2. Discrete Random Variables

Han-You Jeong

School of Electrical and Computer Engineering

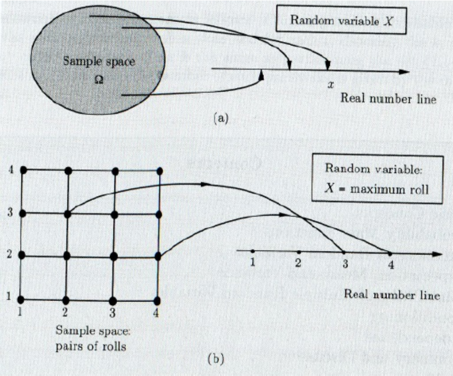
Pusan National University

Probability and Statistics

Outline

What is a Random Variable?

- In many probabilistic models, the outcomes are numerical, e.g., instrument reading or stock prices.
- A random variable is used to associate a particular number with each outcome.
  - We refer to this number as the value of the random variable.



Main Concepts of Random Variables

- A random variable is a real-valued function of the experimental outcome.
- A function of a random variable defines another random variable.
- We can associate with each random variable certain averages of interest, such as the mean and the variance.
- A random variable can be conditioned on an event or on another random variable.
- There is a notion of independence of a random variable from an event or from another random variable.

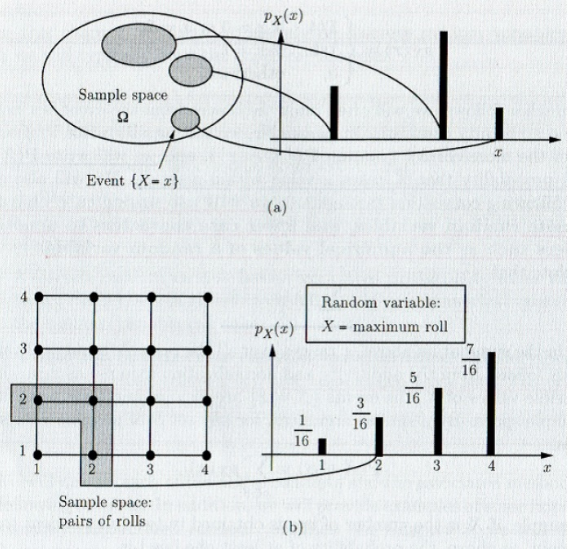
Concepts Related to Discrete Random Variables

- A discrete random variable is a real-valued function of the outcome of the experiment that can take a finite or countably infinite number of values.
- A discrete random variable has an associated probability mass function (PMF), which gives the probability of each numerical value that the random variable can take.
- A function of discrete random variable defines another discrete random variable, whose PMF can be obtained from the PMF of the original random variable.

Probability Mass Functions

- We can characterize a random variable with the probabilities of the values that it can take.
  - For a discrete random variable X, if x is any value of X, the probability mass of x, denoted by $p_X(x)$, is the probability of the event $\{X = x\}$ consisting of all outcomes that give rise to a value of X equal to x, i.e., $p_X(x) = P(\{X = x\})$.
  - We generally use upper case characters to denote random variables, and lower case characters to denote real numbers such as the numerical values of a random variable.
  - $$\sum_x p_X(x) = 1, \text{ and } P(X \in S) = \sum_{X \in S} p_X(x).$$
- Calculation of the PMF of a Random Variable X:
  1. Collect all the possible outcomes that give rise to the event $\{X = x\}$.
  2. Add their probabilities to obtain $p_X(x)$.

Examples of PMF Calculations



The Bernoulli Random Variable

- Consider the toss of a coin, which comes up a head with probability p, and a tail with probability $1 - p$.
- The Bernoulli random variable takes the two values 1 and 0, depending on whether the outcome is a head or a tail:

$$X = \begin{cases} 1, & \text{if a head,} \\ 0, & \text{if a tail.} \end{cases} \quad \text{with its PMF, } p_X(k) = \begin{cases} p, & \text{if } k = 1, \\ 1 - p, & \text{if } k = 0. \end{cases}$$

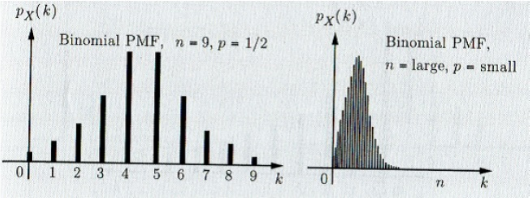
- The Bernoulli random variable is used to model generic probabilistic situations with just two outcomes, e.g.,
  - The state of a telephone at a given time that can be either free or busy.
  - A person who can be either healthy or sick with a certain disease.
  - The preference of a person who can be either for or against a certain political candidate.

The Binomial Random Variable

- A coin is tossed n times, each of which comes up a head with probability p, and a tail with probability $1-p$.
  - Let X be the number of heads in the n-toss sequence.
- We refer to X as a binomial random variable with parameters n and p, whose PMF is given by

$$p_X(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \cdots, n.$$

- From the normalization property, we have $\displaystyle\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1.$



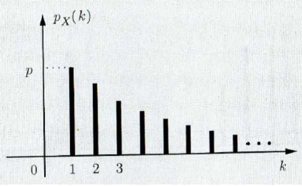
The Geometric Random Variable

- We repeatedly and independently toss a coin with probability of a head equal to p.
- The geometric random variable is the number X of tosses needed for a head to come up for the first time, whose PMF is given by

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

- From the normalization property, we have

$$\sum_{k=1}^{\infty} p_X(k) = p\sum_{k=0}^{\infty}(1-p)^k = p \cdot \frac{1}{1-(1-p)} = 1.$$



The Poisson Random Variable

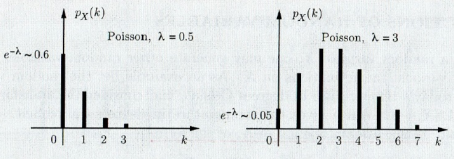
- The Poisson random variable has a PMF given by

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \cdots .$$

- From the normalization property, we have

$$\sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots\right) = e^{-\lambda}e^{\lambda} = 1.$$

- The Poisson random variable can be seen as a binomial random variable with very small p and very large n, such that $p \cdot n = \lambda$.



Functions of Random Variables

- Given a random variable X, consider the transformation $Y = g(X)$.
  - If $Y = g(X) = aX + b$, Y is a linear function of X.
  - We may also consider nonlinear functions, e.g. $Y = \log X$.
- Every outcome in the sample space defines a numerical value x for X and hence also the numerical value $y = g(x)$ for Y.
  - If X is discrete with PMF $p_X(x)$, then Y is also discrete with its PMF $p_Y(y)$, where

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

Ex 2.1 The PMF of a random variable X is given as follows:

$$p_X(x) = \begin{cases} 1/9, & \text{if } x \text{ is an integer in the range } [-4, 4], \\ 0, & \textit{otherwise}; \end{cases}$$

Calculate the PMF p_Y, 1) if $Y = |X|$; and 2) if $Z = X^2$.

Expectation

- The PMF of a random variable X provides us with the probabilities of all the possible values of X.
  - Can we determine a single representative number of X?
- The expectation of X, denoted by $E[X]$, is a weighted (in proportion to probabilities) average of the possible values of X, with PMF p_X, where

$$E[X] = \sum_{\forall x} x p_X(x).$$

  - The expectation can be seen as the center of gravity of the PMF.

Ex 2.2 Consider two independent coin tosses, each with a $3/4$ probability of a head, and let X be the number of heads obtained. Calculate $E[X]$.

Moments, Variance, and the Standard Deviation

- We define the nth moment as $E[X^n]$, the expected value of the random variable X^n.
  - The 1st moment of X is just the mean $E[X]$.
- The variance of random variable X, denoted by $var(X)$, is defined as the expected value of the random variable $(X - E[X])^2$, i.e.,

$$var(X) = E[(X - E[X])^2].$$

  - The variance provides a measure of dispersion of X around its mean.
- The standard deviation of X, denoted by σ_X is defined as the square root of the variance, i.e.,

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

Ex 2.3 In Ex 2.1, calculate the variance of a random variable Z.

Expected Value Rule

- Let X be a random variable with PMF p_X, and let $g(X)$ be a function of X. Then, the expected value of the random variable $g(X)$ is given by

$$E[g(X)] = \sum_{\forall x} g(x) p_X(x).$$

  - Prove the above statement by using a substitution $Y = g(X)$.
  - The variance of a random variable X is given by

$$var(X) = E[(X - E[X])^2] = \sum_{\forall x} (x - E[X])^2 p_X(x).$$

Ex 2.3 (continued) In Ex 2.1, calculate the variance of a random variable X.

Properties of Mean and Variance

- Let X be a random variable and let $Y = aX + b$, where a and b are given scalars. Then,

$$E[Y] = aE[X] + b, \quad var(Y) = a^2 var(X).$$

Ex 2.4 If the weather is good (which happens with probability 0.6), Alice walks the 2 miles to class at a speed of $V = 5$ mph, and otherwise rides her motorcycle at a speed of $V = 30$ mph. What is the mean of the time T to get to class?

Ex 2.5 Calculate the mean and the variance of the Bernoulli random variable with probability p.

Ex 2.6 Calculate the mean and the variance of the discrete uniform random variable in interval $[a, b]$.

Ex 2.7 Calculate the mean and the variance of the Poisson random variable with rate λ.

Decision Making using Expected Values

- Expected values often provides a convenient vehicle for optimizing the choice between several candidate decisions that result in random rewards.

  Ex 2.8 Consider a quiz game where a person is given two questions and must decide which on to answer first. Question 1 (2) will be answered correctly with probability 0.8 (0.5), and the person will then receive as prize \$ 100 (\$ 200). If the first question attempted is answered incorrectly, the quiz terminates; otherwise, the person is allowed to attempt the second question. Which question should be answered first to maximize the expected value of the total prize money received?

Joint PMFs and Marginal PMFs

- Probabilistic models often involve several random variables, e.g., workloads of routers in a network and test results of a drug in a medical diagnosis.
  - All of these random variables are associated with the same experiment, sample space, and probability law, and their values may relate in interesting ways.
- The joint PMF of random variables X and Y, denoted by $p_{X,Y}$, captures the probabilities of the values that X and Y can take, i.e.,

$$p_{X,Y}(x,y) = P(X = x, Y = y), \text{ and } P((X,Y) \in A) = \sum_{\forall (x,y) \in A} p_{X,Y}(x,y).$$

- The marginal PMFs, denoted by p_X and p_Y, are the PMFs of a single random variable X and Y, i.e.,

$$p_X(x) = \sum_{\forall y} p_{X,Y}(x,y), \text{ and } p_Y(y) = \sum_{\forall x} p_{X,Y}(x,y).$$

Functions of Multiple Random Variables

- A function $Z = g(X, Y)$ of random variables X and Y defines another random variable whose PMF is obtained from the joint PMF $p_{X,Y}$

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

- The expected value rule naturally extends to the form

$$E[g(X, Y)] = \sum_{\forall x} \sum_{\forall y} g(x, y) p_{X,Y}(x, y).$$

For example $g(X, Y) = aX + bY + c$, we have

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

More than Two Random Variables

- The joint PMF of three random variables X, Y, and Z is defined in analogy with the above as

$$p_{X,Y,Z}(x,y,z) = P(X = x, Y = y, Z = z),$$

for all possible triplets of numerical values (x, y, z).

- The marginal PMFs are analogously obtained by equations such as

$$p_{X,Y}(x,y) = \sum_{\forall z} p_{X,Y,Z}(x,y,z), \text{ and } p_X(x) = \sum_{\forall y}\sum_{\forall z} p_{X,Y,Z}(x,y,z).$$

- The expected value rule for functions is given by

$$E[g(X,Y,Z)] = \sum_{\forall x}\sum_{\forall y}\sum_{\forall z} g(x,y,z)p_{X,Y,Z}(x,y,z).$$

Ex 2.11 Suppose that n people throw their hats in a box and then each picks one hat at random. What is the expected value of X, the number of people that get back their own hat?

Conditioning a Random Variable on an Event

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

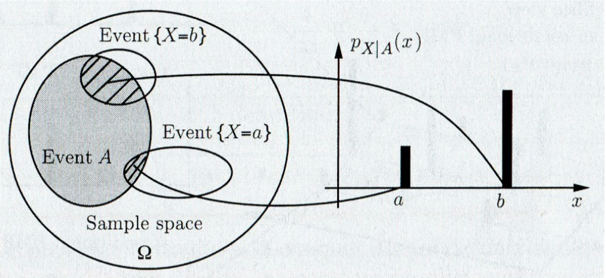
$$p_{X|A}(x) = P(X = x|A) = \frac{P(\{X = x\} \cap A)}{P(A)}.$$

- Note that the events $\{X = x\} \cap A$ are disjoint for different values of x, their union is A, and therefore,

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

- Combining the above two formulas, we have

$$\sum_{\forall x} p_{X|A}(x) = 1.$$



Ex 2.13 A student will take a certain test repeatedly, up to a maximum of n times, each time with a probability p of passing, independent of the number of previous attempts. What is the PMF of the number of attempts, given that the student passes the test?

Conditioning one Random Variable on Another

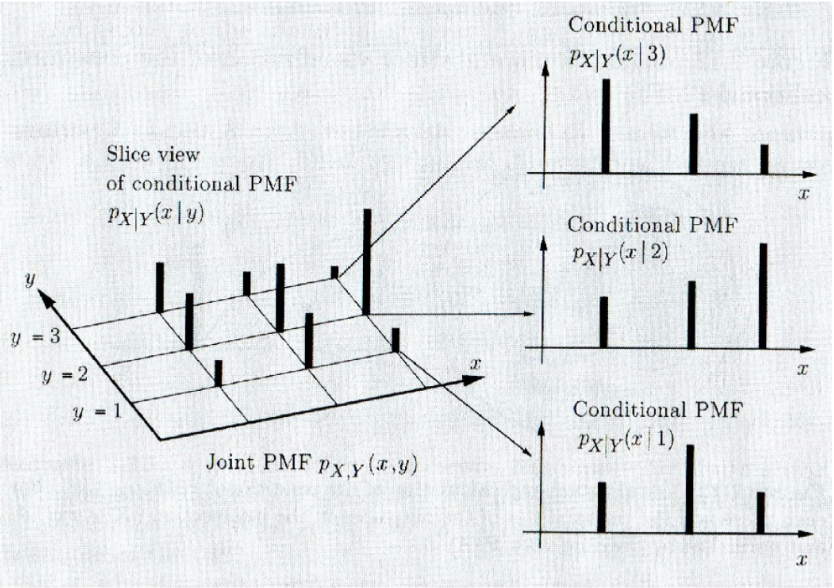
- Consider two random variables X and Y associated with the same experiment.
- If we know that the value of Y is some particular y [with $p_Y(y) > 0$], this provides partial knowledge about the value of X.
- The conditional PMF $p_{X|Y}$ of X given Y, which is specializing the definition of $p_{X|A}$ to events A of the form $\{Y = y\}$:

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

- The conditional PMF is convenient for the calculation of the joint PMF and the marginal PMF, i.e.,

$$p_{X,Y}(x,y) = p_Y(y)p_{X|Y}(x|y) = p_X(x)p_{Y|X}(y|x),$$

$$p_X(x) = \sum_{\forall y} p_{X,Y}(x,y) = \sum_{\forall y} p_Y(y)p_{X|Y}(x|y).$$


Slice view
of conditional PMF
$p_{X|Y}(x \mid y)$
y
$y = 3$
$y = 2$
$y = 1$
x
Joint PMF $p_{X,Y}(x,y)$
Conditional PMF
$p_{X|Y}(x \mid 3)$
x
Conditional PMF
$p_{X|Y}(x \mid 2)$
x
Conditional PMF
$p_{X|Y}(x \mid 1)$
x

An Example

Ex 2.14 Prof. May often has her facts wrong, and answers each of her students' questions incorrectly with probability $1/4$, independent of other questions. In each lecture, May is asked 0, 1, or 2 questions with equal probability $1/3$. Let X and Y be the number of questions May is asked and the number of questions she answers wrong in a given lecture, respectively. What is the probability of at least one wrong answer?

Conditional Expectation

- A conditional expectation is the same as an ordinary expectation, except that it refers to the new universe, and all probabilities and PMFs are replaced by their conditional counterparts.
  - Let X and Y be random variables associated with the same experiment.
  - The conditional expectation of X given an event A with $P(A) > 0$ is

$$E[X|A] = \sum_{\forall x} x p_{X|A}(x), \text{ and } E[g(X)|A] = \sum_{\forall x} g(x) p_{X|A}(x).$$

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

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

Ex 2.16 Messages transmitted by a computer in Boston through a data network are destined for New York with probability 0.5, for Chicago with probability 0.3, and for San Francisco with probability 0.2. The transit time X of a message is random whose means are 0.05, 0.1, and 0.3 seconds for NY, CH, and SF, respectively. Calculate $E[X]$.

Total Expectation Theorem

- The total expectation theorem means that the unconditional average can be obtained by averaging the conditional averages.
  - Let $A_1, \cdots, A_n$ be disjoint events that form a partition of the sample space with $P(A_i) > 0$ for all i. For all events B with $P(A_i \cap B) > 0$ for all i, we have
$$E[X] = \sum_{i=1}^{n} P(A_i)E[X|A_i], \text{ and } E[X|B] = \sum_{i=1}^{n} P(A_i|B)E[X|A_i \cap B].$$
  - We also have
$$E[X] = \sum_{\forall y} p_Y(y)E[X|Y = y].$$

Ex 2.18 You are handed two envelopes, and you are told that one of them contains m times as much money as the other. ($m \in Z$ and $m > 1$) You open one of the envelopes and look at the amount inside. You may now keep this amount, or you may switch envelopes and keep the amount in the other envelope. What is the best strategy?

Independence of a Random Variable from an Event

- The idea of independence is that knowing the occurrence of the conditioning event provides no new information on the value of the random variable.
- Random variable X is independent of the event A if

$$P(X=x, A) = P(A)P(X=x|A) = P(X=x)P(A) = p_X(x)P(A), \ \forall x.$$

In other words, $p_{X|A}(x) = p_X(x)$.

Ex 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. Is X independent of A?

Independence of Random Variables

- Two random variables X and Y are independent if

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

In other words,

$$p_{X|Y}(x|y) = p_X(x), \text{ for all } y \text{ with } p_Y(y) > 0 \text{ and all } x.$$

- X and Y are conditionally independent, given a positive probability event A, if

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

Equivalently, we have

$$p_{X|Y,A}(x|y) = p_{X|A}(x), \text{ for all } x \text{ and } y \text{ such that } p_{Y|A}(y) > 0.$$

- If X and Y are independent random variables, then

$$E[XY] = E[X]E[Y], \text{ and } E[g(X)h(Y)] = E[g(X)]E[h(Y)].$$

$$var(X+Y) = var(X) + var(Y).$$

Independence of Several Random Variables

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

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

  - Any three random variables of the form $f(X)$, $g(Y)$, and $h(Z)$ are independent.
  - Any two random variables of the form $g(X,Y)$ and $h(Z)$ are independent.
  - Two random variables of the form $g(X,Y)$ and $h(Y,Z)$ are usually not independent, because they are both affected by Y.
- If $X_1, X_2, \cdots, X_n$ are independent random variables, then

$$var(X_1 + X_2 + \cdots + X_n) = var(X_1) + var(X_2) + \cdots + var(X_n).$$

Ex 2.21 We wish to estimate the approval rating of a president by asking n persons drawn at random from the voters. Let $X_i = 1$, if the ith person approves the president's performance; otherwise, $X_i = 0$. Calculate the mean and the variance of the sample mean.

Homework # 2

- Section 2.2 Probabilistic Mass Functions: Problem 4
- Section 2.3 Functions of Random Variables: Problem 14
- Section 2.4 Expectation, Mean, and Variance: Examples 2.4 and 2.8, Problems 21 and 23
- Section 2.5 Joint PMFs of Multiple Random Variables: Problem 26
- Section 2.6 Conditioning: Examples 2.14 and 2.18, Problem 32
- Section 2.7 Independence: Example 2.21, Problem 40

Thank You