Hi all,

I would like to further clarify the the notations I used in the slides and resolve the confusion. The major resource of confusion, as I said, is because I overload the notation $\cap$ and did tell the difference explicitly when I explained the sum rule.

To resolve the confusion, from now on, let's only use $P(X,Y)$ for joint probability of two random variables X and Y, and reserve $\cup$ and $\cap$ for set operation.

Recall that we use random variable to describe an event in an experiment.

First, about independence. For two random variable X and Y, $P(X,Y) = P(X) \cdot P(Y)$ then X and Y are independent. This applies to any two random variables and they may share the sample space, as long as they describe two different experiments. For example, X describes tossing a dice at 10 AM and Y describes tossing a dice at 10:01 AM.

To understand the sum rule, let's consider two events X and Y that are from the same sample space and describe the same experiment. For example, in the experiment of tossing a dice, we use $X = \{1\}$ represents the event that 1 appears up $Y = \{2\}$ represents the event that 2 appears up Since X and Y are mutually exclusive (cannot happen in the same time), then we have

$$P(X \cup Y) = P(X) + P(Y)$$

where $\cap$ is the union operation, such that $X \cup Y = \{1, 2\}$.

Can we define $P(X,Y)$? The short answer is no. Loosely speaking, this is because X and Y describe the same event from different aspects. In other words, if we know the outcome of X, in the meantime we will also know the answer of Y — there is no randomness.

If we consider a little more complicated case, where $X = \{1, 3\}$ represents the event that 1 or 3 appears up $Y = \{2, 3\}$ represents the event that 2 or 3 appears up where X and Y are not mutually exclusive. In this case, since $X \cup Y = \{1, 2, 3\}$, therefore we have

$$P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$$

with $X \cap Y = \{3\}$