Lecture 10 - Outline

1. Pairs of Random Variables
   - CDFs: joint, marginal, conditional, independence
   - PMFs: joint, marginal, conditional, independence

Chapter 4  Bivariate Random Variables

Two RVs $X$ and $Y$ defined on the same sample space $S$ are called bivariate RVs, Joint RVs, or random vectors (of length 2).

For each outcome $s \in S$ of the random experiment, $X(s)$ and $Y(s)$ are real numbers (a point in the plane, or a vector $(X(s), Y(s))$).

Examples:
- Time and duration of telephone calls.
  - Times may be originated at random
  - Calls at times when the rates are cheaper may be longer
- Work-in-process across two supply chains feeding an assembly operation
- Complex random signal
  - Amplitude and phase
  - Real and imaginary part
- Input and output of a noisy amplifier

Statistical Description of Bivariate RVs $X$ and $Y$

<table>
<thead>
<tr>
<th></th>
<th>CDF</th>
<th>PMF (discrete)</th>
<th>Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PDF (continuous)</td>
<td></td>
</tr>
<tr>
<td>Joint</td>
<td>$F_{X,Y}(x,y)$</td>
<td>$P_X(x,y)$, $F_X(x,y)$</td>
<td>$E[X</td>
</tr>
<tr>
<td>Conditional</td>
<td>$F_{X</td>
<td>Y}(x</td>
<td>y)$</td>
</tr>
</tbody>
</table>

4.1 Joint CDF of Bivariate RVs

Definition

$$F_{X,Y}(x,y) = P\{X \leq x \cap Y \leq y\} = P\{s \in S : X(s) \leq x \text{ and } Y(s) \leq y\} \equiv P[X \leq x, Y \leq y]$$

= Probability that random point lies in shaded area
Properties

• $0 \leq F_{X,Y}(x,y) \leq 1$ for all $x$ and $y$

• $F_{X,Y}(x,y)$ is nondecreasing in $X$ for fixed $Y$

• $F_{X,Y}(x,y)$ is nondecreasing in $Y$ for fixed $X$

• $\lim_{x\to\infty} F_{X,Y}(x,y) = F_X(\infty, y) = 1$

• $\lim_{y\to\infty} F_{X,Y}(x,y) = F_Y(x, \infty) = 0$

• For any $x,y$, $\lim_{x\to-\infty} F_{X,Y}(x,y) = F_X(-\infty, y) = 0$

• $\lim_{y\to-\infty} F_{X,Y}(x,y) = F_Y(x, -\infty) = 0$

• $F_{X,Y}(X,Y)$ is a right-continuous as a function of $X$ and right continuous as a function of $Y$

Probabilities from CDF

$$P\{X > x\} \cup \{Y > y\} = 1 - F_{X,Y}(x,y)$$

$$P\{X \leq x\} \cup \{Y \leq y\} = F_{X}(x) + F_{Y}(y) - F_{X,Y}(x,y)$$

Joint CDF of Independent RVs

Recall

Two events $A, B$ are said to be independent if and only if $P[A \cap B] = P[A]P[B]$


The RVs $X$ and $Y$ are said to be mutually independent if every event $A$ defined by $X$ is independent of every event $B$ defined by $Y$.

$A = \{X \leq x\}; \quad B = \{Y \leq y\}$

$P[A \cap B] = F_{X,Y}(x,y) = P[A]P[B] \quad \text{Independence}$

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

The joint CDF of independent RVs factors into a product of the marginals

Marginal CDFs

Individual CDFs of $X, Y$, called the marginal CDFs, can be obtained from $F_{X,Y}(x,y)$. For any $x, y$:

$$\lim_{y\to\infty} F_{X,Y}(x,y) = F_{X}(x)$$

$$\lim_{x\to\infty} F_{X,Y}(x,y) = F_{Y}(y)$$

If the RVs are independent, the joint CDF $F_{X,Y}(x,y)$ may be determined from the marginal CDFs $F_X(x)$ and $F_Y(y)$.

If the RVs are not independent, the joint CDF may not be determined from the marginal CDFs. Many different CDFs may have the same marginals! Many different surfaces $F_{X,Y}(x,y)$ may have the same limit values $F_{X,Y}(x,\infty)$ and $F_{X,Y}(\infty,y)$.
### 4.2 Joint PMF (Discrete Bivariate RVs)

#### Definition
For discrete RVs $X$ and $Y$ the joint PMF $P_{X,Y}(x, y) = P[X = x \text{ and } Y = y]$ describes a collection of point masses in the plane $\sum_x \sum_y P_{X,Y}(x, y) = 1$

#### Probabilities from Joint PMF
$P((X,Y) \in A] = \text{sum of all probability masses in the region } A$

This also holds if $A$ is a curve (including a straight line as a special case) — just sum up all the probability masses on the curve

#### Marginal PMFs
$P_X(x) = \sum_y P_{X,Y}(x, y)$
$P_Y(y) = \sum_x P_{X,Y}(x, y)$
The marginal PMFs of $X$ and $Y$ may be obtained from the joint PMF by projection (summing) in the $y$ and $x$ directions, respectively.

The joint PMF may not be determined from the marginals (unless the variables are independent)

#### Example
$PX(4) = \frac{3}{20}$
$PY(3) = \frac{9}{20}$

#### Calculation of Conditional PMF from Joint PMF
$P_{X|Y}(x|y) = \frac{P_{X,Y}(x, y)}{P_Y(y)}$

$P_Y(y) = \text{sum of all probability masses lying at } y = y_j$

$P_{X|Y}(x|y) = \text{ratio of the probability mass at } x \text{ to the total mass on the line } y = y_j$

#### Total probability theorem for conditional PMFs
Since the events $(Y = y_1), (Y = y_2), ...$ are mutually disjoint and collectively exhaustive, forming a partition of the sample space $S$, $P_X(x) = \sum_{y_j \in S_y} P_{X|Y}(x|y_j)P_Y(y_j)$

#### Bayes’ Rule applied to PMFs
Bayes’ rule: $P(A|B) = \frac{P(B|A) P(A)}{P(B)}$

$P_{X|Y}(x|y) = \frac{P_{X,Y}(x, y)}{P_Y(y)}$

$P_X(x) = \sum_{y_j \in S_y} P_{X|Y}(x|y_j)P_Y(y_j)$
Independence

$X, Y$ are said to be independent if and only if

$$P_{X,Y}(x,y) = P_X(x)P_Y(y)$$

i.e., joint PMF factors into product of marginals. This factorization also follows from factorization of the joint CDF

Or

$X, Y$ are said to be independent if and only if

$$P_{X|Y}(x|y) = P_X(x)$$

and

$$P_{Y|X}(y|x) = P_Y(y)$$

Examples of Dependent and Independent RVs:

Experiment 1: roll two 6-sided dice

$X(s) = \text{value of first die}$

$Y(s) = 7 - \text{value of first die}$

Experiment 2: roll two 6-sided dice

$X(s) = \text{value of first die}$

$Y(s) = \text{value of second die}$

Both experiments have the same range of outcomes $X(s), Y(s)$ take values in $1...6$

$P_X(n) = 1/6$ for $n = 1, ..., 6$ in both cases!

$P_Y(n) = 1/6$ for $n = 1, ..., 6$ also in both!!!

However, the two experiments are clearly different!

- In Experiment 1, when $X = n$, $Y = 7 - n$ Æ very dependent!
- In Experiment 2, when $X = n$, $Y$ is totally random Æ very independent!
- Knowing $P_X(), P_Y()$ is insufficient to judge dependence or independence!

Example: Embedded RVs

$Y$ is a Poisson RV with parameter $\lambda$.

Conditioned on $Y = n$, $X$ is a binomial RV with parameters $(n, p)$

Determine the Marginal PMF of $X$

$$P_{X|Y}(k|n) = \binom{n}{k} p^k (1 - p)^{n-k}, \ k = 0, ..., n$$

• Conditioned on $Y = n$, $X$ has values from 0 to $n$.
• Conditioned on $X = k$, $Y$ has values from $k$ to $\infty$.
• Since $X \leq Y$ the joint PMF of $X$ and $Y$ is nonzero on a triangular region

Marginal PMF of $X$

$$P_X(3) = \sum_{n=3}^{\infty} \binom{n}{3} p^3 (1 - p)^{n-3} \frac{\lambda^n e^{-\lambda}}{n!}$$

Application: Random Routing

Messages going into a router are sent along path 1 with probability $p$, path 2 with probability $1 - p$.

If one receives a Poisson number of messages $\mu$ with rate $\lambda$:

- the number of messages sent on path 1 is Poisson with rate $p\mu$,
- the number of messages sent on path 2 is Poisson with rate $(1 - p)\mu$.

Application: Geiger counter with finite efficiency

$\alpha$-particles are counted in a Geiger counter

Each particle is detected with probability $p$ and not detected with probability $1 - p$. Detections are independent of each other

If $\mu \alpha$-particles are emitted (Poisson number), the number detected (the count in the Geiger counter) is a binomial RV with parameters $(\mu, p)$

Detector

Bayes' rule

$$P_{Y|X}(y|3) = \frac{P_{X,Y}(3,y)}{P_X(3)}$$

b) Assume $X$ was observed and found to be $X = 3$.

Compute $P_{Y|X}(3,3)$

Conditioned on $X = 3$, $Y = 3$ Æ Poisson with parameter $\lambda(1 - p)$. This is called a shifted Poisson.