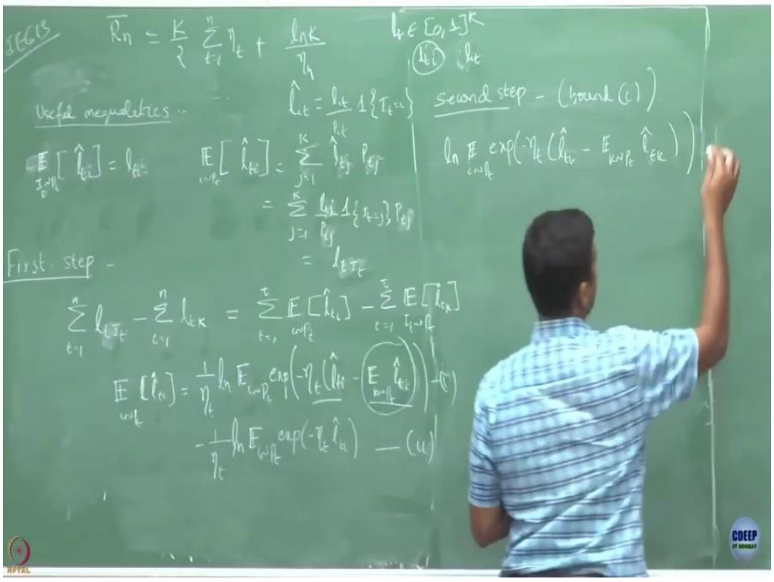
Bandit Algorithm (Online Machine Learning)
Prof. Manjesh Hanawal
Industrial Engineering and Operations Research
Indian Institute of Technology, Bombay

Lecture – 15
Regret Bound of Exp3

So, this proof we are going do it in multiple steps, we just split it into three-four steps that is easier to follow. And again the proof goes along the similar lines as we did in the weighted majority algorithm that is expert prediction with expert advice, but in this case we have to account for the case that I am dealing with the estimators nor the actual loss values ok.

(Refer Slide Time: 00:45)



Our goal is to show $\overline{R_n} \leq \frac{K}{2} \sum_{\{t=1\}}^{n} \eta_t + \frac{\log K}{\eta_n}$ by ok. So, first thing useful inequalities. First if I am going to look into the expectation of $\widehat{l_{it}}$, notice that this l_{it} itself is a random quantity right, because the way it is defined. l_{it} is defined as what is this? So, I_t is the random quantity, so depending on its randomness there is a induce randomness on this l_{it}.

So, now this is a random and if I am take expectation with that randomness. So, this It here is distributed according to P_t right. If I am going to take that what is this quantity is? l_{it} right, we said this is already an unbiased estimator. So, this is going to be l_{it}.

But now if I look into this quantity, but now i treat this to be a random variable i index with which I am going to take this expectation as, then let me treat this value, now what is this quantity? So, let us write down this what is the meaning of this quantity this is going to say that, so this i is a random variable, so i is going to be drawn according to this distribution P_t.

So, i that is why I am going to take summation over j equals to 1 to k, and then it is l_{jt} and then P_{jt} here right. So, this is the meaning of this expectation here. If you are now going to plug back this quantity here, what is this? And it is still I_t equals to j and P_{jt}, right. I have just plug that the value of this estimator here.

So, what is this quantity? So, this indicator remains whenever I_t equals to j; for all other it vanishes, but this I_t is random quantity. So, then this quantity is going to be $l_{I_t t}$. Wherever this guy is there this term remains; everywhere it is going to be 0 ok.

So, with this let us proceed. So, this is we are going to call as first step, I am interested in bounding $l_{I_t t}$. See in terms of the notation I am slightly messing up notice that like what I am doing is when I say l_t, l_t is a vector. And when I say l_{ti}, this is the i th component of this vector. So, sometime I am writing it as l_{it}. For example, here I write. So, let us try to follow the same convention I hope let us take this.

So, here it is just like which component we are going to treat it as random variable right. So, here when you wrote this is your fixing an i, for every i this is your random variable right, now looking at the expectation of this. Now, here, yes this is random variable, but you are looking at a further taking expectation with respect to the i here, so that is why whatever you got here it itself still a random variable right. And you see this why is this is useful, and I write the further steps ok.

So, let us take fix one particular action k, this is the total loss you are going to incur if you are going to deal with kth action or the if you are going to pay the kth action. And this is the total loss you are going to incur if you are going to play as per your algorithm I_t algorithm that says to play I_t around t.

Now, using this notation whatever I have here, what I can write, this is going to be minus. And what is this? This quantity I am going to write it as this, sorry this should be k here, it is fine. I have written this loss difference in this two cumulative loss in terms of their expected value that are induced by the randomness of your algorithm ok.

Now, we are going to write this quantity over here in this following fashion ok. Let me write it first I will discuss ok. So, I did some strange manipulation here ok. First notice that suppose you for time being just forget this part ok; if you just forget this part and basically this and this negate, I may get right, if you for if you do this.

Now, let us take this part notice that this part is an expectation here. So, if you are going to take exponential inside, then this quantity is an already constant because this is already taken to be expectation. So, this expect when you take this exponential on this, this is like a log of a constant, then it log of exponential simply becomes this quantity which is on the left hand side ok. And this is just you can see that this part if you just ignore, because this is going to get nullified with. If you just focus on this part it is nothing but log of exponential of this quantity, but this exponential of this quantity is constant because there is already taken expected value.

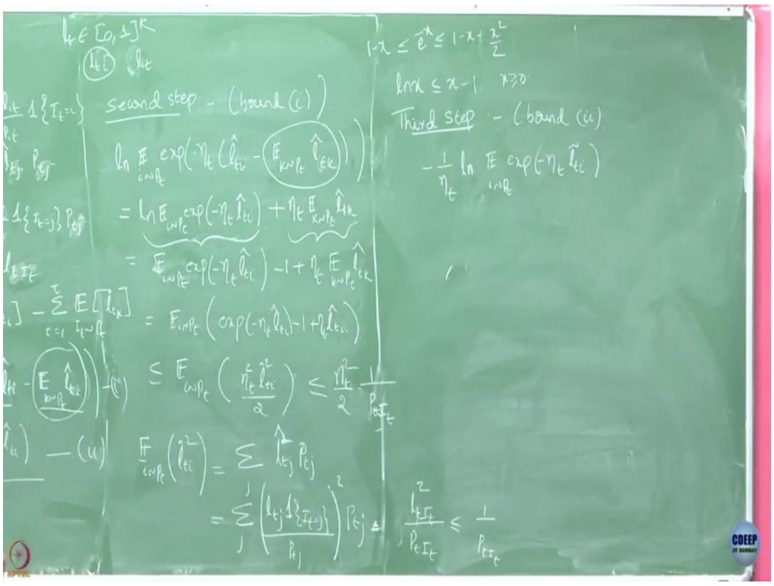
So, this is simply going to be log of expectation of this quantity because this expectation will not matter, because it is already a constant. So, log of exponential will nullify this log will nullify this exponential, what you will end up with simply minus eta times this quantity, this eta and this eta will cancel, and you will just end up with this quantity.

So, why we did this? Why do we did this circus? The circus is to make sure that we express them in terms of their moment generating functions you know moment generating functions or characteristic functions. So, it is basically log of expectation of exponential of that random variable ok. Here basically I have written this l_{ti} to be the moment basically this first part is the moment generating function of this quantity l_{ti} tilde, this is a random guy right. This is basically the moment generating function of that.

And now what is this quantity? This is nothing but the mean even though I have written k here, but this is just an index, but this is nothing but the mean of this quantity. So, I basically subtracting mean from it, and then looking at its moment generating function ok. So, we have expressed this quantity in terms of its moment generating function. Now, we will see that this moment generating function is easier to handle to bound ok.

Now, what we will do? We are going to handle each of this parts separately. So, let me call this as i, and let me call this as ii. Now, we are going to bound. So, the second step is to bound i ok. What is this quantity? Log of we always love to write the things in exponential form, because we get very tight bounds when we write them in exponential. If you remember like in you when we wrote weighted majority algorithm, we had a tight upper and lower bounds on e^{-x}right.

(Refer Slide Time: 14:14)



So, on e^{-x} you have this nice bounds. What was that? $1-x+\frac{x^2}{2}$.

We will use similar things here also ok. Let me say this, what is this quantity. So, this one I am going to write it as and this quantity. So, the first term into the second term, this is going to be plus eta t this term here eta t times I am only going to look into that expectation of k times P_t l_{tk}. I just expanded this.

Student: (Refer Time: 15:31).

So, this is the same argument as I used earlier, and this is like a constant already because expectation is taken. So, log of exponential will nullify, and it will just give you this much ok. Now, let us try to see how to apply these bounds here. This is the log quantity right. So, there is one more bound I am going to look $\log x \leq x-1$. It is true always for let say it is x positive right, this should always hold.

Now, let us apply that logic here. So, if I am going to apply, what is this is going to be exp, so I am going to treat this entire thing as log space is going to be $E_{i\sim P_t}\exp\left(-\eta_t\,\widehat{l_{ti}}\right)-1+\eta_t E_{k\sim P_t}\widehat{l_{ti}}$ Now, I am going to pull this expectation outside from both the terms $E_{i\sim P_t}(\exp\left(-\eta_t\,\widehat{l_{ti}}\right)-1+\eta_t\widehat{l_{ti}})$. So, this quantity will be upper bounded by $E_{i\sim P_t}(\exp(\frac{\eta_t^2\widehat{l_{ti}^2}}{2})$ ok. So, it is fine.

Now, I am going to do anyway $\frac{\eta_t^2}{2}$ is constant, I will pull it out. And now I am going to look at expectation of this. It is a tilde; it is a hat right. So, what is this expectation? Let us compute this expectation. This expectation is nothing $\sum_j \widehat{l_{tj}^2}\, P_{tj}$, this is the definition of this expectation. And now if I am going to replace this quantity by estimator by this definition, this is $\sum_j \left(\frac{l_{tj}\, 1_{\{I_t=j\}}}{P_{tj}}\right)^2 P_{tj}$,.

During some manipulation maybe we can just keep around go fast now. So, this quantity is going to give me what, only this is going to remain only for that I_t everywhere else it is going to cancel. And it is going to give me $\frac{l_{tI_t}^2}{P_{tI_t}}$. Is this fine? Ok. I do not know if I mention that we will assume that this losses are always in the interval between 0, 1 ok. So, if we are going to make that assumption, this guy is going to be 2 P_{tI_t}because this $l_{tI_t}^2$ is upper bounded by 1.

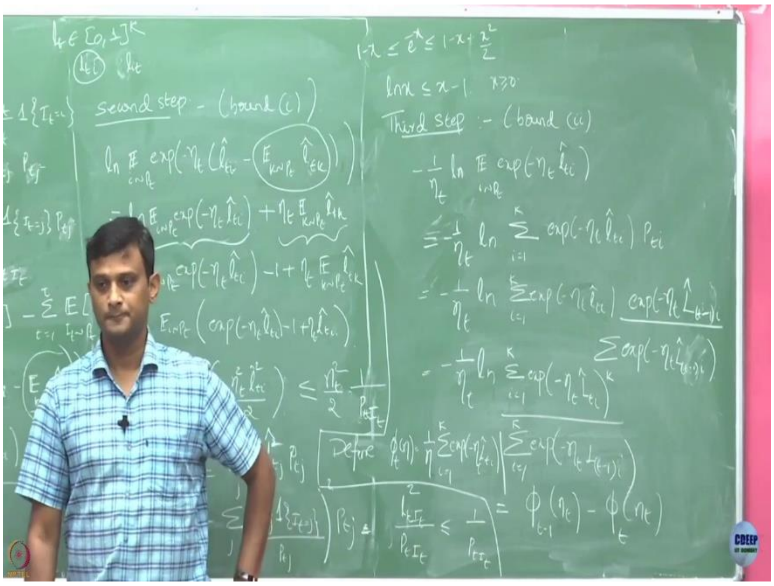
So, fine, what we will actually end up with is, if I am going to substitute this quantity here this guy is $\frac{\eta_t^2}{2\, P_{tI_t}}$ ok. So, this is the second step. What is my third step? Third step is to deal with this guy here. So, let us take $\frac{1}{\eta_t}$, well. So, by the way this the way this proof goes everything looks like what is happening, how the steps are all coming one after another right this manipulation. At least in the adversarial case in this look like steps I mean somebody came up with this, but like these are like standard steps, I mean this way manipulation will some or the other way of manipulation things in this passion will end up giving you the bounds that you are looking for.

See like we are ended up with our regret bound which are like of order square root n right, that means, it is this algorithm is making things learnable. Why? Because if I let n go to infinity I am learning that means, I am doing as good as my benchmark per round if I am allowed large number of rounds. But say what you are learning, you are trying to learn something about which you are clueless. These losses are generated in an arbitrary fashion; you do not know anything about that right. You are not making any assumption. The only assumption is that i made a like they are in the interval 0, 1, but that can also be relaxed by normalizing.

So, you are dealing with a very general scenario coming up which is a what is the right intuition to prove due this process hard, but whatever the way based on this exponential weights based algorithm has develop, and it looks like they have some standard way to go about proving these terms. Even though I am doing a lot of manipulation here, but by enlarge you see that this are kind of similar to what we earlier did it for the weighted majority algorithm. So, it is good like we know these steps.

And if we can under I mean at least if you are conversant with how this proof has gone through maybe in some other setup that we want to prove we can play with this steps and able to come up with a bound ok. And that regard it is important that we know the proof steps for this. Remember it is all math, but apriority is not clear while it has to go in this fashion ok. But do follow with all the steps I am trying writing here, so that later if you have not to prove for something else maybe you should understand this already maybe you can see where to tweak the proof to get bounds for your algorithm ok.

(Refer Slide Time: 23:49)



Now, this one I am going to write it as, so what is this $-\frac{1}{\eta_t}\log\sum_{\{i=1\}}^{k}\exp\big(-\eta_t\ \widehat{l_{ti}}\ \big)P_{ti}$. But now I am going to replace I know what is P_{ti}, now I am going to bring in how this P_{ti} used in a algorithm. So far all the steps I have written they are generic, there is nothing special about EXP algorithm. So, now, if I am going to use this P_{ti} here. So, what is P_{ti}? It has been shown to be I mean it has defined as $\frac{\exp(-\eta_t\,\widehat{L_{ti}})}{\sum_k \exp(-\eta_t\,\widehat{L_{tk}})}$

So, now if I am now I am going to take this numerator, so this quantity, what is $\widehat{L_{ti}}$? $\widehat{L_{ti}}$was defined together cumulative sum till round t right. Now, if I am going to add this as well, ok, sorry, so this at P_t this is defined to be l_t minus till the previous round not till that round. So, if

I am going to add this l_{ti} hat to this quantity, it becomes $\widehat{L_t}$. Is this correct? I have just manipulated this.

Now, I am going to just define this quantity to be $\phi_{t-1}(\eta_t) - \phi_t(\eta_t)$ ok, where I am going to say define. So, if I am going to define $\phi_t(\eta\ \ )$ to be this quantity, then this ratio here I could write it as $\phi_{t-1}(\eta_t) - \phi_t(\eta_t)$. So, is this clear? Ok, fine. So, then maybe for the remaining steps we will do it in the, because this completes my third step; there are just two more steps we will just do it in the next class, ok, fine.