Lecture 4
Common Discrete Distributions

Chao Song

College of Ecology
Lanzhou University

September 14, 2024

Bernoulli distribution

A **Bernoulli trial** is a random experiment, the outcome of which can be classified in one of the two mutually exclusive and exhaustive ways–say, success of failure. Let X be a random variable associated with a Bernoulli trial such that $X = 1$ for success and $X = 0$ for failure, X follows a **Bernoulli distribution**.

Example: Suppose that the probability of germination of a beet seed is 0.8 and the germination of a seed is called a success. If we plant 10 seeds and can assume that the germination of one seed is independent of the germination of another seed. This would correspond to 10 Bernoulli trials with $p = 0.8$.

Bernoulli distribution

The probability mass function of X following a Bernoulli distribution is

$$f(x) = \begin{cases} p, & X = 1 \\ 1 - p, & X = 0 \end{cases},$$

Or more concisely, $f(x) = p^x(1-p)^{1-x}$.

The mean and variance of a Bernoulli distribution is

- $E(X) = 1 \times p + 0 \times (1-p) = p$
- $Var(X) = (1-p)^2 p + (0-p)^2(1-p) = p(1-p)$

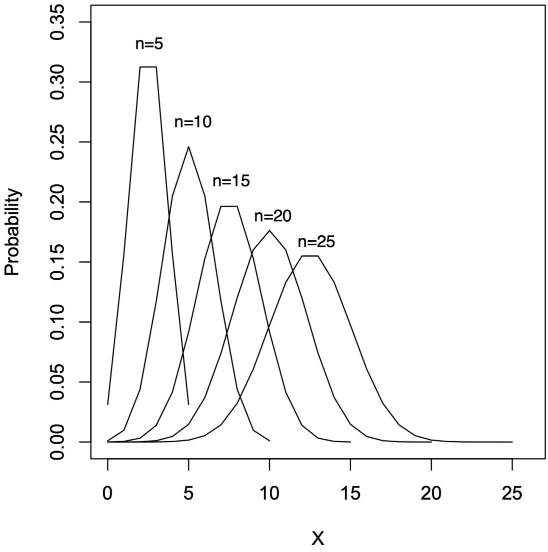
Binomial distribution

In a sequence of Bernoulli trials, we are often interested in the total number of successes, but not the actual order of their occurrences. Let random variable X equal the number of observed successes in n Bernoulli trials.

Binomial distribution: If a random variable X denotes the number of successes in n independent Bernoulli trials, X follows a binomial distribution and its PMF is

$$P(X = k) = \mathbf{C}_n^k p^k (1 - p)^{n-k},\ k = 0, 1, \ldots, n$$

Binomial distribution



Binomial distribution

What is the mean and variance of a binomial distribution?

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \cdot \mathbf{C}_n^x p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{\infty} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= np \sum_{x=1}^{\infty} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x} \\
&= np
\end{aligned}
$$

because $\sum_{x=1}^{\infty} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1}(1-p)^{n-x}$ is the binomial expansion of $(p+1-p)^{n-1}$ and is thus equal to 1.

Binomial distribution

$$
\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot \mathbf{C}_n^x p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{\infty} x(x-1) \cdot \mathbf{C}_n^x p^x (1-p)^{n-x} + \sum_{x=0}^{\infty} x \cdot \mathbf{C}_n^x p^x (1-p)^{n-x} \\
&= \sum_{x=2}^{\infty} x(x-1) \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} + np \\
&= \sum_{x=2}^{\infty} n(n-1)p^2 \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x} + np \\
&= n(n-1)p^2 + np \\
&= n^2p^2 - np^2 + np
\end{aligned}
$$

$$
Var(X) = E(X^2) - [E(X)]^2 = n^2p^2 - np^2 + np - (np)^2 = np(1-p)
$$

Binomial distribution

We can also derive the mean and variance using MGF:

$$
\begin{aligned}
M_X(t) &= E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \mathbf{C}_n^x p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{\infty} \mathbf{C}_n^x (pe^t)^x (1-p)^{n-x} \\
&= (pe^t + 1 - p)^n
\end{aligned}
$$

$$
\begin{aligned}
E(X) &= M'_X(0) = n(pe^t + 1 - p)^{n-1} pe^t \Big|_{t=0} = np \\
E(X^2) &= M''_X(0) \\
&= n(pe^t + 1 - p)^{n-1} pe^t + n(n-1)(pe^t + 1 - p)^{n-2} (pe^t)^2 \Big|_{t=0} \\
&= np + n^2 p^2 - np^2
\end{aligned}
$$

Hypergeometric distribution

A urn contains N balls and K of them are marked. If you randomly select n balls, what is the probability that you get k marked balls?

Let X be the number of marked balls in the n balls one selected,

$$P(X = k) = \frac{\mathbf{C}_K^k \mathbf{C}_{N-K}^{n-k}}{\mathbf{C}_N^n}$$

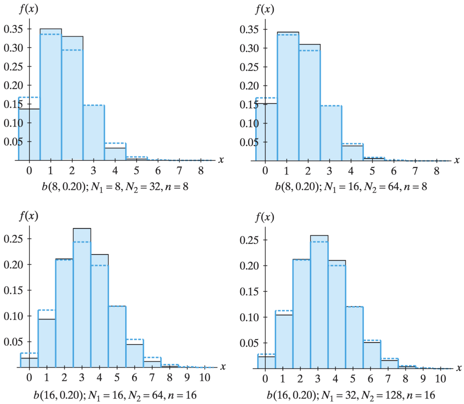
Binomial vs hypergeometric distribution

A urn contains $N1$ white balls and $N2$ of them are marked. Let $p = N_1/(N_1 + N_2)$ and X equal the number of marked balls in a random sample of size n. What is the distribution of X (1) if the sampling is done one at a time with replacement? and (2) if the sampling is one without replacement?

Answer: If sampling is done with replacement, all successive draws are independent. X thus follows a binomial distribution. In contrast, if sampling is done without replacement, one draw will influence the probability of drawing in the next round, we thus have a hypergeometric distribution for X.

Binomial vs hypergeometric distribution

If there are very large number of balls in total compared to the number of balls we draw, i.e., $(N_1 + N_2) \gg n$, hypergeometric distribution and binomial distribution becomes similar.

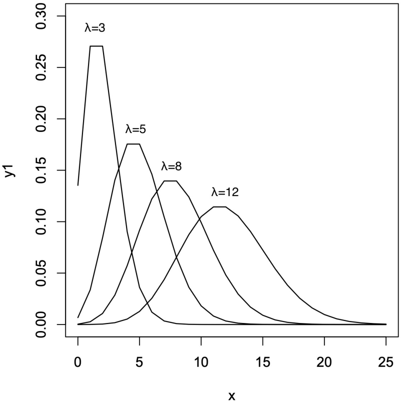


Comparison of binomial and hypergeometric distribution (shaded)

Poisson distribution

Poisson distribution: Let λ be a positive number. A random variable is said to have a Poisson distribution if its probability mass function is

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda},\ k = 0, 1, 2, \ldots$$



Poisson distribution

Let X follows a Poisson distribution with parameter λ. Show that its mean and variance are both λ.

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \cdot \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \sum_{x=1}^{\infty} x \cdot \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda} e^{\lambda} \\
&= \lambda
\end{aligned}
$$

given the power series expansion of exponential function $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$

Poisson distribution

To get the variance of X, we first get $E(X^2)$:

$$\begin{aligned}
E(X^2) &= \sum_{x=0}^{\infty} x^2 \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \sum_{x=0}^{\infty} x(x-1) \frac{\lambda^x}{x!} e^{-\lambda} + \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} e^{-\lambda} \\
&= \sum_{x=2}^{\infty} x(x-1) \frac{\lambda^x}{x!} e^{-\lambda} + \lambda \\
&= \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} + \lambda \\
&= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda \\
&= \lambda^2 + \lambda
\end{aligned}$$

Thus, $Var(X) = E(X^2) - [E(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

Poisson distribution

We can also derive the mean and variance from the MGF:

$$
\begin{aligned}
M_X(t) = E(e^{tX}) &= \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x}{x!} e^{-\lambda} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \\
&= e^{-\lambda} e^{\lambda e^t} \\
&= e^{\lambda e^t - \lambda}
\end{aligned}
$$

$$
\begin{aligned}
E(X) &= M'_X(0) = (e^{\lambda e^t - \lambda} \lambda e^t)\Big|_{t=0} = \lambda \\
E(X^2) &= M''_X(0) = \lambda e^{\lambda e^t - \lambda + t}(\lambda e^t + 1)\Big|_{t=0} = \lambda^2 + \lambda \\
Var(X) &= E(X^2) - [E(X)]^2 = \lambda
\end{aligned}
$$

Poisson distribution

What does a Poisson distributed variable model?

Poisson distribution models the number of events in a time interval t.

- Divide t into n segments such that at most one event occur within a segment;
- Probability of occurrence is $\mu t/n$;
- Number of occurrence is modeled with a binomial distribution.

$$
\begin{aligned}
P(X=k) &= \lim_{n\to\infty} \mathbf{C}_n^k p^k (1-p)^{n-k} \\
&= \lim_{n\to\infty} \frac{n!}{k!(n-k)!} (\frac{\mu t}{n})^k (1-\frac{\mu t}{n})^{n-k} \\
&= \lim_{n\to\infty} \frac{(\mu t)^k}{k!} \frac{n(n-1)\dots(n-k+1)}{n^k} (1-\frac{\mu t}{n})^{-k} (1-\frac{\mu t}{n})^n \\
&= \frac{(\mu t)^k}{k!} e^{-\mu t}
\end{aligned}
$$

Poisson distribution

Poisson distribution is a limiting case of a binomial distribution. Here, $\lambda = \mu t$ is often referred to as the rate parameter of the Poisson distribution.

This derivation gives us a mechanistic insights into when we can use Poisson distribution. When some events occur at a constant rate, we can model the count of event with a Poisson distribution.

Poisson distribution

Example: In a large city, telephone calls to 110 come on the average of two every 3 minutes. If one assumes a Poisson process, what is the probability of five or more calls arriving in a 9-minute period?

Let X denote the number of calls in a 9-minute period. We see that $E(X) = 2 \times 9/3 = 6$. Thus, the PMF of X is

$$P(X = k) = \frac{6^x}{x!} e^{-6}$$

Thus, we have

$$\begin{aligned} P(X \geqslant 5) &= 1 - P(X \leqslant 4) \\ &= 1 - \sum_{x=0}^{4} \frac{6^x}{x!} e^{-6} \\ &= 0.715 \end{aligned}$$

Negative binomial distribution

Negative binomial distribution: In a sequence of independent Bernoulli trials with success probability p, let X be the number of failure until r successes. Then X has a negative binomial distribution with probability mass function

$$P(X = k) = \mathbf{C}_{k+r-1}^{k}(1-p)^{k}p^{r}$$

Why is this called a negative binomial distribution?

Let $q = 1 - p$ and $h(q) = (1-q)^{-r}$. Using Taylor expansion at $q = 0$

$$h(q) = \sum_{k=0}^{\infty} \frac{h^{(k)}(0)}{k!} q^{k} = \sum_{k=0}^{\infty} \mathbf{C}_{k+r-1}^{r-1} q^{k} = \sum_{k=0}^{\infty} \mathbf{C}_{k+r-1}^{k} q^{k}$$

Thus, we can see that the PMF of a negative binomial distribution is the summand of $p^{r}p^{-r}$

Negative binomial distribution

What is the mean and variance of a negative binomial distribution?

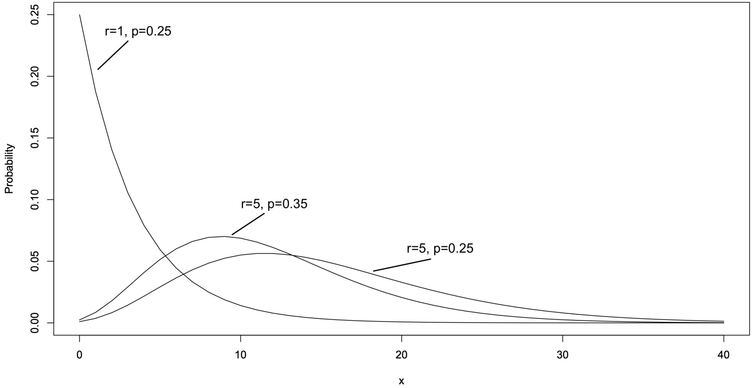
To calculate the mean and variance, we first get the MGF:

$$\begin{aligned} M(t) &= \sum_{k=0}^{\infty} e^{tk} \mathbf{C}_{k+r-1}^{k} (1-p)^k p^r \\ &= p^r \sum_{k=0}^{\infty} \mathbf{C}_{k+r-1}^{k} [(1-p)e^t]^k \\ &= \frac{p^r}{[1-(1-p)e^t]^r} \end{aligned}$$

Using the derivatives of $M(t)$ evaluated at $t = 0$, we get that the mean of X is $\frac{r(1-p)}{p}$ and the variance of X is $\frac{r(1-p)}{p^2}$.

Negative binomial distribution

The negative binomial distribution can take on a variety of shapes, depending on the parameters r and p. An important feature of negative binomial distribution is that its variance is larger than the mean.



Geometric distribution

Geometric distribution: In a sequence of independent Bernoulli trials with success probability p, let X be the total number of failures until we have1 successes, X has a geometric distribution with probability mass function:

$$P(X = x) = (1 - p)^x p$$

Geometric distribution is a special case of negative binomial distribution.

The mean and variance of the geometric distribution is $\frac{1-p}{p}$ and $\frac{1-p}{p^2}$, respectively.

Summary of common discrete distributions

Distribution	Probability mass function	Mean	Variance
Bernoulli	$p^x(1-p)^{1-x}$	p	$p(1-p)$
Binomial	$\mathbf{C}_n^k p^k(1-p)^{n-k}$	np	$np(1-p)$
Poisson	$\frac{\lambda^k}{k!}e^{-\lambda}$	λ	λ
Negative binomial	$\mathbf{C}_{k+r-1}^k(1-p)^k p^r$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$
Geometric	$(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Hypergeometric	$\frac{\mathbf{C}_K^k \mathbf{C}_{N-K}^{n-k}}{\mathbf{C}_N^n}$	$\frac{nK}{N}$	$\frac{nK(N-K)(N-n)}{N^2(N-1)}$