Jinguo Lian
Math437 Notes
January 13, 2016

3 Common Discrete Random Variables

As it turns out, there are some specific distributions that are used over and over in practice, thus they have been given special names. There is a random experiment behind each of these distributions. Since these random experiments model a lot of real life phenomenon, these special distributions are used frequently in different applications. That's why they have been given a name and we devote a section to study them. We will provide PMFs for all of these special random variables, but rather than trying to memorize the PMF, you should understand the random experiment behind each of them. If you understand the random experiments, you can simply derive the PMFs when you need them. Although it might seem that there are a lot of formulas in this section, there are in fact very few new concepts. Do not get intimidated by the large number of formulas, look at each distribution as a practice problem on discrete random variables.

3.1.1 Bernoulli Distribution

A Bernoulli random variable is a random variable that can only take two possible values, usually 0 and 1. This random variable models random experiments that have two possible outcomes, sometimes referred to as "success" and "failure." Here are some examples:

You take a pass-fail exam. You either pass (resulting in X=1) or fail (resulting in X=0).

You toss a coin. The outcome is ether heads or tails.

A child is born. The gender is either male or female.

Formally, the Bernoulli distribution is defined as follows:

Definition 1. *A random variable X is said to be a Bernoulli random variable with parameter p, shown as $X \sim Bernoulli(p)$, if its PMF is given by*

$$P_X(x) = \begin{cases} p & \text{for } x = 1 \\ 1-p & \text{for } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

where $0 < p < 1$.

Remark 2. *1. If $X \sim Bernoulli(p)$, $E(X) = 1 \cdot p + 0 \cdot (1-p) = p$.*

2. If $X \sim Bernoulli(p)$, $Var(X) = (1-p)^2 p + (0-p)^2(1-p) = p(1-p)$.

A Bernoulli random variable is associated with a certain event A. If event A occurs (for example, if you pass the test), then X=1; otherwise X=0. For this reason the Bernoulli random variable, is also called the indicator random variable. In particular, the indicator random variable IA for an event A is defined by

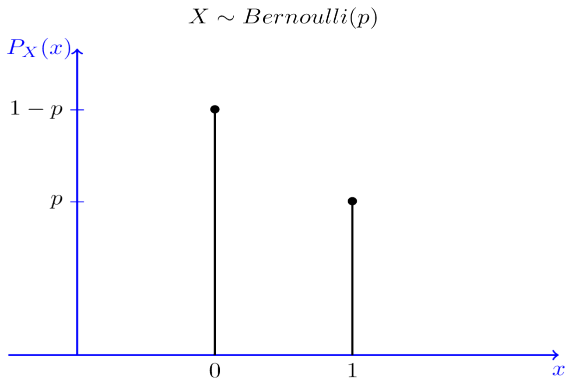

Figure 1: PMF of a Bernoulli(p) random variable.

$$I_A = \begin{cases} 1 & \text{if the event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

The indicator random variable for an event A has Bernoulli distribution with parameter p=P(A), so we can write

$$I_A \sim Bernoulli(P(A))$$

.

3.1.2 Geometric Distribution

The random experiment behind the geometric distribution is as follows. Suppose that I have a coin with $P(H) = p$. I toss the coin until I observe the first heads. We define X as the total number of coin tosses in this experiment. Then X is said to have geometric distribution with parameter p. In other words, you can think of this experiment as repeating independent Bernoulli trials until observing the first success.

The range of X here is $R_X = \{1, 2, 3, ...\}$.

$$P_X(k) = P(X = k) = (1-p)^{k-1}p, \text{ for } k = 1, 2, 3, ...$$

We usually define $q = 1 - p$, so we can write $P_X(k) = pq^{k-1}$, for $k = 1, 2, 3, ...$ To say that a random variable has geometric distribution with parameter p, we write $X \sim Geometric(p)$. More formally, we have the following definition:

Definition 3. *A random variable X is said to be a geometric random variable with parameter p, shown as $X \sim Geometric(p)$, if its PMF is given by*

$$P_X(k) = \begin{cases} p(1-p)^{k-1} & \textit{for } k = 1, 2, 3, ... \\ 0 & \textit{otherwise} \end{cases}$$

where $0 < p < 1$.

Remark 4. *1. If $X \sim Geometric(p)$,*

$$E(X) = p + 2pq + 3pq^2 + ... + kpq^{k-1} + ... = p(1 + 2q + 3q^2 + 4q^3 + ...)$$

$$= p(q + q^2 + q^3 + q^4 + ...)' = p\left(\frac{1}{1-q}\right)' = \frac{p}{(1-q)^2}$$

2. If $X \sim Geometric(p)$, $Var(X) = \frac{q}{p^2}$ the proof is complex.

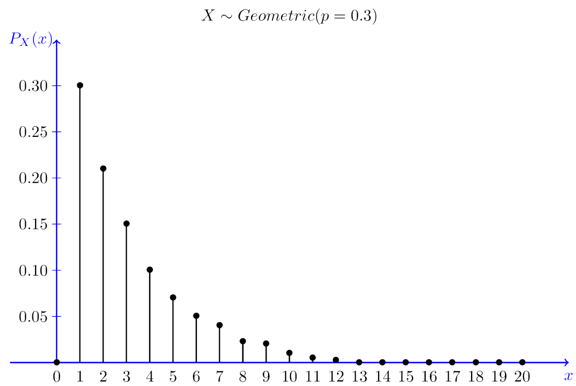

Figure 2: PMF of a Geometric(0.3) random variable.

3.1.3 Binomial Distribution

The random experiment behind the binomial distribution is as follows. Suppose that I have a coin with $P(H) = p$. I toss the coin n times and define X to be the total number of heads that I observe. Then X is binomial with parameter n and p, and we write $X \sim Binomial(n, p)$. The range of X in this case is $R_X = \{0, 1, 2, ..., n\}$. As we have seen in previous Section, the PMF of X in this case is given by binomial formula

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \text{ for } k = 0, 1, 2, ..., n.$$

Definition 5. *A random variable X is said to be a binomial random variable with parameters n and p, shown as $X \sim Binomial(n, p)$, if its PMF is given by*

$$P_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{for } k = 0, 1, 2, \cdots, n \\ 0 & \text{otherwise} \end{cases}$$

where $0 < p < 1$.

Remark 6. *Binomial random variable as a sum of Bernoulli random variables. Here is a useful way of thinking about a binomial random variable. Note that a Binomial(n,p) random variable can be obtained by n independent coin tosses. If we think of each coin toss as a Bernoulli(p) random variable, the Binomial(n,p) random variable is a sum of n independent Bernoulli(p) random variables. This is stated more precisely in the following lemma.*

Lemma 7. *If $X_1, X_2, ..., X_n$ are independent Bernoulli(p) random variables, then the random variable X defined by $X = X_1 + X_2 + ... + X_n$ has a Binomial(n,p) distribution.*

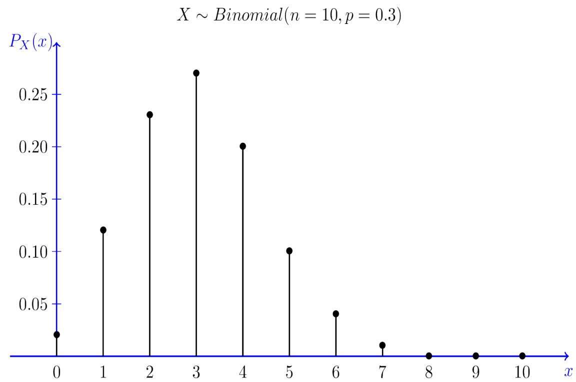

Figure 3: PMF of a Binomial(10,0.3) random variable.

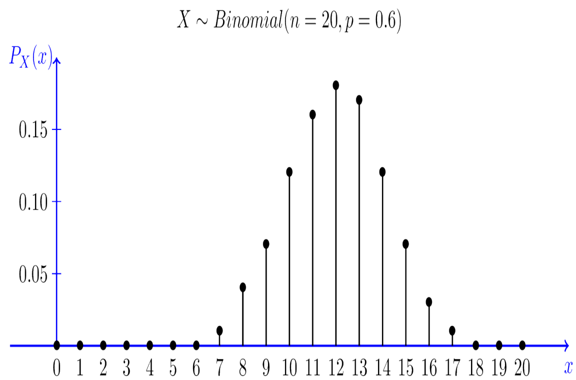

Figure 4: PMF of a Binomial(20,0.6) random variable.

Remark 8. *1. If $X \sim Binomial(p)$, $E(X) = np$.*

2. If $X \sim Binomial(p)$, $Var(X) = np(1-p)$.

To generate a random variable $X \sim Binomial(n, p)$, we can toss a coin n times and count the number of heads. Counting the number of heads is exactly the same as finding $X_1 + X_2 + ... + X_n$, where each X_i is equal to one if the corresponding coin toss results in heads and zero otherwise. This interpretation of binomial random variables is sometimes very helpful. Let's look at an example.

Example 9. *Let $X \sim Binomial(n, p)$ and $Y \sim Binomial(m, p)$ be two independent random variables. Define a new random variable as $Z = X + Y$. Find the PMF of Z.*

Solution: *Since $X \sim Binomial(n, p)$, we can think of X as the number of heads in n independent coin tosses, i.e., we can write*

$$X = X_1 + X_2 + ... + X_n$$

, where the X_i's are independent Bernoulli(p) random variables. Similarly, since $Y \sim Binomial(m, p)$, we can think of Y as the number of heads in m independent coin tosses, i.e., we can write

$$Y = Y_1 + Y_2 + ... + Y_m$$

, where the Yj's are independent Bernoulli(p) random variables. Thus, the random variable $Z = X + Y$ will be the total number of heads in n+m independent coin tosses:

$$Z = X + Y = X_1 + X_2 + ... + X_n + Y_1 + Y_2 + ... + Y_m$$

, where the X_i's and Y_j's are independent Bernoulli(p) random variables. Thus, Z is a binomial random variable with parameters m+n and p, i.e., Binomial$(m+n, p)$. Therefore, the PMF of Z is

$$P_Z(k) = \begin{cases} \binom{m+n}{k} p^k (1-p)^{m+n-k} & \text{for } k = 0, 1, 2, 3, ..., m+n \\ 0 & \text{otherwise} \end{cases}$$

Negative Binomial (Pascal) Distribution

Suppose that I have a coin with $P(H) = p$. I toss the coin until I observe m heads, where $m \in \mathbb{N}$. We define X as the total number of coin tosses in this experiment. Then X is said to have Pascal distribution with parameter m and p. We write $X \sim Pascal(m, p)$. Note that $Pascal(1, p) = Geometric(p)$. Note that by our definition the range of X is given by $R_X = \{m, m+1, m+2, m+3, ...\}$.

Let us derive the PMF of a $Pascal(m, p)$ random variable X. Suppose that I toss the coin until I observe m heads, and X is defined as the total number of coin tosses in this experiment. To find the probability of the event $A = \{X = k\}$, we argue as follows. By definition, event A can be written as $A = B \cap C$, where B is the event that we observe $m-1$ heads (successes) in the first $k-1$ trials, and C is the event that we observe a heads in the k^{th} trial. Note that B and C are independent events because they are related to different independent trials (coin tosses). Thus we can write

$$P(A) = P(B \cap C) = P(B)P(C).$$

Now, we have $P(C) = p$. Note also that $P(B)$ is the probability that I observe observe m1 heads in the k1 coin tosses. This probability is given by the binomial formula, in particular

$$P(B) = \binom{k-1}{m-1} p^{m-1} (1-p)^{\left((k-1)-(m-1)\right)} = \binom{k-1}{m-1} p^{m-1} (1-p)^{k-m}.$$

Definition 10. *A random variable X is said to be a Pascal random variable with parameters m and p, shown as $X \sim Pascal(m, p)$, if its PMF is given by*

$$P_X(k) = \begin{cases} \binom{k-1}{m-1} p^m (1-p)^{k-m} & \text{for } k = m, m+1, m+2, m+3, ... \\ 0 & \text{otherwise} \end{cases}$$

where $0 < p < 1$.

Remark 11. *Negative Binomial Distribution is m copies of Geometric distribution, therefore,*

1. $E(X) = \frac{mq}{p}$
2. $Var(X) = \frac{mq}{p^2}$

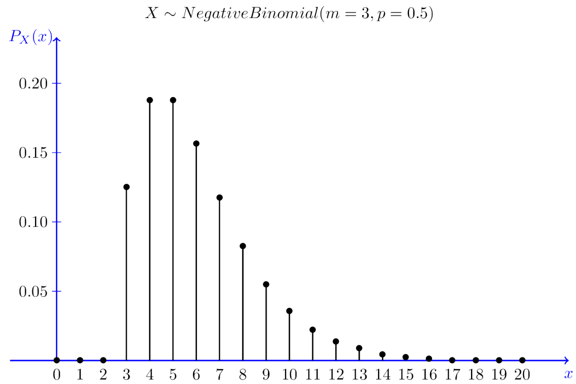

Figure 5: Pascal(3,0.5) (negative binomial) random variable.

3.1.4 Hypergeometric Distribution

Here is the random experiment behind the hypergeometric distribution. You have a bag that contains b blue marbles and r red marbles. You choose $k \leq b+r$ marbles at random (without replacement). Let X be the number of blue marbles in your sample. By this definition, we have $X \leq \min(k, b)$. Also, the number of red marbles in your sample must be less than or equal to r, so we conclude $X \geq \max(0, k-r)$. Therefore, the range of X is given by

$$R_X = \{\max(0, k-r), \max(0, k-r)+1, \max(0, k-r)+2, ..., \min(k, b)\}.$$

To find $P_X(x)$, note that the total number of ways to choose k marbles from b+r marbles is $\binom{b+r}{k}$. The total number of ways to choose x blue marbles and $k-x$ red marbles is $\binom{b}{x}\binom{r}{k-x}$. Thus, we have

$$P_X(x) = \frac{\binom{b}{x}\binom{r}{k-x}}{\binom{b+r}{k}}, \qquad \text{for } x \in R_X.$$

Definition 12. *A random variable X is said to be a Hypergeometric random variable with parameters b,r and k, shown as $X \sim Hypergeometric(b, r, k)$, if its range is*

$$R_X = \{\max(0, kr), \max(0, kr)+1, \max(0, kr)+2, ..., \min(k, b)\}$$

, and its PMF is given by

$$P_X(x) = \begin{cases} \frac{\binom{b}{x}\binom{r}{k-x}}{\binom{b+r}{k}} & \textit{for } x \in R_X \\ 0 & \textit{otherwise} \end{cases}$$

Remark 13. *1.* $E(X) = k\frac{b}{b+r}$

2. $Var(X) = k(\frac{b}{b+r})(1 - \frac{b}{b+r})(\frac{b+r-k}{b+r-1})$

3.1.5 Poisson Distribution

The Poisson distribution is one of the most widely used probability distributions. It is usually used in scenarios where we are counting the occurrences of certain events in an interval of time or space. In practice, it is often an approximation of a real-life random variable. Here is an example of a scenario where a Poisson random variable might be used. Suppose that we are counting the number of customers who visit a certain store from 1pm to 2pm. Based on data from previous days, we know that on average $\lambda = 15$ customers visit the store. Of course, there will be more customers some days and fewer on others. Here, we may model the random variable X showing the number customers as a Poisson random variable with parameter $\lambda = 15$. Let us introduce the Poisson PMF first, and then we will talk about more examples and interpretations of this distribution.

Definition 14. *A random variable X is said to be a Poisson random variable with parameter λ, shown as $X \sim Poisson(\lambda)$, if its range is $R_X = \{0, 1, 2, 3, ...\}$, and its PMF is given by*

$$P_X(k) = \begin{cases} \frac{e^{-\lambda}\lambda^k}{k!} & \text{for } k \in R_X \\ 0 & \text{otherwise} \end{cases}$$

Before going any further, let's check that this is a valid PMF. First, we note that $P_X(k) \geq 0$ for all k. Next, we need to check $\sum_{k \in R_X} P_X(k) = 1$. To do that, let us first remember the Taylor series for e^x, $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$. Now we can write

$$\begin{aligned} \sum_{k \in R_X} P_X(k) &= \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} e^{\lambda} = 1 \end{aligned}$$

Remark 15. $E(X) = Var(X) = \lambda$.

Example 16. *A busy intersection is the scene of many traffic accidents. An analyst studies data on the accidents and concludes that accidents occurs there at "an average rate of $\lambda = 2$ per month". This does not means there are 2 accidents in each month. In any given number of accidents X in a month is a random variable. The Possion distribution can be used to find the probabilities $P(X = k)$ in terms of k and λ, the average rate.*

3.1.6 Poisson as an approximation for binomial

The Poisson distribution can be viewed as the limit of binomial distribution. Suppose $X \sim Binomial(n, p)$ where n is very large and p is very small. In particular, assume that $\lambda = np$ is a positive constant. We show that the PMF of X can be approximated by the PMF of a $Poisson(\lambda)$ random variable. The importance of this is that Poisson PMF is much easier to compute than the binomial. Let us state this as a theorem.

Assumption 1: The probability of exactly one accident in a small time interval of length is approximately λt.

Assumption 2: Accident occur independently in time intervals which do not intersect.

Theorem 17. *Let* $X \sim Binomial(n, p = \frac{\lambda}{n})$*, where* $\lambda > 0$ *is fixed. Then for any* $k \in \{0, 1, 2, ...\}$*, we have*

$$\lim_{n \to \infty} P_X(k) = \frac{e^{-\lambda} \lambda^k}{k!}.$$

3.1.7 The Discrete Uniform Distribution

When we roll a single fair die and observe the number X that came up, the sample space is $S = \{1, 2, 3, 4, 5, 6\}$ and each of the outcomes was equally likely with probability 1/6. The random variable X is said to have a discrete uniform distribution on 1,2,,3,4,5,6.

Definition 18 (Discrete Uniform Distribution on 1,2,...,n)**.**

1. $P(x) = \frac{1}{n}, x = 1, 2, 3, 4, ..., n$
2. $E(X) = \frac{n+1}{2}$
3. $Var(X) = \frac{n^2-1}{12}$.

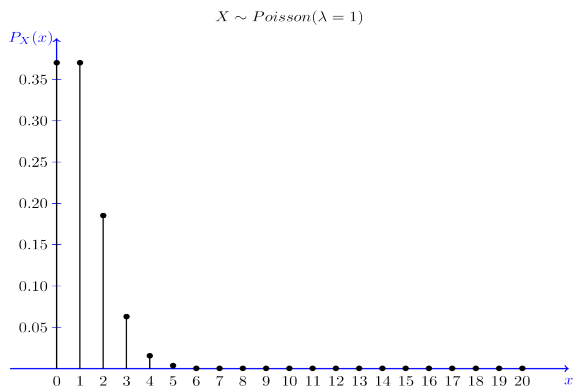

Figure 6: PMF of a Poisson(1) random variable.

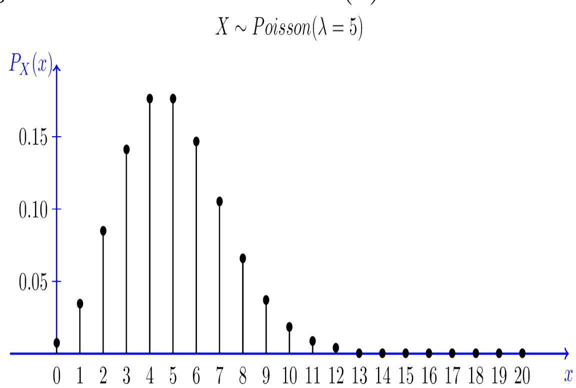

Figure 7: PMF of a Poisson(5) random variable.

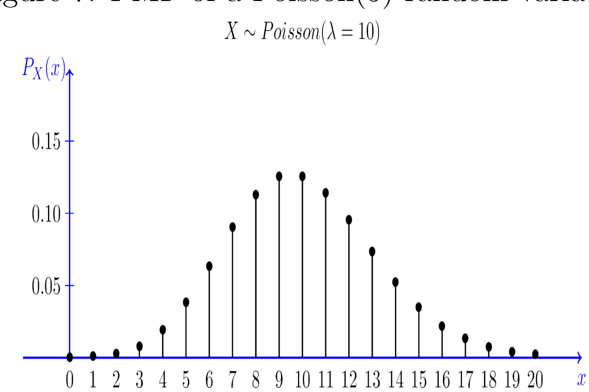

Figure 8: PMF of a Poisson(10) random variable.