Joint PMF and Conditioning of Multiple Random Variables

John N. Tsitsiklis

November 17, 2017

Joint PMF of Multiple Random Variables

Definition 1 (Joint PMF)**.** Let X and Y be two random variables associated with the same experiment. Then the *joint PMF* $p_{X,y}$ of X and Y is defined by

$$p_{X,Y}(x, y) = \mathbf{P}(X = x, Y = y)$$

Definition 2 (Marginal PMF)**.** The *marginal PMF*s of X and Y can be obtained from the joint PMF, using the formulas

$$p_X(x) = \sum_y p_{X,Y}(x, y), \quad p_Y(y) = \sum_x p_{X,Y}(x, y)$$

Remark 1. If A is the set of all pairs that have a certain property, then

$$\mathbf{P}(A) = \sum_{(x,y) \in A} p_{X,Y}(x, y)$$

Remark 2. A function $g(X, Y)$ of X and Y defines another random variable, and

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

Remark 3. If $g(X, Y) = aX + bY + c$, we have

$$\mathbf{E}[aX + bY + c] = a\mathbf{E}[X] + b\mathbf{E}[Y] + c$$

Example 1 (The Hat Problem)**.** Suppose that n people throw their hats in a box and then each picks one hat at random. What is the expected value of X, the number of people that get back their own hat?

Solution For the ith person, let's define the random varible X_i such that

$$X_i = \begin{cases} 1 & \text{if the person picks his own hat,} \\ 0 & \text{otherwise.} \end{cases}$$

Since $\mathbf{P}(X = 1) = \frac{1}{n}$, we have

$$\mathbf{E}[X_i] = 1 \cdot \frac{1}{n} + 0 \cdot (1 - \frac{1}{n}) = \frac{1}{n}$$

We now have

$$X = X_1 + X_2 + \cdots + X_n$$

so that

$$\mathbf{E}[X] = \mathbf{E}[X_1] + \mathbf{E}[X_2] + \cdots + \mathbf{E}[X_n] = n \cdot \frac{1}{n} = 1$$

■

Conditioning

Conditioning a Random Varible on an Event

Definition 3 (Conditional PMF)**.** The *Conditional PMF* of a random variable X, conditioned on a particular event A with $\mathbf{P}(A) > 0$, is defined by

$$p_{X|A}(x) = \mathbf{P}(X = x \mid A) = \frac{\mathbf{P}(\{X = x\} \cap A)}{\mathbf{P}(A)}$$

Remark 4. The events $\{X = x\} \cap A$ are disjoint for different values of x, their union is A, hence

$$\mathbf{P}(A) = \sum_x \mathbf{P}(\{X = x\} \cap A)$$

Combining the above formulas we have

$$\sum_x \mathbf{P}(X \mid A)\, x = 1$$

So $p_{X|A}$ is a legitimate PMF.

Remark 5. If $A_1, \ldots, A_n$ are disjoint events that form a partition of the sample space, with $\mathbf{P}(A_i) > 0$ for all i, then

$$p_X(x) = \sum_{i=1}^{n} \mathbf{P}(A_i)\, p_{X|A_i}(x)$$

Also for any event B, with $\mathbf{P}(A_i \cap B) > 0$ for all i, we have

$$p_{X|B}(x) = \sum_{i=1}^{n} \mathbf{P}(A_i \mid B)\, p_{X|A_i \cap B}(x)$$

Conditioning one Random Variable on Another

Definition 4. Let X and Y be two random variables associated with the same experiment. The *Conditional PMF* $p_{X|Y}$ of X given Y is defined by

$$p_{X|Y}(x|y) = \mathbf{P}(X = x \mid Y = y)$$

Remark 6.

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

Conditional Expectation

Let X and Y be random variables associated with the same experiment

- The *conditional expectation* of X given an event A with $\mathbf{P}(A) > 0$, is defined by
$$\mathbf{E}[X \mid A] = \sum_x x p_{X|A}(x)$$
For a function $g(X)$, we have
$$\mathbf{E}[g(X) \mid A] = \sum_x g(x) p_{X|A}(x)$$
- The conditional expectation of X given a value y of Y is defined by
$$\mathbf{E}[X \mid Y = y] = \sum_x x p_{X|Y}(x|y)$$
- If $A_1, \ldots, A_n$ are disjoint events that form a partition of the sample space, with $\mathbf{P}(A_i) > 0$ for all i, then
$$\mathbf{E}[X] = \sum_{i=1}^{n} \mathbf{P}(A_i) \mathbf{E}[X \mid A_i]$$
Also for any event B, with $\mathbf{P}(A_i \cap B) > 0$ for all i, we have
$$\mathbf{E}[X \mid B] = \sum_{i=1}^{n} \mathbf{P}(A_i \mid B) \mathbf{E}[X \mid A_i \cap B]$$
- We have
$$\mathbf{E}[X] = \sum_y p_Y(y) \mathbf{E}[X \mid Y = y]$$

Example 2 (D. Bernoulli's problem of joint lives)**.** Consider $2m$ persons forming m couples who live together at a given time. Suppose that at some later time the probability of each person being alive is p, independent of other persons. At that later time, let A be the number of persons that are alive and let S be the number of couples in which both partners are alive. For any survivor number a, find $\mathbf{E}[S \mid A = a]$

Solution Let X_j be the indicator of the event that the jth couple survives. Then

$$\begin{aligned}
\mathbf{E}[S \mid A = a] &= \mathbf{E}\left[\sum_{j=1}^{m} X_j\right] \\
&= m\mathbf{E}[X_1] \\
&= m\mathbf{P}(X_1 = 1)
\end{aligned}$$

because the chance of survival is the same for every couple. Now we can choose the a survivors in $\binom{2m}{a}$ ways, and the number of these in which the first couple is remain alive is $\binom{2m-2}{a-2}$. Since all outcomes are equally likely we have

$$\mathbf{P}(X_1 = 1) = \frac{\binom{2m-2}{a-2}}{\binom{2m}{a}} = \frac{a(a-1)}{2m(2m-1)}$$

Hence we have

$$\mathbf{E}[S \mid A = a] = m \cdot \frac{a(a-1)}{2m(2m-1)} = \frac{a(a-1)}{2(2m-1)}$$

An alternative solution:

Let X_i be the random variable taking the value 1 or 0 depending on whether the first partner of the ith couple has survived or not. Let Y_i be the corresponding random variable for the second partner of the ith couple. Then, we have $S = \sum_{i=1}^{m} X_i Y_i$, and by using the total expectation theorem,

$$\begin{aligned}
\mathbf{E}[S \mid A = a] &= \sum_{i=1}^{m} \mathbf{E}[X_i Y_i \mid A = a] \\
&= m\mathbf{E}[X_1 Y_1 \mid A = a] \\
&= m\mathbf{P}(X_1 Y_1 = 1 \mid A = a)
\end{aligned}$$

Also

$$\mathbf{P}(X_1 Y_1 = 1 \mid A = a) = \mathbf{P}(Y_1 = 1 \mid X_1 = 1 \cap A = a) \cdot \mathbf{P}(X_1 = 1 \mid A = a)$$

We have

$$\mathbf{P}(Y_1 = 1 \mid X_1 = 1 \cap A = a) = \frac{a-1}{2m-1}, \quad \mathbf{P}(X_1 = 1 \mid A = a) = \frac{a}{2m}$$

Thus

$$\begin{aligned}
\mathbf{E}[S \mid A = a] &= m \cdot \frac{a-1}{2m-1} \cdot \frac{a}{2m} \\
&= \frac{a(a-1)}{2(2m-1)}
\end{aligned}$$

Note that $\mathbf{E}[S \mid A = a]$ does not depend on p. ■

Example 3. A coin that has probability of heads equal to p is tossed successively and independently until a head comes twice in a row or a tail comes twice in a row. Find the expected value of the number of tosses.

Solution Let X be number of tosses until the game is over and let H_k (or T_k) be the event that a head (or a tail, respectively) comes at the kth toss. Since H_1 and T_1 form a partition of the sample space, by total expectation theorem, we have

$$\mathbf{E}[X] = p\mathbf{E}[X \mid H_1] + (1-p)\mathbf{E}[X \mid T_1]$$

Using the total expectation theorem again, we have

$$\begin{aligned}\mathbf{E}[X \mid H_1] &= p\mathbf{E}[X \mid H_1 \cap H_2] + (1-p)\mathbf{E}[X \mid H_1 \cap T_2] \\ &= 2p + (1-p)(1 + \mathbf{E}[X \mid T_1])\end{aligned}$$

Since $\mathbf{E}[X \mid H_1 \cap H_2] = 1$ as game ends. Also $\mathbf{E}[X \mid H_1 \cap T_2] = 1 + \mathbf{E}[X \mid T_1]$ since the last toss only matters and the game proceeds. Similarly we have

$$\mathbf{E}[X \mid T_1] = 2(1-p) + p(1 + \mathbf{E}[X \mid H_1])$$

Combining the above equations we have

$$\mathbf{E}[X \mid H_1] = \frac{2 + (1-p)^2}{1 - p(1-p)}$$

And

$$\mathbf{E}[X \mid T_1] = \frac{2 + p^2}{1 - p(1-p)}$$

Thus

$$\mathbf{E}[X] = \frac{2 + p(1-p)}{1 - p(1-p)}$$

■