Delay Models in Data Networks

Chapter 3

Introduction and Goals

- Average end-end delays
  - Most important performance measure in a data network
  - Sum of delays on each link
  - Strongly influences choice of flow control and routing algorithms
  - Focus: subnet → network layer
- Components of link delay:
  - Processing delay
  - Queuing delay
  - Transmission delay
  - Propagation delay
- Our focus: queuing and transmission delays

Goals:

- Analysis of system (network) load and performance characteristics
  - Response time
  - Throughput
- Performance tradeoffs are often not intuitive
  - Queuing theory often makes analysis very straightforward
- Examples:
  - How to analyze changes in network workloads? Should I add new terminals? How much?
  - What percentage of calls will be blocked? Adding more lines would solve the problem?

Outline of Topics

- Review of probability theory and random variables
- Review of stochastic processes
- Queuing models
- Little's theorem
- Analysis of single queue models
- Network of queues
- Jackson's theorem

Background Needed for Queuing Analyses

- Basic probability theory
  - Basic definitions
  - Conditional probability
- Random variable concepts
  - Discrete random variables and their characteristics
  - Continuous random variables and their characteristics
  - Joint distributions
- Stochastic processes
  - Poisson processes
  - Markov processes
- Goal of this review:
  - Provide a brief review of topics useful to us
  - Set stage for queuing analyses → delay models in data networks

Review of Probability Theory and Random Variables

Basic Definitions

- Experiment
  - A process whose outcome is not known with probability 1
- Sample space
  - Ω = set of all possible outcomes
- Sample point
  - Outcomes themselves are sample points in Ω
- Example:
  - A coin flipping experiment
  - $\Omega = \{H, T\}$
- Example:
  - Tossing a die:
  - $\Omega = \{1, 2, \ldots, 6\}$

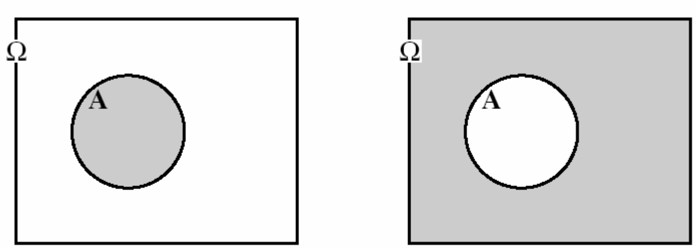
Axioms of Probability

- Common axioms:
  1. $0 \leq Pr[A] \leq 1$ for each event A
  2. $Pr[\Omega] = 1$
  3. $Pr[\varnothing] = 0$
  4. $Pr[A \cup B] = Pr[A] + Pr[B]$ if A and B are mutually exclusive

- Important rules
  1. Pr[Á] = 1 - Pr[A]
  2. $Pr[A \cap B] = 0$ if A and B are mutually exclusive
  3. $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$
  4. $Pr[A \cup B \cup C] = Pr[A] + Pr[B] + Pr[C] - Pr[A \cap B] - Pr[A \cap C] - Pr[B \cap C] + Pr[A \cap B \cap C]$

- Can be understood through Venn Diagrams

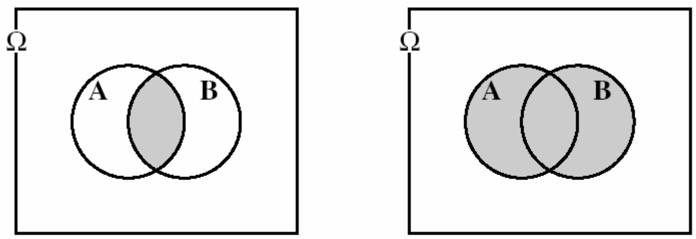
Venn Diagrams



(a) A (b) NOT A
$\overline{A}$

Complementation

Venn Diagrams

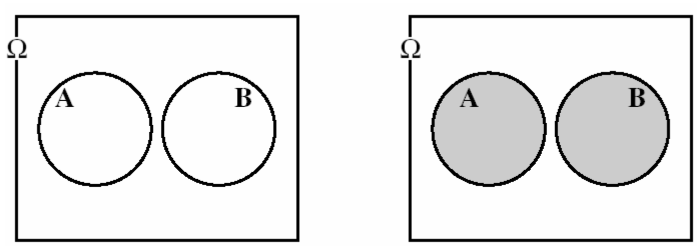


(c) A AND B
$A \cap B$

(d) A OR B
$A \cup B$

Intersection

Venn Diagrams

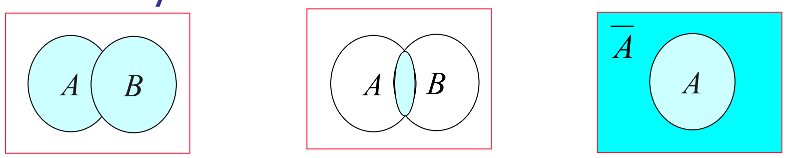


(e) A AND B
$A \cap B$

(f) A OR B
$A \cup B$

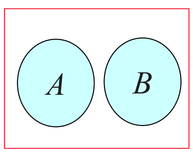
Mutual Exclusivity

Mutually Exclusive Events



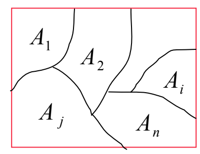
$A \cup B$ $\quad$ $A \cap B$ $\quad$ $\bar{A}$

- If $A \cap B = \phi$, the empty set, then A and B are said to be mutually exclusive (M.E).
- A partition of Ω is a collection of mutually exclusive subsets of Ω such that their union is Ω.



$A \cap B = \phi$

$$A_i \cap A_j = \phi, \quad \text{and} \quad \bigcup_{i=1} A_i = \Omega.$$



Conditional Probability

- The conditional probability of an event A, given that event B has occurred is:

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

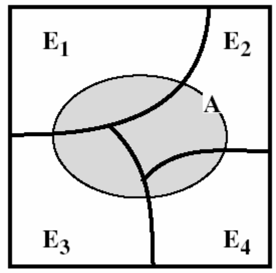
When $\Pr[B] \neq 0$

- When A and B are independent events
  - $\Pr[A \cap B] = \Pr[A]\Pr[B]$

Total Probability

- Given a set of mutually exclusive events $E_1, E_2, \ldots, E_n$ covering all possible outcomes, and
  $$E_i \cap E_j = \phi, \quad \text{and} \quad \bigcup_{i=1}^{n} E_i = \Omega.$$
- Given an arbitrary event A, then:

$$\Pr[A] = \sum_{i=1}^{n} \Pr[A \mid E_i] \Pr[E_i]$$



- Baye's theorem is applied when entire sample space can be subdivided into `n` mutually exclusive events

Baye's Theorem

- Let $E_1, E_2, \ldots, E_n$ covering all possible outcomes, such that
  $$E_i \cap E_j = \phi, \quad \text{and} \quad \bigcup_{i=1}^{n} E_i = \Omega.$$
  Let `A` be an event such that P[`A`] ≠ 0. Then for any other event E_j where $j = 1, 2, \ldots, n$

$$\Pr[E_j \mid A] = \frac{\Pr[A \mid E_j] \Pr[E_j]}{\sum_{i=1}^{n} \Pr[A \mid E_i] \Pr[E_i]}$$

- Qualitatively, Baye's theorem provides the probability that an event really occurred provided the evidence in favor of it
  - Pr[A] represents the a-priori probability of event A

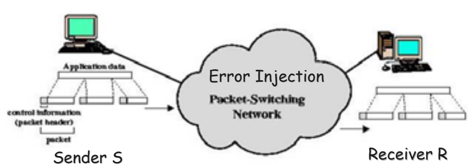
Example: Baye's Theorem

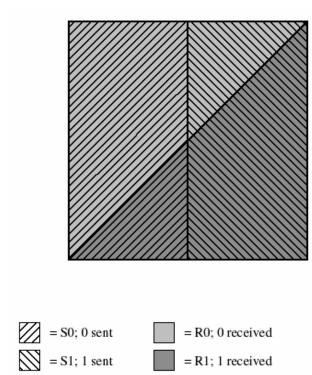
- Hit & run accident involving a taxi
- 85% of taxis are yellow, 15% are blue
- Eyewitness reported that the taxi involved in the accident was blue (event WB)
- Data shows that eyewitnesses are correct on car color 80% of the time
- What is the probability that the taxi was blue?

$$\Pr[Blue|WB] = \frac{\Pr[WB|Blue]\,\Pr[Blue]}{\Pr[WB|Blue]\,\Pr[Blue] + \Pr[WB|Yellow]\,\Pr[Yellow]}$$

$$= \frac{(0.8)(0.15)}{(0.8)(0.15) + (0.2)(0.85)} = 0.41$$

Bayes's Theorem: Another Example





- Network injects errors (flips bits)
- Assume Pr[S1] = Pr[S0] = p = 0.5
- Assume Pr[R1] = Pr[R0] = (1-p) = 0.5
- Given error injection, such that $\Pr[R0|S1] = p_a$ and $\Pr[R1|S0] = p_b$, then :

$$\Pr[S1|R0] = \frac{\Pr[R0|S1].\Pr[S1]}{\Pr[R0|S1].\Pr[S1] + \Pr[R0|S0].\Pr[S0]} = \frac{p_a p}{p_a p + (1-p_b)(1-p)}$$

Random Variables

- Association of real numbers with events
- A random variable X is a function that assigns a real number to every outcome in a sample space, and satisfies the following conditions:
  1. The set $\{X \le x\}$ is an event for every x
  2. $Pr[X= \infty] = Pr[X = -\infty] = 0$
- A random variable can be one of two types:
  - Continuous or
  - Discrete

Random Variables (Cont'd)

- A variable is called a random variable (RV) if it takes one of a specified set of values with a specified probability
- Formally, a non-negative integer-valued random variable X has:
  - A range in {0, 1, 2, 3, …. } and
  - A probability mass function $p_k = P\{X=k\}$ for k=0, 1, 2, …
    - We want to allow $k=\infty$ as X might represent waiting time; thus:

$$p_{\infty} = P\{X = \infty\} = 1 - \sum_{k=0}^{\infty} p_k = 1 - P\{X < \infty\} = 1 - \text{Finite Probabilty}$$

- Important property of a finite RV (i.e., $p_{\infty}=0$):

$$\sum_{k=0}^{\infty} p_k = 1$$

Random Variables: Examples

- Example: assigning a value to each outcome of an experiment
  - Consider the experiment of rolling a pair of dice
  - Ω = {(1,1), (1,2), …, (6,6)}
  - X is a random variable corresponding to the sum of a pair (i,j), that is `X = i + j` for all `i`, `j` in the range of 1, …, 6.
  - For example, `X` = 7 for (4, 3) pair
- Example:
  - Consider flipping of two coins
  - Ω = {(H, H), (H, T), (T, H), (T, T)}
  - X is a random variable corresponding to the number of heads
  - Then, X = 1 for either of (H, T) or (T, H) outcomes

Characteristics of a Random Variable

- Continuous random variables can be described by:
  - Either a distribution function or
  - A density function
- Discrete random variables are described by a probability function $P_x(k)$
- Random variable characteristics:
  - Mean value
  - Second moment
  - Variance
  - Standard deviation

CDF and PDF of a RV

- The Cumulative Distribution Function (CDF):
  CDF of a random variable maps a given value *x* to the probability of the variable taking a values less than or equal to *x*

$$F(X) = P\{X \le x\}$$

- Probability Density Function (PDF):
  PDF of a RV X is given as a derivative of its CDF: $f(X) = \frac{dF(X)}{dX}$

  Given a PDF f(x), the probability of X to be in the interval (x_1, x_2) is given as:

$$P\{x_1 < X \le x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(X)dX$$

Characteristics of a CDF

- The CDF of a RV X is represented as F(x) and is defined for each real number x as:

$$F(x) = P(X \le x) \text{ for } -\infty < x < \infty$$

- Properties:
  1. F(x) is in range of (0, 1): $0 \le F(x) \le 1, \forall x$
  2. F(x) is non-decreasing: If $x_1 < x_2$ then $F(x_1) \le F(x_2)$
  3. Limiting values: $\lim_{x \to \infty} F(x) = 1$ and $\lim_{x \to -\infty} F(x) = 0$

Probability Mass Function (PMF)

- For discrete RVs, the CDF is not continuous
  - Cannot be differenciated
- PMF is used instead of PDF for discrete RVs
- Let X be a discrete RV that can take n distinct values:
  - $X \in \{x_1, x_2, \ldots, x_n\}$ with probabilities $\{p_1, p_2, \ldots, p_n\}$, respectively
  - The PMF maps x_i to p_i as:
    $f(x_i) = p_i$
- The probability of X being in the interval (x_1, x_2) can be computed as:

$$P\{x_1 < X \le x_2\} = F(x_2) - F(x_1) = \sum_{\substack{i \\ x_1 < X \le x_2}} p_i$$

Discrete RV

- Definition:
  - X s a discrete RV if it can take on at most a countable number of values $\{x_1, x_2, \ldots, x_n\}$ such that
    $p(x_i) = P(X = x_i)$ for $i = 1,2,\ldots.$ and $\sum_{i=1}^{\infty} p(x_i) = 1$
- PMF of a discrete RV X:
  - If $I = [a, b]$ such that $a \le b$ and a, b are real number, then
    $P(X \in I) = \sum_{a \le x_i \le b} p(x_i)$
- CDF of a discrete random variable X is given as:
  $F(x) = \sum_{x_i \le x} p(x_i), \forall -\infty < x < \infty$
- Necessary and sufficient conditions for a function f(x) to be a discrete density:

$$f(x) \ge 0 \text{ and } \sum_{\text{all } x} f(x) = 1$$

Example: Discrete RV

- Inventory system
  - Let X be the size of demand for a product (a discrete RV)
  - X = 1, 2, 3, 4 with probabilities 1/6, 1/3, 1/3, and 1/6, respectively
- PMF:
  - P(X=1) = 1/6 = p(1)
  - P(X=2) = 1/3 = p(2)
  - P(X=3) = 1/3 = p(3)
  - P(X=4) = 1/6 = p(4)
- CDF: $F(x) = 0 + \sum_{x_i \leq 4} p(x_i) + 0$
  - P[X≤1] = p(1) = 1/6
  - P[X≤2] = p(1) + p(2) = ½
  - P[X≤3] = p(1) + p(2) + P(3) = 5/6
  - P[X≤4] = 1

Continuous RVs

- Definition:
  - A random variable X is continuous if there exists a non-negative function f(x) such that for any set of real numbers B

    $$1)\ P(X \in B) = \int_B f(x)dx \text{ and}$$

    $$2)\ \int_{-\infty}^{\infty} f(x)dx = 1$$

    f(x) is called Probability Density Function (PDF) for X
  - f(x) is not a probability as in the case of a PMF
- For any real number x:

  $$P(X = x) = P(X \in [x,x]) = \int_x^x f(y)dy = 0$$
- Interpretation of f(x):

  $$\text{For } \Delta x > 0,\ P(X \in [x, x+\Delta x]) = \int_x^{x+\Delta x} f(y)dy$$

Continuous RVs: CDF

- The distribution function F(x) of a continuous RV X is given by:

$$F(x) = P(X \in [-\infty, x]) = \int_{-\infty}^{x} f(y)dy \ , \forall -\infty < x < \infty$$

  - Then, f(x) = F′(x)

- If `I=(a, b)` where `a, b` $\in \Re$ such that `a<b`, then

$$P(x \in I) = \int_{a}^{b} f(y)dy = F(b) - F(a)$$

- If Δx>0 then

$$P(X \in [x, x+\Delta x]) = \int_{x}^{x+\Delta x} f(y)dy = F(x+\Delta x) - F(x)$$

Example: Continuous RV

- A uniform RV on interval [0, 1] has PDF:

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

- Also, if $0 \le x \le 1$ then the CDF is computed as:

$$F(x) = \int_{0}^{x} f(y)dy = \int_{0}^{x} 1.dy = x$$

- If $0 \le x \le x+\Delta x \le 1$ then

$$P(X \in [x, x+\Delta x]) = \int_{x}^{x+\Delta x} f(y)dy = F(x+\Delta x) - F(x)$$

$$= x + \Delta x - x = \Delta x$$

Another Example: Continuous RV

- An exponential RV with mean β and PDF given as:

$$f(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}} \text{ for } x \geq 0$$

- Then, the CDF is derived as:

$$\begin{aligned} F(x) &= \int_0^x f(y)dy = \int_0^x \frac{1}{\beta} e^{-y/\beta} dy \\ &= \frac{1}{\beta} \left| \beta e^{-y/\beta} \right|_0^x = e^0 - e^{-x/\beta} \\ &= 1 - e^{-x/\beta} \text{ for } x \geq 0 \end{aligned}$$

Joint Distributions

- Instead of a single RV, we may have to deal with n RVs simultaneously: $X_1, X_2, \ldots, X_n$
  - Example: service times in a queue $S_1, S_2, \ldots, S_n$
  - Example: Delays in a queue $D_1, D_2, \ldots, D_n$
- Let n=2
  - If X, Y are discrete RVs, then let $p(x, y) = P(X = x, Y = y) \quad \forall x, y$ then p(x, y) is called joint probability mass function of X and Y
- Marginal PMFs of X and Y:
  - If X and Y are independent:

$$p(x, y) = p_x(x) p_y(y) \quad \forall x, y$$

where

$$p_X(x) = \sum_y p(x, y) \quad \text{and}$$

$$p_Y(y) = \sum_x p(x, y)$$

Example: Joint Distributions

- Suppose that X and Y are jointly discrete RVs with

$$p(x,y)=\begin{cases}\frac{xy}{27} & \text{for } x=1,2;\, y=2,3,4\\ 0 & \text{Otherwise}\end{cases}$$

- Then, we can calculate marginal PMFs as:

$$p_X(x)=\sum_{y=2}^{4}\frac{xy}{27}=\frac{2x}{27}+\frac{3x}{27}+\frac{4x}{27}=\frac{x}{3} \text{ for } x=1,2$$

$$p_Y(x)=\sum_{x=1}^{2}\frac{xy}{27}=\frac{y}{27}+\frac{2y}{27}=\frac{y}{9} \text{ for } y=2,3,4$$

$$\because p_X(x)p_Y(y)=\frac{xy}{27}=p(x,y)\ \ \forall x,y$$

$$\Rightarrow X \text{ and } Y \text{ are independent}$$

Continuous Joint Distribution

- Definition:
  - X and Y are jointly continuous if there exists a non-negative function f(x,y), called joint PDF of X and Y, such that for all sets of real numbers A and B,

    $$P(X\in A, Y\in B)=\int_B\int_A f(x,y)dxdy$$
- Independence:
  - X and Y are independent if

    $$f(x,y)=f_X(x)f_Y(y)\quad \forall x,y$$

    where marginal PDFs of X and Y are given as:

    $$f_X(x)=\int_{-\infty}^{\infty}f(x,y)dy$$

    $$f_Y(y)=\int_{-\infty}^{\infty}f(x,y)dx$$

Example: Continuous Joint Distribution

- X and Y are jointly continuous RVs with

$$f(x,y)=\begin{cases}24xy & \text{for } x\ge 0, y\ge 0, \text{and } x+y\le 1\\ 0 & \text{Otherwise}\end{cases}$$

- Then, marginal PDFs are given as:

$$f_X(x)=\int_0^{1-x}24xy\,dy=12x(1-x)^2 \quad \text{for } 0\le x\le 1$$

$$f_Y(y)=\int_0^{1-y}24xy\,dx=12y(1-y)^2 \quad \text{for } 0\le y\le 1$$

$$\because f(\tfrac{1}{2},\tfrac{1}{2})=24(1/2)(1/2)=6\neq f_X(1/2)f_Y(1/2)=9/4$$

$$\therefore X \text{ and } Y \text{ are NOT independent}$$

Mean or Expected Values of a RV

- For a discrete RV X, we have already seen that:

$$\text{Mean } \mu=E(X)=\sum_{j=1}^{\infty}x_j p_X(x_j)$$

- For a continuous RV X:

$$\text{Mean } \mu=E(X)=\int_{-\infty}^{\infty}xf(x)dx$$

- Properties:

$$(1)\ E[CX]=C\,E[X] \text{ for } C\in\Re \text{ and is a constant}$$

$$(2)\ E\left[\sum_{i=1}^{n}C_iX_i\right]=\sum_{i=1}^{n}C_iE[x_i] \text{ even when } x_i\text{'s are dependent}$$

Example: Expected Values

- Discrete case:
  - Let X be a discrete RV such that

$$X = \begin{cases} 1 & \text{w.p.} \quad 1/6 \\ 2 & \text{w.p.} \quad 1/3 \\ 3 & \text{w.p.} \quad 1/3 \\ 4 & \text{w.p.} \quad 1/6 \end{cases}$$

  - Then, μ = E[X] = 1 (1/6) + 2(1/3) + 3(1/3) + 4(1/6) = 5/2
- Continuous case:
  - `X ~ U[0,1]` with PDF `f(x)` given as:

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

  - Then

$$\mu = E[X] = \int_0^1 x dx = 1/2$$

Variance of a RV

- For a RV X and its mean μ, $(X-\mu)^2$ represents the square of distance between them
- The expected value of above quantity is called the variance of X and is given as:

$$Var(X) = E[(X-\mu)^2] = \sum_{i=1}^{n} p_i (X_i - \mu)^2 = \int_{-\infty}^{\infty} (X_i - \mu)^2 f(X) dX$$

  - The summation is for discrete RVs while integral is used for continuous RVs
  - Variance is traditionally denoted by σ^2
- The square root of variance is known as standard deviation and is denoted by σ

Properties of Variance

- Variance of a RV X is given as

$$\sigma^2 = E\left[(X_i - \mu_i)^2\right] = E[X_i^2] - \mu_i^2$$

- `Var(X)` ≥ 0
- `Var(CX)` = `C`^{2} `Var(X)`
- Also,

$$Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var(X_i) \text{ if } X_i\text{'s are independent}$$

Example: Variance

- Discrete RV
  - `X` = 1, 2, 3, 4 w.p. 1/6, 1/3, 1/3, and 1/6, respectively
  - Find `E[X`^{2}`]`:

$$E[X^2] = \sum_{i=1}^{4} x^2 p_x = 43/6$$

  - Then calculate `Var(X) = E[X`^{2}`] -` μ^2 = 43/6 – (5/2)^2 = 11/12
- Continuous RV
  - `X ~ U[0,1]` with pdf `f(x)` = 1 for $0 \leq x \leq 1$ and 0 elsewhere

$$E[X^2] = \int_0^1 x^2 f(x)dx = 1/3$$

  - Then calculate `Var(X) = E[X`^{2}`] -` μ^2 = 1/3 – (1/2)^2 = 1/12

C.O.V. and Convariance of a RV

- The ratio of the standard deviation to the mean is called the Coefficient of Variation (C.O.V.):

$$\text{C.O.V.} = \frac{\text{Standard deviation}}{\text{Mean}} = \frac{\sigma}{\mu}$$

- Covariance:
  Given two RVs X and Y with means μ_X and μ_Y, their covariance is:

$$Cov(X, Y) = \sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - E(X)E(Y)$$

  - For independent variables, covariance is zero as E(XY) = E(X)E(Y)
  - Zero covariance does not necessarily mean independent variables

Correlation Coefficient of a RV

- The normalized value of covariance is called the correlation coefficient or simply the correlation

$$Correlation(X, Y) = \rho_{XY} = \frac{\sigma^2{}_{XY}}{\sigma_X \sigma_Y}$$

- The correlation always lies between –1 and 1

Mean and Variance of Sums of RVs

- Let $X_1, X_2, \ldots, X_n$ be n RVs. If $a_1, a_2, \ldots, a_n$ are n arbitrary constants (weights), then

$$E(a_1X_1 + a_2X_2 + \ldots + a_nX_n) = a_1E(X_1) + a_2E(X_2) + \ldots + a_nE(X_n)$$

- For independent variables:

$$Var(a_1X_1 + a_2X_2 + \ldots + a_nX_n) = a_1^2Var(X_1) + \ldots + a_n^2Var(X_n)$$

Commonly Used Types of RVs

- Bernoulli
- Geometric
- Binomial
- Exponential
- Poisson

Bernoulli RVs

- It is the simplest discrete distribution
- A Bernoulli RV X can take only two values, usually denoted as failure (X=0) and success (X=1)
  - If, P{success} = P{X=1} = p; then
    P{failure} = P{X=0} = q = 1-p
  - PMF: $$f(X) = \begin{cases} 1-p & \text{if X=0} \\ p & \text{if X=1} \\ 0 & \text{Otherwise} \end{cases}$$
  - Mean: E(X) = p
  - Variance: Var(X) = p(1-p)
- Examples of models using Bernoulli RVs:
  - A computer system is up or down
  - A network packet reaches its destination or is lost

Geometric RVs

- This is the distribution of number of trials up to and including the first success in a sequence of Bernoulli trials
- Key characteristics:

$$\text{PMF: } f(X) = (1-p)^{X-1}p \qquad \text{for } X = 1, 2, \ldots, \infty$$

$$\text{CDF: } F(X) = 1-(1-p)^{X}$$

$$\text{Mean: } E(X) = \frac{1}{p}$$

$$\text{Variance: } Var(X) = \frac{1-p}{p^2}$$

- Geometric distribution is the discrete equivalent of exponential distribution
  - It has memoryless property
  - Results of past attempts does not affect the outcome of future attempts
- Examples:
  - Number of local queries to a database between successive accesses to the remote database
  - Number of packets successfully transmitted between those requiring retransmission

Memoryless Property of Geometric Distribution

- Memoryless

$$P[X>x+\Delta x \mid X>x]=P[X>\Delta x] \text{ for all } x,\Delta x \geq 0$$
$$P[X>\Delta x]=P[X>x+\Delta x, X>x]/\,P[X>x]$$
$$P[X>x+\Delta x]= P[X>x]P[X>\Delta x]$$

- $P[X>n]$ $= p(1-p)^n[1+(1-p)+(1-p)^2+\ldots]$
  $= p(1-p)^n\{1/[1-(1-p)]\} = (1-p)^n$

- $P[X>n+m, X>n]/P[X>n]=(1-p)^{n+m}/(1-p)^n=(1-p)^m$
  $\Rightarrow P[X>n+m \mid X>n]=P[X>m]$

Exponential RV

- It is the only continuous distribution with memoryless property
- Key characteristics:

PDF: $f(\tau) = \lambda e^{-\lambda\tau}$ for $\frac{1}{\lambda} > 0$ and $0 \leq \tau \leq \infty$

CDF: $F(\tau) = 1 - e^{-\lambda\tau}$

Mean: $E(\tau) = \frac{1}{\lambda}$

Variance: $Var(\tau) = \frac{1}{\lambda^2}$

- Examples:
  - Time between successive requests to a device
  - Time between failures of a device

Exponential RV (Cont'd)

- Exponential RV T is the continuous version of Geometric RV X: $P_X(i)=p(1-p)^{i-1}$
- $P_T(t=i/n)=p(1-p)^{i-1}=p(1-p)^{nt-1}$
- Assume $n\to\infty$, $p\to 0$, and $np\to\lambda, 1/n\to\Delta t$
  $P_T(t=i/n) \to (\lambda/n)(1-\lambda/n)^{nt} \to \lambda e^{-\lambda t}\Delta t$
- $P[T>t] = \int_t^{\infty}\lambda e^{-\lambda t}dt = e^{-\lambda t}$
- $P[T>t+\Delta t \mid T>t] = P[T>t+\Delta t]/P[T>t]$
  $= e^{-\lambda(t+\Delta t)}/e^{-\lambda t} = e^{-\lambda\Delta t} = P[T>\Delta t]$
- $E[T] = \int_0^{\infty}\lambda e^{-\lambda t}t\,dt = 1/\lambda$
- $Var[T] = E[T^2] - E[T]^2 = 1/\lambda^2$

Binomial RV

- The number of successes X in a sequence of n Bernoulli trials has a binomial distribution
- It has two parameters:
  - p = probability of success in a trial, 0<p<1
  - n = number of trials; n must be a positive integer
- Key characteristics:
  - PMF: $f(X) = \binom{n}{X}p^X(1-p)^{n-X}$ for $X = 0, 1, ..., n$
  - Mean: E(X) = np
  - Variance: Var(X) = np(1-p)
- Examples:
  - Number of processors that are up in a multiprocessor system
  - Number of packets that reach destination without loss

Binomial RV (Cont'd)

- *n* independent trails.
  - X represents the number of success that occur in *n* trails. X is a Binomial RV with parameters (n,p).
- Using CRC to test *n* frames:
  - Each frame is error-free with probability *p*, independent of the results of other tests.
  - X equals the number of error-free frames in the tests.
- Binomial RV denotes the probability of x success among n trails
- Question: Compute E[Y] and Var[Y]?

Binomial RV (Cont'd)

- $E[Y] = E[X_1+X_2+\ldots+X_n]$
  $= E[X_1]+E[X_2]+\ldots+E[X_n] = np$
- $Var[Y] = Var[X_1+X_2+\ldots+X_n]$
  $= Var[X_1]+Var[X_2]+\ldots+Var[X_n] = np(1-p)$
- If $n\rightarrow\infty$, $p\rightarrow 0$, and $np\rightarrow\alpha$, then a Binomial RV approximates a Poisson RV.

Poisson RV

- It is a limiting form of binomial RV
- Probability model of a Poisson RV: phenomena that occur randomly in time
- Key characteristics:
  - PMF: $f(x) = P\{X = x\} = \lambda^x \frac{e^{-\lambda}}{x!}$ for $X = 0, 1, 2, \ldots, \infty$
  - Mean: $E(X) = \lambda$
  - Variance: $Var(X) = \lambda$
- Examples:
  - Arrival of information requests at a WWW server
  - Initiation of telephone calls

Poisson RV (Cont'd)

- An arrival
- A Poisson model:
  - An average arrival rate λ and a time interval T.
  - The number of arrivals X has a Poisson PMF with $\alpha=\lambda T$.
- Question: Compute E[X] and Var[X]?

- $\phi_X(s) = e^{-\alpha}[1+(\alpha e^s/1!)+(\alpha^2 e^{2s}/2!)+\ldots]$
  $= e^{-\alpha}e^{\alpha e^s} = e^{\alpha(e^s-1)}$
- $E[X] = \phi'_X(s)\big|_{s=0} = \alpha$
- $Var[X] =\{\phi''_X(s)-[\phi'_X(s)]^2\}\big|_{s=0} = \alpha$
- Relationship between Poisson RV and Binomial RV?

Binomial RV ⇒ Poisson RV

- Binomial RV ($n \to \infty$, $p \to 0$, and $np \to \alpha$)

$$P_{Binomial}[x=i] = \binom{n}{i} p^i (1-p)^{n-i}$$

$$= \{n!/[(n-i)!\,i!]\}(\alpha/n)^i (1-\alpha/n)^{n-i}$$

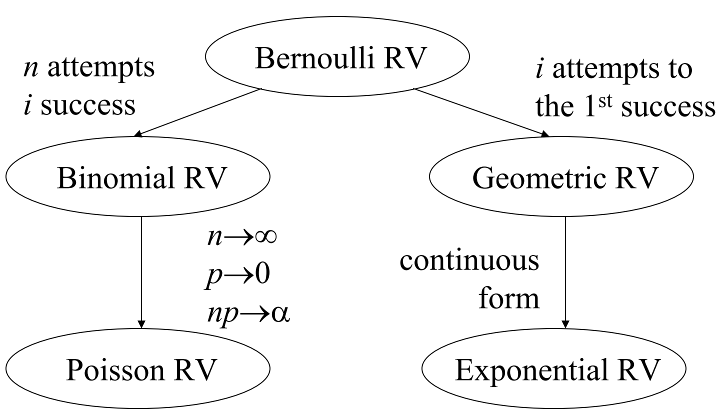
$$= [n(n-1)\ldots(n-i+1)/n^i](\alpha^i/i!)(1-\alpha/n)^n/(1-\alpha/n)^i$$

$$\to 1 \qquad\qquad \to e^{-\alpha} \qquad \to 1$$

$\Rightarrow P_{Binomial}[x=i] \to e^{-\alpha}\alpha^i/i!$

- $E_{Binomial}[X]=np \quad \to \quad E_{Poisson}[X]=\alpha$
- $Var_{Binomial}[X]=np(1-p) \to \quad Var_{Poisson}[X]=\alpha$

Relationship of RVs



Probability Distributions—Relevance to Networks

- Service times of queues (t_{trans}) in packet switching routers can be effectively modeled as exponential
- Arrival pattern of packets at a router is often Poisson in nature
  - and, arrival interval is exponential
- Central Limit Theorem: the distribution of a very large number of independent RVs is approximately normal, independent of individual distributions

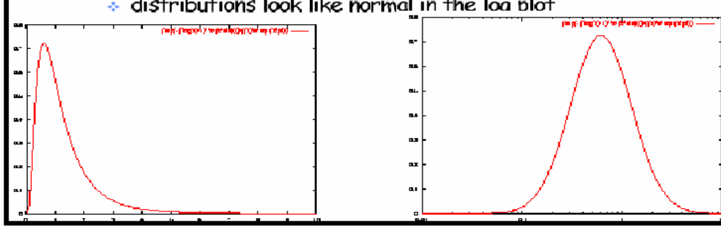
Probability Functions

Log Normal dists

- A random variable X is lognormally distributed if log(X) is normally distributed:

$$f(x) = \frac{e^{-(\log(x-\mu)/m)^2/(2\sigma^2)}}{(x-\mu)\sigma\sqrt{2\pi}}$$

- where μ is the shape parameter, σ is the location parameter and m is the scale parameter.
  - distributions look like normal in the log plot



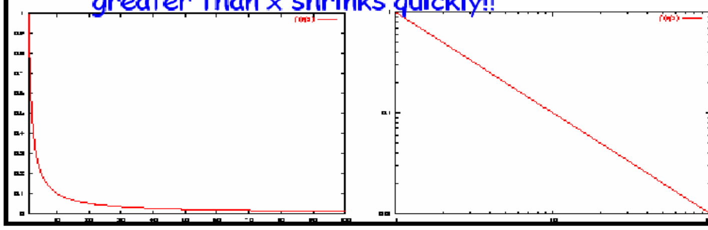
Probability Functions

Power Laws

- Experiments that have trials whose distributions have the general form:

$$P[X > x] \approx c\, x^{-a} \quad x \to \infty,\ a, c > 0$$

- As x increases, the probability that X is greater than x shrinks quickly!!



Probability Functions

Power Laws

- Pareto: the simplest power law distribution

$$f(x) = \frac{ab^a}{x^{a+1}}, \quad x \geq b, \quad a > 0, \quad b > 0$$

⇒ cumulative distribution function (cdf)

$$P[X \leq x] = 1 - \frac{b^a}{x^a}$$

⇒ complementary cumulative distribution function (ccdf)

$$P[X > x] = \frac{b^a}{x^a}$$

Probability Functions

Properties of the Pareto

- a determines the mean and variance
- b determines minimum value x can take

$$E[x] = \begin{cases} \infty & a \le 1 \\ \dfrac{ab}{(1-a)} & a > 1 \end{cases}$$

$$\text{var}[x] = \begin{cases} \infty & a \le 2 \\ \dfrac{ab^2}{(a-1)^2(a-2)} & a > 2 \end{cases}$$

Note: $1<a<2 \Rightarrow$ finite mean, *infinite variance*!

Probability Functions

Pareto

- Pareto studied the distribution of incomes.
  - Few people have a lot of money, many are poor.
- Pareto distributions are concerned with how many events are larger than x: this number is an inverse power of x.

$$P[X > x] = \frac{b^a}{x^a}$$

Probability Functions

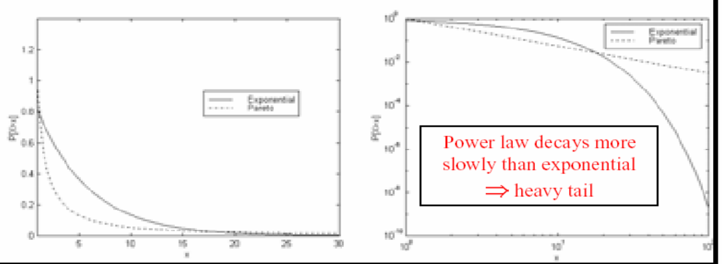
"Signature" of a Power Law

- Consider a log-log plot of $P[X>x]$ vs. x

$$P[X > x] = \frac{b^a}{x^a}$$

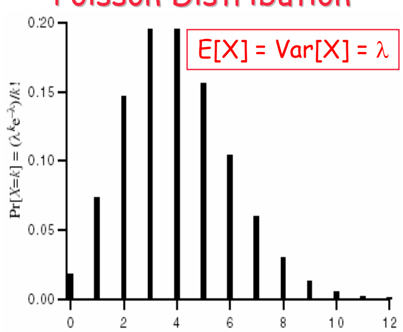
$$\log(P[X > x]) = \log\left(\frac{b^a}{x^a}\right) = -a\log(x) + a\log(b)$$

- Straight line with slope $-a$



Probability Distributions

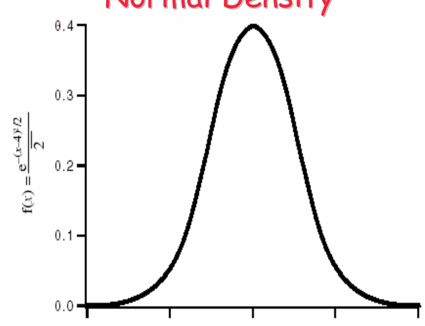
Poisson Distribution



(c) Poisson distribution (λ = 1)

$$Pr[X=k] = \frac{\lambda^k}{k!} e^{-\lambda}$$

Normal Density



(d) Normal probability density (μ = 4, σ = 1)

$$f(x) = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$$