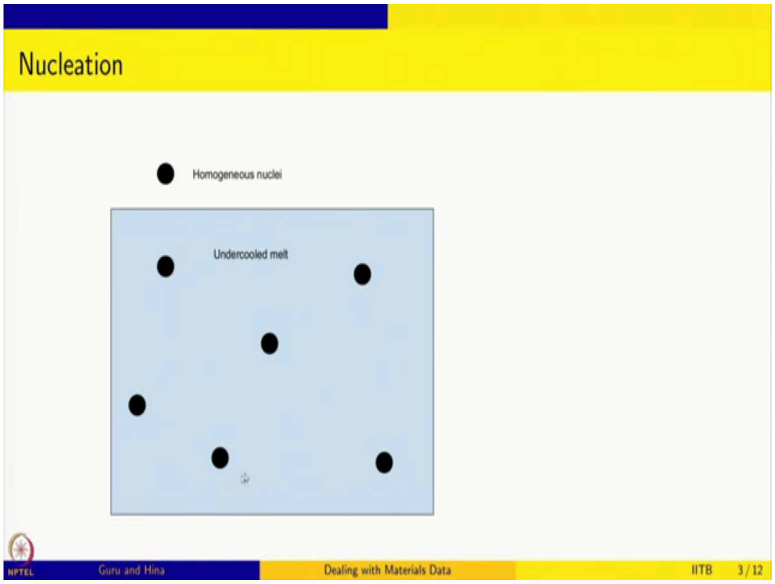


Module: Probability distributions

Nucleation and Poisson distribution

And in this session we are going to talk about the nucleation process and how it is related to the Poisson distribution nucleation is a process which is very-very important in material science and engineering.

(Refer Slide Time: 0:46)



Nucleation

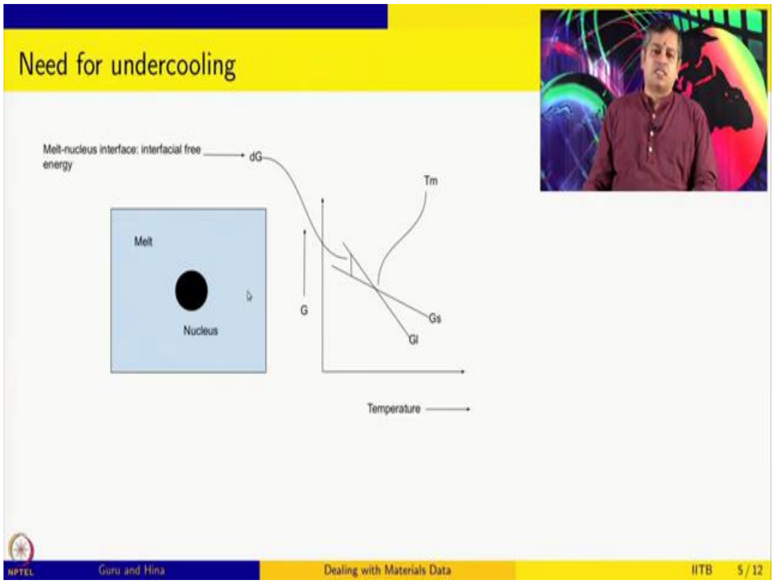
● Homogeneous nuclei

Undercooled melt

Guru and Hina | Dealing with Materials Data | IITB 3/12

Here is an example of an under cooled melt and out of which nuclei is forming. There are actually several nuclei that have found and they are willing to grow and that is how this undercooled melt is going to transform into a complete solid. What we are interested is in the process of the formation of this nuclei, and we want to have the right statistics to describe this process. So in order to describe the process again, we have to understand exactly what is happening physically and once we know that we can translate this problem in mathematical terms to the correct statistics, and then solve the problem.

(Refer Slide Time: 1:30)



Need for undercooling

Guru and Hina | Dealing with Materials Data | IITB 5 / 12

So, let us find out what this is all about. So, when does melt turn into a solid? We all have seen this that you keep water inside the freezer it turns into ice, you keep ice outside in the room it turns into water. So that is the driving force for water to be in liquid form or in solid form and that depends on temperature. So, as a function of temperature if you actually plot what is known as Gibbs free energy, then you can define what is this transformation temperature in this case it is a melting temperature, that is decided by the free energy.

So, above this temperature we find that that the liquid free energy is lower as compared to the solid free energy. So, above this temperature the system would prefer to be a liquid, below this temperature we see that the solid has lower free energy then the liquid. So, the system would prefer to be a solid. So, the melting temperature is that temperature at which these solid and liquid have the same free energy.

So, it is basically the intersection point of the free energies as a function of temperature and at the point of intersection is where we define the transition temperature In this case, because it is solid liquid transition I am calling it as a melting temperature or freezing temperature. You can also think of other transformations and this definition is more general than for this specific scenario.

So, in every case, you will find that the driving force for transformation is given by these free energy changes, the system is trying to minimize the free energy so it would like to be on this line and then on this line it does not want to continue here, because there is another state that is available with lower free energy. So, the system would prefer to ship to that. So, if you are heating the atom, heating the system, the solid will turn into a melt and if you are cooling, the liquid will turn into a solid.

So, we are looking at this process of cooling the liquid and it is turning into solid. But you can notice that there is no driving force at all for transformation at the melting temperature because the two free energies are equal. In other words, the thermodynamic definition of melting point says that if you keep a solid and its melt at this temperature, it will have a tendency neither to melt nor to solidify.

So, there will be no change in the solid melt interface at all. it will neither move towards the solid not moved towards the melt and it will remain where it is and in fact, if you go slightly lower also, it is not going to be favorable for the liquid to turn into solid unless it reaches a critical value in terms of temperature or critical under cooling as we call it, so, we call it under cooling because you have to take the liquid below the melting temperature to some extent before the liquid turns into solid.

You would expect this to be a spontaneous process, but it does not happen at this temperature you have to be below how much below you are below the melting temperature is what is known as the under cooling and a critical under cooling is needed for the liquid to turn into solid or some solid region like this to form in the melt. Why is that? So, that is what is explained in this figure.

So, you have the liquid the free energy of the solid free energy this is the melting point and you need to be in temperature below the melting temperature to some value and this difference is what is known as under cooling and this under cooling has to reach a critical value that is because at this temperature for example, it is favorable; let us consider this point at this temperature for example, it is favorable for the liquid to turn into solid, because solid has lower free energy liquid has higher free energy.

But if it forms this nucleus from the melt, there is the nucleus. So, that will have the solid free energy the, and so much of melt has turned into solid. So, there will be so much have gained in terms of energy, but there is also this interface that it has to form between the melt and this solid.

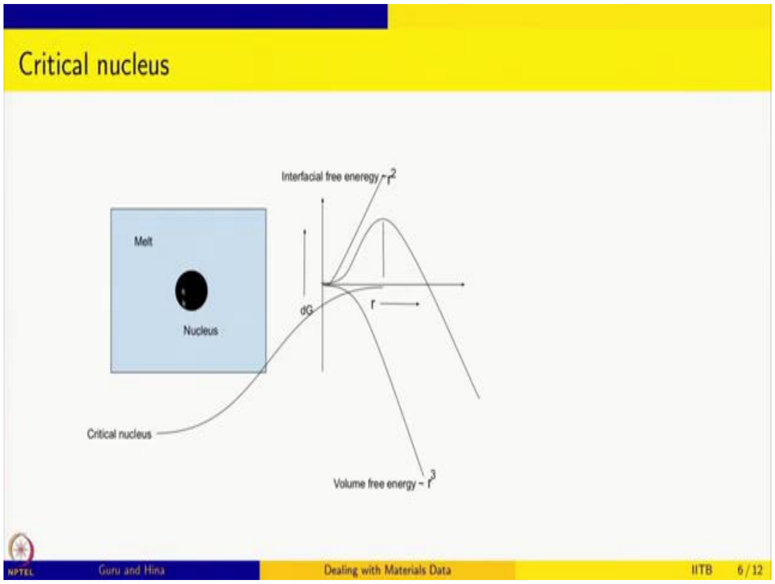
This interface cost energy for the system. So, unless the gain in energy can compensate for this, the system is not going to prefer to form this solid from the melt.

So, the critical under cooling is that temperature where you see that this gain in free energy is more than sufficient to pay for the interface cost. In that case, because it will gain so much energy and it will pay for this and it will still have gain in energy. It is going to grow. So, that is when this actually is called as nucleus. So, there is a critical under cooling that has to be achieved. So, that you will achieve critical change in free energy which will compensate for the cost you have to pay for making this interfaces.

So, till that is reached you will not have it and in all of this we are assuming that the solid is forming in the melt homogeneously that is it is not for example, forming at the mold walls are it is not forming on any impurities or any other particles that are there in the melt. If suppose there is a particle in the melt and the nucleation happens on the particle, there the map particle melt interface will already have an interfacial energy if suppose that is more than the solid, the melt interfacial energy that will have its own driving force.

So the critical free energy change mighty not valid in such cases. So, you have to also incorporate the interfacial energy associated with the melt particle interface or melt impurity interface are the mold wall melt interfacial energy and so on. So, that is why we are not considering this case we are considering nucleation from the melt and such nucleation is known as homogeneous nucleation

(Refer Slide Time: 8:29)



Critical nucleus

So, what is the critical nucleus size and so, that can again be derived. If suppose I assume that the nucleus is spherical particle, then you can see that the surface area will go as square of the radius of the particle and interfacial energy will be proportional to the surface area. So, interfacial energy will go as r square. On the other hand, volume of radius R of sphere will go as r cube and the free energy decreases that it gets because of turning so, much of melt into a solid will go as r cube.

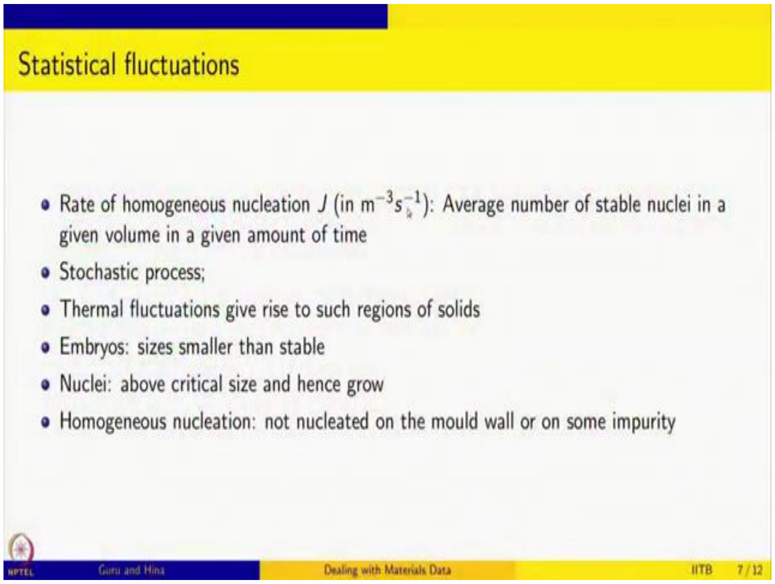
For very small r, r square will still be dominant and so r cube will not be favorable, but as the particle size becomes larger and larger, you will see that this is the point at which the contribution that comes from the volume free energy will more than compensate for what it has to pay in terms of the interfacial free energy.

So, beyond this point, so this is the changeover point at which you will have the critical nuclear size. So, if particle of this size exists, it is at a size where further increase in size is going to decrease the overall system free energy. If it is anywhere, any decrease in the solid size is what is going to decrease the overall system free energy. So, in these stages it is not preferred to have these solid particles anything here is neither stable nor unstable and beyond this these particles become stable.

So, this is the critical size of the nucleus and this is a critical dG star as we call it. So, this is obtained by writing analytical expressions assuming that this sphere. So, typically, or you can assume any shape as long as you describe in terms of the shape how the interfacial free energy and volume free energy change, and you have to minimize this quantities or take the first derivative equate it to zero and the solution you get is r star, you can show that r star is two times gamma by delta G so that delta Gv is the overall volume free energy gain that you get by turning the melt into solid and gamma is the interfacial energy.

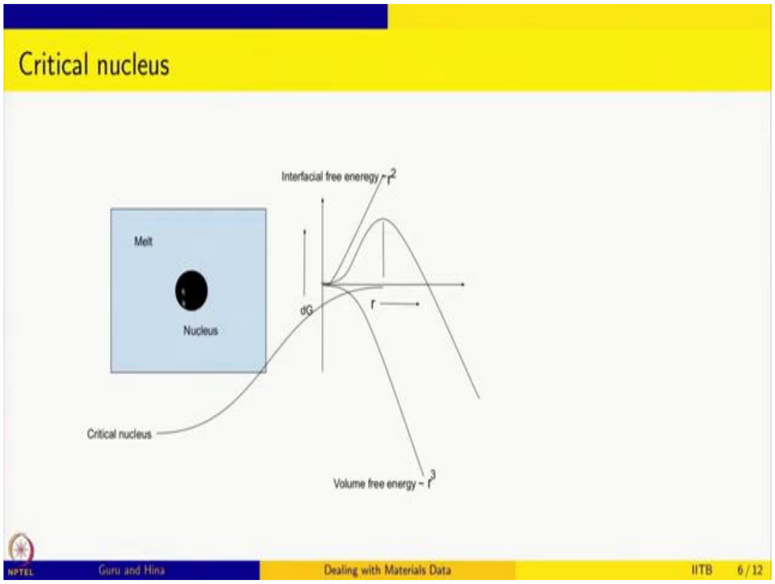
So, it also comes from what is known as Gibbs Thompson equation which says that r is the, If suppose r is the radius, then add 2 gamma by r is the excess free energy associated with that interface, so you have to compensate for it. So delta Gv happens to be two gamma by r. So, in any case, so, you, you can see that there is a critical size and beyond that size the nucleus will grow and this process of turning the melt into a solid by formation of such a stable particles is what is known as nucleation. Specifically we are looking at homogeneous nucleation.

(Refer Slide Time: 11:48)



Statistical fluctuations

- Rate of homogeneous nucleation J (in $\mathrm{m}^{-3}s^{-1}$): Average number of stable nuclei in a given volume in a given amount of time
- Stochastic process;
- Thermal fluctuations give rise to such regions of solids
- Embryos: sizes smaller than stable
- Nuclei: above critical size and hence grow
- Homogeneous nucleation: not nucleated on the mould wall or on some impurity



Critical nucleus

Now, the rate of homogeneous nucleation j in meter cube per second units is the average number of stable nuclei in a given volume in a given amount of time. Now, the form of nucleus, nuclei is a stochastic process. What gives rise to such particles? In the first place is the thermal fluctuations. Because of thermal fluctuations, different sized solid like particles are always forming in the melt.

They are forming in fact, because they are based on the thermal fluctuations they are forming lower temperatures, but if suppose your temperature happens to be lower than the melting temperature, then when the system accesses those states, it realizes that it can decrease its free energy. But if suppose it accesses these sizes, then at these sizes, it is going to find that it is not favorable.

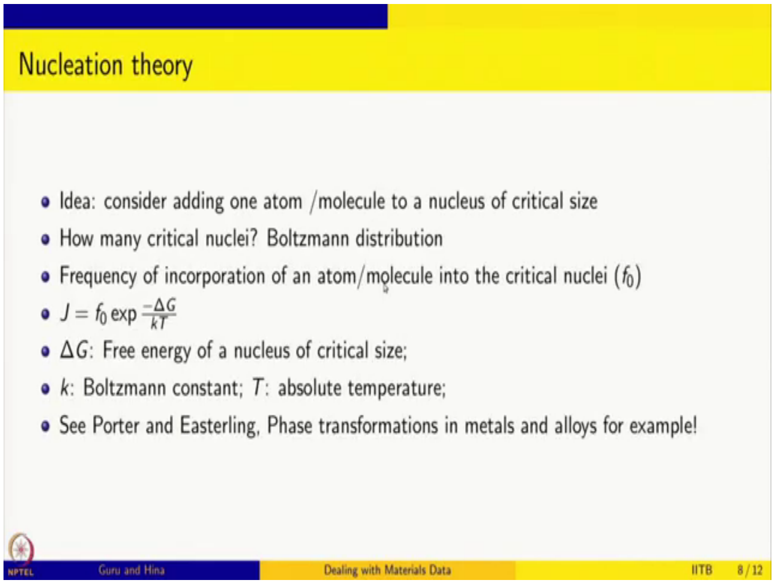
So, they are willing to shrink back, but if suppose, it happens to access this size, then it knows that it had already decreased the system energy. So, it will prefer those states and so, it will continue to do that. So, this is what is happening and so, these thermal fluctuations that give rise to some regions of solids, if the size of these particles happened to be smaller than the stable size then they are called embryos they keep forming and they keep dissolving back into the melt and they are called nuclei if they are about the critical size and so they can grow.

So, homogeneous nucleation is what we have assumed that we have not assumed that the nucleation is on mold wall or impurity or anything like that. So, in the case of homogeneous nucleation if that is the rate of nucleation how do we get this rate? So, the idea so, in principle you have to consider the formation of different sized particles and their dissolution and are building up one by one, you know the, the particle is growing in size, and at every stage it has definite probability to shrink and some finite probability to grow and you can track this and find out when it actually goes over the critical size, that is one approach.

But there is another idea which says that, “Ok, because of thermal fluctuations, what is the probability of finding a particle of exact critical size r star.” And now to this solid if you add one atom or molecule it will become a nuclei, ok. So, if you know the probability or frequency of incorporation of one atom or molecule into critical nucleus and that is of some f0, then the thermal fluctuations which give rise to these solid particles in the melt that is described by the Boltzmann distribution.

So, like I said earlier we are also going to talk about Boltzmann distribution and, and the other statistical distributions like for me direct and cosine stand at some point, but this Boltzmann distribution is described by exponential minus del G by kT, kT where k is the Boltzmann constant and T is absolute temperature, del G is the free energy of the nucleus of critical size in r schematic. This is del G and this is r star.

(Refer Slide Time: 15:12)

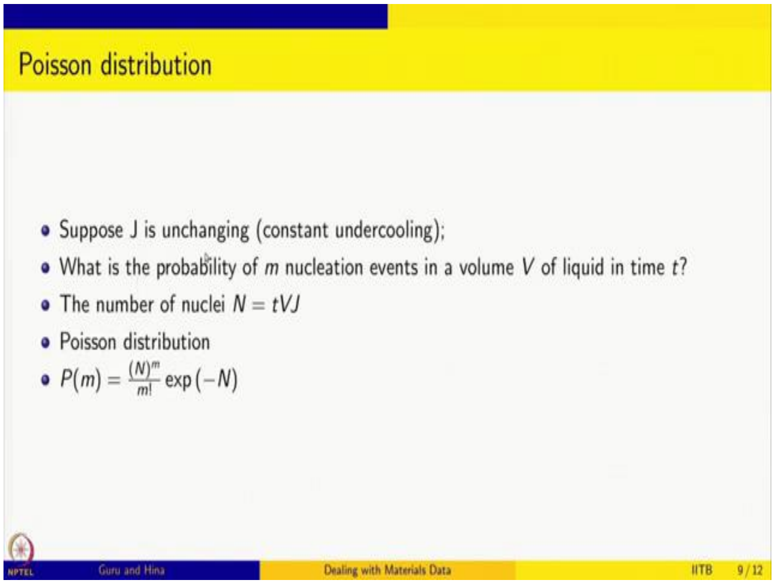


Nucleation theory

- Idea: consider adding one atom /molecule to a nucleus of critical size
- How many critical nuclei? Boltzmann distribution
- Frequency of incorporation of an atom/molecule into the critical nuclei (f_0)
- $J = f_0 \exp \frac{-\Delta G}{kT}$
- ΔG: Free energy of a nucleus of critical size;
- k: Boltzmann constant; T: absolute temperature;
- See Porter and Easterling, Phase transformations in metals and alloys for example!

$$J = f_0 \exp(-\frac{\Delta G}{kT})$$

(Refer Slide Time: 15:33)



Poisson distribution

- Suppose J is unchanging (constant undercooling);
- What is the probability of m nucleation events in a volume V of liquid in time t?
- The number of nuclei $N = tVJ$
- Poisson distribution
- $P(m) = \frac{(N)^m}{m!} \exp(-N)$

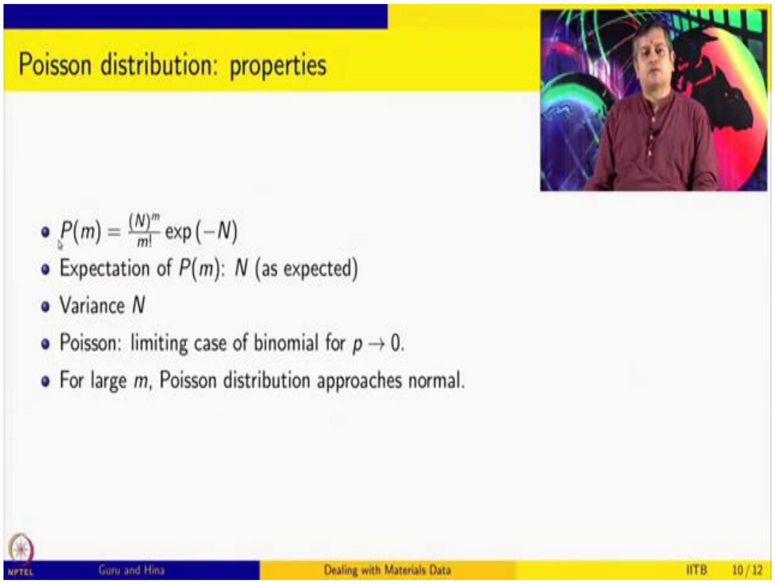
Now, suppose J is unchanging that there is constant under cooling. So, this is again another approximation, this is not true always, because if you form a solid and if there is some latent heat change, then the temperatures can change and so on and so forth. So, there are lots of things that can happen, so, the intercooling might not remain constant, but if you assume that under cooling remains constant, so the homogeneous nucleation rate reminds a constant.

Then we are asking the question what is the probability of m nucleation events in a volume V of liquid in time t? That is the question we are asking. So, the number of nuclear N will be t times V times J, because J is per meter cube per second. So, if you multiply by the volume and time you will get the number of nuclei that will nucleate in volume V at in time t, and we are trying to get this probability.

So, this will be given by probability distribution which is known as Poisson distribution. So, Poisson distribution is also distribution that you will find, for example, for accidentals and errors and things like that. So, we will also look at one such example later.

$$P(m) = \frac{(N)^m}{m!} \exp(-N)$$

(Refer Slide Time: 17:26)



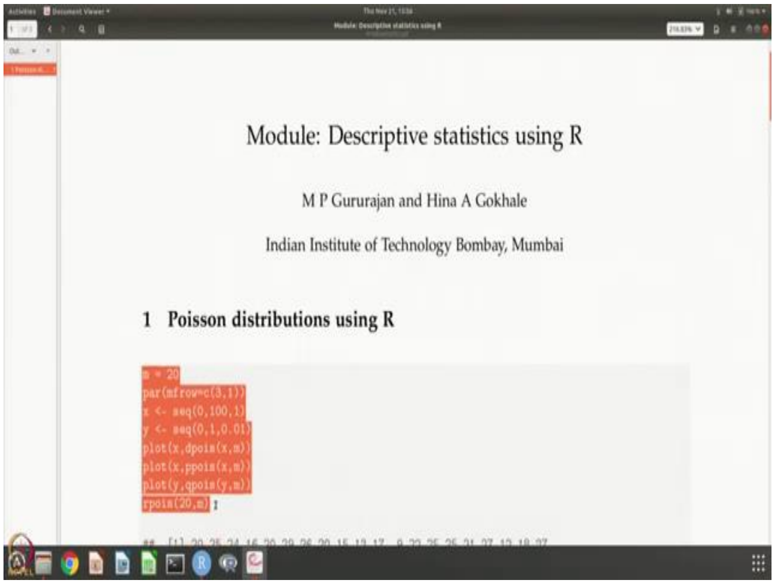
Poisson distribution: properties

- $P(m) = \frac{(N)^m}{m!}\exp(-N)$
- Expectation of $P(m)$: N (as expected)
- Variance N
- Poisson: limiting case of binomial for $p \to 0$.
- For large m, Poisson distribution approaches normal.

So, this is the probability distribution for Poisson distribution function it so happens that the expectation value for the Poisson distribution and the variance both happened to be n ok and can be thought of as a limiting case for binomial for the probability tending towards zero and for large m, the Poisson distribution actually approaches normal distribution.

So, the Poisson distribution using rpois is the command and dpois, ppo1is, qpois and rpois. As usual they give probability density, cumulative distribution function, quantile function and rpois actually gives you the random variants of poison distribution.

(Refer Slide Time: 18:21)



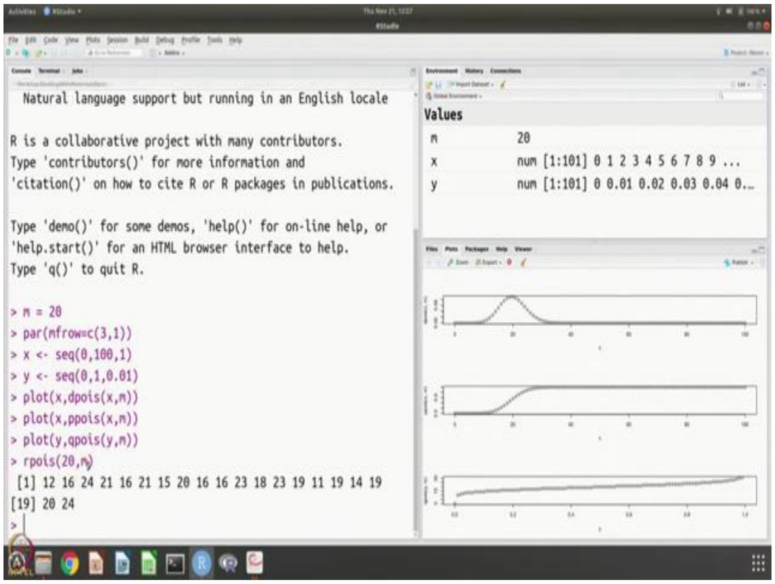
Module: Descriptive statistics using R

M P Gururajan and Hina A Gokhale

Indian Institute of Technology Bombay, Mumbai

1 Poisson distributions using R

```
m = 20
par(mfrow=c(3,1))
x <- seq(0,100,1)
y <- seq(0,1,0.01)
plot(x,dpois(x,m))
plot(x,ppois(x,m))
plot(y,qpois(y,m))
rpois(20,m)
```



```
  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> m = 20
> par(mfrow=c(3,1))
> x <- seq(0,100,1)
> y <- seq(0,1,0.01)
> plot(x,dpois(x,m))
> plot(x,ppois(x,m))
> plot(y,qpois(y,m))
> rpois(20,m)
 [1] 12 16 24 21 16 21 15 20 16 16 23 18 23 19 11 19 14 19
[19] 20 24
>
```

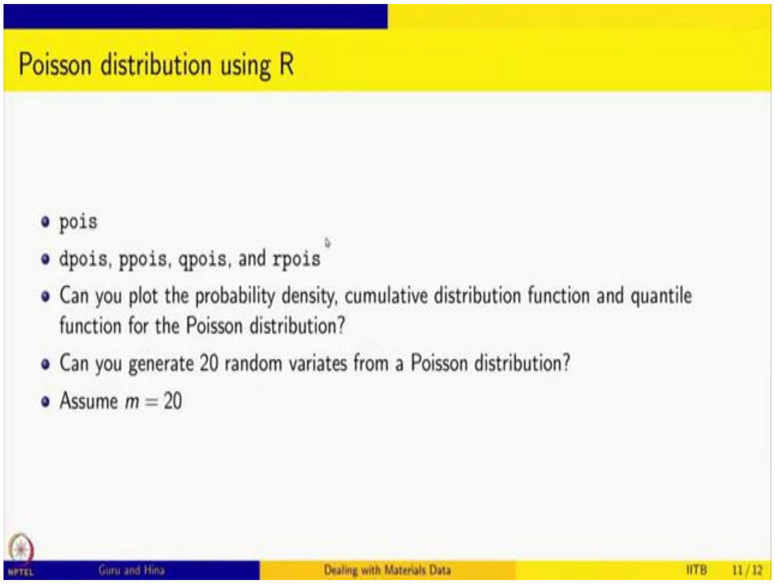
Values

```
m        20
x        num [1:101] 0 1 2 3 4 5 6 7 8 9 ...
y        num [1:101] 0 0.01 0.02 0.03 0.04 0...
```

Let us assume m to be 20 and try to do this commands using R. So, M is 20 and again we are going to make three plots and those are for the density cumulative distribution and the quantile function and the Poisson distribution requires just the parameter. So, the here is the property distribution here is the cumulative distribution and here is the quantile function. Of course, you can also get the random variates and the way to get the random variates is just to say, rpois.

So, I want 20 random variates and the V is the parameter so you can give, so you have 12, 16, 24, etc., of the random areas which are picked from the Poisson distribution.

(Refer Slide Time: 19:49)

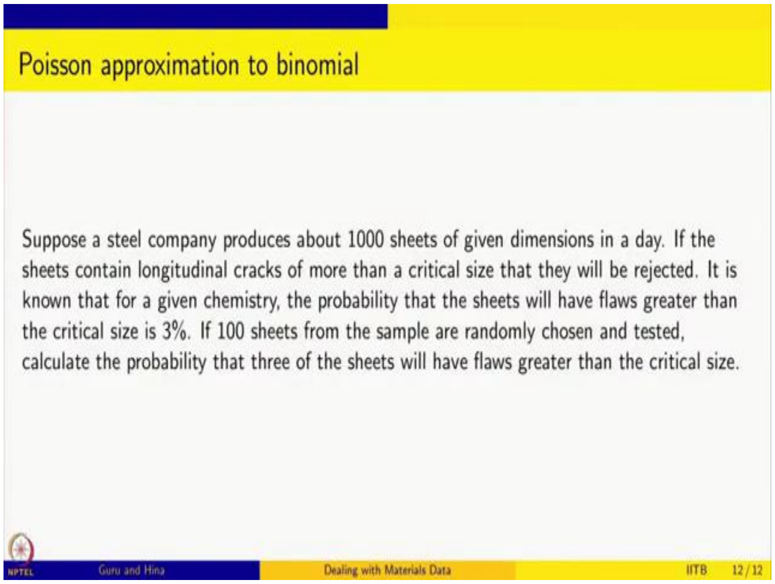


Poisson distribution using R

- `pois`
- `dpois`, `ppois`, `qpois`, and `rpois`
- Can you plot the probability density, cumulative distribution function and quantile function for the Poisson distribution?
- Can you generate 20 random variates from a Poisson distribution?
- Assume $m = 20$

So, let us go back. So, this is how you deal with by now I am assuming that you are comfortable with these commands, because there is a third or fourth time we are using similar commands and so let us also look at another problem. So we looked at the nucleation problem, and we found that nucleation events can be described by the Poisson distribution for homogeneous nucleation assuming constant nucleation rate, there is also another problem set of problems where Poisson distribution is useful or it can be used an approximation for binomial distribution.

(Refer Slide Time: 20:21)



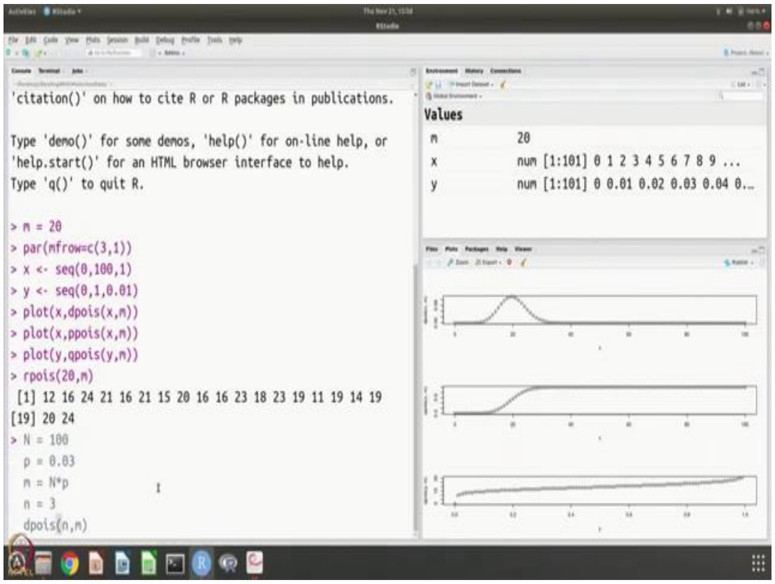
Poisson approximation to binomial

Suppose a steel company produces about 1000 sheets of given dimensions in a day. If the sheets contain longitudinal cracks of more than a critical size that they will be rejected. It is known that for a given chemistry, the probability that the sheets will have flaws greater than the critical size is 3%. If 100 sheets from the sample are randomly chosen and tested, calculate the probability that three of the sheets will have flaws greater than the critical size.

And here is one example, suppose a steel company produces thousand sheets of a given dimension in a day, and if the sheets contain some longitudinal cracks of more than a critical size, they will be rejected and for a given chemistry, if you make sheets out of this composition, we know that the probability of sheets having flaws greater than a critical size is about 3 percent let us say.

Now, if you pick 100 sheets of samples randomly from these thousand sheets that you make, what is the probability that three sheets will have flaws greater than the critical size? So, this is the question and of course, you can think of it as a binomial distribution problem and you can also calculate it as a approximate by Poisson.

(Refer Slide Time: 21:14)

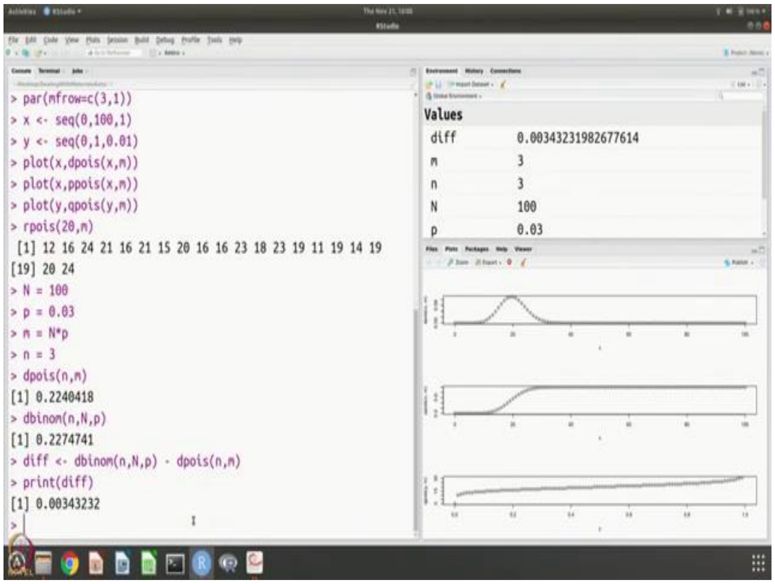


```
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> m = 20
> par(mfrow=c(3,1))
> x <- seq(0,100,1)
> y <- seq(0,1,0.01)
> plot(x,dpois(x,m))
> plot(x,ppois(x,m))
> plot(y,qpois(y,m))
> rpois(20,m)
 [1] 12 16 24 21 16 21 15 20 16 16 23 18 23 19 11 19 14 19
[19] 20 24
> N = 100
  p = 0.03
  m = N*p
  n = 3
  dpois(n,m)
```

```
Values
 m          20
 x          num [1:101] 0 1 2 3 4 5 6 7 8 9 ...
 y          num [1:101] 0 0.01 0.02 0.03 0.04 0.…
```



```
> par(mfrow=c(3,1))
> x <- seq(0,100,1)
> y <- seq(0,1,0.01)
> plot(x,dpois(x,m))
> plot(x,ppois(x,m))
> plot(y,qpois(y,m))
> rpois(20,m)
 [1] 12 16 24 21 16 21 15 20 16 16 23 18 23 19 11 19 14 19
[19] 20 24
> N = 100
> p = 0.03
> m = N*p
> n = 3
> dpois(n,m)
[1] 0.2240418
> dbinom(n,N,p)
[1] 0.2274741
> diff <- dbinom(n,N,p) - dpois(n,m)
> print(diff)
[1] 0.00343232
>
```

```
Values
diff    0.00343231982677614
m       3
n       3
N       100
p       0.03
```

So let us do that. So, what do we have we have 100 and the probability is 0.03 and so if you have 100, then at least three you expected to fail, and you are looking for probability for, when you test 100, 3 of them will be failing and so the Poisson distribution for 3 failures and for the probability, the mean, 3, right in this case, this also happens to be 3 is 0.224.

Now, you can also calculate this quantity as a binomial. What is the binomials and that is here, right, you can calculate the total is 100, and you are looking for three out of it with a probability of 0.03. So, that is 0.227. So, you can see that the difference between the two is at the third decimal place, some 03 or something. So you can actually calculate the difference between this and this, ok.

So, it is 0.003. So you can see and you can see here also so the binomial distribution as p tends to 0 you can think of it as approaching the Poisson distribution. So, because of which for small p, you can use the Poisson distribution to calculate these probabilities. So, this is one of the properties of the Poisson and, and we will also look at some other properties later after we do the normal distribution.

So, we have found that this is the last of the discrete distributions, we have looked at Poisson distribution, and this distribution has lots of use. It is very important in studies like nucleation and it is also useful for finding out defects in a manufacturing process, for example, and it can also be used for approximating the binomial for very small values of p, the probability that goes into the binomial distribution function.

So, to summarize, we have looked at probability distributions, specifically we have looked at discrete probability distributions, we have looked at Bernoulli trials, binomial, negative binomial, hypergeometric, and Poisson distributions and we have looked at least two very important cases, where these distributions are of use. One of them is a microscopy tactic to get the composition of very small volumes.

So, that is the atom probe technique, where we found that negative binomial hypergeometric, and things like that are useful and where we also did some error analysis along with the distribution. So, it is an interesting mix of problem and we have now also looked at Poisson distribution and how it is useful for understanding things like nucleation. So, with this we are going to conclude the discrete distribution scenario.

There is one more uniform distribution which can be either described r containers. So, we will come back to it at some later point and now we are going to move to continuous probability distributions of which are the most important is of course normal. So, we are going to start with normal in the next session. Thank you.