Lecture 6:Important Discrete Distributions

February 11, 2019

The PMF P_X of a random variable tells us pretty much everything we need to know about X. Often we define a random variable by giving its PMF, without making reference to the underlying sample space on which the random variable is defined. You have already seen all the distributions described below in another context, with the exception of the Poisson distribution.

1 Bernoulli distribution

There is not one Bernoulli distribution–you get a different PMF for every value of a parameter p, where $0 \leq p \leq 1$. A random variable X with this distribution satisfies

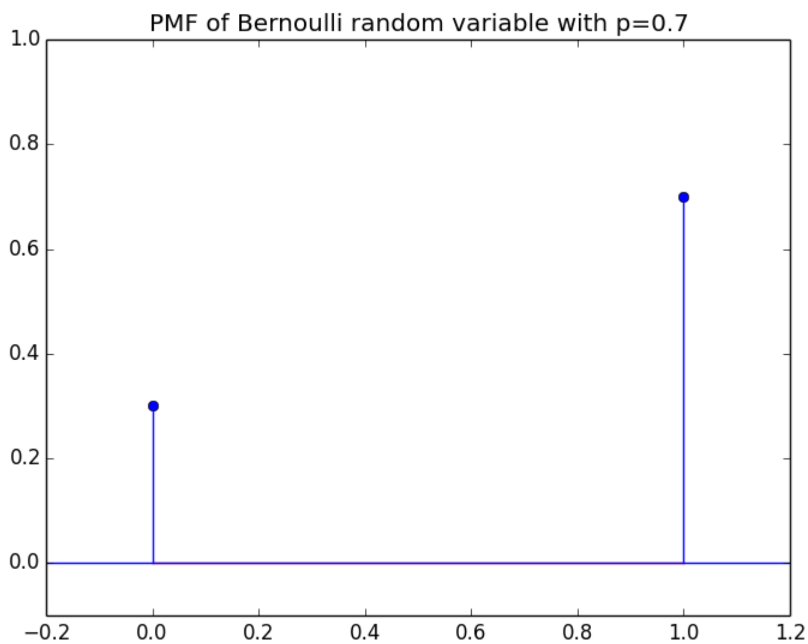
$$P_X(0) = 1 - p, P_X(1) = p.$$

(And 0 elsewhere; from here on out, if we don't specify the value of a PMF P_X at $a \in \mathbf{R}$, then $P_X(a) = 0$.) A random variable X with this distribution: Flip a biased coin with heads probability P, and set $X = 1$ if the result is heads and $X = 0$ otherwise.

2 Binomial distribution

Here there are two parameters: A real number $0 \leq p \leq 1$, and a positive integer n. If $0 \leq k \leq n$, then

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$



PMF of Bernoulli random variable with p=0.7
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2

A random variable X that has this distribution: the number of heads on n successive tosses of a biased coin with heads probability p.

Ordinarily, when we introduce a PMF in this fashion, we need to make sure that its values are all nonnegative and add to 1. Our previous probability calculations about tosses of a biased coin insure this, but here is an independent, super-simple proof: By the binomial theorem:

$$\begin{aligned}\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k} &= (p+(1-p))^n \\ &= 1^n \\ &= 1.\end{aligned}$$

Another way to look at this: A binomial random variable X with parameters p, n, is the sum of n mutually independent Bernoulli variables

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

each with parameter p. To see why this gives the correct distribution, note that $P(X = k)$ is the sum of all terms of the form

$$P((X_1 = a_1) \wedge (X_2 = a_2) \wedge \cdots \wedge (X_n = a_n)),$$

where exactly k of the a_i are 1, and $n-k$ are 0. By independence, each such term is equal to the product

$$P(X_1 = a_1)\cdots P(X_n = a_n) = p^k(1-p)^{n-k}.$$

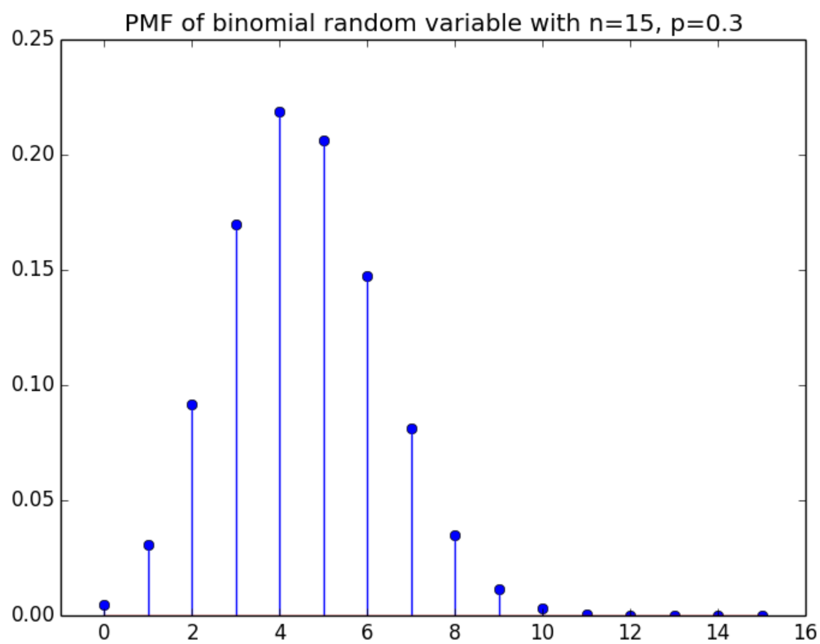
As we have seen before, there are $\binom{n}{k}$ ways to choose the sequence $(a_1, \ldots, a_n)$ of 0's and 1's with exactly k occurrences of 1, so this gives the formula for $P(X = k)$.

3 Geometric distribution

There is a single parameter $0 \leq p \leq 1$. If k is a positive integer, then

$$P_X(k) = (1-p)^{k-1}p.$$

A random variable that has this distribution: X is the number of flips of a biased coin with heads probability p until heads appears. (So, for instance, $X = 1$ if the first toss is heads, $X = 3$ if the first two are tails and the third is heads.) Note that this has a nonzero value at all positive integers k.


PMF of binomial random variable with n=15, p=0.3
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
14
16

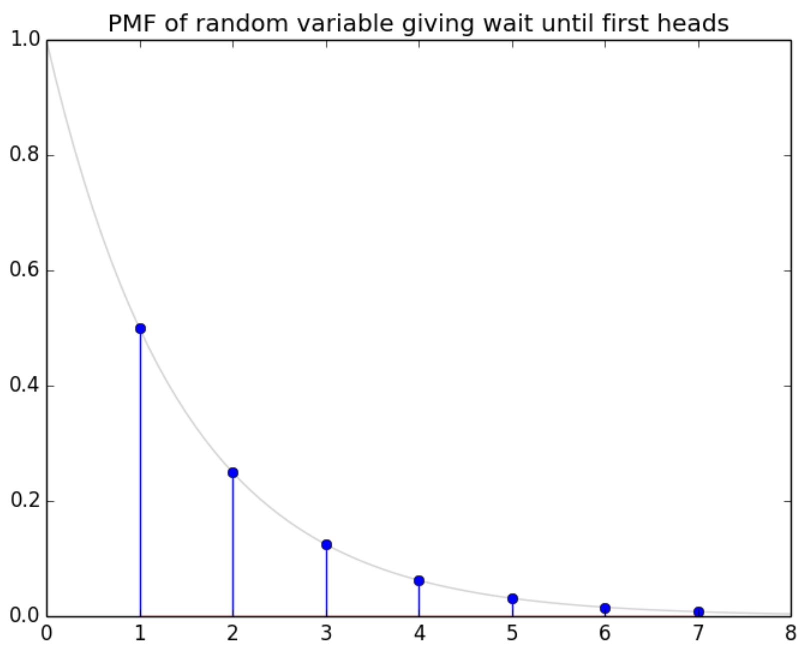


Figure 1: Geometric Distribution

4 Hypergeometric distribution

I don't know why this is called 'hypergeometric'!

To set the stage: Pick 5 cards without replacement from a shuffled deck, and let X denote the number of hearts in the hand. What is $P(X = k)$? Of course, we have to have $0 \leq k \leq 5$ to get a nonzero probability. The answer is

$$\frac{\binom{13}{k} \cdot \binom{39}{5-k}}{\binom{52}{5}.}$$

This requires a bit of explaining: To construct a 5-card hand with exactly k hearts, we specify two things: Which k of the 13 hearts in the deck we choose: there are $\binom{13}{k}$ ways to do this. And, which $5 - k$ of the 39 non-hearts we choose: there are $\binom{39}{5-k}$ ways to do this. That explains the numerator—the denominator is just the total number of 5-card hands, as we've seen many times. A particular case—when you compute the probability of a flush in hearts, the second factor in the numerator becomes 1, so you don't have to write it.

In general, a hypergeometric distribution has three parameters: a positive integer N (52 in the above example), $0 \leq K \leq N$ (13 in the above example), and $0 \leq n \leq n$ (5 in the example). If $0 \leq k \leq n$ then

$$P_X(k) = \frac{\binom{K}{k} \cdot \binom{N-K}{n-k}}{\binom{N}{n}}.$$

A random variable with this distribution is given by the following problem: An urn contains N balls, K of which are black. Reach in and pick out n balls. Let X be the number of black balls (our cards problem is just one instance of this).

If a fixed proportion $p = K/N$ of the balls are black, and if N is allowed to increase toward ∞, while n stays fixed, then this essentially turns into a sampling with replacement problem, since removing n balls from the urn barely changes the proportion of black balls. Thus the hypergeometric distribution approaches the binomial distribution with parameters n, p as a limit. There is a pretty simple formal proof of this fact, but the intuition should suffice.

5 Poisson distribution

Let $\lambda > 0$. We set

$$P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}.$$

The fact that this forms a probability distribution follows from the power series expansion of e^λ:

$$e^\lambda = 1 + \lambda + \frac{\lambda^2}{2!} + \cdots .$$

Where does the Poisson distribution occur 'in nature'? Consider a binomial random variable Y with parameters n, p where p is quite small, at values k such that $k << n$. Then $P(Y = k)$ is very well approximated by $P(X = k)$, where X is a Poisson random variable with $\lambda = np$. The plot below shows the PMF of a binomial random variable with $n = 500$, $p = 0.01$, and a Poisson random variable with $\lambda = 5$. This approximation makes it easy to compute the values of these probabilities without having to resort to evaluating very large binomial coefficients.

For example, suppose 500 shots are fired at a target. It is known that only about one in one hundred shots hit the bulls-eye. What is the probability of getting three or more bulls-eyes? Here we compute probability of the complementary event– that there are 0,1 or 2 bulls-eyes. In the Poisson approximation this is given by

$$e^{-\lambda}(1 + \lambda + \lambda^2/2) = 18.5e^{-5} \approx 0.1247,$$

so the probability of at least three bulls-eyes is about 0.8755. The exact answer is about 0.8766.

The fine print: Of course terms like 'quite small' and '$k << n$' are vague. We will provide an explanation of why this approximation works, and in so doing get a good estimate of the error:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \frac{n \cdot (n-1) \cdots (n-(k-1))}{k!} \cdot p^k (1-p)^n (1-p)^{-k}.$$

We factor n out of each of the k factors in the numerator and get

$$\frac{(np)^k}{k!} \cdot (1-p)^n \cdot (1-\frac{1}{n})(1-\frac{2}{n}) \cdots (1-\frac{k-1}{n}) \cdot (1-p)^{-k}.$$

If we write $\lambda = np$ and use the approximation $1 - p \approx e^{-p}$ for small p, we get

$$\frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot (1-\frac{1}{n})(1-\frac{2}{n}) \cdots (1-\frac{k-1}{n}) \cdot e^{-pk}.$$

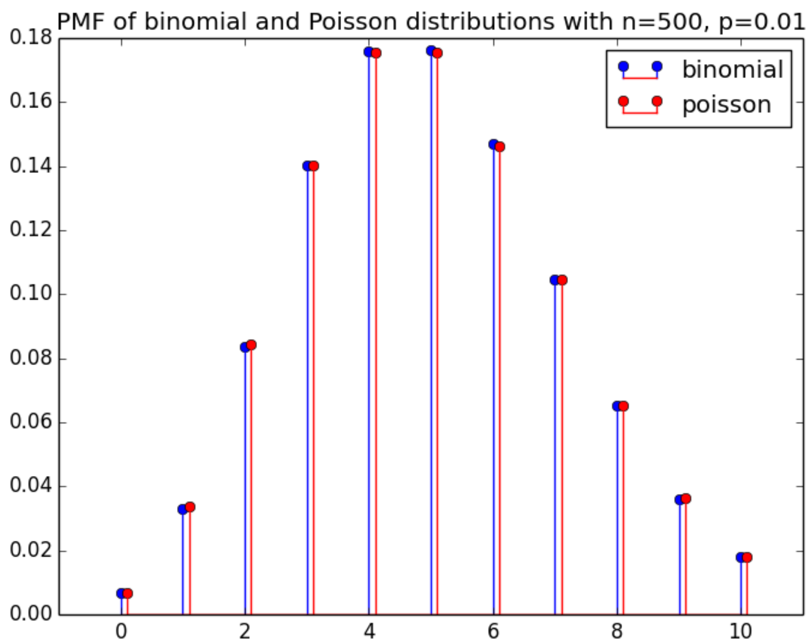
As we saw when we did the birthday problem, the product

$$(1-\frac{1}{n})(1-\frac{2}{n}) \cdots (1-\frac{k-1}{n})$$

is about $e^{\frac{-k^2}{2n}}$ when k is much smaller than n, so we get

$$\frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot e^{pk - \frac{k^2}{2n}}.$$

This is the Poisson distribution multiplied by an error factor of $e^{pk - \frac{k^2}{2n}}$. The point is that this factor is very close to 1. For instance, with the parameters $n = 500, p = 0.01$ given above and $k \leq 10$, the factor is never larger than 1.025, which means a relative error of 2.5%.


PMF of binomial and Poisson distributions with n=500, p=0.01
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
2
4
6
8
10
binomial
poisson