University of Southampton

SESA6085 – Advanced Aerospace Engineering Management

Lecture 2

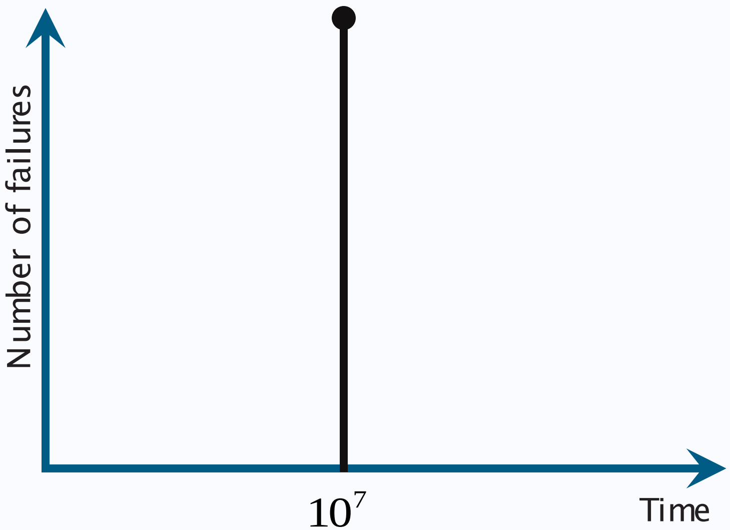
2024-2025

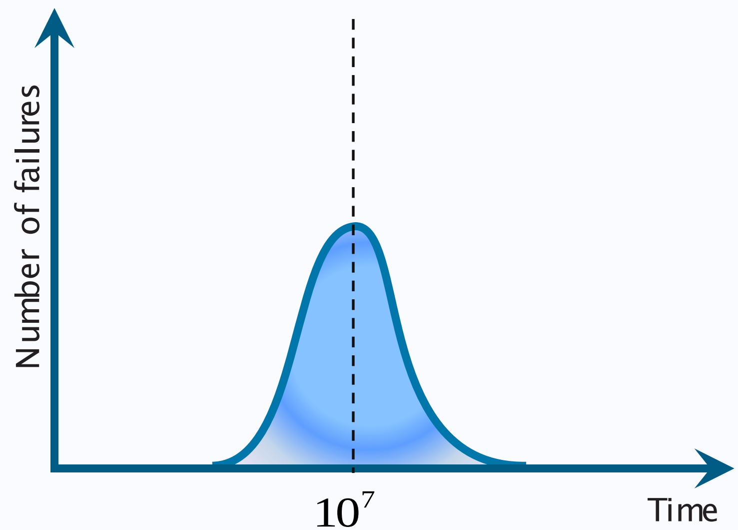
Dr David Toal

Continuous Variations

Question

- A power supply unit fails at an average rate of once per 10 million hours. Which distribution is more likely? Why?

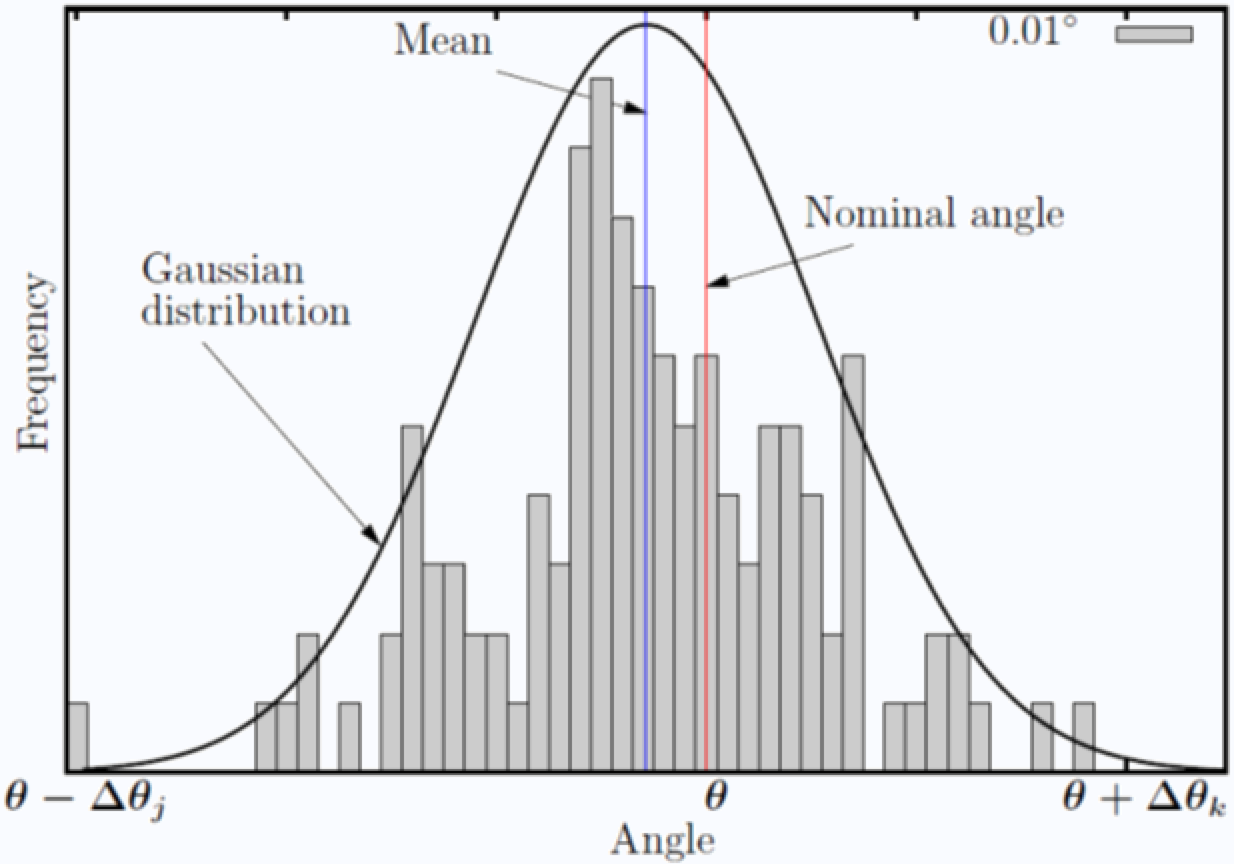


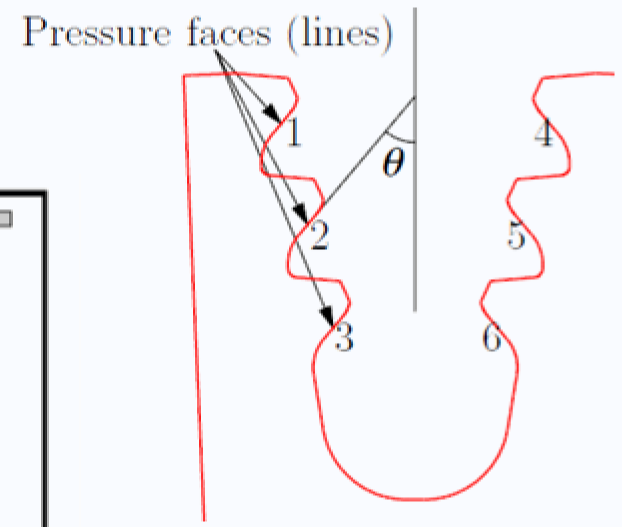


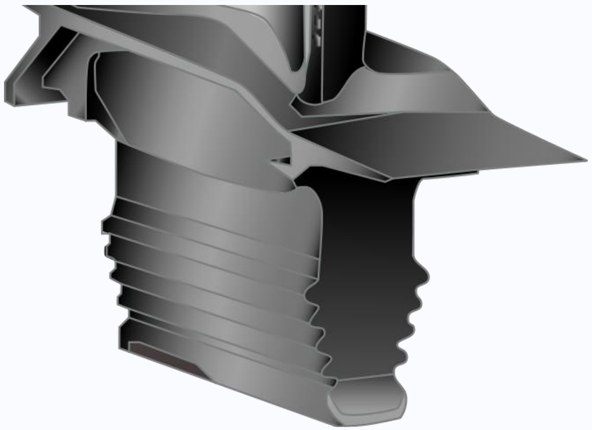
Variations

- Variations of parameters in engineering applications (machined dimensions, material strengths, transistor gains, resistor values, temperatures, pressures, etc.) can be described in two ways:
  1. State minimum and maximum values (i.e. tolerances): This provides no information of the actual shape of the distribution
  2. Describe the nature of the statistical distribution of the parameter, e.g. using data derived from measurements and manufacturing data
- Variations are everywhere and modelling their distribution is very important
  - Time or cycles – useful for reliability
  - Distance, force, stress – useful for robust design

Variations In A Gas Turbine[1]





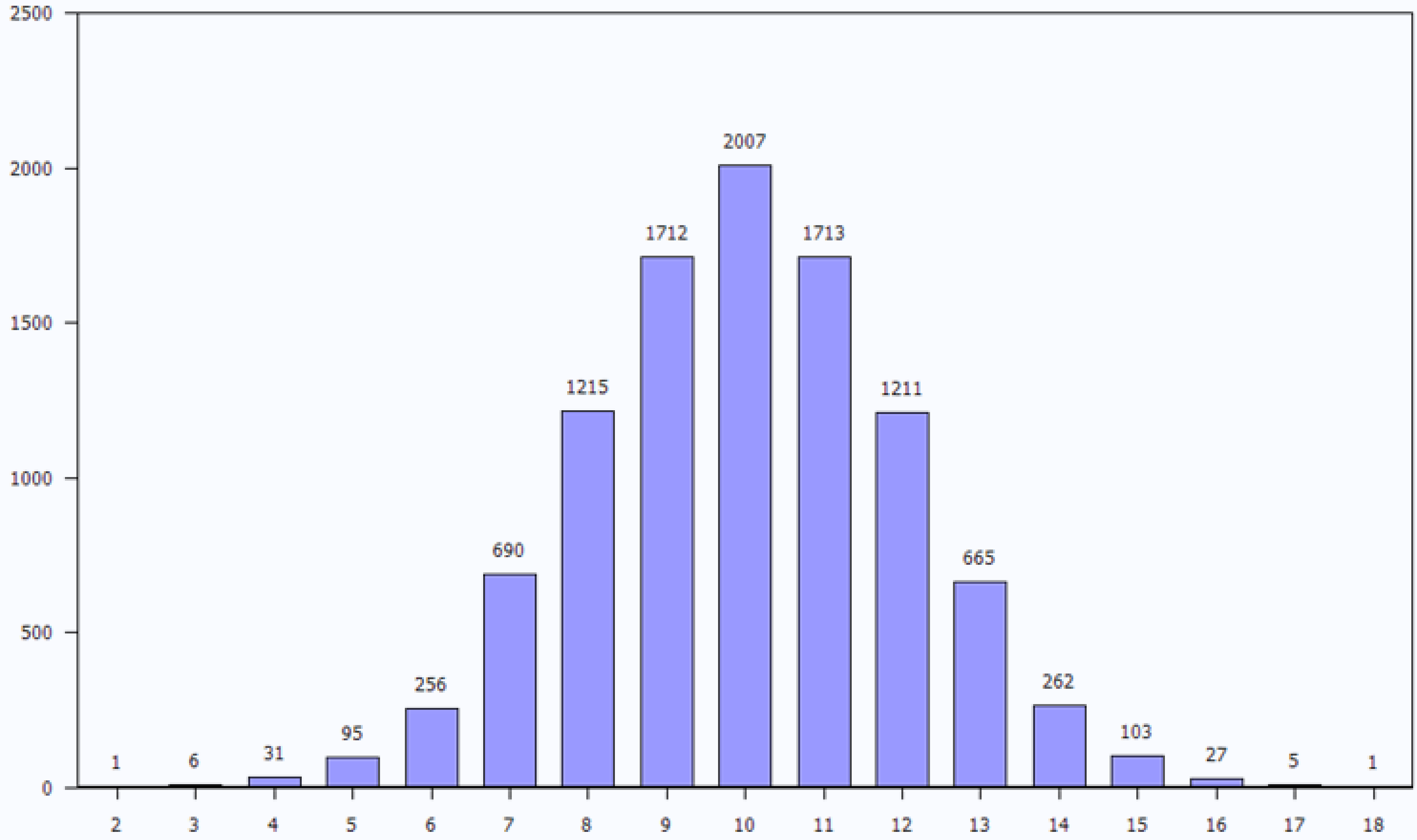


[1] Deshpande et al. "Geometric parameterisation of firtree joints in gas turbine discs considering manufacturing variability", 3rd Aircraft Structural Design Conference, 2012

Variations In The Gym….

Frequency Histograms

- Histograms offer a discrete representation of a continuous variation and should be a familiar concept



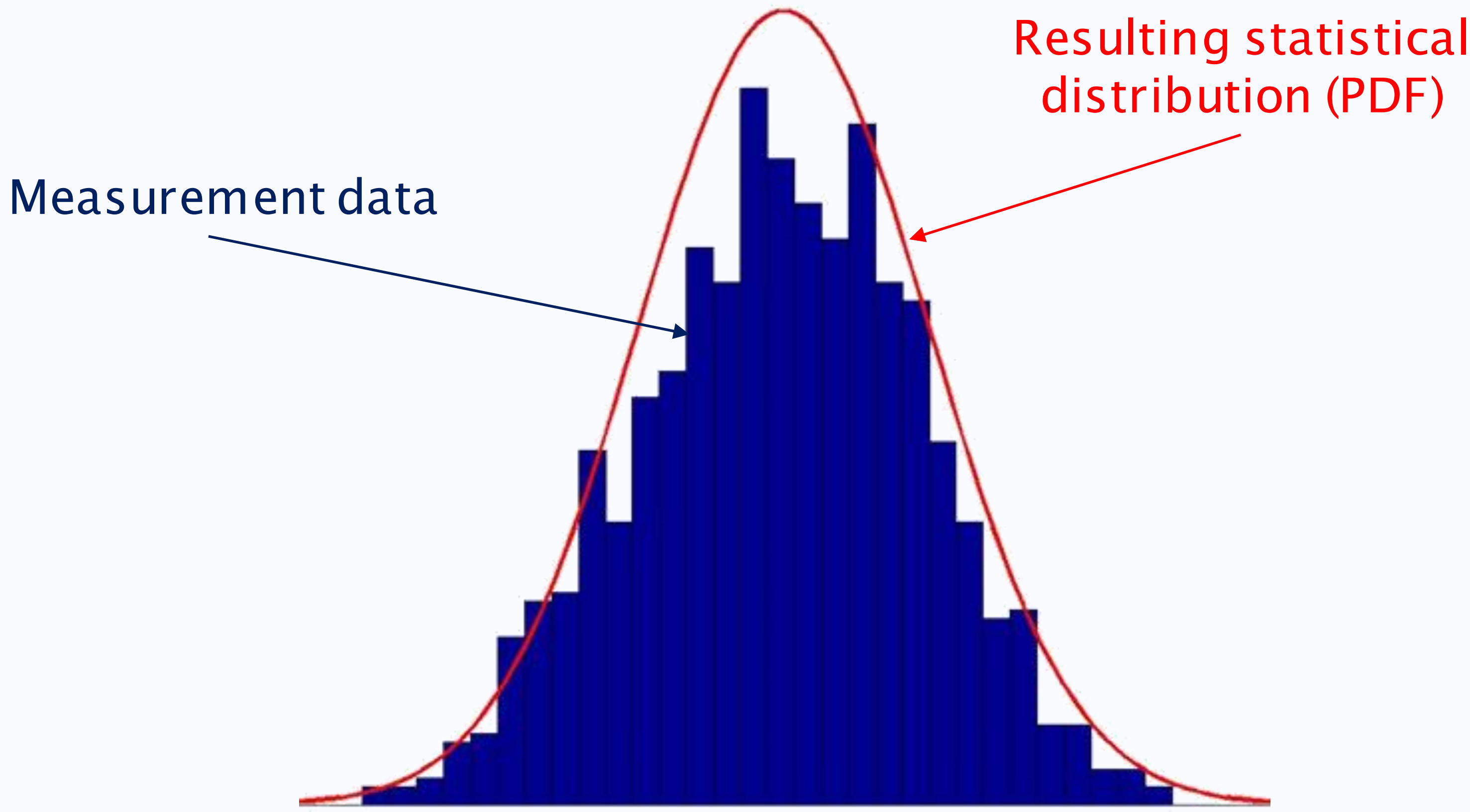
Probability Density Function (PDF)

- As the number of samples increases and the measurement interval decreases, the histogram approaches a continuous curve which describes the population *probability density function* (PDF).
- The area under the PDF curve is equal to unity.

$$\int_{-\infty}^{\infty} f(t)\ \mathrm{d}t = 1$$

- In reliability terms this is the probability that a component fails in the period [*t* to *t*+*dt*]

Probability Density Function (PDF)



N.B. Measurement data and statistical distribution are not to the same scale

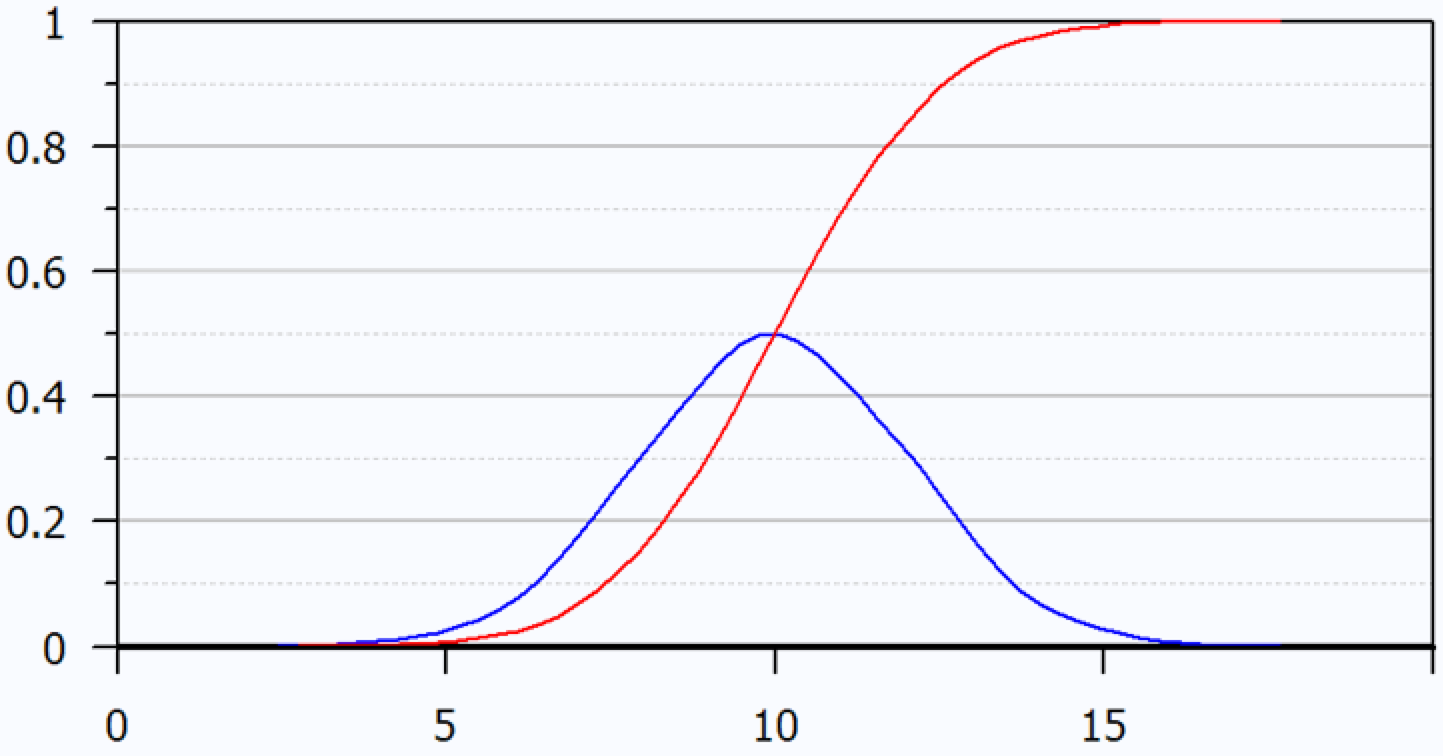
Cumulative Distribution Function (CDF)

- The *cumulative distribution function* (CDF) gives the probability that a measured value falls between -∞ and t_1

$$F(t) = \int_{-\infty}^{t_1} f(t)\,\mathrm{d}t$$

- In reliability terms this gives the probability that a component fails between [-∞, t_1]

PDF & CDF

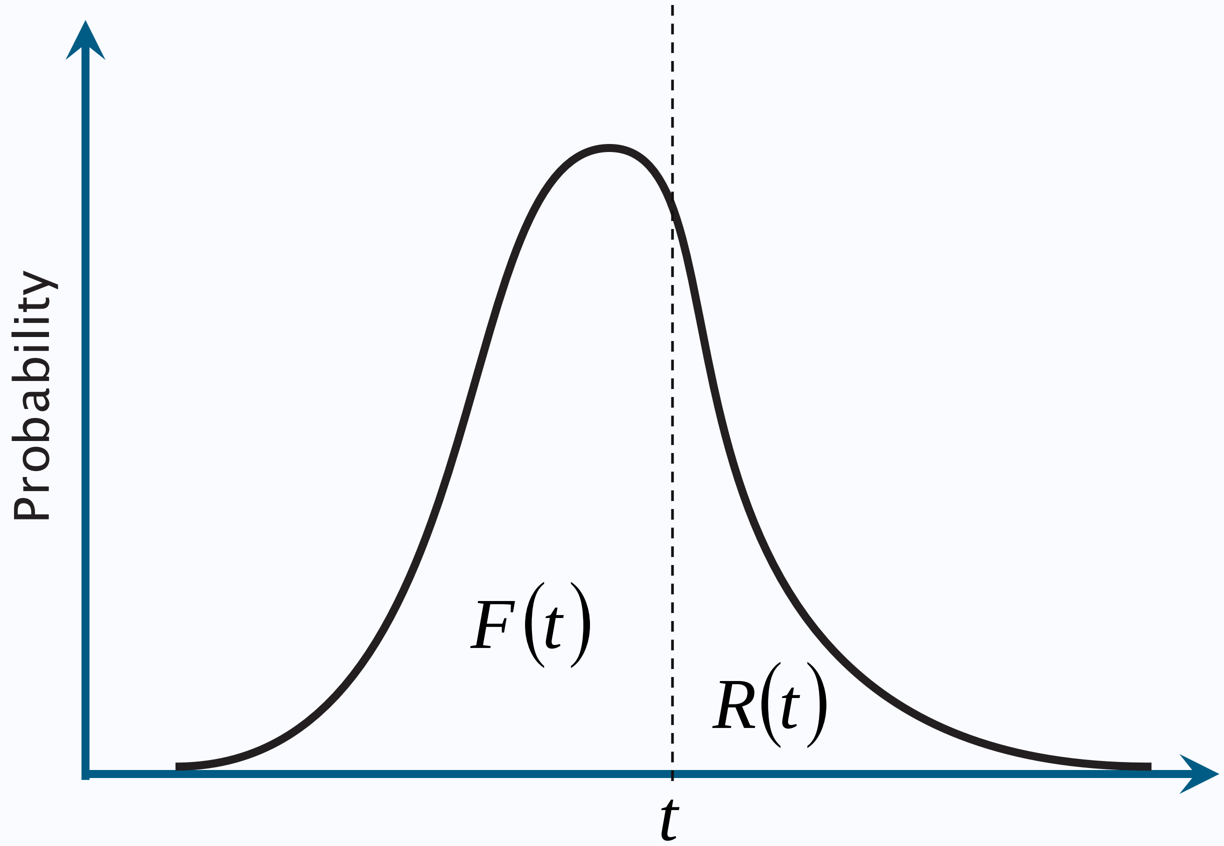


Reliability Function

- Reliability engineering concerns with the probability that an item will survive for a stated interval, i.e. that there is no failure in the interval (0 to t_1).
- This is given by the reliability function $R(t)$

$$R(t) = 1 - F(t) = 1 - \int_{-\infty}^{t_1} f(t)\ \mathrm{d}x \equiv \int_{x}^{\infty} f(t)\ \mathrm{d}x$$

Reliability Function



Hazard Function

- The hazard function or hazard rate *h(t)* is the conditional probability of failure in the interval (t, $t+\mathrm{d}t$), given that there was no failure by t:
- This can be thought of as the probability of imminent failure at t or an indicator of the "proneness" to failure after t

$$h(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)}$$

- N.B the hazard function is not a density or probability but a measure of "risk" i.e. the greater the hazard between t_1 and t_2 the greater the risk of failure

Cumulative Hazard Function

- The cumulative hazard function $H(t)$ is given by

$$H(t) = \int_{-\infty}^{t_1} h(t)\,\mathrm{d}t = \int_{-\infty}^{t_1} \frac{f(t)}{1 - F(t)}\,\mathrm{d}t$$

- The greater the value of $H(t)$ the greater the risk of failure by t_1

f(t), F(t), R(t), h(t) & H(t)

- It should be noted that *f*, *F*, *R*, *h* and *H* give mathematically equivalent descriptions and given any of these functions it is possible to deduce the other four
- For example, because,

$$H(t) = \int_{-\infty}^{t} h(t)\,\mathrm{dt} = \int_{-\infty}^{t} \frac{f(t)}{1 - F(t)}\,\mathrm{d}t$$

- Then,

$$H(t) = -\ln(1 - F(t)) = -\ln(R(t))$$

Continuous Distribution Functions

Continuous Distribution Functions

- There are a considerable number of continuous distribution functions each with a different form
- Let's consider a few of the important ones…

- We will assume that the continuous variable of interest is denoted t
- From a reliability perspective time is very important, this does not mean that t cannot be replaced by any other parameter of interest

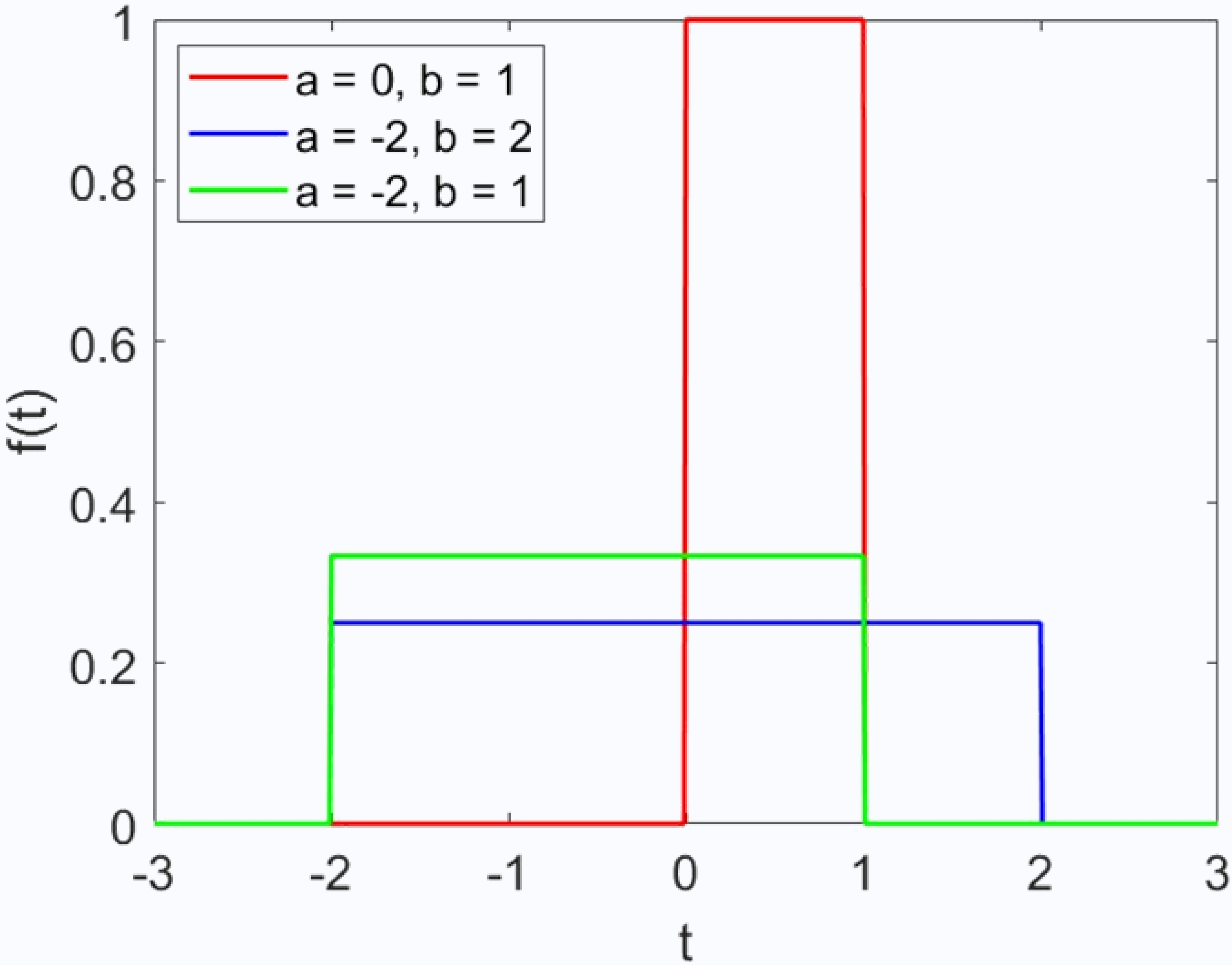
Uniform Distribution Function

- Simplest of all distributions functions
  - Otherwise known as a rectangular distribution

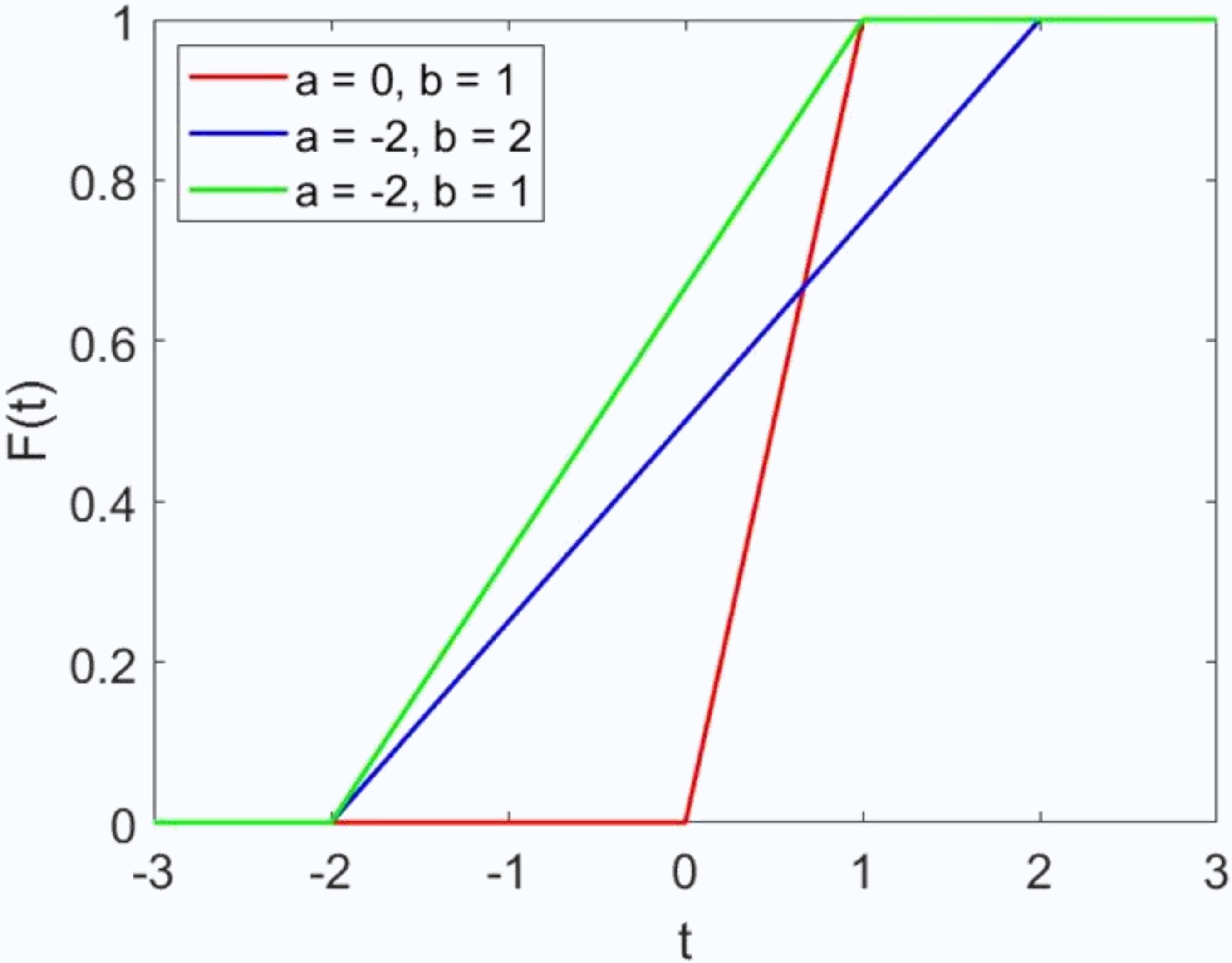
$$f(t) = \begin{cases} \dfrac{1}{b-a} & t \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

$$F(t) = \begin{cases} 0 & t < a \\ \dfrac{t-a}{b-a} & t \in [a, b] \\ 1 & t > b \end{cases}$$

Uniform Distribution Function



Uniform Distribution Function

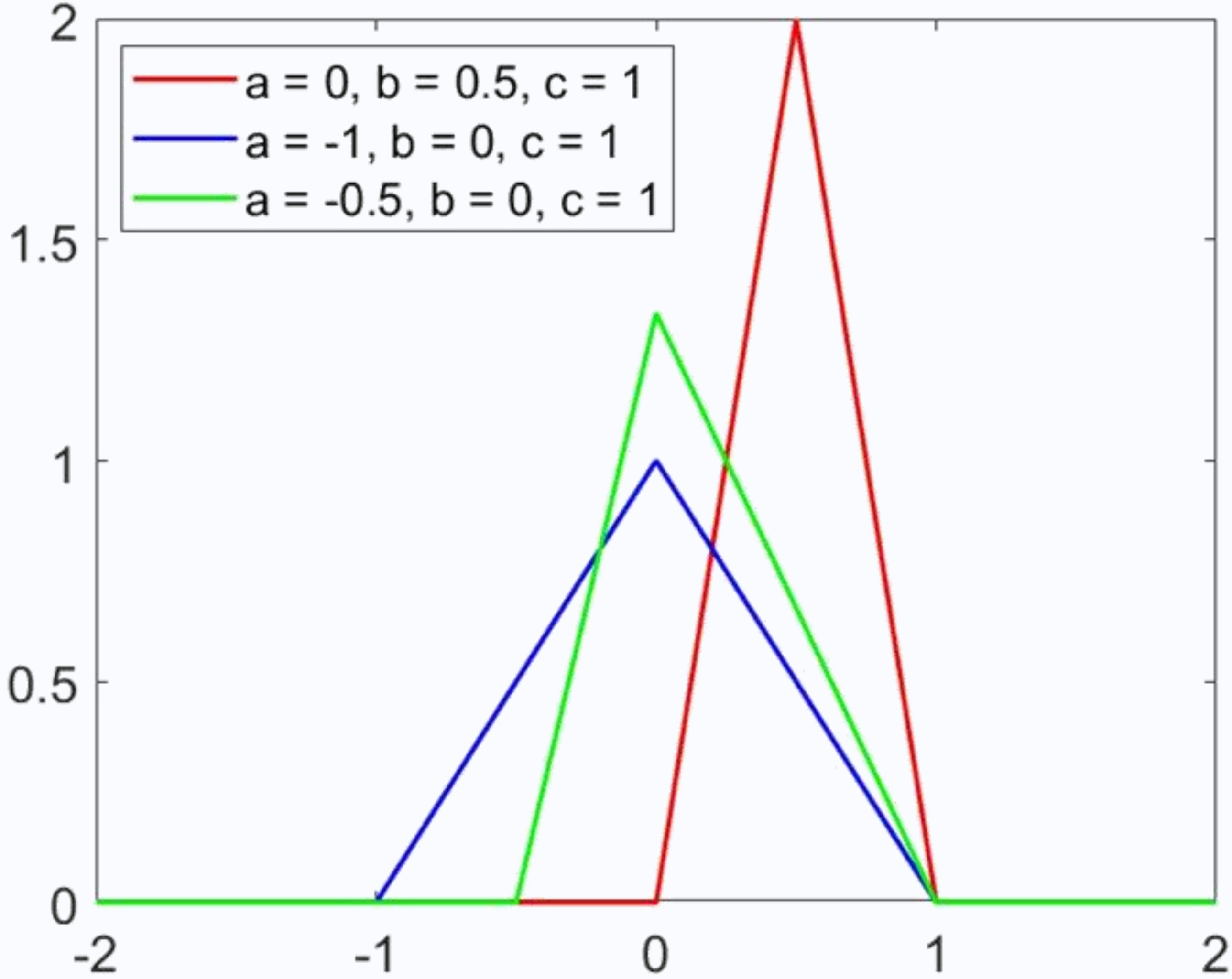


Triangular Distribution Function

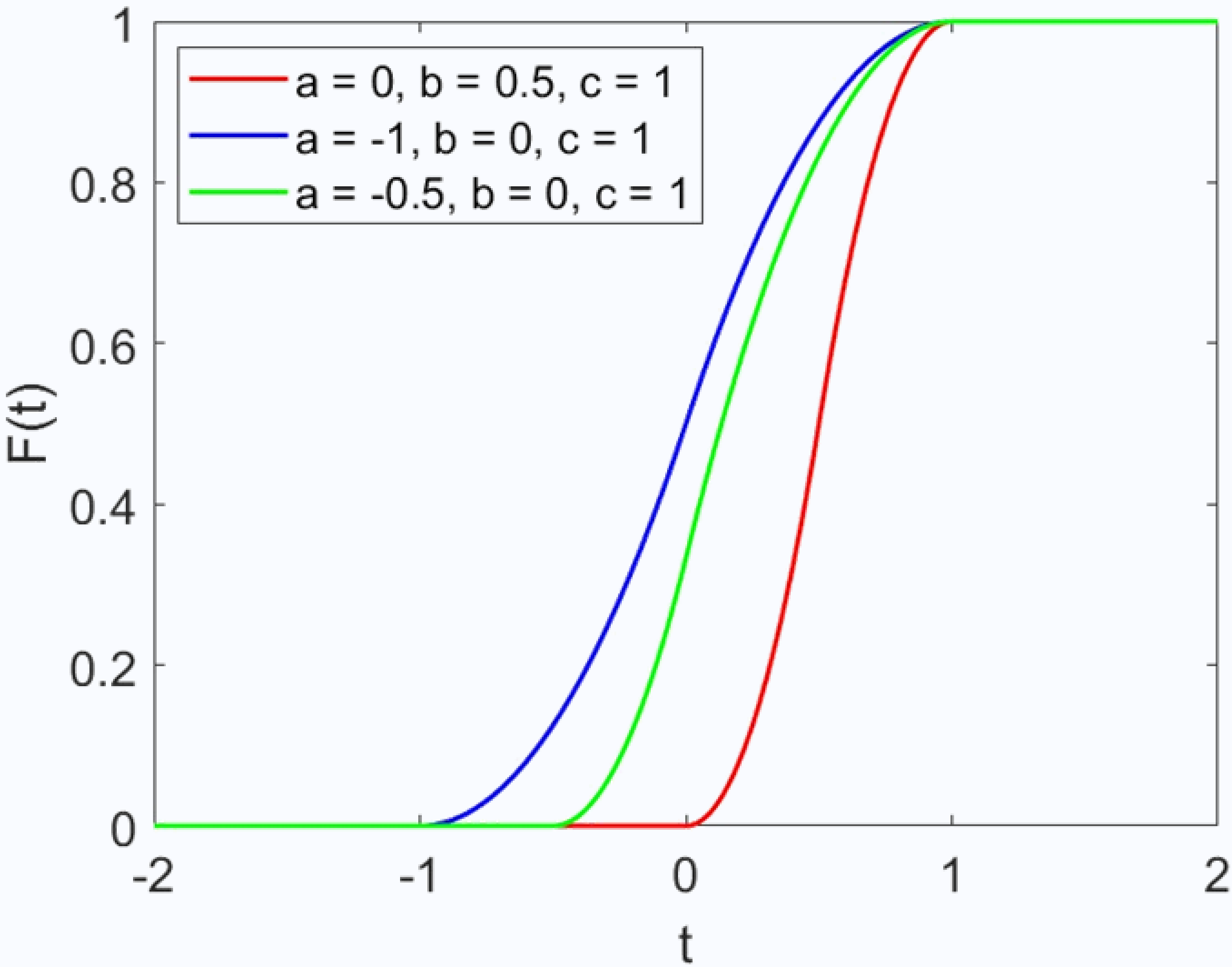
- PDF defined by a minimum (*a*), peak (*b*) and maximum (*c*)

$$f(t) = \begin{cases} \dfrac{2(t-a)}{(c-a)(b-a)} & a \le t \le b \\ \dfrac{2(c-t)}{(c-a)(c-b)} & b < t \le c \\ 0 & t < a, t > c \end{cases}$$

Triangular Distribution Function



Triangular Distribution Function



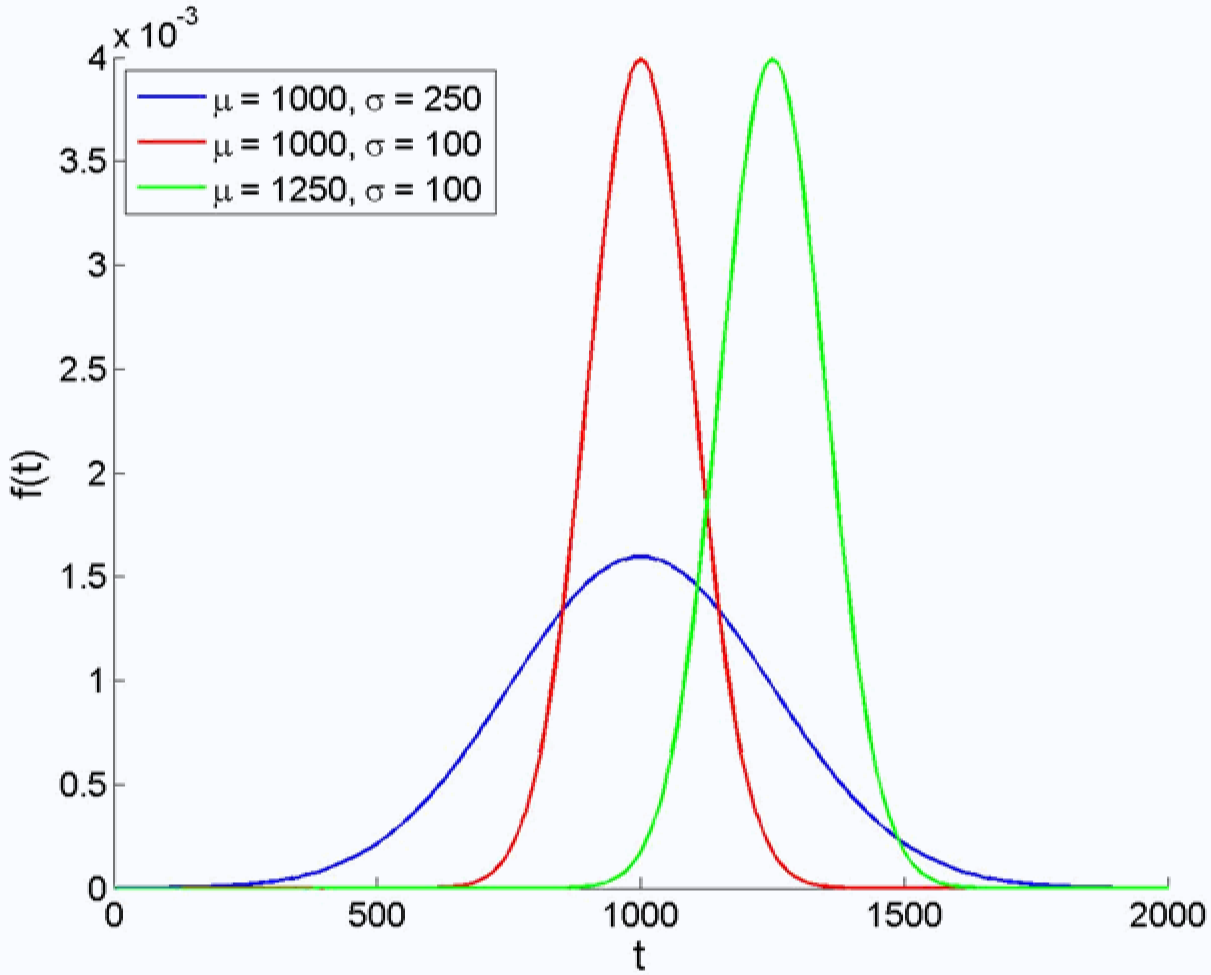
Gaussian (s-normal) Distribution

- Generally the most widely used model of variation
- The PDF is:

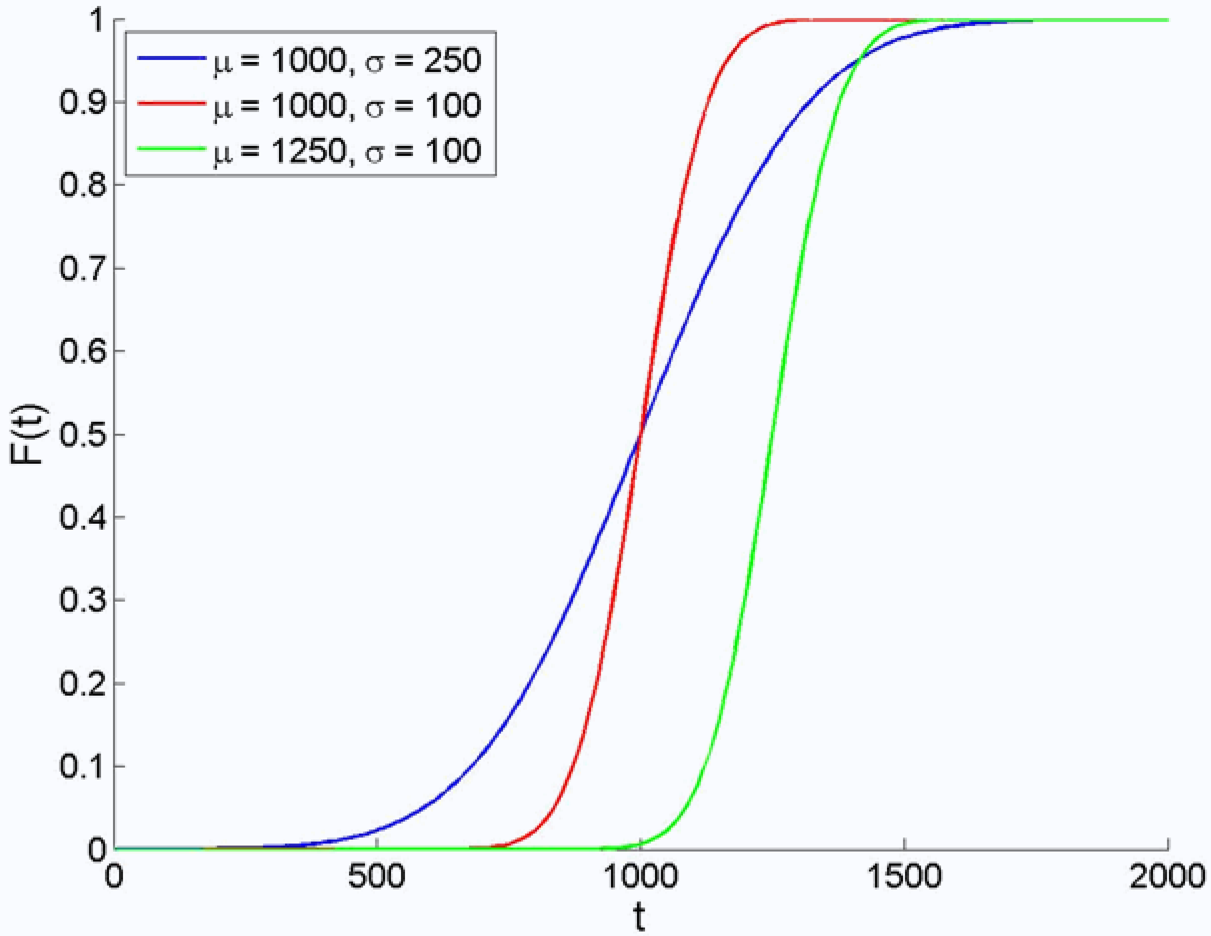
$$f(t) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right]$$

- μ - location parameter, mean
- σ – scaling parameter, standard deviation

Gaussian Distribution – PDF

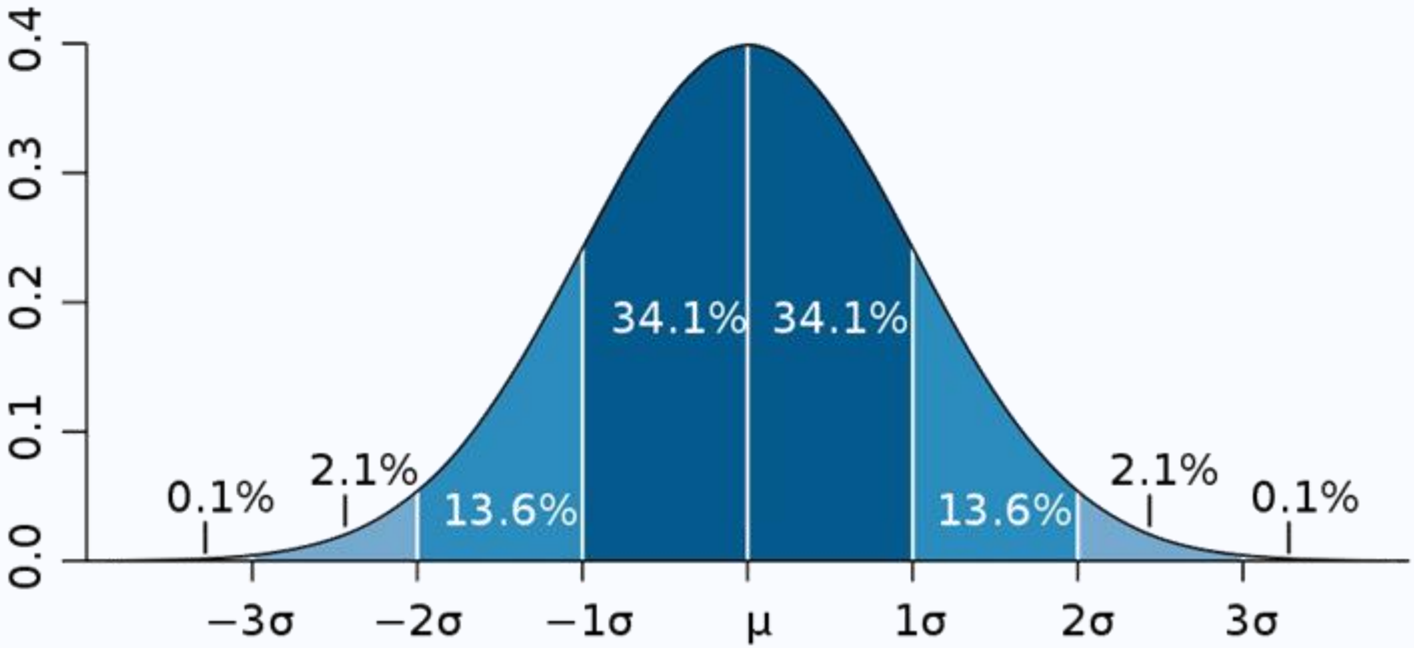


Gaussian Distribution – CDF



Gaussian Distribution

- A useful result is that lying within 1 standard deviation of the mean is 68.26% of the population (34.13% either side)
  - 95.44% within 2 standard deviations
  - 99.74% within 3 standard deviations



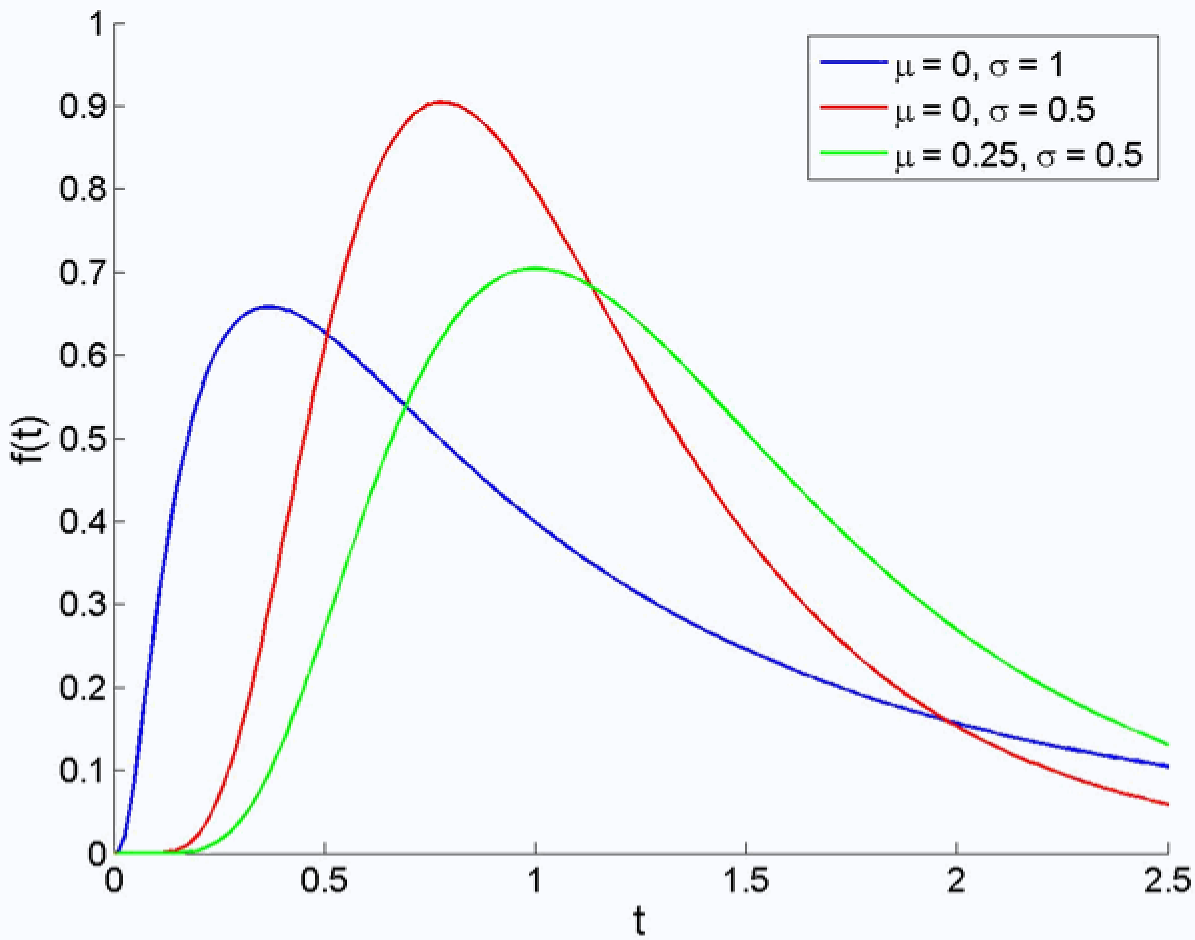
Lognormal Distribution

- More versatile than Gaussian distribution
- Has a wider range of shapes and it is often a better fit to reliability data, such as items that suffer from wearout.
- Lognormal PDF is:

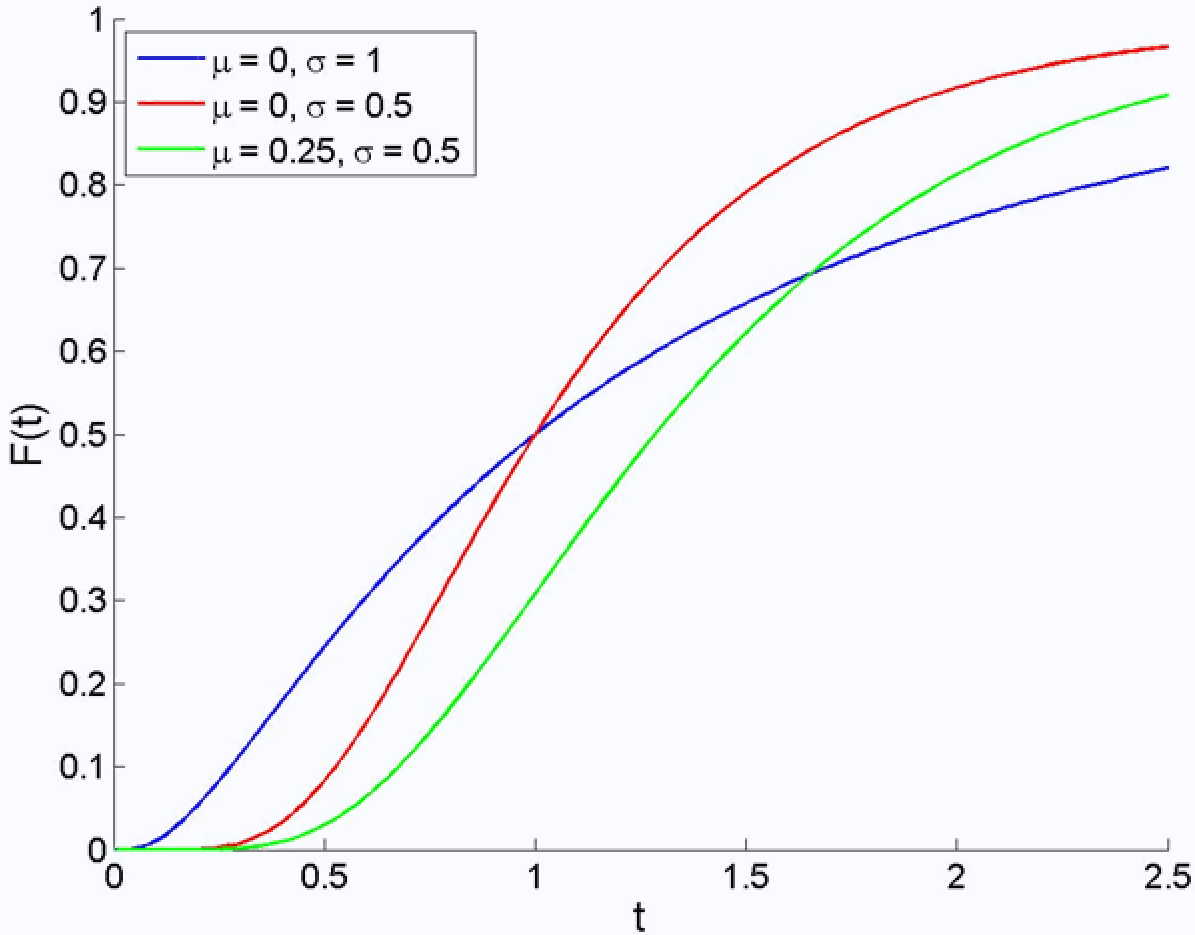
$$f(t) = \begin{cases} \dfrac{1}{t\sigma(2\pi)^{1/2}} \exp\left[-\dfrac{1}{2}\left(\dfrac{\ln t - \mu}{\sigma}\right)^2\right] & (\text{for } t \geq 0) \\ 0 & (\text{for } t < 0) \end{cases}$$

- μ - location parameter
- σ – scaling parameter

Lognormal Distribution - PDF



Lognormal Distribution - CDF



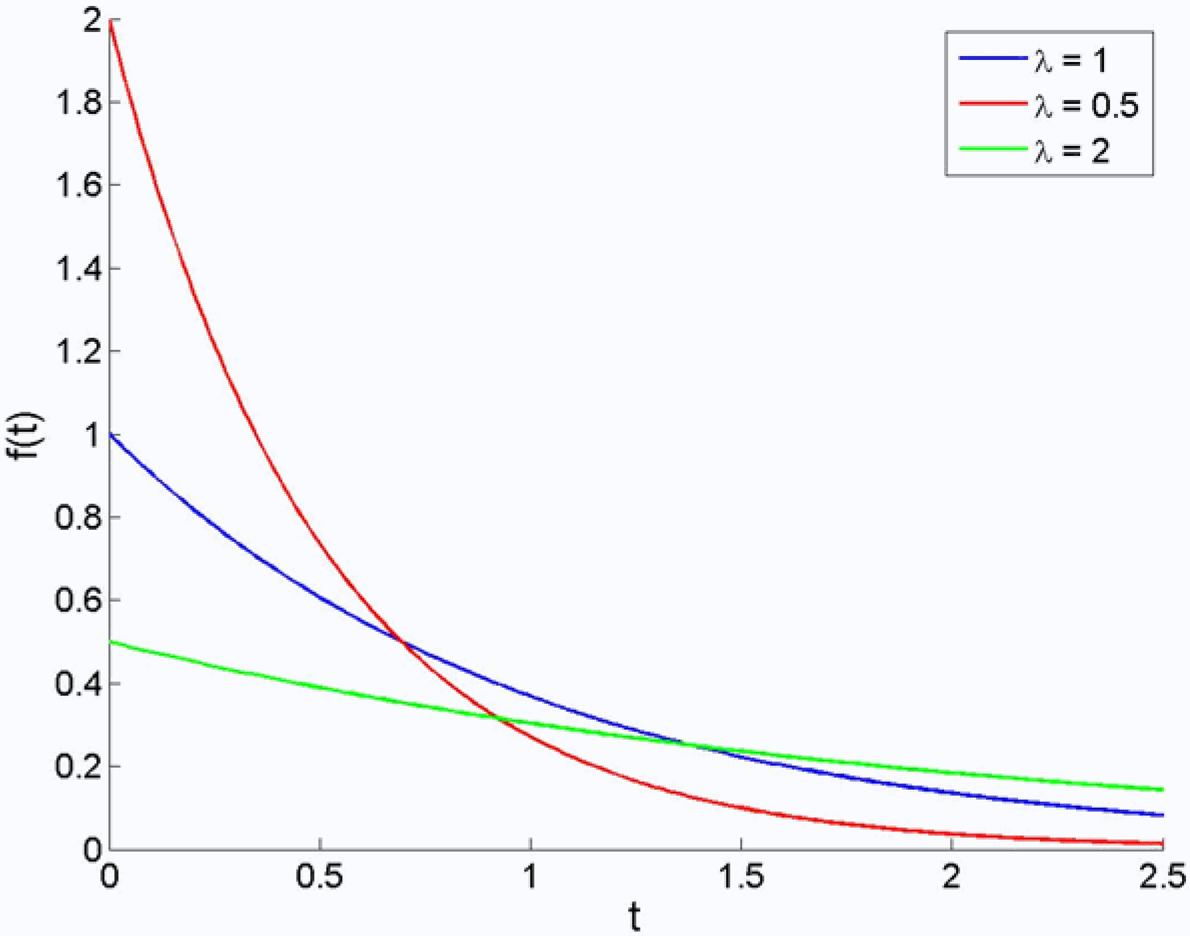
Exponential Distribution

- The exponential distribution describes the situation wherein the hazard rate is constant
- The PDF is:

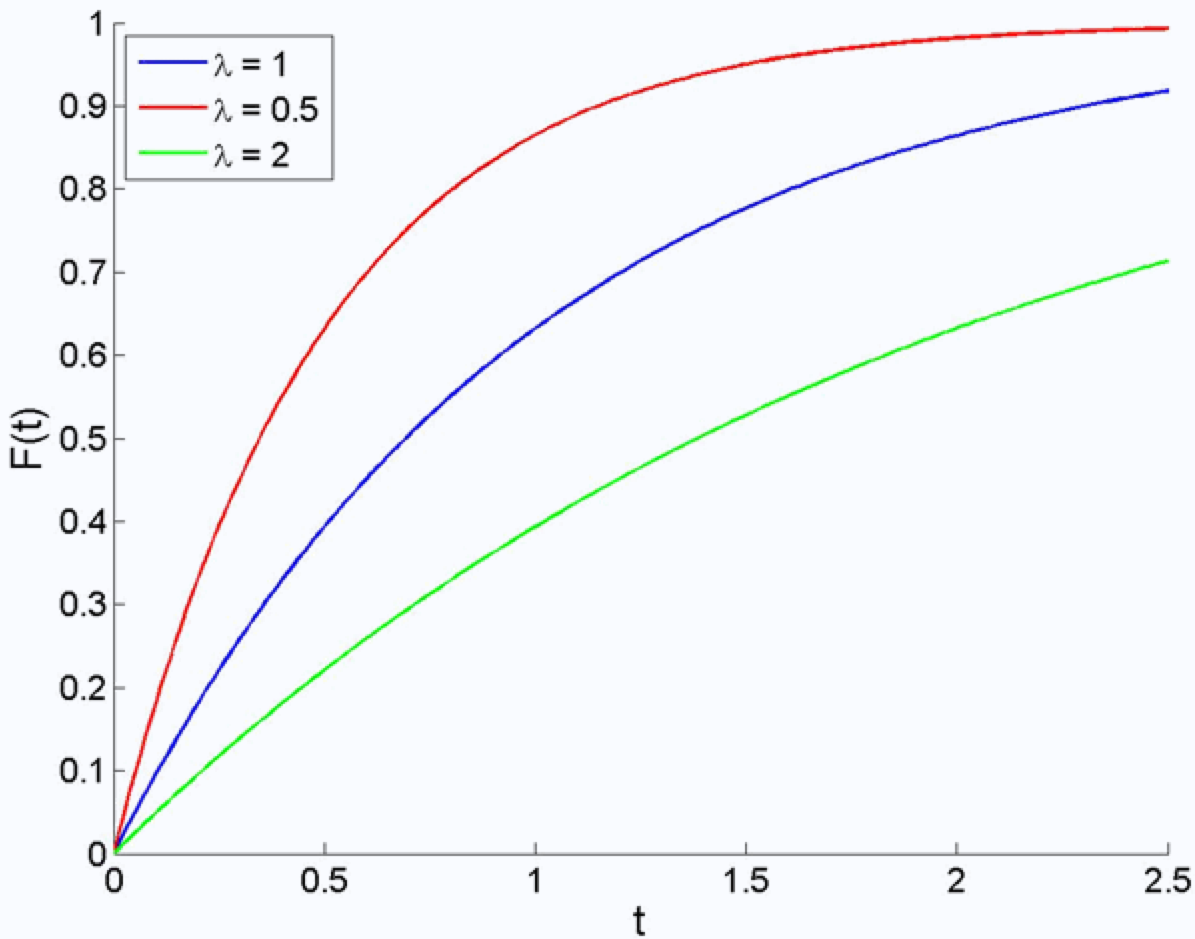
$$f(t) = \begin{cases} \lambda \exp(-\lambda t) & (\text{for } t \geq 0) \\ 0 & (\text{for } t < 0) \end{cases}$$

- λ – Scaling parameter
- λ is also the constant hazard rate
- $^1/_\lambda$ is often termed the mean time to failure (MTTF) in reliability literature

Exponential Distribution - PDF



Exponential Distribution - CDF



Weibull Distribution

- The Weibull distribution is widely used in reliability work as it is very flexible
- The Weibull PDF is

$$f(t) = \begin{cases} \dfrac{\beta}{\eta^{\beta}} t^{\beta-1} \exp\left[-\left(\dfrac{t}{\eta}\right)^{\beta}\right] & (\text{for } t \geq 0) \\ 0 & (\text{for } t < 0) \end{cases}$$

- β – shape parameter
- η – scaling parameter
- η is sometimes referred to as the characteristic life the point at which 63.2% of the population will have failed

Weibull Distribution

- The Weibull CDF is

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]$$

- The hazard rate is

$$\frac{\beta}{\eta^{\beta}} t^{\beta - 1}$$

Weibull Distribution - Flexibility

- When β=1, the Weibull distribution gives a constant hazard rate (exponential reliability function)
- When β>1 $\Rightarrow$ increasing hazard rate
- When β<1 $\Rightarrow$ decreasing hazard rate
- When β=3.5, the Weibull distribution approximates the normal distribution

Three Parameter Weibull Distribution

- The Weibull distribution is often modified to include an additional location parameter, γ
- Useful when failures, for example, only start after a finite amount of time

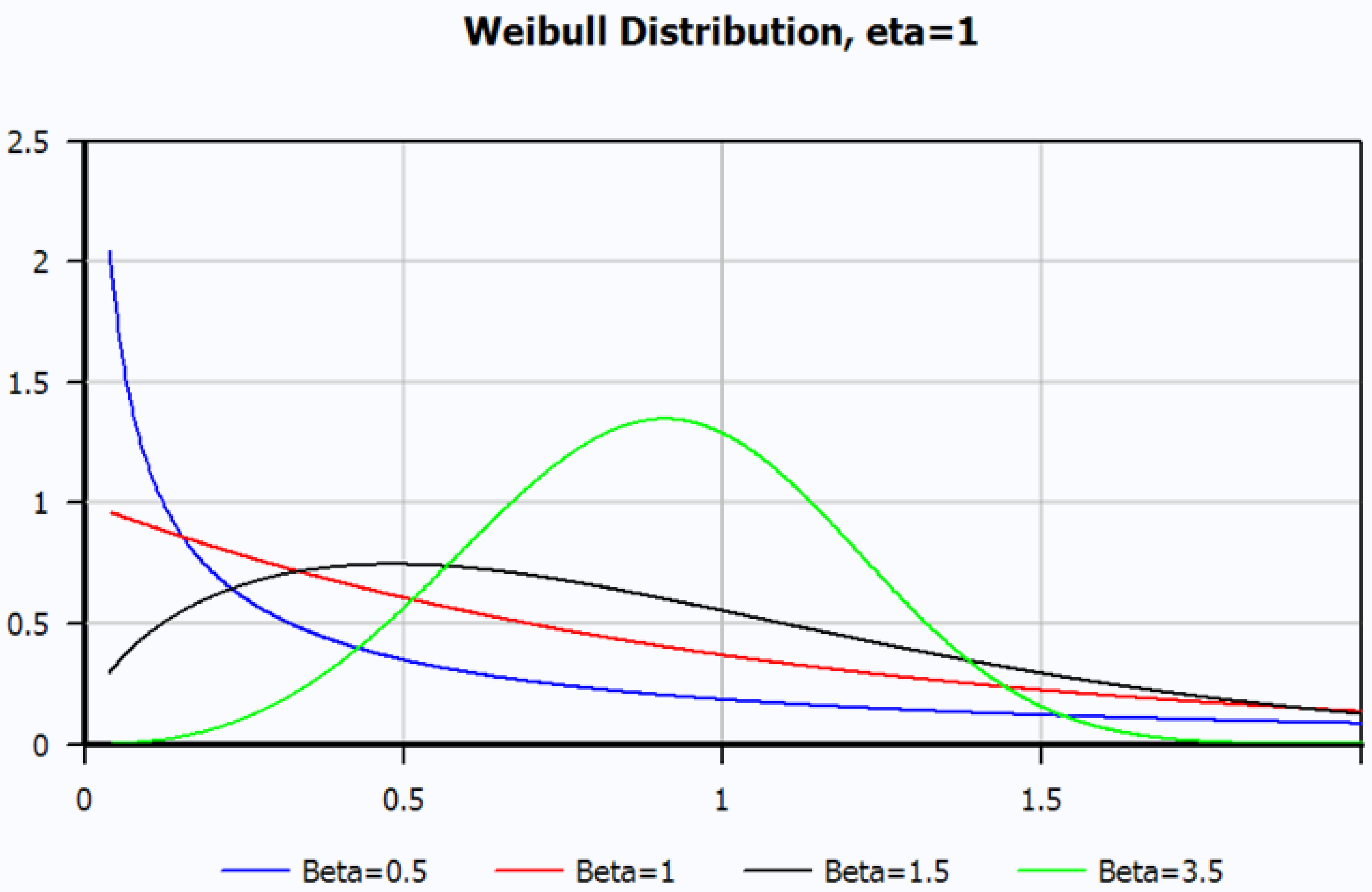
$$f(t) = \begin{cases} \dfrac{\beta}{\eta^{\beta}}(t-\gamma)^{\beta-1}\exp\left[-\left(\dfrac{t-\gamma}{\eta}\right)^{\beta}\right] & (\text{for } t \geq 0) \\ 0 & (\text{for } t < 0) \end{cases}$$

Three Parameter Weibull Distribution

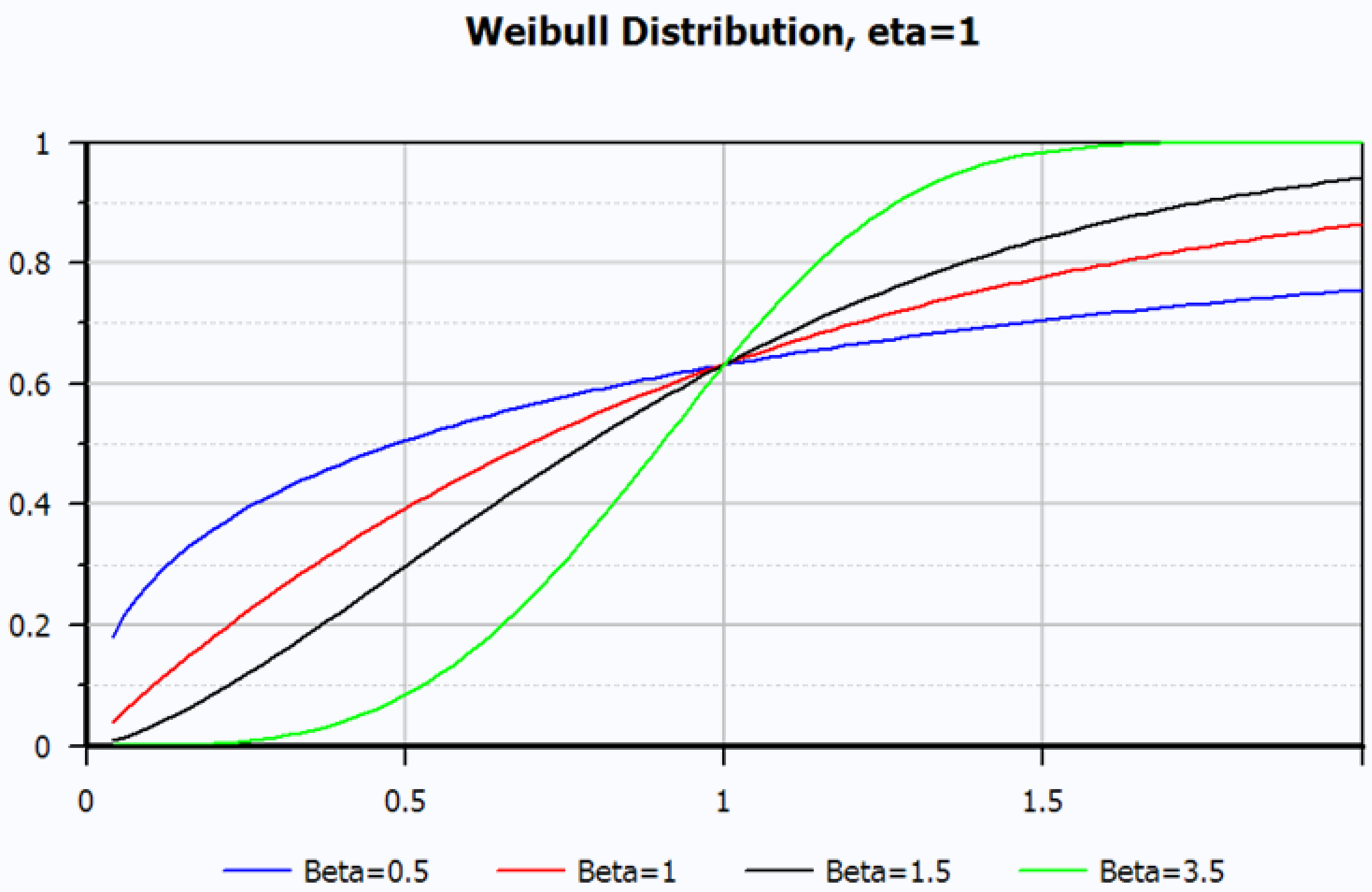
- The CDF is also translated, and becomes

$$F(t) = 1 - \exp\left[-\left(\frac{t-\gamma}{\eta}\right)^{\beta}\right]$$

Weibull Distribution - PDF



Weibull Distribution - CDF



Other Distribution Functions

(Not examinable)

Rayleigh Distribution

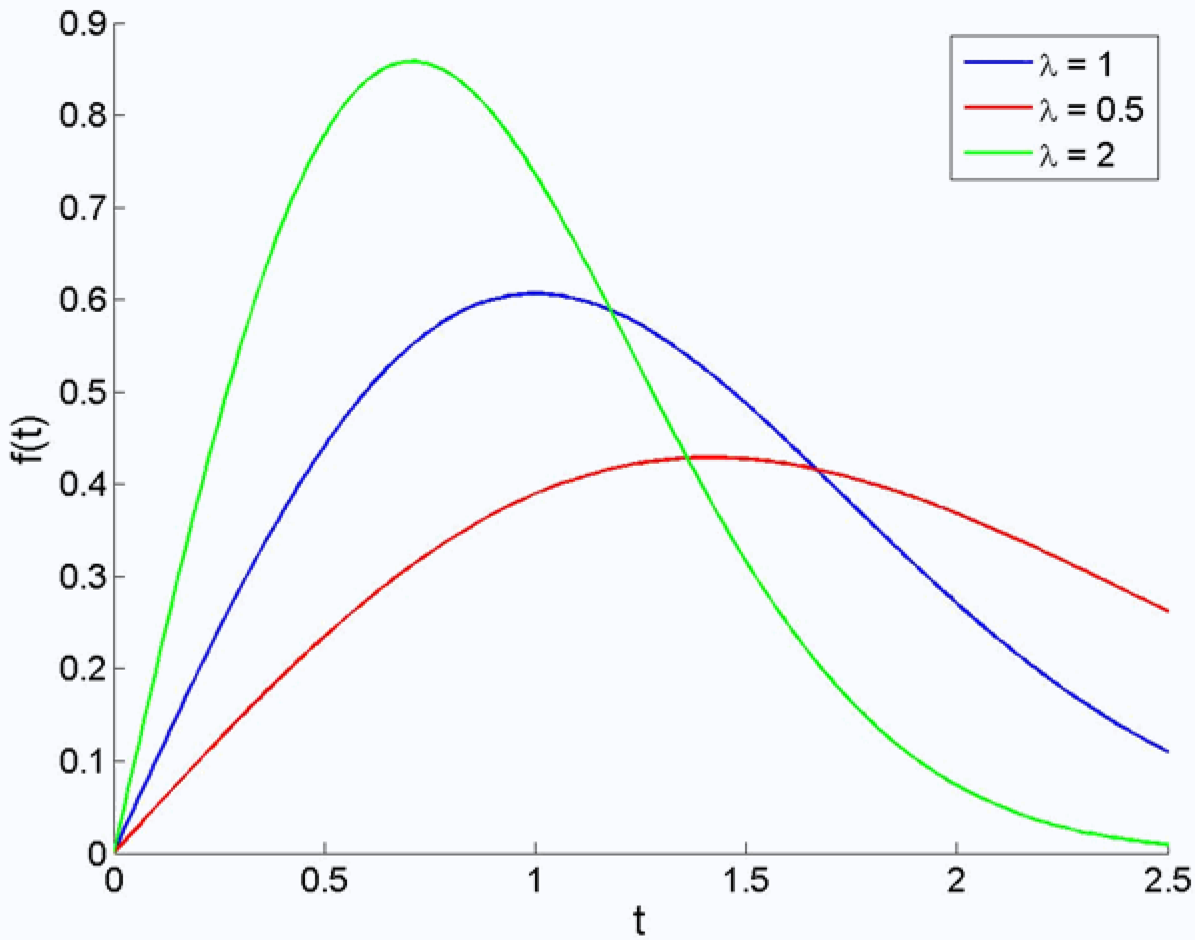
- Linearly increasing hazard rate
- Similar in form to the exponential distribution
- λ is a constant

$$f(t) = \lambda t \exp(-{}^{\lambda t^2}/_{2})$$

$$F(t) = 1 - \exp(-{}^{\lambda t^2}/_{2})$$

$$h(t) = \lambda t$$

Rayleigh Distribution



Gamma Distribution

- Similar to the Weibull distribution it covers a wider range of distributions

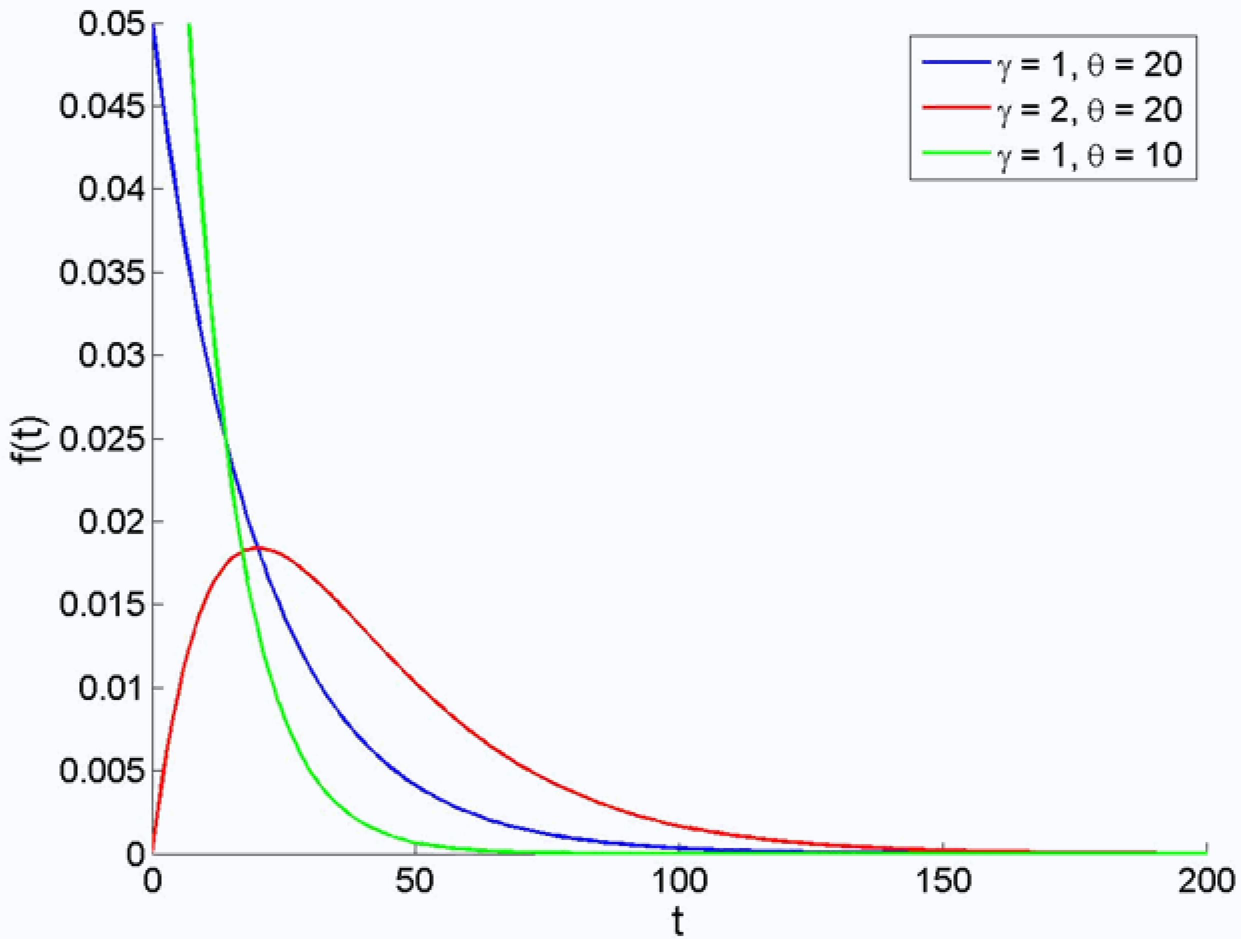
$$f(t) = \frac{t^{\gamma-1}}{\theta^{\gamma}\Gamma(\gamma)} \exp\left(-\frac{t}{\theta}\right)$$

- Where Γ is the so-called Gamma function
  - When $0 < \gamma < 1$ the failure rate decreases from infinity
  - When $\gamma = 1$ the failure rate is constant and equals $1/\theta$
  - When $\gamma > 1$ the failure rate increases to infinity
- The CDF is:

$$F(t) = \int_0^t \frac{t^{\gamma-1}}{\theta^{\gamma}\Gamma(\gamma)} \exp\left(-\frac{t}{\theta}\right) dt = I\left(\frac{t}{\theta}, \gamma\right)$$

- Where I is the incomplete Gamma function

Gamma Distribution



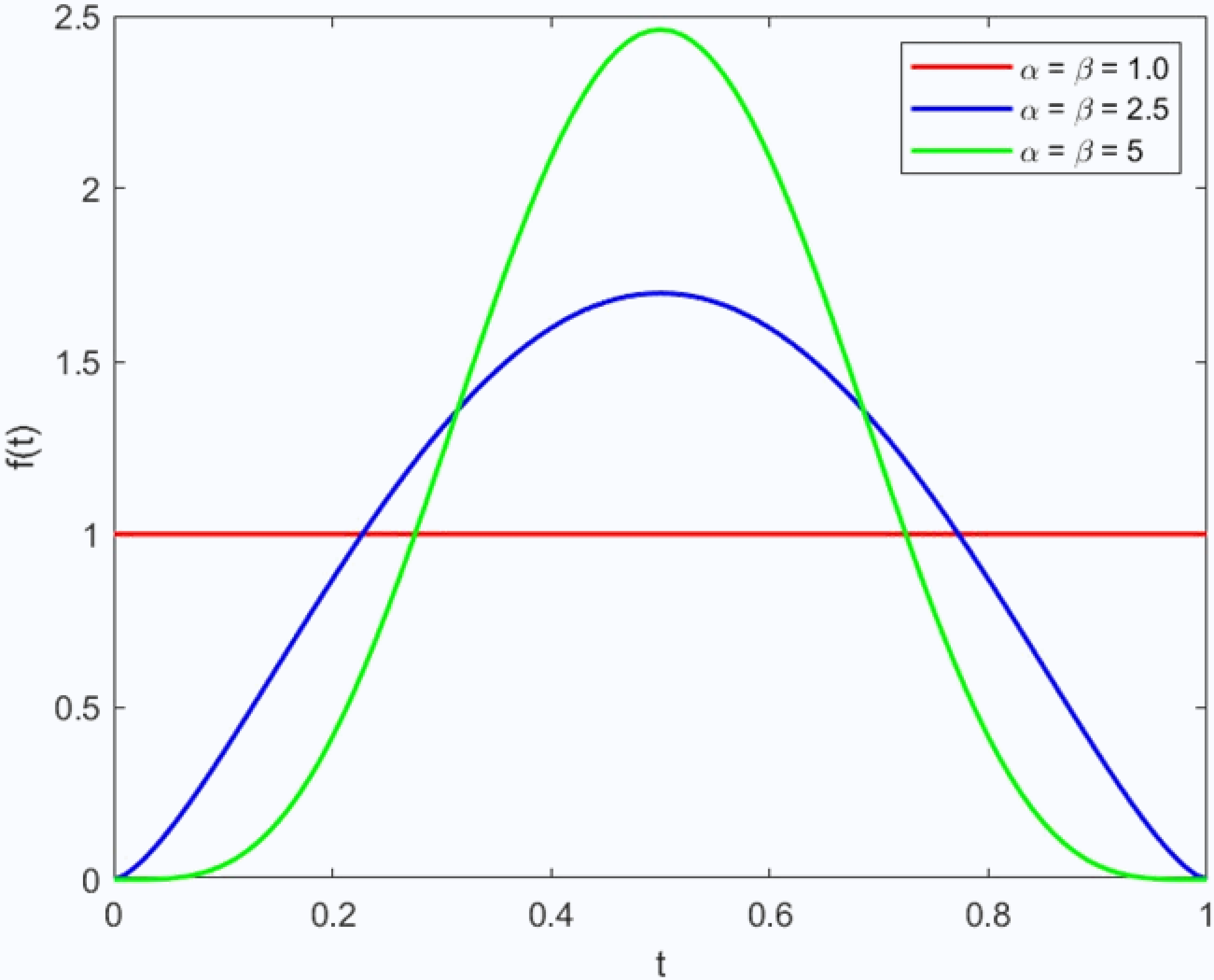
Beta Distribution

- Used to describe components where life is limited to a finite interval
- Or where events are limited to a finite interval e.g. project management

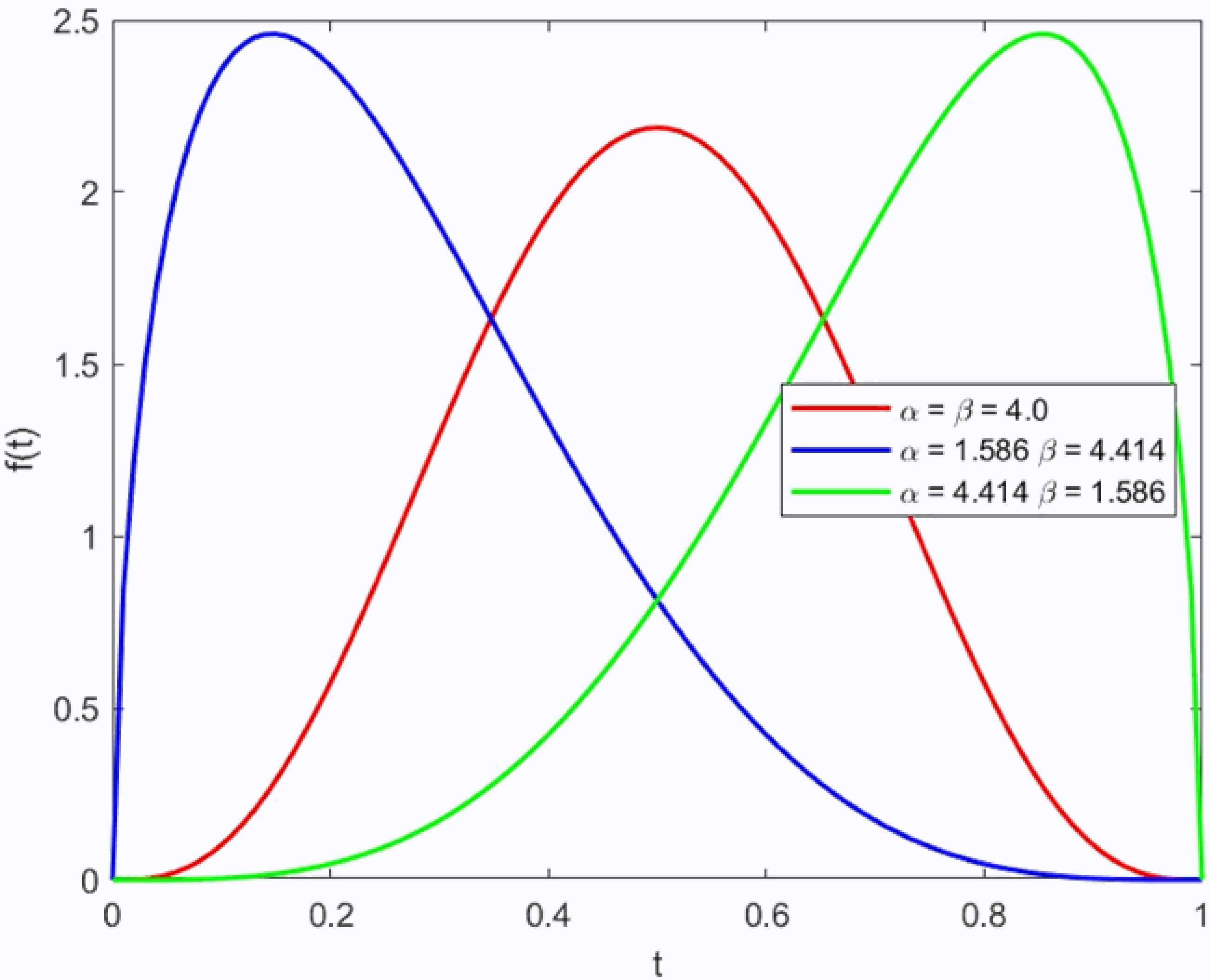
$$f(t) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha-1}(1-t)^{\beta-1} & (0 < t < 1) \\ 0 & (\text{otherwise}) \end{cases}$$

  - $\alpha > 0$
  - $\beta > 0$
- There is no closed-form solution for *F(t)* etc.

Beta Distribution



Beta Distribution



Even More Distributions

- Inverse gamma distribution
- Log-logistic distribution
- Birnbaum-Saunders distribution
- …

- Of course there is nothing to stop you from coming up with your own distribution function as long as…

$$\int_a^b f(t)\ \mathrm{d}t = 1$$

There Are Lots of Distributions….

- Don’t be this person…

Computing Continuous Functions

Excel

Distribution Name	CDF, F(t)
Uniform Distribution	N/A
Triangular Distribution	N/A
Normal Distribution	NORM.DIST(t,μ,σ,*TRUE*)
Lognormal Distribution	LOGNORM.DIST(t,μ,σ,*TRUE*)
Exponential Distribution	EXPON.DIST(t,λ,*TRUE*)
Weibull Distribution	WEIBULL.DIST(t,β,η,*TRUE*)

- Replace “TRUE” with “FALSE” for the PDFs
- Subtract γ from the data to obtain a 3 parameter Weibull

Matlab

- Use the `makedist` function to create the desired distribution using the distribution's name e.g.

  `makedist('NAME','Parameter1',Parameter_Value…)`

- The `pdf` and `cdf` functions can then be used to evaluate

Distribution Name	Name
Uniform Distribution	`'Uniform'`
Triangular Distribution	`'Triangular'`
Normal Distribution	`'Normal'`
Lognormal Distribution	`'Lognormal'`
Exponential Distribution	`'Exponential'`
Weibull Distribution	`'Weibull'`

- For more information type `help makedist`

Python

- Within the `scipy.stats` library contains a number of distribution functions with methods for evaluating the pdf, cdf and generating random numbers

Distribution Name	Name
Uniform Distribution	`scipy.stats.uniform`
Triangular Distribution	`scipy.stats.triang`
Normal Distribution	`scipy.stats.norm`
Lognormal Distribution	`scipy.stats.lognorm`
Exponential Distribution	`scipy.stats.expon`
Weibull Distribution	`scipy.stats.dweibull`

A Warning

- No matter what software you use always check the formulation is as you expect it to be
- Some common pitfalls include:
  - Different parameter names or ordering e.g. b & c for a triangular distribution can often be swapped around
  - Different parameter definitions e.g. some software expects λ and others $^{1}/_{\lambda}$ for an exponential distribution
- Be particularly careful with some of the scipy implementations, the location and scaling parameters can be confusing

Discrete Variations

Discrete Distribution Functions

- Just as there are continuous distributions there are also discrete distribution functions
- Binomial distribution
  - Two outcomes (pass/fail)

$$f(x) = \frac{n!}{x!\,(n-x)!} p^x q^{(n-x)}$$

  - We will see the binomial coefficient appear throughout the module

$$\frac{n!}{x!\,(n-x)!} \equiv \binom{n}{x}$$

  - Calculates the number of combinations of x given n

Discrete Distribution Functions

- Poisson Distribution
  - Event occurring at a constant rate

$$f(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \; (k = 0, 1, 2, \ldots)$$

  - Can approximate the Binomial distribution
- Hypergeometric distribution
  - Models the probability when there is no replacement

$$f(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

  - Consider a box of N things, M of which are defective and we draw out n of those things without replacement

University of
Southampton