ECE 2521: Analysis of Stochastic Processes

Lecture 6

Department of Electrical and Computer Engineering
University of Pittsburgh

October, 6^{th} 2021

Azime Can-Cimino

Correlation and Covariance

- Given any two random variables X and Y, we are often interested in understanding their relationship
- Their relationship can be examined using their joint PDF
- .... and also using parameters such as their correlation, covariance, and correlation coefficient
- **Correlation** of two random variables X and Y is:

$$R_{X,Y} = E[XY]$$

Note: If $R_{X,Y} = 0$, then X and Y are *orthogonal*.

- **Covariance** of two random variables X and Y is:

$$\begin{aligned} C_{X,Y} &= \text{Cov}[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] \\ &= E[XY] - \mu_X\mu_Y = R_{X,Y} - \mu_X\mu_Y \end{aligned}$$

Correlation Coefficient

- **Correlation Coefficient** of two random variables X and Y:

$$\rho_{X,Y} = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{R_{X,Y} - \mu_X \mu_Y}{\sigma_X \sigma_Y}$$

- The units of the correlation and covariance are the product of the units of X and Y:

  For example, if X has units of volts and Y has units of seconds, then their correlation and covariance have units of volts-seconds. The correlation coeffient $\rho_{X,Y}$ is dimensionless.

- The correlation coefficient is bounded: $-1 \leq \rho_{X,Y} \leq 1$.
- If $\rho_{X,Y}$ is close to ± 1, then X and Y are highly correlated.
- If $\rho_{X,Y} = 0$, then X and Y are **uncorrelated**.
- If X and Y are independent, then they are also uncorrelated. But the converse is not always true.

Example 1

- Random variables L and T have joint PMF:

$P_{L,T}(l,t)$	$t = 40$ sec	$t = 60$ sec
$l = 1$ page	0.15	0.1
$l = 2$ pages	0.3	0.2
$l = 3$ pages	0.15	0.1

- Find the following quantities:
  (1) $E[L]$ and $\text{Var}[L]$.
  (2) $E[T]$ and $\text{Var}[T]$.
  (3) The correlation $R_{L,T} = E[LT]$
  (4) The covariance $C_{L,T} = \text{Cov}[L, T]$
  (5) The correlation coefficient $\rho_{L,T}$

Example 1 - Solution

It is helpful to first make a table that includes the marginal PMFs.

$P_{L,T}(l,t)$	$t=40$	$t=60$	$P_L(l)$
$l=1$	0.15	0.1	0.25
$l=2$	0.3	0.2	0.5
$l=3$	0.15	0.1	0.25
$P_T(t)$	0.6	0.4	

(1) The expected value of L is

$$E[L] = 1(0.25) + 2(0.5) + 3(0.25) = 2.$$

Since the second moment of L is

$$E\left[L^2\right] = 1^2(0.25) + 2^2(0.5) + 3^2(0.25) = 4.5,$$

the variance of L is

$$\mathrm{Var}[L] = E\left[L^2\right] - (E[L])^2 = 0.5.$$

Example 1 - Solution (continued)

(2) The expected value of T is

$$E[T] = 40(0.6) + 60(0.4) = 48.$$

$$E\left[T^2\right] = 40^2(0.6) + 60^2(0.4) = 2400.$$

$$\mathrm{Var}[T] = E\left[T^2\right] - (E[T])^2 = 2400 - 48^2 = 96.$$

(3) The correlation is

$$\begin{aligned} E[LT] &= \sum_{t=40,60} \sum_{l=1}^{3} lt P_{LT}(lt) \\ &= 1(40)(0.15) + 2(40)(0.3) + 3(40)(0.15) \\ &\quad + 1(60)(0.1) + 2(60)(0.2) + 3(60)(0.1) = 96 \end{aligned}$$

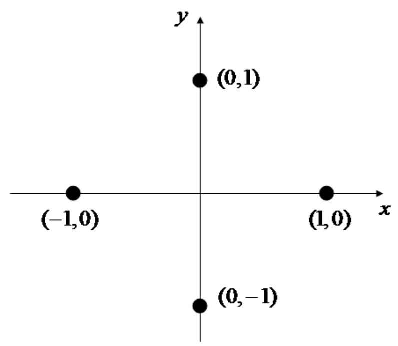
(4) The covariance of L and T is

$$\mathrm{Cov}[L, T] = E[LT] - E[L]E[T] = 96 - 2(48) = 0$$

(5) Since $\mathrm{Cov}[L, T] = 0$, the correlation coefficient is $\rho_{L,T} = 0$.

Example 2 (Uncorrelated but dependent random variables)

- X and Y are two discrete random variables with the joint PMF shown below with probability $1/4$ at each point
- Joint probability mass function for discrete random variables X and Y that are dependent but uncorrelated



- X and Y are symmetric about 0:

$$E[X] = E[Y] = 0$$

$$E[XY] = 0$$

$$\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = 0$$

- The random variables X and Y are *uncorrelated*, but X and Y are *dependent* because, if we know $X = 1$ occurred, then we also know that $Y = 0$ occurred.

Expected value of Functions of Two Random Variables

- In many situations, we are only interested in the expected value of a derived random variable $W = g(X, Y)$
- In this case, we do not need to derive the probability model for W since its expected value can be calculated directly from the joint PDF or joint PMF of X and Y
- If X and Y are discrete, then:

$$E[W] = E[g(X, Y)] = \sum_{x}\sum_{y} g(x, y) p_{X,Y}(x, y)$$

- If X and Y are continuous, then:

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

Example

- If the joint pdf $f_{XY}(x, y)$ is given as below find the expected value of $W = XY$:

$$f_{X,Y}(x, y) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise .} \end{cases}$$

Mean and Variance of a Sum of Two Random Variables

- Let X and Y be two general random variables, and Z is their sum:

$$Z = X + Y$$

- The mean of Z is:

$$\mu_Z = E[Z] = E[X+Y] = E[X] + E[Y] = \mu_X + \mu_Y$$

- The variance of Z is:

$$\begin{aligned} \sigma_Z^2 = \mathrm{Var}(X+Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y) \\ &= \sigma_X^2 + \sigma_Y^2 + 2\rho_{X,Y}\sigma_X\sigma_Y \end{aligned}$$

Mean and Variance of a Sum of Two Random Variables

Proof If X and Y are discrete RVs, the mean of their sum Z is:

$$\begin{aligned} E[Z] = E[X+Y] &= \sum_x \sum_y (x+y) p_{X,Y}(x,y) \\ &= \sum_x x \sum_y p_{X,Y}(x,y) + \sum_y y \sum_x p_{X,Y}(x,y) \\ &= \sum_x x p_X(x) + \sum_y y p_Y(y) \\ &= E[X] + E[Y] \end{aligned}$$

- The derivation is analogous if X and Y are continuous RVs with the sums above replaced by integrals.

Mean and Variance of a Sum of Two Random Variables

Proof The variance of Z is:

$$\begin{aligned}
\sigma_Z^2 &= E[(Z-\mu_Z)^2] = E[(X+Y-(\mu_X+\mu_Y))^2] \\
&= E[((X-\mu_X)+(Y-\mu_Y))^2] \\
&= E[(X-\mu_X)^2+(Y-\mu_Y)^2+2(X-\mu_X)(Y-\mu_Y)] \\
&= \text{Var}(X)+\text{Var}(Y)+2\text{Cov}(X,Y)
\end{aligned}$$

Note: If X and Y are uncorrelated (including independent), then:

$$\text{Var}(X+Y) = \text{Var}(X)+\text{Var}(Y)$$

Mean and Variance of a Sum of Multiple Random Variables

- For multiple general random variables, X_1, X_2, ..., X_N, the mean of their sum is:

$$E[\sum_{i=1}^{N} X_i] = \sum_{i=1}^{N} E[X_i]$$

- The variance of their sum is:

$$\text{Var}[\sum_{i=1}^{N} X_i] = \sum_{i=1}^{N} \text{Var}[X_i] + 2\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \text{Cov}[X_i, X_j]$$

Note: If X_1, X_2, ..., X_N are uncorrelated and independent, then:

$$\text{Var}[\sum_{i=1}^{N} X_i] = \sum_{i=1}^{N} \text{Var}[X_i]$$

Conditioning by an Event

- We consider the probability model for two or more random variables given the knowledge that some event B has occured
- **Conditional Joint Probability Mass Function**
  - For two discrete random variables X and Y and a conditioning event B ($\text{Prob}(B) > 0$) the conditional joint PMF of X and Y:

$$\begin{aligned} p_{X,Y|B}(x) &= \text{Prob}(X = x \text{ and } Y = y|B) \\ &= \begin{cases} \frac{p_{X,Y}(x,y)}{\text{Prob}(B)} & \text{if } (x,y) \in B \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

  - The conditional joint PMF is non-zero for a pair (x, y), if (x, y) is contained in the conditioning event B, and zero otherwise
  - Satisfies the axioms of probability:
    (1) Non-negativity: $p_{X,Y|B}(x, y) \geq 0$.
    (2) Normalization: $\sum_{(x,y)\in B}\sum p_{X,Y|B}(x, y) = 1$

Conditioning by an Event

- **Conditional Joint Probability Density Function**
  - For two continuous random variable X and Y and a conditioning event B ($\text{Prob}(B) > 0$) the conditional joint PDF of X and Y:

$$f_{X,Y|B}(x,y) = \begin{cases} \frac{f_{X,Y}(x,y)}{\text{Prob}(B)} & \text{if } (x,y) \in B \\ 0 & \text{otherwise.} \end{cases}$$

  - Satisfies the axioms of probability:
    (1) Non-negativity: $f_{X,Y|B}(x,y) \geq 0$.
    (2) Normalization: $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y|B}(x,y)dxdy = 1$

Conditional Expectations

- The **conditional expected value** of a function of two random variables $W = g(X, Y)$ given condition B is:

$$\begin{aligned} E[W|B] &= E[g(X,Y)|B] \\ &= \begin{cases} \sum\sum_{(x,y)\in B} g(x,y) p_{X,Y|B}(x,y) & \text{if } X, Y \text{ discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y|B}(x,y)\,dxdy & \text{if } X, Y \text{ continuous} \end{cases} \end{aligned}$$

- Then **conditional variance** of W is computed by:

$$\mathrm{Var}[W|B] = E[W^2|B] - (E[W|B])^2.$$

Conditioning by a Discrete Random Variable

- In some situations involving two random variables X and Y, we may have partial knowledge of the occurrence of one of the random variables, say we know $Y = y$ occurred. Then we can derive a conditional probability distribution for X given Y.
- For an event $Y = y$, where $\text{Prob}(y) > 0$, the conditional PMF of X given $Y = y$ is:

$$p_{X|Y}(x|y) = \text{Prob}(X = x|Y = y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

- Similarly, the conditional PMF of Y given $X = x$ is:

$$p_{Y|X}(y|x) = \text{Prob}(Y = y|X = x) = \frac{p_{X,Y}(x,y)}{p_X(x)}$$

- The joint PMF in terms of their conditional PMFs:

$$p_{X,Y}(x,y) = p_{X|Y}(x|y)p_Y(y) = p_{Y|X}(y|x)p_X(x)$$

Conditioning by a Discrete Random Variable

- The conditional PMF satisfies the axioms of probability:
  (1) Non-negativity: $p_{X|Y}(x|y) \geq 0$.
  (2) Normalization: $\sum_{x \in S_X} p_{X|Y}(x|y) = 1$
- For fixed $Y = y$, the distribution $p_{X|Y}(x|y)$ can be viewed as a renormalized slice of the joint PMF $p_{X,Y}(x,y)$ along $Y = y$
- The **conditional expected value of** X **given** $Y = y$:

$$E[X|Y = y] = \sum_x x p_{X|Y}(x|y)$$

- Let $g(X, Y)$ be a function of the two random variables. The **conditional expected value of** $g(X, Y)$ **given** $Y = y$ is:

$$E[g(X, Y)|Y = y] = \sum_x g(x, y) p_{X|Y}(x|y)$$

Conditioning by a Continuous Random Variable

- For any y where $f_Y(y) > 0$, the conditional PDF of X given $Y = y$:

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

- The joint PDF in terms of the conditional PDFs:

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

- The conditional PDF satisfies the axioms of probability:
  (1) Non-negativity: $f_{X|Y}(x|y) \geq 0$.
  (2) Normalization: $\int_{-\infty}^{\infty} f_{X|Y}(x|y)dx = 1$
- For fixed $Y = y$, the distribution $f_{X|Y}(x|y)$ can be viewed as a renormalized slice of the joint PDF $f_{X,Y}(x,y)$ along $Y = y$

Conditioning by a Continuous Random Variable

- The **conditional expected value of** X **given** $Y = y$ is:

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx$$

- Let $g(X, Y)$ be a function of the two continuous random variables X and Y. The **conditional expected value of** $g(X, Y)$ **given** $Y = y$ is:

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

Example (Continuous)

- The joint PDF of two random variables X and Y is:

$$f_{X,Y}(x,y) = \begin{cases} 1/\pi & \text{if } x^2 + y^2 \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

(1) Find the conditional PDF $f_{Y|X}(y|x)$.

(2) Compute and plot the conditional PDFs $f_{Y|X}(y|0)$ and $f_{Y|X}(y|1/2)$.

(3) What is the expected value $E[Y|X = 0]$?

(4) What is the conditional variance $\text{Var}[Y|X = 0]$?

Law of Iterated Expectations (Total Expectation Theorem)

- The conditional expected value $E[X|Y=y]$ is a function of random variable Y since it is conditioned on a specific value of $Y=y$
- The unconditional expected value for random variable X can be obtained from the conditional expected value via: $E[X] = E_Y[E_X[X|Y=y]]$

Proof If X and Y are discrete:

$$\begin{aligned} E_Y[E_X[X|Y=y]] &= \sum_{y\in S_Y} p_Y(y)\Big(\sum_{x\in S_X} x p_{X|Y}(x|y)\Big) = \sum_{x\in S_X}\sum_{y\in S_Y} x p_{X|Y}(x|y) p_Y(y) \\ &= \sum_{x\in S_X} x \sum_{y\in S_Y} p_{X,Y}(x,y) = \sum_{x\in S_X} x p_X(x) = E[X] \end{aligned}$$

- If X and Y are continuous:

$$\begin{aligned} E_Y[E_X[X|Y=y]] &= \int_{-\infty}^{\infty} f_Y(y)\Big(\int_{-\infty}^{\infty} x f_{X|Y}(x|y)dx\Big)dy = \int_{-\infty}^{\infty} x\Big(\int_{-\infty}^{\infty} f_{X|Y}(x|y) f_Y(y) dy\Big)dx \\ &= \int_{-\infty}^{\infty} x\Big(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy\Big)dx = \int_{-\infty}^{\infty} x f_X(x)dx = E[X] \end{aligned}$$

Example

- Break a stick of length L twice at uniformly chosen random points
- Let Y be the length of the stick at the first break point and X be the length of the stick at the second break point
- What is $E[X]$?

Solution Given that the stick has length $Y = y$ after the first breakage, the conditional expected value of the stick length at the second break point is:

$$E[X|Y = y] = \frac{y}{2}$$

Note that the above expected value is a function of y. The unconditional expected value of the stick length at the second break point is then:

$$E[X] = E[E[X|Y = y]] = E[\frac{y}{2}] = \frac{E[y]}{2} = \frac{L}{4}$$

Independent Random Variables

- Two **discrete** random variables X and Y are independent iff:

$$p_{X,Y}(x,y) = p_X(x)p_Y(y)$$

$$p_{X|Y}(x|y) = p_X(x)$$

$$p_{Y|X}(y|x) = p_Y(y)$$

- Two **continuous** random variables X and Y are independent iff:

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

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

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

- CDF relationship for independent random variables:

$$F_{X,Y}(x,y) = F_X(x)F_Y(y)$$

Independent Random Variables

- When X and Y are independent:
  - The conditional PMF or conditional PDF of X given $Y = y$ is the same for all $y \in S_Y$, and
  - The conditional PMF or the conditional PDF of Y given $X = x$ is the same for all $x \in S_X$.
- When X and Y are independent, learning that $Y = y$ provides no information about X, and learning that $X = x$ provides no information about Y.
- For independent random variables X and Y,
  (a) $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$
  (b) $R_{X,Y} = E[XY] = E[X]E[Y]$
  (c) $\text{Cov}[X, Y] = \rho_{X,Y} = 0$
  (d) $\text{Var}(aX + bY) = a^2\text{Var}(X) + b^2\text{Var}(Y)$

Example

(1) Random variables X and Y have the following joint PMF:

$p_{X,Y}(x,y)$	$y=0$	$y=1$	$y=2$
$x=0$	0.01	0	0
$x=1$	0.09	0.09	0
$x=2$	0	0	0.81

- Are X and Y independent?

(2) Random variables Q and G have the following joint PMF:

$p_{Q,G}(q,g)$	$g=0$	$g=1$	$g=2$	$g=3$
$q=0$	0.06	0.18	0.24	0.12
$q=1$	0.04	0.12	0.16	0.08

- Are Q and G independent?