Conditional distributions (discrete case)

Definition:

Let X and Y be two discrete random variables on sample space Ω. X takes values $x_1, x_2, \ldots$ and Y takes values $y_1, y_2, \ldots$

The conditional probability mass function $p_{X|Y}(x \mid y)$ of X given $Y = y$ is defined by

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

provided $p_Y(y) > 0$.

Example (from last lecture)

Suppose the joint p.m.f. of X and Y is given by the table

X \ Y	10	20	30
0	$\frac{2}{18}$	$\frac{2}{18}$	0
1	$\frac{3}{18}$	$\frac{4}{18}$	$\frac{3}{18}$
2	0	0	$\frac{4}{18}$

The conditional p.m.f. of X given $Y=y$:

Conditioned on $Y=10$,

$$P_{X|Y}(0|10) = \frac{P_{X,Y}(0,10)}{P_Y(10)} = \frac{\frac{2}{18}}{\frac{5}{18}} = \frac{2}{5}$$

$$P_{X|Y}(1|10) = \frac{P_{X,Y}(1,10)}{P_Y(10)} = \frac{\frac{3}{18}}{\frac{5}{18}} = \frac{3}{5}$$

$$P_{X|Y}(2|10) = \frac{P_{X,Y}(2,10)}{P_Y(10)} = 0.$$

Conditioned on $Y=20$,

$P_{X|Y}(0|20) = \frac{\frac{2}{18}}{\frac{6}{18}} = \frac{1}{3}$, $P_{X|Y}(1|20) = \frac{\frac{4}{18}}{\frac{6}{18}} = \frac{2}{3}$. $P_{X|Y}(2|20)=0$.

Conditioned on $Y=30$,

$P_{X|Y}(0|30) = 0$, $P_{X|Y}(1|30) = \frac{\frac{3}{18}}{\frac{7}{18}} = \frac{3}{7}$, $P_{X|Y}(2|30) = \frac{4}{7}$.

Basic properties of conditional distribution

① Suppose $P(Y=y)>0$. The $P_{X|Y}(\cdot \mid y)$ is a probability mass function.

Because $\sum_i P_{X|Y}(x_i|y) = \sum_i P_{X,Y}(x_i, y) \big/ P_Y(y)$,

and $P_Y(y) = \sum_i P_{X,Y}(x_i, y)$, therefore

$$\sum_i P_{X|Y}(x_i|y) = 1.$$

Notation: since $P_{X|Y}(\cdot \mid y)$ is a p.m.f., the notation

$$P_{X|Y}(\cdot \mid y) \sim \left(\begin{array}{c}\text{name of a} \\ \text{distribution}\end{array}\right)$$

means that the p.m.f. $P_{X|Y}(\cdot \mid y)$ is the same as the p.m.f. of the distribution described on the right hand side.

(This notation is not standard, it really means that the conditional distribution of X given $Y=y$ is as described).

② If we know P_Y and $P_{X|Y}(\cdot|\cdot)$, then the p.m.f. of X can be calculated as

$$P_X(x) = \sum_{y_i : P_Y(y_i)>0} P_{X|Y}(x|y_i)\, P_Y(y_i).$$

Example 1

This sentence means $N \sim \text{Poisson}(\lambda)$

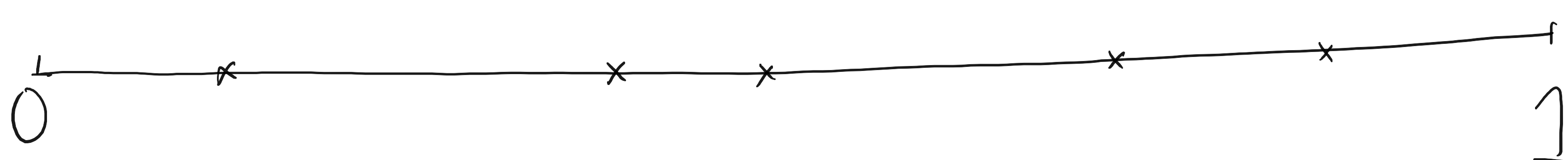
The conditional distribution of X given $N=k$ is $\text{Binomial}(p,k)$, for $k \in \{0,1,2,\ldots\}$

Suppose X has binomial distribution with parameter p and N, where N Poisson distribution with parameter (λ).

What is the (marginal) distribution of X?

Intuitive picture of what the dist. of X should be:

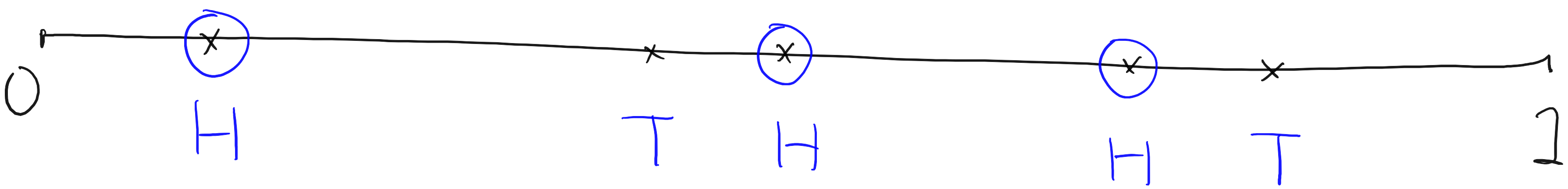
First sample N, N counts the number of arrivals in $[0,1]$, where the waiting time between arrivals is $\text{Exponential}(\lambda)$



$k=5$ in this picture

Conditioned on $N=k$, $\text{Binomial}(p,k)$ counts # of successes if we perform k independent $\text{Bernoulli}(p)$ trials.

You can think of it as flipping k coins with head prob. p, and keep the j^{th} point if the j-th coin flip is H



The R.V. X counts number of points that are kept after this Bernoulli selection scheme.

X is still a Poisson R.V., with parameter λp.

(this is called a thinning of Poisson process, you can figure out the parameter λp by considering the mean of X)

- The rigorous calculation using formula in basic property ② above can be found on page 48-49 in Pinsky-Karlin.

$$N \sim \text{Binomial}(q, m)$$

Example 2

The conditional distribution of X given $N = k$ is $\text{Binomial}(p, k)$ that is

$$p_{X|N}(j \mid k) = \binom{k}{j} p^j (1-p)^{k-j}, \quad j \in \{0, 1, \cdots, k\}.$$

Let X have binomial distribution with parameter p and N, where N has a binomial distribution with parameters q and m. What is the marginal distribution of X?

Intuitive explanation:

We first sample N. Since it's $\text{Binomial}(q, m)$, it counts successes in m i.i.d. Bernoulli (q) trials

H prob. q

N = # of H in m flips

H H T T H T T H

Now conditioned on $N = k$, to sample X which has conditional distribution $\text{Binomial}(p, k)$ we take another k i.i.d. Bernoulli (p) trials and X count the number of successes.

Effectively it's the same as performing two independent rounds of m Bernoulli trials,

$X_1, X_2, \dots, X_m$ success parameter p

$Y_1, Y_2, \dots, Y_m$ Suceess parameter q

and point j is counted as a final success if in both round it is a success. X counts # of final successes.

The probability for both X_i, Y_i to be success is

$$\mathbb{P}(X_i Y_i = 1) = P(X_i = 1) P(Y_i = 1) = pq.$$

Therefore $X \sim \text{Binomial}(m, pq)$.

Calculation:

By description X can only take value in $\{0, 1, 2, \dots, m\}$

$$P_X(j) = \sum_{k=0}^{m} P_{X|N}(j|k)\, P_N(k)$$

note that $P_{X|N}(j|k) = 0$ if $j > k$ by def. of binomial dist.

plug in p.m.f.
$$= \sum_{k=j}^{m} \frac{\cancel{k!}}{(k-j)!\, j!} p^j (1-p)^{k-j} \frac{m!}{(m-k)!\, \cancel{k!}} q^k (1-q)^{m-k}$$

change summation variable to $i = k-j$
$$= \sum_{i=0}^{m-j} \frac{1}{i!\, j!} p^j (1-p)^i \frac{m!}{(m-i-j)!} q^{i+j} (1-q)^{m-j-i}$$

$$= \frac{m!}{j!} (pq)^j \sum_{i=0}^{m-j} \frac{1}{i!\,(m-i-j)!} \left((1-p)q\right)^i (1-q)^{m-j-i}$$

$$= \binom{m}{j} (pq)^j \sum_{i=0}^{m-j} \binom{m-j}{i} (q-pq)^i (1-q)^{m-j-i}$$

$$= \binom{m}{j} (pq)^j (1-pq)^{m-j}$$

in the last step we used the binomial expansion formula.