EEL 5544 Noise in Linear Systems Lecture 14

JOINT PROBABILITY MASS FUNCTIONS

DEFN If X and Y are discrete random variables, the *joint probability mass function* for X and Y is

$$P_{X,Y}(x,y) = P(X = x, Y = y).$$

Example: Flipping 2 coins (continued)

For the previous example of flipping two coins, we previously found all the values of the joint pmf, which is given by

$$P_{X,Y} = \begin{cases} 1/4, & (x,y) \in \{(0,1),(1,0)\} \\ 1/2, & (x,y) = (1,1) \\ 0, & \text{otherwise} \end{cases}$$

STATISTICAL INDEPENDENCE

- Recall that if $A, B \in \mathcal{F}$, then A and B are s.i. iff

$$P(A \cap B) = P(A)P(B).$$

- How should we generalize this concept to random variables?
- Would at least like

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = P(a_1 < X \leq a_2)P(b_1 < Y \leq b_2) \tag{1}$$

  for all $a_1 \leq a_2$, $b_1 \leq b_2$

- Writing (1) in terms of the dist. fcns yields

$$\begin{aligned} F_{XY}(a_2,b_2) - F_{XY}(a_1,b_2) - F_{XY}(a_2,b_1) + F_{XY}(a_1,b_1) \\ = [F_X(a_2) - F_X(a_1)]\,[F_Y(b_2) - F_Y(b_1)] \\ = F_X(a_2)F_Y(b_2) - F_X(a_1)F_Y(b_2) - F_X(a_2)F_Y(b_1) + F_X(a_1)F_Y(b_1) \end{aligned} \tag{2}$$

- Comparing the LHS and RHS of (2), a sufficient condition is

DEFN Random variables X and Y are s.i. iff

- What does this imply about the density functions?

$$\begin{aligned} f_{XY}(x,y) &= \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y) \\ &= \left[\frac{\partial}{\partial x} F_X(x)\right]\left[\frac{\partial}{\partial y} F_Y(y)\right] \\ &= f_X(x) f_Y(y) \end{aligned}$$

Examples: Are X and Y s.i. in the previous two examples?

CONDITIONING WITH MULTIPLE RVS

- Consider our previous definition of conditional probability.
- Given events A, B, with $P(B) > 0$, then

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

- To extend our concepts of conditional prob. to multiple RVS, consider the events

$$\begin{aligned} A &= \{X \leq x\} \\ B &= \{Y \in C\}, \end{aligned}$$

  where $P(B) \neq 0$.
- Then

$$P(A|B) = P(X \leq x | Y \in C) = \frac{P(X \leq x, Y \in C)}{P(Y \in C)}$$

- We will write this as a conditional distribution function

$$F_X(x|Y \in C) = \frac{P(X \leq x, Y \in C)}{P(Y \in C)},$$

which also admits a conditional density function

$$f_X(x|Y \in C) = \frac{d}{dx} F_X(x|Y \in C).$$

SPECIAL CASES OF COND. DISTS AND DENSITIES

1. **Point conditioning** (generally most useful case)

First, suppose $P(Y = y) > 0$.Then

$$F_{X|Y}(x|y) = F_X(x|Y = y) = \frac{P(X \leq x, Y = y)}{P(Y = y)}.$$

In particular, if Y is a discrete RV and we condition on $Y = y_k$, then

$$F_{X|Y}(x|y_k) = \frac{P(X \leq x, Y = y_k)}{P_Y(y_k)}.$$

If both X and Y are discrete RVs, then

$$P_{X|Y}(x_j|y_k) = \frac{P_{X,Y}(x_j, y_k)}{P_Y(y_k)}.$$

If X and Y are s.i., then

$$P_{X|Y}(x_j|y_k) =$$

as expected.

We often use point conditioning with continuous RVs for which

$$\begin{aligned} P(Y = y) &= 0 \\ P(X \leq x, Y = y) &= 0 \end{aligned}$$

In this case, the conditional distribution function is defined as a limit:

$$
\begin{aligned}
F_{X|Y}(x|Y=y) &= \lim_{h\to 0} F_X(x|y \le Y \le y+h) \\
&= \lim_{h\to 0} \frac{P(X \le x, y \le Y \le Y+h)}{P(y \le Y \le y+h)} \\
&= \lim_{h\to 0} \frac{\int_{-\infty}^{x} \int_{y}^{y+h} f_{XY}(u,t)\,dt\,du}{\int_{y}^{y+h} f_Y(v)\,dv}
\end{aligned}
$$

We can apply the mean-value theorem to this expression to get

$$
F_{X|Y}(x|Y=y) = \lim_{h\to 0} \frac{\int_{-\infty}^{x} h f_{XY}(u,t')\,du}{h f_Y(t'')},
$$

for some $y \le t' \le y+h$ and some $y \le t'' \le y+h$.

Then in the limit,

$$
F_{X|Y}(x|Y=y) =
$$

More often, we find the conditional density function, which has a nicer form:

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

2. **Non-point conditioning**

   Another common conditional distribution is given by

$$
F_X(x|Y \le y) =
$$

   when $F_Y(y) > 0$.

<u>CONDITIONAL JOINT PROBABILITIES ARE ALSO POSSIBLE:</u>

EX:

$$
F_{XY}(x,y|x_1 < X \le x_2) = \frac{P\left(X \le x, Y \le y, x_1 < X \le x_2\right)}{P(x_1 < X \le x_2)}
$$

The denominator can be expressed directly in terms of the marginal distribution function for X

Consider the regions of X in the numerator:
$X \leq x \cap x_1 < X \leq x_2$
Use number line in x to determine regions:

- $x \leq x_1$

- $x_1 < x \leq x_2$

- $x > x_2$

$$F_{XY}(x, y|x_1 < X \leq x_2) = \begin{cases} 0, & x \leq x_1 \\ \frac{P(x_1 < X \leq x, Y \leq y)}{P(x_1 < X \leq x_2)}, & x_1 < x \leq x_2 \\ \frac{P(x_1 < X \leq x_2, Y \leq y)}{P(x_1 < X \leq x_2)}, & x > x_2 \end{cases}$$

This conditional joint distribution function can be written as:

$$F_{XY}(x, y|x_1 \leq X \leq x_2) = \begin{cases} 0, & x \leq x_1 \\ \frac{F_{XY}(x,y) - F_{XY}(x_1,y)}{F_X(x_2) - F_X(x_1)}, & x_1 < x \leq x_2 \\ \frac{F_{XY}(x_2,y) - F_{XY}(x_1,y)}{F_X(x_2) - F_X(x_1)}, & x > x_2 \end{cases}$$

<u>TOTAL PROBABILITY AND BAYES' THEOREM</u>

(for point conditioning)

The conditional density for Y given $X = x$ is

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

$$\Rightarrow f_{XY}(x, y) = \tag{3}$$

(4)

We can find the marginal density of Y as

$$f_Y(y) = \tag{5}$$

By applying (3) to (5), we get

DEFN The *Total Probability Law for point conditioning*:

$$f_Y(y) = \tag{7}$$

To calculate $f_{X|Y}(x|y)$, we can apply (3) to the definition for $f_{X|Y}(x|y)$ to get

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

Then applying (7) yields

DEFN *Bayes' Rule for point conditioning*:

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