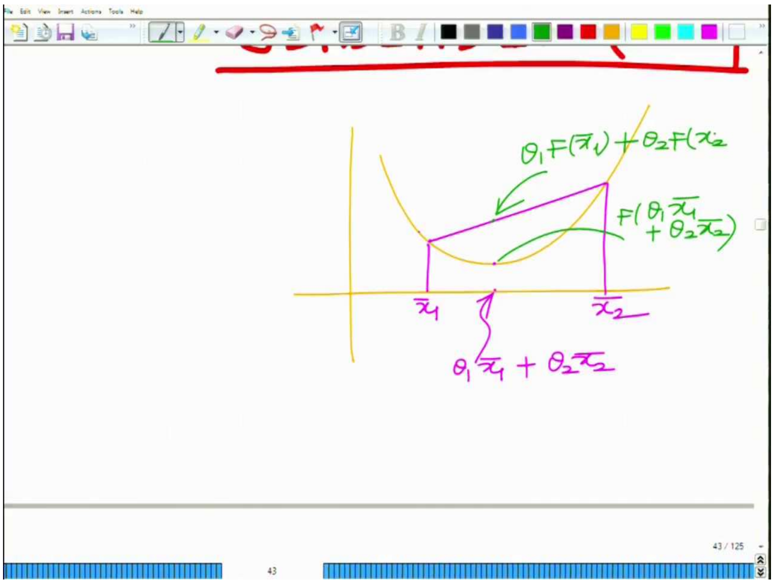
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture - 27
Jensen's Inequality and Practical Application: BER calculation in Wired and Wireless Scenario

Hello. Welcome to another module in this massive open online course. Let us discuss Jensen's inequality which has significant applications in various fields.

(Refer Slide Time: 01:17)



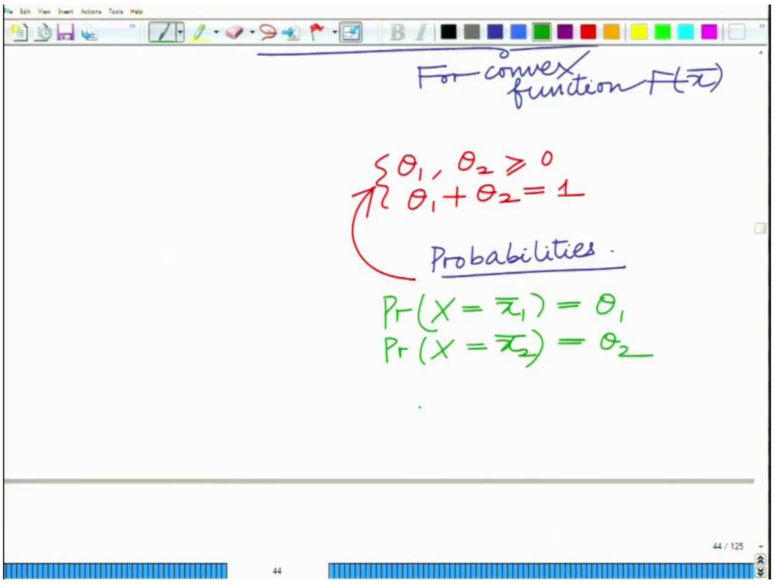
So, first let us look at the definition of the convex function. Consider a function $F(\bar{x})$ where $\bar{x}$ is a vector. Let the domain of this function $F(\bar{x})$ is a convex set. Thus function $F(\bar{x})$ is also convex, if it satisfies the following property.

$$F\left(\theta \bar{x}_1 + (1-\theta)\bar{x}_2\right) \le \theta F(\bar{x}_1) + (1-\theta) F(\bar{x}_2)$$

Where the two points $\bar{x}_1$ and $\bar{x}_2$ belongs to the domain of function $F(\bar{x})$ and θ is a scalar quantity such that $0 \le \theta \le 1$. This can be pictorially represented in the following figure. If θ_1 and θ_2 are set as $\theta_1 = \theta$ and $\theta_2 = 1-\theta$, then this convex function satisfies

$$F(\theta_1 \bar{x}_1 + \theta_2 \bar{x}_2) \le \theta_1 F(\bar{x}_1) + \theta_2 F(\bar{x}_2)$$

(Refer Slide Time: 03:50)



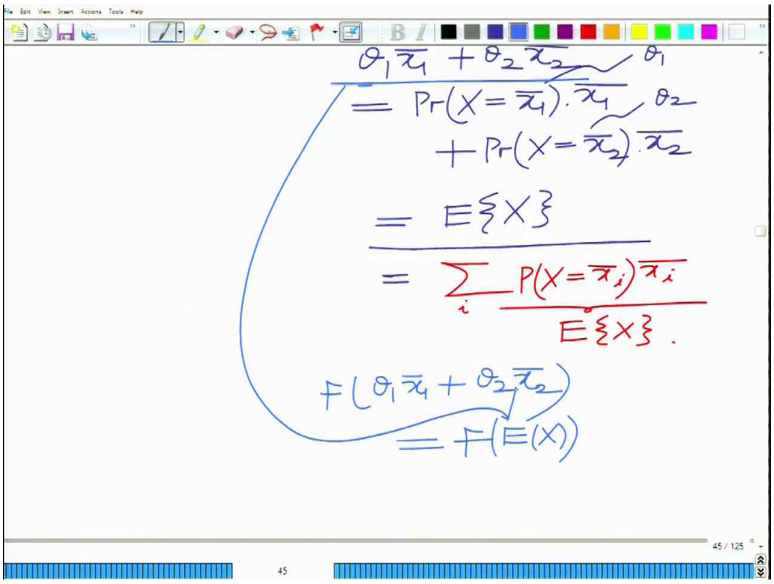
So these θ_1 and θ_2 are non-negative scalar and also the sum of these two is equal to 1.

$\theta_1 \theta_2 \geq 0$ and $\theta_1 + \theta_2 = 1$

Thus these θ_1 and θ_2 can be taken as the probabilities. So, consider a distribution probability of random variable X as

$$\Pr\left(X = \bar{x}_1\right) = \theta_1$$
$$\Pr\left(X = \bar{x}_2\right) = \theta_2$$

(Refer Slide Time: 05:01)



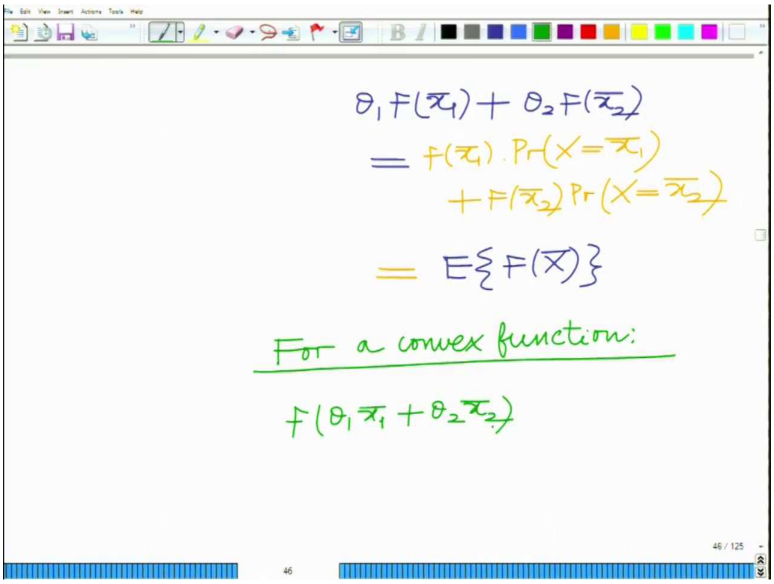
Therefore, the quantity

$$
\begin{aligned}
\theta_1 \bar{x}_1 + \theta_2 \bar{x}_2 &= \Pr(X = \bar{x}_1) \cdot \bar{x}_1 + \Pr(X = \bar{x}_2) \cdot \bar{x}_2 \\
&= \sum_i P(X = \bar{x}_i) \cdot \bar{x}_i \\
&= E\{X\}
\end{aligned}
$$

And hence

$$
F(\theta_1 \bar{x}_1 + \theta_2 \bar{x}_2) = F(E\{X\})
$$

(Refer Slide Time: 07:10)



On the other hand, the quantity

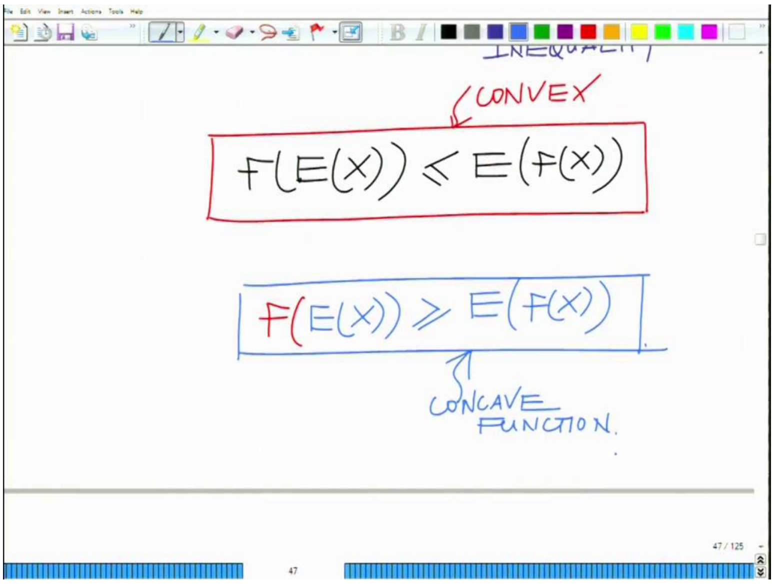
$$
\begin{aligned}
\theta_1 F\left(\bar{x}_1\right)+\theta_2 F\left(\bar{x}_2\right) &= \Pr\left(X=\bar{x}_1\right)\cdot F\left(\bar{x}_1\right)+\Pr\left(X=\bar{x}_2\right)\cdot F\left(\bar{x}_2\right)\\
&=\sum_i P\left(X=\bar{x}_i\right)\cdot F\left(\bar{x}_i\right)\\
&=E\left\{F\left(X\right)\right\}
\end{aligned}
$$

Therefore, for a convex function

$$
\begin{aligned}
F\left(\theta_1\bar{x}_1+\theta_2\bar{x}_2\right) &\le \theta_1 F\left(\bar{x}_1\right)+\theta_2 F\left(\bar{x}_2\right)\\
F\left(E\left\{X\right\}\right) &\le E\left\{F\left(X\right)\right\}
\end{aligned}
$$

This is known as Jensen's inequality for a convex function. It states that for a convex function, the function of an expected value of a random variable is less than expected value of the function of that random variable. This is an important and very handy tool and it is frequently used in signal processing, communication and information theory.

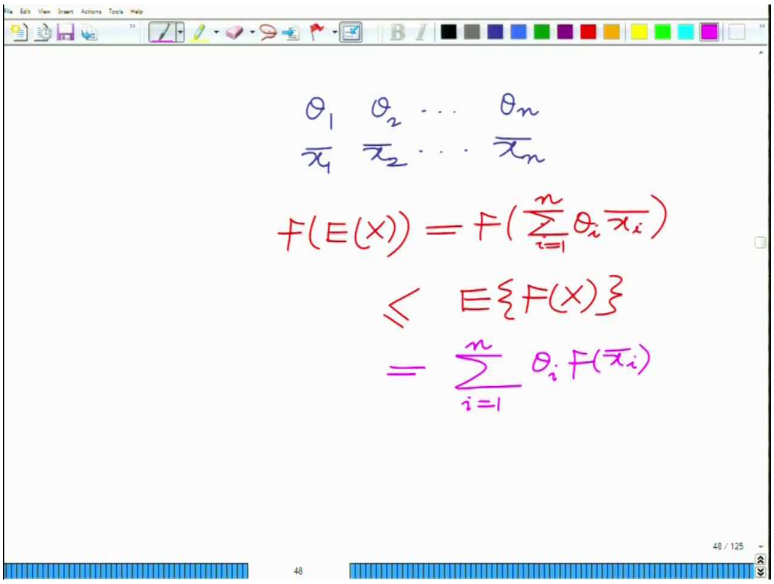
(Refer Slide Time: 10:28)



Similarly, as the cord lies below the curve of a concave function; therefore for a concave function

$$F\left(E\{X\}\right) \geq E\{F(X)\}$$

(Refer Slide Time: 11:50)



Also this will be generalized as follows. Consider an *n*-dimensional scenario. There are n number of θ as $\theta_1, \theta_2, \cdots, \theta_n$ corresponding to $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n$. Therefore;

$$F\left(E\{X\}\right) = F\left(\sum_{i=1}^{n} \theta_i \bar{x}_i\right)$$

And

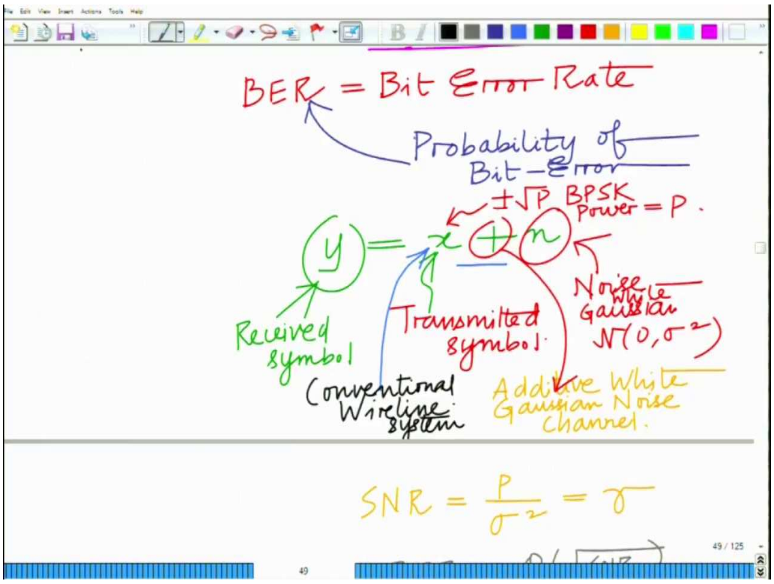
$$E\{F(X)\} = \sum_{i=1}^{n} \theta_i F(\bar{x}_i)$$

And thus the Jensen's inequality for the general case is as follows.

$$F\left(E\{X\}\right) \le E\{F(X)\}$$
$$F\left(\sum_{i=1}^{n} \theta_i \bar{x}_i\right) \le \sum_{i=1}^{n} \theta_i F(\bar{x}_i)$$

It holds even for a continuous random variable X. Let us discuss a practical application of this in the context of a communication system.

(Refer Slide Time: 14:50)

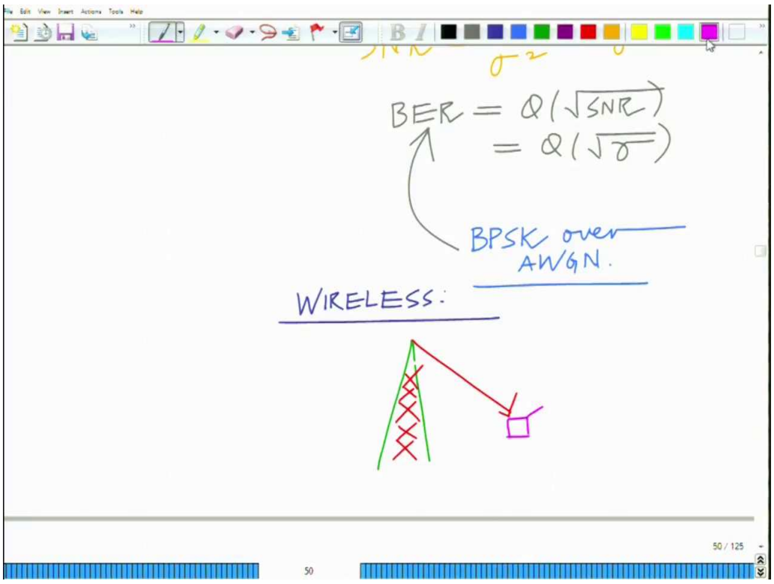


Consider a communication model as

$$y = x + n$$

This is a conventional wireline system where x is the transmitted symbol, y is the received symbol and n is the white Gaussian noise with zero mean and σ^2 variance. Let us discuss about the Bit Error Rate (BER) of this transmission. Basically, BER denotes the probability with which a bit is received in error over a communication channel. Transmitted symbol x can typically be a Binary Phase Shift Keying (BPSK) symbol denoting the square root of the power i.e. $\sqrt{P}$. So, this is an AWGN channel with BPSK signal transmission.

(Refer Slide Time: 16:36)



The SNR is the signal power divided by noise power. Let us denote this by γ.

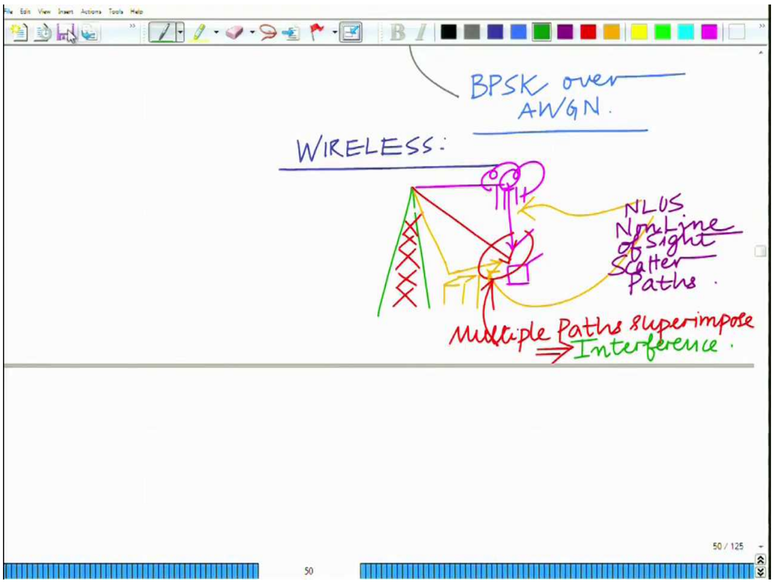
$$SNR = \frac{P}{\sigma^2} = \gamma$$

Also the bit error rate of an additive white BPSK transmission is given by following expression.

$$\begin{aligned} \text{BER} &= Q\left(\sqrt{\text{SNR}}\right) \\ &= Q\left(\sqrt{\gamma}\right) \end{aligned}$$

However, this AWGN model is generally associated with a conventional digital communication system in which there is a wired medium between the transmitter and the receiver. This is also known as a wired channel. A common example of such channel is conventional telephone which uses the twisted copper pair or a coaxial cable.

(Refer Slide Time: 19:14)

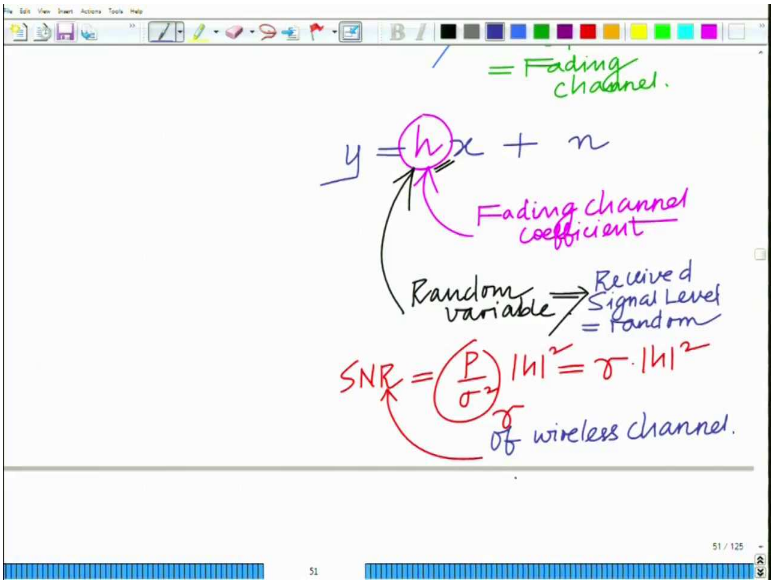


On the other hand, in a wireless system, there is a base station which is transmitting to a mobile in the cell mostly in a direct path. Also there happen several reflections which creates scattered paths. The large buildings or structures which generate such reflections are commonly known as the scatterers. So, these scatterers give rise to the non-line of sights (NLOS) signals. So when these LOS and NLOS signals superimpose, this creates multipath environment which leads to interference. This interfence is both constructive as well as destructive.

Also, the moment this interference is added into the scenario, it introduces uncertainty in the received signal level. The signal level can dip if the interference is destructive or the signal level can rise if the interference is constructive. So, in general, the signal level or power level of the received signal varies with time.

Therefore, because of this multipath reflection phenomenon, the resulting interference leads to a time varying power for the received signal in a wireless system and this process is termed as fading and such wireless channel is known as a fading channel.

(Refer Slide Time: 22:38)



For the wireless channel, let us use a different model which considers fading.

$$y = hx + n$$

Where x is the transmitted symbol, y is the received symbol and n is the white Gaussian noise with zero mean and σ^2 variance. The extra term h is the fading channel coefficient which is a random variable. This implies that the received power or signal level is random and therefore the SNR is influenced by this fading channel coefficient.

So, the SNR of the wireless channel is

$$SNR = \frac{P}{\sigma^2}|h|^2 = \gamma|h|^2$$

And therefore one can look at the resulting bit error rate depending on the SNR of the fading wireless channel and apply Jensen's inequality to derive a suitable conclusion. So, let us continue this discussion in the subsequent module.