Statistical Methods 105

Department of Statistics and Operations Research

January 21, 2019

Chapter 1
Discrete random variable

Plan

1. Discrete probability distributions
2. Some Discrete probability Distributions
   - Discrete Uniform Random Variable
   - Binomial Distribution
3. Hypergeometric Distribution
4. Poisson Distribution

Plan

1. Discrete probability distributions
2. Some Discrete probability Distributions
   - Discrete Uniform Random Variable
   - Binomial Distribution
3. Hypergeometric Distribution
4. Poisson Distribution

Definition (Probability function)

The set of ordered pairs $(x, f(x))$ is a probability function, probability mass function, or probability distribution of the discrete random variable X if, for each possible outcome x,

1. $f(x) \geq 0$,
2. $\sum_{x \in X} f(x) = 1$,
3. $P(X = x) = f(x)$.

Definition (cumulative distribution function)

The cumulative distribution function $F(x)$ of a discrete random variable X with probability distribution $f(x)$ is

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad \textit{for} \quad -\infty < x < +\infty.$$

Definition (Mean of a Random Variable)

Let X be a random variable with probability distribution $f(x)$. The mean, or expected value, of X is

$$\mu = E(x) = \sum_{x} x\, f(x).$$

Example

A lot containing 7 components is sampled by a quality inspector; the lot contains 4 good components and 3 defective components. A sample of 3 is taken by the inspector. Find the expected value of the number of good components in this sample.

Solution

Assume X represents the number of good components in the sample. The probability distribution of X is

$$f(x) = \frac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}}, \ x = 1, 2, 3. \tag{1}$$

Using the formula (1), we obtain

$$f(0) = \frac{1}{35},\ f(1) = \frac{12}{35},\ f(2) = \frac{18}{35}\ and\ f(3) = \frac{4}{35}.$$

Therefore,

$$\mu = E(x) = 0 * f(0) + 1 * f(1) + 2 * f(2) + 3 * f(3) = \frac{12}{7} = 1.7.$$

Thus, if a sample of size 3 is selected at random over and over again from a lot of 4 good components and 3 defective components, it will contain, on average, 1.7 good components.

Theorem

Let X be a random variable with probability distribution $f(x)$. The expected value of the random variable $g(X)$ is

$$\mu_{g(X)} = E[g(X)] = \sum_x g(x)f(x).$$

Example

Suppose that the number of cars X that pass through a car wash between 4:00 P.M. and 5:00 P.M. on any sunny Friday has the following probability distribution:

x	4	5	6	7	8	9
$f(x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Let $g(X) = 2X + 1$ represent the amount of money, in dollars, paid to the attendant by the manager. Find the attendant's expected earnings for this particular time period.

Solution

Simple calculations yield

x	4	5	6	7	8	9
$f(x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$
$g(x)$	9	11	13	15	17	19
$f(x)g(x)$	$\frac{9}{12}$	$\frac{11}{12}$	$\frac{13}{4}$	$\frac{15}{4}$	$\frac{17}{6}$	$\frac{19}{6}$

Therefore, the attendant's expected earnings for this particular time period is equal to:

$$E[g(X)] = \frac{9}{12} + \frac{11}{12} + \frac{13}{4} + \frac{15}{4} + \frac{17}{6} + \frac{19}{6} \approx 14.67.$$

Example

Let X be a random variable with probability distribution as follows:

x	0	1	2	3
$f(x)$	$\frac{1}{3}$	$\frac{1}{2}$	0	$\frac{1}{6}$

Find the expected value of $Y = (X - 1)^2$.

Solution

Simple calculations yield

x	0	1	2	3
$f(x)$	$\frac{1}{3}$	$\frac{1}{2}$	0	$\frac{1}{6}$
$g(x)$	1	0	1	4
$f(x)g(x)$	$\frac{1}{3}$	0	0	$\frac{2}{3}$

Therefore, the expected value of Y is equal to:

$$E(Y) = E[g(X)] = 1.$$

Theorems (Variance of Random Variable)

Let X be a random variable with probability distribution $f(x)$ and mean μ. The variance of X is

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x).$$

The positive square root of the variance, σ, is called the standard deviation of X.

Example

Calculate the variance of $g(X) = 2X + 3$, where X is a random variable with probability distribution:

x	0	1	2	3
$f(x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$

Solution

Simple calculations yield

x	0	1	2	3
$f(x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$
$g(x)$	3	5	7	9
$f(x)g(x)$	$\frac{3}{4}$	$\frac{5}{8}$	$\frac{7}{2}$	$\frac{9}{8}$

Therefore, The expected value of $g(X)$ is equal to

$$E[g(X)] = \frac{3}{4} + \frac{5}{8} + \frac{7}{2} + \frac{9}{8} = 6$$

So, the variance of $g(X) = 2X + 3$ is equal to

$$\sigma^2 = (3-6)^2 * \frac{1}{4} + (5-6)^2 * \frac{1}{8} + (7-6)^2 * \frac{1}{2} + (9-6)^2 * \frac{1}{8} = 4,$$

and the standard deviation of of $g(X)$ is equal to: $\sigma = \sqrt{4} = 2$.

Plan

1 Discrete probability distributions

2 Some Discrete probability Distributions
- Discrete Uniform Random Variable
- Binomial Distribution

3 Hypergeometric Distribution

4 Poisson Distribution

Definition (Discrete Uniform Random Variable)

A random variable X is called discrete uniform if it has a finite number of possible values, say $x_1, x_2, \ldots, x_n$ and

$$P(X = x_i) = \begin{cases} \frac{1}{n}, & \text{for all } 1 \leq i \leq n \\ 0, & \text{elsewhere.} \end{cases}$$

Note: n is called the parameter of the distribution.

Example

Experiment: tossing a balanced die.

- Sample space: $S = \{1, 2, 3, 4, 5, 6\}$.
- Each sample point of S occurs with the same probability $\frac{1}{6}$.
- Let $X =$ the number observed when tossing a balanced die.

Solution

The probability distribution of X is:

$$P(X = x) = \begin{cases} \frac{1}{6}, & \text{for all } 1 \leq x \leq 6 \\ 0, & \text{elsewhere.} \end{cases}$$

Definition (Bernouilli Process)

Strictly speaking, the Bernoulli process must possess the following properties:

1. The experiment consists of repeated trials.
2. Each trial results in an outcome that may be classified as a success or a failure.
3. The probability of success, denoted by p, remains constant from trial to trial.
4. The repeated trials are independent.

Definition (Binomial Distribution)

A Bernoulli trial can result in a success with probability p and a failure with probability $q = 1 - p$. Then the probability distribution of the binomial random variable X, the number of successes in n independent trials, is

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, \; x = 0, 1, 2 \ldots, n.$$

Example

The probability that a certain kind of component will survive a shock test is $3/4$. Find the probability that exactly 2 of the next 4 components tested survive.

Solution

Let X the number of components that will survive a shock test. Assuming that the tests are independent and $p = \frac{3}{4}$ for each of the 4 tests, then X is a binomial distribution Binomial$(4, \frac{3}{4})$ or $B(4, \frac{3}{4})$. Hence,

$$P(X = 2) = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 \approx 0.21,$$

and

$$P(X = 0) = \binom{4}{0}\left(\frac{3}{4}\right)^0\left(\frac{1}{4}\right)^4 = 0.0625,$$

and

$$P(X = 5) = 0.$$

Example

The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that

1. at least 10 survive,
2. from 3 to 8 survive,
3. exactly 5 survive?

Solution

1)

$$\begin{aligned} P(X \geq 10) &= 1 - P(X < 10) \\ &= 1 - \sum_{x=0}^{9} \binom{15}{x} (0.4)^{x} (0.6)^{15-x} \\ &= 1 - 0.9662 = 0.0338 \end{aligned}$$

2)

$$P(3 \leq X \leq 8) = \sum_{x=3}^{8} \binom{15}{x} (0.4)^x (0.6)^{15-x} = 0.8779.$$

3)

$$P(X = 5) = \binom{15}{5} (0.4)^5 (0.6)^{15-5} = 0.1859.$$

Theorem

The mean and variance of the binomial distribution $B(n, p)$ are

$$\mu = n\,p \text{ and } \sigma^2 = n\,p\,q.$$

Plan

1. Discrete probability distributions
2. Some Discrete probability Distributions
   - Discrete Uniform Random Variable
   - Binomial Distribution
3. Hypergeometric Distribution
4. Poisson Distribution

3) Hypergeometric Distribution

Definition

The probability distribution of the hypergeometric random variable X, the number of successes in a random sample of size n selected from N items of which K are labeled success and $N - K$ labeled failure, is

$$h = (x, N, n, K) = P(X = x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}.$$

Theorem

The mean and variance of the hypergeometric distribution $h(N, K, n)$ are

$$\mu = n\frac{K}{N} \text{ and } \sigma^2 = n\frac{K}{N}\left(1 - n\frac{K}{N}\right)\frac{N-n}{N-1}.$$

Example

Lots of 40 components each are deemed unacceptable if they contain 3 or more defectives. The procedure for sampling a lot is to select 5 components at random and to reject the lot if a defective is found. What is the probability that exactly 1 defective is found in the sample if there are 3 defectives in the entire lot?

Solution

Using the hypergeometric distribution with $n = 5$, $N = 40$, $k = 3$, and $x = 1$, we find the probability of obtaining 1 defective to:

$$h(1, 40, 5, 3) = \frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}} = 0.3011.$$

Theorem (Approximation)

If n is small compared to K, then a binomial distribution $B(n, p = \frac{K}{N})$ can be used to approximate the hypergeometric distribution $h(N, n, K)$.

Example

A manufacturer of automobile tires reports that among a shipment of 5000 sent to a local distributor, 1000 are slightly blemished. If one purchases 10 of these tires at random from the distributor, what is the probability that exactly 3 are blemished?

Solution

Since $K = 1000$ is large relative to the sample size $n = 10$, we shall approximate the desired probability by using the binomial distribution. The probability of obtaining a blemished tire is 0.2. Therefore, the probability of obtaining exactly 3 blemished tires is

$$h(3, 5000, 10, 1000) \approx \binom{10}{3} (0.2)^3 (0.8)^7 = 0.2013.$$

Plan

1. Discrete probability distributions
2. Some Discrete probability Distributions
   - Discrete Uniform Random Variable
   - Binomial Distribution
3. Hypergeometric Distribution
4. Poisson Distribution

Definition

Let X the number of outcomes occurring during a given time interval. X is called a Poisson random variable, with parameter λ, when its probability distribution is given by

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

where λ is the average number of outcomes.

Example

During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

Solution

Using the Poisson distribution with $x = 6$ and $\lambda = 4$, we have

$$p(6, 4) = e^{-4}\frac{4^6}{6!} = 0.1041.$$

Theorem

If a random variable X has a Poisson distribution. Then both the mean and the variance of X are λ.

$$\mu = \lambda \textit{ and } \sigma^2 = \lambda$$

Theorem (Approximation)

Let X be a binomial random variable with probability distribution $B(n, p)$. When n is large ($n \to +\infty$), and p small ($p \to 0$), then the poisson distribution can be used to approximate the binomial distribution $B(n, p)$ by taking $\lambda = np$.

Example

In a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and accidents are independent of each other.

1. What is the probability that in any given period of 400 days there will be an accident on one day?
2. What is the probability that there are at most three days with an accident?

Solution

Let X be a binomial random variable with $n = 400$ and $p = 0.005$. Thus, $np = 2$. Use the Poisson approximation,

1. $$P(X = 1) = e^{-2}2^{1} = 0.271$$
.
2. $$P(X \leq 3) = \sum_{x=0}^{3} e^{-2}\frac{2^x}{x!} = 0.857$$
.