What is Functional Data Analysis and What does it Assume?

Frequentist statistics assumes that the data collected are random samples collected from an unobservable probability distribution with fixed parameters. It is interested with singular observations/vectors of observations and their associated probabilities of appearing in a sample.

Bayesian Statistics assumes that data are collected are random samples from an unobservable probability distribution where the parameters of the distribution are also unobservable and therefore open to randomness, they are treated as being sampled from probability distributions themselves such that information from samples drawn previously from similar distributions can be used to inform our estimate of what the given parameter is for the sample we have just drawn.

Functional Data Analysis is not concerned with singular observations or vectors of data but rather with smooth curves known as functions in this case. This makes intuitively makes sense for some forms of data e.g. the height of a person over a time continuously grows due to some underlying process or the temperature of a city over time may have some seasonal pattern (<https://towardsdatascience.com/introduction-to-functional-data-analysis-fda-c9d298e8fcf5>)

. The main thing about these functional data is that theoretically this data could be measured at any point in time (treated as a continuous variable) e.g. the temperature of a city could be measured at any point in time. Some data that would not fit this interpretation would be the value of a coin flip as this can not continuously be measured and takes discrete values or the mean income of people who replied to a census as while this data is continuous it can only be measured/only exists when a census takes place e.g. there is not some underlying function that has a continuous support from which we can sample . If we however considered this data to represent the mean income of the entire population then we could consider this to be functional data as this is a quantity that always exists regardless of when it is sampled and it can theoretically be sampled at any point in time.

Within the previously stated frameworks our observed data would be a vector of observations of values from these functions. Within a functional framework our data is a smooth process that describes how the temperature changes over time. This removes the fixed number (dimensionality) nature of the observations and making the number of observations effectively infinitely dimensional. Theoretically the data we have observed are arbitrary as we could have chose to observe any time point in this process and could continue to if the process is still running e.g temperature system with no noticable reason for change.

The observation of a complete function is clearly impossible due to the nature of a function requiring an infinite number of observations for every time point we would wish to measure a quantity at. Therefore we must estimate these functions. This estimation process is described in the next chapter. If we assume that all of our observations have no error associated in them e.g they are direct draws from the function that describes the variation with no error then estimating the function is done through interpolation. However if it is assumed that there is error in the observations then the process is known as smoothing.

Sometimes data effectively comes to us as functions/curves e.g when tracking the movement of a particle/ handwriting data

Also some advanced sensing equipment can estimate curves easily as they can take measurements at incredibly fine intervals

**Actual Writing**

To answer this question lets first think about what data a statistician is typically tasked with analysing. Blood pressure readings on the first 4 days after someone has taken a drug, the amount in kilogrammes of a certain pollutant that is in the air in different cities or the number of children that take the bus to school in different cities are good examples of the type of data a statistician may be asked to analyse. These would usually be stored in a data matrix. Taking the first example of blood pressure readings these may be given in a matrix like the one given in expression below

To understand the theoretical concept of functional data analysis it is first good to consider the types of data a statistician may typically work with in a typical/non-functional setting of data analysis. Some good examples of data a statistician may typically be asked to analyse