Chapter 3. Discrete Random Variables

Review

- 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
$$\sum_i p_i = 1.$$
- Graphic description: (1) probability function $\{p_i\}$, (2) cdf F.
- Tabular form:

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

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"):

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

Example: A player bet an amount against the house. With probability p, he will win and double-up. Otherwise lose his bet. For what value of p it is a fair game? When it favors the house? (The intuitive interpretation of expected value, a hint on *Law of Large Numbers*)

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"):

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

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

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

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

Proposition:

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

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

Examples

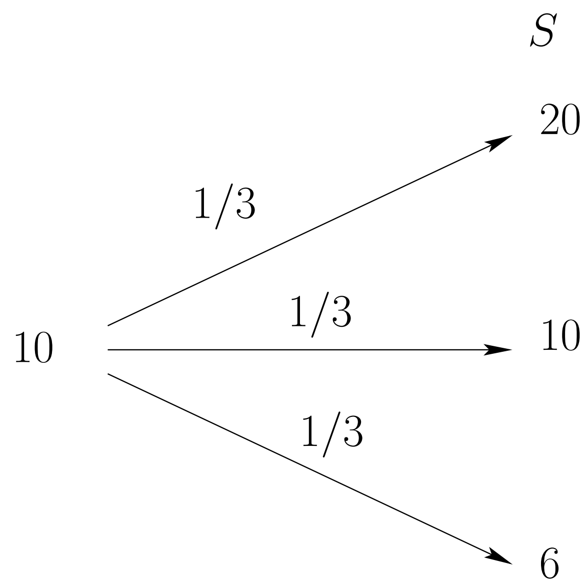
1. Suppose one word is randomly selected from the sentence THE GIRL PUT ON HER BEAUTIFUL RED HAT. X denote the number of letters in the word that is selected.

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

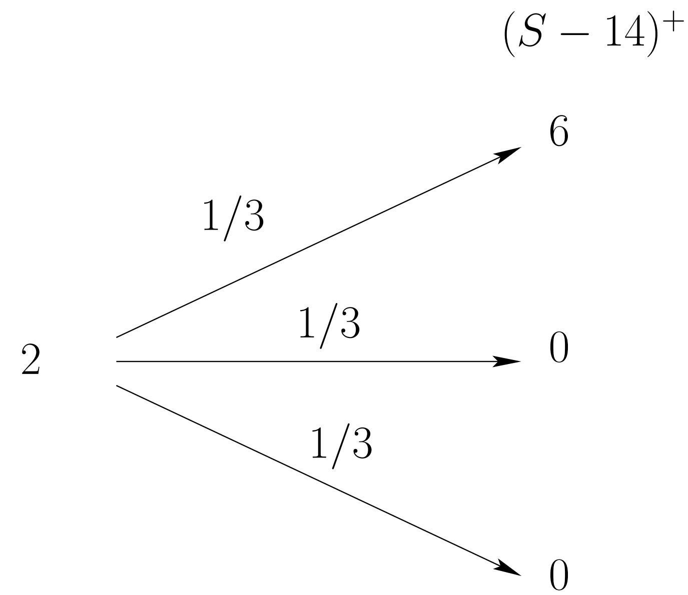
$$Y \doteq \frac{X - \mu}{\sigma}.$$

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

3. (Portfolio Optimization) Consider a stock whose price today is \$10. At the end of year, its price S has the following distribution.



Stock price



Option

A stock call option is priced at \$2 today, with payoff $(S - 14)^+$. You have in total \$10K for investment. The goal is to minimize the variance of your portfolio at the end of the year while the expected return is at least 20%.

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 $X \sim B(n; p)$.

1. Does this really define a probability distribution? Answer: Yes.
2. What is the physical meaning of this distribution? Binomial Experiment.
   (a) The experiment consists of n identical trials.
   (b) The outcome of each trial is dichotomous: success (S) or failure (F).
   (c) The probability of success p. The probability of failure $q = 1 - p$.
   (d) The trials are independent.
   (e) The total number of success has binomial distribution $B(n; p)$.

Example

1. Toss a fair n times. Number of heads? Number of tails?
2. 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})
3. Toss a coin n times, and get k heads. Given this, what is the probability that the first toss is heads?

Some 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).

Expectation and Variance of $B(n; p)$

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

$$\begin{aligned} E[X] &= np \\ \mathrm{Var}[X] &= np(1-p). \end{aligned}$$

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].

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).

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

$\hat{p}$ 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 distribution

- 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.

1. Does this really define a probability distribution? Answer: Yes.
2. What is the physical meaning of this distribution?
   (a) An experiment consists of a random number of trials.
   (b) Each trial has dichotomous outcomes: Success (S) or Failure (F).
   (c) The probability of success p. The probability of failure $q = 1 - p$.
   (d) The trials are independent.
   (e) The total number of trials needed for the first success has geometric distribution with probability of success p.

Expectation and Variance of Geometric Distribution

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.

Examples

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

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

2. 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?

Maximum likelihood estimate (MLE)

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 distribution

- 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 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

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

1. Split the unit interval into n intervals of equal length.
2. The occurrence of the event is independent from interval to interval.
3. On each interval of length s (small),

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

4. Total number of events is approximately $B(n; \lambda/n)$.
5. Let n go to infinity.

Expectation and Variance of Geometric Distribution

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

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

Proof. Direct computation.

Example

1. The number of automobile accidents at a certain intersection in a month has approximately a Poisson distribution with a mean 4. What is the probability that no accidents occur in a week? (1 month = 4 weeks)

2. (Winning State Lottery) Suppose 1 out of 10,000 tickets is a winning ticket in a state lottery. How many tickets one needs to purchase so that he/she has at least 50% chance to have a winning ticket? (Use Poisson approximation)