PARTNERS Summer School
Mathematics & Statistics
Sessions 8, 9: Probability and Statistics: Queues

Malcolm Farrow

July 2010

Stochastic processes and queueing systems

- Queues: e.g.:

Stochastic processes and queueing systems

- Queues: e.g.:
  - Post office

Stochastic processes and queueing systems

- Queues: e.g.:
  - Post office
  - Supermarket checkout

Stochastic processes and queueing systems

- Queues: e.g.:
  - Post office
  - Supermarket checkout
  - Road junction

Stochastic processes and queueing systems

- Queues: e.g.:
  - Post office
  - Supermarket checkout
  - Road junction
  - Jobs in a factory

Stochastic processes and queueing systems

- Queues: e.g.:
  - Post office
  - Supermarket checkout
  - Road junction
  - Jobs in a factory
  - Hospital waiting list

Stochastic processes and queueing systems

- Queues: e.g.:
  - Post office
  - Supermarket checkout
  - Road junction
  - Jobs in a factory
  - Hospital waiting list
  - Internet traffic

Stochastic processes and queueing systems

- Queues: e.g.:
  - Post office
  - Supermarket checkout
  - Road junction
  - Jobs in a factory
  - Hospital waiting list
  - Internet traffic
- We usual model queues as *stochastic processes* because the "random" behaviour is important.

Stochastic processes and queueing systems

- A *stochastic process* is a system whose development, in time or space, is governed, at least partially, by probabilistic laws.

Stochastic processes and queueing systems

- A *stochastic process* is a system whose development, in time or space, is governed, at least partially, by probabilistic laws.
- Probability — I assume some basic ideas of probability. If this causes a problem, please let me know.

Stochastic processes and queueing systems

- A *stochastic process* is a system whose development, in time or space, is governed, at least partially, by probabilistic laws.
- Probability — I assume some basic ideas of probability. If this causes a problem, please let me know.
- We use stochastic models for queues because features of the behaviour depend on the “randomness.”

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

- Probabilistic laws govern:

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

- Probabilistic laws govern:
  - The time, T, between successive arrivals which is a random variable.

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

- Probabilistic laws govern:
  - The time, T, between successive arrivals which is a random variable.
  - The service time, S, of a customer. This is another random variable.

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

- Probabilistic laws govern:
  - The time, T, between successive arrivals which is a random variable.
  - The service time, S, of a customer. This is another random variable.
- Deterministic laws:

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

- Probabilistic laws govern:
  - The time, T, between successive arrivals which is a random variable.
  - The service time, S, of a customer. This is another random variable.
- Deterministic laws:
  - The queue discipline: first come, first served.

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

- Probabilistic laws govern:
  - The time, T, between successive arrivals which is a random variable.
  - The service time, S, of a customer. This is another random variable.
- Deterministic laws:
  - The queue discipline: first come, first served.
  - The number of servers which is 1 in this case.

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

- Probabilistic laws govern:
  - The time, T, between successive arrivals which is a random variable.
  - The service time, S, of a customer. This is another random variable.
- Deterministic laws:
  - The queue discipline: first come, first served.
  - The number of servers which is 1 in this case.
- Properties of interest:

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

- Probabilistic laws govern:
  - The time, T, between successive arrivals which is a random variable.
  - The service time, S, of a customer. This is another random variable.
- Deterministic laws:
  - The queue discipline: first come, first served.
  - The number of servers which is 1 in this case.
- Properties of interest:
  - The number, $N(t)$, of customers present in the queue at time t.

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

- Probabilistic laws govern:
  - The time, T, between successive arrivals which is a random variable.
  - The service time, S, of a customer. This is another random variable.
- Deterministic laws:
  - The queue discipline: first come, first served.
  - The number of servers which is 1 in this case.
- Properties of interest:
  - The number, $N(t)$, of customers present in the queue at time t.
  - The waiting time, W_n, of customer number n.

Simple queueing system

Customers arrive at a service point and are served in order of arrival.

- Probabilistic laws govern:
  - The time, T, between successive arrivals which is a random variable.
  - The service time, S, of a customer. This is another random variable.
- Deterministic laws:
  - The queue discipline: first come, first served.
  - The number of servers which is 1 in this case.
- Properties of interest:
  - The number, $N(t)$, of customers present in the queue at time t.
  - The waiting time, W_n, of customer number n.
  - The proportion of time, p, when the server is idle.

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals
    1.1 What is the probability distribution for the inter-arrival times?

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals
   1.1 What is the probability distribution for the inter-arrival times?
   1.2 Are the inter-arrival times independent of the state of the system?

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals
   1.1 What is the probability distribution for the inter-arrival times?
   1.2 Are the inter-arrival times independent of the state of the system?
2. Service mechanism

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals
    1.1 What is the probability distribution for the inter-arrival times?
    1.2 Are the inter-arrival times independent of the state of the system?
2. Service mechanism
    2.1 How many servers are there?

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals
   1.1 What is the probability distribution for the inter-arrival times?
   1.2 Are the inter-arrival times independent of the state of the system?
2. Service mechanism
   2.1 How many servers are there?
   2.2 What is the pattern of service times? What is the probability distribution of service times? Are the service times independent of the state of the system and of each other?

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. <u>Pattern of arrivals</u>
   1.1 What is the probability distribution for the inter-arrival times?
   1.2 Are the inter-arrival times independent of the state of the system?
2. <u>Service mechanism</u>
   2.1 How many servers are there?
   2.2 What is the pattern of service times? What is the probability distribution of service times? Are the service times independent of the state of the system and of each other?
   2.3 Service availability. Could there be, for example, random interruptions to the queue service?

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals
   1.1 What is the probability distribution for the inter-arrival times?
   1.2 Are the inter-arrival times independent of the state of the system?
2. Service mechanism
   2.1 How many servers are there?
   2.2 What is the pattern of service times? What is the probability distribution of service times? Are the service times independent of the state of the system and of each other?
   2.3 Service availability. Could there be, for example, random interruptions to the queue service?
3. Queue discipline

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals
   1.1 What is the probability distribution for the inter-arrival times?
   1.2 Are the inter-arrival times independent of the state of the system?
2. Service mechanism
   2.1 How many servers are there?
   2.2 What is the pattern of service times? What is the probability distribution of service times? Are the service times independent of the state of the system and of each other?
   2.3 Service availability. Could there be, for example, random interruptions to the queue service?
3. Queue discipline
   3.1 First come first served.

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals
   1.1 What is the probability distribution for the inter-arrival times?
   1.2 Are the inter-arrival times independent of the state of the system?
2. Service mechanism
   2.1 How many servers are there?
   2.2 What is the pattern of service times? What is the probability distribution of service times? Are the service times independent of the state of the system and of each other?
   2.3 Service availability. Could there be, for example, random interruptions to the queue service?
3. Queue discipline
   3.1 First come first served.
   3.2 Last come first served.

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals
   1.1 What is the probability distribution for the inter-arrival times?
   1.2 Are the inter-arrival times independent of the state of the system?
2. Service mechanism
   2.1 How many servers are there?
   2.2 What is the pattern of service times? What is the probability distribution of service times? Are the service times independent of the state of the system and of each other?
   2.3 Service availability. Could there be, for example, random interruptions to the queue service?
3. Queue discipline
   3.1 First come first served.
   3.2 Last come first served.
   3.3 Priority systems.

Queueing Systems

A queueing system is a system in which stochastic fluctuations can cause congestion (and therefore the development of queues). In describing a queueing system there are three essential features:

1. Pattern of arrivals
   1.1 What is the probability distribution for the inter-arrival times?
   1.2 Are the inter-arrival times independent of the state of the system?
2. Service mechanism
   2.1 How many servers are there?
   2.2 What is the pattern of service times? What is the probability distribution of service times? Are the service times independent of the state of the system and of each other?
   2.3 Service availability. Could there be, for example, random interruptions to the queue service?
3. Queue discipline
   3.1 First come first served.
   3.2 Last come first served.
   3.3 Priority systems.
   3.4 Parallel queues, for example supermarket checkout.

Arrival mechanism: Poisson process

A commonly used model for arrivals is called a *Poisson process*. This is a very simple but useful model. The basic idea is that arrivals occur "completely at random in time". Examples where we might use a Poisson process model include:

- "Clicks" on a Geiger counter.
- Vehicles passing on a (quiet) road.
- Arrivals of telephone calls at an exchange.
- Accidents.

We are often interested in $N(t)$, the number of arrivals occurring in a time interval of length t. Let $p_n(t) = \Pr\{N(t) = n\}$.

Arrival mechanism: Poisson process

We make three assumptions:

Arrival mechanism: Poisson process

We make three assumptions:

1. The process is *stationary.* That is the distribution of the number of arrivals in the interval $(t_0, t_0 + \delta t)$ is the same as that for the interval $(t_0 + \tau, t_0 + \tau + \delta t)$.

Arrival mechanism: Poisson process

We make three assumptions:

1. The process is *stationary*. That is the distribution of the number of arrivals in the interval $(t_0, t_0 + \delta t)$ is the same as that for the interval $(t_0 + \tau, t_0 + \tau + \delta t)$.
2. There is a constant rate λ, representing the mean number of arrivals per unit time, such that for a short time interval of length δt

$$\begin{aligned} p_0(\delta t) &\approx 1 - \lambda \delta t \\ p_1(\delta t) &\approx \lambda \delta t \\ p_n(\delta t) &\approx 0 \end{aligned}$$

for $n > 1$, where the approximation is to the order of δt.

Arrival mechanism: Poisson process

We make three assumptions:

1. The process is *stationary*. That is the distribution of the number of arrivals in the interval $(t_0, t_0 + \delta t)$ is the same as that for the interval $(t_0 + \tau, t_0 + \tau + \delta t)$.
2. There is a constant rate λ, representing the mean number of arrivals per unit time, such that for a short time interval of length δt

$$\begin{aligned} p_0(\delta t) &\approx 1 - \lambda \delta t \\ p_1(\delta t) &\approx \lambda \delta t \\ p_n(\delta t) &\approx 0 \end{aligned}$$

   for $n > 1$, where the approximation is to the order of δt.
3. The numbers of arrivals in non-overlapping time intervals are independent.

Arrival mechanism: Poisson process

As an approximation, imagine dividing time up into sections, each of length δt. Suppose that the probability that an arrival occurs in a particular section is $\lambda \delta t$. In the time interval $(0, t)$ there are $t/\delta t$ sections. The number of sections with events in them has a *binomial distribution* and the probability of j events in $(0, t)$ is

$$
\begin{aligned}
\tilde{p}_j &\approx \binom{t/\delta t}{j} (\lambda \delta t)^j (1 - \lambda \delta t)^{t/\delta t - j} \\
&\approx \binom{t/\delta t}{j} \left(\frac{\lambda \delta t}{1 - \lambda \delta t} \right)^j (1 - \lambda \delta t)^{t/\delta t} \\
&\approx \frac{\lambda^j}{j!} t(t - \delta t)(t - 2\delta t) \cdots (t - \{j-1\}\delta t)(1 - \lambda \delta t)^{t/\delta t - j}
\end{aligned}
$$

Arrival mechanism: Poisson process

It is easily shown that

$$\tilde{p}_j \to \frac{\exp(-\lambda t)(\lambda t)^j}{j!}$$

as $\delta t \to 0$.
Thus $N(t)$, the number of arrivals in the time interval $(0, t)$, has a *Poisson distribution* with mean λt.

Arrival mechanism: Poisson process

We can prove formally that the number of arrivals in a time interval of length t has a Poisson distribution with mean λt. If this number is $N(t)$ then

$$\Pr\{N(t) = j\} = \frac{e^{-\lambda t}(\lambda t)^j}{j!}$$

for $j = 0, 1, 2, \ldots$.

Arrival mechanism: Time till next arrival

- Consider the waiting time, T, till the next arrival.

Arrival mechanism: Time till next arrival

- Consider the waiting time, T, till the next arrival.
- Suppose we start waiting at time 0. Here 0 may be a fixed point, the time of a previous arrival or whatever. It does not matter.

Arrival mechanism: Time till next arrival

- Consider the waiting time, T, till the next arrival.
- Suppose we start waiting at time 0. Here 0 may be a fixed point, the time of a previous arrival or whatever. It does not matter.
- The probability that the arrival has not happened by time t is

$$\Pr\{N(t) = 0\} = \exp(-\lambda t).$$

This is the probability that the waiting time is at least t.

Arrival mechanism: Time till next arrival

- Consider the waiting time, T, till the next arrival.
- Suppose we start waiting at time 0. Here 0 may be a fixed point, the time of a previous arrival or whatever. It does not matter.
- The probability that the arrival has not happened by time t is

$$\Pr\{N(t) = 0\} = \exp(-\lambda t).$$

  This is the probability that the waiting time is at least t.
- Thus the *distribution function* of T, that is the probability that $T < t$, is

$$F_T(t) = 1 - \exp(-\lambda t).$$

Arrival mechanism: Time till next arrival

- Consider the waiting time, T, till the next arrival.
- Suppose we start waiting at time 0. Here 0 may be a fixed point, the time of a previous arrival or whatever. It does not matter.
- The probability that the arrival has not happened by time t is

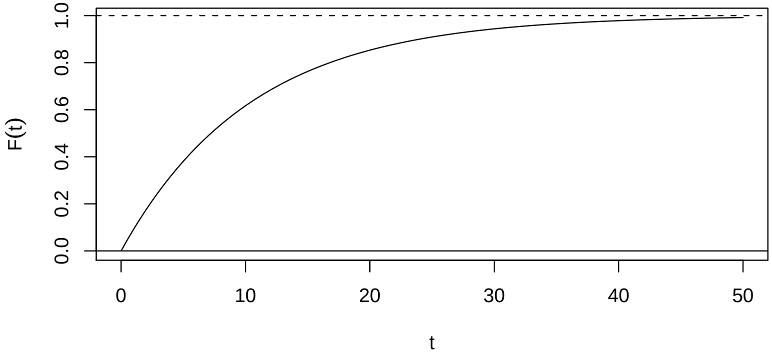
$$\Pr\{N(t) = 0\} = \exp(-\lambda t).$$

  This is the probability that the waiting time is at least t.
- Thus the *distribution function* of T, that is the probability that $T < t$, is

$$F_T(t) = 1 - \exp(-\lambda t).$$

- This is the distribution function of an *exponential distribution* with parameter λ.

Distribution function of inter-arrival time



Probability density function

If a continuous random variable T has distribution function $F_T(t)$ then

$$f_T(t) = \frac{d}{dt} F_T(t)$$

is called the *probability density function* (or *pdf*) of T.
Note that

$$\int_{-\infty}^{\infty} f_T(t)\ dt = 1$$

— the total probability.
For our exponential distribution

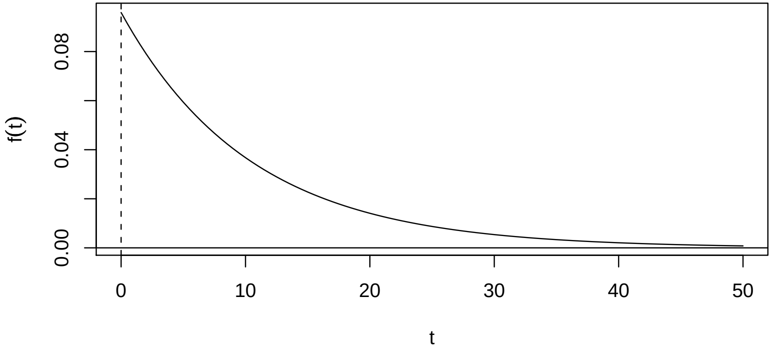
$$F_T(t) = 1 - \exp(-\lambda t)$$

so

$$f_T(t) = \lambda \exp(-\lambda t)$$

$(0 < t < \infty)$.

Probability density function of inter-arrival time



Example: Road vehicle headways

The times of arrivals of motor vehicles passing a point in Chester Road, Sunderland going East.

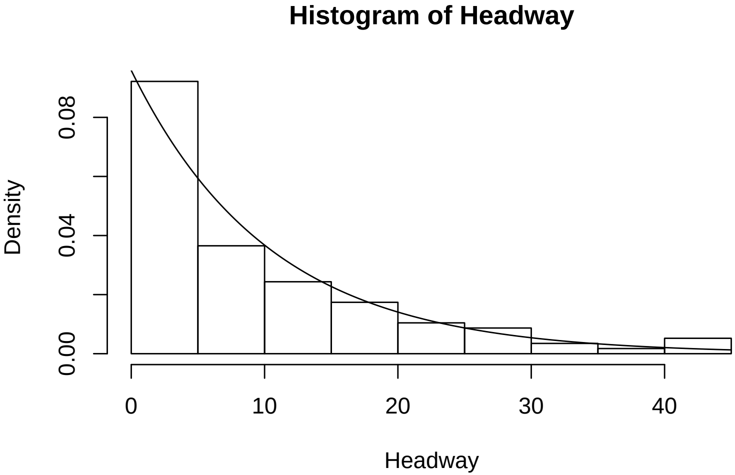
```
12.40   3, 6, 9,15,24,28,30
12.41   7,12,14,16,21,24,30,50
12.42   9,22,28,46,53
12.43  22,25,35,38,58
12.44   2, 5, 8,10,14,17,27,30,45
12.45   3,46
12.46  13,42,51
12.47   0, 9,11,18,23,26,39,51

 .............................

12.59   2,31,34,38,54,59
```

Example: Road vehicle headways

These can be converted to time intervals (in seconds)between vehicles. (The first value is the time till the first arrival, which has the same distribution under our model).

3	3	3	6	9	4	2	37	5	2	2
5	3	6	20	19	13	6	18	7	29	3
10	3	20	4	3	3	2	4	3	10	3
15	18	43	27	29	9	9	9	2	7	5
3	13	12	44	4	16	13	2	9	1	13
12	1	2	16	26	4	13	13	4	14	3
31	5	2	5	5	20	1	1	7	21	3
4	25	22	14	2	4	11	15	8	41	18
3	10	7	3	5	21	6	5	22	2	2
5	7	3	24	3	7	32	3	19	2	1
6	29	3	4	16	5					

Example: Road vehicle headways



Time to second arrival

- The second arrival has not occurred before time t if $N(t) < 2$.

Time to second arrival

- The second arrival has not occurred before time t if $N(t) < 2$.
- Now

$$\Pr\{N(t) < 2\} = \Pr\{N(t) = 0\} + \Pr\{N(t) = 1\} = e^{-\lambda t} + \lambda t e^{-\lambda t}$$

Time to second arrival

- The second arrival has not occurred before time t if $N(t) < 2$.
- Now

$$\Pr\{N(t) < 2\} = \Pr\{N(t) = 0\} + \Pr\{N(t) = 1\} = e^{-\lambda t} + \lambda t e^{-\lambda t}$$

- Hence the distribution function is

$$F_2(t) = 1 - (1 + \lambda t)e^{-\lambda t}$$

Time to second arrival

- The second arrival has not occurred before time t if $N(t) < 2$.
- Now

$$\Pr\{N(t) < 2\} = \Pr\{N(t) = 0\} + \Pr\{N(t) = 1\} = e^{-\lambda t} + \lambda t e^{-\lambda t}$$

- Hence the distribution function is

$$F_2(t) = 1 - (1 + \lambda t)e^{-\lambda t}$$

- Hence the probability density function is

$$\begin{aligned} f_2(t) &= \frac{d}{dt} F_2(t) \\ &= \lambda(1 + \lambda t)e^{-\lambda t} - \lambda e^{-\lambda t} \\ &= \lambda^2 t e^{-\lambda t} \end{aligned}$$

Gamma distribution

If a continuous random variable T has probability density function

$$f_g(t) = \begin{cases} 0 & (t < 0) \\ \frac{\lambda^\alpha t^{\alpha-1} e^{-\lambda t}}{\Gamma(\alpha)} & (t \geq 0) \end{cases}$$

where α and λ are positive parameters, then we say that T has a *gamma*(α, λ) *distribution*.
Here $\Gamma(\alpha)$ denotes the *gamma function*. This has the property that, if x is a positive number, then

$$\Gamma(x) = (x-1)\Gamma(x-1).$$

If n is a positive integer, then

$$\Gamma(n) = (n-1)!.$$

Gamma distribution

Gamma pdf:

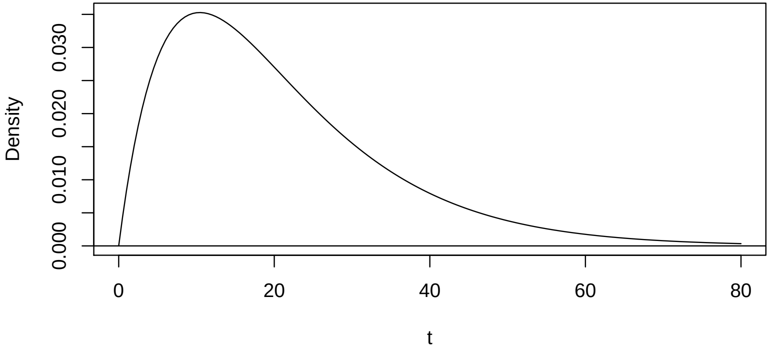
$$f_g(t) = \frac{\lambda^\alpha t^{\alpha-1} e^{-\lambda t}}{\Gamma(\alpha)}$$

The time to the second arrival has pdf

$$f_2(t) = \frac{\lambda^2 t^{2-1} e^{-\lambda t}}{1!} = \frac{\lambda^2 t^{2-1} e^{-\lambda t}}{\Gamma(2)}.$$

So this is a gamma$(2,\ \lambda)$ distribution.
Notice that an exponential(λ) distribution is a gamma$(1,\ \lambda)$ distribution.

Time to second arrival



Mean of a continuous random variable

A continuous random variable X has pdf $f_X(x)$.
Its *mean* or *expectation* (if it exists) is

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x \; f_X(x) \; dx.$$

Similarly, if $g(X)$ is some function of X, then the expectation of $g(X)$ is

$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) \; f_X(x) \; dx.$$

Means: Example 1

$$T \sim \text{exponential}(\lambda)$$

$$\begin{aligned}
\mathrm{E}(T) &= \int_0^\infty t\lambda e^{-\lambda t}\ dt \\
&= [-te^{-\lambda t}]_0^\infty + \int_0^\infty e^{-\lambda t}\ dt \\
&= \left[-\frac{1}{\lambda}e^{-\lambda t}\right]_0^\infty \\
&= \frac{1}{\lambda}
\end{aligned}$$

Means: Example 2

$$T \sim \text{gamma}(\alpha,\ \lambda)$$

$$\begin{aligned}
\mathrm{E}(T) &= \int_0^\infty \frac{t\lambda^\alpha t^{\alpha-1}e^{-\lambda t}}{\Gamma(\alpha)}\,dt \\
&= \int_0^\infty \frac{\lambda^\alpha t^{\alpha+1-1}e^{-\lambda t}}{\Gamma(\alpha)}\,dt \\
&= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}\frac{1}{\lambda}\int_0^\infty \frac{\lambda^{\alpha+1} t^{\alpha+1-1}e^{-\lambda t}}{\Gamma(\alpha+1)}\,dt \\
&= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}\frac{1}{\lambda} \\
&= \frac{\alpha}{\lambda}
\end{aligned}$$

Means: Example 2

$$T \sim \mathrm{gamma}(\alpha,\ \lambda)$$

$$\begin{aligned}
\mathrm{E}(T^2) &= \int_0^\infty \frac{t^2 \lambda^\alpha t^{\alpha-1} e^{-\lambda t}}{\Gamma(\alpha)}\, dt \\
&= \int_0^\infty \frac{\lambda^\alpha t^{\alpha+2-1} e^{-\lambda t}}{\Gamma(\alpha)}\, dt \\
&= \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \frac{1}{\lambda^2} \int_0^\infty \frac{\lambda^{\alpha+2} t^{\alpha+2-1} e^{-\lambda t}}{\Gamma(\alpha+2)}\, dt \\
&= \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \frac{1}{\lambda^2} \\
&= \frac{(\alpha+1)\alpha}{\lambda^2}
\end{aligned}$$

Variance of a random variable

The *variance* of a quantitative random variable X (if it exists) is

$$
\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}\left\{[X - \mathrm{E}(X)]^2\right\} \\
&= \mathrm{E}\left\{X^2 - 2X\mathrm{E}(X) + [\mathrm{E}(X)]^2\right\} \\
&= \mathrm{E}(X^2) - 2\mathrm{E}(X)\mathrm{E}(X) + [\mathrm{E}(X)]^2 \\
&= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2
\end{aligned}
$$

Variance: Example

$$T \sim \mathrm{gamma}(\alpha,\ \lambda)$$

Variance: Example

$$T \sim \mathrm{gamma}(\alpha,\ \lambda)$$

- $$\mathrm{E}(T) = \frac{\alpha}{\lambda}$$

Variance: Example

$$T \sim \mathrm{gamma}(\alpha,\ \lambda)$$

- $$\mathrm{E}(T) = \frac{\alpha}{\lambda}$$
- $$\mathrm{E}(T^2) = \frac{(\alpha + 1)\alpha}{\lambda^2}$$

Variance: Example

$$T \sim \mathrm{gamma}(\alpha,\ \lambda)$$

- $$\mathrm{E}(T) = \frac{\alpha}{\lambda}$$
- $$\mathrm{E}(T^2) = \frac{(\alpha+1)\alpha}{\lambda^2}$$
- So
$$\mathrm{Var}(T) = \frac{(\alpha+1)\alpha}{\lambda^2} - \left\{\frac{\alpha}{\lambda}\right\}^2 = \frac{\alpha}{\lambda^2}$$

Normal distribution

- The *normal distribution*, sometimes called the *Gaussian distribution* is very important in probability and statistics. It is a continuous distribution.

Normal distribution

- The *normal distribution*, sometimes called the *Gaussian distribution* is very important in probability and statistics. It is a continuous distribution.
- It has two parameters, the mean and the variance, often written $\mu,\ \sigma^2$. If X has a normal distribution with mean μ and variance σ^2 we write

$$X \sim N(\mu,\ \sigma^2$$

The range of X is $-\infty < X < \infty$.

Normal distribution

- The *normal distribution*, sometimes called the *Gaussian distribution* is very important in probability and statistics. It is a continuous distribution.
- It has two parameters, the mean and the variance, often written $\mu,\ \sigma^2$. If X has a normal distribution with mean μ and variance σ^2 we write

$$X \sim N(\mu,\ \sigma^2$$

The range of X is $-\infty < X < \infty$.
- The probability density function is

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \qquad (-\infty < x < \infty).$$

Normal distribution

- The *normal distribution*, sometimes called the *Gaussian distribution* is very important in probability and statistics. It is a continuous distribution.
- It has two parameters, the mean and the variance, often written $\mu,\ \sigma^2$. If X has a normal distribution with mean μ and variance σ^2 we write

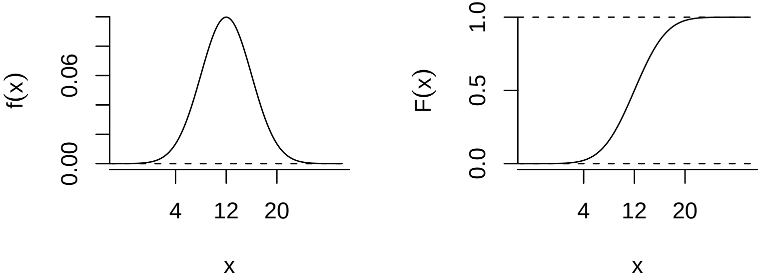
$$X \sim N(\mu,\ \sigma^2$$

  The range of X is $-\infty < X < \infty$.
- The probability density function is

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \qquad (-\infty < x < \infty).$$

- The distribution function can not be written explicitly.

Normal distribution



Probability density function $f(x)$ and distribution function $F(x)$ for a normal distribution with mean $\mu = 12$ and variance $\sigma^2 = 16$.

The standard normal distribution

If $\mu = 0$ and $\sigma^2 = 1$ we have a *standard normal distribution*. If Z has a standard normal distribution we write $Z \sim N(0, 1)$.

- Probability density function:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

- Distribution function:

$$\Phi(z) = \int_{-\infty}^{z} \phi(u)\ du$$

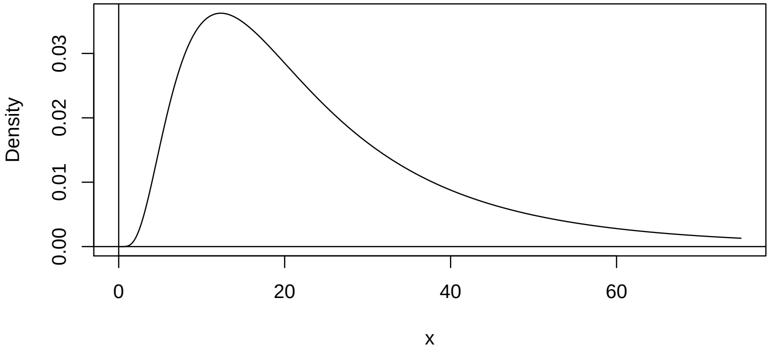
The lognormal distribution

- Sometimes a variable does not have a normal distribution but some *transformation* of the variable does.
- For example, if X must be positive, eg a service time, then perhaps $\ln(X)$ has a normal distribution.

$$0 < X < \infty \;\Leftrightarrow\; -\infty < \ln(X) < \infty$$

- If $Y \sim N(\mu,\ \sigma^2)$ and $X = e^Y$ (so $Y = \ln(X)$) then we say that X has a lognormal$(\mu,\ \sigma^2)$ distribution, $X \sim \mathrm{lognormal}(\mu,\ \sigma^2)$.

The lognormal distribution



Probability density function for a lognormal distribution with $\mu = 3$ and $\sigma^2 = 0.49$ ($\sigma = 0.7$).

The lognormal distribution

If $X \sim \mathrm{lognormal}(\mu,\ \sigma^2)$ what is the mean of X?
If $X \sim \mathrm{lognormal}(\mu,\ \sigma^2)$ and $Y = \ln(X)$ then $\mathrm{E}(X) = \mathrm{E}(e^Y)$.
Hence

$$\begin{aligned}
\mathrm{E}(X) = \mathrm{E}(e^Y) &= \int_{-\infty}^{\infty} e^y (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}(y-\mu)^2\right\}\ dy \\
&= \int_{-\infty}^{\infty} (2\pi\sigma^2)^{-1/2} \exp\left\{y - \frac{1}{2\sigma^2}(y-\mu)^2\right\}\ dy
\end{aligned}$$

The lognormal distribution

$$
\begin{aligned}
y - \frac{1}{2\sigma^2}(y-\mu)^2 &= -\frac{1}{2\sigma^2}[y^2 - 2\mu y - 2\sigma^2 y + \mu^2] \\
&= -\frac{1}{2\sigma^2}[y^2 - 2(\mu+\sigma^2)y + (\mu+\sigma^2)^2 \\
&\quad -(\mu+\sigma^2)^2 + \mu^2] \\
&= -\frac{1}{2\sigma^2}[\{y - (\mu+\sigma^2)\}^2 - (\mu+\sigma^2)^2 + \mu^2] \\
&= -\frac{1}{2\sigma^2}[\{y - (\mu+\sigma^2)\}^2 - 2\mu\sigma^2 - \sigma^4] \\
&= -\frac{1}{2\sigma^2}(y-\theta)^2 + \mu + \frac{\sigma^2}{2}
\end{aligned}
$$

where $\theta = \mu + \sigma^2$.

The lognormal distribution

Hence

$$\begin{aligned}
\mathrm{E}(X) &= \int_{-\infty}^{\infty} (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{1}{2\sigma^2}(y-\theta)^2\right\} \, dy \;\; e^{\mu+\sigma^2/2} \\
&= e^{\mu+\sigma^2/2}
\end{aligned}$$

Estimation: Method of moments

- A simple way to *estimate* the values of parameters of a distribution if we have data is to equate sample moments (mean, variance) with theoretical moments and solve.

Estimation: Method of moments

- A simple way to *estimate* the values of parameters of a distribution if we have data is to equate sample moments (mean, variance) with theoretical moments and solve.
- Suppose we have $n = 50$ independent observations $y_1, \ldots, y_{50}$ from $N(\mu,\ \sigma^2)$:

```
5.82 3.96 3.57 5.00 ...... 5.32 5.95 5.64 6.66 3.33
```

Estimation: Method of moments

- A simple way to *estimate* the values of parameters of a distribution if we have data is to equate sample moments (mean, variance) with theoretical moments and solve.
- Suppose we have $n = 50$ independent observations $y_1, \ldots, y_{50}$ from $N(\mu,\ \sigma^2)$:

  `5.82 3.96 3.57 5.00 ...... 5.32 5.95 5.64 6.66 3.33`

- Sample mean:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i = \frac{268.3}{50} = 5.366.$$

Estimation: Method of moments

- A simple way to *estimate* the values of parameters of a distribution if we have data is to equate sample moments (mean, variance) with theoretical moments and solve.
- Suppose we have $n = 50$ independent observations $y_1, \ldots, y_{50}$ from $N(\mu,\ \sigma^2)$:

  `5.82 3.96 3.57 5.00 ...... 5.32 5.95 5.64 6.66 3.33`

- Sample mean:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{268.3}{50} = 5.366.$$

- We *estimate* μ with $\hat{\mu} = \bar{y} = 5.366$.

Estimation: Method of moments

- Sample variance:

$$
\begin{aligned}
s^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n-1}\left\{\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right\} \\
&= \frac{1}{49}\{1503.931 - 50 \times 5.366^2\} = 1.3109
\end{aligned}
$$

Estimation: Method of moments

- Sample variance:

$$
\begin{aligned}
s^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n-1}\left\{\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right\} \\
&= \frac{1}{49}\{1503.931 - 50 \times 5.366^2\} = 1.3109
\end{aligned}
$$

- We *estimate* σ^2 with $\hat{\sigma}^2 = s^2 = 1.3109$ and estimate the standard deviation σ with $s = \sqrt{s^2} = 1.1449$.

Estimation: Lognormal distribution

In the case of data from a lognormal distribution we can saimply take logs of the data and then proceed as for a normal distribution.
Eg: $X_1, \ldots, X_{50} \sim \mathrm{lognormal}(\mu,\ \sigma^2)$.
Data: $336.97,\ 52.46,\ 35.52, 148.41, \ldots$
Take logs: $5.82,\ 3.96,\ 3.57,\ 5.00, \ldots$ as before.

Simulation

- One way to investigate stochastic system behaviour is to *simulate*.
- Quicker and cheaper than experimenting on the real thing!
- Generate artificial *pseudo-random* inter-arrival times and service times.
- Build up the events and see what happens.
- Possibly repeat many times.
- Actually more useful in more complicated systems.

Simulation: Example

As an example, consider use of a computer network.

- Users log on as a Poisson process. Inter-arrival times are independent exponential(λ) variables.
  Choose $\lambda = 0.1$. So mean inter-arrival time is 10.

Simulation: Example

As an example, consider use of a computer network.

- Users log on as a Poisson process. Inter-arrival times are independent exponential(λ) variables.
  Choose $\lambda = 0.1$. So mean inter-arrival time is 10.
- Service times: Users stay logged on for random lengths of time, eg lognormal($\mu,\ \sigma^2$).
  Choose $\mu = 3.5,\ \sigma^2 = 1.0$ so $\mu + \sigma^2/2 = 4.0$ and the mean service time is $e^4 = 54.6$.

Simulation: Example

As an example, consider use of a computer network.

- Users log on as a Poisson process. Inter-arrival times are independent exponential(λ) variables.
  Choose $\lambda = 0.1$. So mean inter-arrival time is 10.
- Service times: Users stay logged on for random lengths of time, eg lognormal($\mu,\ \sigma^2$).
  Choose $\mu = 3.5,\ \sigma^2 = 1.0$ so $\mu + \sigma^2/2 = 4.0$ and the mean service time is $e^4 = 54.6$.
- We assume that there is unlimited capacity.

Simulation: Example

As an example, consider use of a computer network.

- Users log on as a Poisson process. Inter-arrival times are independent exponential(λ) variables.
  Choose $\lambda = 0.1$. So mean inter-arrival time is 10.
- Service times: Users stay logged on for random lengths of time, eg lognormal($\mu,\ \sigma^2$).
  Choose $\mu = 3.5,\ \sigma^2 = 1.0$ so $\mu + \sigma^2/2 = 4.0$ and the mean service time is $e^4 = 54.6$.
- We assume that there is unlimited capacity.
- Let us start the simulation with an empty system.

Simulation: Example

As an example, consider use of a computer network.

- Users log on as a Poisson process. Inter-arrival times are independent exponential(λ) variables.
  Choose $\lambda = 0.1$. So mean inter-arrival time is 10.
- Service times: Users stay logged on for random lengths of time, eg lognormal($\mu,\ \sigma^2$).
  Choose $\mu = 3.5,\ \sigma^2 = 1.0$ so $\mu + \sigma^2/2 = 4.0$ and the mean service time is $e^4 = 54.6$.
- We assume that there is unlimited capacity.
- Let us start the simulation with an empty system.
- Generate inter-arrival and service times by computer. (In the case of lognormal service times we might generate normal random variables and then exponentiate).

Simulation: Inter-arrival and service times

User	Inter-arrival time	Arrival time	Service time	Service ends
1	15.7		44.2	
2	10.4		24.9	
3	3.9		83.0	
4	12.7		19.2	
5	7.8		41.8	
6	1.0		16.9	
7	28.2		37.7	
8	10.4		23.1	
9	13.4		83.8	
10	0.3		49.6	
11	3.5		11.3	
12	7.7		6.5	

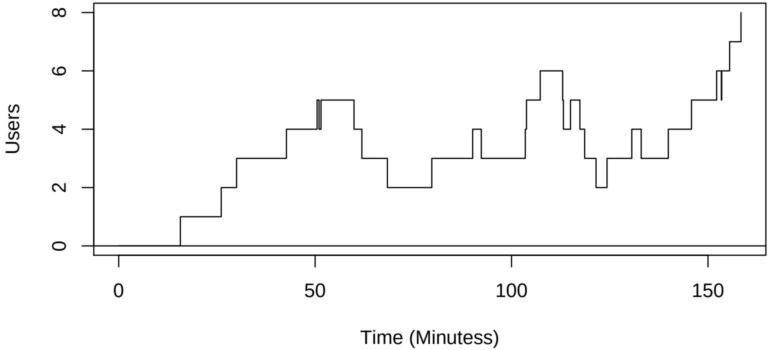
Simulation: Service start and end times

User	Inter-arrival time	Arrival time	Service time	Service ends
1	15.7	15.7	44.2	59.9
2	10.4	26.1	24.9	51.0
3	3.9	30.0	83.0	113.0
4	12.7	42.7	19.2	61.9
5	7.8	50.5	41.8	92.3
6	1.0	51.5	16.9	68.4
7	28.2	79.7	37.7	117.4
8	10.4	90.1	23.1	113.2
9	13.4	103.5	83.8	187.3
10	0.3	103.8	49.6	153.4
11	3.5	107.3	11.3	118.6
12	7.7	115.0	6.5	121.5

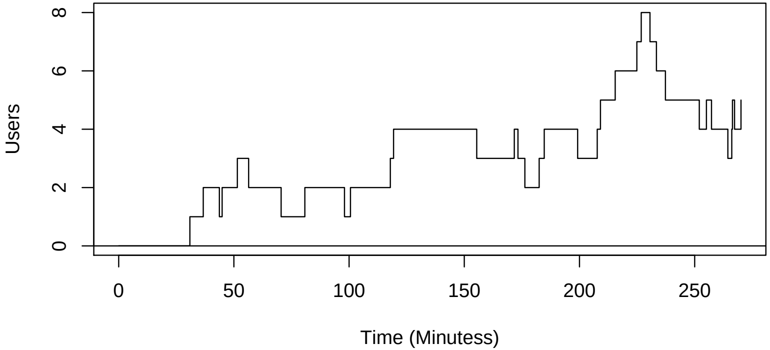
Simulation: Events and user numbers

Event	Time	Change	Number of users
	0.0	0	0
1	15.7	1	1
2	26.1	1	2
3	30.0	1	3
4	42.7	1	4
5	50.5	1	5
6	51.0	-1	4
7	51.5	1	5
8	59.9	-1	4
9	61.9	-1	3
10	68.4	-1	2
11	79.7	1	3

Simulation: Graph



Simulation: Second run



Simulation: Change to $\mu = 5$