Modelling and Simulation
Introduction

Contents

1 Logistics

- Instructor: Ahmed H. Zahran
- Lecture: Saturdays 1:30 PM
- Grading
  - 20% Midterm
  - 20 % Project
  - 60% Final
- References

  - Dimitri P. Bertsekas and John N. Tsitsiklis, “Introduction to Probability”
  - Kishor Trivedi, “Probability and Statistics with Reliability, Queuing, and Computer Systems Applications”
  - M. F. Neuts “Matrix Geometric Solutions in Stochastic Models”
  - B. R. Havercort, “Performance of Computer Communication Systems, a model based approach”
  - Selected papers

- Course Content
  - Introduction
  - Probability basics revisited
  - System Simulation
  - Discrete Markov Chain
  - Continuous time Markov Chain
  - Markov Reward Models
  - Phase-Type distribution
  - Probabilistic Complex analysis (time permitting)
  - Renewal Process (time permitting)

2 System Performance Evaluation

2.1 Definitions

System: a collection of components which are organized and interact in order to fulfill a common task

Performance: The degree to which a system or component accomplishes its designated functions within given constraints, such as speed, accuracy, or memory usage.

2.2 Performance Evaluation Approaches

- Measurements
  - System unavailable during design phase
- Discrete Event Simulation (DES)
  - Time consuming task for accurate results.
- Analysis
  - Oversimplified assumptions
    * Is the model logically correct, complete, or overly detailed?

* Are the distributional assumptions justified?
* How sensitive are the results to simplifications in the distributional assumptions?
* Are other stochastic properties, such as independence assumptions, valid?
* Is the model represented on the appropriate level?

3 Probability Basics revived :)

3.1 Probability Models

- Sample Space: is the set of all possible outcomes of an experiment or random trial. It is often denoted S, Ω, or U (for "universe")
- Event: subset of the sample space
- Algebra of Events. Revise any probability book
- Probability Axioms

$$(A1): \text{ For any event } A,\ 1 > P(A) \geq 0$$
$$(A2): \ P(S) = 1$$
$$(A3): \ P(A \cup B) = P(A) + P(B) \ (\text{if } A \cap B = \emptyset)$$

Conditional Probability P(A|B): prob. that A occurs, given that 'B' has occurred

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$

A and B are said to be ***mutually independent***, iff,

$$P(A|B) = P(A)$$

Hence

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A) \\ &= P(A)P(B) \end{aligned}$$

- Independence Vs. Mutually Exclusive
  - the probability of the union of mutually exclusive events is the sum of their probabilities
  - the probability of the intersection of two mutually independent events is the product of their probabilities

3.2 Discrete Random Variables

- A **random variable (RV)** X is a mapping (function) from the sample space S to the set of real numbers

$$X : S \to \Re : \text{ that is, } X(s) \in \Re \quad \forall s \in S.$$

- If image(X) finite or countable infinite, X is a ***Discrete*** RV
- A_x: set of all sample points such that,

$$\{s|X(s) = x\}$$

$$P(A_x) = P(X = x) = P(s|X(s) = x) = \sum_{X(s)=x} P(s)$$

- **Probability Mass Function (PMF)**

$$p_X(x) = P(X = x) = \sum_{X(s)=x} P(s)$$

  - PMF Properties

$$0 \leq p_X(x) \leq 1 \ \ \forall x \in R$$
$$\sum_{x \in R} p_X(x) = 1$$

- Cumulative Distribution Function (CDF)

$$\begin{aligned} F_X(t) &= P(-\infty < X \leq t) = P(X \leq t) \\ &= \sum_{x \leq t} p_X(x) \end{aligned}$$

- CDF Properties

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

$$0 \leq F(x) \leq 1 \text{ for } -\infty < x < \infty$$

$F(x)$ is monotone in x and $x_1 \leq x_2$, then
$P(-\infty < X \leq x_1) \leq P(-\infty < X \leq x_2)$

$$\lim_{x \to -\infty} F(x) = 0 \text{ and } \lim_{x \to \infty} F(x) = 1$$

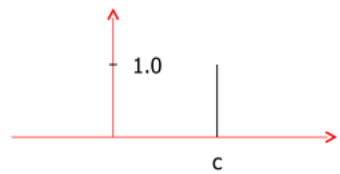
If $x_{i+1} > x_i$, then, $F(x_{i+1}) = F(x_i) + p_X(x_{i+1})$

3.3 Common Discrete RV

3.3.1 Constant Random Variable

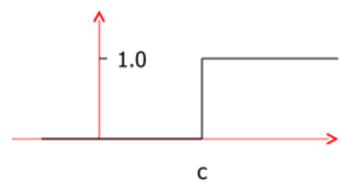
- PMF

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



- CDF

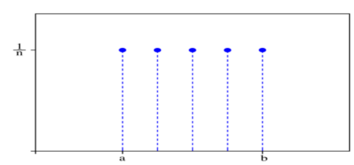
$$F_X(t) = \begin{cases} 0 & \text{if } x < c \\ 1 & \text{otherwise} \end{cases}$$



3.3.2 Discrete Uniform RV

- Discrete RV X that assumes n discrete value with equal probability 1/n
- Discrete uniform pmf

$$p_X(x_i) = \begin{cases} \frac{1}{n}, & \text{if } x_i \in \text{image of } X \\ 0 & \text{otherwise} \end{cases}$$

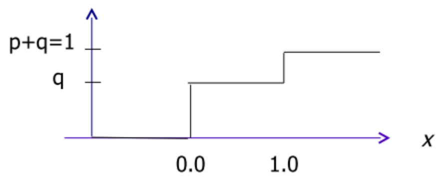


- Discrete uniform distribution function
  X can take integer values 1, 2, ..., n, then

$$F(x) = \sum_{i=1} p_X(i) = \frac{\lfloor x \rfloor}{n}$$

3.3.3 Bernoulli RV

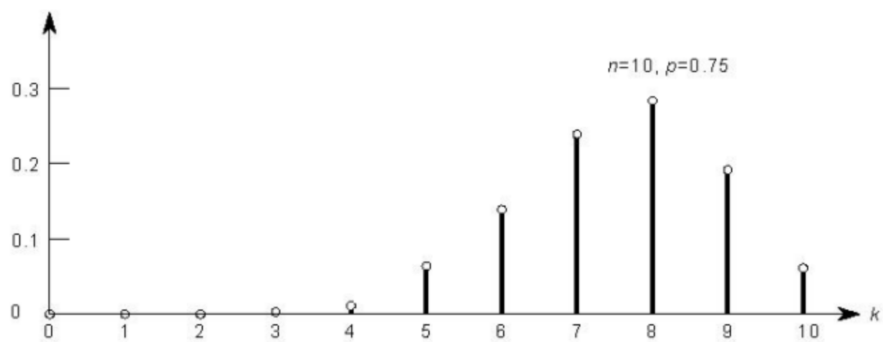
- generated by experiments that has a binary valued outcome, e.g. {0,1}, {Success, Failure}
- The experiment is named Bernoulli trial **(BT)** and the generated RV is called Bernoulli RV
- PMF
  - P(X=1)=p
  - P(X=0)=(1-p)
- CDF
  - $F(x) = \begin{cases} 0 & x < 0 \\ q & 0 \le x < 1 \\ 1 & x > 1 \end{cases}$



3.3.4 Binomial Random Variable

- Generated from n Bernoulli trials.
- RV Y_n: no. of successes in n BTs
- PMF b(k;n,p)

$$\begin{aligned} p_k &= P(Y_n = k) = p_{Y_n}(k) \\ &= \begin{cases} \binom{n}{k} p^k (1-p)^{n-k}, & 0 \le k \le n \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$



- CDF

$$B(t;n,p) = F_{Y_n}(t) = \sum_{i=0}^{\lfloor t \rfloor} \binom{n}{i} p^i (1-p)^{n-i}$$

Example: Packet Transmission

- The bit is transmitted successfully with prob p
- The probability of no error transmission for a packet containing N bits is

$$P_{errorFree} = p^n$$

- The probability of recovering packets with at most m erroneous bits using forward error correction (FEC) is

$$P_{recovery} = \sum_{i=0}^{m} \left(\begin{array}{c} n \\ m \end{array} \right) p^{n-m} (1-p)^m$$

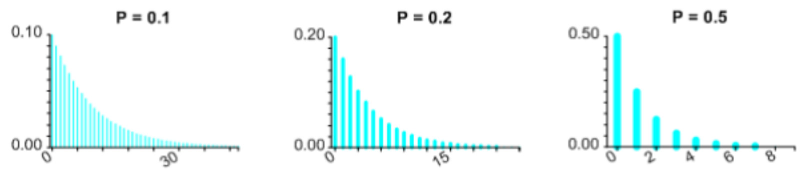
3.3.5 Geometric RV

- Number of independent trials upto and including the 1st success.

Define RV $Z(\in S)$: sample point $0^{i-1}1 = i$

$$p_Z(i) = q^{i-1}p = p(1-p)^{i-1}, \; i = 1, 2, 3, \dots$$

$$F_Z(t) = \sum_{i=1}^{\lfloor t \rfloor} p(1-p)^{i-1} = 1 - (1-p)^{\lfloor t \rfloor}, \; t \geq 0$$



- Geometric RV is the **only** discrete distribution that exhibits MEMORYLESS property.
- given that the first success has not yet occurred, the conditional probability distribution of the number of additional trials does not depend on how many failures have been observed.

$$P(Y = i | Z > n) = \frac{P(Z = n+i, Z > n)}{P(Z > n)}$$
$$= pq^{i-1} = p_Z(i)$$

- **Example:** number or loops until exiting
- Number of failures until the first success $\rightarrow$ modified geometric distribution

3.3.6 Poisson RV

- RV such as "no. of arrivals in an interval (0,t)"

$$f(k, \lambda t) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}, \; k = 0, 1, 2, \dots$$

- The Poisson distribution can be derived as a limiting case to the binomial distribution as the number of trials goes to infinity and the expected number of successes remains fixed.
- Poisson RV can be used as an approximation of the binomial distribution if n is sufficiently large and p is sufficiently small.
- There is a rule of thumb stating that the Poisson distribution is a good approximation of the binomial distribution if n is at least 20 and p is smaller than or equal to 0.05, and an excellent approximation if n $\geq$ 100 and np $\leq$ 10
- In a small interval, Δt, prob. of new arrival= $\lambda\Delta$t.
- pmf b(k;n, λt/n) [Binomial]

$$b(k; n, \frac{\lambda t}{n}) = \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k}, k = 0, 1, \ldots, n$$

- CDF B(k;n, λt/n)= $\sum_k b(k; n; \frac{\lambda t}{n})$

$$\lim_{n \to \infty} \frac{n!}{(n-k)!k!} \left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k} \quad lim_{h \to 0}(1+h)^h = e$$

Poisson Process

- Poisson process is a stochastic process which counts the number of events and the time that these events occur in a given time interval.
- The number of arrivals in the given time interval is a Poisson RV.
- The time between each pair of consecutive events has an exponential distribution with parameter λ. **(prove)**
- Sum of two Poisson RV is a new Poisson RV.
- Poisson process is used as a typical arrival process (call arrival, Requests for individual documents on a web server, packet arrival, ...)

3.4 Probability Generating Function (PGF)

- Letting, P(X=k)=pk , PGF of X is defined by

$$G_X(z) = \sum_{k=0}^{\infty} p_k z^k = p_0 + p_1 z^1 + p_2 z^2 + \ldots + p_k z^k + \ldots$$

- Can be used to estimate the moments of any RV.
- **Theorem 1:** If two RVs have the same PGF, then they have the same distribution (One-to-one mapping: pmf (or CDF) ⇔PGF)
- **Theorem 2:** The PGF of a sum of RVs is the product of their individual PGFs

Example: proving the sum of two independent Poisson RV is another Poisson RV

Proof:

the PGF od Poisson Random Variable can be expressed as

$G_X(z) = \sum_{k=0}^{\infty} p_k z^k = \sum_{k=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^k}{k!} z^k = e^{-\lambda t}[e^{\lambda t z}] = e^{\lambda t(z-1)}$

Since the PGF of the sum of two random variables (X and Y) is the product of their individual PGFs

then the PGF of the sum (M) can be expressed as

$G_M(z) = G_X * G_Y = e^{\lambda t(z-1)} * e^{\eta t(z-1)} = e^{(\lambda+\eta)t(z-1)}$

Note that $G_M(z)$ has the same form of Poisson RV PGF. Then M is a Poisson RV with parameter $(\lambda + \eta)$

3.5 Maximum of two RVs

Let Z=max{X,Y}

$$\begin{aligned} P_Z(z) &= P(max[X,Y] < z) \\ &= P(X < z, Y < z) \\ &= \underbrace{P(X < z)P(Y < z)}_{by\, independence} \end{aligned}$$

- The CDF of the maximum is the product of the CDF of involved ***independent*** Rvs
- Determine the distribution of the minimum of two RVs.

3.6 Homework

- Prove that geometric distribution is memoryless.
- Derive the PGF of discussed Rvs
- Show that Sum of two Binomial RVs is another Binomial distribution.
- Show that the min of two geometric RV is another geometric RV.