Notes: Joint Probability and Independence for Discrete RV's

CS 3130: Probability and Statistics for Engineers

October 19, 2017

Sometimes we are interested in looking at the probabilities of multiple outcomes simultaneously. As we will see, this is particularly interesting if the multiple variables have some relationship or dependency on each other. Examples:

- Repeated experiments (e.g., flipping a coin multiple times)
- Collecting multiple variables (e.g., temperature, barometric pressure, wind speed)
- Studying the relationship between two variables (e.g., is smoking related to incidents of lung cancer?)
- Repeated measurements of the same quantity (with error)

Joint Probability Mass Functions Let X and Y be two discrete random variables. Remember we can describe events (subsets of the sample space) by the notation $\{X = a\}$, meaning "the set of all outcomes that result in X being equal to a". Earlier when we talked about probability, we discussed the probability of intersections of events A, B, with notation $P(A \cap B)$. Using the two random variables X and Y, we can define two events $X = a$ and $Y = b$ and talk about their intersection $\{X = a\} \cap \{Y = b\}$. Remember, in English this event is "X is equal to a *and* Y is equal to b". The **joint probability mass function** for X and Y is defined as

$$p_{X,Y}(a, b) = P(X = a, Y = b) = P(\{X = a\} \cap \{Y = b\})$$

The joint probability is a true probability function, and thus must satisfy the usual rules:

$$\begin{aligned} p_{X,Y}(a, b) &\geq 0, && \text{for all } a, b \\ \sum_i \sum_j p_{X,Y}(a_i, b_j) &= 1, && \text{where } a_i, b_j \text{ are all possible outcomes for } X \text{ and } Y \end{aligned}$$

<u>Example:</u> You are sending a binary message over a wireless network. Each bit sent has some probability of being corrupted. Let S be a binary random variable representing a sent bit and R be a binary random variable representing the corresponding received bit. The joint pmf can be specified by four probabilities, like so:

		R = 0	R = 1
S	0	0.45	0.08
	1	0.06	0.41

Table of $P(S = a, R = b)$.

This table should be read as follows: $P(S = 0, R = 0) = 0.45$ means "the probability of sending a 0 and receiving a 0 is 45%", etc.

Example: Let X and Y be the outcome of two dice rolls. Then the pmf $p_{X,Y}(a, b) = \frac{1}{36}$ for all possible combinations of a and b. The table is a 6×6 table with all entries equalling $\frac{1}{36}$.

From the joint pmf $p_{X,Y}$, we can recover the pmf's for X and Y, p_X and p_Y. First, notice that the events $Y = b_i$ for each possible outcome b_i are all *disjoint* events. Therefore, we can use the law of total probability:

$$P(X = a) = \sum_i P(\{X = a\} \cap \{Y = b_i\}) = \sum_i p_{X,Y}(a, b_i)$$

The probability $P(X = a)$ is called the **marginal** probability of X, or we say we got $P(X = a)$ by **marginalizing** over the random variable Y.

Example: In the R, S example above, we can compute the marginal probabilities of R by summing down the columns (keeping R fixed, and summing over different S outcomes). Likewise, we can compute the marginal probabilities of S by summing across the rows (keeping S fixed, and summing over different R outcomes). We get the following:

$$P(R = 0) = 0.51, \quad P(R = 1) = 0.49, \quad P(S = 0) = 0.53, \quad P(S = 1) = 0.47$$

In-class Exercise: Say you flip two fair coins. Let H be the total number of heads and B be the binary number the two coins represent (if heads is a 1 and tails is a 0). Write down the table for the pmf $p_{H,B}$ and the marginal probabilities p_H, p_B.

Conditional Probability. We've seen joint probabilities are just the same as using the intersection of events. Therefore, our definition of conditional probability can also be rephrased in terms of the joint pmf of two random variables X and Y:

$$\begin{aligned} P(X = a \,|\, Y = b) &= \frac{P(\{X = a\} \cap \{Y = b\})}{P(Y = b)} \\ &= \frac{p_{X,Y}(a, b)}{p_Y(b)} \end{aligned}$$

In-class Exercise: In the "send/receive" example above, compute the probability of receiving a 1, given that a 1 was sent. Also, compute the probability that a 0 was sent, given that you received a 0.

In-class Exercise: In the coin flip example above, compute the conditional probabilities for H given B and also for B given H.

Independence. Remember our equivalent rules for two events A and B to be independent:

$$P(A, B) = P(A)P(B), \quad \text{or } P(A|B) = P(A), \quad \text{or } P(B|A) = P(B).$$

These can now all be written for random variables. We say two random variables X, Y are **independent** if for all possible outcomes $X = a$ and $Y = b$,

$$p_{X,Y}(a, b) = p_X(a)\, p_Y(b).$$

Notice this is equivalent to our first definition of independence for events, with the added rule that it must hold for *all* possible outcomes for X and Y. In other words, X and Y are independent of each other if all the events defined by X are independent from all the events defined by Y.

The independence definition above says only that events of the form $\{X = a\}$ and $\{Y = b\}$ are independent. However, we can see that this implies that *any* two events $A = \{a_i\}$ and $B = \{b_j\}$ defined using X and Y are also independent:

$$
\begin{aligned}
P(\{X \in A\} \cap \{Y \in B\}) &= \sum_i P(\{X = a_i\} \cap \{Y \in B\}) && \text{applying total probability to } A \\
&= \sum_i \sum_j P(\{X = a_i\} \cap \{Y = b_j\}) && \text{applying total probability to } B \\
&= \sum_i \sum_j P(X = a_i) P(Y = b_j) && \text{using definition of independence} \\
&= \sum_i P(X = a_i) \sum_j P(Y = b_j) && \text{first term doesn't depend on } j \\
&= \sum_i P(X = a_i)\, P(Y \in B) && \text{summing over } j \\
&= P(X \in A)\, P(Y \in B) && \text{summing over } i
\end{aligned}
$$

In-class Exercise: Determine for the "send/receive" and coin flip examples above whether the two random variables are independent of each other.

In-class Exercise: Let X and Y be two dice rolls. Verify that $X \in \{1, 3, 4\}$ is independent of $Y \in \{1, 2\}$.