Chapter 4: Random Variables

Xiugang Wu

University of Delaware

Spring, 2020

Outline

- Random Variables
- Discrete Random Variables
- Expected Value
- Expectation of a Function
- Variance
- Bernoulli and Binomial Random Variables
- Poisson Random Variables
- Geometric Random Variables

Outline

- Random Variables
- Discrete Random Variables
- Expected Value
- Expectation of a Function
- Variance
- Bernoulli and Binomial Random Variables
- Poisson Random Variables
- Geometric Random Variables

Random Variables

Frequently, we are mainly interested in some function of the outcome rather than the outcome itself. These real-valued functions defined on the sample space are known as random variables (RV's).

Example: Toss three fair coins. Let Y denote the number of heads happening. Then Y is a RV taking one of the values 0, 1, 2, 3 with respective probabilities:

$$P(Y = 0) = P(\{\text{TTT}\}) = \frac{1}{8}$$

$$P(Y = 1) = P(\{\text{TTH, THT, HTT}\}) = \frac{3}{8}$$

$$P(Y = 2) = P(\{\text{HHT, HTH, THH}\}) = \frac{3}{8}$$

$$P(Y = 3) = P(\{\text{HHH}\}) = \frac{1}{8}$$

and

$$\sum_{i=0}^{3} P(Y = i) = 1.$$

Random Variables

Example: Independent trials of flipping a coin with probability p of taking up heads are continuously performed until a heads occurs. Let X denote the number of times the coin is flipped. Then X is a RV taking values on $\{1, 2, 3, 4, \ldots\}$. We have

$$
\begin{aligned}
P(X = 1) &= P(\{\text{H}\}) = p \\
P(X = 2) &= P(\{\text{TH}\}) = (1 - p)p \\
P(X = 3) &= P(\{\text{TTH}\}) = (1 - p)^2 p \\
&\vdots \\
P(X = n) &= P(\{\text{TT}\cdots\text{TH}\}) = (1 - p)^{n-1} p
\end{aligned}
$$

and

$$\sum_{n=1}^{\infty} P(X = n) = \sum_{n=1}^{\infty} (1 - p)^{n-1} p = \frac{p}{1 - (1 - p)} = 1.$$

Outline

- Random Variables
- Discrete Random Variables
- Expected Value
- Expectation of a Function
- Variance
- Bernoulli and Binomial Random Variables
- Poisson Random Variables
- Geometric Random Variables

Discrete Random Variables

A RV that can take on at most a countable number of possible values is said to be discrete.

For a discrete RV X, the probability mass function (PMF) of X, denoted by $p_X(a)$ or simply $p(a)$, is defined as

$$p(a) = P(X = a).$$

Example: If X must assume one of the values $x_1, x_2, x_3, \ldots$, then we have $p(x_i) \geq 0$, for $i = 1, 2, \ldots$ and $p(x) = 0$ for all the other values of x, and

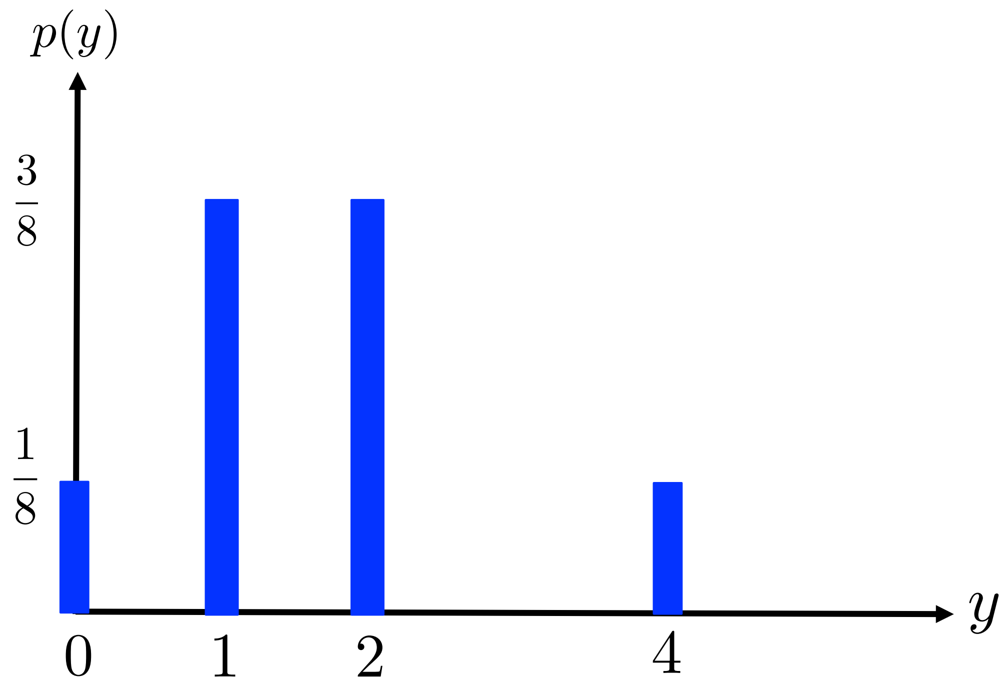
$$\sum_{i=1}^{\infty} p(x_i) = 1.$$

Example

Example: Consider a RV Y with

$$Y = \begin{cases} 0 & \text{with probability } \frac{1}{8} \\ 1 & \text{with probability } \frac{3}{8} \\ 2 & \text{with probability } \frac{3}{8} \\ 4 & \text{with probability } \frac{1}{8} \end{cases}$$

The PMF of Y can be demonstrated using the following diagram:



Example

Problem: The PMF of a RV X is given by

$$p(x) = \frac{c\lambda^x}{x!}, \; x = 0, 1, 2, \ldots$$

where λ is some positive number, and $p(x) = 0$ for other values. 1) Find the value of c; 2) Find $P(X = 0)$ and $P(X > 2)$.

Solution: 1) Since $p(x)$ is a PMF, we have

$$1 = \sum_{x=1}^{\infty} p(x) = \sum_{x=1}^{\infty} \frac{c\lambda^x}{x!} = c \sum_{x=1}^{\infty} \frac{\lambda^x}{x!} = ce^{\lambda}$$

and therefore $c = e^{-\lambda}$.

2)

$$P(X = 0) = p(0) = \frac{e^{-\lambda}\lambda^0}{0!} = e^{-\lambda}$$

and

$$P(X > 2) = \sum_{x=3}^{\infty} p(x) = 1 - \sum_{x=0}^{2} p(x) = 1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2 e^{-\lambda}}{2}$$

Cumulative Distribution Function

The cumulative distribution function (CDF) of a discrete RV X, denoted by $F_X(a)$ or simply $F(a)$, can be expressed in terms of PMF $p(x)$ by

$$F(a) = \sum_{x \le a} p(x).$$

Later we will generalize CDF to continuous RV's. Indeed, note that PMF only exists for discrete RV's, but CDF can be defined for any RV.

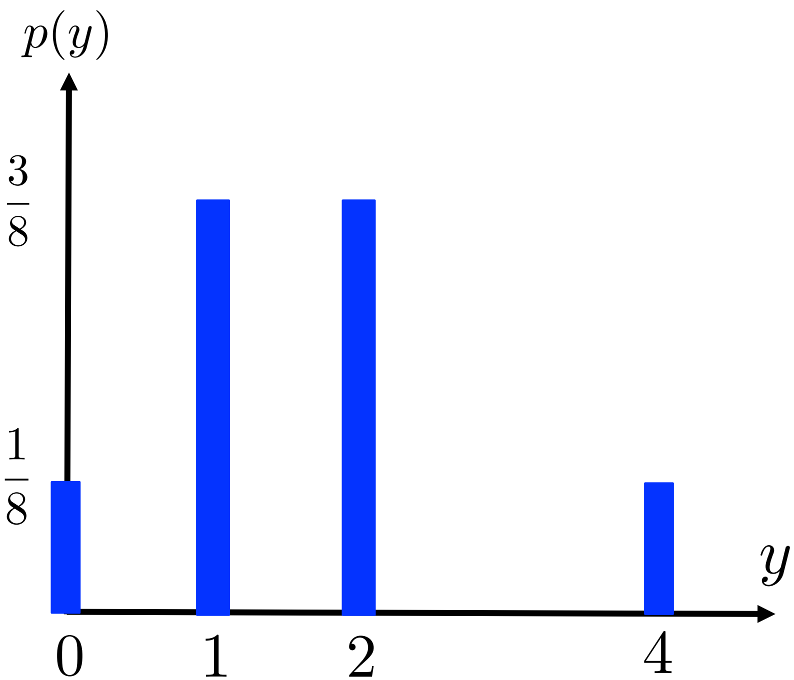
If Y is a discrete RV, then its CDF is a step function. For example, if Y has PMF given by

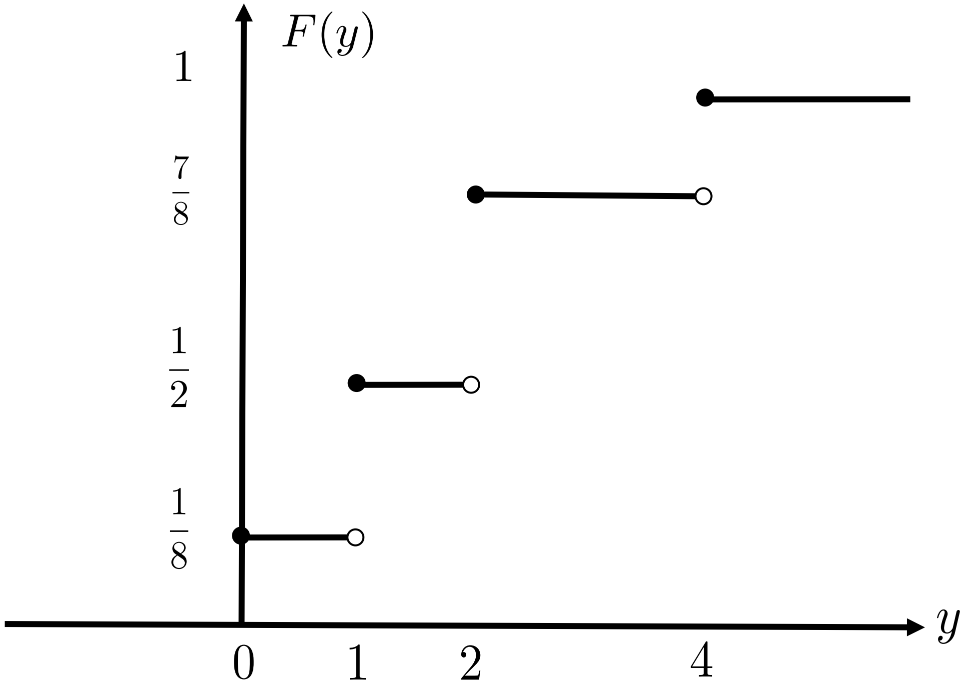
$$p(0) = \frac{1}{8},\ p(1) = \frac{3}{8},\ p(2) = \frac{3}{8},\ p(4) = \frac{1}{8}$$

then its CDF is given by

$$F(a) = \begin{cases} 0 & a < 0 \\ \frac{1}{8} & 0 \le a < 1 \\ \frac{1}{2} & 1 \le a < 2 \\ \frac{7}{8} & 2 \le a < 4 \\ 1 & a \ge 4 \end{cases}$$

Cumulative Distribution Function





$F(y)$ is a right-continuous but not left-continuos function.

Outline

- Random Variables
- Discrete Random Variables
- Expected Value
- Expectation of a Function
- Variance
- Bernoulli and Binomial Random Variables
- Poisson Random Variables
- Geometric Random Variables

Expected Value

If X is a discrete RV having a PMF $p(x)$, then the expectation or expected value of X is

$$\mathrm{E}[X] = \sum_{x:\ p(x)>0} x \cdot p(x)$$

For example, if the PMF of X is $p(0) = p(1) = \frac{1}{2}$, then

$$\mathrm{E}[X] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}.$$

If the PMF of X is $p(0) = \frac{1}{3}, p(1) = \frac{2}{3}$, then

$$\mathrm{E}[X] = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}.$$

The concept of expectation is analogous to the physical concept of the center of mass. Think of it in diagram!

Example

Suppose X is the outcome when we roll a fair die. Then the expectation of X is

$$\mathrm{E}[X] = \sum_{i=1}^{6} i \cdot p(i) = \frac{1}{6} \sum_{i=1}^{6} i = \frac{7}{2}$$

Let an indicator variable for the event A be defined as

$$I_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs} \end{cases}$$

Then the expectation of I_A is

$$\mathrm{E}[I_A] = 1 \cdot p(1) + 0 \cdot p(0) = P(A)$$

Outline

- Random Variables
- Discrete Random Variables
- Expected Value
- Expectation of a Function
- Variance
- Bernoulli and Binomial Random Variables
- Poisson Random Variables
- Geometric Random Variables

Expectation of Function

Given a RV X with PMF $p(x)$, we can also calculate the expectation of some function g of X. In particular, we have

$$\mathrm{E}[g(X)] = \sum_{x:\ p(x) \geq 0} g(x) \cdot p(x)$$

Corollary: If a, b are constant and X is a RV, then $\mathrm{E}[aX + b] = a\mathrm{E}[X] + b$.

Example: Suppose X has PMF given by $p(0) = p(1) = \frac{1}{2}$, and $g(x) = x^2$. Then

$$\mathrm{E}[g(X)] = g(0) \cdot p(0) + g(1) \cdot p(1) = \frac{1}{2}$$

In general, the expectation

$$\mathrm{E}[X^n] = \sum_{x:\ p(x) \geq 0} x^n \cdot p(x)$$

is called the n-th moment of X. In particular, the expectation of X is also called the first moment, or the mean of X.

Outline

- Random Variables
- Discrete Random Variables
- Expected Value
- Expectation of a Function
- Variance
- Bernoulli and Binomial Random Variables
- Poisson Random Variables
- Geometric Random Variables

Variance

Two RV's can have the same expectation, but behave (quite) differently in their distributions. For example, consider RV's Y and Z, where the pmf for Y is $p_Y(1) = p_Y(-1) = 0.5$ and the pmf for Z is $p_Z(100) = p_Z(-100) = 0.5$. Both Y and Z have the same expectation, i.e. $\mathrm{E}[Y] = \mathrm{E}[Z] = 0$, but obviously they have very different spread of distributions.

To measure the spread of the distribution of X, we will consider the average squared deviation of X from its mean $\mathrm{E}[X] = \mu$, and call this measure the variance of X, i.e.,

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

The square root of $\mathrm{Var}(X)$ is called the standard deviation of X, i.e.,

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

Fact: Note that we have

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

Also, $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$ because

$$\mathrm{Var}(aX + b) = \mathrm{E}[(aX + b - \mathrm{E}[aX + b])^2] = \mathrm{E}[a^2(X - \mathrm{E}[X])^2] = a^2\mathrm{Var}(X).$$

Example

For the previous example of Y and Z, we have

$$\mathrm{Var}(Y) = \mathrm{E}[Y^2] - \mathrm{E}[Y]^2 = \mathrm{E}[Y^2] = (-1)^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{2} = 1$$

and

$$\mathrm{Var}(Z) = \mathrm{E}[Z^2] - \mathrm{E}[Z]^2 = \mathrm{E}[Z^2] = (-100)^2 \cdot \frac{1}{2} + 100^2 \cdot \frac{1}{2} = 10000.$$

Another example: If X denotes the outcome of rolling a fair die, then $\mathrm{E}[X] = \frac{7}{2}$ and

$$\mathrm{E}[X^2] = \sum_{i=1}^{6} i^2 p(i) = \frac{91}{6},$$

and therefore

$$\mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

Outline

- Random Variables
- Discrete Random Variables
- Expected Value
- Expectation of a Function
- Variance
- Bernoulli and Binomial Random Variables
- Poisson Random Variables
- Geometric Random Variables

Bernoulli and Binomial RV

Bernoulli RV (success and failure):

$$X = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1-p \end{cases}$$

PMF: $p(1) = p$ and $p(0) = 1 - p$.

Expectation: $\mathrm{E}[X] = p$. Second Moment: $\mathrm{E}[X^2] = p$.

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

Binomial RV: Now suppose that we do n independent trials (each time being success w.p. p and failure w.p. $1-p$). Let X represent the number of successes in the n trials. Then X is said to be a binomial RV with parameter (n, p), denoted by $X \sim \mathrm{Binomial}(n, p)$; in particular the case of $(1, p)$ is Bernoulli.

PMF: $p(i) = \binom{n}{i} p^i (1-p)^{n-i}$ where $0 \le i \le n$; the shape of PMF is interesting

Expectation: $\mathrm{E}[X] = np$. Variance: $\mathrm{Var}(X) = np(1-p)$.

Outline

- Random Variables
- Discrete Random Variables
- Expected Value
- Expectation of a Function
- Variance
- Bernoulli and Binomial Random Variables
- Poisson Random Variables
- Geometric Random Variables

Poisson and Geometric RV

Poisson RV: A RV X taking on one of the values $0, 1, 2, \ldots$, is said to be a Poisson RV with parameter $\lambda > 0$ if its PMF is given by

$$p(i) = e^{-\lambda}\frac{\lambda^i}{i!}, i = 0, 1, 2, \ldots$$

The parameter λ affects the shape of PMF for $X \sim \text{Poisson}(\lambda)$

Expectation: $\mathrm{E}[X] = \lambda$. Second moment: $\mathrm{E}[X^2] = \lambda^2 + \lambda$.

Variance: $\mathrm{Var}(X) = \lambda$.

Geometric RV: Perform independent trials until a success occurs. The number of times we do the trials is a Geometric RV.

PMF: $p(i) = p(1-p)^{i-1}$

Expectation: $\mathrm{E}[X] = \frac{1}{p}$. Variance: $\mathrm{Var}(X) = \frac{1-p}{p^2}$.