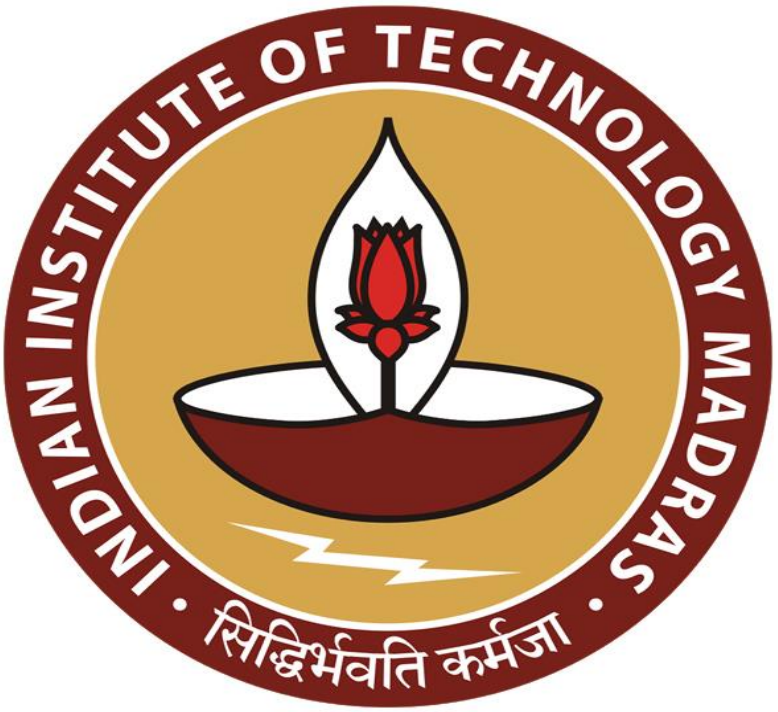
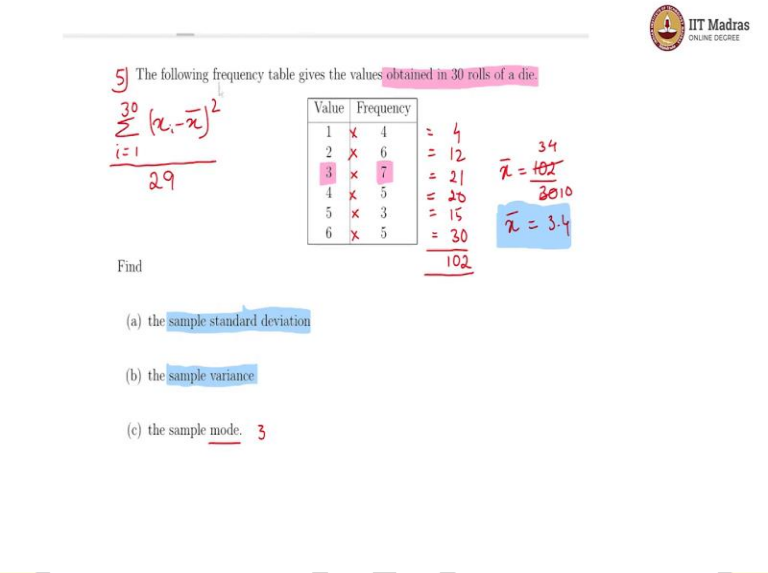


IIT Madras

ONLINE DEGREE

Statistics for Data Science - 1
Professor Usha Mohan
Prathyush P (Support team)
Department of Management Studies
Indian Institute of Technology, Madras
Week - 03
Tutorial - 05

(Refer Slide Time: 0:14)



In our fifth question, there is this frequency table given to us and these are values obtained in 30 rolls of a die. So, the value 1 was achieved 4 times, value 2 was achieved 6 times and so on. Now, they want us to find the sample standard deviation, sample variance and sample mode. We should be doing this in the opposite order because the mode is the simplest, so I would like to finish that first, clearly 7 is the maximum frequency and 3 is the value with the maximum frequency.

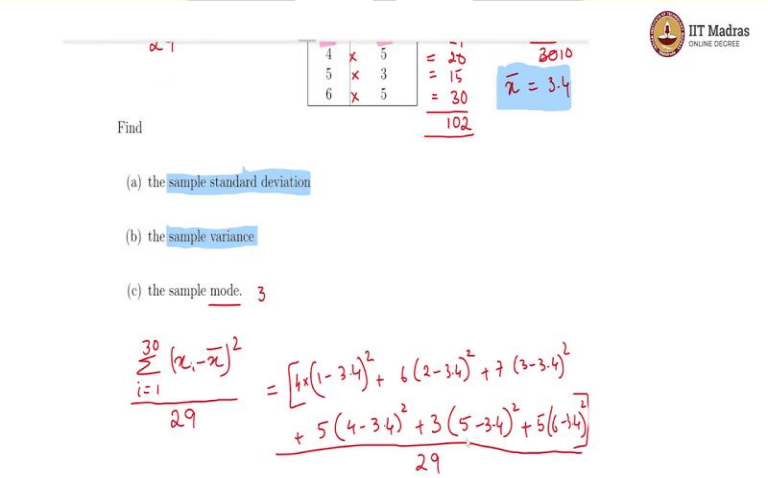
So, our mode is 3, simple, and then sample standard deviation is simply the square root of the sample variance, so let us calculate the sample variance first and for calculating the sample variance we first have to calculate the mean. So, for calculating the mean, what we need to observe here is the mean is for the values that are being

obtained on the die, which means one has come four times, so that contribution is 1 × 4 which is 4 but 2 has come 6 times.

So, the contribution from 2 would be 12 and this is going to give us 21. 3 has come 7 times. 4 is going to give us 20. 5 is giving us 15 and lastly 6 is giving us 30. So, this is an individual contributions to the total sum and the total sum would be this, 16 + 21 is 37, 37 + 20 is 57, 57 + 15 is 72, 72 + 30 is 102. So, the mean is, let us call mean $\bar{x}$ is 102 ÷ 30 which gives us this is 3 tens and this is 3 × 3 and 4. So, 3.4 is the mean, so this is our mean and for calculating variance what we do is we now do $\sum_{i=1}^{30}(x_i - \bar{x})$.

So, for each term we are going to look at the observation, we are going to look at its difference from the deviation from the mean and we are going to square it and divide by 29 that is 30 - 1, when we are calculating the sample variance and sample standard deviation we use n - 1 in the denominator not n.

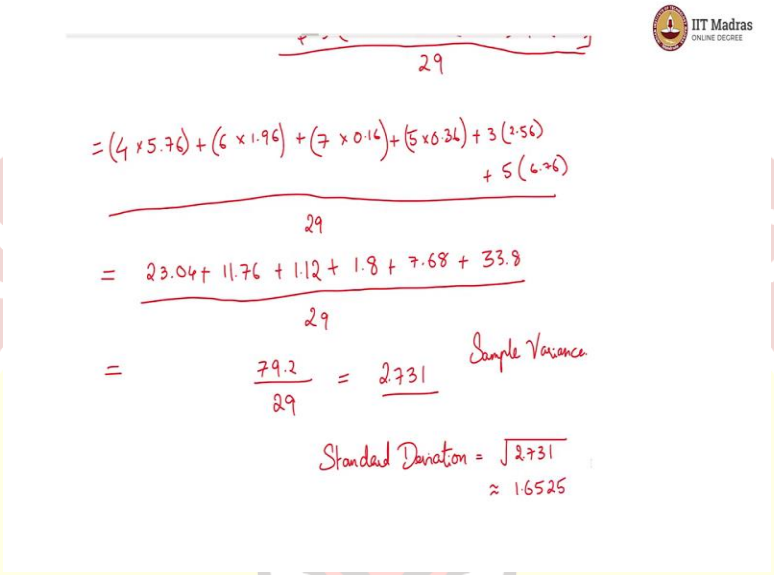
(Refer Slide Time: 3:24)



So this in order to calculate this quantity, we are essentially going to do $(1 - 3.4)^2$, how many times is this going to come? 4 times. So, $4(1 - 3.4)^2$, plus we have 2 coming up 6 times, so $6(2 - 3.4)^2 + 7(3 - 3.4)^2 + 5(4 - 3.4)^2 + 3(5 - 3.4) + 5(6 - 3.4)^2$. And this whole thing this entire thing is to be divided by 29.

$$\frac{4(1-3.4)^2 + 6(2-3.4)^2 + 7(3-3.4)^2 + 5(4-3.4)^2 + 3(5-3.4) + 5(6-3.4)^2}{29}$$

(Refer Slide Time: 4:35)



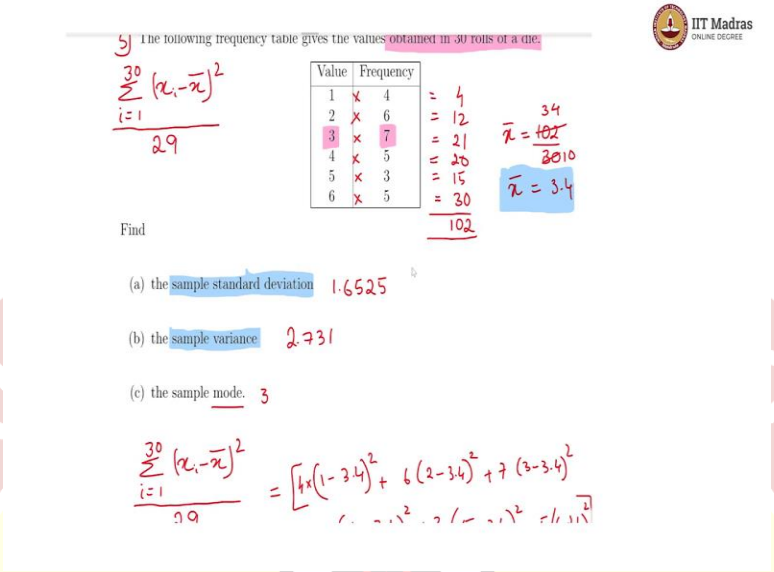
So, let us calculate that, we get $4[(2.4)^2 = 5.76] + 6[(1.4)^2 = 1.96] + 7[(0.4)^2 = 0.16]$. I will put them in brackets in order to avoid confusion, plus $5[(0.6)^2 = 0.36] + 3[(1.6)^2 = 2.56] + 5[(2.6)^2 = 6.76]$.

Again $\frac{4[5.76] + 6[1.96] + 7[0.16] + 5[0.36] + 3[2.56] + 5[6.76]}{29}$ and simplifying further we get, $4 \times 5.76 = 23.04$ plus $6 \times 1.96 = 11.76$ plus $7 \times 0.16 = 1.12$ plus $5 \times 0.36 = 1.8$ plus $3 \times 2.56 = 7.68$ plus $5 \times 6.76 = 33.8$ divided by the whole thing, divided by 29 which gives us further 23.04 + 11.76 = 34.8, 34.8 + 1.12 = 35.92, 35.92 + 1.8 = 37.72, 37.72 + 7.68 = 45.4, 45.4 + 33.8 = 79.2.

$\frac{79.2}{29} = 2.731$

This is our sample variance and standard deviation is merely the square root of this, so standard deviation is $\sqrt{2.731}$ which is roughly, this is also roughly I only approximated to 3 decimals anyway this is roughly 1.6525.

(Refer Slide Time: 8:03)



So, our answers would be sample variance is 2.731 and sample standard deviation is 1.6525.