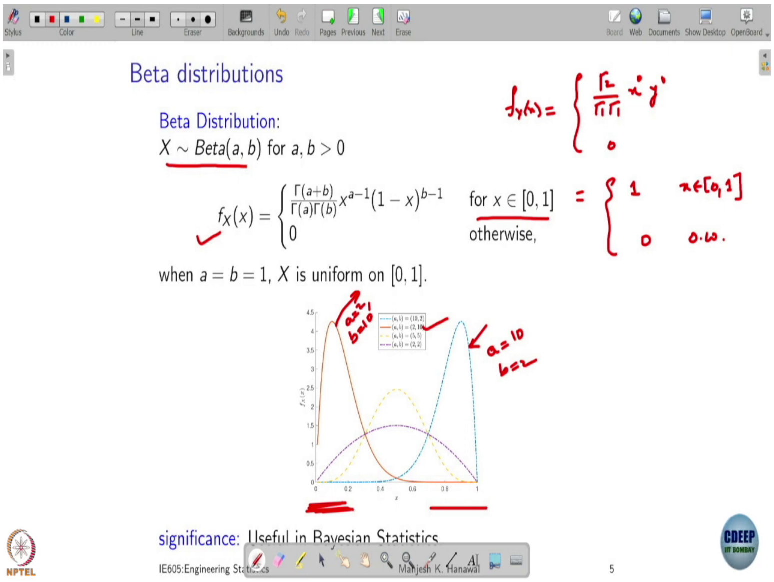
Engineering Statistics
Professor Manjesh Hanawal
Department of Industrial Engineering and Operation
Indian Institute of Technology, Bombay
Week 4
Lecture 20
Beta distributions and Exponential families

(Refer Slide Time: 00:14)



Now, another distribution is called a beta distribution. This is another expansions we are doing, like gamma, with gamma distribution we broadened our scope of distributions like similarly we do beta distributions. Beta distributions, we are going to denote beta a, b, there are 2 parameters a and b and this PDF is defined like this.

And notice that for beta distribution it is only defined in the positive interval but restricted to 0, 1. Now, if you make this a and b and same as 1, this, suppose now put a and b as same what you are going to get, f of x equals to gamma 2 gamma 1 gamma 1 x0 and y0. So, what is the value of gamma 2?

Student: 1.

Professor: 1. So, and what is the value of gamma 1? 1. So, this is like 1 if x is between 0, 1 and this is like 0 otherwise. What this distribution corresponds to?

Student:Uniform [0, 1].

Professor: Uniform 0, 1. So, by putting a and b equals to same to 1 you have recovered uniform distribution. Like that you can choose any distributions you like in this. And again, that is what like I have put some special cases here. The first thing I put is like let us take this one a, b 2, 2 and a, b 5, 5. So, in these 2 cases I have put both a and b to the same value. If you see that when a, b are 2, 2 it is like this.

And if I am increasing the value of a, b to 5, this is like a kind of narrowing down. The peak is always at the middle 0.5, and similarly, you can imagine that if I increase a, b value to 10 it may be looking more and more concentrated around that point. Now, the case where a and b are not different like I have again considered a 2 symmetry 2 cases when a is 2 and b equals to 10 you can read the value on the screen. Let me write here. So, here...

Now, consider this case, this is corresponding to this curve. In this curve a is 2 and b is 10. And this curve here is the, here a is 10 and b is 2. So, notice that when they are similar, the a and b values are not the same some kind of skewness is happening in the peak of this curve like when b is large the peak is mostly towards your left and when a is large the peak is towards your right. And when a and b are same, the peak is exactly in the middle.

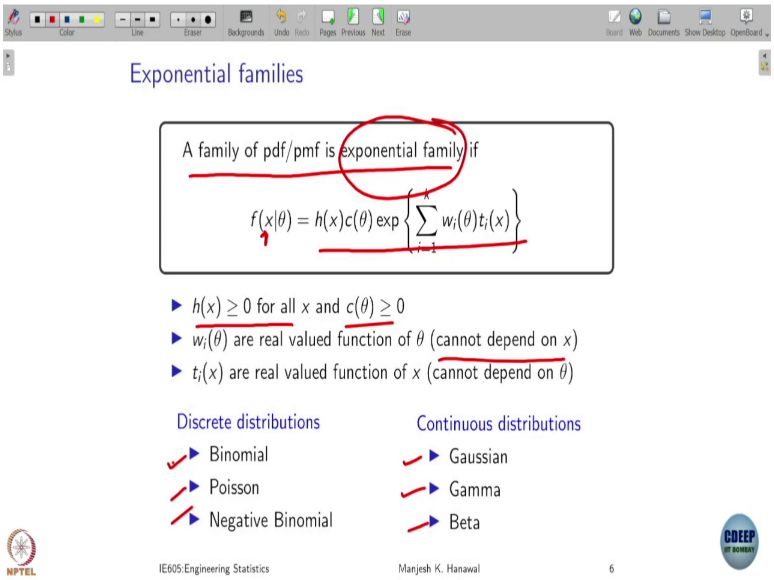
And as a very special case when a and b are equals to 1 it is like a flat curve, no peak is there, because we just showed that is a uniform in this. So, again this is like a generalizing thing. Like from uniform distributions now we have parameterized and you are able to capture so many different distributions.

And this is often used in Bayesian statistics, like I said earlier, like if I do not have any prior information, I am going to take uniform like everything is equally likely. But if you have initially some prior information that smaller values are going more likely than the larger values. What you are going to do? You are going to take b as larger than a or a larger than b?

Student: b larger than a.

Professor: b larger than a like in this case. Like here I told you like the ones which are smaller are more likely that is why here b is going to be larger. On the other hand, if you have some prior information that larger values are going to be more likely then you will put take a to be larger than b. Next.

(Refer Slide Time: 05:22)



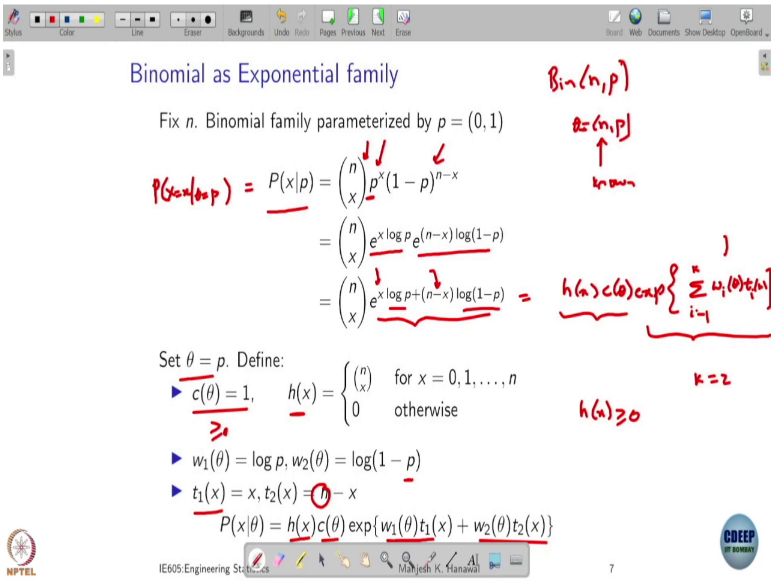
Even though we talked about so many different distributions and we talked about their various parameters, it actually happens that most of the distributions we have talked so far, they can put in a very compact way. And they all belong to one special class of distributions called as exponential families. Now, let us say, let us write our generic probability density function, let us say I have been given a distribution which is parametrized by theta, and that I am going to express in this format.

Now, let us try to decipher what I have written here. There are 2 things h function, c function, w function and ti functions. So, h is simply a function here which depends on the point. Notice that here deliberately I have written given theta. So, this distribution this PDF is a parameterized. This depends on a given theta and now we have to define it for all possible values of x. Now, this is, this h function is going to be positive for all x.

And the parameter, for that parameter, that c of theta function is a positive valued. And now, this wi here, this wi is a function of theta and this is real valued. And this, when I write it like this, this wi cannot depend on xi, this wi only function of theta. And this last one ti this is again a really valued function, but this only depends on theta but not on x. If a PDF function parameterize by theta I can write like this, this is called exponential family.

We are not hardcoding what is should be h, c, wi and ti. We are just saying that h of x should be positive for all x, c of theta should be positive and wi should be just depends on theta not on x and ti should depend just on x and not on theta. Now, you will see that all the distributions that we talked in the discrete case and all the continuous one that we have talked so far, Gaussian, gamma and beta they actually fall into exponential family. They can all be expressed in this format for some values of h, c, ti and wi. So, this exponential families that is why pretty handy because it covers large distributions that we have already studied.

(Refer Slide Time: 08:33)



Now, let us look why is that is the case, why is that we are saying that binomial belongs to exponential. Now, let us try to see these binomial distributions. I am able to write it in the form distribution associated with the, that is the common template of the PDF for exponential distribution. So, binomial is what, n comma p. So, the parameters in this case theta are n and p. So, now assume that among these 2 components, this one component is known.

So, my theta parameter is actually only p. This is just to simplify things. Now, if you write your PMF we already know what is the PMF of a binomial distribution. This is, what is this, this is basically like a probability that X equals to x given that your parameter theta is p. So, what are basically saying that my random variable takes value x under this parameter p that is n choose x p to the power x 1 minus p n minus x.

This is the definition of binomial random variable, or this is the probability mass function of your binomial distribution. Now, let us manipulate this. Can I write p power x as e to the power x log p?

Student: Yes.

Professor: Nothing changes. This I have also written like this. And now, this product of 2 exponentials I have just written as exponential and taken e to the power some of the exponents. This is also correct. This is just a property of exponents. Now, let us see that. I can write it in this form that I wished. What was that? That is h of x, c of theta, then exp, my i equals to 1 to, let us, how, what was the index we use, i and then we said wi of, it was theta, theta and ti of x.

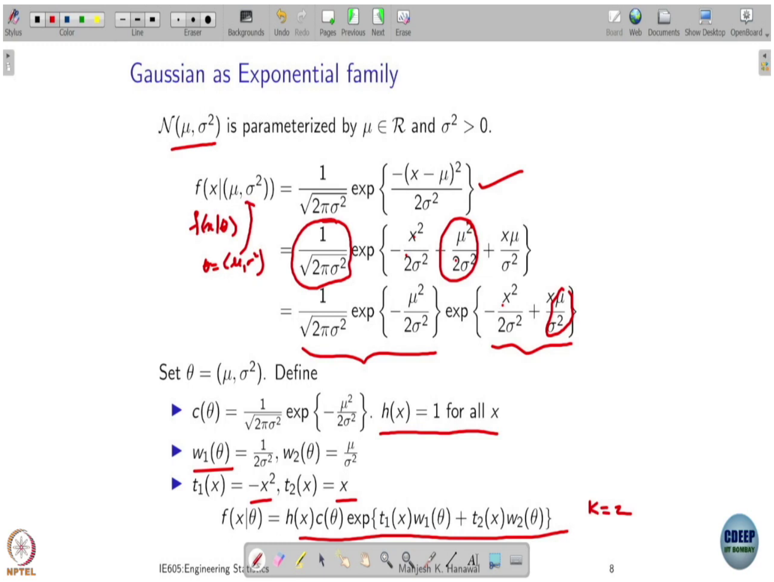
Let us say this can be represented as, we said K let us say, this can be represented as in this format. Now, let us say my parameter theta is p. Now, first decide, so, exponential, so, this exponential, we will map it to this exponential terms within, but now, first, let us the, this part, I have to map it to the factors multiplying by exponential. So, h of x c of theta is simply n choose x, but here n is known only thing x this.

So, then in that case, I will choose h of x equals to simply this n choose x and this is true for x0 1 to n, this being a binomial I will choose x0 1 to up to n and 0 otherwise. And this requirement that my h of x has to be greater than or equal to 0 now that is fulfilled. This guy is positive. Now, for c of theta, this c of theta has to depend only on theta. Now here, there is nothing so it just take c of theta to 1.

And this is also positive quantity that requirements c of theta is positive is also 1. Now, let us see that, now let us focus on the terms inside exponential. I know that w1, so each of these wi's has to depend only on theta. What I will do is for that these things, I am going to now define w1 theta to be log p and w2 theta to be log 1 minus p. And now both w and w2 depends only on the parameter p they are not dependent on x.

And now if to define t1x, t1x I will take this x and t2x I will take it as this n minus x. Now, this only depends on x and not on the parameter. Notice that n is a parameter, but we are assumed it to be fixed. So, the parameter is only p and this does not depend on this p. Now, have I put it in the format that I want? And here K is 2 because I am adding only 2 terms in the exponent. Now, that is why I have now put this probability mass function as h of x, c of theta in this form, this is the required form for it to be called exponential family.

(Refer Slide Time: 14:18)



So, that was for a discrete case. Let us now look into one continuous case. So, for this we will take Gaussian, so Gaussian is going to be parameterized by two values here mu and sigma square, so theta here is that is what like this fx given theta and this theta is mu and sigma square here. That is what we have written. We know that its PDF is like this. Now, let us see this could be expressed in the exponential family.

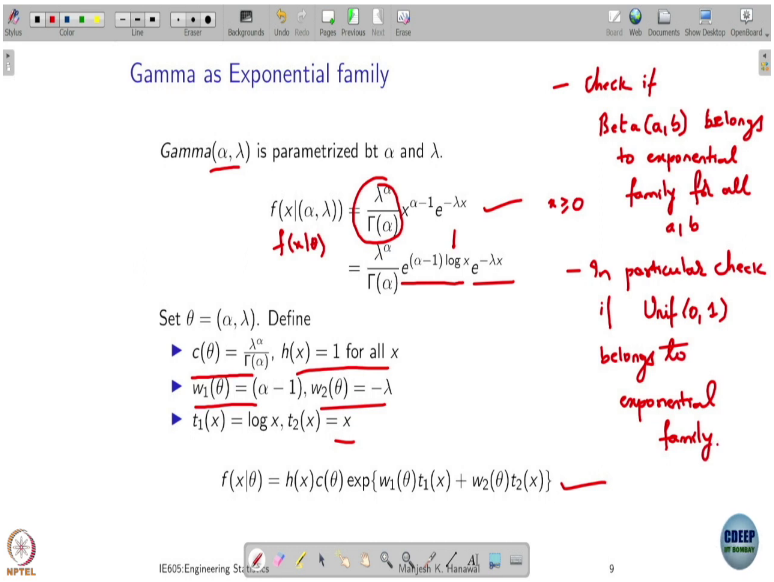
So, to do that we have first I have simplified this exponent inside, I have expanded the square, you will get this. Now, let us see this term here depends on sigma squared, which is my parameter. So, I will take this c of theta. Now, one more thing I have done, when I expanded this, this square mu squared 2 sigma squared is there, so I have also pulled out this. So, this entire thing here, it depends only on the parameter mu and sigma square and not on x.

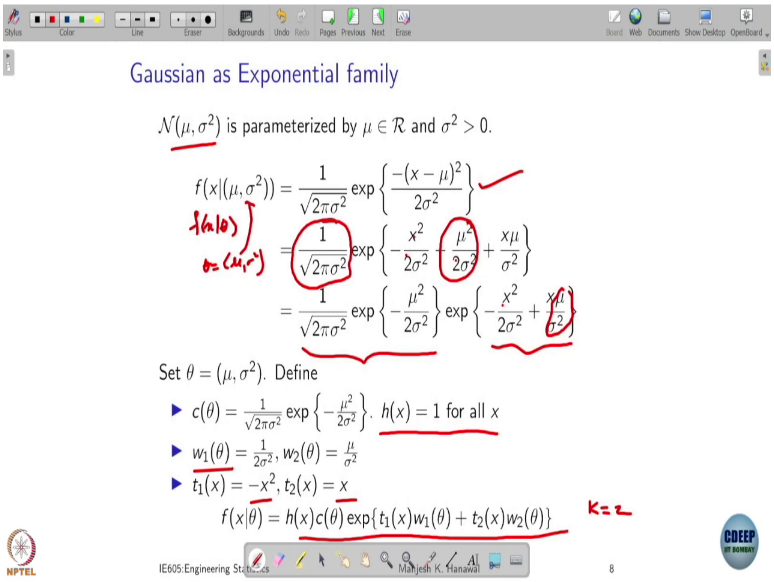
So that is what I am going to take c of theta to be this entire quantity, this quantity. And I will simply define h of x to be 1 for all x. So, notice that when I pulled out this, this did not depend on x, but the other terms, this and this, this term depends both on x and sigma square and this depends on x and parameters.

So, what all the, wherever the things both depend on x as well as the parameter I kept it inside my exponential and where it only depends on the parameter I have pulled out and this helped me the simplification. And now, if you now look into this part I can take mu and theta to be this 1 by sigma square here and mu to theta to be mu by sigma square part this portion here and take t1x to be minus x square and t2x to be x, this portion x.

Now, if you notice that with this definition, I am now able to write f of x given theta exactly like this, again for k equals to 2 this is what is the definition of exponential family. So, that is what like we have again verified that the Gaussian is also belongs to exponential family.

(Refer Slide Time: 17:26)





Now, quickly check whether this gamma function we just discussed today is also belongs to exponential family. Now, we said gamma is parametrized by alpha and lambda and this is its function when x is greater than or equals to 0. And when x is less than 0, we also have to handle a case where it is 0 by appropriately defining. Now, what I have done here, this first thing we have kept it same, but now x to the power alpha minus 1 I have rewritten as e to the power alpha minus 1 log x. Can I do this?

This is a property of exponents and log that I have used and e to the power lambda x I have written. And now let us see if I can come up with a h function, c function, ti function and wi function as required in my exponential family. So, one can define c of theta to be alpha to the power, lambda to the power alpha gamma alpha h of x equals to 1. Now, w1 of theta to be alpha minus 1, w2 to be minus alpha. So, here w function only depends on your parameters.

And t1x can you take as log x and t2x you can take it as simply x. Now, if you could expand this, now, this is, this f of x given theta where theta is alpha lambda you are able to write it like this. So, this is again exponential family. Now, an exercise for you check if beta a, b is belongs to exponential family for all a, b. And see that in particular check if uniform 0, 1 belongs to exponential family. So, please check this, do not skip it.