Probability Theory and Mathematical Statistics

Miskolc, 2025.

Dr. Tamás Glavosits

Lecture 5

Famous Discrete Random Variables

1. Mathematical Tools

The Generating Function of a Random Variable

Definition

Let ξ be a discrete random variable that only takes nonnegative integer values. Define the function $G_\xi : D \subseteq \mathbb{C} \to \mathbb{C}$ as

$$G_\xi(z) \doteq \sum_{k \in \mathcal{R}_\xi} p_k z^k = \sum_{k \in \mathcal{R}_\xi} \mathbb{P}(\xi = k) z^k \qquad (z \in D).$$

Properties of the Generating Function

Theorem

The most important properties of the generating function are:

- $p_k = \frac{G_\xi^{(k)}(0)}{k!}$ *for all* $k = 0, 1, 2, \dots$.
- *If the random variable* ξ *has an expectation, then it can be computed using the generating function as*

$$\mathbb{E}(\xi) = G_\xi'(1).$$

- *If there exisits varinace of random variable* ξ, *then*

$$\mathbb{D}^2(\xi) = G_\xi'(1) + G_\xi''(1) - [G_\xi'(1)]^2.$$

- *If* ξ_1 *and* ξ_2 *are independent, then*

$$G_{\xi_1+\xi_2}(z) = G_{\xi_1}(z)G_{\xi_2}(z).$$

Convolution of Distributions

Definition

Let ξ_1 and ξ_2 be independent random variables that only take positive integer values. Let

$$p_i := \mathbb{P}(\xi_1 = x_i), \qquad q_j := \mathbb{P}(\xi_2 = y_j) \qquad (i, j = 0, 1, 2, \dots)$$

Then

$$\mathbb{P}(\xi_1 + \xi_2 = k) = \sum_{i=0}^{k} p_i q_{k-i} \qquad (k = 0, 1, 2, \dots),$$

which is called the **convolution of the distributions** (p_i) **and** (q_j).

The Formula for the Sum of a Geometric Series

If $a \in \mathbb{R}$, $q \in \mathbb{R}$, $|q| < 1$, then the series

$$\sum_{k=0}^{\infty} aq^k = a + aq + aq^2 + \dots$$

is called a **geometric series**. The **formula for the sum of a geometric series** is easily obtained as

$$\sum_{k=0}^{\infty} aq^k = \lim_{n\to\infty} \sum_{k=0}^{n-1} aq^k = \lim_{n\to\infty} a\frac{1-q^{n-1}}{1-q} = \frac{a}{1-q}.$$

2. Famous Discrete Distributions

Bernoulli Distribution

Definition

The Bernoulli distribution with parameter p corresponds to tossing a coin with head probability $p \in]0, 1[$. The coin is tossed once, and the random variable ξ takes the value 1 if we get heads, and 0 if tails. Notation: $\xi \sim \mathcal{B}(p)$.

Bernoulli Distribution: Distribution, Expectation, and Variance

- The distribution of ξ is:

x_i	0	1
p_i	$1-p$	p

Theorem

If ξ is a Bernoulli random variable with parameter p, then its

- ***expectation and variance:***

$$\mathbb{E}(\xi) = p, \qquad \mathbb{D}^2(\xi) = p(1-p).$$

The Geometric Distribution

Definition

The geometric distribution has two versions. In both cases, we perform a sequence of independent trials and observe the occurrence of an event A with probability p. The random variable ξ is called a **geometric distribution with parameter p** if

Version A: ξ_A denotes the number of trials required for the first successful occurrence of event A.

Version B: ξ_B denotes the number of trials before the first successful occurrence of event A.

Geometric Distribution: Distribution, Expectation, and Variance

Theorem

Preserving the above notation, let ξ_A be a geometric random variable with respect to version A.

- ***distribution:*** *of ξ_A is*

$$\mathbb{P}(\xi_A = k) := p_k = p(1-p)^k \qquad (k = 0, 1, 2, \dots);$$

- ***expectation and variance:*** *of ξ_A is*

$$\mathbb{E}(\xi_A) = \frac{1}{p}; \qquad \mathbb{D}^2(\xi_A) = \frac{1-p}{p^2}$$

Geometric Distribution: Distribution, Expectation, and Variance

Theorem

Preserving the above notation, let ξ_B be a geometric random variable with respect to version B.

- ***distribution:*** *of ξ_B is*

$$\mathbb{P}(\xi_B = k) := p_k = p(1-p)^k \qquad (k = 0, 1, 2, \dots);$$

- ***expectation and variance:*** *of ξ_B is*

$$\mathbb{E}(\xi_B) = \frac{1}{p} - 1; \qquad \mathbb{D}^2(\xi_B) = \frac{1-p}{p^2}$$

Negative binomial distribution

Definition

We perform a series of independent experiments in which we observe the number of occurrences of an event A with probability p. Let ξ denote the number of trials required for the observed event A to occur the r-th time (where $r \in \mathbb{Z}_+$ is a fixed number). Then we say that ξ is a **negative binomial distributed random variable with parameters r and p**. Another name for the negative binomial distribution is the **Pascal distribution**.
Notation: $\xi \sim \mathcal{NB}(r, p)$.

Remark on the negative binomial distribution

Remark

1. When we talk about the number of occurrences of an event A with probability p, we should always think of tossing a coin, where
$$\mathbb{P}(\text{heads}) = p, \qquad \mathbb{P}(\text{tails}) = 1 - p.$$
So, most of the discrete random variables we examine can be derived from a Bernoulli distribution.
2. A random variable $\xi \sim \mathcal{NB}(r, p)$ is an independent sum of r geometric distributed random variables with parameter p, that is, if $\xi \sim \mathcal{NB}(r, p)$, then there exist random variables with geometric distributions $\xi_1, \xi_2, \ldots \xi_r \sim \mathrm{Geom}(p)$ (Version A) such that
$$\xi = \xi_1 + \xi_2 + \cdots + \xi_r.$$

Distribution, expected value, and variance of the negative binomial distribution

Theorem

Let $\xi \sim \mathcal{NB}(r,p)$, then

- ***the distribution of ξ is:***

$$\mathbb{P}(\xi = r+k) = \binom{r+k-1}{r-1} p^r (1-p)^k \qquad (k = 0,1,2,\dots)$$

(from we can see that $\mathcal{R}_\xi = \{r, r+1, r+2, \dots\}$).

- ***The expected value and variance of ξ is:***

$$\mathbb{E}(\xi) = \frac{r}{p}, \qquad \mathbb{D}^2(\xi) = \frac{r(1-p)}{p^2}$$

The Binomial Distribution

Definition

Let ξ denote the number of occurrences of an event A with probability p in n independent trials. Then we say that ξ has a **binomial distribution with parameters** n **and** p, denoted as $\xi \sim \mathcal{B}(n, p)$.

Remark on the Binomial Distribution

Remark

A random variable $\xi \sim \mathcal{B}(n, p)$ is the sum of n independent Bernoulli distributed random variables with p parameters, that is, there exist $\xi_1, \xi_2, \ldots \xi_n$ independent Bernoulli distributed random variables with p parameters such that

$$\xi = \xi_1 + \xi_2 + \cdots + \xi_n.$$

Binomial Distribution: Distribution, Expectation, and Variance

Theorem

If $\xi \sim \mathcal{B}(n, p)$, then

- *the **distribution** of ξ is*

$$p_k = \binom{n}{k} p^k (1-p)^{n-k} \qquad (k = 0, 1, \ldots, n);$$

- *the **expectation and variance** of ξ is*

$$\mathbb{E}(\xi) = np, \qquad \mathbb{D}^2(\xi) = np(1-p).$$

The Hypergeometric Distribution

Definition

We draw balls from an urn without replacement. Using the usual notation:

- N is the total number of balls,
- s is the number of red balls,
- $N - s$ is the number of white balls, and
- we draw n balls, among which the drown balls k are red.
- Let ξ denote the number of red balls drawn.

Then we say that the random variable ξ has a **hypergeometric distribution**.

Hypergeometric Distribution: Distribution, Expectation, and Variance

Theorem

Preserving the above notation let ξ be a hypergeometric random variable.

- *The **distribution** of ξ is*

$$p_k = \frac{\binom{s}{k}\binom{N-s}{n-k}}{\binom{N}{n}}, \quad (\max\{0, n-(N-s)\} \leqslant k \leqslant \min\{s,n\});$$

- *The **expectation and variance** of ξ is*

$$\mathbb{E}(\xi) = n\frac{s}{N}, \qquad \mathbb{D}^2(\xi) = n\frac{N-n}{N-1}\frac{s}{N}\left(1-\frac{s}{N}\right).$$

The Poisson Distribution

Definition

Let $\lambda > 0$ be a fixed real number. We say that the random variable ξ has a **Poisson distribution** with parameter λ if its distribution is

$$p_k = \mathbb{P}(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (k = 0, 1, 2, \dots).$$

Notation: $\xi \sim \mathrm{Poiss}(\lambda)$.

Poisson Distribution: Expectation and Variance

Theorem

Let $\xi \sim \mathrm{Poiss}(\lambda)$ be a random variable. The ***expectation and variance*** *of ξ is*

$$\mathbb{E}(\xi) = \lambda \qquad \mathbb{D}^2(\xi) = \lambda$$

Limit Theorems

1. The hypergeometric distribution can be approximated by the binomial for large N, that is, if ξ_k is hypergeometric with parameters N_k and s_k such that $N_k \to \infty$ and $\frac{s_k}{N_k} \to p \in]0, 1[$, moreover $n_k = n$ is fixed, then $\xi_k \sim \mathcal{B}(n, p)$.
2. The Poisson distribution can be approximated by the Binomial distribution for large n, that is, if $\xi_k \sim \mathcal{B}(n_k, p_k)$ such that $n_k \to \infty$ and $n_k p_k \to \lambda > 0$, then $\xi_k \sim \mathrm{Poiss}(\lambda)$.

End of Lecture 5