$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Bayes' rule:

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

$$P(B_j|A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_i P(A|B_i)P(B_i)}$$

- Random process is stationary if the joint distribution of any set of samples does not depend on the placement of the time origin

$$F_{X(t_1),\ldots,X(t_k)}(x_1,\ldots,x_k) = F_{X_{(t_1+\tau)},\ldots,X_{(t_k+\tau)}}(x_1,\ldots,x_k)$$

- Wide sense stationary if

$$m_X(t) = m \;\; \forall t \quad \text{and} \quad C_x(t_1,t_2) = C_X(t_1 - t_2)$$

Bernoulli Trials

- Random experiments has two possible outcomes, success and failure and respective probability p and q with $p + q = 1$
- Sequence of n independent experiment - sequence of Bernoulli trials
- S_n sample space of an experiment, $S_n - \{2^n$ tuples of 0's and 1's$\}$
- the probability of obtaining exactly k successes in n trials

$$p(k) = \binom{n}{k} p^k q^{n-k} \quad k = 0, 1, \ldots, n$$

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$$

Poisson distribution

- arrival of jobs to a computing center for the time interval $[0, t]$
- interval divided into n subintervals of very short duration $\delta = t/n$
- the probability of more than one event occurrence in a subinterval is negligible compared to the probability of observing one or zero events
- an event, occurs or not in a subinterval is independent
- probability of an event occurrence in each subinterval is p then the expected number of event occurrences in the interval $[0, t]$ is np
- events occur at a rate of λ events per second, the average number of events in the interval $[0, t]$ is λt, so $\lambda t = np$
- n interval constituting a sequence of Bernoulli trials with the probability of success $p = \lambda t/n$

- the probability of k arrivals in a total of n intervals

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

  as $n \longrightarrow \infty$ the binomial distribution approaches a Poisson distribution with parameter λt

$$p(N(t)=k) = \frac{(\lambda t)^k}{k!}e^{-\lambda t}$$

- Application modeled as Poisson distributed
  – congestion, queuing, number of jobs arriving, the number of jobs completing service, the number of messages transmitted through a communication channel in a fixed interval of time

Exponential Random Variable

- T time between event occurrences in a Poisson process
- The probability that the intervent time T exceeds t seconds is equivalent to no event occuring in t seconds

$$\begin{aligned} P[T > t] &= P[\text{no events in } t \text{ seconds}] = (1-p)^n \\ &= (1 - \frac{\lambda t}{n})^n \rightarrow e^{-\lambda t} \;\text{ as }\; n \rightarrow \infty \end{aligned}$$

- T is an exponential random variable with parameter λ
- the intervent times in a Poisson process from an *iid sequence of exponential random variables with mean* $1/\lambda$
- *sum of* n iid exponential random variables has an Erlang distribution; if T_j denote the iid exponential interarrival times

$$S_n = T_1 + T_2 + \cdots + T_n$$

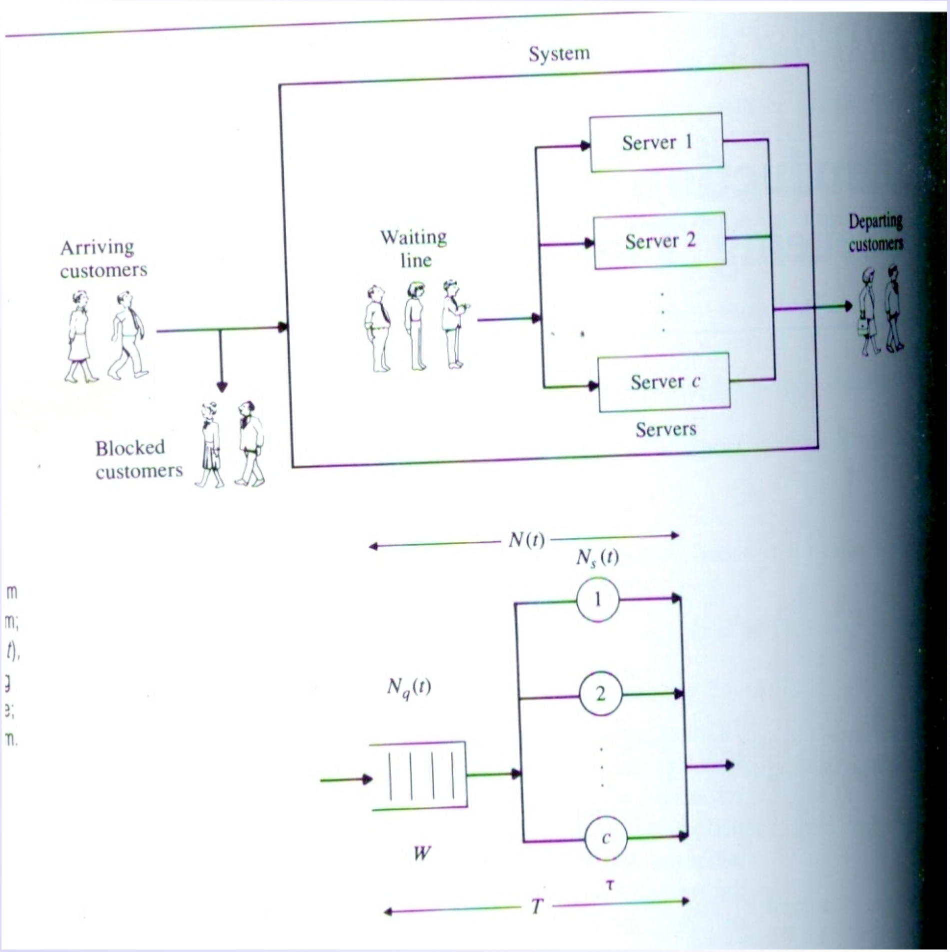
- Exponential random variable satisfies the memoryless property

$$P[X > t + h | X > t] = P[X > h]$$

- left side is probability of having to wait at least h additional seconds given that one has already been waiting t seconds
- right side is the probability of waiting at least h seconds when one first begins to wait, i.e. the probability of waiting at least an additional h seconds is the same regardless of how long one has already been waiting
- Application of Exponential distribution
  - Time between two successive job arrivals to a computing server
  - Service time at a server in a queuing network
  - Time to failure (lifetime) of a component
  - Time required to repair a component that has malfunctioned

Queueing Theory

- Users - Resources - time period of resource use - queue
- Random nature of demand behavior of customers implies that probabilistic measures such as average delay, average throughput etc. are required to assess the performance of the system
- arrival time of the ith customer S_i and arrival rate is λ
- ith customer seeking a service will require τ_i seconds of service time
- a limited number of waiting spaces and if no room customers are turned away called ”blocked” at rate λ_b
- Queue or service discipline specifies the order in which customers are selected from the queue and allowed into service
- Waiting time W_i; time elapses from the arrival time of the ith customer until the time when it enters service; total delay $T_i = W_i + \tau_i$


System
Server 1
Server 2
Server c
Servers
Arriving customers
Waiting line
Departing customers
Blocked customers
N(t)
$N_s(t)$
1
2
c
$N_q(t)$
W
τ
T

- performance of the system is given by the statistics of the waiting time W and T
- the proportion of customers that are blocked λ_b/λ
- the proportion of time that each server is utilized and the rate at which customers are services by the system, $\lambda_d = \lambda - \lambda_b$
- These are function of the number of customers in the system at time t and the number of customers in queue at time t
- $a/b/m/K$ is used to describe queueing system
  - a type of arrival process
  - b service time distribution
  - m number of servers
  - K maximum number of customers allowed in the system at any time

- if a is given by M, then the arrival process is Poisson and the interarrival times are independent, identically distributed exponential random variables
- if b is given by M, then the service times are iid exponential random variables
- if b is given by D, then the service times are constant (deterministic)
- if b is given by G, then the service times are iid according to some general distribution