Discrete Probability Distributions
03/09/2025

Unlike continuous random variables that are often the outcomes from measurements, discrete random variables stem from counting. They are important in quality control, demographics, and social sciences. This article provides a short summary of discrete probability statistics.

The Bernoulli Process

The Bernolli process is a mental “experiment” for probabilistic thinking, by selecting a number of samples from a population called trials, with the following restrictions:

a) There are n trials and n is a fixed number
b) Each trial results in one of only two possible outcomes. One is called “success” and the other “failure” just for convenience
c) The probability of success on each trial is a constant, p
d) The trials are independent unless otherwise noted (see hypergeometric distribution)

Uniform Distribution

The probability of each trial is a constant, typically one over the dimension of the population space, e.g., an unbiased six-faced die or a fair coin.

$$u(x; N) = \frac{1}{N}$$

The mean and variance of the uniform distribution are:

$$\mu = \frac{N+1}{2} \; and \; \sigma^2 = \frac{(N+1)(N-1)}{12}$$

Example: What is the probability of drawing a “face card (J, Q, and K)” from a deck of cards? There are $N = 52$ cards among which 12 are face cards, each with a probability of $1/N$. Based on the addition rule, the probability of drawing a face card is $12/52 \; or \; 3/13$.

Geometric Distribution

Independent trials within a population space conclude the probability of success p, and failure, $1 - p$; the probability of the number of trials, x, that results in the first success is:

$$g(x; p) = p(1-p)^{x-1}$$

This is understood as the experiment that first selects $(x-1)$ failures immediately followed by success. The mean and variance of the distribution are:

$$\mu = \frac{1}{p} \; and \; \sigma^2 = \frac{1-p}{p}$$

Example: A race track publishes the winning (1st, 2nd, or 3rd place) probability by their favorite horse as 0.67. How many races must one bet to be 95% sure of winning something? Using the addition rule, the cumulative probability of betting n times is computed as below:

$$P(x=n) = \sum_{i=1}^{n} p(1-p)^{i-1} = p\frac{1-(1-p)^n}{p} = 1-(1-p)^n$$

Set $P(x=n) = 0.95$, solving for $n = 2.7$, or 3 bets.

Hypergeometric Distribution

From a population of size N with k successes, one may randomly select n samples that contain x successes. This is known as a "hypergeometric experiment" since the samples are selected from the population without replacement (therefore not independent). The probability of the experiment can be understood as selecting x from k successes and $n-x$ from $N-k$ failures over the total number of ways of selecting n samples from the population of dimension N:

$$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

The mean and variance of hypergeometric distribution are:

$$\mu = \frac{nk}{N} \; and \; \sigma^2 = \frac{nk(N-n)(N-k)}{N^2(N-1)}$$

Example: Five cards are drawn at random without replacement from a deck of cards. What is the probability that exactly 3 are face cards?

$$h(3; 52, 5, 12) = \frac{\binom{12}{3}\binom{40}{2}}{\binom{52}{5}} = 0.066$$

Binomial Distribution

The binomial distribution is a direct result of the Bernoulli trials for a success probability of p and failure $1-p$. This is the general case of geometric distribution for there are $C(n,x)$ ways of having x successes and $(n-x)$ failures.

$$b(x;n,p) = \binom{n}{x} p^x (1-p)^{n-x}$$

The mean and variance of binomial distribution are:

$$\mu = np \ and \ \sigma^2 = np(1-p)$$

Example: The employee population of a major employer includes 30% minorities. 15 people are selected at random from the employees to form a committee. What is the probability that exactly 3 committee members are of minorities?

$$b(3;15,0.3) = \binom{15}{3} 0.3^3 0.7^{12} = 0.17$$

Multinomial Distribution

In a given trial that results in k outcomes (events) with probabilities $p_1, p_2, \ldots, p_k$, the probability distribution of the random variables, $x_1, x_2, \ldots, x_k$, representing the the number of occurrences of the k outcomes in n independent trials is

$$m(x_1, x_2, \ldots, x_k; p_1, p_2, \ldots, p_k, n) = \binom{n}{x_1, x_2, \ldots, x_k} p_1^{x_1} p_2^{x_2}, \ldots, p_k^{x_k}$$

Example: The US households are 67% white, 14% hispanic, 13% black, and 6% other races. What is the probability that of 10 households selected, 6 are white, 2 are hispanic, 1 is black and 1 is from other races?

$$m(6,2,1,1;0.67,0.14,0.13,0.06,10) = \binom{10}{6,2,1,1} 0.67^6 0.14^2 0.13^1 0.06^1 = 0.035$$

Multivariate Hypergeometric Distribution

Likewise, the probability distribution of the random variables $x_1, x_2, \ldots, x_k$, representing the the number of elements selected from k subsets in a sample size of n is:

$$mh(x_1, x_2, \ldots, x_k; a_1, a_2, \ldots, a_k, N, n) = \frac{\binom{a_1}{x_1}\binom{a_2}{x_2}\cdots\binom{a_k}{x_k}}{\binom{N}{n}}$$

Example: From a box of 10 oranges, 4 apples, and 6 bananas, a random sample of 8 pieces of fruit is selected. What is the probability that 4 oranges, 1 apple and 3 bananas are in the sample?

$$mh(4, 1, 3; 10, 4, 6, 20, 8) = \frac{\binom{10}{4}\binom{4}{1}\binom{6}{3}}{\binom{20}{8}} = 0.133$$

Negative Binomial Distribution

The negative binomial probability distribution of the random variable, X, for the number of the trial on which the kth success occurs is:

$$b^*(x; k, p) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

The mean and variance of the negative binomial distribution are:

$$\mu = \frac{k}{p} \ and \ \sigma^2 = \frac{k(1-p)}{p}$$

The name of negative binomial distribution is for the reason that the distribution can be written in an alternative form:

$$b^*(x; k, p) = \binom{-k}{x-k} p^k (1-p)^{x-k}$$

Example: A telemarketer makes a sale with probability 0.2. What is the probability that the telemarketer makes the second sale on the fifth call?

$$b^*(5; 2, 0.2) = \binom{4}{1} 0.2^2 0.8^3 = 0.082$$

Binomial Distribution as a Limiting Case of Hypergeometric Distribution

Let $p = k/N$. So that $k = Np \ and \ N - k = N(1-p)$, at $N \to \infty$:

$$\lim_{N\to\infty} h(x;N,n,k) = \lim_{N\to\infty} \frac{\binom{Np}{x}\binom{N(1-p)}{n-x}}{\binom{N}{n}} = \binom{n}{x} p^x (1-p)^{n-x} = b(x;n,p)$$

The limiting case is a good approximation for hypergeometric distribution when $n/N \leq 0.05$.

Example: In 2018, 83% of US households had internet access. If six US households were randomly selected then, what is the probability that the number of households with internet access was exactly 4?

$$b(4;6,0.83) = \binom{6}{4} 0.83^4 0.17^2 = 0.206$$

Poisson Distribution

The probability distribution of the Poisson random variable, X, representing the number of outcomes occurring during a given time interval or specified region is:

$$p(x;\lambda) = \frac{\lambda^x}{x!} e^{-\lambda}$$

Where λ is the average number of outcomes per unit time, distance, area or volume.

The mean and variance of Poisson distribution are:

$$\mu = \lambda \; and \; \sigma^2 = \lambda$$

Example: The average number of lion spotting is 5 on a popular safari tour. What is the probability that no lion is seen on the next tour?

$$p(0;5) = \frac{5^0}{0!} e^{-5} = 0.0067$$

Poisson Distribution as a Limiting Case of Binomial Distribution

Let $np = \lambda \; at \; n \to \infty$.

$$\lim_{n\to\infty} b(x;n,p) = p(x;\lambda)$$

Example: A tricycle company received complaints on a brake problem one in 10,000 products sold. During a quality audit, 200 tricycles were selected from the production for detailed examination. What is the probability that 5 tricycles have brake problems?

$$b(5; 200, 0.0001) = \binom{200}{5} 0.0001^5 0.9999^{195} = 0$$

Using Poisson approximation with $\lambda = np$:

$$p(5; 200 \times 0.0001) = \frac{0.02^5}{5!} e^{-0.02} = 0$$

Concluding Remarks

Discrete probability distributions are important foundations for studying statistics. Although multivariate distributions often become complex, most of the distributions may be approximated by others when either the population size or the sample size is large. For example, the binomial distribution can be used to approximate the hypergeometric distribution when the population size is large. Also, Poisson distribution can be used to approximate the binomial distribution when the sample size is large. In addition, when $np \geq 5 \ and \ n(1-p) \geq 5$, the normal distribution (a continuous distribution) may be used to approximate the binomial distribution with so-called "continuity corrections." Therefore, the majority of practical work may be accomplished with the normal distribution using standard tools. From time to time, nonetheless, when the population size or the sample size is small, one must work with the discrete distributions in order to be mathematically rigorous.