RANDOM SIGNALS

Chapter II

Probability Theory and Random Variables

INSTITUT NATIONAL DES SCIENCES APPLIQUÉES TOULOUSE
LAAS CNRS

Objective of Chapter II

- Recall some basic facts concerning Probability theory and Random variables
- Introduce the main notions for the following chapters

1. Axioms of Probability Theory

- Consider a set Ω called a *Probability Space*. For example, consider the rolling of a die, we have

$$\Omega = \{\omega_1 = \text{Face } 1, \omega_2 = \text{Face } 2, \ldots, \omega_6 = \text{Face } 6\}$$

- When the die is rolled, we obtain an element $\omega \in \Omega$.
- We can consider $\Omega_1 = \{\omega_2\} \subset \Omega$ is the event $\text{Face } 2$ turns up when the die is rolled.
- We can also consider $\Omega_2 = \{\omega_2, \omega_4, \omega_6\} \subset \Omega$ is the event $\text{Face } 2$, $\text{Face } 4$ or $\text{Face } 6$ turns up
- Subsets Ω_i are events and two events Ω_1 and Ω_2 are called *mutually exclusive if* $\Omega_1 \cap \Omega_2 = \emptyset$
- To manage events, we can use the set theory notation, intersection $\Omega_1 \cap \Omega_2$, union $\Omega_1 \cup \Omega_2 \ldots$

1. Axioms of Probability Theory

- We assign probabilities to events via a *probability function* $P(\bullet)$ defined on the set of all subsets of Ω denoted $\mathcal{P}(\Omega)$. $P(\bullet)$ satisfies the following axioms

i) $P(\Omega_i) \geqslant 0, \ \forall \Omega_i \in \mathcal{P}(\Omega)$

ii) $P(\Omega) = 1$

iii) If $\Omega_i \cap \Omega_j = \emptyset, \ i \neq j, \ \ i,j = 1, \ldots, n$, then

$$P(\Omega_1 \cup \Omega_2 \cup \ldots, \cup \Omega_n) = P(\Omega_1) + P(\Omega_2) + \ldots + P(\Omega_n)$$

iv) If $\Omega_i \cap \Omega_j = \emptyset, \ i \neq j, \ \ i,j = 1, \ldots, n, \ldots$, then

$$P(\Omega_1 \cup \Omega_2 \cup \ldots, \cup \Omega_n \cup \ldots) = P(\Omega_1) + P(\Omega_2) + \ldots + P(\Omega_n) + \ldots$$

- The triplet $(\Omega, \mathcal{P}(\Omega), P)$ is called a *Random Experiment*.
- For general sets Ω (for example $\mathbb{R}$), some care has to be taken to define an appropriate class of events. This class has to be a *σ-algebra*. In general $\mathcal{P}(\Omega)$ is not a σ-algebra

1. Axioms of Probability Theory

From the above axioms, it is possible to show some special results.

- $P(\emptyset) = 0$
- $P(\Omega_1) = 1 - P\left(\complement_{\Omega}\Omega_1\right)$
- $P(\Omega_1 \cup \Omega_2) = P(\Omega_1) + P(\Omega_2) - P(\Omega_1 \cap \Omega_2) \leqslant P(\Omega_1) + P(\Omega_2)$
- If $\Omega_1 \subset \Omega_2$ then $P(\Omega_2) = P(\Omega_1) + P\left(\Omega_2 \cap \complement_{\Omega}\Omega_1\right) \geqslant P(\Omega_1)$

2. Independence

Definition

Two events Ω_1 and Ω_2 are *independent* if and only if $P(\Omega_1 \cap \Omega_2) = P(\Omega_1).P(\Omega_2)$

- A statement about one event does not affect the statement about the other
- This definition can be extended for more than two events
- When two events are not independent, a statement about one event affects a statement about the other and we can define the conditional probability.

Definition

Consider two events Ω_1 and Ω_2 such that $P(\Omega_2) \neq 0$. The conditional probability of Ω_1 given Ω_2 is defined as

$$P(\Omega_1|\Omega_2) = \frac{P(\Omega_1 \cap \Omega_2)}{P(\Omega_2)}$$

The conditional probability allows the introduction of an a priori knowledge about an event in the evaluation of the probability of another one

3. Random Variables: distribution function

Definition

A fonction $X(\bullet)$ defined on Ω and whose values are in E is called a random variable if for every $x \in E$ the inequality

$$X(\omega) \leqslant x$$

defines a subset of Ω whose probability is defined.

- If $E = \mathbb{R}$, X is called a *real-valued random variable*
- When E is countable, X is called a *discrete random variable*

Definition

The distribution function of a random variable X is the function from E to $[0, 1]$ defined by

$$F_X(x) = P(X \leqslant x) = P(\{\omega \in \Omega : X(\omega) \leqslant x\})$$

3. Random Variables: distribution function

The main properties of the distribution function are

- $F_X(x)$ is monotone nondecreasing i.e. $F_X(x_1) \leqslant F_X(x_2)$ if $x_1 < x_2$
- $\lim_{x \to \infty} F_X(x) = 1$ and $\lim_{x \to -\infty} F_X(x) = 0$
- $P(x_1 < X \leqslant x_2) = F_X(x_2) - F_X(x_1)$
- If $x_1 = x - \epsilon$ and $x_2 = x$, then $P(x - \epsilon < X \leqslant x) = F_X(x) - F_X(x - \epsilon)$. Letting $\epsilon \to 0$, we have
$$P(X = x) = F_X(x) - F_X(x_-)$$
- When E is countable, from the above analysis, F_X is discontinuous at all $x \in E$ where the jump equals $P(X = x)$

3. Random Variables: density function

When the random variable is continuous, we can define its density function.

Definition

If the random variable X is continuous, there exists a density function f_X defined as

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du$$

If $F_X(x)$ is continuous at all x, we have

$$f_X(x) = \frac{d}{dt} F_X(x)$$

at all x where the derivative exists.

3. Random Variables: density function

PROPERTIES

- Since F_X is monotonic, we have $f_X(x) \geqslant 0$
- From the properties of F_X, we also have

$$\int_{-\infty}^{\infty} f_X(u)du = 1$$

- If $x_1 \leqslant x_2$,

$$F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(u)du$$

- We also have

$$P(\{x \leqslant X(\omega) \leqslant x + dx\}) = f_X(x)dx$$

and $P(\{X(\omega) = x\}) = 0$

3. Random Variables: density function

EXAMPLE 1

Consider that a telephone call occurs at random in the time interval $[0, T]$. Let $\Omega = [0, T]$ and consider that

$$P(t_1 \leqslant \omega \leqslant t_2) = \frac{t_2 - t_1}{T},\ t_1, t_2 \in [0, T] \text{ (a call in the interval } [t_1, t_2])$$

Consider the random variable $X(\omega) = \omega$

- If $x > T$ then $\{X(\omega) \leqslant x\} = \Omega$ and $P(\{X(\omega) \leqslant x\}) = F_X(x) = 1$
- If $x < 0$ then $\{X(\omega) \leqslant x\} = \emptyset$ and $P(\{X(\omega) \leqslant x\}) = F_X(x) = 0$
- If $0 \leqslant x \leqslant T$ then $\{X(\omega) \leqslant x\} = \{0 \leqslant \omega \leqslant x\}$ and $P(\{X(\omega) \leqslant x\}) = F_X(x) = x/T$
- The density function is obtained by differentiation

$$f_X(x) = \begin{cases} 0, & x < 0 \\ 1/T, & 0 < x < T \\ 0, & x > T \end{cases}$$

$F_X(x)$ not differentiable at 0 and T

3. Random Variables: density function

EXAMPLE 2

- Consider the experiment of tossing a coin. We have $\Omega = \{p, f\}$, p:heads, f:tails
- $\mathcal{P}(\Omega) = \{\emptyset, \{p\}, \{f\}, \Omega\}$ and $P(\{f\}) = p_1$, $P(\{p\}) = p_2$, $p_1 + p_2 = 1$
- Define the random variable $X(\omega)$ by $X(p) = 1$ and $X(f) = 0$
- To determine the distribution function, remark that
  - If $x \geqslant 1$ then $\{X(\omega) \leqslant x\} = \{X(\omega) = 0\} \cup \{X(\omega) = 1\} = \Omega \Rightarrow F_X(x) = 1$
  - If $x < 0$ then $\{X(\omega) \leqslant x\} = \emptyset \Rightarrow F_X(x) = 0$
  - If $0 \leqslant x < 1$ then $\{X(\omega) \leqslant x\} = \{X(\omega) = 0\} = \{f\} \Rightarrow F_X(x) = p_1$
  - To unify discrete and continuous random variables, we can obtain the density function by differentiation introducing the Dirac delta function δ. We have

$$f_X(x) = p_1\delta(x) + p_2\delta(x-1)$$

and

$$\int_{-\infty}^{\infty} f_X(u)du = p_1 + p_2 = 1$$

$$\left(\int_{-\infty}^{\infty} f(u)\delta(u-u_0)du = f(u_0)\right)$$

3. Random Variables: Expectation and moments

Definition

- If X is a random variable, we define (when it exists) the expectation (also called average or first moment) by

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

- More generally if $y = g(x)$, $g(\bullet)$ being a fixed real function, we have

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

  In particular case when $g(x) = x^n$, the indicator is called the *nth moment*.

- The nth central moments are defined as

$$E[(X - E[X])^n] = \int_{-\infty}^{\infty} (x - E[x])^n f_X(x) dx$$

3. Random Variables: Expectation and moments

- $E[X^2]$ is called the mean square value of X
- $E[(X - E[X])^2] = var[X]$ is the variance
- $\sigma[X] = (var[X])^{1/2}$ is the standard deviation. It measures the dispersion about the mean of X
- If X is discrete taking values x_i with probabilities p_i, $i = 1, \ldots, N$, then

$$f_X(x) = \sum_{i=0}^{N} p_i \delta(x - x_i)$$

and

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx = \sum_{i=0}^{N} x_i^n p_i$$

3. Random Variables: Expectation and moments

PROPERTIES

- $E[X]$ is a linear operator i.e. $E[aX + bY] = aE[X] + bE[Y], \forall\, a, b$
- If a is a constant $E[a] = a$
- $var[aX + b] = a^2 var[X], \forall\, a, b$
- $var[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$. (Huyghens's Theorem)
- $\min_a E[(X - a)^2] = E[(X - E[X])^2] = var[X]$

A random variable may also be specified in terms of its characteristic function defined as

$$\Phi_X(u) = E[e^{juX}] = \int_{-\infty}^{\infty} f_X(x) e^{jux} dx \text{ (Inverse Fourier transform of } f_X)$$

and

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_X(u) e^{-jux} dx$$

In particular

$$E[X^n] = j^{-n} \frac{d^n \Phi_X}{du^n}(0) \text{ and } \Phi_X(u) = \sum_{k=0}^{\infty} \frac{j^k E[X^k]}{k!} u^k$$

3. Random Variables: Examples

Bernoulli distribution

- $\Omega = \{\omega_1, \omega_2\}$ and $X(\Omega) = \{0, 1\}$
- $P(X = 1) = p$ and $P(X = 0) = 1 - p$, $p \in [0, 1]$
- The distribution is given by

$$F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - p, & 0 \leqslant x < 1 \\ 1, & x \geqslant 1 \end{cases}$$

- The density function is

$$f_X(x) = (1 - p)\delta(x) + p\delta(x - 1)$$

- $E[X] = p$ and $var[X] = p(1 - p)$
- $\Phi_X(u) = 1 - p + pe^{ju}$

3. Random Variables: Examples

Binomial distribution

- $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, $X(\Omega) = \{0, 1, \dots, n\}$, $X = \sum_{i=1}^{n} X_i$

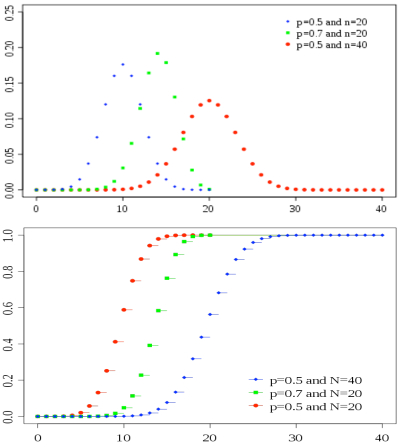
  (X_i : Bernoulli distributed and mutually independent)

- $P(X = k) = C_n^k p^k (1-p)^{n-k}$ $k = 1, \dots, n$, $n \in \mathbb{N}^*$, $p \in [0, 1]$
- The density function is

$$f_X(x) = \sum_{k=1}^{n} P(X = k)\delta(x - k)$$

- $E[X] = np$ and $var[X] = np(1-p)$
- $\Phi_X(u) = (1 - p + pe^{ju})^n$

3. Random Variables: Examples



Binomial random variable: density and distribution functions

3. Random Variables: Examples

Poisson distribution

- $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$, $X(\Omega) = \mathbb{N}^*$
- The probabilities are

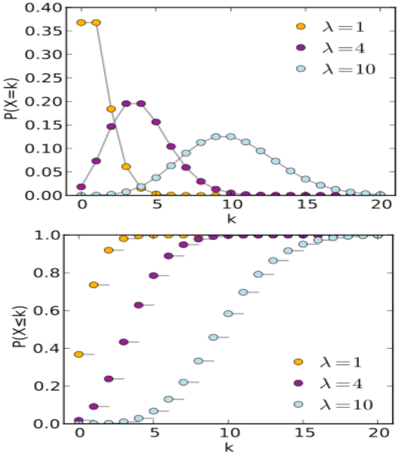
$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \lambda : \text{ positive parameter}$$

- The density function is

$$f_X(x) = \sum_{k=1}^{\infty} P(X = k)\delta(x - k)$$

- $E[X] = \lambda$ and $var[X] = \lambda$
- $\Phi_X(u) = e^{\lambda(e^{ju} - 1)}$

3. Random Variables: Examples



Poisson random variable: density and distribution functions

3. Random Variables: Examples

Uniform distribution (continuous)

- $X(\Omega) = \mathbb{R}$
- The density function is given by

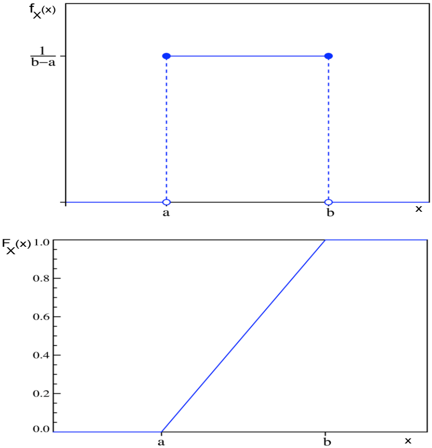
$$f_X(x) = \begin{cases} 0, & x < a \\ \dfrac{1}{b-a}, & a \leqslant x \leqslant b \\ 0, & x > b \end{cases} \qquad a, b \in \mathbb{R}$$

- The distribution function is

$$F_X(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \leqslant x \leqslant b \\ 1, & x > b \end{cases}$$

- $E[X] = (a+b)/2$ and $var[X] = (b-a)^2/12$
- $\Phi_X(u) = \dfrac{e^{jbu} - e^{jbu}}{it(b-a)}$

3. Random Variables: Examples



Uniform random variable: density and distribution functions

3. Random Variables: Examples

Gaussian (Normal) distribution (continuous)

- $X(\Omega) = \mathbb{R}$
- The density function is given by

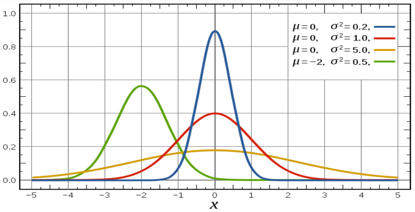
$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \mu, \sigma \in \mathbb{R}^+$$

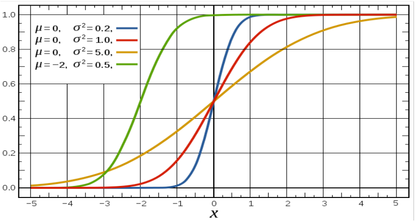
- The distribution function is

$$F_X(x) = P(X \leqslant x) = \frac{1}{2} + \mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right), \quad \left(\mathrm{erf}(u) = \frac{2}{\sqrt{\pi}}\int_{-\infty}^{u} e^{-u^2} du\right)$$

- $E[X] = \mu$ and $var[X] = \sigma$
- $\Phi_X(u) = e^{j\mu u - \frac{\sigma^2 u^2}{2}}$

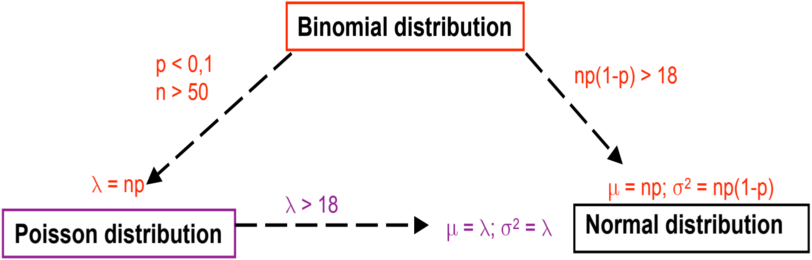
3. Random Variables: Examples





Normal random variable: density and distribution functions

3. Random Variables: Approximations



Approximations between distributions

3. Random Variables: Tchebycheff's Inequatlity

Theorem (Tchebycheff)

Consider a random variable X*. We have*

$$P(E[X] - \epsilon < X < E[X] + \epsilon) \geqslant 1 - \frac{{\sigma_X}^2}{\epsilon^2}$$

For any random variable, if $\sigma_X \ll \epsilon$, the probability that X takes values in the interval $[E[X] - \epsilon, E[X] + \epsilon]$ is equal to 1.

4. Pair of Random Variables

In some situations, it could be necessary to consider several random variables. For the sequel, consider a pair of random variables will be necessary.

Definition

- Consider two random variables X and Y. We say that they are *jointly distributed* if the are defined on the same probability space.
- They may be characterized by their *joint distribution function*

$$F_{XY}(x,y) = P(\{X(\omega) \leqslant x\} \cap \{Y(\omega) \leqslant y\})$$

  or their *joint density function*

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(u,w)dudw$$

- It follows that

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}}{\partial x \partial y}(x,y)$$

4. Pair of Random Variables

PROPERTIES

- $F_{XY}(-\infty, y) = F_{XY}(x, -\infty) = F_{XY}(-\infty, -\infty) = 0 \quad \forall x, y$
- $F_{XY}(\infty, \infty) = 1$, $F_{XY}(x, y) \geqslant 0$ and $f_{XY}(x, y) \geqslant 0 \quad \forall x, y$
- If $x_2 > x_1$ then $F_{XY}(x_2, y) - F_{XY}(x_1, y) = P(\{x_1 < X(\omega) \leqslant x_2\} \cap \{Y(\omega) \leqslant y\})$
- $P(\{x < X(\omega) \leqslant x + dx\} \cap \{y < Y(\omega) \leqslant y + dy\}) = f_{XY}(x, y)dxdy$
- We have
$$\iint f_{XY}(x, y)dxdy = 1$$
- If X and Y are discrete random variables, the density function is given by
$$f_{XY}(x, y) = \sum_i \sum_j p_{ij}\delta(x - x_i)\delta(y - y_j) \text{ with } p_{ij} = P(X = x_i \text{ and } Y = y_j)$$
- If X and Y are independent the joint density and distribution functions satisfy
$$f_{XY}(x, y) = f_X(x)f_Y(y) \text{ and } F_{XY}(x, y) = F_X(x)F_Y(y)$$

4. Pair of Random Variables

We can also extend the moments to the case of a pair of random variables

Definition

- If X and Y are random variables, we define the (n, m)-*moment* by

$$E[X^nY^m] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^n y^m f_{XY}(x, y)dxdy$$

- More generally if $y = g(x, y))$, $g(\bullet, \bullet)$ being a fixed function, we have

$$E[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f_{XY}(x, y)dxdy$$

- The (n, m)-central moments are defined as

$$E[(X - E[X])^n(Y - E[Y])^m] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - E[X])^n(y - E[Y])^m f_{XY}(x, y)dxdy$$

4. Pair of Random Variables

PROPERTIES

- $E[X]$ is the $(1,0)$-moment, $E[Y]$ is the $(0,1)$-moment and the $(0,0)$-moment is equal to 1
- The$(1,1)$-central moment is important. It is called the *covariance* and is denoted by $cov[X,Y]$
- We have $cov[X,Y] = E[XY] - E[X]E[Y]$ and $cov[X,X] = \sigma_X^2 = var[X]$
- $E[XY]$ is a scalar product. When $E[XY] = 0$, we say that X and Y are *orthogonal*
- The ratio

$$\rho[X,Y] = \frac{cov[X,Y]}{\sigma_X \sigma_Y}$$

defines the correlation coefficient of X and Y. We have $|\rho[X,Y]| \leqslant 1$
- When $\rho[X,Y] = 0$, X and Y are said *uncorrelated*
- When $|\rho[X,Y]| = 1 \Leftrightarrow Y = aX + b$ with probability 1 and

$$a = \frac{cov[X,Y]}{var[X]} \text{ and } b = E[Y] - \frac{cov[X,Y]E[X]}{var[X]}$$

- We can also define the characteristic function as $\Phi_{YX}(u_1,u_2) = E[e^{i(u_1X+u_2Y)}]$
- X and Y independent $\Rightarrow E[XY] = E[X]E[Y]$, $cov[X,Y] = 0$.The converse is in general false
- X and Y independent $\Leftrightarrow \Phi_{YX}(u_1,u_2) = \Phi_X(u_1)\Phi_Y(u_2)$

5. Conditional Probabilities and Expectations

Definition

- Consider two random variables X and Y. We define the conditional density function $f_{X|Y}(x|y)$ of x given $\{Y(\omega) = y\}$ for all x and y such that $f_Y(y) > 0$ by

$$f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{f_{XY}(x,y)}{\int f_{XY}(x,y)dx}$$

were $f_Y(y)$ is called the *marginal* density function of Y

- Reversing the role of X and Y, we obtain

$$f_{XY}(x,y) = f_{Y|X}(y|x)f_X(x)$$

- Substituting, we obtain the *Bayes' rule*

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)}$$

PROPERTIES

Let X and Y be jointly distributed random variables. We have

- $f_{X|Y}(x|y) \geqslant 0$
- $\int f_{X|Y}(x|y)dx = 1$
- $f_{X|Y}(x|y)$ depends of the realizations of Y, it is itself a random variable. More precisely, it is a function of the random variable Y
- $f_X(x) = E[f_{X|Y}(x|Y)] = \int f_{X|Y}(x|y) f_Y(y) dy$
- $f_{X|Y}(x|y) = f_X(x)$ if X and Y are independent

5. Conditional Probabilities and Expectations

Definition

- The conditional expectation of the random variable X, given the random variable Y is defined by

$$E[X|Y] = \int x f_{X|Y}(x|y)dx$$

$E[X|Y] =$ is a random variable because it depends of the realizations of Y

- The conditional variance matrix is then defined by

$$\sigma^2_{X|Y} = E[(X - E[X|Y])^2|Y]$$

Note that $\sigma^2_{X|Y}$ is also a random variable.

PROPERTIES

Let X, Y, Z be jointly distributed random variables and $g(\bullet)$ a scalar-valued function. Then

- $E[X|Y] = E[X]$ if X and Y are independent
- $E[X] = E[E[X|Y]]$
- $E[g(Y)X|Y] = g(Y)E[X|Y]$
- $E[g(Y)X] = E[g(Y)E[X|Y]]$
- $E[a|Y] = a \ \forall a$
- $E[g(Y)|Y] = g(Y)$
- $E[aX + bY|Z] = aE[X|Z] + bE[Y|Z] \ \forall a, b$