TA Lecture 04 - Expectation

April 3,7

School of Information Science and Technology,
ShanghaiTech University

上海科技大学
ShanghaiTech University

Outline

Expectation

Definition

The *expected value* (also called the *expectation* or *mean*) of a discrete r.v. X whose distinct possible values are $x_1, x_2, \cdots$ is defined by

$$E(X) = \sum_{j=1}^{\infty} x_j P(X = x_j)$$

If the support is finite, then this is replaced by a finite sum. We can also write

$$E(X) = \sum_{x} \underbrace{x}_{\text{value}} \underbrace{P(X = x)}_{\text{PMF at } x}$$

where the sum is over the support of X.

Expectation

The expected value of a sum of r.v.s is the sum of the individual expected values.

Theorem

For any r.v.s X, Y and any constant c,

$$E(X+Y) = E(X) + E(Y)$$

$$E(cX) = cE(X)$$

Expectation

Theorem

If X is a discrete r.v. and g is a function from $\mathbb{R}$ to $\mathbb{R}$, then

$$E(g(X)) = \sum_{x} g(x) P(X = x)$$

where the sum is taken over all possible values of X.

Expectation

Theorem

Let X be a nonnegative integer-valued r.v. Let F be the CDF of X, and $G(x) = 1 - F(x) = P(X > x)$. The function G is called the survival function of X. Then

$$E(X) = \sum_{n=0}^{\infty} G(n)$$

That is, we can obtain the expectation of X by summing up the survival function (or, stated otherwise, summing up tail probabilities of the distribution).

Indicator

Theorem

There is a one-to-one correspondence between events and indicator r.v.s, and the probability of an event A is the expected value of its indicator r.v. I_A:

$$P(A) = E(I_A).$$

Indicator

Let A and B be events. Then the following properties hold.

1. $(I_A)^k = I_A$ for any positive integer k.
2. $I_{A^c} = 1 - I_A$.
3. $I_{A \cap B} = I_A I_B$.
4. $I_{A \cup B} = I_A + I_B - I_A I_B$.

Indicator

- Given n events $A_1, \ldots, A_n$ and indicators $I_j, j = 1, \ldots, n$.
- $X = \sum_{j=1}^{n} I_j$: the number of events that occur
- $\binom{X}{2} = \sum_{i<j} I_i I_j$: the number of pairs of distinct events that occur
- $E(\binom{X}{2}) = \sum_{i<j} P(A_i \cap A_j)$
    - $E(X^2) = 2\sum_{i<j} P(A_i \cap A_j) + E(X)$.
    - $Var(X) = 2\sum_{i<j} P(A_i \cap A_j) + E(X) - (E(X))^2$.

Variance

Definition

The variance of an r.v. X is

$$\mathrm{Var}(X) = E(X - EX)^2 .$$

The square root of the variance is called the *standard deviation (SD)*:

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}.$$

Variance

- For any r.v. X, $\mathrm{Var}(X) = E(X^2) - (EX)^2$.
- $Var(X + c) = Var(X)$ for any constant c.
- $Var(cX) = c^2 Var(X)$ for any constant c.
- If X and Y are independent, then $Var(X + Y) = Var(X) + Var(Y)$.
- $Var(X) \geq 0$ with equality if and only if $P(X = a) = 1$ for some constant a.

PGF and Moments

Definition

The *probability generating function* (PGF) of a nonnegative integer-valued r.v. X with PMF $p_k = P(X = k)$ is the generating function of the PMF. By LOTUS, this is

$$E\left(t^X\right) = \sum_{k=0}^{\infty} p_k t^k.$$

The PGF converges to a value in $[-1, 1]$ for all t in $[-1, 1]$ since $\sum_{k=0}^{\infty} p_k = 1$ and $\left|p_k t^k\right| \leq p_k$ for $|t| \leq 1$.

PGF and Moments

Let X be a nonnegative integer-valued r.v. with PMF $p_k = P(X = k)$, and the PGF of X is $g(t) = \sum_{k=0}^{\infty} p_k t^k$, we have

- $E(X) = g'(t)|_{t=1}$
- $E(X(X-1)) = g''(t)|_{t=1}$

Probability Model: Coupon Collector

Suppose there are n types of toys, which you are collecting one by one, with the goal of getting a complete set. When collecting toys, the toy types are random (as is sometimes the case, for example, with toys included in cereal boxes or included with kids' meals from a fast food restaurant). Assume that each time you collect a toy, it is equally likely to be any of the n types. Let N denote the number of toys needed until you have a complete set. Find $E(N)$ and $Var(N)$.

Probability Model: Pattern Matching

Suppose a coin with probability p for heads is tossed repeatedly, and we obtain a sequence of H and T (H denotes Head and T denotes Tail). Let N denote the number of toss to observe the first occurrence of the pattern "HH". Find $E(N)$ and $Var(N)$.

Outline

Main Contents Recap

HW Problems

More Exercices

Problem 1

Let X have PMF

$$P(X = k) = cp^k/k \ for \ k = 1, 2, ...,$$

where p is a parameter with $0 < p < 1$ and c is a normalizing constant. We have $c = -1/log(1 - p)$, as seen from the Taylor series

$$-log(1 - p) = p + \frac{p^2}{2} + \frac{p^3}{3} + \dots$$

This distribution is called the Logarithmic distribution (because of the log in the above Taylor series), and has often been used in ecology. Find the mean and variance of X.

Problem 1 Solution

Problem 2

Let a random variable X satisfies Hypergeometric distribution with parameters w, b, n.

(a) Find $E[\binom{X}{2}]$.

(b) Use the result of (a) to find the variance of X.

Problem 2 Solution

Problem 3

Suppose there are n types of toys, which you are collecting one by one, with the goal of getting a complete set. When collecting toys, the toy types are random. Assume that each time you collect a toy, it is equally likely to be any of the n types. Let N denote the number of toys needed until you have a complete set. Find Var(N).

Problem 3 Solution

Problem 4

Suppose there are n types of toys, which you are collecting one by one, with the goal of getting a complete set. When collecting toys, the toy types are random. Assume that each time you collect a toy, independently of past types collected, it is type j with probability p_j, and $\sum_{j=1}^{n} p_j = 1$. Let N denote the number of different types of toys that appear among the first m collected toys. Find E(N) and Var(N)

Problem 4 Solution

Problem 5

People are arriving at a party one at a time. While waiting for more people to arrive they entertain themselves by comparing their birthdays. Let X be the number of people needed to obtain a birthday match, i.e., before person X arrives there are no two people with the same birthday, but when person X arrives there is a match.

Assume for this problem that there are 365 days in a year, all equally likely. By the result of the birthday problem form Chapter 1, for 23 people there is a 50.7% chance of a birthday match (and for 22 people there is a less than 50% chance). But this has to do with the median of X; we also want to know the mean of X, and in this problem we will find it, and see how it compares with 23.

Problem 5 Continued

(a) A median of an r.v. Y is a value m for which $P(Y \leq m) \geq 1/2$ and $P(Y \geq m) \geq 1/2$. Every distribution has a median, but for some distributions it is not unique. Show that 23 is the unique median of X.

Problem 5 Solution

Problem 5 Continued

(b) Show that $X = I_1 + I_2 + \cdots + I_{366}$, where I_j is the indicator r.v. for the event $X \geq j$. Then find $E(X)$ in terms of p_j's defined by $p_1 = p_2 = 1$ and for $3 \leq j \leq 366$,

$$p_j = (1 - \frac{1}{365})(1 - \frac{2}{365}) \cdots (1 - \frac{j-2}{365})$$

(c) Compute $E(X)$ numerically.

Problem 5 Solution

Problem 5 Continued

(d) Find the variance of X, both in terms of the p_j's and numerically. Hint:What is I_i^2, and what is $I_i I_j$ for $i < j$? Use this to simplify the expansion

$$X^2 = I_1^2 + \cdots + I_{366}^2 + 2\sum_{j=2}^{366}\sum_{i=1}^{j-1} I_i I_j.$$

Problem 5 Solution

Problem 6

Suppose there are n types of toys, which you are collecting one by one, with the goal of getting a complete set. When collecting toys, the toy types are random. Assume that each time you collect a toy, independently of past types collected, it is type j with probability p_j, and $\sum_{j=1}^{n} p_j = 1$. Let N denote the number of toys needed until you have a complete set. Find E(N) and Var(N).

Problem 6 Solution

Outline

BH CH4 #63

A group of n people play "Secret Santa": each puts his or her name on a slip of paper in a hat, picks a name randomly from the hat (without replacement), and then buys a gift for that person. Unfortunately, they overlook the possibility of drawing one's own name, so some may have to buy gifts for themselves (on the bright side, some may like self-selected gifts better). Assume $n \geq 2$.

(a) Find the expected number of people who pick their own names.

(b) Find the expected number of pairs of people, A and B, such that A picks B 's name and B picks A 's name (where $A \neq B$ and order doesn't matter).

(c) Denote X as the number of people who pick their own names. What is the approximate distribution of X if n is large (specify the parameter value or values)? What does $P(X = 0)$ converge to as $n \to \infty$?