Applied Statistics

Recitation #1: Expectation

March 4, 2022

Expectation

Definition

The **expected value** or **mean** or **first moment** of X is defined to be

$$E(X) = \sum_{x} x P_x(x)$$

assuming that the sum is well-defined.

We can think of the expectation as the average of a very large number of independent draws from the distribution (IID draws).

Properties of Expectation

- $E(a) = a$
- $E(aX) = aE(X)$
- $E(aX + b) = aE(X) + b$
- Linearity: $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$
- LOTUS: $E(g(x)) = \sum_x g(x) P_x(x)$

Compute the Expectation of a Bernoulli random variable.

Definition

A random variable X that only takes two values 0 (failure) and 1 (success) with $P(X = 1) = p$ has the Bernoulli distribution with parameter p.

Example: You flip a coin one time. X is the result of the coin flip. You can define "Heads" as success and "Tails" as failure (or the other way around, it does not matter) Then X follows a Bernoulli distribution with parameter p.

$$P_X(x) = p^x(1-p)^{1-x}$$

- Let $X \sim Bernoulli(p)$. $E(X) =?$

Now let's try to compute this expectation empirically in Python

1. Draw random samples X from a Bernoulli variable:
2. Compute the frequency and relative frequency of $X = 0, X = 1$.
3. Compute the mean.

Linearity of Expectation

Linearity of expectation is the property that the expected value of the sum of random variables is equal to the sum of their individual expected values, regardless of whether they are independent.

Theorem

Let $X_1, \ldots, X_n$ be a set of random variables. Then

$$E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n]$$

Prove it for the case of two discrete variables.

Question

Suppose that that everyone in the lab gets up, goes outside, and then and sits back down at random (i.e., all seating arrangements are equally likely). What is the expected number of people that return to their original seat?

Question

Suppose that that everyone in the lab gets up, goes outside, and then and sits back down at random (i.e., all seating arrangements are equally likely). What is the expected number of people that return to their original seat?

Let's compute this empirically in Python

Let $\pi = [1, \ldots, n]$ be the original seat assignment. (i.e, person 1 is on seat 1), and π' be the final seat assignment (a permutation of *pi*.

1. Sample n permutations of π at random.
2. Count the number of people who have returned to their seats (number of places i where $\pi(i) = \pi'(i)$).
3. Repeat 1000 times and compute the mean.

Question

Suppose that that everyone in the lab gets up, goes outside, and then and sits back down at random (i.e., all seating arrangements are equally likely). What is the expected number of people that return to their original seat?

1. Let X_i be a Bernoulli r.v. denoting that the i-th person returns to their seat. What is the probability distribution of X_i?
2. What is the expected number of people who return to their original seat? (Use the linearity of expectation).
3. Compare with your empirical estimate.

Binomial Distribution

Imagine you toss the fair coin 5 times. Some possible outcomes are 00000, 1000, 01100 etc. Let X be the number of successes (heads) after the 5 times. Then X follows a Binomial distribution with parameters (5, 0.5).

Definition

The binomial distribution with parameters n and p is the discrete probability distribution of the number of successes in a sequence of n independent experiments, each with a binary outcome: success (with probability p) or failure (with probability $q = 1 - p$). The pmf of the binomial distribution is

$$P_X(x) = \binom{n}{x} p^x (1-p)^{1-x} \text{ for } x = 0, \ldots, n,$$

Question

Suppose $Y \sim Binomial(8, .6)$.

1. Run a simulation with 1000 trials to estimate P(Y = 6) and $P(Y \leq 6)$.
2. Use the formula for binomial probabilities to compute $P(Y = 6)$ exactly.
3. A friend has a coin with probability .6 of heads. She proposes the following gambling game: You will toss it 10 times and count the number of heads. The amount you win or lose on k heads is given by $k^2 - 7k$. (a) Plot the payoff function. (b) Run a simulation to decide if this is a good bet.

Geometric Distribution

Imagine you toss a coin until you get heads. Let X be trial of the first success. What is the pmf of X?