Chapter 2

Discrete Random Variables

Peng-Hua Wang

Graduate Institute of Communication Engineering

National Taipei University

Chapter Contents

2.1 Basic Concepts

Concepts

- For an experiment, a **random variable** is a particular number associated with each outcome.
- "Mathematically, a random variable is a **real-valued function** of the experimental outcome."
- We can assign probabilities to values of a random variable.
- When do we use random variables?
  - Outcomes are numerical: dice roll, stock prices, ...
  - Outcomes are not numerical, but associated with some numerical values: average grade point of randomly selected student, ...

Example

- A sequence of 3 tosses of a coin
- The outcomes are $\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$. These 3-long sequences of heads and tails are not random variables. (why?)
- The number of heads in the sequence is a random variable.
  - Let X be The number of heads. We have

$$X(HHH) = 3, \quad X(HHT) = 2, \quad X(HTH) = 2, \quad X(HTT) = 1$$

$$X(THH) = 2, \quad X(THT) = 1, \quad X(TTH) = 1, \quad X(TTT) = 0$$

$$\mathrm{P}(X = 3) = \mathrm{P}(HHH)$$

$$\mathrm{P}(X = 2) = \mathrm{P}(HHT) + \mathrm{P}(HTH) + \mathrm{P}(THH)$$

Example

- Deterministic function of a random variable is also a random variable.
  - Let $Y = X^2$ is the square function of X.
  - $\mathrm{P}(Y = 4) = \mathrm{P}(X = 2) = \mathrm{P}(HHT) + \mathrm{P}(HTH) + \mathrm{P}(THH)$

More concepts

- Random variables are real-valued functions of the experimental outcome.
- Deterministic functions of a random variable are also random variables.
- Each random variable can be associated with certain "averages", such as the mean and the variance.
- A random variable can be **conditioned** on an event or on another random variable.
- We can define **independence** between random variables.

More concepts

- If the range of a random variable (all values that it can take) is either finite or countably infinite, it is called a **discrete** random variables.
- If the range of a random variable is uncountably infinite, it is not discrete.
  - Select a number a from the interval $[0, 1]$.
  - The random variable $X = a^2$ is not discrete.
  - The random variable

$$Y = \begin{cases} 1, & a \geq 0.5 \\ 0, & a < 0.5 \end{cases}$$

    is discrete.
- We focus on discrete random variables in this chapter.

2.2 Probability Mass Functions

PMFs

- For a discrete random variable X, the **probability mass function (PMF)** $p_X(x)$ is the probability of the event $\{X = x\}$:

$$p_X(x) = \mathrm{P}(\{X = x\})$$

- For example, toss of two fair coins. Let X be the number of heads obtained. The PMF of X is

$$\begin{aligned} p_X(0) &= 1/4 \\ p_X(1) &= 1/2 \\ p_X(2) &= 1/4 \\ p_X(x) &= 0, \quad \text{otherwise.} \end{aligned}$$

Basic Properties

- Let $\mathcal{S}$ be the set of all possible values of a random variable X.

$$\sum_{x\in\mathcal{S}} p_X(x) = 1$$

- Let $\mathcal{A}$ be a set of some values of a random variable X.

$$\mathrm{P}(x \in \mathcal{A}) = \sum_{x\in\mathcal{A}} p_X(x)$$

- For example, toss of two fair coins. Let X be the number of heads.

$$\mathrm{P}(X > 0) = p_X(1) + p_X(2) = 3/4.$$

Bernoulli Random Variables

- Bernoulli random variable X, $X = 0$ or 1 is defined by

$$p_X(0) = 1 - p, \quad p_X(1) = p, \quad 0 \leq p \leq 1.$$

- We can use Bernoulli rv for modeling a coin toss. p is the probability of head. $X = 1$ means a head obtained.

Binomial Random Variables

- Binomial random variable X, $X = 0, 1, \ldots n$ is defined by

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n, \quad 0 \le p \le 1.$$

- We can use binomial rv for modeling the number of heads in n coin tosses. p is the probability of head. $X = k$ means k heads obtained.
- $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$

Geometric Random Variables

- Toss a coin repeatedly. Let X be the number of tosses until a head comes up. The PMF of X is

$$p_X(k) = p(1-p)^{k-1}, \quad k = 1, 2, \ldots \quad 0 \le p \le 1.$$

- X is called a geometric rv.
- $\sum_{k=1}^{\infty} p_X(k) = \sum_{k=1}^{\infty} p(1-p)^{k-1} = 1.$

Poisson Random Variables

- Let λ be the average typos per n words, or the "typo rate." Then $p = \lambda/n$ be the "type probability." Let X be the number of typos in n words. We know that X is a binomial rv.

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- If n is large but λ remains fixed (i.e., p is very small), we can prove that

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \to e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, ...$$

- This is called the Poisson random variable.

Poisson Random Variables

- $\sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = 1$
- We can use Poisson rv for modeling
  - The number of miss-spelled words.
  - The number of cars involved in accidents in a city on a given day

2.3 Functions of Random Variables

Functions of Random Variables

- Let X is a random variable. We can generate another random variable Y by transform $Y = g(X)$
- If X is a random variable, then $Y = g(X)$ is also a random variable. We can calculate the PMF of Y from the PMF of X.

$$p_Y(y) = \sum_{\{x|y=g(x)\}} p_X(x)$$

Example 2.1

- Let X is a uniform random variable. $X = -4, -3, ...3, 4$. Find the PMF of X.
- Let $Y = |X|$. Find the PMF of Y.
- Let $Z = X^2$. Find the PMF of Z.

2.4 Expectation, Mean, and Variance

Mean

- The PMF of a random variable X provides us with all information about X.
- If we want to obtain a summary of X, we can use the **expected value** or called the **mean** of X. Defined by

$$E[X] = \sum_x x p_X(x).$$

- The expected value is a weighted average of all possible values of X. The weighting coefficients are the corresponding probability.

Mean

- The means of some random variables do not exists, or more precise, are not well-defined.
- The mean is well-defined if

$$\sum_x |x| p_X(x) < \infty.$$

- Example: $p_X(2^k) = 2^{-k}, k = 1, 2, ...$
- Example: $p_X(2^k) = p_X(-2^k) = 2^{-k}, k = 2, 3, ...$ This PMF is symmetric.

Example 2.2

- Two independent coin tosses, each with a $3/4$ probability of a head.
- Let X be the number of heads obtained.
- Binomial random variable with parameters $n = 2$ and $p = 3/4$
- Find the PMF of X.
- Find $E[X]$

Mean of Function of a Random Variable

- Let $Y = g(X)$ where X and Y are random variables.
- $E[X] = \sum_x x p_X(x)$.
- $E[Y] = E[g(X)] = \sum_x g(x) p_X(x)$.
- We do not need to calculate the PMF of Y.
- Example.
  - Let X is a uniform random variable.
    $X = -4, -3, ...3, 4$. Find $E[X]$.
  - Let $Y = |X|$. Find $E[Y]$.
  - Let $Z = X^2$. Find $E[Z]$.

Optimality of Mean

- Let X is a random variable with PMF $p_X(x)$.
- We want to find a number c to summarize X. That is, the error between c and the values of X should be minimized.
- We can use squared difference between c and values of X as a measure of error.
- That is, we should find a constant c to minimize $E[(X-c)^2]$.

Optimality of Mean

- The answer $c = E[X]$. (proof).
- The corresponding minimized error $E[(X - E[X])^2]$ is called the **variance** of X.
- That is, $E[x]$ is the minimized-mean-squared estimate (MMSE) of X which have the minimized mean-squared error (MSE).

Variance

- Definition.

$$\mathrm{var}(X) = E[(X - E[X])^2]$$

- Standard deviation

$$\sigma_X = \sqrt{\mathrm{var}(X)}$$

- In general, $E[X^n]$ is called the nth moment of X. Mean is the first moment.

Example 2.3

- Let X is a uniform random variable. $X = -4, -3, ...3, 4$. Find $\text{var}[X]$.

Properties

- If, $Y = aX + b$, then $E[Y] = aE[X] + b$ and $\text{var}(Y) = a^2\text{var}(X)$
- $\text{var}(X) = E[X^2] - (E[X])^2$

Example 2.4. Average Speed

- If the weather is good (with probability 0.6), Alice walks the 2 km to class at a speed of $V = 5$ km/hour, and otherwise rides her motorcycle at a speed of $V = 30$ km/hour.
- What is the mean of the time T to get to class?
- $E[T] = E[2/V] \neq E[2/V]$

Example 2.5. Mean and Variance of the Bernou

- Tossing a coin which comes up a head with probability p and a tail with probability $1 - p$. Let X be the associated rv.
- Find $E[X]$, $E[X^2]$ and $\mathrm{var}(X)$

Example 2.6. Discrete Uniform RV

- What is the mean and variance associated with a roll of a fair six-sided die?

Example 2.7. The Mean of the Poisson

- Find $E[X]$ of a Poisson rv with pmf

$$P_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, \cdots$$

2.5 Joint PMFs of Multiple Random Variables

Definition

- Two discrete random variables X and Y associated with the same experiment.
- If (x, y) is a pair of possible values of X and Y, the probability mass function (x, y) is the probability of the event $\{X = x, Y = y\}$:

$$\begin{aligned} P_{X,Y}(x, y) &= \mathrm{P}(\{X = x, Y = y\}) \\ &= \mathrm{P}(\{X = x\} \cap \{Y = y\}) \\ &= \mathrm{P}(\{X = x\} \text{ and } \{Y = y\}) \end{aligned}$$

Marginal PMFs

- We can calculate the PMFs of X and Y by

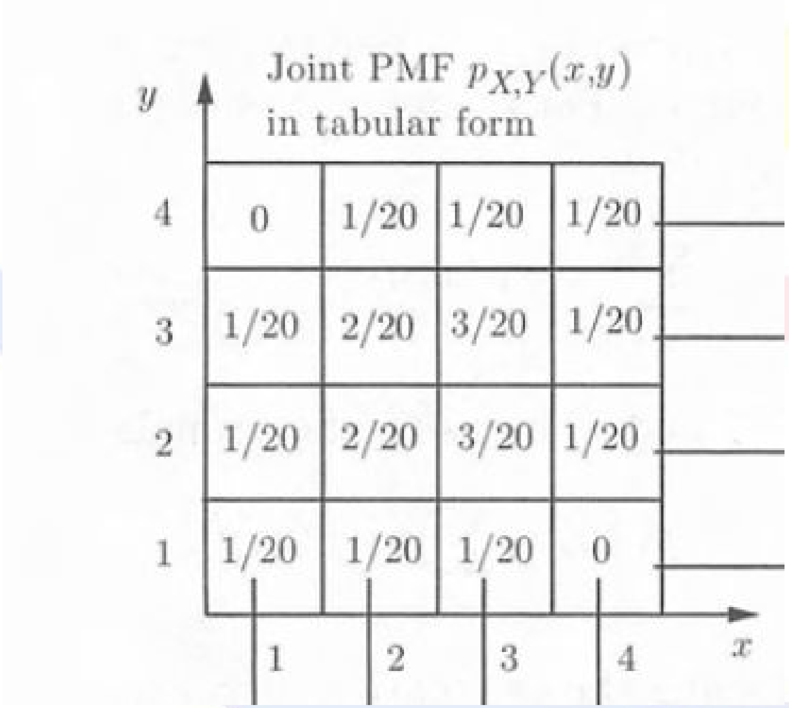
$$P_X(x) = \sum_y p_{X,Y}(x,y)$$

$$P_Y(y) = \sum_x p_{X,Y}(x,y)$$

- We sometimes refer to $P_X(x)$ and $p_Y(y)$as the marginal PMFs.

Example 2.9

The following figure show the joint PMF of two RVs X and Y. Find their marginal PMFs.



$y \backslash x$	1	2	3	4
4	0	1/20	1/20	1/20
3	1/20	2/20	3/20	1/20
2	1/20	2/20	3/20	1/20
1	1/20	1/20	1/20	0

Functions of Multiple RVs

- Functions of Multiple RVs

$$Z = g(X,Y) \Rightarrow p_Z(z) = \sum_{\{(x,y)|z=g(x,y)\}} p_{X,Y}(x,y)$$

- Mean of Functions of Multiple RVs

$$E[g(X,Y)] = \sum_x \sum_y g(x,y) p_{X,Y}(x,y)$$

- Linearity

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$

Example 2.9

The following figure show the joint PMF of two RVs X and Y. Find the PMF of $Z = X + 2Y$ and $E[Z]$

Joint PMF $p_{X,Y}(x,y)$ in tabular form

y \ x	1	2	3	4
4	0	1/20	1/20	1/20
3	1/20	2/20	3/20	1/20
2	1/20	2/20	3/20	1/20
1	1/20	1/20	1/20	0

More than Two Random Variables

- Joint PMF

$$p_{X_1,X_2,...X_n}(x_1, x_2, ...x_n) = \mathrm{P}(\{X_1 = x_1, X_2 = x_2, ...X_n = x_n\})$$

- Linearity

$$E[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \cdots + a_nE[X_n]$$

Example 2.10. Mean of the Binomial RV

- Y is a binomial RV with parameters (n, p). Then Y is the sum of n Bernoulli RVs

$$Y = X_1 + X_2 + \cdots + X_n$$

where $\mathrm{P}(X = 1) = p$.

- The mean of Y

$$E[Y] = E[X_1] + E[X_2] + \cdots + E[X_n] = np.$$

Example 2.11. The Hat Problem.

- n people throw their hats in a box and then each picks one hat at random.
- Each hat can be picked by only one person, and each assignment of hats to persons is equally likely. (independence)
- What is the expected value of X, the number of people that get back their own hat?
- Let a random variable $X_i = 1$ if the ith person selects his/her own hat, and $X_i = 0$ otherwise.

$$\mathrm{P}(X_i = 1) = 1/n, \quad E[X_i] = 1/n$$

- $X = X_1 + X_2 + \cdots + X_n$

2.6 Conditioning

Conditioning an RV on an Event

- The conditional PMF of a random variable X, conditioned on a particular event A with $\mathrm{P}(A) > 0$, is defined by

$$p_{X|A}(x) \triangleq \mathrm{P}(X = x|A) = \frac{\mathrm{P}(\{X = x\} \cap A)}{\mathrm{P}(A)}$$

- The event $\mathrm{P}(\{X = x\} \cap A)$ is disjoint for distinct x, therefore

$$\mathrm{P}(A) = \sum_{x} \mathrm{P}(\{X = x\} \cap A)$$

Example 2.12

- Let X be the roll of a fair six-sided die and let A be the event that the roll is an even number. Find $p_{X|A}(x)$.

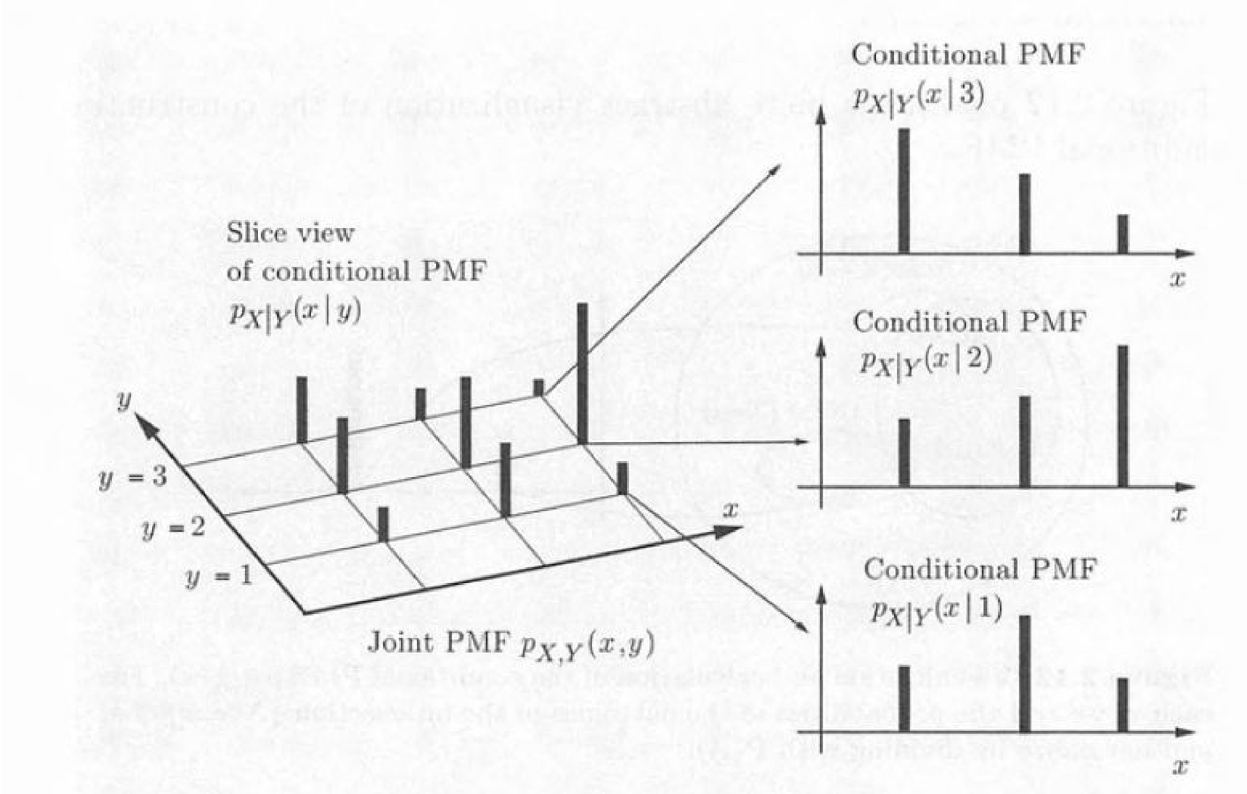
Conditioning an RV on Another

- Let X and Y be two random variables associated with the same experiment. The conditional PMF $p_{X|Y}(x|y)$ is defined by

$$p_{X|Y}(x|y) \triangleq \mathrm{P}(X = x|Y = y) = \frac{\mathrm{P}(X = x, Y = y)}{\mathrm{P}(Y = y)}$$

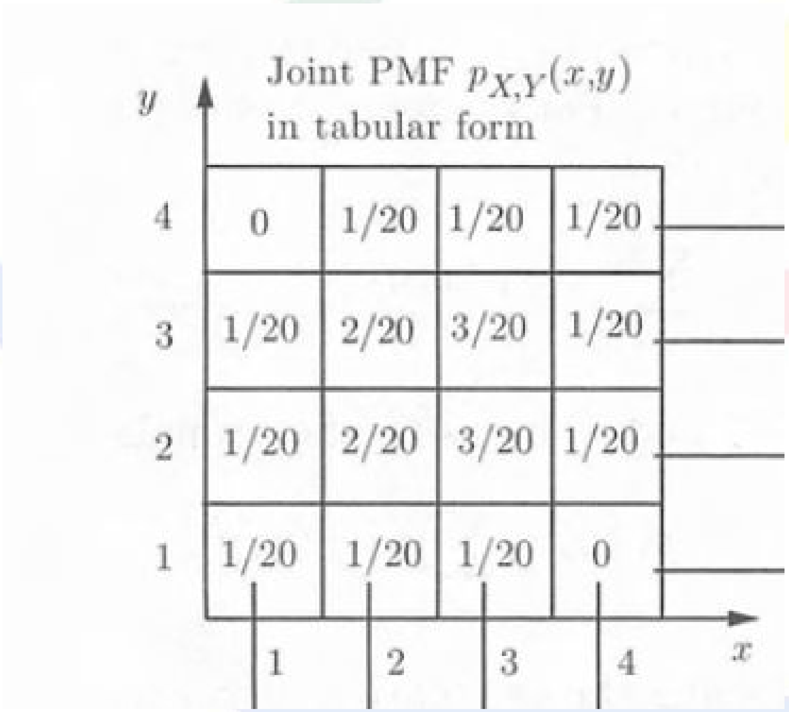
$$= \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

Visualization



Example 2.9 revisited

The following figure show the joint PMF of two RVs X and Y. Find $p_{X|Y}(x|y)$ and $p_{Y|X}(y|x)$



Joint PMF $p_{X,Y}(x,y)$ in tabular form

y \ x	1	2	3	4
4	0	1/20	1/20	1/20
3	1/20	2/20	3/20	1/20
2	1/20	2/20	3/20	1/20
1	1/20	1/20	1/20	0

Conditional Expectation

- The conditional expectation of X given an event A with $P(A) > 0$, is defined by

$$E[X|A] = \sum_x x p_{X|A}(x|A)$$

- For function $g(X)$,

$$E[g(X)|A] = \sum_x g(x) p_{X|A}(x|A)$$

- The conditional expectation of X given a value y of Y is defined by

$$E[X|Y = y] = \sum_x x p_{X|Y}(x|Y = y)$$

Conditional Expectation

- If $A_1, \cdots A_n$ form a partition of the sample space with $P(A_i) > 0$ for all i, then

$$E[X] = \sum_i E[X|A_i]\mathrm{P}(A_i)$$

and

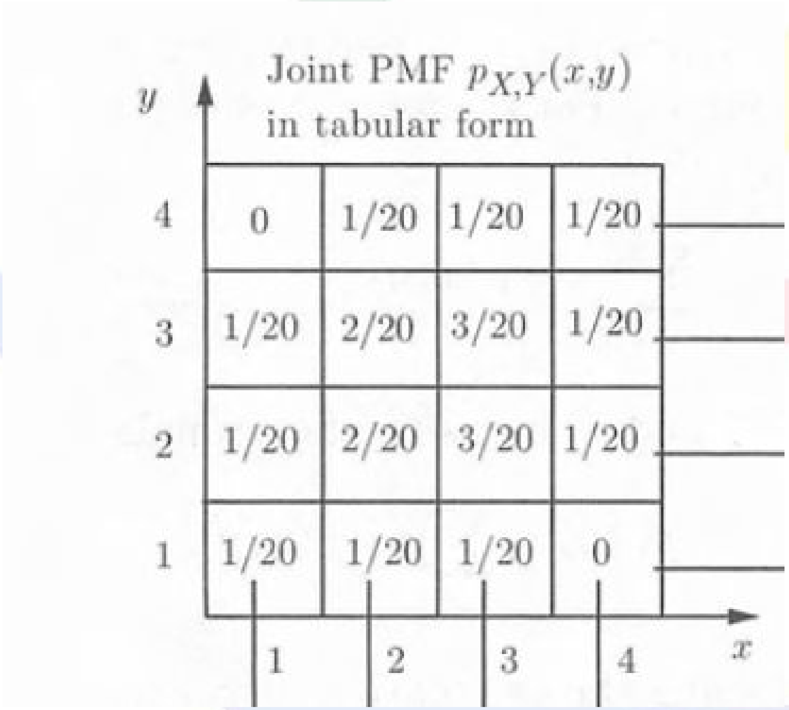
$$E[X|B] = \sum_i E[X|A_i \cap B]\mathrm{P}(A_i|B)$$

- "Total expectation theorem"

$$E[X] = \sum_y E[X|Y = y]\mathrm{P}_Y(y)$$

Example 2.9 revisited

The following figure show the joint PMF of two RVs X and Y. Find $E[X|X < Y]$, $E[X^2|X < Y]$, $\mathrm{Var}(X|X < Y)$, $E[X|Y]$


Joint PMF $p_{X,Y}(x,y)$ in tabular form

$y \backslash x$	1	2	3	4
4	0	1/20	1/20	1/20
3	1/20	2/20	3/20	1/20
2	1/20	2/20	3/20	1/20
1	1/20	1/20	1/20	0

2.7 Independence

Independence of a Random Variable from an I

- The random variable X is independent of the event A if

$$\mathrm{P}(X = x, A) = \mathrm{P}(X = x)\mathrm{P}(A) = p_X(x)\mathrm{P}(A), \quad \text{for all } x$$

- Since $\mathrm{P}(X = x, A) = p_{X|A}(x)p(A)$, the definition of independence is equivalent to

$$p_{X|A}(x) = p_X(x), \quad \text{for all } x.$$

Example 2.19

Consider two independent tosses of a fair coin. Let X be the number of heads and let A be the event that the number of heads is even. Are X and A independent?

Independence of Random Variables

- Two random variables X and Y are independent if

$$p_{X,Y}(x,y) = p_X(x)p_Y(y), \quad \text{for all } x, y.$$

- Since $p_{X,Y}(x,y) = p_{X|Y}(x|y)p_Y(y)$, the definition of independence is equivalent to

$$p_{X|Y}(x|y) = p_X(x), \quad \text{for all } x, y.$$

- If X and Y are independent, then
  - $E[XY] = E[X]E[Y]$
  - $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$
  - $\text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$

Independence of Several Random Variables

- Three random variables X, Y and Z are independent if

$$p_{X,Y,Z}(x,y,z) = p_X(x)p_Y(y)p_Z(z), \quad \text{for all } x,y,z.$$

  - $f(X)$, $g(Y)$, and $h(Z)$, are independent.
  - $g(X,Y)$ and $h(Z)$ are independent.
  - In general, $g(X,Y)$ and $h(Y,Z)$ are NOT independent.
- If $X_1, X_2, \ldots, X_n$ are independent random variables, then

$$\text{var}(X_1 + X_2 + \cdots + X_n) = \text{var}(X_1) + \text{var}(X_2) + \cdots \text{var}(X_n)$$

Example 2.20. Variance of Binomial RV

- If Y is a binomial rv with parameters (n, p), then

$$Y = X_1 + X_2 + \cdots + X_n$$

  where $X_1, X_2, \ldots, X_n$ are independent Bernoulli RVs with parameters p
- $E[X_k] = p, E[X_k^2] = p, \mathrm{Var}(X_k) = p - p^2 = p(1-p)$
- $\mathrm{Var}(Y) = np(1-p)$

Variance of Poisson RV

- If Y is a binomial rv with parameters (n, p), then

$$E[Y] = np, \quad \text{Var}(Y) = np(1-p) = np - np^2$$

- Poisson RV is the limiting case of $n \to \infty, p \to 0, np = \lambda$
- For this limiting case, we have

$$E[Y] = \lambda, \quad \text{Var}(Y) = np(1-p) = \lambda - p\lambda = \lambda$$