NATCOR
STOCHASTIC MODELLING
(preliminary reading)

<u>Introduction to Probability & Random Variables</u>

<u>IMPORTANT</u>

This material is designed to be viewed is two ways:

1. **In ‘Slide Show’ format, to take advantage of the animations;**
2. **In ‘Notes Page’ format, so that you can see the accompanying notes.**

You might find it helps to print out the latter (to refer to) whilst viewing the former.

Overview

Basic Probability

- Concepts
- Laws and Notation
- Conditional Probability
- Total Law of Probability

Discrete Random Variables

- Definition
- Probability Mass Function
- Mean and Variance
- Conditional Expectation
- Poisson Distribution

Continuous Random Variables

- Definition
- Probability Density Function
- Probability as an Integral
- Mean and Variance
- Conditional Expectation
- Exponential Distribution
- Normal Distribution

BASIC PROBABILITY CONCEPTS: 'Experiments' and 'Events'

Experiment: an action whose outcome is uncertain (roll a die)

Sample Space: set of all possible outcomes of an experiment (S = {1, 2, 3, 4, 5, 6})

Event: a subset of outcomes that is of interest to us (e.g. event 1 = even number, event 2 = higher than 4, event 3 = throw a 2, etc)

Probability: measure of how likely an event is to occur (between 0 and 1)

How to measure probability

•**Probability:** measure of how likely an event is to occur (between 0 and 1)

- Classical Definition
  - Calculate, assuming events equally likely
- Relative Frequency Approach
  - Doing an experiment, using historical data
- Subjective/Bayesian Probability
  - Trusting instinct, using judgement

LAWS & NOTATION:

<u>EX:</u> As part of a survey on whether to ban smoking inside parliament, 1000 politicians were interviewed, 600 of which were from the Yellow Party and 400 were from the Blue party. Results of the survey are as follows:

	Ban Smoking	Not Ban Smoking	
Yellow	250	350	**600**
Blue	250	150	**400**
	500	**500**	**1000**

'Basic' Events

Probability that a person chosen randomly (i.e. everyone has an equal chance of selection) from the survey is a member of the Yellow Party (Y):

$$P(Y) = 600/1000 = 0.6$$

Probability that a person chosen randomly from the survey would ban smoking (Ban):

$$P(Ban) = 500/1000 = 0.5$$

‘Combined’ Events 1 (*Intersection* of events)

Probability that a person chosen randomly from the survey is a member of the Yellow Party (Y) and would ban smoking (Ban), which is denoted by:

$$P(Y \cap Ban) = 250/1000 = 0.25$$

Probability that a person chosen randomly from the survey would ban smoking (*Ban*) and is a member of the Yellow Party (Y), denoted by:

$$P(Ban \cap Y) = 250/1000 = 0.25$$

Clearly the intersection of events does not depend on their order of listing.

‘Combined’ Events 2
(*Union* of events)

Probability that a person chosen is ***either*** from the Yellow Party (Y) ***or*** would ban smoking (Ban):

$$P(Y \cup Ban) = \frac{250 + 350 + 250}{1000} = 0.85$$

So here we have seen that:

$$P(Y \cup Ban) = P(Y) + P(Ban) - P(Y \cap Ban)$$

This is an example of the Addition Law for Probabilities, see next slide …………….

Addition Law for Probabilities (in general):

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

A special case is when Events A and B are **mutually exclusive** , i.e. they cannot both happen, in which case the addition law simplifies to:

$$P(A \cup B) = P(A) + P(B)$$

'Combined' Events 3
(Conditional events)

The <u>Conditional Probability</u> that a randomly chosen person would ban smoking (Ban) ***given*** that he/she is from the Yellow Party (Y):

$$P(Ban \mid Y) = \frac{250}{600} = \frac{5}{12}$$

i.e.:

$$P(Ban|Y) = \frac{P(Y \cap Ban)}{P(Y)}$$

Or by simple rearrangement:

$$P(Y \cap Ban) = P(Ban|Y) \times P(Y)$$

This is an example of the Multiplication Law for Probabilities, see next slide …………….

Multiplication Law for Probabilities (in general)

$$P(B/A) = \frac{P(B \cap A)}{P(A)}$$ or $$P(B \cap A) = P(B/A) \times P(A)$$

A special case is when Events A and B are **independent,**
i.e. the occurrence of A has no influence on the probability of B [and vice versa)]

i.e. P(B/A) = P(B) [and P(A/B)=P(A)]

in which case the multiplication law simplifies to:

$$P(B \cap A) = P(B) \times P(A)$$

NOTE: This notion of independence can be extended to more than 2 events.

Law of Total Probability

<u>EX:</u> The results of the survey on whether to ban smoking inside parliament could be reported as:

$P(Y) = 0.6$, $P(Ban/Y) = 5/12$

$P(B) = 0.4$, $P(Ban/B) = 5/8$

Now $P(Ban) = P(Ban \cap Y) + P(Ban \cap B)$ [*as $Ban \cap Y$ and $Ban \cap B$ are mutually exclusive & the only two ways that Ban can occur*]

$= P(Ban/Y) \times P(Y) + P(Ban/B) \times P(B)$ [*by multiplication rule*]

$= 5/12 \times 0.6 + 5/8 \times 0.4$

$= 0.25 + 0.25 = 0.5$

This is an example of the Law of Total Probability, see next slide …………….

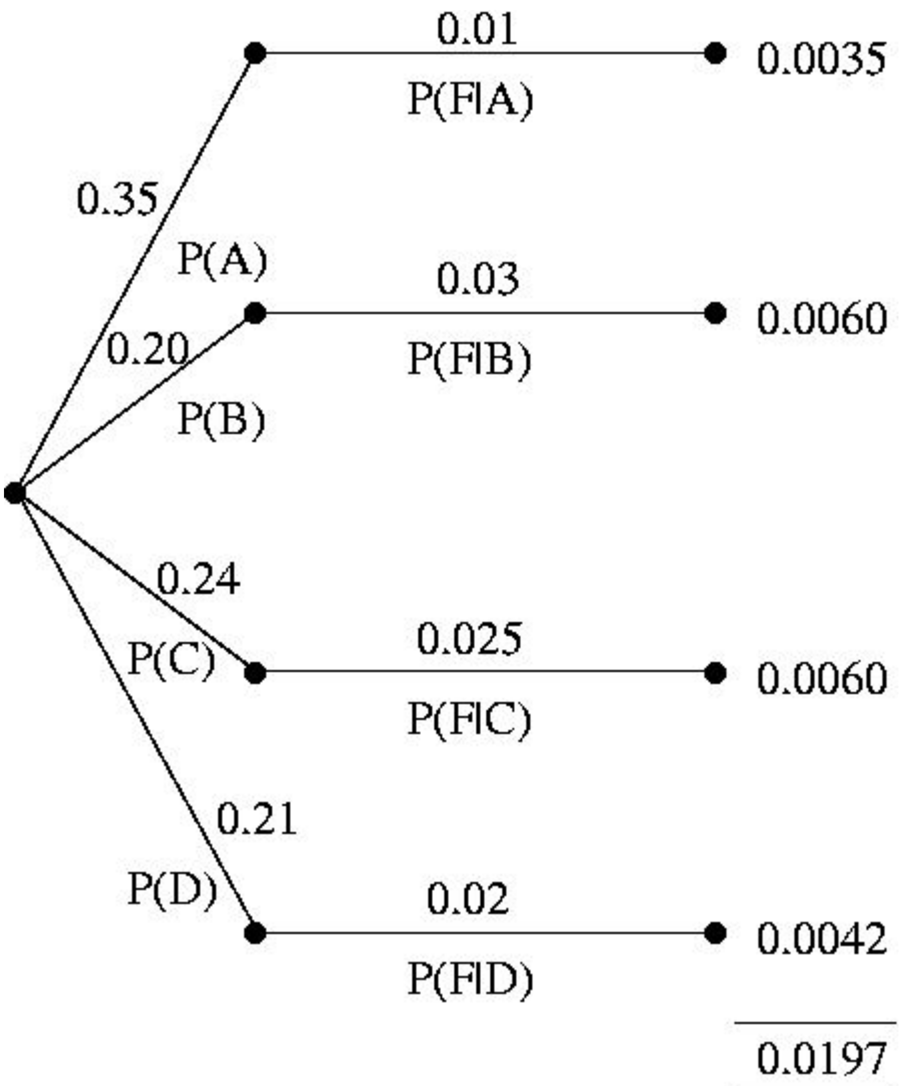
Law of Total Probability in general

Suppose events A_1, A_2, A_3, … A_n are mutually exclusive and complete (i.e. one of them must occur), then the probability of another event (B) can be calculated by weighting the conditional probabilities of B, i.e:

$$P(B) = P(B/A_1) \times P(A_1) + P(B/A_2) \times P(A_2) + \ldots P(B/A_n) \times P(A_n)$$

See another example on next slide …..

Tree Diagrams also help think about probabilities



In a factory, a brand of chocolates is packed into boxes on four production lines A, B, C, D. Records show that a small percentage of boxes are not packed properly for sale as follows

Line	A	B	C	D
% Faulty	1	3	2.5	2
% Output	35	20	24	21

What is the probability that a box chosen at random from the factory's output is faulty?

Discrete Random Variables

- A random variable is a numerical description of the outcome of an experiment.

- When the values are ‘discrete’, (e.g. value shown when die is thrown, first number drawn in lottery, daily number of patients attending A&E in Lancaster), the random variable is <u>discrete.</u>

- { For Later ***When the values are ‘continuous’ , (e.g. height of randomly selected student, temperature in Lancaster, time between patients arriving at A&E), the random variable is <u>continuous</u>.}

Probability Mass Function (pmf) of a discrete random variable

- The pmf, p(x), of a discrete random variable (X) simply lists the probabilities of each possible value, x, of the random variable.

- E.g. Sum of values when two dice are thrown:

X=	2	3	4	5	6	7	8	9	10	11	12
p(x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Expected value and variance of discrete random variables

•The expected value, or mean, of a discrete random variable is defined as:

$$\mu = E(X) = \sum xp(x)$$

•The variance of a discrete random variable is a measure of variability and is defined as:

$$\sigma^2 = E[(X-\mu)^2] = \sum (x-\mu)^2 p(x)$$

•N.B. σ is referred to as the *standard deviation* of X.

•E.g. Sum of values when two dice are thrown:

X=	2	3	4	5	6	7	8	9	10	11	12
p(x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

$$\mu = \mathrm{E}(X) = 2\left(\frac{1}{36}\right) + 3\left(\frac{2}{36}\right) + \cdots + 12\left(\frac{1}{36}\right) = 7$$

$$\sigma^2 = \mathrm{Var}(X) = (2-7)^2\left(\frac{1}{36}\right) + (3-7)^2\left(\frac{2}{36}\right) + \cdots + (12-7)^2\left(\frac{1}{36}\right) = 5.83$$

$$\sigma = 5.83^{0.5} = 2.42$$

<u>Expected value</u> and <u>variance</u> of <u>combinations</u> of discrete random variables

- If X and Y are random variables and a and b are constants:

$$\mathrm{E}(aX + bY) = a\mathrm{E}(X) + b\mathrm{E}(Y);$$

- If X and Y are also <u>independent</u> random variables (i.e. the value of X has no influence on the value of Y (e.g. two throws of a die))

$$\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$$

and

$$\mathrm{Var}\,(aX + bY) = a^2\,\mathrm{Var}(X) + c^2\,\mathrm{Var}(Y).$$

Law of Total Probability for Expected Values of a discrete random variable

Suppose events A_1, A_2, A_3, … A_n are mutually exclusive and complete (i.e. one of them must occur), then the expected value of a r.v. X can be calculated by weighting the conditional expected values of X, i.e.:

$$E(X) = E(X/A_1) \,.\, P(A_1) + E(X/A_2) \,.\, P(A_2) + \ldots E(X/A_n) \,.\, P(A_n)$$

- For example if 20% of the working population work at home and the remaining 80% travel an average of 10 miles to work, the overall average distance travelled to work is:

  E(travel to work distance) = 0 x 0.2 + 10 x 0.8 = 8 miles.

Poisson Random Variable

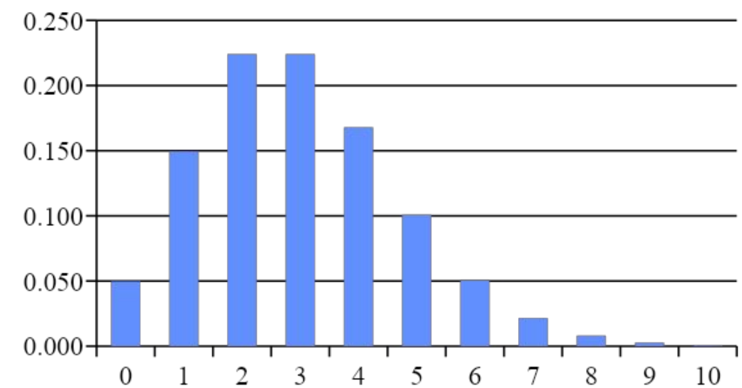
- The most important discrete random variable in stochastic modelling is the <u>Poisson</u> random variable.

- A Poisson distribution with mean β has pmf:

$$p(x) = \frac{e^{-\beta}\beta^{x}}{x!}$$

and its variance $\sigma^2 = \beta$.

e.g. p(x) for Poisson r.v. with □=3.5



•And Poisson probabilities can be obtained directly from the formula, statistical tables or computer software.

Poisson Random Variable

- **The General Theory:**

- When 'events' of interest occur '***at random'*** at rate λ per unit time;
- No. of events in period of time T has a Poisson distribution with mean λT

- Events in real stochastic processes (e.g. arrivals of customers at a bank, calls to a call centre, patients to an A&E department, breakdowns in equipment) occur 'at random' when there are a large number of potential customers/ callers/ patients/ components each with independent probabilities of arriving/ calling/ falling ill/ breaking.

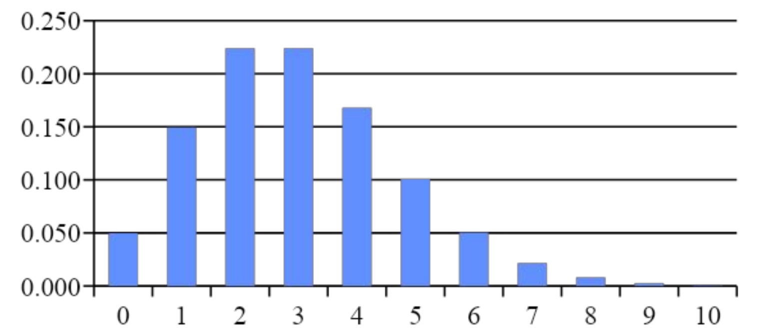
Poisson Random Variable

- **Example:**

- If arrivals of customers to a bank are ***at random***, at an average rate of 0.6 per minute, then:

- *No. of arrivals in 5 minute period has Poisson distribution with mean = 3.*

x	p(x)
0	0.050
1	0.149
2	0.224
3	0.224
4	0.168
5	0.101
6	0.050
7	0.022
8	0.008
9	0.003
……	……



Continuous Random Variables

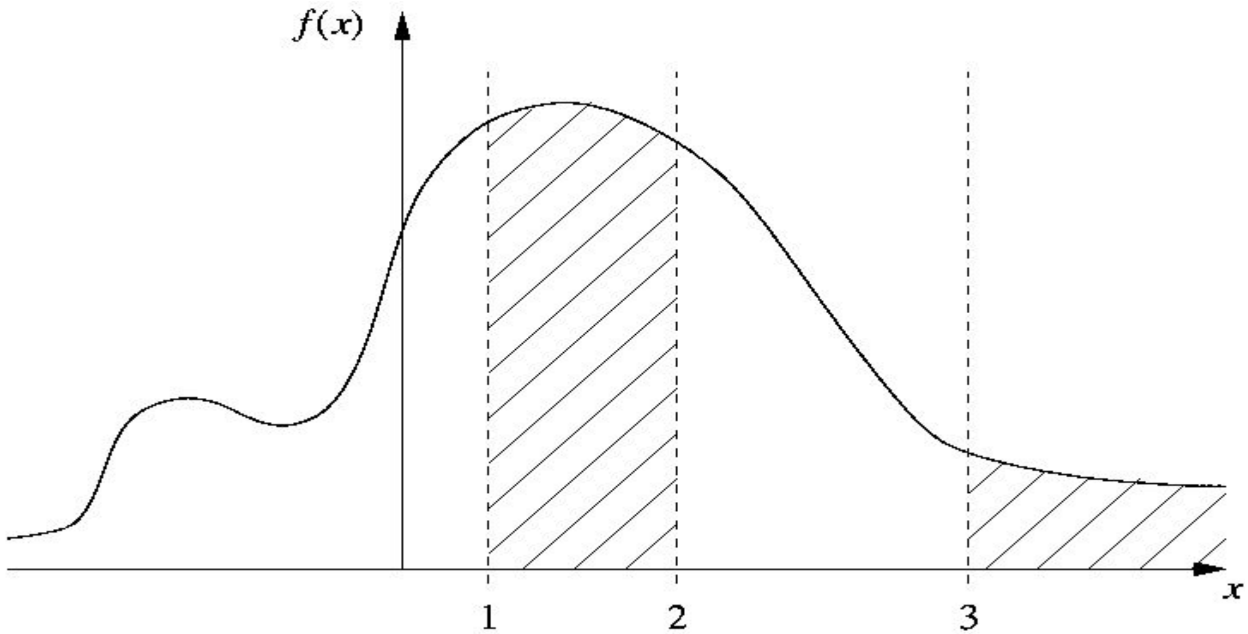
- A random variable is a numerical description of the outcome of an experiment.

- When the values are ‘continuous’ , (e.g. height of randomly selected student, temperature in Lancaster, time between patients arriving at A&E), the random variable is <u>continuous</u>.

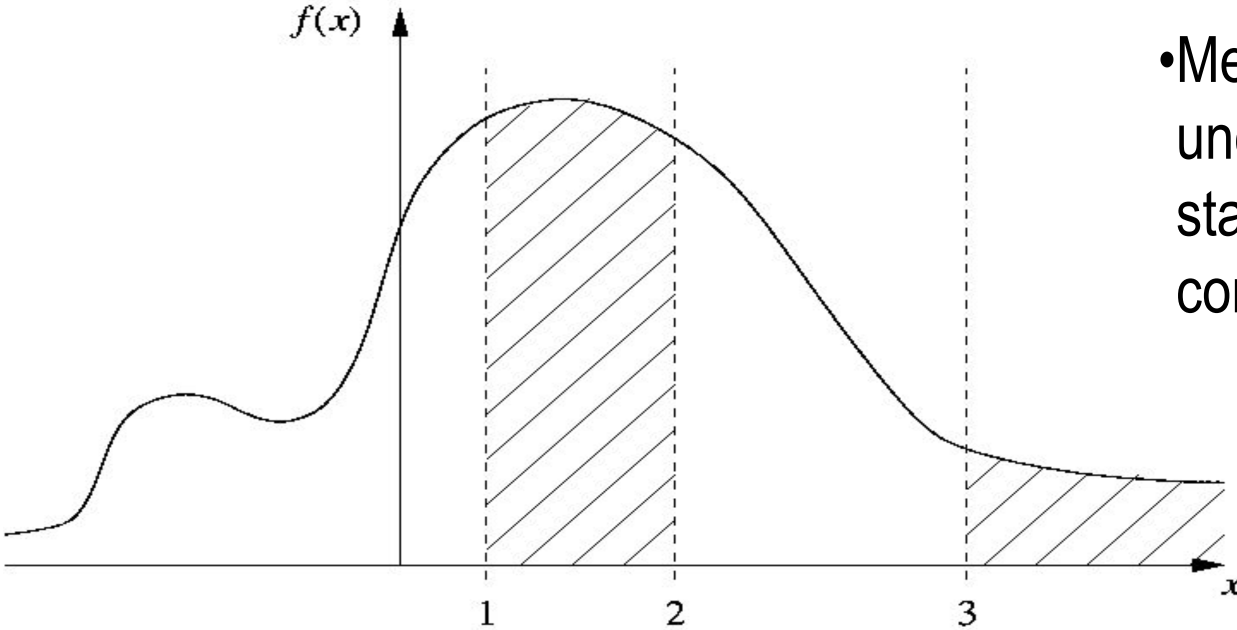
Probability Density Functions (p.d.f.)

•The random behaviour of a continuous random variable X is captured by the ***probability density function*** (p.d.f.) $f(x)$, which is as follows:

- $f(x)$ is never negative
- the total area under $f(x)$ is one (total probability)



Probabilities involving X are obtained by determining the area under the pdf, i.e. $f(x)$, for the appropriate range of values − e.g.



- Methods to find areas under pdf are integration, statistical tables and computer software

$$P(1 < X < 2) = \int_1^2 f(x)dx = \text{Area under } f(x) \text{ between } x = 1 \text{ and } x = 2.$$

$$P(X \geq 3) = \int_3^\infty f(x)dx = \text{Area under } f(x) \text{ beyond } x = 3.$$

Expected value and variance of continuous random variables

- The expected value, or mean, of a continuous random variable is defined as:

$$\mu = \mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x) dx$$

- The variance of a continuous random variable is a measure of variability and is defined as:

$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$$

- N.B. σ is referred to as the *standard deviation of* X.

Expected value and variance of combinations of continuous random variables

NB. Exactly same as for discrete r.v.

- If X and Y are random variables and a and b are constants:

$$E(aX + bY) = aE(X) + bE(Y);$$

- If X and Y are also independent random variables (i.e. the value of X has no influence on the value of Y (e.g. two throws of a die))

$$E(XY) = E(X)E(Y)$$

and

$$\text{Var}\,(aX + bY) = a^2\,\text{Var}(X) + c^2\,\text{Var}(Y).$$

Law of Total Probability for Expected Values of a continuous random variable

NB. Exactly same as for discrete r.v.

Suppose events A_1, A_2, A_3, … A_n are mutually exclusive and complete (i.e. one of them must occur), then the expected value of a r.v. X can be calculated by weighting the conditional expected values of X, i.e:

$$E(X) = E(X/A_1) \,.\, P(A_1) + E(X/A_2) \,.\, P(A_2) + \ldots E(X/A_n) \,.\, P(A_n)$$

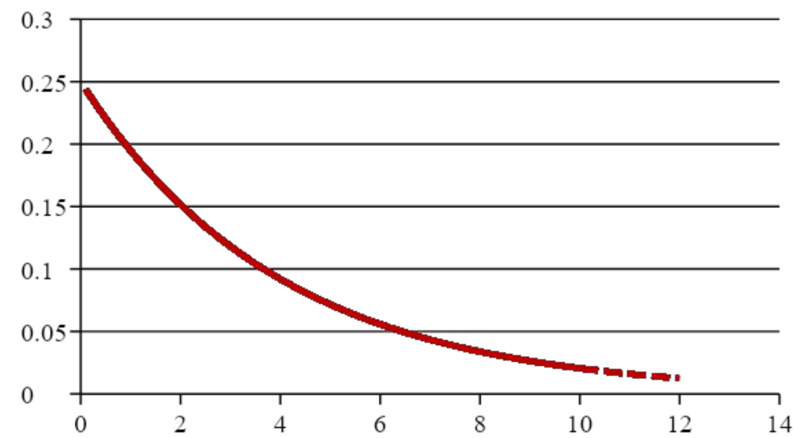
Exponential Random Variable

- The most important continuous random variable in stochastic modelling is the <u>Exponential</u> random variable.
- Exponential distribution with mean $1/\gamma$ has pdf:

$$f(x) = \gamma e^{-\gamma x}$$

- and its variance $\sigma^2 = (1/\gamma)^2$, i.e. variance = mean^2

E.g. *f(x)* for Exponential r.v. with mean=4.



•And areas under curve follow from simple formula:

$$P(a < X < b) = \int_a^b f(x)dx = e^{-a\gamma} - e^{-b\gamma}$$

Exponential Random Variable

- **<u>The General Theory:</u>**

- When 'events' of interest occur '***at random'*** at rate λ per unit time (as is common in real stochastic processes – see earlier note);

- The time between events has an Exponential distribution with mean $1/\lambda$.

- And the time to the next event has an Exponential distribution with mean $1/\lambda$, whether or not an event has just occurred. [This is the <u>memoryless property</u> of the Exponential distribution – and is counter-intuitive!].

- Conversely, if the gaps between events are independent and from an exponential distribution with mean $1/\lambda$, the events occur '***at random'*** at rate λ per unit time.

Exponential Example

THEORY

- IF Events occur '***at random***', at rate A per unit time.

THEN: Time between events has an Exponential distribution, with mean of 1/A.

AND: Time to next event has an Exponential distribution, with mean of 1/A.

EXAMPLE

- IF Arrivals to bank occur '***at random***', at rate of 0.6 per minute.

THEN: Time between arrivals has an Exponential distribution, with mean = 1.667 minutes.

AND: Time to next arrival has an Exponential distribution, with mean of 1.667 minutes.

Normal Random Variable

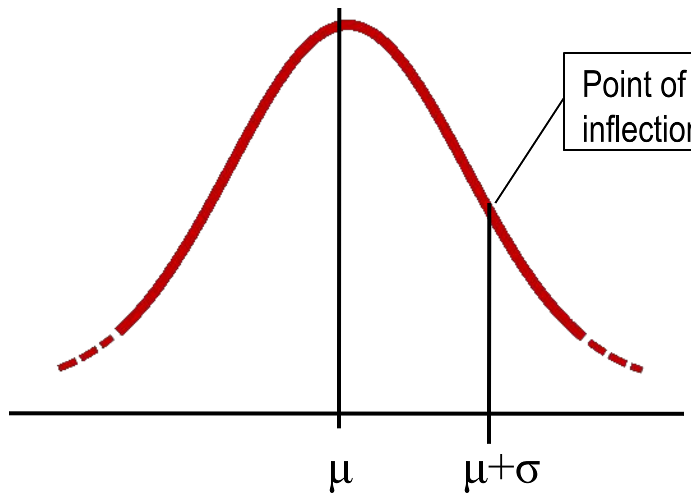
- The most important continuous random variable in statistics is the Normal random variable.

- Normal distribution with mean μ and variance σ^2 has pdf:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{[-(x-\mu)^2/2\sigma^2]}$$

& hence its standard deviation is σ.

E.g. *f(x)* for Normal r.v. with mean=μ and stan dev σ..



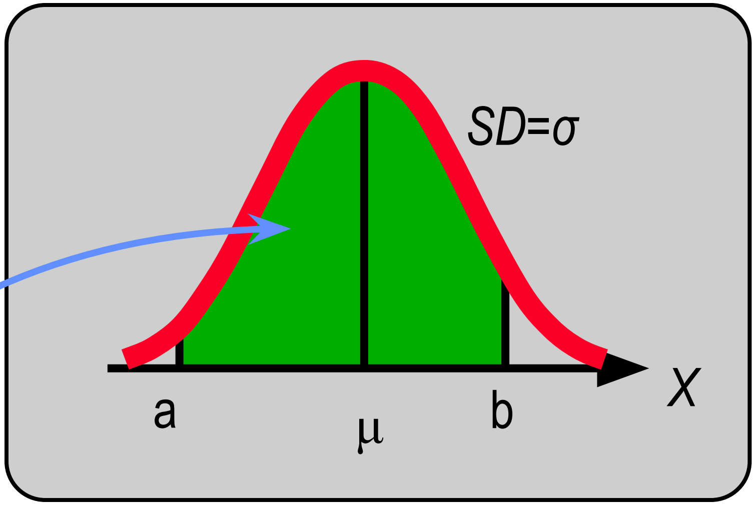
•And areas under curve are obtained from statistical tables or from computer software.

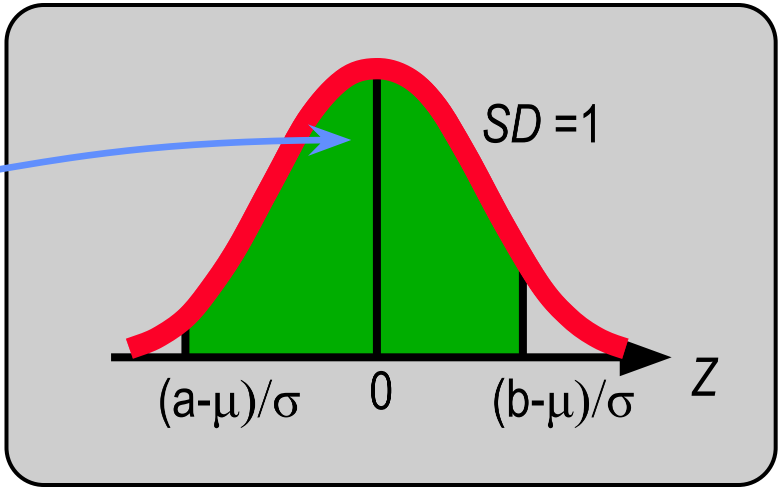
Standardised Normal Random Variable

- Areas under <u>any Normal curve</u> (mean=μ & SD=σ) can be found by finding the 'equivalent' area under the <u>standard Normal curve</u> (mean=0 & SD=1), i.e.:

$$P(a < X < B)$$

$$= P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right)$$





Normal Random Variables

Important (in general) because:

- Many naturally occurring r.v.'s have a Normal distribution, e.g. weights, heights …

- Many useful statistics behave as Normal r.v.'s, even if the r.v.'s from which they derive are not Normal, e.g. Central Limit Theorem.