Probability Review

Georgios Varsamopoulos

School of Computing and Informatics

Arizona State University
Basic Probability concepts
Basic concepts and terminology

• Basic concepts by example: dice rolling
  – (Intuitively) coin tossing and dice rolling are Stochastic Processes
  – The outcome of a dice rolling is an Event
  – The values of the outcomes can be expressed as a Random Variable, which assumes a new value with each event.
  – The likelihood of an event (or instance) of a random variable is expressed as a real number in [0,1].
    • Dice: Random variable
    • instances or values: 1,2,3,4,5,6
    • Events: Dice=1, Dice=2, Dice=3, Dice=4, Dice=5, Dice=6
    • Pr{Dice=1}=0.1666
    • Pr{Dice=2}=0.1666
  – The outcome of the dice is independent of any previous rolls (given that the constitution of the dice is not altered)
Basic Probability Properties

• The sum of probabilities of all possible values of a random variable is exactly 1.
  – \( \Pr\{\text{Coin=HEADS}\} + \Pr\{\text{Coin=TAILS}\} = 1 \)

• The probability of the union of two events of the same variable is the sum of the separate probabilities of the events
  – \( \Pr\{\text{Dice}=1 \cup \text{Dice}=2 \} = \Pr\{\text{Dice}=1\} + \Pr\{\text{Dice}=2\} = \frac{1}{3} \)
Properties of two or more random variables

• The tossing of two or more coins (such that each does not touch any of the other(s) ) simultaneously is called *(statistically) independent processes* (so are the related variables and events).

• The probability of the intersection of two or more independent events is the product of the separate probabilities:
  
  $\Pr\{\text{Coin1}=\text{HEADS} \cap \text{Coin2}=\text{HEADS} \} = \Pr\{\text{Coin1}=\text{HEADS}\} \cdot \Pr\{\text{Coin2}=\text{HEADS}\}$
Conditional Probabilities

• The likelihood of an event can change if the knowledge and occurrence of another event exists.

\[
\Pr\{A \mid B\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}}, \quad \Pr\{B\} > 0
\]

• Notation:

• Usually we use conditional probabilities like this:

\[
\Pr\{A\} = \sum_i \Pr\{A \mid B_i\} \Pr\{B_i\}
\]
Conditional Probability Example

• Problem statement (ball picking):
  – There are 3 identical sacks with colored balls. Sack A has 1 red and 4 green balls, sack B has 2 red and 3 green balls and sack C has 4 red and 1 green balls. We choose 1 sack equally randomly and pull a ball while blind-folded. What is the probability that we chose a red ball?

• Thinking:
  – If we pick sack A, there is 0.2 probability that we get a red ball
  – If we pick sack B, there is 0.4 probability that we get a red ball
  – If we pick sack C, there is 0.8 probability that we get a red ball
  – For each sack, there is 0.333 probability that we choose that sack

• Result ( \( R \) stands for “picking a red ball”):

\[
Pr\{R\} = Pr\{R \mid A\} Pr\{A\} + Pr\{R \mid B\} Pr\{B\} + Pr\{R \mid C\} Pr\{C\} = \\
= 0.2 \cdot 0.333 + 0.4 \cdot 0.333 + 0.8 \cdot 0.333 = 0.4662
\]
Moments and expected values

- **$m$th moment:**
  \[ E[ X^m ] = \sum x_i^m \Pr\{ X = x_i \} \]

- **Expected value** is the 1st moment: \( \mu_X = E[ X ] \)

- **$m$th central moment:**
  \[ E[ (X - \mu_X)^m ] = \sum (x_i - \mu_X)^m \Pr\{ X = x_i \} \]

- 2nd central moment is called **variance**
  \( \text{Var}[X], \sigma^2_X \)
  \[ \text{Var}[ X ] = E[ X^2 ] - E^2[ X ] \]
Discrete Distributions
Geometric Distribution

• Based on random variables that can have 2 outcomes (\(A\) and \(\neg A\)): \(\Pr\{A\}, \Pr\{\neg A\} = 1 - \Pr\{A\}\)

• Expresses the probability of number of trials needed to obtain the event \(A\).
  — Example: what is the probability that we need \(k\) dice rolls in order to obtain a 6?

• Formula: (for the dice example, \(p = \frac{1}{6}\))
  \[
  \Pr\{Y = k\} = p(1 - p)^{k-1}
  \]
Geometric distribution example

- A wireless network protocol uses a stop-and-wait transmission policy. Each packet has a probability $p_E$ of being corrupted or lost.
- What is the probability that the protocol will need 3 transmissions to send a packet successfully?
  - Solution: 3 transmissions is 2 failures and 1 success, therefore:
    $$\Pr\{Y = 3\} = (1 - p_E) p_E^2$$
- What is the average number of transmissions needed per packet?
  $$E[Y] = \sum_{i=1}^{\infty} i \Pr\{Y = i\} = \sum_{i=1}^{\infty} i (1 - p_E) p_E^{i-1} = (1 - p_E) \sum_{i=1}^{\infty} i p_E^{i-1} = (1 - p_E) \frac{1}{(1 - p_E)^2} = \frac{1}{1 - p_E}$$
**Binomial Distribution**

- Expresses the probability of some events occurring out of a larger set of events
  
  - Example: we roll the dice $n$ times. What is the probability that we get $k$ 6’s?

- Formula: (for the dice example, $p=\frac{1}{6}$)

$$
\Pr\{Y = k\} = \frac{n!}{k!(n-k)!} \ p^k \ (1 - p)^{n-k}
$$
Binomial Distribution Example

• Every packet has \( n \) bits. There is a probability \( p_B \) that a bit gets corrupted.

• What is the probability that a packet has exactly 1 corrupted bit?
  \[
  \Pr\{Y = 1\} = \frac{n!}{1!(n-1)!} p_B^1 (1 - p_B)^{n-1} = \frac{n!}{(n-1)!} p_B (1 - p_B)^{n-1}
  \]

• What is the probability that a packet is not corrupted?
  \[
  \Pr\{Y = 0\} = \frac{n!}{0!(n)!} p_B^0 (1 - p_B)^n = (1 - p_B)^n
  \]

• What is the probability that a packet is corrupted?
  \[
  p_E = 1 - \Pr\{Y = 0\} = 1 - (1 - p_B)^n
  \]
Poisson Process
Poisson Distribution

• Poisson Distribution is the limit of Binomial when \( n \to \infty \) and \( p \to 0 \), while the product \( np \) is constant. In such a case, assessing or using the probability of a specific event makes little sense, yet it makes sense to use the “rate” of occurrence (\( \lambda = np \)).
  
  – Example: if the average number of falling stars observed every night is \( \lambda \) then what is the probability that we observe \( k \) stars fall in a night?

• Formula:

\[
\Pr\{Y = k\} = \frac{\lambda^k e^{-\lambda}}{k!}
\]
Poisson Distribution Example

• In a bank branch, a rechargeable Bluetooth device sends a “ping” packet to a computer every time a customer enters the door.
• Customers arrive with Poisson distribution of $\lambda$ customers per day.
• The Bluetooth device has a battery capacity of $m$ Joules. Every packet takes $n$ Joules, therefore the device can send $u=\frac{m}{n}$ packets before it runs out of battery.
• Assuming that the device starts fully charged in the morning, what is the probability that it runs out of energy by the end of the day?

\[
\Pr\{Y \geq u\} = 1 - \Pr\{Y < u\} = 1 - \sum_{k=0}^{u-1} \frac{\lambda^k e^{-\lambda}}{k!}
\]
Poisson Process

• It is an extension to the Poisson distribution in time
• It expresses the number of events in a period of time given the rate of occurrence.
• Formula

\[ P_m(t) = \frac{(\lambda t)^m e^{-\lambda t}}{m!} \]

• A Poisson process with parameter \( \lambda \) has expected value \( \lambda \) and variance \( \lambda \).
• The time intervals between two events of a Poisson process have an exponential distribution (with mean \( 1/\lambda \)).
Poisson dist. example problem

- An interrupt service unit takes $t_o$ sec to service an interrupt before it can accept a new one. Interrupt arrivals follow a Poisson process with an average of $\lambda$ interrupts/sec.

- What is the probability that an interrupt is lost?
  - An interrupt is lost if 2 or more interrupts arrive within a period of $t_o$ sec.

$$\Pr\{Y \geq 2\} = 1 - \Pr\{Y < 2\} = 1 - P_0(t_o) - P_1(t_o) =$$

$$1 - \frac{\lambda t_o^0 e^{-\lambda t_o}}{0!} - \frac{(\lambda t_o)^1 e^{-\lambda t_o}}{1!} = 1 - \left(1 + \lambda t_o\right) e^{-\lambda t_o}$$
Problem 1

• We roll a regular dice and we toss a coin as many times as the roll indicates.
  – What is the probability that we get no tails?
  – What is the probability that we get 3 heads?

• Using the same process, we are asked to get at least 6 tails in total.
  – If we don’t get 6 tails with the first roll, we roll again and repeat as many times as needed to get at least 6 tails.
  – What is the probability that we need more than 1 rolls to get 6 tails?
Problem 2

• A process has to send 1000 bytes over a wireless link using a stop-and-wait protocol. The payload size per packet is 10 bytes. The bit-wise error probability is $10^{-3}$.
  
  – Find the expected number of total packet transmissions needed to transfer the 1000 bytes.
Queuing Systems
Little’s Law: 
Average number of tasks in system = 
Average arrival rate \times system time per task 

\[ L = \lambda W \]
What is the system performance? E.g.

- average number of tasks in system?
- average number of tasks in queue?
- average turnaround time in whole system?
- average waiting time in the queue?
Different Queuing Models

• M/M/1
  – Arrival: Poisson process
  – Processing: Poisson process
  – Number of servers: 1

• M/M/m
  – Arrival: Poisson process
  – Processing: Poisson process
  – Number of servers: $m$

• M/G/1
  – Arrival: Poisson process
  – Processing: generalized (arbitrary) process
  – Number of servers: 1

• G/M/1, G/M/m, G/G/1... etc.
Results for M/M/1 Queue

- Probability that the system is busy:
  \[ \rho = \frac{\lambda}{\mu} \]

- Average number of tasks in the system:
  \[ L = \frac{\rho}{1 - \rho} \]
Problem 1

- If $\lambda$ and $\mu$ are the arrival and service rate, respectively, in a M/M/1 queue:
  - What is the average system time?
  - What is the average queue length?
  - What is the average wait (queuing) time?
  - What is the probability that the system is busy?
  - What is the probability that the system is empty at the arrival of a task?