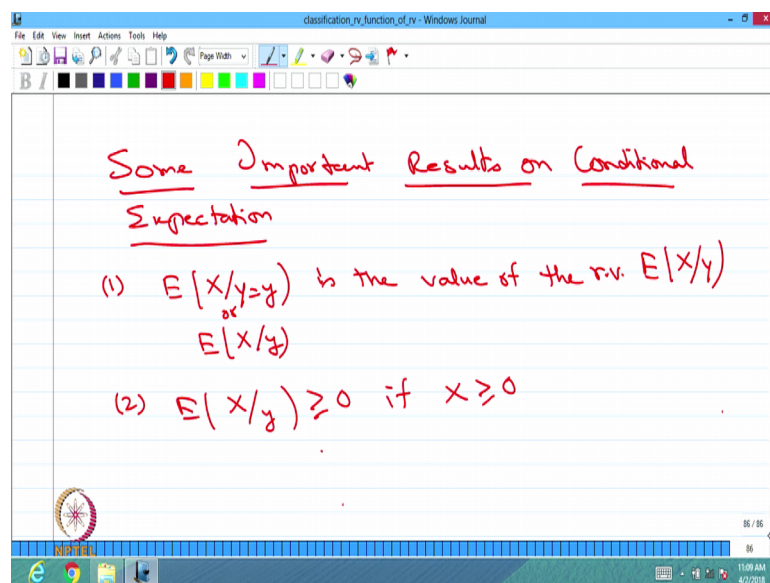


**Introduction to Probability Theory and Stochastic Processes**  
**Prof. S. Dharmaraja**  
**Department of Mathematics**  
**Indian Institute of Technology, Delhi**

**Lecture – 41**

Now, we are going to give few important results on the condition expectation.

(Refer Slide Time: 00:10)



Some important results on conditional expectation, first important result expectation of  $X$  given the random variable takes a value small  $y$  is the value of the random variable expectation of  $X$  given capital  $Y$ . So, sometimes we write this in the form of expectation of  $X$  given small  $y$ , either we write like this or we write expectation of  $X$  given small  $y$ , that is the value of the random variable expectation  $X$  given  $y$ .

The second result in the expectation of  $X$  given small  $y$  is always going to be greater than or equal to 0, whenever the random variable is going to be greater than or equal to 0 that means, whenever the random variable takes a values greater or equal to 0, or the probability of  $X$  greater than or equal to 0 is 1 in that case the expectation of the conditional expectation of  $X$  given other random variable is also going to be greater or equal to 0.

(Refer Slide Time: 02:19)

(3)  $E(1/y) = 1$

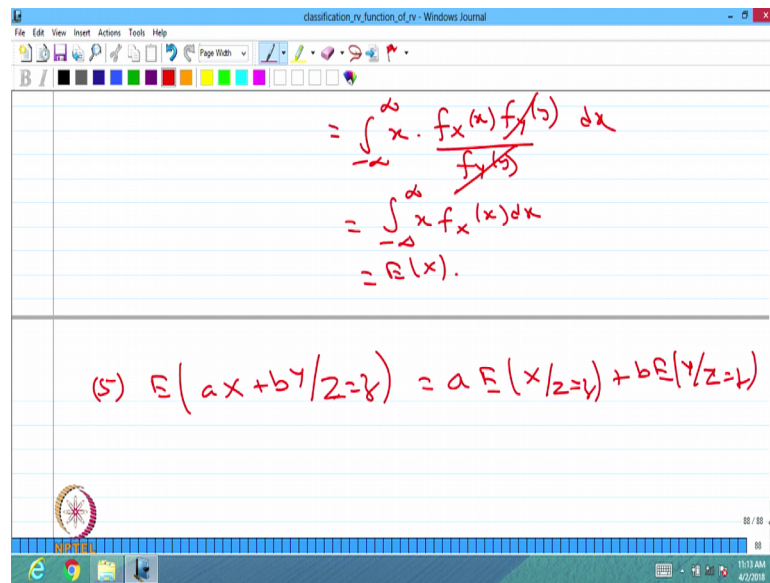
(4) If  $X$  and  $Y$  are independent rvs  
 $E(X/y) = E(X)$ .

i.e.  $E(X/y) = \int_{-\infty}^{\infty} x \frac{f_{X,Y}(x,y)}{f_Y(y)} dx$   
 $= \int_{-\infty}^{\infty} x \cdot \frac{f_X(x) f_Y(y)}{f_Y(y)} dx$

The third result, if you compute some constant given the other random variable, that is always going to be 1, next one if 2 random variables are independent, if 2 random variables are independent, then the conditional expectation same as the expectation of that random variable.

Because these two random variables are independent, you can prove it, we can prove it by considering both the random variables are discrete, both the random variables are continuous you can easily prove it; that means, that is suppose I consider both the random variables are of the continuous type; that means, expectation of  $X$  given  $y$  that is nothing, but minus infinity to infinity  $x$  times, conditional is same as joint divided by marginal. Again since these two random variables are independent, I can rewrite the product of marginal and it cancels out.

(Refer Slide Time: 04:08)



The screenshot shows a Windows Journal window titled "classification\_of\_function\_of\_random\_variable". The window contains two lines of handwritten mathematical derivations in red ink on a lined background. The first line shows the derivation of the expectation of a function of a random variable,  $E(g(x))$ , starting with the integral  $\int_{-\infty}^{\infty} x \cdot \frac{f_x(x) f_y(y)}{f_{xy}(x,y)} dx$ , which simplifies to  $\int_{-\infty}^{\infty} x f_x(x) dx$ , and finally to  $E(x)$ . The second line shows the linearity of expectation:  $(5) E(ax + by | Z=z) = a E(x | Z=z) + b E(y | Z=z)$ . The window also shows a standard Windows taskbar at the bottom with the Start button, taskbar icons, and system tray.

So, you will get minus infinity to infinity  $x$  times the probability density function of  $x$ , that you know that is same as expectation of  $X$ .

Here I have considered both the random variables are continuous type random variables, even if you consider both the random variables are of the discrete type again, you can prove it in the same way. So, if two random variables are independent then the condition expectation is same as the original expectation.

Next result, suppose you go for finding the conditional expectation of some constant times  $aX$  plus  $bY$ , with the other random variable  $Z$  takes a value small; that means, now I am considering three random variables  $X$   $Y$   $Z$ . So, expectation of some constant times one random variable plus constant times, another random variable given the third random variable take some value, that is same as constant is out, the expectation of  $X$  given the random variable takes a value  $Z$ , plus constant is out the condition expectation of  $Y$  given  $Z$  takes a value  $Z$ . So, this is true for all  $a$  and  $b$  both are real this also can be proved.

(Refer Slide Time: 05:53)

classfication\_xy\_function\_of\_xy - Windows Journal

File Edit View Insert Actions Tools Help

B I [color palette]

$= E(x)$ .

(5)  $E(ax + by | Z=z) = aE(x|Z=z) + bE(y|Z=z)$

(6)  $E(E(x|Y)) =$

LHS  $\int_{-\infty}^{\infty} E(x|Y=y) f_Y(y) dy$ .

88 / 88

11:15 AM 4/2/2018

The sixth property I said expectation of  $X$  given small  $y$  is a value of the random variable. Therefore, you can go for finding expectation of expectation of  $X$  given capital  $Y$ , expectation of  $X$  given capital  $Y$ , that is a random variable, we are finding the expectation for the condition expectation; that means, you can prove it you can prove it for considering both the random variables are discrete and continuous.

So, let us prove it then we will write down what is the answer. So, the left hand side is going to be minus infinity to infinity, by considering both the random variables are of the continuous type, that is expectation of  $X$  given this value multiplied by the probability density function of  $y$ .

(Refer Slide Time: 07:00)

The image shows a Windows Journal window with the title 'classification\_xy\_function\_of\_xy - Windows Journal'. The content consists of three lines of handwritten mathematical equations in red ink on a lined background:

$$= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} x f_{x/y}(x/y) dx \right) f_y(y) dy$$


---


$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot \frac{f_{x,y}(x,y)}{f_y(y)} f_y(y) dx dy$$

$$= \int_{-\infty}^{\infty} x f_x(x) dx \quad \left( \because \int_{-\infty}^{\infty} f_{x,y}(x,y) dy = f_x(x) \right)$$

By using the definition of condition expectation I can expand. So, the outer 1 is minus infinity to infinity  $f_y dy$ . Now, I am just expanding what is a condition expectation of  $x$  given  $y$  that is nothing, but minus infinity to infinity  $X$  times the conditional probability density function of so, this is inner integration and outer integration is minus infinity to infinity  $f$  of  $y$   $dy$ .

That is same as Now, I can expand now I can expand the condition probability density function that is nothing, but minus infinity to infinity minus infinity to infinity  $x$  times the joint probability density function, divided by marginal probability density function and I can rearrange the probability density function  $dx dy$ .

So, I can cancel these two and, also I can use minus infinity to infinity  $x$  times. So, I can take it out minus infinity to infinity  $x$  integration of minus infinity to infinity  $f$  of  $x$  comma  $y$  comma  $y$   $dy$  is nothing, but the probability density function of  $x$  minus infinity to infinity integration, joint probability density function with respect to  $dy$ , that is nothing, but the probability density function of  $x$ . Then  $dx$ , that is since minus infinity to infinity probability joint probability density function of  $x$  and  $y$  with respect to  $y$ , that is going to be probability density function of  $x$ .

(Refer Slide Time: 09:26)

A screenshot of a Windows Journal window titled "classification\_rv\_function\_of\_rv - Windows Journal". The window contains handwritten mathematical derivations in red ink on a lined background. The derivations show the expectation of a random variable X as a double integral over the joint probability density function, which is then simplified by integrating out the y variable to show it is equal to the expectation of X.

$$\begin{aligned} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot \frac{f_{x,y}(x,y)}{f_y(y)} f_y(y) dx dy \\ &= \int_{-\infty}^{\infty} x f_x(x) dx \quad \left( \because \int_{-\infty}^{\infty} f_{x,y}(x,y) dy = f_x(x) \right) \\ &= E(X) \end{aligned}$$

So, this is nothing but expectation of X provided it exist.

(Refer Slide Time: 09:35)

A screenshot of a Windows Journal window titled "classification\_rv\_function\_of\_rv - Windows Journal". The window contains handwritten mathematical derivations in red ink on a lined background. It shows two properties of conditional expectation: (5) the linearity property and (6) the law of iterated expectations. The law of iterated expectations is proven by showing that the left-hand side, which involves a double integral over the joint density function, simplifies to the expectation of X.

$$\begin{aligned} (5) \quad E(ax + by | Z=z) &= aE(x|Z=z) + bE(y|Z=z) \\ (6) \quad E(E(x|Y)) &= E(x) \\ \text{LHS} &= \int_{-\infty}^{\infty} E(x|Y=y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} x f_{x|y}(x|y) dx \right) f_Y(y) dy \end{aligned}$$

That means the expectation of expectation of X given Y is nothing, but expectation of X this is a very important result the left hand side involves the random variable Y whereas, the right hand side is free from the random variable Y; that means, if you want to compute the expectation of X, we can always find the another random variable for that random variable, you find out the conditional expectation of X with respect to the sort of

dummy random variable. After finding the conditional expectation, you find out the expectation of the conditional expectation that is going to be the expectation of  $X$ .

That means you can choose any random variable  $y$ , as long as the conditional expectation you can able to compute, then find out the expectation of conditional expectation is going to be the original expectation.

So, this is a very important result, whenever it is very difficult to find out the expectation of one random variable, but you can always relate that random variable with another random variable, then compute the conditional expectation. If that process is easy comparing to finding the expectation of the random variable, then you can use this result and find out the expectation of the random variable  $X$ , that is same as expectation of conditional expectation  $X$  given  $Y$ . So, we are going to use this property to compute the expectation for randoms.

(Refer Slide Time: 11:30)

Random Sum

$$S_N = X_1 + \dots + X_N$$

$N$  is a discrete type RV

$X_i$  - iid rvs  
 $N \perp X_i$  are independent rvs

$$E(S_N) = E(E(S_N/N))$$

$$= E(N) E(X_1) \quad \text{Wald's Equation}$$

If you recall we have discussed the random some long back that is we have used  $S$  suffix  $N$  as the random sum, as a sum of  $N$  random variables, where  $N$  is discrete type random variable  $N$  is a discrete type in particular it is a positive integer valued random variable.

So, if you want to find out the expectation of  $S_N$  that is nothing, but expectation of expectation of  $S_N$  given  $N$  that means, first you fix the value of capital  $N$ , then find out the conditional expectation then, compute the expectation for that that is nothing, but

since we made in the random sum problem we made the assumption of  $X_i$ 's are independent, which will be  $x_i$ 's or i i d random variables  $x_i$  or i i d random variables as well as  $n$  and  $x_i$ 's are independent. Therefore, you can compute that is equal to expectation of  $E_N$  multiplied by expectation of  $X_1$  all are i i d random variable.

Therefore any one random variable expectation of  $X$  multiplied by the expectation of  $N$  that is going to be the expectation of the random sum. We have already discussed the random sum much earlier now, in that time we have discussed only the distribution of the random sum. Now, we are discussing the expectation of the random sum that is same as expectation of  $N$ , multiplied by the expectation of  $X_1$ . Since all the  $X_i$ 's are independent random variable. This is called a Wald's equation using the earlier property we can find out the expectation of the random sum is the product of expectation of  $N$  with expectation of  $X_1$ .

(Refer Slide Time: 14:06)

The screenshot shows a Windows Journal window with the following content:

(6)  $E(h(X)) = E(E(h(X)/Y))$

(7)  $E(X h(Y)/Y=y) = h(y) E(X/Y=y)$

The next property number 6, expectation of so, this can be extended for a function of random variable expectation of  $h$  of  $X$ , also can be computed in the same way expectation of conditional expectation of  $h$  of  $X$ , given the random variable  $Y$ . As long as  $h$  is a Borel measurable function therefore,  $h$  of  $X$  is a random variable. So, the expectation of  $h$  of  $X$  is same as expectation of condition expectation of  $h$  of  $X$  given  $Y$ ; that means, we are extending the previous result with the function of random variable



In the 7th result expectation of  $X$  times  $h$  of  $Y$  given  $Y$  that is same as  $h$  of  $y$  expectation of  $X$  given  $Y$ ,  $X$  multiplied by  $h$  of  $y$  given  $Y$  takes a value small  $y$ , that is same as  $h$  of  $y$  into conditional expectation of  $X$  given  $Y$  takes a value small  $y$ ; that means, when you made it given  $Y$  takes a value small  $y$  the  $h$  of  $y$  no more random variable,  $Y$  is a random variable  $h$  of  $y$  is a function of a random variable, where  $h$  is a Borel measurable function therefore,  $h$  of  $y$  is a random variable.

So, when  $Y$  takes a value small  $y$ ,  $h$  of capital  $Y$  is no random it is going to be treated as a constant. So, constant is out while computing the expectation. So,  $h$  of  $y$  times the conditional expectation, there are some more properties with the conditional expectation that is with respect to the sigma field and so, on. So, for this course we stop it with the this many results.

(Refer Slide Time: 16:32)

Conditional Variance

$$\text{Var}(X/Y=y) = E(X^2/Y=y) - (E(X/Y=y))^2$$

$$\text{Var}(X) = \text{Var}(E(X/Y=y)) + E(\text{Var}(X/Y=y))$$

provided  $E(X^2) < \infty$ .

Now, we will go for the next topic is called conditional variance, the way we discussed the conditional distribution first then conditional expectation. So, now, we are going to define the conditional variance that is nothing, but variance of we are going to define, conditional variance as the variance of  $X$  given the other random variable takes a value small  $y$ , that is same as it. Similar to the variance formula, but now it is a conditional variance that is same as expectation of  $X$  square given  $Y$  takes a value small  $y$ , minus the conditional expectation of  $X$  given small  $y$  the whole square.

When you compute the variance of  $X$  and that is expectation of  $X$  square minus expectation of  $X$  whole square the same way, one can define the conditional variance as a variance of  $X$  given  $Y$  takes a value  $y$  that is nothing, but the expectation of  $X$  square given  $Y$  takes a value  $y$  minus conditional expectation of  $X$  given  $Y$ .

We have already given the result for conditional expectation of  $h$  of  $x$ . So, now, we are going to find out the variance of the random variable  $X$ , using the conditional expectation and the conditional variance, that is the variance of conditional expectation plus expectation of the conditional variance provided the second order moment that is finite.

So, the way we are finding the expectation of  $X$  using the conditional expectation, one can find the variance of  $X$  also, in terms of the conditional expectation and the conditional variance; that means, for any random variable variance of  $X$  can be computed by finding the conditional expectation and the conditional variance first, then apply the variance and the expectation with the vice versa, then summation becomes the variance of  $X$ .

(Refer Slide Time: 19:18)

Handwritten mathematical derivations on a digital notepad:

$$S_N = X_1 + X_2 + \dots + X_N$$

$$E(S_N) = E(N) E(X_1)$$

$$\text{Var}(S_N) = \text{Var}(E(S_N)/N=n) + E(\text{Var}(S_N)/N=n)$$

$$\text{Var}(E(S_N)/N=n) = \text{Var}(N E(X_1))$$

We are going to use this result to compute the variance of a random sum that is a random sum is defined, sum of  $i$  i d random variable and,  $N$  is a discrete type positive integer random variable, which is independent of  $x$  i's. So, we have already got it the expectation of  $S_N$  that is expectation of  $N$  multiplied by expectation of  $X_1$ . Now, we are going to

find out variance of the random sum using the previous result that is going to be expectation of variance of the conditional expectation of  $E(S_N | N)$  given  $N$  takes a value  $n$  plus, the expectation of the conditional variance of  $E(S_N | N)$  given  $N$  takes value  $n$ .

So, one can find the variance of random sum using the conditional expectation and, the conditional variance, computing the expectation for the conditional variance and, variance for the conditional expectation that sum quantity is going to be the variance of random sum.

We have already the results for we have already the results for variance of expectation of  $S_N$  given  $N$  takes a value  $n$  that is nothing, but the variance of the expectation of  $S_N$  given  $n$  takes a value,  $N$  that is nothing, but expectation of finite  $N$  random variables expectation that is nothing, but  $N$  times the expectation of  $X_1$ . Since they are identical, we can make it  $N$  times expectation of  $N$ .

(Refer Slide Time: 21:35)

The image shows a screenshot of a software window titled "classification\_rv\_function\_of\_rv - Windows Journal". The window contains handwritten mathematical derivations in red ink on a lined background. The derivations are as follows:

$$\begin{aligned} \text{Now} \quad & \text{Var}(S_N | N=n) = n \text{Var}(X_1) \\ & E(\text{Var}(S_N | N=n)) = E(N \text{Var}(X_1)) \\ & = \text{Var}(X_1) E(N). \\ \text{Var}(S_N) &= [E(X_1)]^2 \text{Var}(N) + \text{Var}(X_1) \cdot E(N) \end{aligned}$$

That is same as that is same as since we are going for variance of  $N$  expectation of  $X_1$  expectation of  $X_1$ , you have to treat it as a constant. So, when you are computing the variance of  $N$ ,  $N$  is a random variable here and this is a constant therefore, expectation of  $X_1$  that whole square multiplied by variance of  $N$  ok.

So, this is a one quantity for variance of conditional expectation. So, we got variance of expectation now, you have to go for conditional variance, then you have to go for

expectation. So, first we have to find out variance of the conditional  $S_N$  that is going to be when  $N$  is fixed. So, you are trying to find out the variance of finite  $N$   $X_i$ 's; that means, all are independent random variable therefore, it is going to be  $n$  times variance of  $X_1$ .

Therefore, the expectation of conditional variance so, that is same as expectation of  $N$  variance of  $X_1$ , variance of  $X_1$  is a constant. So, constant is out. So, that is variance of  $X_1$  expectation of  $N$ .

Therefore adding both the quantities adding both the quantities variance of  $S_N$  is going to be the earlier answer is it is expectation of  $X_1$  square, variance of  $N$  into variance of  $N$  plus the second quantity is variance of  $X_1$  times expectation of  $N$ . So, this is going to be the variance of the random sum. So, once you know the distribution of  $X_i$ 's and distribution of  $N$ , one can find the distribution of random sum. And if you know the mean and variance of  $X_i | Z_N$ , you can find out the mean and variance of the random sum also.