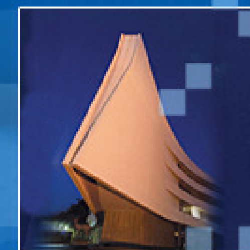
ERP 420

Research Project

Operational Laws, Probability, and stochastic processes



Dr A. De Freitas

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

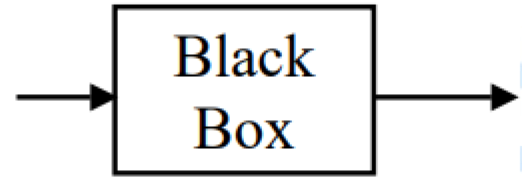
References

- “The Art of Computer Systems Performance Analysis” by Raj Jain
  - http://www.eletrica.ufpr.br/pedroso/2014/TE816/Art_Of_Computer_Systems_Performance_Analysis_Techniques_For_Experimental_Measurements_Simulation_And_Modeling-Raj_Jain.pdf
- “Computer Systems Analysis Using Queuing Network Models”, Edward Lazowska, John Zahorjan, Scott Graham, and Kenneth Sevcik
  - http://homes.cs.washington.edu/~lazowska/qsp/

Operational Laws

- Relationships that do not require any assumptions about the distribution of service times or inter-arrival times.
- Identified originally by Buzen (1976) and later extended by Denning and Buzen (1978).
- **Operational** $\Rightarrow$ Directly measured.
- **Operationally testable assumptions**
  $\Rightarrow$ assumptions that can be verified by measurements.
  - For example, whether number of arrivals is equal to the number of completions?
  - This assumption, called job flow balance, is operationally testable.
  - A set of observed service times is or is not a sequence of independent random variables is not is not operationally testable.

Operational Quantities



- Quantities that can be directly measured during a finite observation period.
- T = Observation interval $\quad A_i$ = number of arrivals
- C_i = number of completions $\quad B_i$ = busy time B_i

$$\text{Arrival Rate } \lambda_i = \frac{\text{Number of arrivals}}{\text{Time}} = \frac{A_i}{T}$$

$$\text{Throughput } X_i = \frac{\text{Number of completions}}{\text{Time}} = \frac{C_i}{T}$$

$$\text{Utilization } U_i = \frac{\text{Busy Time}}{\text{Total Time}} = \frac{B_i}{T}$$

$$\text{Mean service time } S_i = \frac{\text{Total time Served}}{\text{Number served}} = \frac{B_i}{C_i}$$

Utilization Law

$$
\begin{aligned}
\text{Utilization } U_i &= \frac{\text{Busy Time}}{\text{Total Time}} = \frac{B_i}{T} \\
&= \frac{C_i}{T} \times \frac{B_i}{C_i} = \frac{\text{Completions}}{\text{Time}} \times \frac{\text{Busy Time}}{\text{Completions}} \\
&= \text{Throughput} \times \text{Mean Service Time} = X_i S_i
\end{aligned}
$$

- This is one of the operational laws
- Operational laws are similar to the elementary laws of motion For example,

$$d = \frac{1}{2} a t^2$$

- Notice that distance *d*, acceleration *a*, and time *t* are **operational quantities**. No need to consider them as expected values of random variables or to assume a distribution.

Example

- Bandwidth of a communication link is 56,000 bps. This link is used to transmit 1500-byte packets that flow through the link at a rate of 3 packets/sec. What is the utilization of the link?

Example

- Bandwidth of a communication link is 56,000 bps. This link is used to transmit 1500-byte packets that flow through the link at a rate of 3 packets/sec. What is the utilization of the link?
- S0:= time to transmit a packet on the link(transmission delay)

$$S_0 = \frac{1500 \times 8bits}{56000bps} = 0.214\sec$$

$$U_0 = X_0 S_0 = 3 \times 0.214 = 0.642$$

Example

- A disk is serving 50 requests/sec; each request requires 0.005 seconds of service.
- 1) What is the Utilization?
- 2) Maximum possible service rate?

Example

- A disk is serving 50 requests/sec; each request requires 0.005 seconds of service.
- 1) What is the Utilization?
- 2) Maximum possible service rate?
- 1) U0 = X0 × S0 = 50/s × 0.005 s = 0.25 = 25%
- 2) $U_0 = 1$

$$X_0^{\max} = \frac{U_0}{S_0} = \frac{1}{0.005 \quad \sec} = 200 req / \sec$$

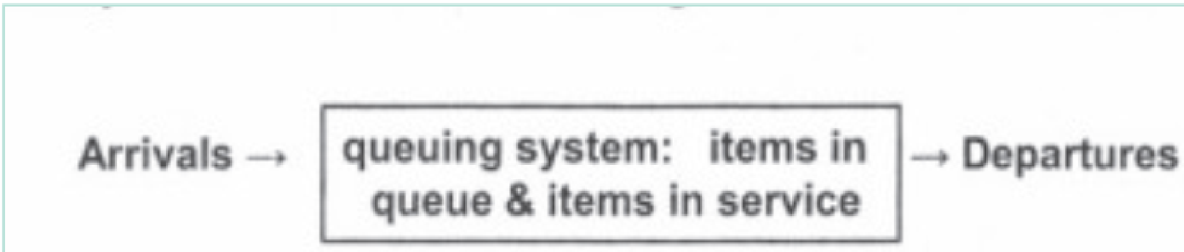
Example

- A router forwards 100 packets/second onto a link. The transmission time (i.e., time to put packets onto the link), on average, is 1 ms. What is the link utilization?
- What is link capacity?

Example

- A router forwards 100 packets/second onto a link. The transmission time (i.e., time to put packets onto the link), on average, is 1 ms. What is the link utilization?
- Link throughput: $X0 = 100$ packets/sec
- Service time/transmission time: $S0 = 0.001$ sec
- Utilization: $U0 = X0 \times S0 = 0.1 = 10\%$
- Link capacity? Set $Uo = 1$ and solve for $X0 \Rightarrow U=XS \Rightarrow 1=0.001X \Rightarrow X=1000$ packets per second

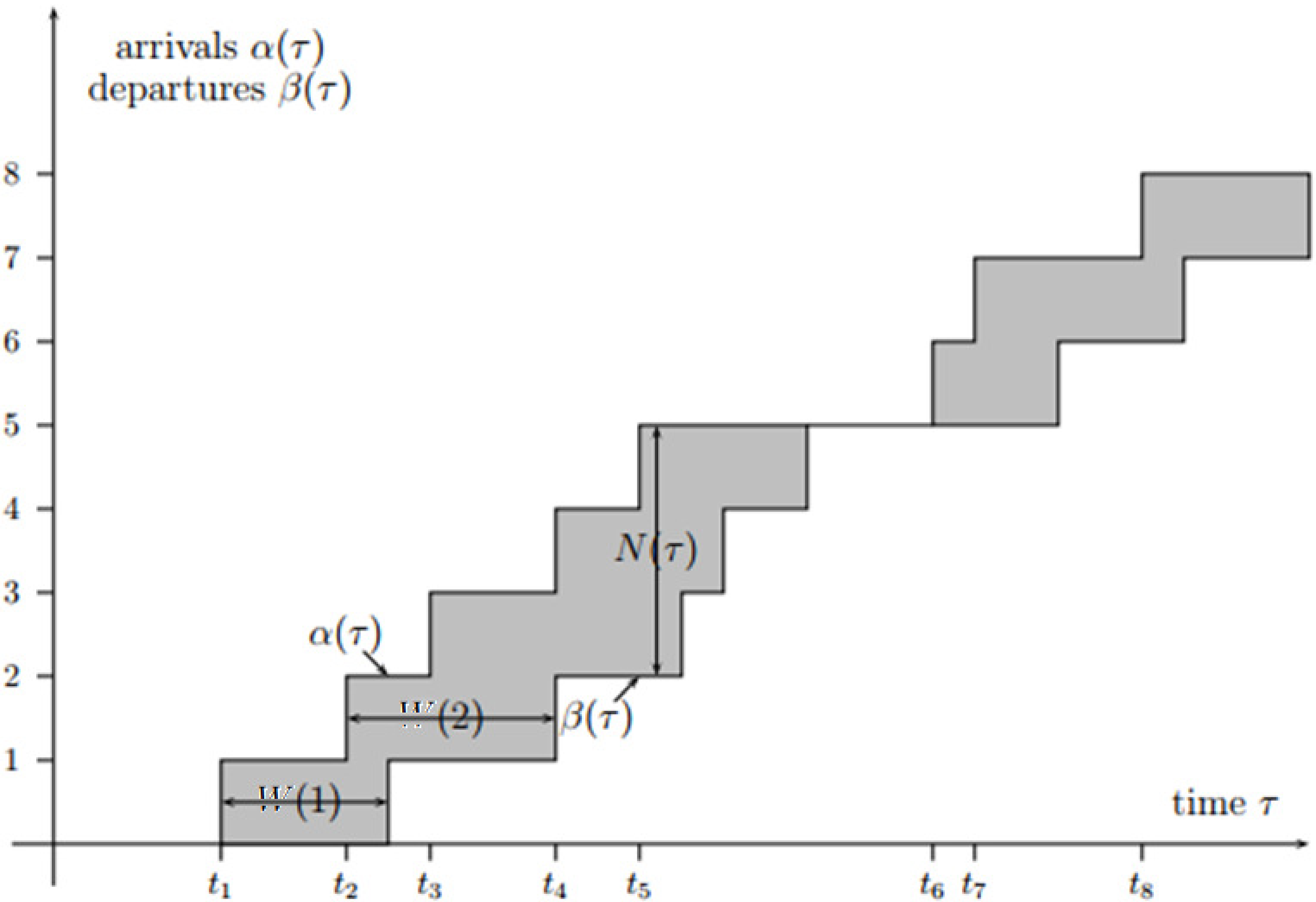
Little’s Law



- L =average number of items in the queuing system,
- W = average waiting time in the system for an item, and
- λ =average number of items arriving per unit time or arrival rate or throughput (total number of completions per time= throughput).
- **Mean number of packets in system = arrival rate x mean response time**

$$L = \lambda W = XW$$

Little's Law Proof



UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Example

- A restaurant processes on average 1500 customers per day (=15 hours). On average, there are 50 customers waiting to place an order, waiting for an order to arrive or eating. A) What is average time in the restaurant.
- B) Out of the 50 customers, 40 customers on the average are eating. What is average wait at the counter?

Example

- A restaurant processes on average 1500 customers per day (=15 hours). On average, there are 50 customers waiting to place an order, waiting for an order to arrive or eating. What is average time in the restaurant.
- λ = 1500 customers/day = 100 customers/hour; L = 50 customers;
- (Average time in the restaurant)W = L/λ = 50/100 = 1/2 hours,

Example

- B) Out of the 50 customers, 40 customers on the average are eating. What is average wait at the counter?
- $\lambda = 100$, $L = 50 - 40 = 10$ customers at the service counter;
- $W = L/\lambda = 10/100$ hours = 0.1 hours= 6 minutes average wait at the counter

References

- J. Virtamo 38.3143 Queueing Theory / Poisson process, https://www.netlab.tkk.fi/opetus/s383143/

- Banks, Carson, Nelson & Nicol, Discrete-Event System Simulation (5th Edition), Pearson, 2009

Random Variables

We are often more interested in a some number associated with the experiment rather than the outcome itself.

Example 1. The number of heads in tossing coin rather than the sequence of heads/tails

A real-valued random variable X is a mapping
$X: \quad \mathcal{S} \mapsto \mathcal{R}$
which associates the real number $X(e)$ to each outcome $e \in \mathcal{S}$.

Example 2. The number of heads in three consecutive tossings of a coin (head = **h**, tail=**t** (tail))

e	$X(e)$
hhh	3
hht	2
hth	2
htt	1
thh	2
tht	1
tth	1
ttt	0

- The values of X are "drawn" by "drawing" e
- e represents a "lottery ticket", on which the value of X is written

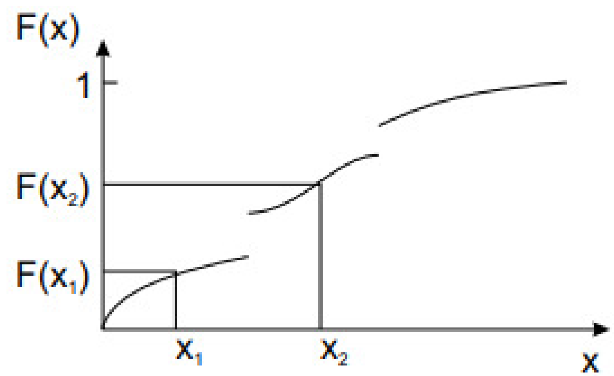
The image of a random variable X

$$\boxed{\mathcal{S}_X = \{x \in \mathcal{R} \mid X(e) = x,\ e \in \mathcal{S}\}}$$ (complete set of values X can take)

- may be finite or countably infinite: discrete random variable
- uncountably infinite: continuous random variable

Distribution function (cdf, cumulative distribution function)

$$\boxed{F(x) = \mathrm{P}\{X \leq x\}}$$

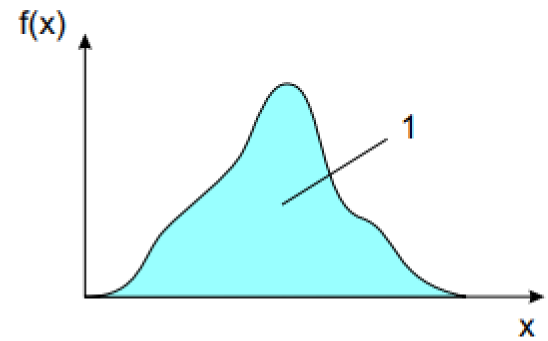


The probability of an interval

$$\boxed{\mathrm{P}\{x_1 < X \leq x_2\} = F(x_2) - F(x_1)}$$

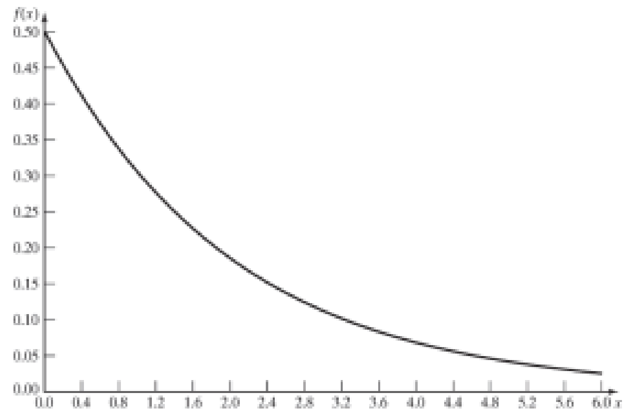
Continuous random variable: probability density function (pdf)

$$\boxed{f(x) = \frac{dF(x)}{dx} = \lim_{dx \to 0} \frac{\mathrm{P}\{x < X \leq x + dx\}}{dx}}$$



- Example: Life of an inspection device is given by X, a continuous random variable with pdf:

$$f(x) = \begin{cases} \frac{1}{2} e^{-x/2}, & \mathrm{x} \geq 0 \\ 0, & \text{otherwise} \end{cases}$$



  - X has an exponential distribution with mean 2 years
  - Probability that the device's life is between 2 and 3 years is:

$$P(2 \leq x \leq 3) = \frac{1}{2} \int_2^3 e^{-x/2} dx = 0.14$$

Discrete random variable

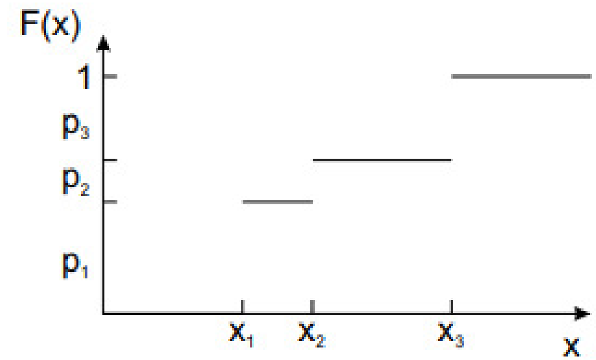
The set of values a discrete random variable X can take is either finite or countably infinite, $X \in \{x_1, x_2, \ldots\}$.

With these are associated the point probabilities

$$p_i = \mathrm{P}\{X = x_i\}$$

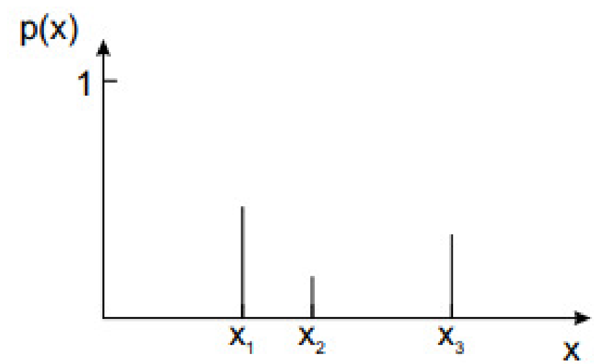
which define the discrete distribution

The distribution function is a step function, which has jumps of height p_i at points x_i.



Probability mass function (pmf)

$$p(x) = \mathrm{P}\{X = x\} = \begin{cases} p_i & \text{when } x = x_i \\ 0, & \text{otherwise} \end{cases}$$



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Parameters of distributions

Expectation

Denoted by $\mathrm{E}[X] = \bar{X}$

Continuous distribution: $\boxed{\mathrm{E}[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx}$

Discrete distribution: $\boxed{\mathrm{E}[X] = \sum_i x_i p_i}$

Variance

Denoted by $\mathrm{V}[X]$ (also $\mathrm{Var}[X]$)

$$\boxed{\mathrm{V}[X] = \mathrm{E}[(X - \bar{X})^2]} = \mathrm{E}[X^2] - \mathrm{E}[X]^2$$

Poisson Distribution

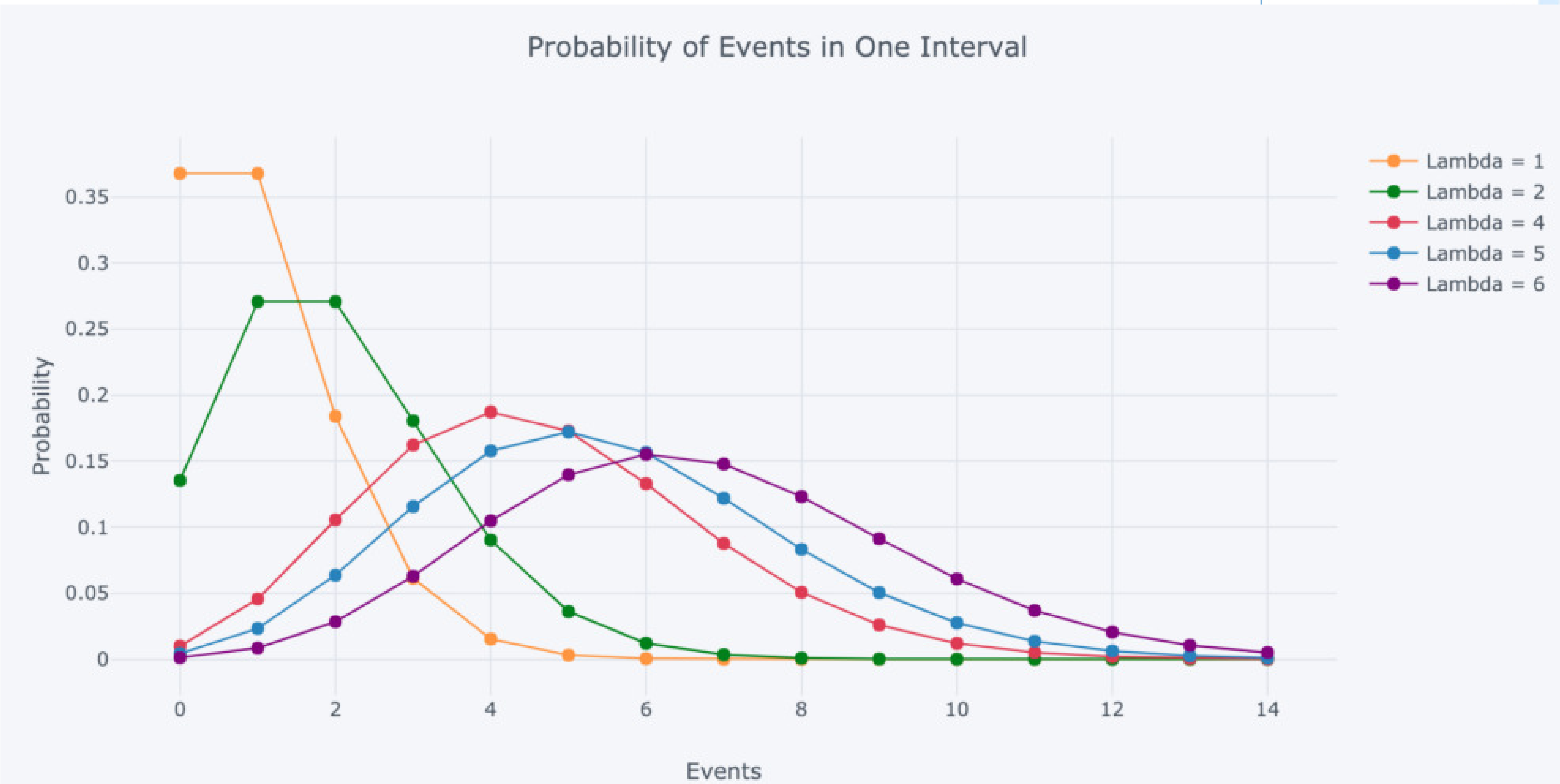
- Poisson distribution describes many random processes quite well and is mathematically quite simple.
  - where $\alpha > 0$, pdf (probability density function) and cdf (cumulative distribution function) are:

$$p(x) = \begin{cases} \dfrac{e^{-\alpha}\alpha^{x}}{x!}, & x = 0,1,\ldots \\ 0, & \text{otherwise} \end{cases}$$

$$F(x) = \sum_{i=0}^{x} \frac{e^{-\alpha}\alpha^{i}}{i!}$$

  - *Expected value& variance:*
  - $E(X) = \alpha = V(X)$

Poisson Distribution



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Example

- Example: A computer repair person is "beeped" each time there is a call for service. The number of beeps per hour ~ Poisson(α = 2 per hour).
  - A) The probability of three beeps in the next hour?
  - B) The probability of two or more beeps in a 1-hour period?

Example

- Example: A computer repair person is called each time there is a call for service. The number of calls per hour ~ Poisson(α = 2 per hour).
  - A) The probability of three calls in the next hour:

    $p(3) = e^{-2}2^3/3! = 0.18$

    or, $p(3) = F(3) - F(2) = 0.857-0.677=0.18$
  - B) The probability of two or more calls in a 1-hour period:

    $$\begin{aligned} p(2 \text{ or more}) &= 1 - p(0) - p(1) \\ &= 1 - F(1) \\ &= 0.594 \end{aligned}$$

Stochastic Process

Let ξ denote the random outcome of an experiment. To every such outcome suppose a waveform $X(t,\xi)$ is assigned.

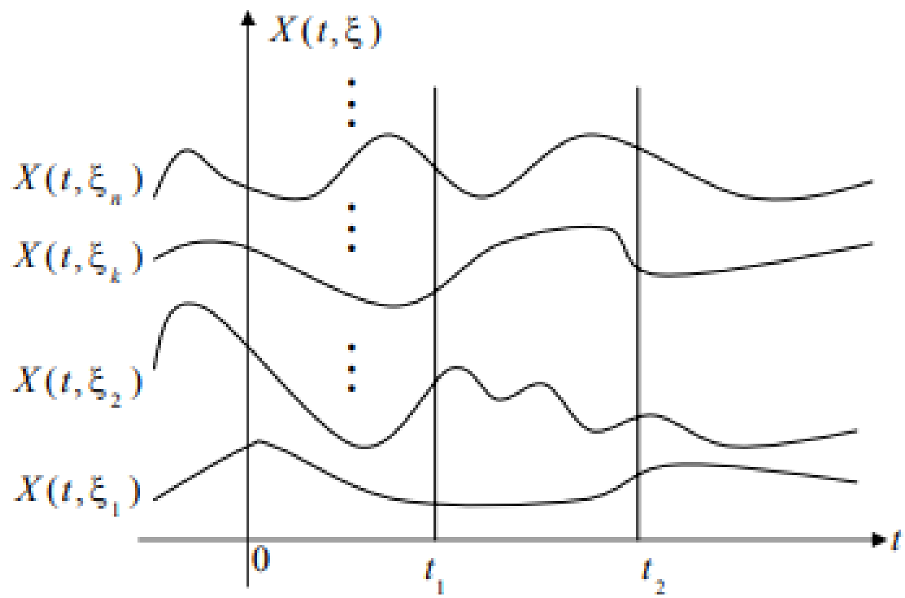


Fig. 14.1

The collection of such waveforms form a stochastic process. The set of $\{\xi_k\}$ and the time index t can be continuous or discrete (countably infinite or finite) as well.

For fixed $\xi_i \in S$ (the set of all experimental outcomes), $X(t,\xi)$ is a specific time function.

For fixed t,

$$X_1 = X(t_1, \xi_i)$$

is a random variable. The ensemble of all such realizations $X(t,\xi)$ over time represents the stochastic process $X(t)$.

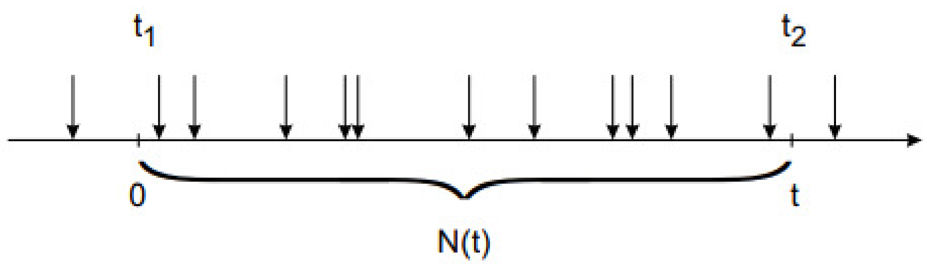
Poisson process

General

Poisson process is one of the most important models used in queueing theory.

- Often the arrival process of customers can be described by a Poisson process.
- In teletraffic theory the "customers" may be calls or packets. Poisson process is a viable model when the calls or packets originate from a large population of independent users.

In the following it is instructive to think that the Poisson process we consider represents discrete arrivals (of e.g. calls or packets).



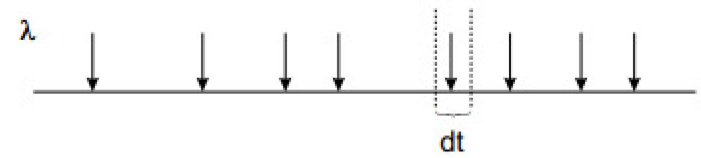
Mathematically the process is described by the so called counter process N_t or $N(t)$. The counter tells the number of arrivals that have occurred in the interval $(0, t)$ or, more generally, in the interval (t_1, t_2).

$$\begin{cases} N(t) = \text{number of arrivals in the interval } (0,t) & \text{(the stochastic process we consider)} \\ N(t_1,t_2) = \text{number of arrival in the interval } (t_1,t_2) & \text{(the increment process } N(t_2)-N(t_1)) \end{cases}$$

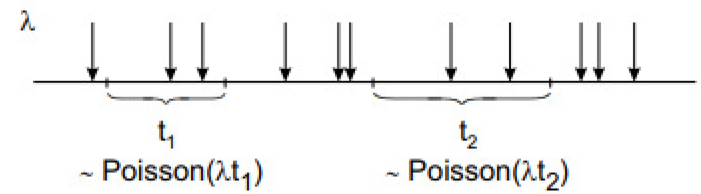
Definition

The Poisson process can be defined in three different (but equivalent) ways:

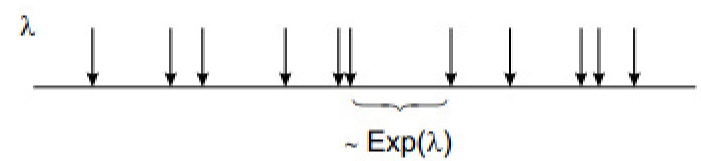
1. Poisson process is a pure birth process:
   In an infinitesimal time interval dt there may occur only one arrival. This happens with the probability λdt independent of arrivals outside the interval.



2. The number of arrivals $N(t)$ in a finite interval of length t obeys the Poisson(λt) distribution,
   $$\mathrm{P}\{N(t) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$
   Moreover, the number of arrivals $N(t_1, t_2)$ and $N(t_3, t_4)$ in non-overlapping intervals $(t_1 \leq t_2 \leq t_3 \leq t_4)$ are independent.



3. The interarrival times are independent and obey the Exp(λ) distribution:
   $$\mathrm{P}\{\text{interarrival time} > t\} = e^{-\lambda t}$$



Poisson Process

- There is also a formula for the mean (expected value), and variance (δ^2)

$$\mu = \lambda t$$

$$\delta^2 = \lambda t$$

- Equal mean and variance: $E[N(t)] = V[N(t)] = \lambda t$