Lecture 3
Discrete Random Variables

Chao Song

College of Ecology
Lanzhou University

September 12, 2023

Motivating examples

Example 1: Let the random experiment be throwing a die. The sample space associated with this experiment is $S = \{1, 2, 3, 4, 5, 6\}$, with elements of S indicating the number of spots on the side facing up. Let X be a function such that $X(s) = s$. Now, X is a real-valued function that has the outcome space S as its domain and $\{0, 1, 2, 3, 4, 5, 6\}$ as its space.

Example 2: A rat is selected at random from a cage and its sex is determined. The sample space is thus $S = \{female, male\} = \{F, M\}$. Let X be a function that has the outcome space S as its domain and the set of real numbers $\{x : x = 0, 1\}$ as its range.

Definition of random variables

Definition: Given a random experiment with an sample space S, a function X that assigns one and only one real number $X(s) = x$ to each element s in S is called a random variable. The space of X is the set of real number $\{x : X(s) = x, s \in S\}$

Example 1: Let the random experiment be throwing a die. The sample space associated with this experiment is $S = \{1, 2, 3, 4, 5, 6\}$, with elements of S indicating the number of spots on the side facing up. Let $X(s) = s$, the space of the random variable X is $\{1, 2, 3, 4, 5, 6\}$.

Example 2: A rat is selected at random from a cage and its sex is determined. The sample space is $S = \{female, male\}$. Let X be a function such that $X(F) = 0$ and $X(M) = 1$. X is a random variable with space $\{0, 1\}$

Definition of random variables

A few remarks on the definition of random variable:

- Intuitively, we may view random variable as a quantity whose value is determined by the outcome of an random experiment. For practical purpose, this intuitive interpretation of random variable is sufficient;
- Rigorously, a random variable is a function that maps the outcome of a random experiment to real numbers. This is mainly for mathematical rigor.
- Roughly speaking, because probability is a measure mapping events to unit interval. The argument of probability is events. Thus, if we are going to define probability for random variable, we must be able to interpret $\{X \leqslant x\}$ as an event.
- How to map outcome of random experiment to a real number is not a trivial mathematical question. In practice, the choice is often made based on intuition or convenience.

Discrete random variables

Discrete random variable: a random variable is discrete if it only takes values that are in some countable subsets $\{x_1, x_2, \ldots\}$ of real number.

- Number of heads in 10 coin flips;
- Number of coin flips until we have two heads;
- Species richness in a country;
- Number of students late to this class each week.

Probability mass function

Definition: The probability mass function (PMF) of a discrete random variable X is the function $f(x) : \mathbb{R} \to [0, 1]$ given by $f(x) = P(X = x)$.

Properties of probability mass function:

- $0 \leqslant f(x) \leqslant 1$ for all x;
- $f(x) = 0$ if $x \notin \{x_1, x_2, \dots\}$;
- $\sum_x f(x) = 1$.

Probability mass function

Example: Let X be the number of heads when tossing two fair coins. What is the probability mass function for random variable X?

Answer: possible number of heads are 0, 1, 2. The PMF of X is

- $P(X = 0) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$;
- $P(X = 1) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}$;
- $P(X = 2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$;

Cumulative distribution function

Definition: The cumulative distribution function (CDF) of a random variable X is the function $F(x) : \mathbb{R} \to [0, 1]$ given by $F(x) = P(X \leqslant x)$

Properties of cumulative distribution function:

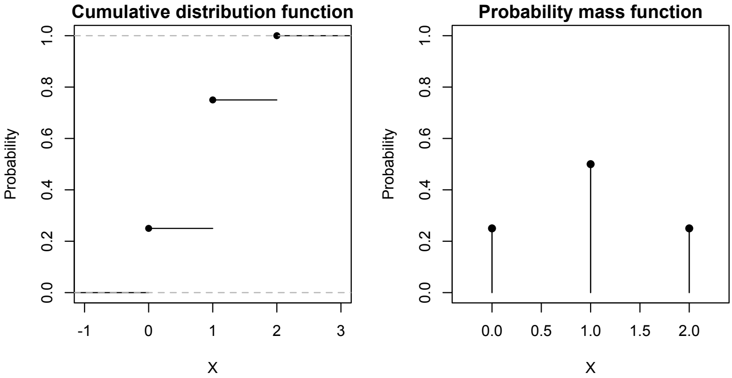
- $F(x)$ is a non-decreasing function: if $x < y$, then $F(x) \leqslant F(y)$;
- $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to -\infty} F(x) = 1$;
- $F(x)$ is right-continuous.

Proposition: Consider real numbers x and y with $x < y$, then

- $P(X > x) = 1 - F(x)$;
- $P(x < X \leqslant y) = F(y) - F(x)$;

Visualize PMF and CDF

Example: Let X be the number of heads when tossing two fair coins. Possible values for X are 0, 1, and 2. The PMF and CDF of X are:



Mathematical expectation

In addition to PMF and CDF, which fully characterize the distribution of a random variable, **mathematical expectation** is an important concept in summarizing characteristics of distribution of probability.

Definition: if $f(x)$ is the probability mass function of the discrete random variable X with space S, and if the summation $\sum_{x \in X} u(x)f(x)$ exists, then the sum is called the mathematical expectation or the expected value of $u(x)$, and it is denoted $E[u(x)]$

Mathematical expectation

Example: Let X be the number of heads when tossing two coins. What is the expected value of X? If one gets two points for each head, what is the expected value of points?

Answer: The expected value of number of heads is

$$E(X) = 0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} = 1.$$

Let $u(x)$ be the points one get after tossing two coins, $u(x) = 2x$, then

$$E[u(x)] = (2 \times 0) \times \frac{1}{4} + (2 \times 1) \times \frac{1}{2} + (2 \times 2) \times \frac{1}{4} = 2.$$

Properties of mathematical expectation

When exists, the mathematical expectation satisfies the following properties:

- If c is a constant, then $E(c) = c$;
- If c is a constant, $E[cu(x)] = cE[u(x)]$;
- if c_1 and c_2 are constant,
  $E[c_1 u_1(x) + c_2 u_2(x)] = c_1 E[u_1(x)] + c_2 E[u_2(x)]$.

The above properties arise from the fact that **mathematical expectation is a linear operation**. Thus nonlinear operations cannot be applied the same way. For example, $E(x^2) \neq [E(x)]^2$ in general.

Mean and variance

Mean and **variance** are special cases of the mathematical expectation. Let X be a discrete random variable with probability mass function $f(x)$

- Mean: $\mu = E(X) = \sum_{x \in S} x f(x)$;
- Variance: $\sigma^2 = Var(X) = E[(X - \mu)^2] = \sum_{x \in S} (x - \mu)^2 f(x)$

Mean and variance

Let X be a random variance with mean μ and variance σ. Its variance can be calculated as $\sigma^2 = E(X^2) - \mu^2$

Proof:

$$\begin{aligned}
\sigma^2 &= E[(x-\mu)^2] = E[X^2 - 2\mu X + \mu^2] \\
&= E(X^2) - 2\mu E(X) + \mu^2 \\
&= E(X^2) - \mu^2
\end{aligned}$$

Mean and variance

Properties of mean and variance: Let X be a random variable with mean μ and variance σ^2. Let a and b be constants. What is the mean and variance of $aX + b$?

Based on the property of mathematical expectation, we have

- $E(aX + b) = aE(X) + b = a\mu + b$;
- $Var(aX + b) = E[(aX + b - a\mu - b)^2] = E[a^2(X - \mu)^2] = a^2\sigma^2$

Moment

The mean x_i is the distance of that point from the origin. In mechanics, the product of a distance and its weight is called a moment, so $x_i f(x_i)$ is a moment having a moment arm of length x_i. The sum of these products would be the moment of the system of distance and weights.

Definition: For a random variable with probability mass function $f(x)$, we define $\Sigma_{x \in S}(x - a)f(x)$ as the first moment about a. More generally, we call $\Sigma_{x \in S}(x - a)^n f(x)$ the nth moment of X about a.

Moment generating function

Definition: Let X be a random variable. We define the moment generating function of X to be

$$m_X(t) = E(e^{tX})$$

Moment generating function, as its name suggests, can be used to find moments of a random variable.

$$\frac{d}{dt} m_X(t) = E(Xe^{tX}),$$

which when we evaluate at $t = 0$ becomes $E(X)$. More generally, the nth derivative of $m_X(t)$ evaluated at zero is the expected value of X^n, i.e., $m^{(n)}(0) = E(X^n)$

Moment generating function

The moment generating function determines the distribution of X.

If the space of S is $\{b_1, b_2 \ldots\}$, the moment generating function is given by the expansion

$$M(t) = e^{tb_1} f(b_1) + e^{tb_2} f(b_2) + e^{tb_3} f(b_3) + \cdots$$

Thus the coefficients of e^{tb_i} is the probability

$$f(b_i) = P(X = b_i)$$

If two random variables have two probability mass functions $f(x)$ and $g(y)$ and the same space S, and if their moment generating functions are equal:

$$e^{tb_1} f(b_1) + e^{tb_2} f(b_2) + \cdots = e^{tb_1} g(b_1) + e^{tb_2} g(b_2) + \cdots$$

It follows that $f(b_i) = g(b_i)$ must hold.

Moment generating function

Example: Suppose random variable has a probability mass function

$$f(x) = q^{x-1}p, \ x = 1, 2, 3, \ldots$$

What is the moment generating function of X? What is the mean of X?

Answer: The moment generating function of X is

$$\begin{aligned} M(t) = E(e^{tX} &= \sum_{x=1}^{\infty} e^{tx} q^{x-1} p \\ &= \frac{p}{q} \sum_{x=1}^{\infty} (qe^t)^x \\ &= \frac{p}{q} \sum_{x=1}^{\infty} (qe^t) + (qe^t)^2 + (qe^t)^3 + \cdots \\ &= \frac{p}{q} \frac{qe^t}{1 - qe^t} = \frac{pe^t}{1 - qe^t} \end{aligned}$$

Moment generating function

We use the derivatives of the moment generating function to calculate the mean:

$$M'(t) = \frac{(1 - qe^t)pe^t - pe^t(-qe^t)}{(1 - qe^t)^2}$$
$$= \frac{pe^t}{(1 - qe^t)^2}$$

Evaluating $M'(t)$ at 0, we have:

$$E(X) = M'(0) = \frac{p}{1 - q}$$