EECS 240: Random Processes

Thomas Ketseoglou

UCI

September 24, 2015

Outline

Review

- What is probability? What is conditional probability?
- How can it be used in ECE?
- What is random variable (r.v.)?
- What is a Gaussian r.v.? What is discrete r.v, what is continuous r.v.?
- What is pmf. CDF, pdf?
- What are independent r.v.'s? What are uncorrelated ones? If $X, \ Y$ are independent r.v.'s, what is the pdf of $Z = X + Y$?
- What is a random process (r.pr.)?
- What is a Wide Sense Stationary (W.S.S.) process?
- What is an ergodic r.pr.?
- What is a Gaussian r.pr.?
- What is a Gaussian random vector (R.V.)?
- What is a multi-dimensional p.d.f.?
- What is expectation?
- What is conditional expectation?

What is probability? (cont'd)

- Probability is a number assigned to sets called events. The higher the value of probability, the higher the likelihood of the event.
- An outcome of an experiment is any possible observation of that experiment.
- **Sample Space:** The sample space of an experiment is the finest grain, mutually exclusive collectively exhaustive set of all possible outcomes.
- **Event:** An event is practically an set containing outcomes of an experiment on which a probability can be assigned.

What is probability? (cont'd)

Definition

A probability space $(P, \ S, \ \mathcal{S})$ where S is the sample space (i.e. all possible outcomes), P is the probability (the likelihood), and $\mathcal{S}$ is the set of all the events (subsets over which we assign probabilities) needs to obey three axioms:

1. $P(A) \geq 0$
2. $P(S) = 1$
3. For any finite collection of disjoint events, i.e. $A_{i}{}_{i=1}^{n}$ s.t. for any $i \neq j, \ i, \ j = 1, \cdots, n | A_i \cap A_j = \emptyset$, we have

$$P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

What is a random variable?

Definition

A random variable is a function X from the probability space to $\mathbb{R}$. Depending on the range of this function, we have:

1. discrete r.v. if the range is finite, or countable, and
2. continuous r.v. if the range is continuous.

What is a Gaussian, or normal r.v.?

Definition

A continuous r.v. is normal if its pdf is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{(x-m)^2}{2\sigma^2})$$

for $x \in \mathbb{R}$.

- The mean of X is m.
- The variance of X is σ^2.

What is a r.pr.?

- A random process, $X(t)$ is a function of time s.t. for each $t = t_0$, $X(t_0)$ is a r.v.
- Examples include: Noise, received power in wireless, an image (2-D), fading.

Wide Sense Stationarity

- An r.pr. $X(t)$ is WSS iff:
  1. $E(X(t)) = \textit{constant}$ (independent of t).
  2. $R_X(t,\ \tau) = E(X(t+\tau)X(t)) = R_X(\tau)$ for $\forall t,\ \tau$
- The power spectral density of a W.S.S. $X(t)$ is then $S_X(f) = \mathcal{F}\{R_X(\tau)\}$

Ergodicity

- A WSS random process $X(t)$ is ergodic, iff a statistical mean that relates to the process equals the corresponding time average. For example a process $X(t)$ is mean ergodic iff:

$$E(X(t)) = \lim_{T \to +\infty} \frac{1}{2T} \int_{-T}^{+T} X(t)dt$$

Independence, uncorrelatedness, etc.

Definition

Two r.v.'s are independent iff $f_{X,\ Y}(x,\ y) = f_X(x) f_Y(y)$.

Definition

Two r.v.'s are uncorrelated iff $E(XY) = E(X)E(Y)$

- Independence $\Rightarrow$ Uncorrelatedness, but not vice versa.
- **Reminder:** The correlation coefficient, $\rho = \frac{E((X-m_X)(Y-m_Y))}{\sigma_X \sigma_Y} = \frac{E(XY)-E(X)E(Y)}{\sigma_X \sigma_Y}$ is always in $[-1,\ +1]$.
- For jointly Gaussian r.v.'s Independence $\Leftrightarrow$ uncorrelatedness.

Independence, uncorrelatedness, etc.

- If $X, \ Y$ are independent, then $Z + X + Y$ has pdf given by the convolution of the X and Y pdf's, i.e.

$$f_Z(z) = f_X(x) * f_Y(y) = \int_{-\infty}^{+\infty} f_X(z-u) f_Y(u) du$$

pmf, CDF, PDF, pdf

Definition

The probability mass function of a d.r.v. is the probability of X at each x_i, $P_X(x_i)$.

Definition

The cumulative distribution function (CDF) (also called PDF), $F_X(x)$ is a function defined as follows:

$$F_X(x) = P(X \leq x)$$

Definition

The probability density function of a c.r.v., $f_X(x)$ is defined as follows:

$$f_X(x) = \frac{dF_X(x)}{dx}$$

Multi-dimensional pdf

- $f_{X_1,X_2,\cdots,X_n}(x_1,\cdots,x_n)$ is a multidimensional pdf if it describes the probabilistic description of $\mathbf{X} = X_1,\cdots,X_n$. This means that:
  1. $f(x_1,\cdots,x_n) \geq 0 \ \ \forall \ \ x_1,\cdots,x_n$
  2. $P(\mathbf{X} \in \mathcal{A}) = \int_{\mathcal{A}} f(_1,\cdots,x_n)dx_1\cdots dx_n$

Expectations

- For an r.v., X with pdf $f_X(x)$

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

- For any function of X, $g(X)$

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

This is an extremely useful fact!

Conditional Expectations

- $E(X|Y) = \int_{-\infty}^{+\infty} x f_{X|Y}(x|y) dx$ where $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$ is the conditional pdf of $X|Y$.
- $E(g(X)|Y) = \int_{-\infty}^{+\infty} g(x) f_{X|Y}(x|y) dx$ where $f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$. This is also extremely useful!

Important Property of Expectation: Chain Rule

Theorem

Let $g(X_1, X_2, \cdots, X_N)$ a function of r.v.'s $X_1, \cdots, X_N$. Then

$$E(g(X_1, \cdots, X_N)) = E\{E(g(X_1, \cdots, X_N)|X_2, \cdots, X_N)\}$$

. Note that the outer expectation is over $x_2, \cdots, x_N$, while the inner one is over x_1.

Proof.

The inner expectation is equal to $\int_{-\infty}^{+\infty} g(x_1, \cdots, x_N) f(x_1|x_2, \cdots, x_N) dx_1$. Taking the expectation of this w.r.t. $X_2, \cdots, X_N$ results into

$$\int \cdots \int_{-\infty}^{+\infty} f(x_2, \cdots, x_N) dx_2 \cdots dx_N \int_{-\infty}^{+\infty} g(x_1, \cdots, x_N) f(x_1|x_2, \cdots, x_N) dx_1$$

By inserting all terms into the inner integral we get that the integral equals

$$\int \cdots \int_{-\infty}^{+\infty} f(x_2, \cdots, x_N) f(x_1|x_2, \cdots, x_N) g(X_1, \cdots, X_N) dx_1 \cdots dx_N$$

which gives the desired result as $f(x_2, \cdots, x_N) f(x_1|x_2, \cdots, x_N) = f(x_1, x_2, \cdots, x_N)$. The applications of this theorem are endless. We will see one application to pair-wise Gaussian r.v.'s. $\square$

Jointly Gaussian r.v.'s

Definition

$X, \ Y$ are jointly Gaussian if for some $\rho, \ m_X, \ m_Y, \ \sigma_X, \ \sigma_Y$ (under what conditions on their values?) the joint pdf is given as:

$$f_{X,Y}(x, \ y) = \frac{1}{2\pi\sigma_X\sigma_Y(1-\rho^2)^{1/2}} \times$$

$$\exp\left(-\frac{\sigma_Y^2(x-m_X)^2 + \sigma_X^2(1-m_Y)^2 - 2\rho\sigma_X\sigma_Y(x-m_X)(y-m_Y)}{2\sigma_X^2\sigma_Y^2(1-\rho^2)}\right)$$

Jointly Gaussian r.v.'s (cont'd)

It can be proved that the following facts are true:

1. $E(X) = m_X, \; Var(X) = \sigma_X^2$, X is Gaussian. Similarly for Y.

Jointly Gaussian r.v.'s (cont'd)

It can be proved that the following facts are true:

1. $E(X) = m_X,\ Var(X) = \sigma_X^2$, X is Gaussian. Similarly for Y.
2. $E(X|Y) = m_X + \rho\frac{\sigma_X}{\sigma_Y}(Y - m_Y)$, $Var(X|Y) = \sigma_X^2(1-\rho^2)$, and $X|Y$ is Gaussian. Similarly for $Y|X$.

Jointly Gaussian r.v.'s (cont'd)

It can be proved that the following facts are true:

1. $E(X) = m_X,\ Var(X) = \sigma_X^2$, X is Gaussian. Similarly for Y.
2. $E(X|Y) = m_X + \rho\frac{\sigma_X}{\sigma_Y}(Y - m_Y)$, $Var(X|Y) = \sigma_X^2(1 - \rho^2)$, and $X|Y$ is Gaussian. Similarly for $Y|X$.
3. ρ is the actual correlation coefficient between $X,\ Y$.

Jointly Gaussian r.v.'s (cont'd)

It can be proved that the following facts are true:

1. $E(X) = m_X,\ Var(X) = \sigma_X^2$, X is Gaussian. Similarly for Y.
2. $E(X|Y) = m_X + \rho\frac{\sigma_X}{\sigma_Y}(Y - m_Y)$, $Var(X|Y) = \sigma_X^2(1-\rho^2)$, and $X|Y$ is Gaussian. Similarly for $Y|X$.
3. ρ is the actual correlation coefficient between $X,\ Y$.

Proof.

$\square$

Jointly Gaussian r.v.'s (cont'd)

It can be proved that the following facts are true:

1. $E(X) = m_X,\ Var(X) = \sigma_X^2$, X is Gaussian. Similarly for Y.
2. $E(X|Y) = m_X + \rho\frac{\sigma_X}{\sigma_Y}(Y - m_Y)$, $Var(X|Y) = \sigma_X^2(1-\rho^2)$, and $X|Y$ is Gaussian. Similarly for $Y|X$.
3. ρ is the actual correlation coefficient between $X,\ Y$.

Proof.

1. First rewrite the joint pdf in a form similar to $f_{X|Y}(x|y)f_Y(y)$ as follows:

$$\frac{1}{\sqrt{2\pi}\sigma_X(1-\rho^2)^{1/2}} \exp\left(-\frac{((x-m_X)\sigma_Y - \rho(y-m_Y)\sigma_X)^2}{2\sigma_X^2\sigma_Y^2(1-\rho^2)}\right) \times$$
$$\frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left(-\frac{(y-m_Y)^2}{2\sigma_Y^2}\right)$$

□

Proofs (cont'd)

Integrating w.r.t. x to get the pdf of Y gives that $f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left(-\frac{(y-m_Y)^2}{2\sigma_Y^2}\right)$. Then, $f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi}\sigma_X(1-\rho^2)^{1/2}} \exp\left(-\frac{((x-m_X)\sigma_Y-\rho(y-m_Y)\sigma_X)^2}{2\sigma_X^2\sigma_Y^2(1-\rho^2)}\right)$.

- This $f_{X|Y}(x|y)$ is Gaussian. Why?

Proofs (cont'd)

Integrating w.r.t. x to get the pdf of Y gives that $f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left(-\frac{(y-m_Y)^2}{2\sigma_Y^2}\right)$. Then, $f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi}\sigma_X(1-\rho^2)^{1/2}} \exp\left(-\frac{((x-m_X)\sigma_Y-\rho(y-m_Y)\sigma_X)^2}{2\sigma_X^2\sigma_Y^2(1-\rho^2)}\right)$.

- This $f_{X|Y}(x|y)$ is Gaussian. Why?
- Rewrite this as $f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi}\sigma_X(1-\rho^2)^{1/2}} \exp\left(-\frac{(x-(m_X+\rho(y-m_Y)\sigma_X))^2}{2\sigma_X^2(1-\rho^2)}\right)$. Then, $X|Y$ is Gaussian with the required mean, variance.

Proofs (cont'd)

Integrating w.r.t. x to get the pdf of Y gives that $f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left(-\frac{(y-m_Y)^2}{2\sigma_Y^2}\right)$. Then,

$f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi}\sigma_X(1-\rho^2)^{1/2}} \exp\left(-\frac{((x-m_X)\sigma_Y-\rho(y-m_Y)\sigma_X)^2}{2\sigma_X^2\sigma_Y^2(1-\rho^2)}\right)$.

- This $f_{X|Y}(x|y)$ is Gaussian. Why?
- Rewrite this as $f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi}\sigma_X(1-\rho^2)^{1/2}} \exp\left(-\frac{(x-(m_X+\rho(y-m_Y)\sigma_X))^2}{2\sigma_X^2(1-\rho^2)}\right)$. Then, $X|Y$ is Gaussian with the required mean, variance.
- Can you prove that ρ is the correlation coefficient of X, Y, by using the chain rule? Give it a try!

Minimum Mean Square Error Estimation (MMSEE)

- The MMSEE of X given Y is always the conditional mean $E(X|Y)$.

Minimum Mean Square Error Estimation (MMSEE)

- The MMSEE of X given Y is always the conditional mean $E(X|Y)$.
- By definition, this is the estimate of X given Y that minimizes the mean square error over all the estimates of X given Y, $g(Y)$:

$$E(X|Y) = \arg\min_{g(Y)} E_{X|Y} \{X - g(Y)\}^2$$

Minimum Mean Square Error Estimation (MMSEE)

- The MMSEE of X given Y is always the conditional mean $E(X|Y)$.
- By definition, this is the estimate of X given Y that minimizes the mean square error over all the estimates of X given Y, $g(Y)$:

$$E(X|Y) = \arg\min_{g(Y)} E_{X|Y}\left\{X - g(Y)\right\}^2$$

- Let's prove this!

Minimum Mean Square Error Estimation (MMSEE)

- The MMSEE of X given Y is always the conditional mean $E(X|Y)$.
- By definition, this is the estimate of X given Y that minimizes the mean square error over all the estimates of X given Y, $g(Y)$:

$$E(X|Y) = \arg\min_{g(Y)} E_{X|Y}\left\{X - g(Y)\right\}^2$$

- Let's prove this!
- The mean square error of any estimate given Y is $E_{X|Y=y}(g(y) - X)^2 = \int_{-\infty}^{+\infty}(g(y) - X)^2 f_{X|Y=y}(X|Y = y)dx$ This is a convex function of $g(x)$. Taking the derivative of the MSE w.r.t. $g(x)$ and setting it to zero gives the desired result as follows:

$$\frac{\theta}{\theta g(x)} \int_{-\infty}^{+\infty} (g(y) - X)^2 f_{X|Y=y}(X|Y = y)dx = 0$$

is equivalent to
$g(x) = \int_{-\infty}^{+\infty}(X) f_{X|Y=y}(X|Y = y)dx = E(X|Y = y)$.