Lecture 4. Some Discrete Probability Distribution

Vu Nguyen Son Tung

Faculty of information technology
HANU

The Bernoulli Process

An experiment often consists of repeated trials, each with two possible outcomes that may be labeled success or failure. We may choose to define either outcome as a success. The process is referred to as a **Bernoulli process**. Each trial is called a **Bernoulli trial**.

The Bernoulli Process

An experiment often consists of repeated trials, each with two possible outcomes that may be labeled success or failure. We may choose to define either outcome as a success. The process is referred to as a **Bernoulli process**. Each trial is called a **Bernoulli trial**.
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.

Binomial Distribution

The number X of successes in n Bernoulli trials is called **a binomial random variable**. The probability distribution of this discrete random variable is called **the binomial distribution**, and its values will be denoted by $B(x; n, p)$ since they depend on the number of trials and the probability of a success on a given trial.

Binomial Distribution

The number X of successes in n Bernoulli trials is called **a binomial random variable**. The probability distribution of this discrete random variable is called **the binomial distribution**, and its values will be denoted by $B(x; n, p)$ since they depend on the number of trials and the probability of a success on a given trial.

Bernoulli's Formula

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

$$B(x; n, p) = C_n^x \, p^x q^{n-x}, \qquad x = 0, 1, 2, \ldots, n.$$

Binomial Distribution

The number X of successes in n Bernoulli trials is called **a binomial random variable**. The probability distribution of this discrete random variable is called **the binomial distribution**, and its values will be denoted by $B(x; n, p)$ since they depend on the number of trials and the probability of a success on a given trial.

Bernoulli's Formula

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

$$B(x; n, p) = C_n^x \, p^x q^{n-x}, \qquad x = 0, 1, 2, \ldots, n.$$

Note:

- Since $p + q = 1$, we have $1 = (p+q)^n = \sum_{x=0}^{n} C_n^x \, p^x q^{n-x}$.

Binomial Distribution

The number X of successes in n Bernoulli trials is called **a binomial random variable**. The probability distribution of this discrete random variable is called **the binomial distribution**, and its values will be denoted by $B(x; n, p)$ since they depend on the number of trials and the probability of a success on a given trial.

Bernoulli's Formula

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

$$B(x; n, p) = C_n^x\, p^x q^{n-x}, \qquad x = 0, 1, 2, \ldots, n.$$

Note:

- Since $p + q = 1$, we have $1 = (p + q)^n = \sum_{x=0}^{n} C_n^x\, p^x q^{n-x}$.
- Frequently, we are interested in problems where it is necessary to find $P(X < k)$ or $P(a \leqslant X \leqslant b)$. Binomial sums:

  $P(X < k + 1) = B(k; n, p) = \sum_{x=0}^{k} C_n^x\, p^x q^{n-x}$ are given in the Table.

Mean and Variance

It would seem reasonable to assume that the mean and variance of a binomial random variable also depend on the values assumed by parameters n, p, and q. Indeed, this is true, and in the following theorem we derive general formulas that can be used to compute the mean and variance of any binomial random variable as functions of n, p, and q.

Mean and Variance

It would seem reasonable to assume that the mean and variance of a binomial random variable also depend on the values assumed by parameters n, p, and q. Indeed, this is true, and in the following theorem we derive general formulas that can be used to compute the mean and variance of any binomial random variable as functions of n, p, and q.

Theorem

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

$$\mu = np, \quad \sigma^2 = npq.$$

Mean and Variance

It would seem reasonable to assume that the mean and variance of a binomial random variable also depend on the values assumed by parameters n, p, and q. Indeed, this is true, and in the following theorem we derive general formulas that can be used to compute the mean and variance of any binomial random variable as functions of n, p, and q.

Theorem

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

$$\mu = np, \quad \sigma^2 = npq.$$

Areas of Application

The binomial distribution finds applications in many scientific fields. An industrial engineer is keenly interested in the "proportion defective" in an industrial process. Often, quality control measures and sampling schemes for processes are based on the binomial distribution. The binomial distribution is also used extensively for medical and military applications. In both fields, a success or failure result is important. For example, "cure" or "no cure" is important in pharmaceutical work, and "hit" or "miss" is often the interpretation of the result of firing a guided missile.

Poisson Distributions

Experiments yielding numerical values of X, the number of outcomes occurring during a given time interval or in a specified region, are called **Poisson experiments**. The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year.

For example, a Poisson experiment can generate observations for the random variable X representing the number of telephone calls received per hour by an office, or the number of days school is closed due to snow during the winter.

The specified region could be a line segment, an area, a volume, or a piece of material.

Poisson Distributions

Experiments yielding numerical values of X, the number of outcomes occurring during a given time interval or in a specified region, are called **Poisson experiments**. The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year.
For example, a Poisson experiment can generate observations for the random variable X representing the number of telephone calls received per hour by an office, or the number of days school is closed due to snow during the winter. The specified region could be a line segment, an area, a volume, or a piece of material.

Poisson Process and properties. A Poisson experiment is derived from the Poisson process, which has the properties:

1. The number of outcomes occurring in one time interval or specified region of space is independent of the number that occur in any other disjoint time interval or region.
2. The probability that a single outcome will occur during a very short time interval or in a small region is proportional to the length of the time interval or the size of the region and does not depend on the number of outcomes occurring outside this time interval or region.
3. The probability that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.

Poisson Distributions. The number X of outcomes occurring during a Poisson experiment is called a **Poisson random variable**, and its probability distribution is called the **Poisson distribution**.

Poisson Distributions. The number X of outcomes occurring during a Poisson experiment is called a **Poisson random variable**, and its probability distribution is called the **Poisson distribution**.

The probability distribution of the Poisson random variable X, representing the number of outcomes occurring in a given time interval or specified region denoted by t, is

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

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

Poisson Distributions. The number X of outcomes occurring during a Poisson experiment is called a **Poisson random variable**, and its probability distribution is called the **Poisson distribution**.

The probability distribution of the Poisson random variable X, representing the number of outcomes occurring in a given time interval or specified region denoted by t, is

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

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

Note: Table A.2 contains Poisson probability sums:

$$P(r, \lambda t) = \sum_{x=0}^{r} p(x; \lambda t),$$

for selected values of λt ranging from 0.1 to 18.0.

Poisson Distributions. The number X of outcomes occurring during a Poisson experiment is called a **Poisson random variable**, and its probability distribution is called the **Poisson distribution**.

The probability distribution of the Poisson random variable X, representing the number of outcomes occurring in a given time interval or specified region denoted by t, is

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

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

Note: Table A.2 contains Poisson probability sums:

$$P(r, \lambda t) = \sum_{x=0}^{r} p(x; \lambda t),$$

for selected values of λt ranging from 0.1 to 18.0.

Theorem

Both the mean and the variance of the Poisson distribution $p(x; \lambda t)$ are λt.

Approximation of Binomial Distribution by a Poisson Distribution

In the case of the binomial, if n is quite large and p is small (is close to 0), the Poisson distribution can be used, with $\mu = np$, to approximate binomial probabilities.

Approximation of Binomial Distribution by a Poisson Distribution

In the case of the binomial, if n is quite large and p is small (is close to 0), the Poisson distribution can be used, with $\mu = np$, to approximate binomial probabilities.
If p is close to 1, we can still use the Poisson distribution to approximate binomial probabilities by interchanging what we have defined to be a success and a failure, thereby changing p to a value close to 0.

Approximation of Binomial Distribution by a Poisson Distribution

In the case of the binomial, if n is quite large and p is small (is close to 0), the Poisson distribution can be used, with $\mu = np$, to approximate binomial probabilities.
If p is close to 1, we can still use the Poisson distribution to approximate binomial probabilities by interchanging what we have defined to be a success and a failure, thereby changing p to a value close to 0.

Theorem

Let X be a binomial random variable with probability distribution $b(x; n, p)$. When $x \longrightarrow \infty$, $p \longrightarrow 0$, and $np \longrightarrow \mu$ remains constant,

$$b(x; n, p) \longrightarrow p(x; \mu), \quad \text{as} \quad n \longrightarrow \infty.$$

THANK YOU FOR YOUR ATTENTION!