Introduction to Probability and Statistics

Prof. Mohammed-Slim Alouini & Prof. Hesham ElSawy
slim.alouini@kaust.edu.sa & hesham.elsawy@kfupm.edu.sa

Eng. Chaouki Bem Issaid & Eng. Lama Niyazi
chaouki.benissaid@kaust.edu.sa & lama.niyazi@kaust.edu.sa

STC Academy, Riyadh, KSA

23 June to 4 July 2019

Lecture Objectives

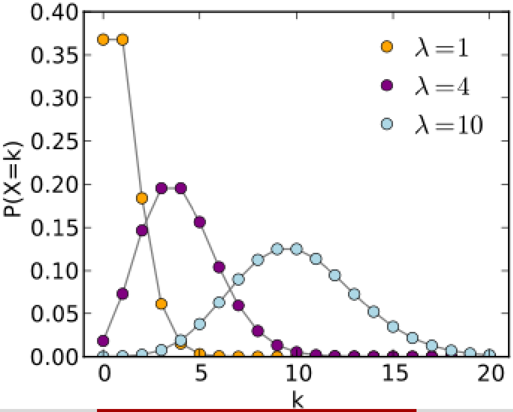
- Yesterday
  - Introduction to probability
  - Random experiments
  - Outcome spaces, events, and probability laws
  - Conditional probability and Baye's theorem
  - Independence and combined experimenters
  - Counting and combinatorics.
- Today
  - Discrete random variables

Outcome of Random Experiment

- Events of random experiments can be
  - descriptive
  - numeric
- Examples
  - Tossing a coin twice $\{HH, HT, TH, TT\}$ descriptive
  - Selecting cards from a 52-deck of cards descriptive
  - Drawing balls from boxes descriptive
  - Rolling dice numeric
  - Rotating pointers numeric

Outcome of Random Experiment

- It is favorable to map descriptive events into numerical values
  - Calculate expectations
  - Define distributions
  - Make transformations

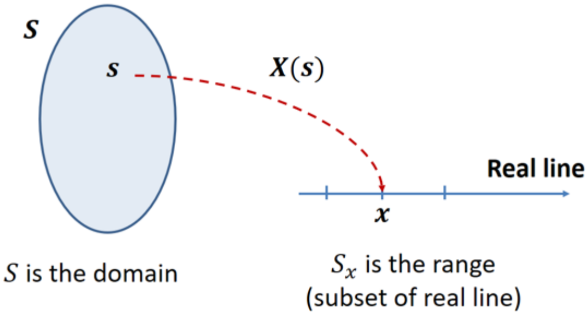


Random Variable

Random Variable

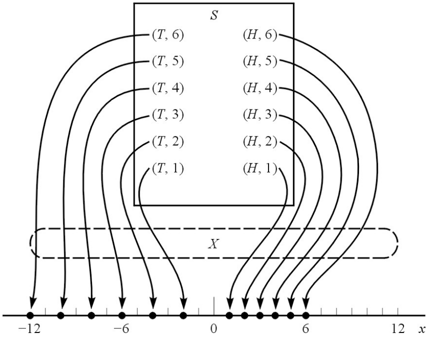
A random variable X is a function that assigns a real number $x \in \mathbb{R}$, such that $X(s) = x$, to each outcome $s \in \mathcal{S}$.

$$X : \mathcal{S} \to \mathbb{R}$$



Example

- Consider combined experiment of tossing a coin and rolling a die
- The outcome space is $\mathcal{S} = \{H, T\} \times \{1, 2, 3, 4, 5, 6\}$
- A random variable can be defined via the following mapping
- $X : \{H, T\} \times \{1, 2, 3, 4, 5, 6\} \to \{-12, -10, -8, -6, -4, -2, 1, 2, 3, 4, 5, 6\}$



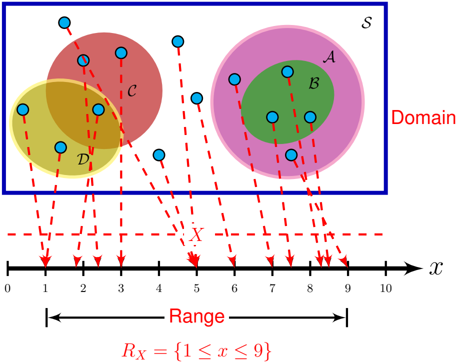
Events

Events

- The random variable is denoted as X and its numeric values (realizations or instantiations) are denoted as x
- Events are specified in the form $\{X \leq x\}$
- Probabilities are assigned to events according to the probability law
  - $\mathbb{P}\{\text{event } \mathcal{A} \text{ occurs}\} \Longrightarrow \mathbb{P}\{X = x\}$
  - $\mathbb{P}\{\text{event } \mathcal{B} \text{ occurs}\} \Longrightarrow \mathbb{P}\{X > x\}$
  - $\mathbb{P}\{\text{event } \mathcal{C} \text{ occurs}\} \Longrightarrow \mathbb{P}\{X < x\}$
  - $\mathbb{P}\{\text{event } \mathcal{D} \text{ occurs}\} \Longrightarrow \mathbb{P}\{x_1 \leq X \leq x_2\}$

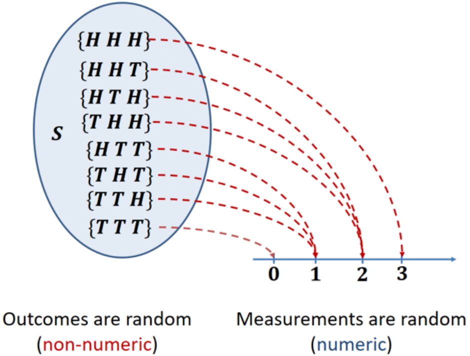
Example

- Consider the following outcome space
- Define a mapping function $X : \mathcal{S} \rightarrow \{1 \leq x \leq 9\}$
- $\mathbb{P}(\mathcal{B}) = \mathbb{P}(7 < X < 9)$
- $\mathbb{P}(\mathcal{A}) = \mathbb{P}(7 \leq X)$
- $\mathbb{P}(\mathcal{C}) = \mathbb{P}(1 < X \leq 3)$
- $\mathbb{P}(\mathcal{D}) = \mathbb{P}(X \leq 2)$
- $\mathbb{P}(\overline{(\mathcal{A} \cup \mathcal{C} \cup \mathcal{D})}) = \mathbb{P}(5 \leq X \leq 6)$
- How about the following events
    - $\mathbb{P}(X < 1)$=$\mathbb{P}(\phi) = 0$
    - $\mathbb{P}(X < 10)$ =$\mathbb{P}(\mathcal{S}) = 1$
    - Both are legitimate events



Example

- Consider a random experiment of tossing a fair coin 3 times
- Define a random variable that counts the number of heads
- $X : \mathcal{S} \to \{0, 1, 2, 3\}$
- $\mathbb{P}(X = 0) = \mathbb{P}(X = 3) = \frac{1}{8}$
- $\mathbb{P}(X = 1) = \mathbb{P}(X = 2) = \frac{3}{8}$



Example

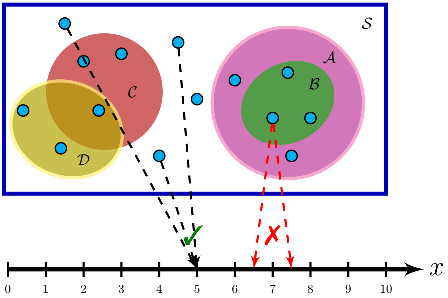
- Find the range of the following random variables
  - Toss a coin 100 times, Let X be the number of heads I observe
  - Toss a coin until the first head. Let Y be the number of coin tosses
  - The random variable T is defined as the time (in hours) from now until the net earthquake in a certain city

Conditions

Conditions for the function to be a random variable

For a function to be a random variable is should satisfy the following conditions

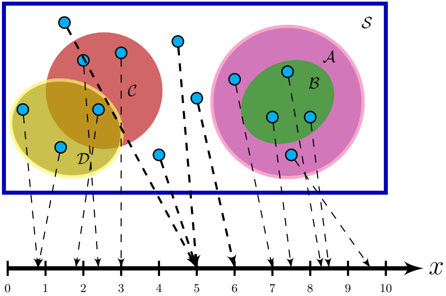
- Every point element $s \in \mathcal{S}$ should correspond to only one value $x \in \mathbb{R}$
- The set $\{X \leq x\}$ should be an event for every x



Conditions

Conditions for the function to be a random variable

- The set $\{X \leq x\}$ should be an event for every x
  - $\mathbb{P}(X < 0) = \mathbb{P}(\phi) = 0$
  - $\mathbb{P}(X < 4.5) = \mathbb{P}(\mathcal{C} \cup \mathcal{D})$
  - $\mathbb{P}(X < 7) = 1 - \mathbb{P}(\mathcal{A})$
  - $\mathbb{P}(X < 10) = 1$



Types of Random Variables

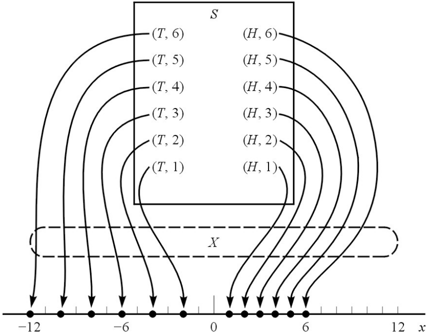
Types of Random Variables

There are three types of random variables

- Discrete
- Continuous
- Mixed

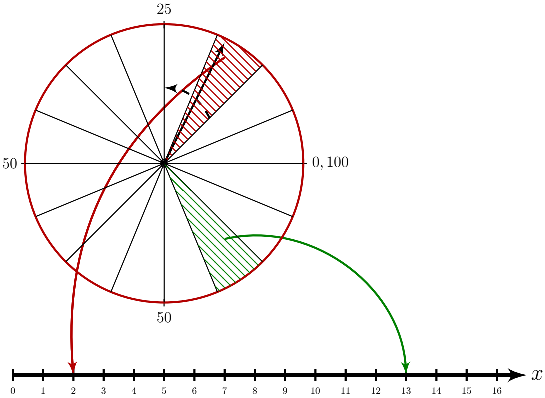
Discrete Random Variables

- A random variable is discrete if its range is countable
- Discrete random variables can be defined on a discrete or continuous outcome space
- Example
  - A discrete random variables defined on a discrete outcome space



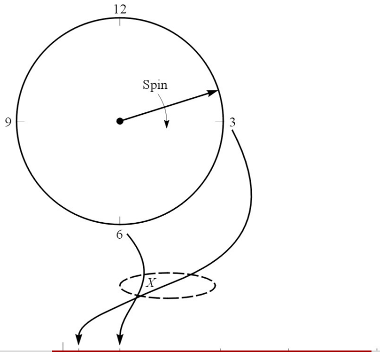
Discrete Random Variables

- A discrete random variable defined on a continuous outcome space
- Domain $\{0 \leq s \leq 100\}$ and Range $\{0, 1, 2, 3, \ldots, 15\}$
- Random variable $X : \{0 \leq s \leq 100\} \rightarrow \{0, 1, 2, 3, \ldots, 15\}$



Continuous Random Variables

- A random variable is continuous if its range is uncountable
- Continuous random variables can only be defined on a continuous outcome space
- Consider a spinning wheel with outcome space $\{S\} = \{0 \leq s \leq 12\}$
- Define $X = s^2$
- Random variable $X : \{0 \leq s \leq 12\} \rightarrow \{0 \leq x \leq 144\}$



Mixed Random Variables

- Mixed random variables have some discrete values and some continuous values
- Example
  - Consider an experiment of rolling a fair die with probability 0.5 or spinning a fair wheel, with a scale of $\{0 \le s \le 6\}$, with probability 0.5
  - The outcome space is $\mathcal{S} = \{0 \le s \le 6\}$
  - The probability law is as follows

$$\mathbb{P}(\{x\}) = \begin{cases} \frac{1}{12} & x = \{1,2,3,4,5,6\} \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \mathbb{P}(2.5 \le x \le 3.5) &= \mathbb{P}(\text{die})\mathbb{P}(x = 3|\text{die}) + \mathbb{P}(\text{wheel})\mathbb{P}(2.5 \le x \le 3.5|\text{wheel}) \\ &= \frac{1}{2} \times \frac{1}{6} + \frac{1}{2} \times \frac{3.5 - 2.5}{6} \\ &= \frac{1}{6} \end{aligned}$$

Progress

- Last section
  - Random variables
- Next section
  - Distribution function
  - Density function

Probability Mass Function

Probability mass function (PMF)

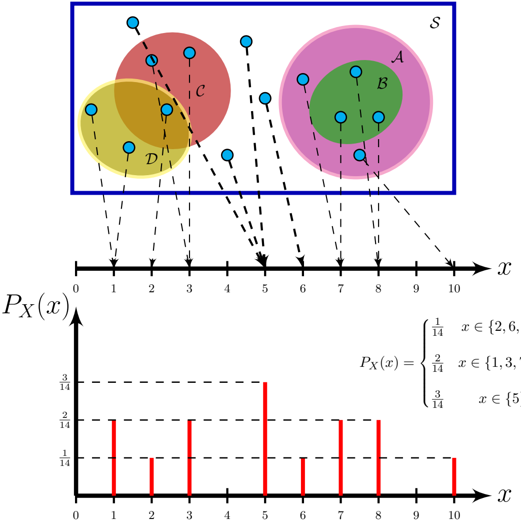
- Since the range of X is countable, it can be listed as

$$R_X = \{x_1, x_2, x_3, \cdots\}$$

- Let the event of $\mathcal{A}_k = \{s \in \mathcal{S} | X(s) = x_k\}$, then the probability of the random variable can be calculated
- The PMF is the probability $\mathbb{P}\{X(s) = x_k\}$ for $k = 1, 2, 3, \ldots$

$$P_X(x_k) = \mathbb{P}(X = x_k) = \mathbb{P}(\mathcal{A}_k)$$

Probability Mass Function



$$P_X(x) = \begin{cases} \frac{1}{14} & x \in \{2, 6, 10\} \\ \frac{2}{14} & x \in \{1, 3, 7, 8\} \\ \frac{3}{14} & x \in \{5\} \end{cases}$$

Probability Mass Function

Probability mass function (PMF)

The PMF satisfies all the probability axioms

- $0 \leq P_X(x) \leq 1$
- $\sum_{x \in R_x} P_X(x) = 1$
- For any event $\mathcal{A} \subset R_X$, the probability $\mathbb{P}(X \in \mathcal{A}) = \sum_{x \in \mathcal{A}} P_X(x)$

Extended notation

- The PMF can be extended on $\mathbb{R}$ follows

$$P_X(x) = \begin{cases} \mathbb{P}(X = x) & x \in R_X \\ 0 & \text{otherwise} \end{cases}$$

Example

- Consider tossing an unfair coin with $\mathbb{P}(H) = p$, where $0 \leq p \leq 1$. The coin is tossed until the first head appears. Let Y be the random variable of the number of tosses. Find the PMF of Y, validate your answer, and calculate $\mathbb{P}(2 \leq Y < 5)$

$$P_Y(1) = \mathbb{P}(Y = 1) = \mathbb{P}(H) = p$$

$$P_Y(2) = \mathbb{P}(Y = 2) = \mathbb{P}(TH) = (1-p)p$$

$$P_Y(3) = \mathbb{P}(Y = 3) = \mathbb{P}(TTH) = (1-p)^2 p$$

$$\vdots$$

$$P_Y(n) = \mathbb{P}(Y = n) = \mathbb{P}(TT \cdots TH) = (1-p)^{n-1} p$$

- Then we have

$$P_Y(y) = \begin{cases} p(1-p)^{y-1} & y \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

Example

- Consider tossing an unfair coin with $\mathbb{P}(H) = p$, where $0 \leq p \leq 1$. The coin is tossed until the first head appears. Let Y be the random variable of the number of tosses. Find the PMF of Y, validate your answer, and calculate $\mathbb{P}(2 \leq Y < 5)$ for $p = 0.5$

$$\sum_{y=1}^{\infty} p(1-p)^{y-1} = p\frac{1}{1-(1-p)} = 1$$

- The probability that $\mathbb{P}(2 \leq Y < 5)$ is given as

$$\begin{aligned}\mathbb{P}(2 \leq Y < 5) &= \sum_{y=2}^{4} p(1-p)^{y-1} \\ &= \frac{1}{2}(\frac{1}{2} + \frac{1}{4} + \frac{1}{8}) \\ &= \frac{7}{16}\end{aligned}$$

Independence

Statistical independence

- Recall that two events $\mathcal{A}$ and $\mathcal{B}$ are statistically independent if and only if (iff):
$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B})$$
- Similarly, two random variables X and Y are independent iff
$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$
- In general, if X and Y are independent, then for all sets $\mathcal{A}$ and $\mathcal{B}$
$$\mathbb{P}(X \in \mathcal{A}, Y \in \mathcal{B}) = \mathbb{P}(X \in \mathcal{A})\mathbb{P}(Y \in \mathcal{B})$$
- In terms of conditional probabilities
$$\mathbb{P}(X \in \mathcal{A} | Y \in \mathcal{B}) = \mathbb{P}(X \in \mathcal{A}) \quad \text{and} \quad \mathbb{P}(Y \in \mathcal{B} | X \in \mathcal{A}) = \mathbb{P}(Y \in \mathcal{B})$$

Example

- Toss a coin four times and let X be the number of heads in the first two tosses and Y be the number of heads in the last two tosses. Find the probability that $\mathbb{P}(X < 2, Y > 1)$

Progress...

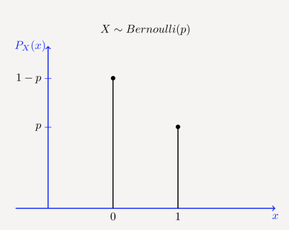
- Last section
  - PMF
  - Independence
- Next section
  - Popular random variables

Bernoulli Distribution

Bernoulli Distribution

- Defines an experiment with only two outcomes: *success*, denoted with 1, and *failure* denoted with 0
- A random variable X is said to be a Bernoulli random variable with parameter p, shown as $X \sim Bernoulli(p)$, if its PMF is given by

$$P_X(x) = \begin{cases} p & x = 1 \\ 1-p & x = 0 \\ 0 & \text{otherwise} \end{cases}$$

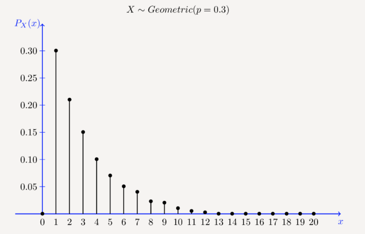


Geometric Distribution

Geometric Distribution

- Consider an experiment with only two outcomes: *success* and *failure*.
- The geometric distribution defines the number of Bernoulli trials (assuming independence) until the first success
- A random variable X is said to be a geometric random variable with parameter p, shown as $X \sim Geometric(p)$, if its PMF is given by

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & x \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$$

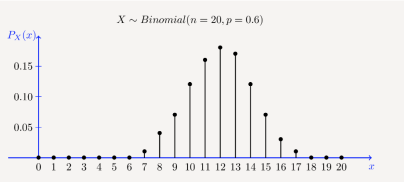


Binomial Distribution

Binomial Distribution

- Consider a Bernoulli experiment with only two outcomes: *success* and *failure*, which is repeated n times
- The binomial distribution defines the number of successes within the n independent Bernoulli trials
- A random variable X is said to be a binomial random variable with parameter p, shown as $X \sim Binomial(p, n)$, if its PMF is given by

$$P_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x \in \{1, 2, \cdots, n\} \\ 0 & \text{otherwise} \end{cases}$$

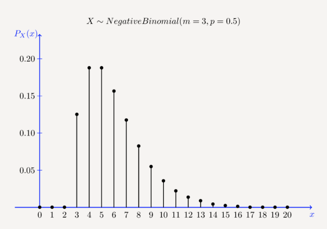


Negative Binomial (Pascal) Distribution

Pascal Distribution

- Consider an experiment with only two outcomes: *success* and *failure*.
- The Pascal distribution defines the number of Bernoulli trials (assuming independence) until we observe m successes
- A random variable X is said to be a Pascal random variable with parameter p, shown as $X \sim Pascal(p, m)$, if its PMF is given by

$$P_X(x) = \begin{cases} \binom{x-1}{m-1} p^m (1-p)^{x-m} & x \in \{m, m+1, m+2, \cdots\} \\ 0 & \text{otherwise} \end{cases}$$



Hypergeometric Distribution

Hypergeometric Distribution

- Consider a bag with b blue marbles and r red marbles where you select $k \leq b + r$ marbles at random
- The Hypergeometric distribution defines the number of blue marbles in your selected sample with the range of $R_X = \{\max(0, k-r) \leq X \leq \min(k, b)\}$
- A random variable X is said to be a Hypergeometric random variable with parameter p, shown as $X \sim Hypergeometric(p, m)$, if its PMF is given by

$$P_X(x) = \begin{cases} \frac{\binom{b}{x}\binom{r}{k-x}}{\binom{b+r}{k}} & x \in R_x \\ 0 & \text{otherwise} \end{cases}$$

The Poisson Random Variable

- Consider and interval T where an event of interest can occur at any time instant
- Let X count the number of times that the event of interest occurs during the interval T
- Then X follows the Poisson distribution
- Such an experiment can be considered as Bernoulli trials with $N \to \infty$ & $p \to 0$ such that $N \times p = \lambda$
- Binomial distribution $\Longrightarrow$ Poisson distribution as

$$\lim_{N \to \infty} N \times p = \lambda$$

- The Poisson distribution used to model
  - Number of telephone calls during a time interval
  - Number of photons emitted from a LED
  - Number of electrons emitted from a small section of a cathode

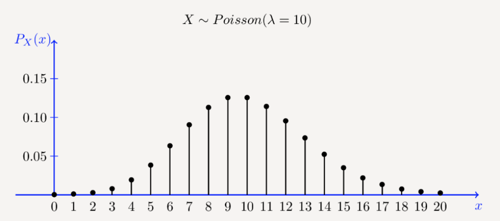
The Poisson Random Variable

Poisson density function

- A random variable X is called Poisson if its density function has the form

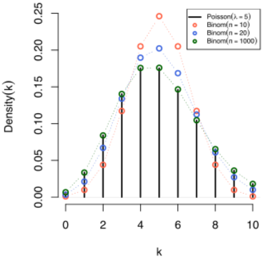
$$P_X(x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!} & x \in \{0, 1, 2, 3, \cdots\} \\ 0 & \text{otherwise} \end{cases}$$

- The Poisson random variable range is $0 \leq x \leq \infty$
- The Poisson density is parameterized with the rate λ

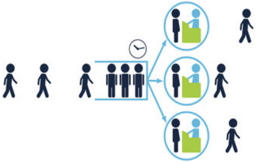


The Poisson Random Variable

- The Binomial to Poisson convergence for $\lambda = Np = 5$



- Poisson distributions is extensively used in queueing theory



Example

Example

- Consider Poisson car arrivals at gas station with rate 50 car per hour
- Each car requires 1 minute to fuel
- Find the probability that the cars queue

- Solution
- The cars will queue if two or more arrivals occurs during one minute
- The arrival rate of cars per minute is $\lambda = \frac{5}{6}$

$$
\begin{aligned}
\mathbb{P}(X > 1) &= \sum_{k=2}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \\
&= 1 - \sum_{k=0}^{1} \frac{\lambda^k e^{-\lambda}}{k!} \\
&= 1 - e^{-\frac{5}{6}} - \frac{5}{6} e^{-\frac{5}{6}} \\
&= 0.2032
\end{aligned}
$$

Progress...

- Last section
  - Popular discrete distributions
- Next section
  - Cumulative distribution function

Cumulative Distribution Function

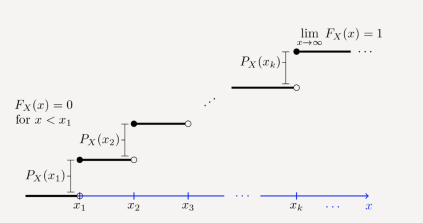
Cumulative Distribution Function (CDF)

- The CDF capture events in the form $\{X \leq x\}$

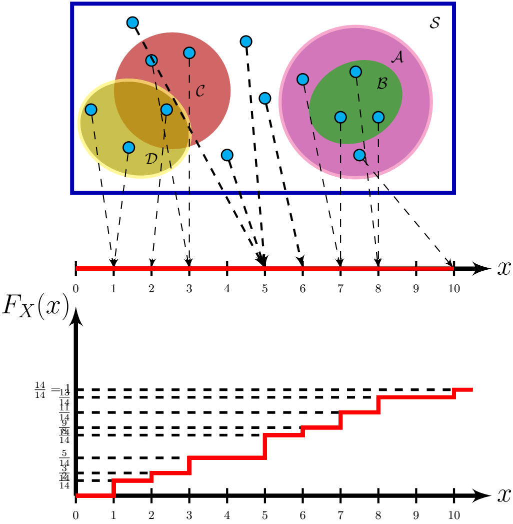
$$F_X(x) = \mathbb{P}(X < x) \quad \text{for all} \quad x \in R$$

- For discrete random variable, the CDF is in the form of staircase that starts at $F_X(-\infty) = 0$, jumps at each $x_k \in R_X$, and is equal to one at the end $F_X(\infty) = 1$

$$F_X(x) = F_X(x_k), \quad \text{for} \quad x_k \leq x < x_{k+1}$$



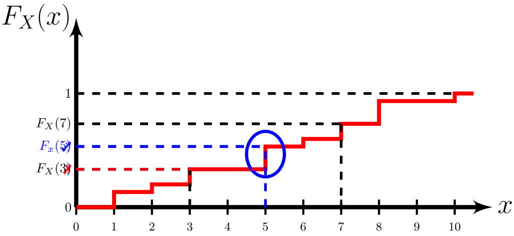
Cumulative Distribution Function



Properties of Distribution Functions

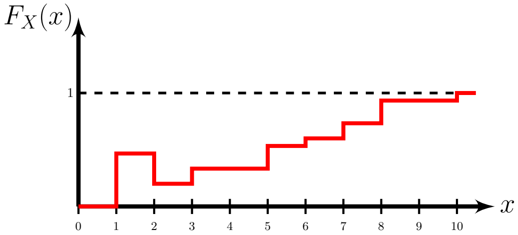
Properties of Distribution Functions

- The CDF has the following properties
  1. $F_X(-\infty) = 0$
  2. $F_X(\infty) = 1$
  3. $0 \leq F_X(x) \leq 1$
  4. $F_X(x_1) \leq F_X(x_2)$ if $x_1 < x_2$
  5. $\mathbb{P}(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$
  6. $F_X(x^+) = F_X(x)$



Exercise

- Is the following function is a valid CDF

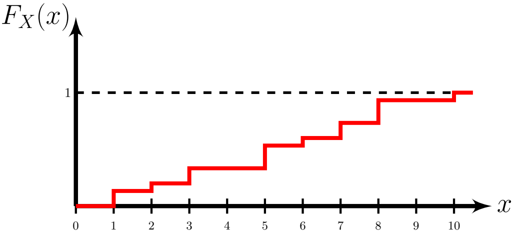


Discrete Random Variable

Distribution Functions for Discrete Random Variables

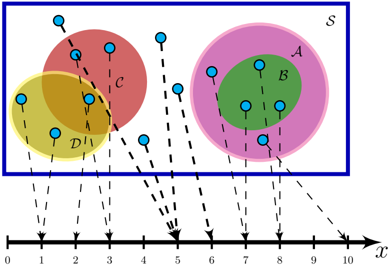
- The CDF can be expressed as

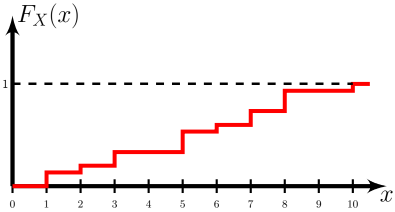
$$F_X(x) = \sum_{x_k \leq x} P_x(x_k)$$

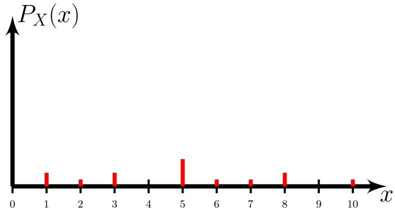


Density Function

- $f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}$
- $F_X(x) = \int_{-\infty}^{\infty} f_X(x)dx$







Example 1

- Consider an experiment of two times coin tossing and let X be the number of observed heads. Find the PMF and CDF of X

$$
\begin{aligned}
P_X(0) &= \tfrac{1}{4} \\
P_X(1) &= \tfrac{1}{2} \\
P_X(2) &= \tfrac{1}{4}
\end{aligned}
$$

- Hence,

$$
P_X(x) \begin{cases} \frac{1}{4}, & x = 0 \\ \frac{1}{2}, & x = 1 \\ \frac{1}{4}, & x = 2 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad F_X(x) \begin{cases} 0, & x < 0 \\ \frac{1}{4}, & 0 \le x < 1 \\ \frac{3}{4}, & 1 \le x < 2 \\ 1, & x \ge 2 \end{cases}
$$

Example 2

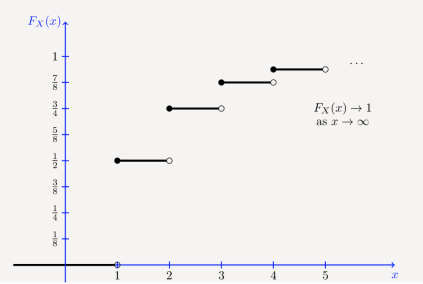
- Let X be a discrete random variable with range $R_X = \{1, 2, 3, \cdots\}$ and PMF

$$P_X(x) \begin{cases} \frac{1}{2^x} & x = \{1, 2, 3, \dots\} \\ 0 & \text{otherwise} \end{cases}$$

- Find
  - The CDF of X
  - The probability $\mathbb{P}(2 < X \leq 5)$
  - The probability $\mathbb{P}(X > 4)$
- The CFD is given by

$$F_X(x) \begin{cases} 0, & x < 1 \\ \frac{2^k - 1}{2^k}, & k \leq x < k + 1 \text{ for } k = 1, 2, 3, \cdots \end{cases}$$

Example 2



- The probability $\mathbb{P}(2 < X \leq 5)$ is given by

$$\mathbb{P}(2 < X \leq 5) = F_X(5) - F_X(2) = \frac{31}{31} - \frac{3}{4} = \frac{7}{32}$$

- The probability $\mathbb{P}(X > 4)$

$$\mathbb{P}(X > 4) = 1 - F_X(4) = 1 - \frac{15}{16} = \frac{1}{16}$$

Progress

- Last section
  - CDF and PDF
- Current section
  - Expectation

Expected value

- Descriptive probabilisitic events are converted to numeric values via random variables mapping
- Numeric values are more favorable so that we can apply functions and transformations to random events
- Expectation is the process of averaging when a random variable is involved, which is denoted as
  - Expectation of X
  - Expected value of X
  - Mean of X
  - Statistical average of X
- The expected value is denoted as $\mathbb{E}[X] = \bar{X}$
- The expectation notation $(\bar{\cdot})$ should not be confused with the complement operator

Example

- Consider that 90 people are randomly selected and the fractions of dollars in their pockets are counted. The following is the outcome

Number	8	12	28	22	15	5
Money	0.18	0.45	0.64	0.72	0.77	0.95

- What is the average money?

$$\text{Average } \$ = 0.18 \times \frac{8}{90} + 0.45 \times \frac{12}{90} + 0.72 \times \frac{22}{90} + 0.77 \times \frac{15}{90} + 0.95 \times \frac{5}{90}$$
$$= \$0.632$$

- Recall probability as a relative frequency
- Define a discrete random variable X as the number cents
- Then, X has the range $\{0 \leq x \leq 100\}$, and

$$\text{Average } \$ = \mathbb{E}[X] = \sum_{i=1}^{100} x_i P_X(x_i)$$

Expected value

Expectation

The mean value of X is calculated as follows

- For discrete random variable

$$\mathbb{E}[X] = \bar{X} = \mu_X = \sum_{x_i \in R_X} x_i P_X(x_i)$$

Example

- Find the expected value of $X \sim Bernoulli(p)$

$$\mathbb{E}[X] = 0 \times P_X(0) + 1 \times P_X(1) = p$$

- Find the expected value of $X \sim Geometric(p)$

$$\mathbb{E}[X] = \sum_{x=1}^{\infty} xp(1-p)^{x-1} = \frac{1}{p}$$

- Find the expected value of $X \sim Poisson(\lambda)$

$$\mathbb{E}[X] = \sum_{x=0}^{\infty} xe^{-\lambda}\frac{\lambda^x}{x!} = \lambda$$

Linearity of the Expectation

Expectation

The expectation is a linear operator. That is

- For discrete random variable X, we have

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

- For a set of random variables $X_1, X_2, \cdots, X_N$

$$\mathbb{E}[X_1 + X_2 + \cdots + X_N] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_N]$$

Example

- Find the expected value of $X \sim Binomial(p, n)$
- The Binomial can be interpreted as the sum of n independent Bernoulli random variables, hence,

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}[X_1 + X_2 + X_3 + \cdots + X_n] \\ &= \mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3] + \cdots + \mathbb{E}[X_n] \\ &= np\end{aligned}$$

- Find the expected value of a $X \sim Pascal(p, m)$
- The Pascal can be interpreted as the sum of m independent geometric random variables, hence,

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}[X_1 + X_2 + X_3 + \cdots + X_m] \\ &= \mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3] + \cdots + \mathbb{E}[X_m] \\ &= \frac{m}{p}\end{aligned}$$

Function of Random variable

Steps to get $P_Y(\cdot)$

1- Determine the range of X and the range of $Y = g(X)$

$$R_Y = \{g(x) | x \in R_X\}$$

2- Find the probability $P_Y(y) = \mathbb{P}(g(X) = y) = \sum_{x:g(x)=y} P_X(x)$

3- Make sure that $P_Y(y)$ is a valid PDF

Example

- Let X be a random variable with $P_X(x) = \frac{1}{5}$ for $k = -1, 0, 1, 2, 3$
- Let $Y = 2|X|$, find the PMF of Y
- The range of Y is $R_Y = \{0, 2, 4, 6\}$
- The PMF of Y is

$$P_Y(y) \begin{cases} \frac{1}{5}, & y = 0, 4, 6 \\ \frac{2}{5}, & y = 2 \\ 0, & \text{otherwise} \end{cases}$$

Expectation of a Function of Random variable

- Let $Y = g(X)$ be a function of a single random variable X
- Then Y is also a random variable

Expectation

- The mean value of $Y = g(X)$ is calculated as follows

$$\mathbb{E}[Y] = \mathbb{E}[g(X)] = \sum_{i=1}^{N} g(x_i)\mathbb{P}(x_i)$$

Moments

Moments around the origin

Averaging over the function $g(X) = X^n$ leads to the n^{th} **moment** for the random variable X, which is given by

$$m_n = \mathbb{E}[X^n] = \sum_{x \in R_X} x_i^n P_X(x)$$

It is clear that $m_0 = 1$ and $m_1 = \bar{X}$.

Moments

Central moments

Averaging over the function $g(X) = (X - X_n)^n$ leads to the n^{th} **central moment** (i.e., around the mean value $\bar{X}$) for the random variable X. The n^{th} central moment is given by

$$\mathbb{E}[(X - \bar{X})^n] = \sum_{x \in R_X} (x - \bar{X})^n P_X(x)$$

It is clear that $\mu_0 = 1$ and $\mu_1 = 0$.

Important Central Moments

Variance

- The ***variance*** is the second central moment μ_2, which measures the expected squared deviation of a random variable from its mean.
- The variance is usually denoted as σ_X^2
- The square root is the variance is denoted as the ***standard deviation***, which measures the spread if $f_X(x)$ round its mean value.
- The variance is calculated as

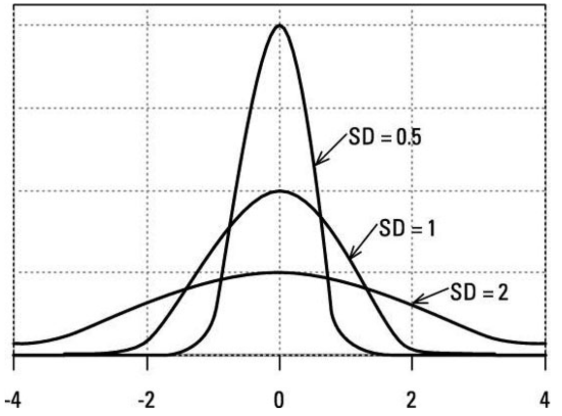
$$\sigma_X^2 = \mathbb{E}[(X - \bar{X})^2] = \sum_{x \in R_X} (x - \bar{X})^2 P_X(x)$$

- Alternatively

$$\begin{aligned}
\sigma_X^2 &= \mathbb{E}[(X - \bar{X})^2] = \mathbb{E}[X^2 - 2X\bar{X} + \bar{X}^2] \\
&= \mathbb{E}[X^2] - 2\mathbb{E}[X]\bar{X} + \bar{X}^2 \\
&= \mathbb{E}[X^2] - \mathbb{E}[X]^2
\end{aligned}$$

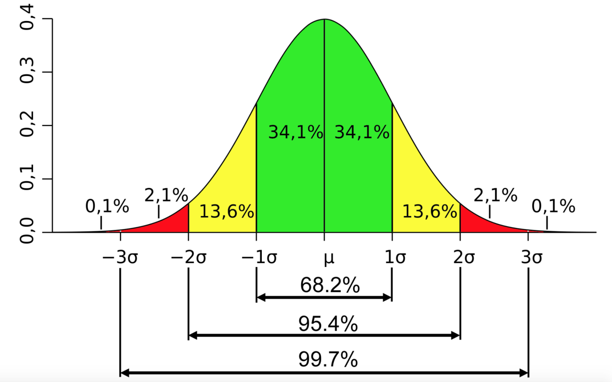
Important Central Moments

- Illustration of standard deviation in normal distribution



Important Central Moments

- Illustration of standard deviation in normal distribution



Variance

Useful results

- Let $Y = aX + b$, then
$$\text{Var}(Y) = a^2\text{Var}(X)$$
- Let $X_1, X_2, X_3, \cdots, X_n$ be indepenent random variables and let $X = X_1 + X_2 + X_3 + \cdots + X_n$, then
$$\text{Var}(X) = \text{Var}(X_1) + \text{Var}(X_2) + \cdots + \text{Var}(X_n)$$

- Find the variance of $X = Binomial(p, n)$
- Var(X)=n p (1-p)

Important Central Moments

Skew and Skewness

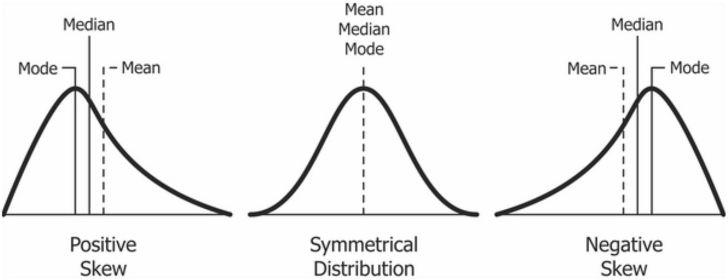
- The ***skew*** is the third central moment μ_3, which measures the asymmetry of $f_X(x)$ about the mean value $\bar{X}$.
- The skew is given by

$$\mu_3 = \mathbb{E}[(X - \bar{X})^3] = \sum_{x \in R_X} (x - \bar{X})^3 P_X(x)$$

- If $P_X(x)$ is symmetric, then $\mu_3 = 0$ (same applies for all μ_n with odd n)
- The normalized third central moment $\frac{\mu_3}{\sigma_X^3}$ is known as the ***skewness*** or ***coefficient of skewness*** of $P_X(x)$

Important Central Moments

- Illustration of skewed densities



Questions?