Probability and Random Process (SWE3026)

Discrete Random Variables

JinYeong Bak
jy.bak@skku.edu
College of Computing, SKKU

H. Pishro-Nik, "Introduction to probability, statistics, and random processes", available at https://www.probabilitycourse.com, Kappa Research LLC, 2014.

Special Distributions

Families of discrete random variable

Bernoulli RVs:

Example. Flip a coin {H,T}. Take an exam {Pass, Fail}.

$$X \sim Bernoulli(p)$$

PMF:

$$P_X(x) = \begin{cases} p & \text{for } x = 1 \\ 1-p & \text{for } x = 0 \end{cases} \quad \text{Range}(X) = \{0, 1\}.$$

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

Special Distributions

Geometric RVs:

$$X \sim Geometric(p)$$

$$R_X = \text{Range}(X) = \{1, 2, 3, \cdots\}$$

Random experiment: consider a coin with $P(H) = p$. Toss the coin repeatedly until the first heads is observed.

$X =$ The total number of coin tosses

$$X \sim Geometric(p)$$

Special Distributions

Definition. **A random variable X is said to be a *geometric* random variable with parameter p , shown as $X \sim Geometric(p)$, if**

$$R_X = \{1, 2, 3, \cdots\} = \mathbb{N},$$

$$P_X(k) = p(1-p)^{k-1}, \qquad k = 1, 2, 3, \ldots$$

Special Distributions

Definition. A random variable X is said to be a **binomial random variable** with parameters n and p $(P(H) = p)$, shown as $X \sim Binomial(n, p)$, if

$$R_X = \{0, 1, 2, ..., n\}$$

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, 2, \cdots, n$$

Special Distributions

Remember:

$$\underbrace{\text{HTH} \;\; \ldots \;\; \text{H}}_{n \text{ times}} \qquad X = k \;\Rightarrow\; \begin{matrix} k \;\; \text{Heads } (H) \\ n-k \;\; \text{Tails } (T) \end{matrix} \longrightarrow \binom{n}{k}$$

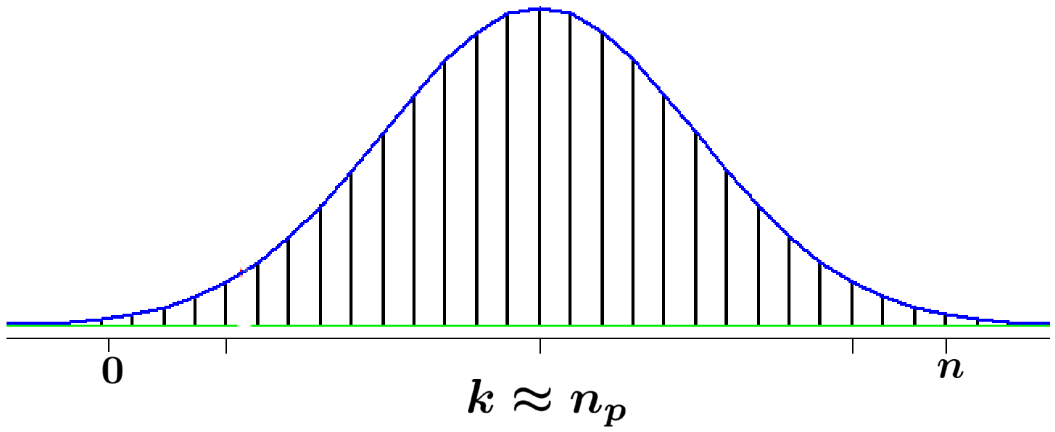
$$\underbrace{\text{HH} \;\; \ldots \; \text{H}}_{k \text{ Heads}}\underbrace{\text{TT} \;\; \ldots \;\; \text{T}}_{n-k \text{ Tails}} \qquad \longrightarrow \; p^k(1-p)^{n-k}$$

If $n = 1$, then $Binomial(1, p) = Binomial(p)$.

Special Distributions

Lemma. If $X_1, X_2, \ldots, X_n$ are independent $Bernoulli(p)$ random variable, then the random variable X define by $X = X_1 + X_2 + \ldots + X_n$, is a $Binomial(n, p)$ RV.

distribution.



Special Distributions

Pascal Distribution (Negative Binomial):

Example. You flip a coin until you observe m heads.

X **: total number of coin toss**

$$R_X = \{m, m+1, m+2, \cdots\}.$$

Find PMF.

- $A = \{X = k\}, \quad \text{or } A = B \cap C,$
- B **is the event that we observe $m-1$ heads in the first $k-1$ trials.**
- C **is the event that we observe a heads in the kth (the last) trial.**

Special Distributions

$$P(A) = P_X(k) = P(X = k)$$

$$P(A) = P(B \cap C) = P(B)P(C), \quad B \text{ and } C \text{ are independent events.}$$

$$P(C) = P(H) = p$$

Using binomial formula, $Binomial(n = k - 1, p)$

$$P(B) = \binom{k-1}{m-1} p^{m-1} (1-p)^{\left((k-1)-(m-1)\right)} = \binom{k-1}{m-1} p^{m-1} (1-p)^{k-m}.$$

$$P(A) = P(B \cap C) = P(B)P(C) = \binom{k-1}{m-1} p^{m} (1-p)^{k-m}.$$

Special Distributions

Definition. A random variable X is said to be a **Pascal random variable** with parameters m and p $(P(H) = p)$, shown as $X \sim Pascal(m, p)$, if

$$P_X(k) = \binom{k-1}{m-1} p^m (1-p)^{k-m} \qquad k = m, m+1, m+2, \ldots$$

Where $0 < p < 1$.

Special Distributions

Hypergeometric Distribution:

Example. You have a bag that contains b blue marbles and r red marbles. You choose $k \leq b + r$ marbles at random (without replacement).

X : The number of blue marbles in your sample

$$P_X(x) = P(\text{you observe } x \text{ blue marbles}) = \frac{|A|}{|S|} = \frac{\binom{b}{x}\binom{r}{k-x}}{\binom{b+r}{k}}.$$

Special Distributions

Poisson Random Variable:

Poisson RVs are used to model

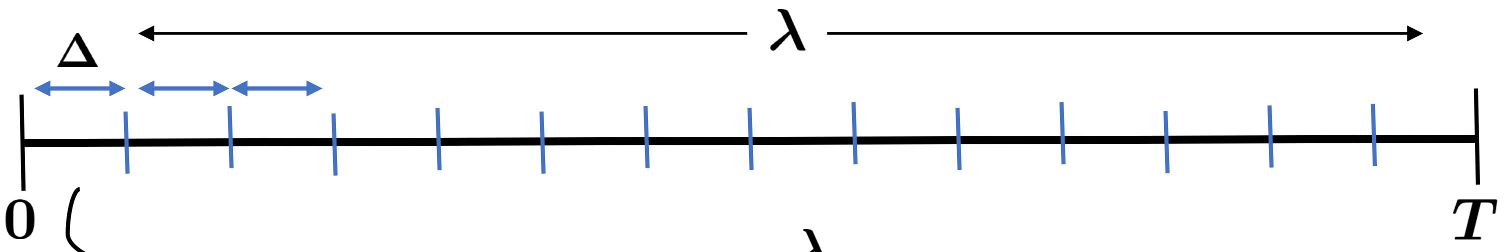
- Arrival of customers at a service facility
- Arrival of information request at a server

Counting the occurrence of certain events in an interval of time or space.

Special Distributions

Arrival of customers in an interval:

λ : the average number of arrivals in that interval



$$X \sim Bernoulli(p = \frac{\lambda}{n})$$

$$\Delta = \frac{T}{n}$$

Special Distributions

X : the total number of customers = $X_1, X_2, \cdots, X_n$

$$X \sim Binomial(n, p = \frac{\lambda}{n})$$

$$P(X = k) = P_X(k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} (\frac{\lambda}{n})^k (1 - \frac{\lambda}{n})^{n-k}.$$

Thus,

$$\lim_{n \to \infty} P_X(k) = \frac{e^{-\lambda} \lambda^k}{k!}.$$

Special Distributions

Definition. A random variable X is said to be a **Poisson random variable** with parameter λ, shown as $X \sim Poisson(\lambda)$, if

$$R_X = \{0, 1, 2, 3, \cdots\},$$

$$P_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}, \qquad k = 0, 1, 2, 3, \ldots$$

Special Distributions

Example. **The number of emails that I get in a weekday can be modeled by a Poisson distribution with an average of 0.2 emails per minute.**

a) **What is the probability that I get no emails in an interval of length 5 minutes?**

b) **What is the probability that I get more than 3 emails in an interval of length 10 minutes?**

Special Distributions

Definition. A random variable X is said to be a **Uniform random variable**, shown as $X \sim Uniform(R_X)$, if

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

$$P_X(x_i) = \frac{1}{|R_X|}.$$

Cumulative Distribution Function (CDF)

Definition. The cumulative distribution function (CDF) of random variable X is defined as

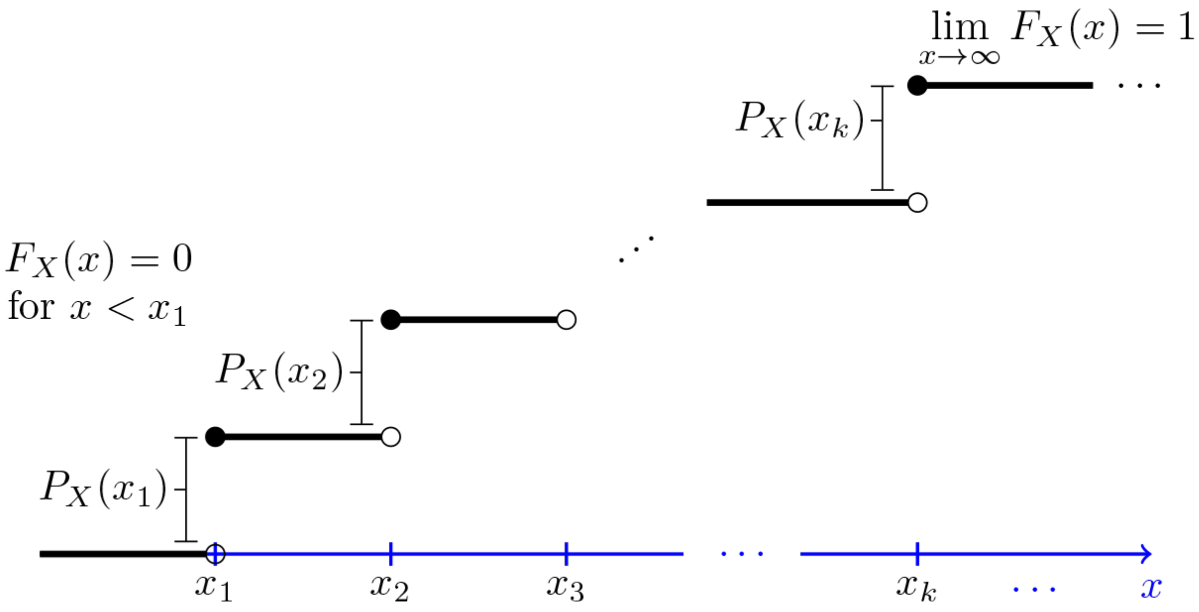
$$F_X(x) = P(X \leq x), \text{ for all } x \in \mathbb{R}.$$

Cumulative Distribution Function (CDF)

Example. **Toss a coin twice, let X be the number of observed heads. Find the CDF of X.**

Cumulative Distribution Function (CDF)

If X is a discrete random variable with range $R_X = \{x_1, x_2, x_3, \ldots\}$, such that $x_1 \leq x_2 \leq x_3 \leq \cdots$.



Cumulative Distribution Function (CDF)

Theorem. **Let X be a discrete random variable with range**

$$R_X = \{x_1, x_2, x_3, \ldots\}.$$

a) $F_X(-\infty) = P(Y < -\infty) = 0, \quad F_X(+\infty) = 1$

b) $y \geq x \;\Rightarrow\; F_X(y) \geq F_X(x)$

c) $x_i \in R_X, \quad F_X(x_i) - F_X(x_i - \epsilon) = P_X(x_i), \;\text{ For } \epsilon > 0 \text{ small enough.}$

d) $x_i \leq x < x_{i+1} \;\Rightarrow\; F_X(x) = F_X(x_i).$

Cumulative Distribution Function (CDF)

For all $a \leq b$, we have

$$P(a < X \leq b) = F_X(b) - F_X(a).$$

Cumulative Distribution Function (CDF)

Proof:

$$P(a < X \le b) = P(X \le b) - P(X \le a),$$

$$\underbrace{P(X \le b)}_{F_X(b)} = \underbrace{P(X \le a)}_{F_X(a)} + P(a < X \le b),$$

$$F_X(a) = P(X \le a) = P(X < a) + P(X = a),$$

$$\Rightarrow \; P(X < a) = F_X(a) - P(X = a).$$

Cumulative Distribution Function (CDF)

Example. Let X be a discrete random variable with range $R_X = \{1, 2, 3, ...\}$. Suppose the PMF of X is given by

$$P_X(k) = \frac{1}{2^k} \text{ for } k = 1, 2, 3, ...$$

a) Find and plot the CDF of X, $F_X(x)$.

b) Find $P(1 < X \leq 3)$.