Chapter 3. Discrete Random Variables

- Discrete random variable: A random variable that can only take finitely many or countably many possible values.

- Distribution: Let $\{x_1, x_2, \ldots\}$ be the possible values of X. Let

$$P(X = x_i) = p_i,$$

where $p_i \geq 0$ and

$$\sum_i p_i = 1.$$

- Tabular form:

x_i	x_1	x_2	$\cdots$
$p(x_i)$	p_1	p_2	$\cdots$

- Graphic description:

  1. Probability function $\{p_i\}$.

  2. Cumulative distribution function (cdf) F:
  $$F(x) = P(X \leq x) = \sum_{\{i:\ x_i \leq x\}} p(x_i).$$
  F is a step function with jumps at x_i and jump size p_i.

Expected Value, Variance, and Standard Deviation

Discrete random variable X. Possible values $\{x_1, x_2, \ldots\}$, and $P(X = x_i) = p_i$.

- Expectation (expected value, mean, "μ"):

$$EX \doteq \sum_i x_i P(X = x_i) = \sum_i x_i p_i.$$

What is the meaning of expected value?

THEOREM: Consider a function h and random variable $h(X)$. Then

$$E[h(X)] = \sum_i h(x_i) P(X = x_i)$$

Corollary: Let a, b be real numbers. Then $E[aX + b] = aEX + b$.

Corollary: Consider functions $h_1, \ldots, h_k$. Then

$$E[h_1(X) + \cdots + h_k(X)] = E[h_1(X)] + \cdots + E[h_k(X)].$$

- Variance ("σ^2") and Standard deviation ("σ"):

$$\text{Var}[X] \doteq E\left[(X - EX)^2\right], \quad \text{Std}[X] \doteq \sqrt{\text{Var}X}.$$

Example (interpretation of variance): A random variable X,

$$P(X = 0) = 1/2, \quad P(X = a) = 1/4, \quad P(X = -a) = 1/4.$$

Proposition: $\text{Var}[X] = 0$ if and only if $P(X = c) = 1$ for some constant c.

Proposition:

$$\text{Var}[X] = E[X^2] - (EX)^2.$$

Proposition: Let a, b be real numbers. Then $\text{Var}[aX + b] = a^2\text{Var}[X]$.

Examples

1. A used-car salesman sells 0, 1, 2, 3, 4, 5, or 6 cars each week with equal probability. Find the average, variance, and standard deviation of the number of cars sold by the salesman each week. If the commission on each car sold is $150, find the average, variance, and standard deviation of weekly commission.

2. An insurance policy pays a dollars if an event A occurs. Suppose $P(A) = p$ with $0 < p < 1$. What should the company charge as premium in order to ensure an expected profit of 0.05a?

3. Consider a random variable X with mean μ and standard deviation σ. Its standardization is

$$Y \doteq \frac{X - \mu}{\sigma}.$$

What is the mean, variance, and standard deviation of Y?

4. Consider two stocks A and B with current price $(2, 2)$. Assuming that by the end of the year, the two stocks' prices will 50-50 be either $(4, 2)$ or $(1, 3)$. You have 100 dollars and want to invest in these two stocks, say x dollars in stock A, and $100 - x$ in stock B.

(a) What is your expected return by the end of the year?

(b) Choose an x to minimize the variance of your return.

Bernoulli and Binomial Distributions

- Bernoulli random variable: A random variable X takes values in $\{0, 1\}$ such that
$$P(X = 1) = p, \quad P(X = 0) = 1 - p.$$

- Binomial random variable: A random variable X that takes values in $\{0, 1, \ldots, n\}$ such that
$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

This binominal distribution is denoted by $B(n; p)$, and we write $X \sim B(n; p)$.

- What is the physical meaning of this distribution? Binomial Experiment.

  1. The experiment consists of n identical trials.
  2. The outcome of each trial is dichotomous: success (S) or failure (F).
  3. The probability of success p. The probability of failure $q = 1 - p$.
  4. The trials are independent.
  5. The total number of success has binomial distribution $B(n; p)$.

  Remark: In the setup of Binomial trials, let Y_i be the outcome of the i-th trial such that $Y_i = 1$ means "S" and $Y_i = 0$ means "F". Then

  1. $Y_1, Y_2, \ldots, Y_n$ are independent identically distributed (iid) Bernoulli random variable with probability of success p.
  2.
  $$X = Y_1 + Y_2 + \cdots + Y_n$$

1. Suppose X is $B(n; p)$. What is the distribution of $Y = n - X$?

2. Toss a fair n times. Number of heads? Number of tails?

3. A jumbo jet has 4 engines that operate independently. Each engine has probability 0.002 of failure. At least 2 operating engines needed for a successful flight. Probability of an unsuccessful flight? (Approx. 3×10^{-7})

4. Two fair dice are thrown n times. Let X be the number of throws in which the number on the first die is smaller than that on the second die. What is the distribution of X?

5. Toss a coin n times, and get k heads. Given this, what is the probability that the first toss is heads?

EXPECTATION AND VARIANCE OF $B(n; p)$

Theorem. If X is a random variable with binomial distribution $B(n; p)$, then

$$E[X] = np$$
$$\text{Var}[X] = np(1 - p).$$

Comment on the proof. Two approaches: (1) Direct computation. (2) Write X in terms of the sum of independent Bernoulli random variables [will come back to this later on after we learn more on independent random variables].

Comments on Random Sampling

1. In a family of size 10, 60% support Republican and 40% support Democrat. Randomly sample 3 members of the family. The number of those support the Republican. Is its distribution $B(3; 0.6)$?

2. In a population of size 300 million, 60% support Republican and 40% support Democrat. Randomly sample 3 members of the population. The number of those support the Republican. Is its distribution $B(3; 0.6)$?

REMARK: Unless specified, the size of population is always much larger compared to the sample size. Samples can be regarded as independent and identically distributed. (This comment applies to general random sampling).

ESTIMATING p: VERY PRELIMINARY DISCUSSION

Illustrative example. Pick a random sample of $n = 100$ Americans, and X is the number of people support Republican. What is your estimate for the percentage of the population that support Republican?

Comment. Denote the quantity we wish to estimate by p. It is a fixed number. X has distribution $B(100; p)$. A natural estimate is to use the sample percentage

$$\hat{p} \doteq \frac{X}{100}.$$

$\hat{p}$ is a random variable. (If X happens to be 50, $\hat{p}$ takes value 0.5. If X happens to be 52, $\hat{p}$ takes value 0.52, etc).

$$E\hat{p} = p, \quad \text{Var}[\hat{p}] = \frac{p(1-p)}{n}.$$

$\hat{p}$ is unbiased and consistent (more on these later).

Maximum likelihood estimate (MLE)

Consider the same example with sample size n, and X is the number of people support Republican. Then X is $B(n; p)$.

MLE: Suppose $X = k$. What value of p makes the actual observation $(X = k)$ most likely?

Solution: Maximizing (with respect to p)

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

amounts to maximizing $k \ln p + (n - k) \ln(1 - p)$. Check that it is maximized at $p = k/n$. The MLE is

$$\hat{p} = \frac{X}{n}.$$

Remark. In general, MLE is not unbiased but consistent. More on MLE later.

- Geometric random variable. A random variable X takes values in $\{1, 2, \ldots\}$ such that
$$P(X = k) = p(1 - p)^{k-1}.$$
  We say X has geometric distribution with probability of success p.

- What is the physical meaning of this distribution?

  1. An experiment consists of a random number of trials.
  2. Each trial has dichotomous outcomes: Success (S) or Failure (F).
  3. The probability of success p. The probability of failure $q = 1 - p$.
  4. The trials are independent.
  5. The total number of trials needed for the first success has geometric distribution with probability of success p.

Theorem. Suppose a random variable X is geometrically distributed with probability of success p. Then

$$E[X] = \frac{1}{p}, \qquad \mathrm{Var}[X] = \frac{1-p}{p^2}.$$

Proof. Direct computation.

1. Suppose X is geometric with probability of success p. Find $P(X > n)$.

2. A pair of identical coins are thrown together independently. Let p be the probability of heads for each coin. What is the distribution of number of throws required for both coins to show heads simultaneously? at least one heads?

3. **Memoryless property**: Suppose X is a geometric random variable with probability of success p. Then for any n and k,

$$P(X > n + k | X > n) = P(X > k).$$

4. Toss a pair of 6-side fair dice. What is the probability that a sum of face value 7 appears before a sum of face value 4?

Consider a sequence of independent Bernoulli trials with probability of success p, and X is the number of trials needed for the first success.

MLE: Suppose $X = k$. What value of p makes the actual observation $(X = k)$ most likely?

SOLUTION: Maximizing (with respect to p)

$$P(X = k) = p(1 - p)^{k-1}.$$

amounts to maximizing $\ln p + (k - 1)\ln(1 - p)$. Check that it is maximized at $p = 1/k$. The MLE is

$$\hat{p} = \frac{1}{X}.$$

Remark. This MLE is not unbiased.

- Poisson random variable. A random variable X takes values in $\{0, 1, 2, \ldots\}$ such that

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

We say X has Poisson distribution with parameter (or, rate) λ.

1. Does this really define a probability distribution? Answer: Yes.

2. What is the physical meaning of this distribution? Answer: Limit of Binomial distributions $B(n; p)$ with $np = \lambda$, as $n \to \infty$.

Remark: Poisson approximation of Binomial $B(n; p)$ when n big, p small, and $\lambda = np < 7$.

ANOTHER STORY OF POISSON DISTRIBUTION

Consider a unit time interval, and X the number of certain events that occur during this interval.

1. The occurrence of the event is independent on non-overlapping intervals.

2. On each interval of length s (small),

$$P(1 \text{ event occurs}) \approx \lambda s$$
$$P(\text{more than 1 event occurs}) \approx 0$$
$$P(0 \text{ event occurs}) \approx 1 - \lambda s.$$

3. Divide the interval into n equal-length subintervals.

4. Total number of occurrences is approximately $B(n; \lambda/n)$.

5. Let n go to infinity.

Total number of occurrences is Poisson with parameter (rate) λ.

1. Suppose in a post office customers come in randomly with rate 10 per hour. The office will be open for 8 hours a day, and 5 days a week. What is the distribution of the total number of customers in (a) an hour; (b) a day; (c) a week?

2. The probability of hitting a target with a single shot is 0.002. In 3000 shots, what is the probability of hitting the target more than two times? (Use Poisson approximation)

EXPECTATION AND VARIANCE OF GEOMETRIC DISTRIBUTION

Theorem. Suppose X has Poisson distribution with parameter λ. Then

$$E[X] = \lambda, \qquad \text{Var}[X] = \lambda.$$

Proof. Direct computation.

1. A student observes a Poisson random variable X with rate λ. If the outcome is x, then the student toss a coin x times where the coin has probability p to be heads. Let Y be the number of total heads.

   (a) Find the distribution of Y.

   (b) Find the distribution of $X - Y$.

   (c) Are Y and $X - Y$ independent?