Lecture notes: Discrete Random Variables

Jie Fu

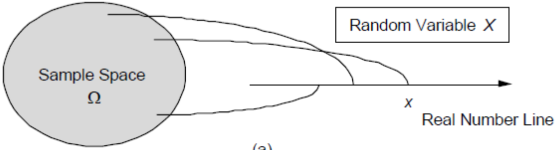
Department of Electrical and Computer Engineering
University of Florida

Outline

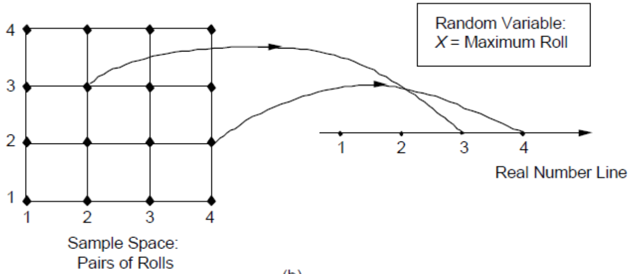
- Random variables.
- Discrete Random variables.
- Probability mass function.
- Cumulative distribution function.
- Three important discrete RVs.

Random variable

A random variable (r.v.) associates a value (a number) to **every possible outcome.**
e.g. In a dice roll, the outcome $\{1, 2, 3, 4, 5, 6\}$ can be directly assigned to the random variable X , where $X = 1, 2, \ldots, 6$.



(a)



(b)

Discrete Random Variables

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.

$$p_X : \mathcal{X} \to [0, 1]$$

where $\mathcal{X}$ is all possible values X can take.

Notation:

$$p_X(x) = P(X = x)$$

Properties: $p_X(x) \geq 0$, $\sum_x P_X(x) = 1$.

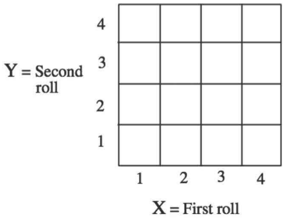
Example

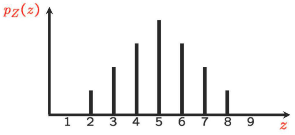
Two independent tosses of a fair coin:
Define discrete RV X: The total number of heads.
The PMF of X is

A function of one or several random variables' is also a random variable.

Two rolls of a 4-sided die:

$$Z = X + Y$$





Calculation of the PMF of a Random Variable For each possible value z of Z:

1. Collect all the possible outcomes that give rise to the event $\{Z = z\}$.
2. Add their probabilities to obtain $p_Z(z)$.

Cumulative Distribution Function (CDF)

If $(\Omega, \mathcal{F}, P)$ is a probability space with X a real discrete RV on Ω, the **Cumulative Distribution Function (CDF)** is denoted as $F_X(x)$ and provides the probability $P(X \leq x)$. In particular, for every x we have

$$F_X(x) = P(X \leq x) = \sum_{k \leq x} p_X(k)$$

Loosely speaking, the CDF $F_X(x)$ "accumulates" probability "up to" the value x.

Example

I toss a coin twice. Let X be the number of observed heads. Find the CDF of X.
Recall the PMF of X:

$$p_X(x) = \begin{cases} \frac{1}{4} & \text{if } x = 0, 2 \\ \frac{1}{2} & \text{if } x = 1 \\ 0 & \text{otherwise.} \end{cases}$$

To find $F_X(x)$

- $x < 0$:
- $x = 0$:
- $x = 1$:
- $x = 2$:
- $x > 2$:

Next, we will introduce some important discrete RV.

- Bernoulli Random Variable
- Binomial Random Variable.
- Geometric Random Variable.

Bernoulli Random Variable

e.g. Medical treatment: Two outcomes in the sample space: $x = 1$ for "success" and $x = 0$ for "failure".

Bernoulli RV: Models a trial that results in success/failure, Heads/Tails, etc.

- A **Bernoulli RV** X takes two values 0 and 1.
- The PMF for a Bernoulli RV X is defined by

$$p_X(x) = P(X = x) = \begin{cases} p, & x = 1 \\ 1 - p, & x = 0 \\ 0, & \text{o.w.} \end{cases}$$

$$X \sim \text{Bernoulli}(p)$$

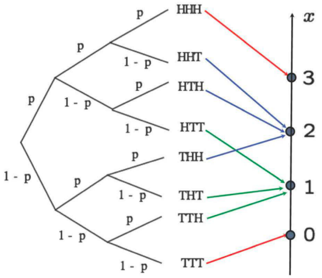
Examples/applications:

- The coin flip is either heads (1) or tail (0).
- whether a bit is 0 or 1, whether a bit is in error, whether a component has failed, whether something has been detected.
- The state of a telephone at a given time that can be either free (0) or busy (1).
- A person who can be either healthy or sick with a certain disease.

The Binomial Random Variable

consider the following experiment:

- A biased coin is tossed **n** times.
- Each toss is independently of prior tosses: Head with probability p; Tail with probability $1 - p$.
- The number X of heads up is a binomial random variable, refer to as the **Binomial** RV **with parameters** n **and** p.



$$X \sim \text{Binomial}(n, p)$$

$$P(X = x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, \ldots, n \\ 0, & \text{o.w.} \end{cases}$$

where $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ is the different combinations of x objects chosen from n objects.

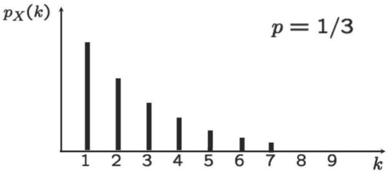
Example: Binomial

A pharmaceutical company is testing a new medication to see if it is effective in treating a certain disease. The company wants to know the probability that the medication will work for at least 80% of patients. To test this, they give the medication to 20 patients and observe whether it works or not.

- Based on the observation that the medicine worked for 5 out of 20 patients, do you think you believe the company's claim that the effective rate is 80%?

Geometric Random Variable

- Experiment: infinitely many independent tosses of a coin $P(Heads) = p$.
- Sample space: Set of infinitte sequences of H and T.
- Random variable X: number of tosses until the first Heads.
- Models of: waiting times; number of trials until a success.



decreases as a **geometric progression** with parameter $1 - p$.

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

what is the probability of no heads ever?

Example

A customer calls a tech support line, and each agent has a 20% chance of successfully resolving the issue. The customer may need to speak to multiple agents before the problem is fixed. Assuming all agents are not sharing the knowledge/experience learned from talking with the customer, how many agents the customer need to talk to before his issue is solved with probability $> 95\%$?

Poisson random variable

Consider the random variable X: The number of typos in a book of n words: - each word can be misspelled with a probability p.

Poisson random variable

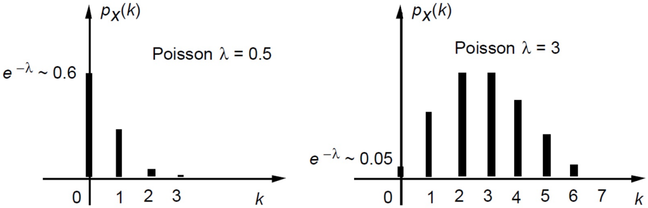
Consider the random variable X: The number of typos in a book of n words: - each word can be misspelled with a probability p.

$$X \sim \text{Poisson}(\lambda)$$

$$P_X(x) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda}, & x = 0, 1, \dots \\ 0, & \text{o.w.} \end{cases}$$

where $\lambda = n \cdot p$, the expected number of occurrences in the interval.
— A Poisson RV with parameter λ is an approximation of binomial RV with parameter $n \gg 0$ and $p \ll 1$.

Poisson random variable



Figure

- $\lambda \leq 1$, monotonically decreasing.
- $\lambda > 1$, first increase and then decrease.

Example

suppose the book has 50,000 words, and each word can be mistyped with a probability $p = 0.2\%$. What is the probability that the book has five typos?

Summary

- Bernoulli: A single trial with two possible outcomes: success (1) with probability p and failure (0) with probability $1 - p$.
- Binomial: The number of successes in n independent Bernoulli trials, each with success probability p.
- Geometric: The number of trials until the first success in repeated independent Bernoulli trials with success probability p.
- Poisson: Counting rare events (e.g., defective items in manufacturing, earthquakes per year). Approximate Binomial with $n \gg 1$ and $p \ll 1$.