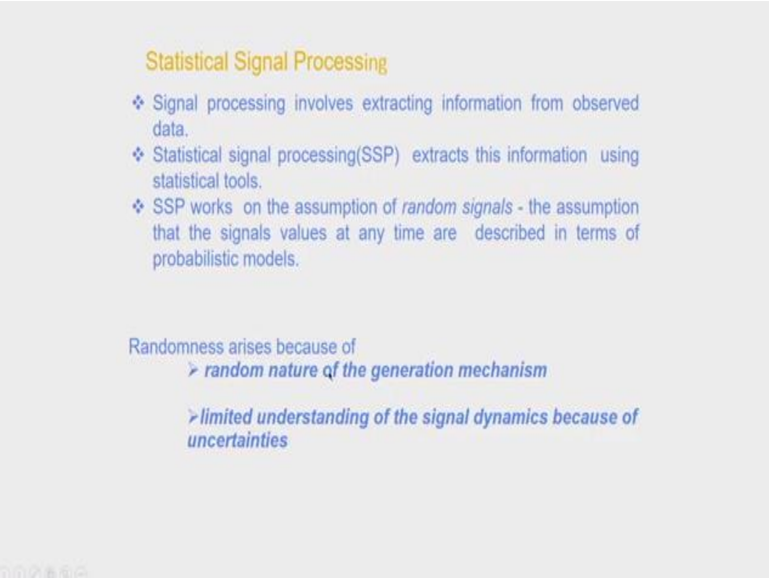
Statistical Signal Processing
Prof. Prabin Kumar Bora
Department of Electronics & Electrical Engineering
Indian Institute of Technology, Guwahati

Lecture No 1
Overview of statistical signal processing

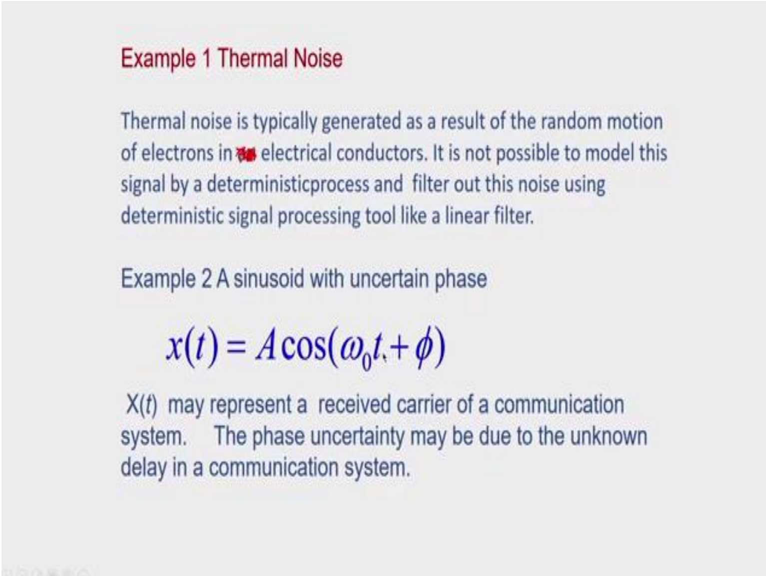
Hello students! In this lecture, I will give a broad outline on statistical signal processing.

(Refer Slide Time: 00:40)



As you know, signal processing involves extracting information from observed data. Statistical signal processing, SSP in brief extract this information is in statistical tools. SSP works on the assumption of random signals. The assumption that the signal values at any point of time are described in terms of probabilistic models, randomness arises because of random nature of the degeneration of the mechanism. Limited understanding of the signal dynamics, because of the uncertainties present in this signal.

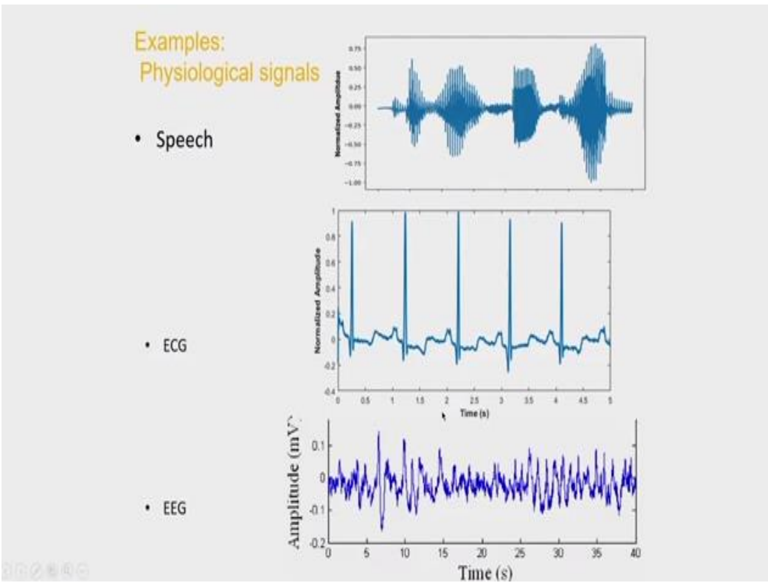
(Refer Slide Time: 01:23)



I will give some examples: example 1 thermal noise. Thermal noise is typically generated as a result of the random motion of electrons in electrical conductors. It is not possible to model this signal by a deterministic process and filter out the noise using deterministic signal processing tool like, a conventional linear filter. Therefore thermal noise is usually modelled as a random signal. Example 2, A sinusoid with uncertain phase.

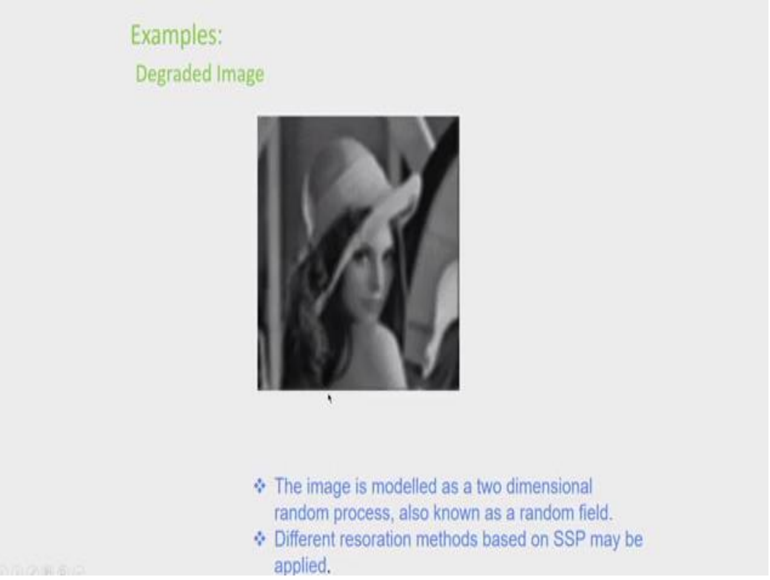
This is the sinusoid x t is equal to Acos omega not t + phi, phi is the uncertain phase. For example, x t may represent a received carrier of a communication system. The phase uncertainty may be due to the unknown delay in the communication system. So in this case, x t will be a random signal and the randomness is because of the unknown delay in this system.

(Refer Slide Time: 02:32)



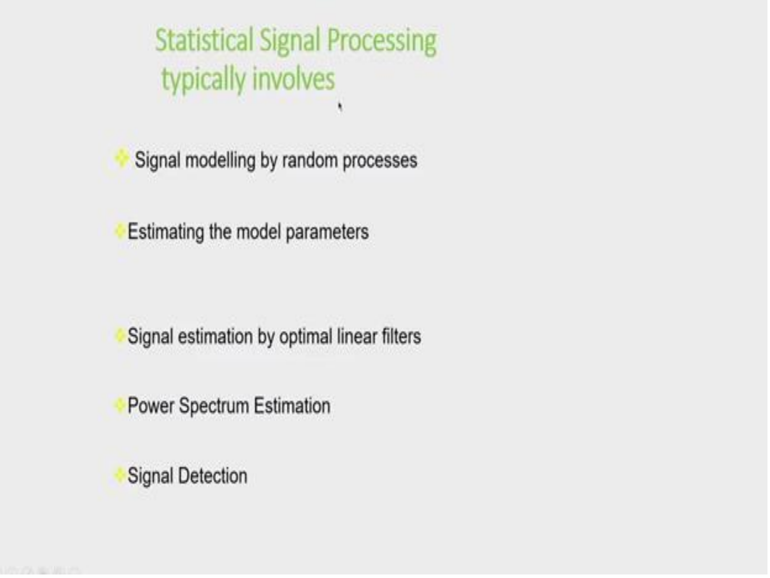
Most of the physiological signals are unknown in nature; they are model as random process. For examples: Speech Signal, This is a speech wave form, this is an ECG wave form, again we have given an EEG wave form. All these wave form can be characterized by using statistical tools.

(Refer Slide Time: 02:57)



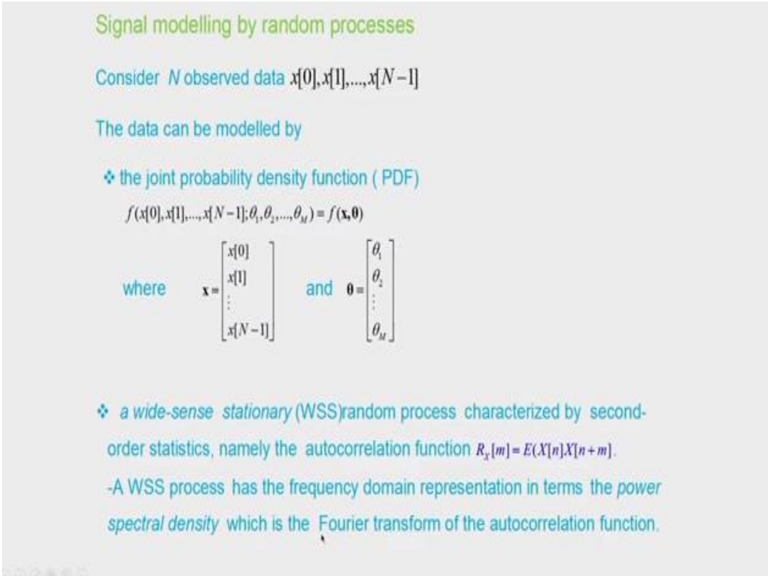
In this example, we saw a degraded image. Now, the image is a two dimensional signal. So, it is modelled as a two dimensional random process, also known as a random field and now in different restoration methods based on SSP may be applied to the blurred image, or get a clean image out of this degraded image.

(Refer Slide Time: 03:24)



Statistical signal processing, typically involves signal modelling by random processes, estimating the model parameters, signal estimation by optimal linear filters, power spectrum estimation and signal detection. These are the typical problems, considered in a normal statistical signal processing course, but there are other topics, which will not provide.

(Refer Slide Time: 03:54)

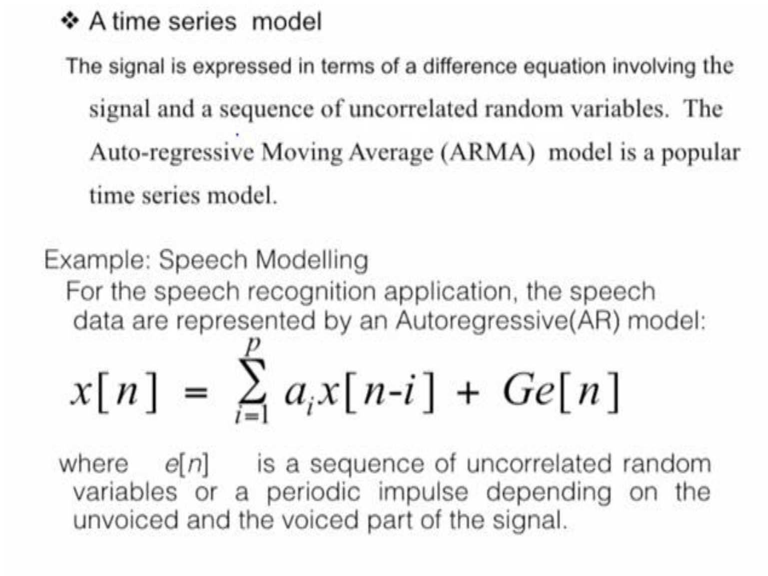


Signal modelling by a random process. Consider N observed data, x0, x1 up to x n-1 and these are the observed data. The data can be modelled by the joint probability density function PDF. This is the description; smaller f is the symbol for joint probability density function. It is a function of x0, x1 up to x n-1, and also parameters theta 1, theta 2 up to theta M, and in vector notation, we denoted it by f of x, theta.

Where x is the data vector comprising x0, x1 up to x n-1. And theta is the parameter vector, comprising of theta 1, theta 2 up to theta M. So that way we can model data by means of its joint probability density function. The data can be also modelled as a wide-sense stationary WSS random process. Such a process is characterized by second order statistics, namely, the autocorrelation function.

Defined as Rx of m is equal to E of x n in to X of n+m, where is the expectation operator. We will discuss about this autocorrelation function in a later class. A WSS process has the frequency domain representation in terms the power spectral density. This is a very important aspect, which is the Fourier transform of autocorrelation function. Therefore we can analyze the data in the frequency domain, in terms of the power spectral density.

(Refer Slide Time: 05:45)

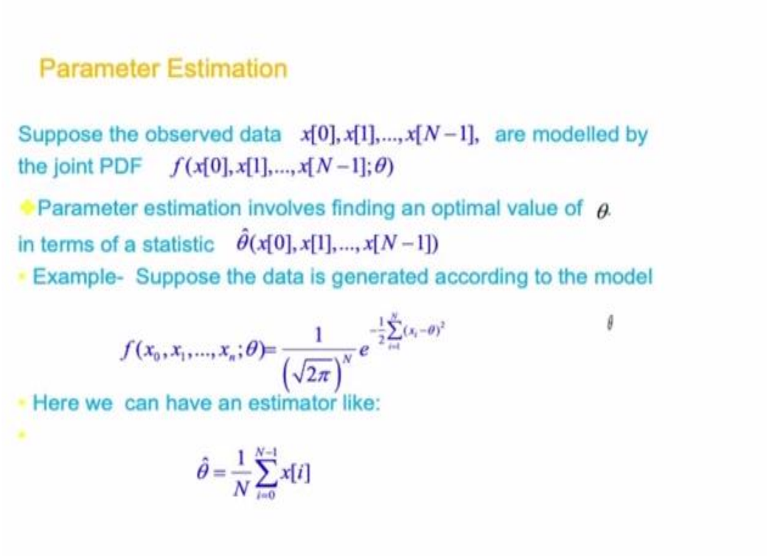


Third model is a time series model. This signal is expressed in terms of a difference equation; you will not a difference equation involving the signal and the sequence of uncorrelated random variables. There will be the signal and its difference and there will be the uncorrelated random variables and their differences. The Auto-regressive Moving Average model ARMA model that is a popular time series model.

Example: Speech modelling, for the speech recognition application, the speech data are represented by an autoregressive model. That is abbreviated as AR model. Here x n is a linear combination of previous values of x n, here x n is modelled as a linear combination of previous values of x n, that is auto-regression, plus some excitation term that is G into e n. where e n is a sequence of uncorrelated random variables or a periodic impulse.

It may be a be an auto uncorrelated random variable, or a impulse depending on the unvoiced or voice part of the signal because these signal as typically it has voice part and unvoiced part and in the unvoiced part this e n will be uncorrelated random variable, and in the case of voice signal en is a periodic impulse. Once we have a model, we have to estimate the parameters of the model.

(Refer Slide Time: 07:33)

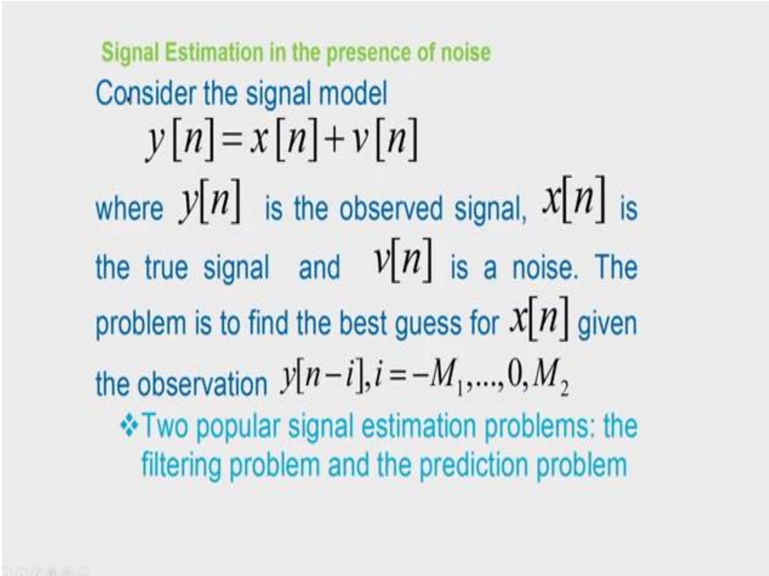


Suppose the observed data x0, x1 up to x n-1 are modelled by the joint PDF. That joint PDF is x0, x1 up to x n-1, in terms of parameter theta. Parameter estimation involves finding an optimal value of theta in terms of a statistic, theta at x0, x1 up to x n-1 any function of data. This theta head is a function of data, which does not involved the unknown parameter is known as a statistic. So parameter estimation involves finding the optimal value of data in terms of the statistics.

Example: Suppose the data is generated according to the model. This is the model, this is joint PDF. It is a Gaussian PDF, and it is given by this expression 1 by route over 2 pi to the power n into e to the power minus half of summation i going from 1 to N of xi- theta whole square. Theta is the unknown parameter. So here the model is given, we have to estimate the unknown parameter theta.

We can have an estimator like, theta head is equal to 1 by N summation i going from 1 to N-1 of x i. So, this estimator will see how to derive this estimator.

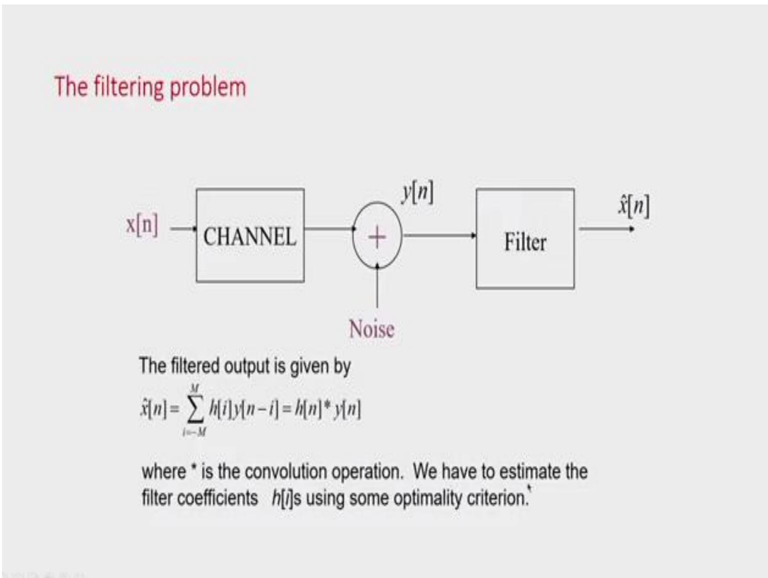
(Refer Slide Time: 09:08)



The third problem considered is the signal estimation in the presence of noise. Consider the signal model, y n is equal to x n + v n. This is the observed signal. y n is the observed signal, x n is the true signal and v n is a noise. Usually it is noise sequence will be uncorrelated. The problem is to find the best guess for x n given the observation y of n-1, i go in from –M1 to M2. Here M1, M2 maybe infinity.

So, how to estimate the true value of the signal from the observed noises signal. There are two popular signal estimation problems will be considering the filtering problem and the prediction problem.

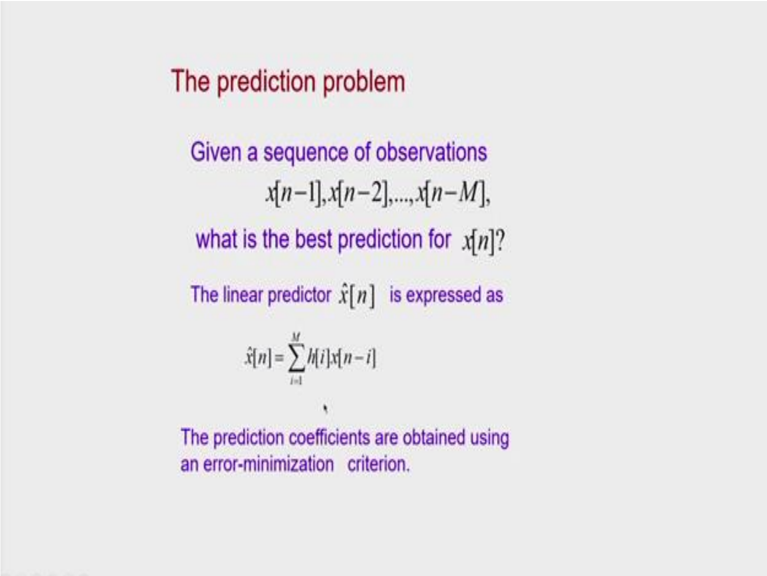
(Refer Slide Time: 10:05)



The filtering problem is like this, we have a signal x n. It is passing through some channel, and getting corrupted by noise. So that way y n will be the modified signal plus noise and we have to use a filter, so that we can get back except set n so the filtered output, this is the filtered output. The filtered output is given by this expression, x set n n is summation h i y n-i, i going from –M+M. And this is the convolution operation.

So we can write as h n star into y n, where star is the convolution operator. So basically, the filter performs the convolution operation. We have to estimate the filter coefficient, h i's using some optimality criterion. That is the problem, studied in inner filtering.

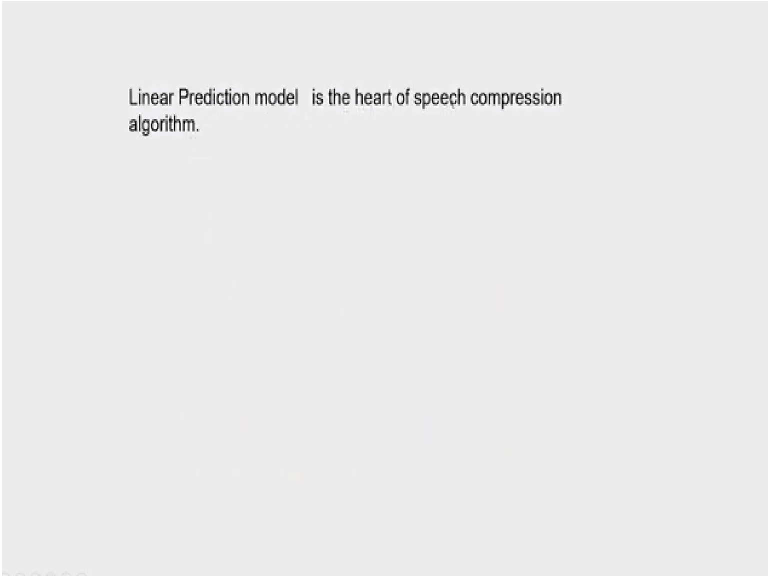
(Refer Slide Time: 11:13)



Next we will consider the prediction problem given a sequence of observations. That is the sequence of data, x n-1, x n-2, up to x n-M, what is the best prediction for x n.? Now reconsidering linear predictor for x set n, this can be expressed as a linear combination of the past data. So x set n is equal to summation h i x n- i going from 1 to M. This is again a convolution operation. The prediction coefficients are obtained using an error minimization criterion.

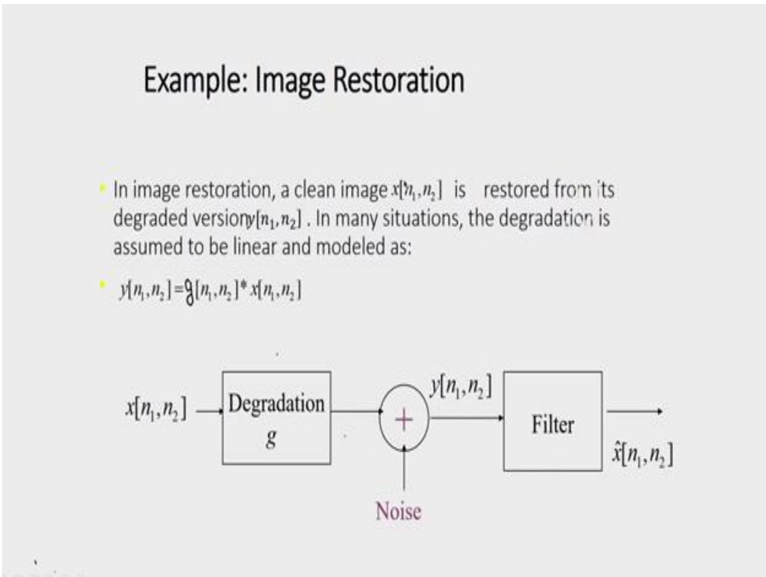
So, like in optimal filtering. This is also a special case of optimal filtering only, so we can apply some optimality criterion to obtain the filter predictor parameters.

(Refer Slide Time: 12:13)



Linear prediction model is the heart of speech compression algorithm. So any speech compression algorithm, for example, there is an algorithm called CLP, Code excited linear prediction. So, this model is based on linear prediction of speech. Another example we will consider image restoration

(Refer Slide Time: 12:41)

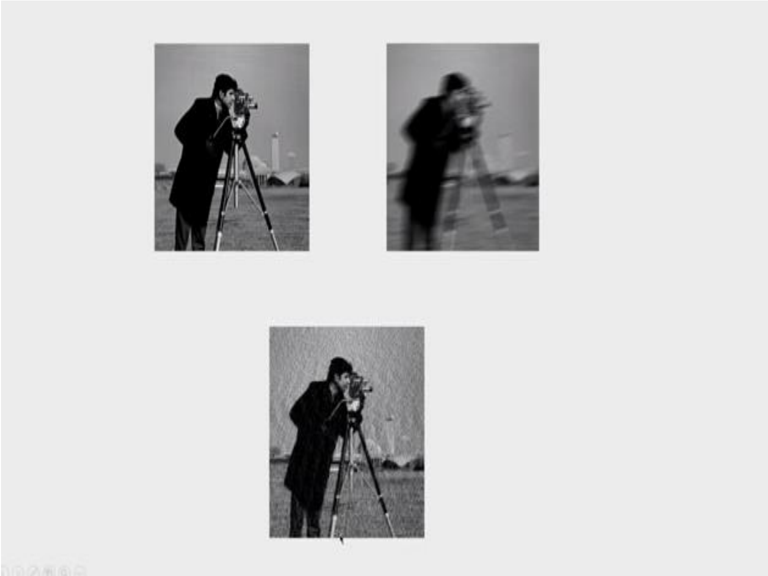


In image restoration, a clean image x n1, n2 is restored from its degraded version y n1, n2. In many situations, the degradation is assumed to be linear and modelled as: y of n1, n2 is equal to g of n1, n2 star x of n1, n2. That is convolution of g n1, n2 and x n1, n2. This is two dimensional convolutions and the convolution operates, we have already described earlier.

And now the model is like this: x n1, n2.

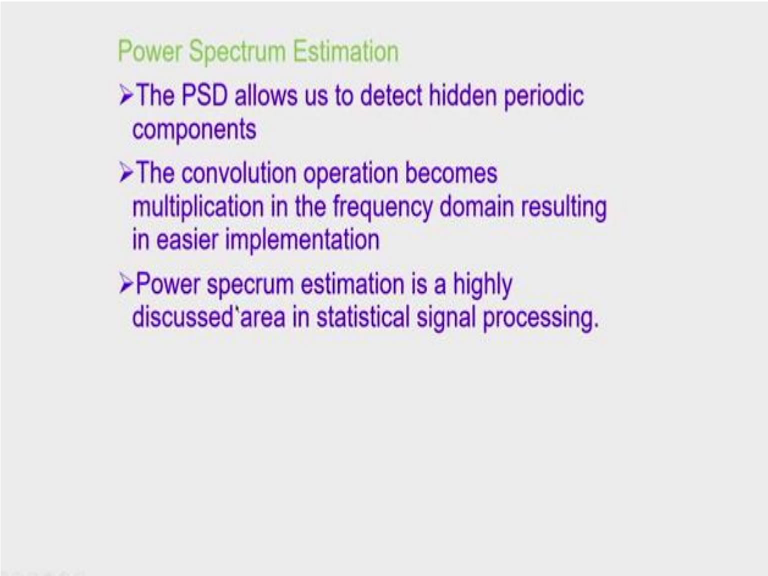
This is the input image. It is subjected to degradation g and some noise is added. Because of we get y n1, n2 and by means of filtering. That is optimal filter; we should get back the original image. So, this model is used in image restoration.

(Refer Slide Time: 13:43)



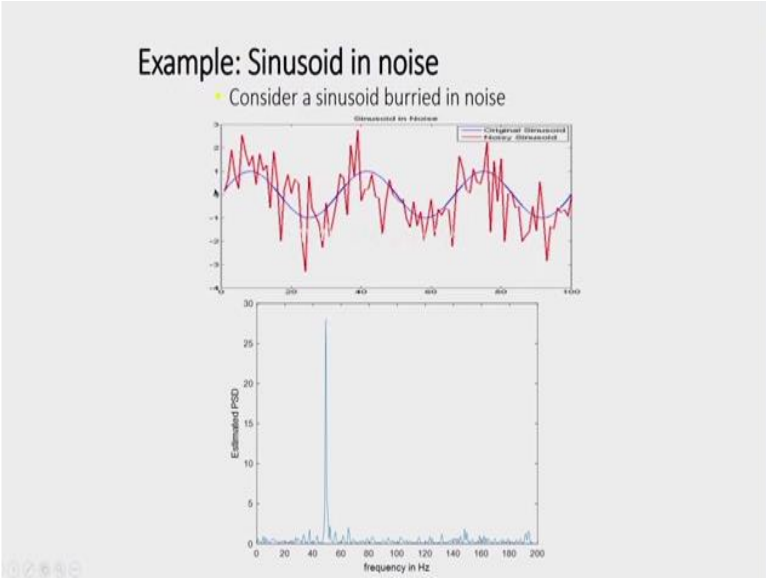
Example is this: This is a famous image known as the cameraman image. And this is a degraded version. And if we apply some signal estimation technique, like inner filtering, we get back this original image.

(Refer Slide Time: 14:01)



Fourth problem considered as the power spectrum estimation. The PSD power spectrum density allows us to detect hidden periodic component in a signal. The convolution operation becomes multiplication in the frequency domain resulting in easier implementation of optimal linear filters in the frequency domain. Power spectrum estimation is highly discussed area in statistical processing. However, in this course, we will not cover this topic.

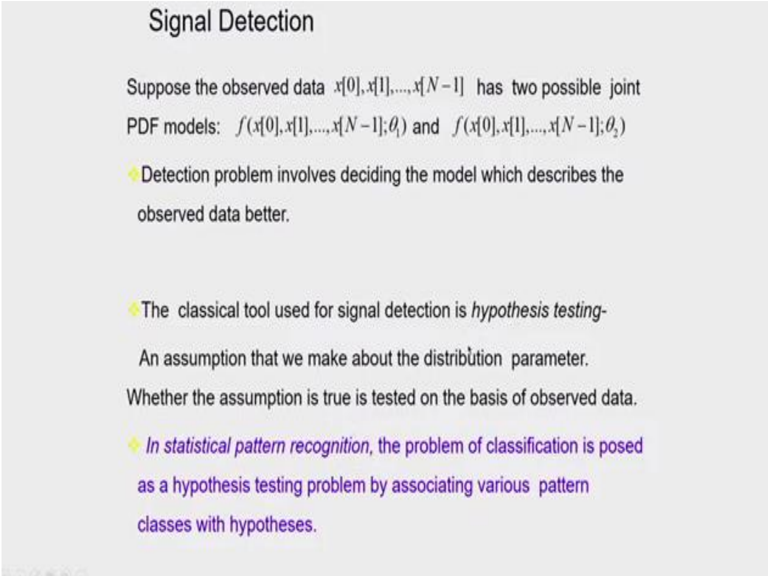
(Refer Slide Time: 14:38)



This is an example: Suppose we have sinusoid. Because of additive noise, the sinusoid is corrupted. Like this, this plot. So, now, there is a noisy signal. This is a noisy signal, and we have to do the power spectral analysis to estimate the power spectrum or the sinusoid in the signal. So, this is estimated PSD. There are techniques available. a lot of technique are there, how to estimate power spectral density.

Then we plot the estimated power spectral density against frequency, we see a peak. That is corresponding to the power line that is 50 hertz. So this is the peak, and that correspond to the 50 hertz sinusoidal.

(Refer Slide Time: 15:35)

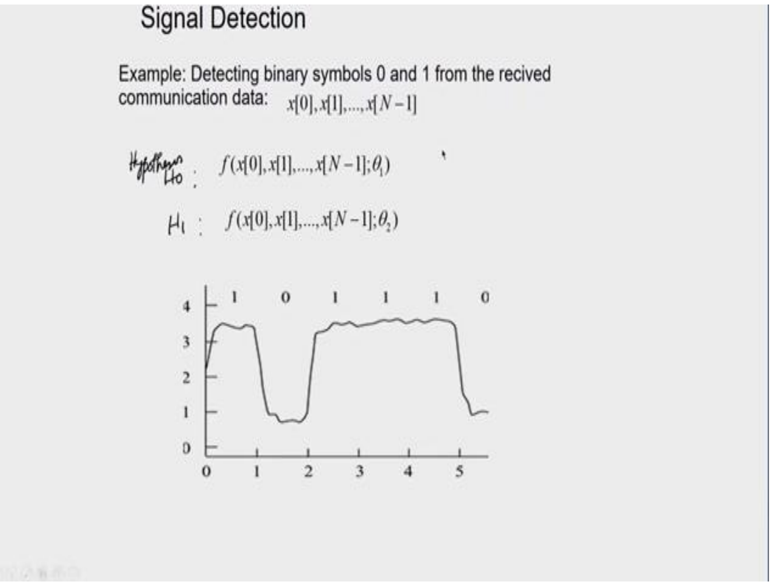


Signal detection, suppose the observed data x0, x1, x2 up to x N-1 has two possible joint PDF models: that is f x0, x1 same distribution with model theta 1, or with model theta 2. We have two distributions and that is what same as of PDF but parameter theta 1 and theta 2 are different. Detection problem involves deciding the model which describes the observed data better. So, suppose we have the observed data.

Now, it may be because of theta 1 or theta 2. But which ever, discuss data better, that will take a model. The classical tool use for signal detection is hypothesis testing. In hypothesis testing, we may make an assumption about the distribution parameter. Whether the assumption is true is tested on the basis of observed data. That is the hypothesis testing. Hypothesis testing has lot of applications in decision theory, particularly in statistical patterns recognition.

The problem of classification is posed as a hypothesis testing problem by associating various pattern classes with the hypothesis. So statistical pattern recognition is widely used in pattern classification such as, speech recognition, character recognition in biomedical signal disease identification, etc. and statistical pattern recognition basically used the tools of hypothesis testing.

(Refer Slide Time: 17:27)

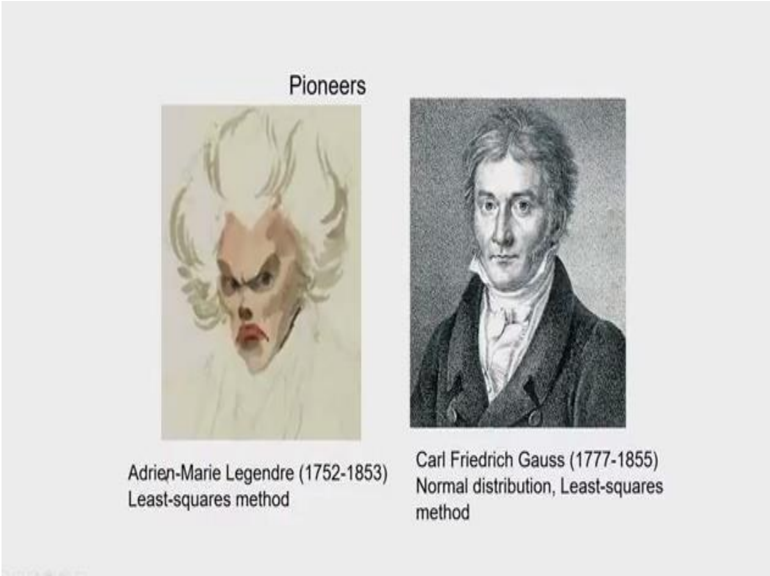


This is one example of signal detection. Detecting binary symbols 0 and 1 from the received communication data: x0, x1, up to x N-1 that is a frame of data. And we have two hypothesis that is hypothesis one, suppose this is hypothesis one. Hypothesis is generally we call it H 0, and H 1 on the basis of these two hypothesis. Now, we will interpret this data, and from analyzing this data, we will say, whether theta 1 is correct or theta 2 is correct. We will decide about that.

And that way, this is one plot with the distribution of suppose this is up to this instant it is 1, and then up to this instant it is 0, like that this we are identifying. We are detecting ones and zeros from this data sequence. So this is one application of signal detection. However, we will not discuss about signal detection. In this course, I will tell something about the Pioneer statistical signal processing.

We can trace back to the least squares method, as they origin of statistical signal processing. Least square method is

(Refer Slide Time: 19:01)



Due to Adrien- Marie Legendre. Here is a no portrait of him is available except this caricature. He published the part least-square with minimizes the square errors that is least-squares method. So, but Carl Friedrich Gauss who is a famous German mathematician, and he was the person who discovered normal distribution

Gauss knew about least squares method earlier than Legendre, but he published it later on with more details. There is a big controversy about who is the inventor of this least squares method. But, both are credited because any development in science or technology is because of different improvements, and both contributed to this least squares method, and that may be considered as the starting of this statistical signal processing. In middle of 20th century, there was lot of development of statistical signal processing.

(Refer Slide Time: 20:17)

And particularly two groups one is American group, headed by a Norbert Wiener and they worked on the optimal filtering problem. In Russia, parallel Andrey Kolmogorov, kinson, etc. Many other people are there, they are all done this problem of optimal filtering problem, and also Kolmogorov is credited with the asymmetric theory of probability.

Probability theory because till Kolmogorov, there was no proper definition of probability and he gave the proper definition of probability, which the axiom that also we will be discussing. And Norbert Wiener is credited with this spectral representation of random signals, as I have told you that second order random process, namely the WSS processes can be represented in terms of frequency.

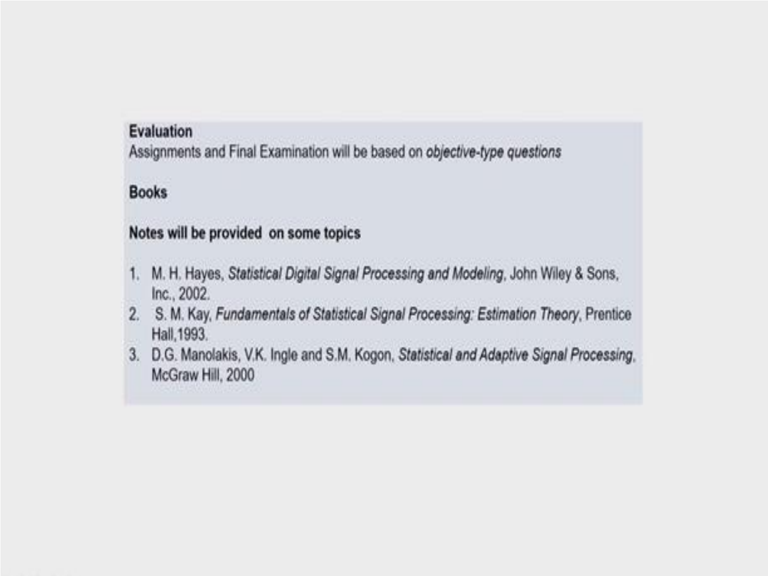
And that is credited to Norbert Wiener, and then many other persons are also kinson for example. Even Einstein knew it, and the Kolmogorov also has contribution, but spectral representation of random signals is mainly attributed to Norbert Wiener. In the optimal filtering theory, there are two great big breakthroughs,

(Refer Slide Time: 21:50)

Namely the Kalman filter which is widely used. And this is due to R E Kalman, he is a famous electrical engineer who has contributed greatly to mathematics and kalman filter is used many application including control and guidance tracking, etc. And another breakthrough is the adaptive filter, and there is a famous algorithm called leastmean square, algorithm and that is due to a Bernard withdrew and another person called Hop.

So, adaptive filter has also revolutionized the application of statistical signal processing and learning the rule of adaptive filter is also used in machine learning. In this course, will evaluate through assignments

(Refer Slide Time: 22:50)



And final examination and all will be based on objective type questions; I will provide notes on the certain selected topics. And my notes will be based on my presentations and notes will be based on these three books. Hayes Statistical digital signal processing and modelling. S M Kay, fundamentals of statistical signal processing: estimation theory. That estimation part will be covered from this book.

Then Manolakis, Ingle and Kogon, statistical and adaptive signal processing. Some part of estimation and adaptive filtering will be covered from this book. Thus, we have gave a brief overview of statistical signal processing. We saw that the different topics discussed under statistical signal processing. In the next class will give the background of statistical signal processing that is probabilistic analysis. Thank you.