2. Statistical Analysis: an approach of modelling turbulence - I

Okay, so let us start the class. Welcome you all again to a new chapter so we will start with, as I said on the theory first, we look into a new chapter called statistical analysis, so the reason I am starting with the theory and this part is you need to understand the theory before you can start to model and we start with this new chapter let us call this theory. Chapter 1 Statistical Analysis. Okay, so the reason I'm taking up this chapter first is we discussed that turbulence can be characterized or defined it in terms of its dissipative, diffusive, you know, it's random and so on. So each of this term we understand in depth before we start to model.

So the first thing that we learned is turbulence is random, chaotic, irregular motion, right? So it's a turbulence is characterized as defined as a random process. Ok, so have you come across any random processes? any examples that you have, or is it only in this course you are going to first encounter random process in the context of fluid mechanics the clue is already here we are looking into statistics now we are also saying turbulence is a random process any turbulent random processes that you have encountered in your laboratory anything that you have measured where it exhibited randomness. So most of you are engineers. So engineers are mostly interested in statistical values, mean, right? Average values.

So for example, you are interested in what's tomorrow's temperature. You will not ask like give me it as a function of time, right? You look into, like, mostly an average value, let us say, in the afternoon, what should be the average temperature value or average, you know, moisture content and so on so engineers are mostly interested in so engineers are mostly interested in mean or average values. Example you have mean temperature right, let us say mean temperature or mean pressure and so on. ok right. So, we look at this random process.

Let us say, for example, you would have gone to, Most mechanical engineers would have had a lab called internal combustion engine lab or some combustion course, or you would have gone to a heat transfer course and so on. You would have measured the temperature inside an internal combustion engine right. So let us say an example problem temperature. Temperature measurement in an IC engine. This is a standard UG lab in many courses.

So, how many samples did you take here? The question is, how many samples? I call it n. How many number of samples did you take before you said this is the value, this is the statistical value, how many samples to take to obtain a mean or any other statistical value 7 to 8? Some are saying 7 to 8 samples. Anybody else? Any other? Only 7 to 8? Nobody took more than that or less than this? 8. Why? Why you took 7, 8 samples and said this is the value? This could be in any laboratory. It need not be in your combustion lab.

It could be in your vibration lab. You are like looking into some noise. Right some decibel levels or oscillation frequency, you would have taken some samples in time or independent samples, and you would have got a mean, so how many samples you took and why you took so many samples, so you are saying okay, there is a small variation when I increase the number of samples the mean is not varying so much but there is an error so there is an error and you are saying it is like decimal 0.1 or something compared to your value correct So, this could be because this is the amount of time you had, others have to do experiments, it could be that your values are not changing so much, you stopped. So, let us do a simple experiment.

This is a everybody can do this and then we will see and understand the true nature of what should be called a mean and which mean is important for us to study turbulence. So this you can do it by yourself. So I am going to give you some dice. So just take these dice and then you flip the dice each of you and then tell me let us take you take 10 samples. you said 7, 8 I am generous ok let us take 10 samples and then you tell me what is the mean that you got ok and we will see what happens from that we will learn just pass it on.

So, flipping a dice, is it a random process, or is it deterministic, including you? This is a random process ok. So, flipping a dice is a random process. This is a simple experiment compared to actually going to lab and we learn a lot from this very simple experiment. So, let us collect 10 samples that is n equal to 10 since you all said 7 to 8 let us take 10 samples and then please also obtain the mean of it set let us call this an experiment 1. So, experiment 1 done by one student.

What are the samples you got? Can you tell me? 10 different samples. 2, 6, 1, 5, 1, 2, 5, 3, 4, 6. And the mean that you got? 3.

5. Mean is? 3.5 from 10 samples. Any other anybody else who have calculated it now? 4. I need the samples also. 5. ok, I will take from her.

5, 6, 1, 3, 4, 6, 1, 6, 4, 1, and what is the mean? 3.7. Anybody else from this side? 5, 2, 2, 1, 3, 6, 4, 6, 4, 2, ok.

3.7, ok. So, have others got similar values? Some have got, ok. So, if I just continue this. So, others have got some other values. So, if I just continue this some have got values like 4.1 and so on 4, some have got 4 and so on.

So, what do you learn from this? Some said you know, you take 7, 8 samples and one said I stopped it because the variation was so small, small compared to whom, right. you right the others have are getting bigger values somebody got let us say this is the engine temperature, right somebody got 3.5, and for you, it was not varying much it was 3.51, maybe 3.49 you said okay this is fine, but the others got 4.

1 right so that means which is the correct one even if you repeat this experiment will you get the same value another 7, 8 samples. So, then, why did you stop at 7 or 8 samples? What

is the existing mean? So, if I take all these samples here. So, you are saying 20 samples. So, you believe if I do 20 samples, let us say n equal to 20 and if I repeat it will I get the same value here more or less when? What should be n then? Who gave that value? Engines are not designed using theory.

There is a range. Yes, there is a range. The one who made that engine would have given the operational range of this temperature is plus or minus this. How did they get? Doing an experiment. Why did it vary? Now, my question is fundamentally is this mean? The question is, is this a mean or a random process? Isn't the mean also a random process here? If I plot this, this is varying, right.

So, if I make a plot, ok. So, let us call these values are going right, right. You have taken 20 samples. So, it is like going up and down like this, correct. And from 20 samples n, from 20 samples if I make it, I am getting, let us say, this is 3.5, let us say, and somebody else would have got because their signal is different, right so they would have got 4.

1, and their signal is going so the mean itself is a random process here then what then how do I define turbulence is random To get to that real randomness, I need to define a mean which is not random anymore. So this is a random process. And what kind of mean is this? This is an arithmetic mean. So, the arithmetic mean is also a random process. The takeaway here from this very simple experiment is arithmetic mean is also a random process.

It is not the true mean. So, then we need to define what is called true mean or an ensemble mean. So, this is a theoretical mean that we are going to obtain or at least we are going to define we will see whether it can be obtained or not because we cannot work with this arithmetic mean. It is changing as you change the number of samples its value is changing since arithmetic mean itself is random I cannot get to the true randomness which is turbulence I want turbulence is defined as random so I want to capture randomness the true essence whatever it is if I go with arithmetic mean it will come the randomness will have error correct so we define what is called a true mean or an ensemble mean so before that the arithmetic mean itself if I define this let us say it is capital X_N, N is there because it is still a function of number of samples. So, I define this you all know what you have done you have

$$X_N = \frac{\sum_{n=1}^{N} x_n}{N}$$

Where x_n is a random event or a random realization. So, the X_N is a random event that occurred you flip the dice you got some value you have basically summed it up and then you averaged correct this is your arithmetic mean definition.

So, now, what we define is called a true mean or an ensemble mean ok. So, this is defined as Xn, or I would call it X. Because it is no longer a function of n. X is equal to in some places also if you have small x we use angle brackets this you see in literature or sometimes an over bar is also used to determine what is an ensemble mean. Here this will be similar to your the only thing is you are going to put a limit here, limit n tends to infinity.

$$X = < x > = \bar{x} = \lim_{n \to \infty} \frac{\sum_{n=1}^{N} x_n}{N}$$

So, this is your true mean, but how to obtain is a question here. It is not just a straightforward. So, you need two conditions to get this, two conditions to satisfy to get true mean the first one is you would need the samples that you take has to be statistically independent So now he was telling that I flip it. I more or less know if I flip it like this, I get it right. But somebody else flipped it differently.

So whether you flip the dice or somebody else flip the dice, these two events should not be statistically dependent. OK, so you need samples must be. statistically independent ok. So, they should not affect each other the way you have to collect the samples. So, now this is not at all possible to do right.

So, if you let us say today you have taken lots of samples maybe 1 million samples and then it is still varying and then you are tired you go home you come back tomorrow and repeat the experiment. So, you are continuing to take more number of samples. There will be sometimes possible or not sometimes it is these samples is not always easy to get statistically independent. One realization will have an effect on the other one. So let's say you often see in literature there is some experimental data coming from one book and then you will see experimental data coming from another book.

There is a small variation. One is done, let's say, in India, and the other one let's say it's done in Europe. So, the ambient conditions has changed. Maybe the experimental setup has changed. Even if you use the same experimental setup, the person doing it is different. So, there are so many things that is going to change here.

So, getting statistically independent samples is not possible. That is the one problem with getting a true mean or an ensemble mean. The second thing is you have to get infinite number of samples. these two things are only theoretically possible. So, therefore, ensemble mean is only theoretically possible.

So, in lab you cannot do this. But we will not go into the lab now because the moment we go into the lab we come back to arithmetic mean which is a random process and therefore I cannot get my true randomness or turbulence and without understanding turbulence I cannot model it. So let us believe that true mean exists, and I have access to it, and I can separate the turbulence from its mean. So let us believe and start the theory and then we will see later what can be done with respect to that one. Okay.