SESA6085

Advanced Aerospace Engineering Management

Author: Yusaf Sultan
Lecturer: David Toal

Word Count: 2940

Contents

List of Figures

List of Tables

Definitions

$\boldsymbol{P(A)}$	Probability of a general event A occurring.	$\boldsymbol{N}$	Total number of equally likely possible outcomes in the sample space.
$\boldsymbol{n}$	Number of favorable outcomes (ways in which event A occurs)	$\boldsymbol{P(AB)}$	Probability of events $\boldsymbol{A}$ **and** $\boldsymbol{B}$ occurring.
$\boldsymbol{P(A+B)}$	Probability of events $\boldsymbol{A}$ **or** $\boldsymbol{B}$ occurring.	$\boldsymbol{P(A\|B)}$	Probability of event $\boldsymbol{A}$ given event $\boldsymbol{B}$ has already occurred
$\boldsymbol{P(A)}$	Probability of event $\boldsymbol{A}$ given event $\boldsymbol{B}$ has already occurred	$\boldsymbol{P\left(\overline{A}\right)}$	Probability of event $\boldsymbol{A}$ not occurring.
s-independent	statistically independent events	**s-dependent**	statistically dependent events
$\boldsymbol{f(t)}$	Probability Distribution Function (PDF)	$\boldsymbol{F(t)}$	Cumulative Distribution Function (CDF)
$\boldsymbol{R(t)}$	Reliability Function	$\boldsymbol{h(t)}$	Hazard Function
$\boldsymbol{\mu}$	Mean (Gaussian location parameter)	$\boldsymbol{\sigma}$	Standard Deviation (Gaussian scaling parameter)
$\boldsymbol{\lambda}$	Exponential Scaling Parameter	$\boldsymbol{\beta}$	Weibull shape parameter
$\boldsymbol{\eta}$	Weibull scaling parameter (characteristic life)	$\boldsymbol{\gamma}$	Weibull location parameter (failure free time)

1. Lecture 1

1.1. Probability Fundamentals, Rules and Notation

The most basic definition of the probability for a general event A occurring **is the ratio of the number of favorable outcomes n to the total number of equally likely possible outcomes N**, this is shown in a mathematical representation in **Eq. 1**.

$$P(A) = \frac{n}{N} \tag{1}$$

Where:
- $\boldsymbol{P(A)}$: The probability of outcome A.
- $\boldsymbol{N}$: Total number of equally likely possible outcomes in the sample space.
- $\boldsymbol{n}$: Number of favorable outcomes (ways in which event A occurs)

Note that **Eq. 1** is only for events of equal probability, for example rolling a dice. Instead if **N is the number of experiments** then **Eq. 2** applies, implying that the larger the number of experiments the closer to **Eq. 1** the probability becomes.

$$P(A) = \lim_{N \to \infty} \left(\frac{n}{N}\right) \tag{2}$$

This module uses the following notation for the probability of combined events, these are:

- $\boldsymbol{P(A)}$: Probability of event $\boldsymbol{A}$ occurring.
- $\boldsymbol{P(AB)}$: Probability of events $\boldsymbol{A}$ **and** $\boldsymbol{B}$ occurring.
- $\boldsymbol{P(A+B)}$: Probability of events $\boldsymbol{A}$ **or** $\boldsymbol{B}$ occurring.
- $\boldsymbol{P(A|B)}$: Probability of event $\boldsymbol{A}$ given event $\boldsymbol{B}$ has already occurred.
- $\boldsymbol{P\left(\overline{A}\right)}$: Probability of event $\boldsymbol{A}$ **not** occurring (note that $P(A) = 1 - P\left(\overline{A}\right)$).

1.2. Statical Independence

If two events are **statistically independent** (s-independent) from one another (meaning that the probability of one event occurring is completely separate from another event happening or not happening), then **Eq. 3** is true.

$$\left.\begin{aligned} P(A|B) = P\left(A|\overline{B}\right) = P(A) \\ P(B|A) = P\left(B|\overline{A}\right) = P(B) \end{aligned}\right\} \text{s-independent} \tag{3}$$

Furthermore, the joint probability of two s-independent events can be represented in the forms shown in **Eq. 4** with the further expressions derived from subbing in **Eq. 3**, **Eq. 4** is also known as the **product or series rule**.

$$P(AB) = P(A)P(B)\} \text{ s-independent} \tag{4}$$

1.3. Statistical Dependence

If two events are instead **statistically dependent** (s-dependent) from one another (the probability of on event happening or not happening **does** have an effect of the probability of another event), then the adjoint probability of these two events is shown in **Eq. 5**

$$\left.\begin{aligned} P(AB) &= P(A)P(B|A) \\ P(AB) &= P(A|B)P(B) \end{aligned}\right\} \text{s-dependent} \tag{5.1}$$

$$P(B|A) = \left.\frac{P(AB)}{P(A)}\right\} \text{s-dependent and } P(A) \neq 0 \tag{5.2}$$

1.4. Probability Fundamentals, Rules and Notation Cont.

Generally speaking the probability of one event **or** another event occurring , whether they are s-dependant or s-independent is given by equation **Eq. 6**.

$$P(A+B) = P(A) + P(B) - P(AB) \tag{6.1}$$

$$P(A+B) = P(A) + P(B) - P(A)P(B)\} \text{ s-independent} \tag{6.2}$$

Note that the $P(AB)$ in **Eq. 6** must be subtracted as it is counted twice in the first two terms.

1.5. Mutual Exclusivity

Two events can be said top be mutually exclusive if they **cannot occur at the same time as one another**. This means that the adjoint probability and or probability can be written in the form shown in **Eq. 7**.

$$\left.\begin{aligned} &P(AB) = 0 \\ &P(A+B) = P(A) + P(B) \end{aligned}\right\} \text{If A and B are Mutually Exclusive} \tag{7}$$

If instead there are **multiple mutually exclusive events** which together yield the probability of another event, then then probability of that event can be written in the form given by **Eq. 8**.

$$P(A) = \textstyle\sum_i P(AB_i) = \sum_i P(A|B_i)P(B_i)\}\text{If A and all Bs are Mutually Exclusive} \tag{8}$$

1.6. Sequence Diagrams

Sequence diagrams act as an easy way of visualizing complex interactions and can be used to calculate overall probabilities, an example of a sequence diagram is shown in **Figure 1**.

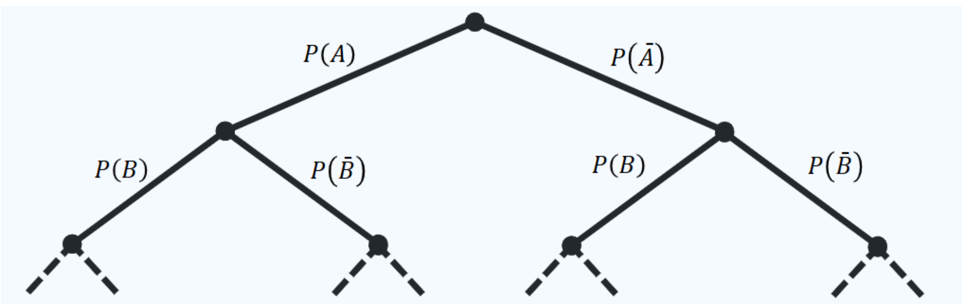

Figure 1: Example of a sequence diagram.

Probabilities down a leg are **and** probabilities and are therefore multiplied given that they are **statistically independent**. **Or** probabilities can be calculated by adding together subsequent probabilities.

1.7. Baye's Theorem

By rearranging **Eq. 5** a simple form of **Baye's theorem** which is shown in **Eq. 9**.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \qquad \text{Given } P(B) \neq 0 \tag{9}$$

Eq. 9 can be further developed by substituting in **Eq. 8** which yields the **generalized Baye's theorem** shown **Eq. 10**.

$$P\left(A_j \mid B\right) = \frac{P(B|A_i)P(A_i)}{\sum_j P\left(B|A_j\right)P\left(A_j\right)} \tag{10}$$

Note that in **Eq. 10** A_j is the jth event effecting the event B. If the probability of event B depends on the probability of event A both happening and not happening then **Eq. 10** simplifies down to a form called the **binary partition** form, shown in **Eq. 11**.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P\left(B|\overline{A}\right)P\left(\overline{A}\right)} \tag{11}$$

2. Lecture 2

2.1. Frequency Histograms

A **frequency histogram** is a type of bar chart which is used to represent the distribution of data, on the **x axis** are bins of data and the **y axis** represents the frequency that occurs within that bin, an example of a frequency h istogram is shown in **Figure 2**.

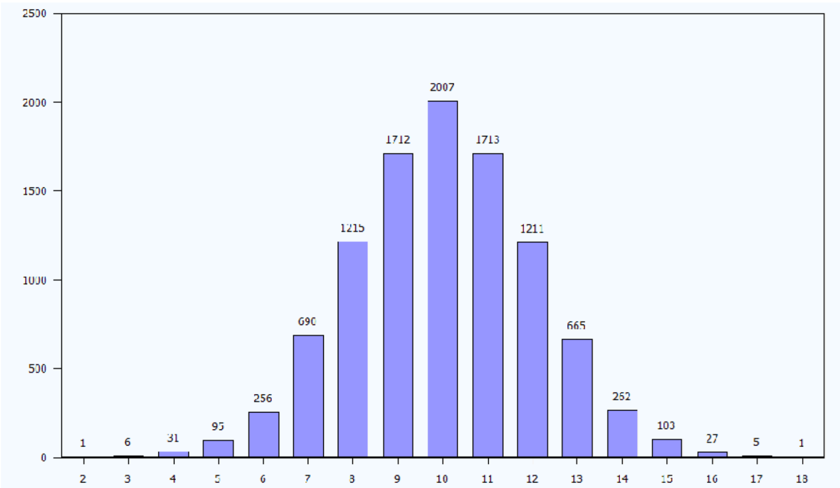

Figure 2: Example of a frequency histogram.

As we increase the number os samples taken, effectively decreasing the width of each bin, then the data will approach a smooth curve.

2.2. Definition of a PDF

As the frequency histogram bin width approaches an infinitesimal width, the histogram approaches a continuous curve known as the **Probability Density Function** (PDF). A PDF has one criteria in that the **area under the curve must be equal to 1**, the mathematical definition of a PDF is shown in **Eq. 12**.

$$\int_{-\infty}^{\infty} f(t)dt = 1 \tag{12}$$

PDFs can be used to find the probability that a certain value t is that value. In terms of reliability engineering its the **probability that a component fails** at the time t.

2.3. Definition of a CDF

A **Cumulative Distribution Function** (CDF) yields the probability that a given value will fall between the limits of $-\infty$ and t_1, its mathematical definition is shown in **Eq. 13**.

$$F(t) = \int_{-\infty}^{t_1} f(t)dt \tag{13}$$

2.4. Reliability Function

Reliability is the probability that a component will survive from a time $t = 0$ to a time $t = t_1$ and its mathematical definition is shown in **Eq. 14**, with a graph depicting the reliability function shown in **Figure 3**.

$$R(t) = 1 - F(t) = 1 - \int_{-\infty}^{t_1} f(t)dx \equiv \int_{x}^{\infty} f(t)dx \tag{14}$$

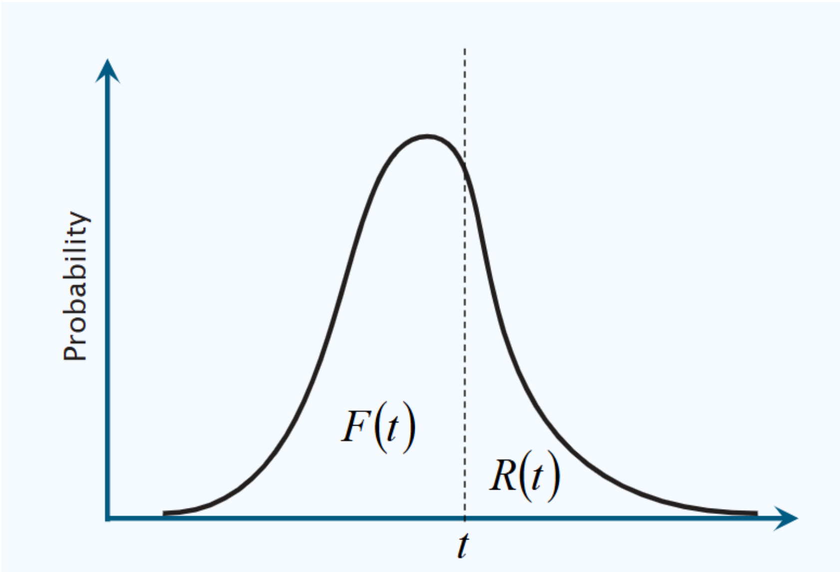

Figure 3: Graph of the reliability function and CDF on a PDF.

2.5. Hazard Function

Also known as the hazard rate the **hazard function** gives the probability of failure at a time t, given that there has not already been a failure. The mathematical definition for the hazard function is shown in **Eq. 15**.

$$h(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - R(t)} \tag{15}$$

The hazard function can be though of as a measure of the probability of eminent failure at t or the proneness of failure after t. Note that there does exist a cumulative hazard function which is not assessed as well as methods to rearrange between all of these functions.

2.6. Continuous Distribution

A continuous PDF is a smooth curve representing how the probability varies with an area under the curve being equal to one. Effectively there is an infinite number of probability distributions as long as they satisfy the conditions set above, some of the most common are shown below.

2.6.1. Uniform Distribution Function

The most simplest distribution function assumes, that the distribution is zero and then one fixed value for a set time period. The PDF and CDF are defined in **Eq. 16**.

$$f(t) = \begin{cases} \frac{1}{b-a} & t \in [a,b] \\ 0 & \text{Otherwise} \end{cases} \tag{16.1}$$

$$F(t) = \begin{cases} 0 & t < a \\ \frac{t-a}{b-a} & t \in [a,b] \\ 1 & t > b \end{cases} \tag{16.2}$$

Where:
- ***a***: Start of non-zero probability.
- ***b***: End of non-zero probability.

The PDF and CDF for a uniform probability distribution are shown in graphically **Figure 4**.

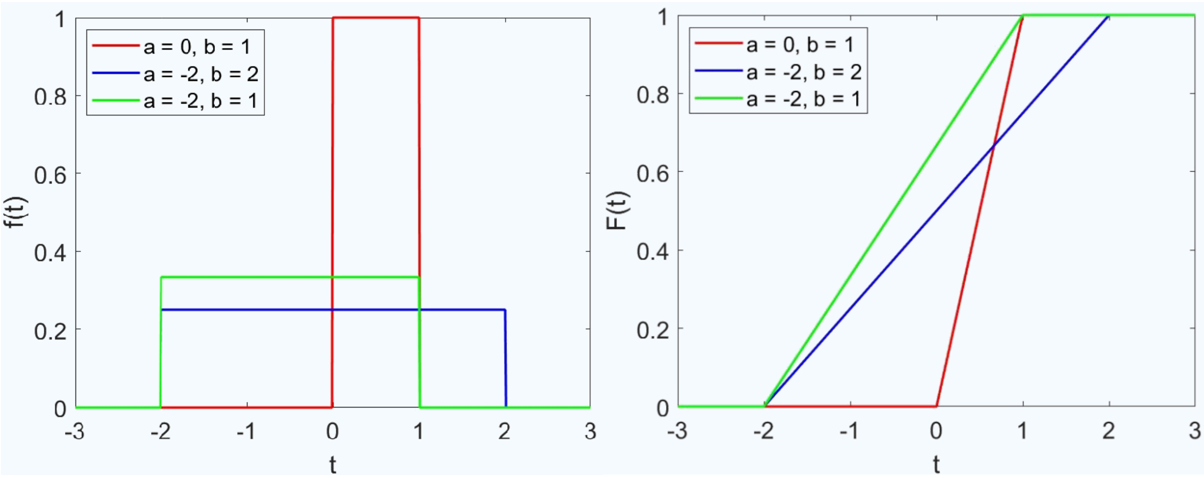

Figure 4: Plots of uniform PDFs [left] and plots of uniform CDFs [right].

2.6.2. Triangular Distribution Functions

Triangular distribution functions are slightly more complex than the aforementioned uniform distribution functions. Their PDF and CDF are shown in **Eq. 17**.

$$f(t) = \begin{cases} \frac{2(t-a)}{(c-a)(b-a)} & a \leq t \leq b \\ \frac{2(c-t)}{(c-a)(c-b)} & b \leq t \leq c \\ 0 & t < a, t > c \end{cases} \tag{17.1}$$

$$F(t) = \begin{cases} 0 & t < a \\ \frac{(t-a)^2}{(b-a)*(c-a)} & a \leq t \leq b \\ 1 - \frac{(c-t)^2}{(c-a)*(c-b)} & b \leq t \leq c \\ 1 & t > c \end{cases} \tag{17.2}$$

Where:
- ***a***: Start of non-zero probability.
- ***b***: Probability peak.
- ***c***: End of non-zero probability.

The PDF and CDF for a triangular probability distribution are shown graphically in **Figure 5**.

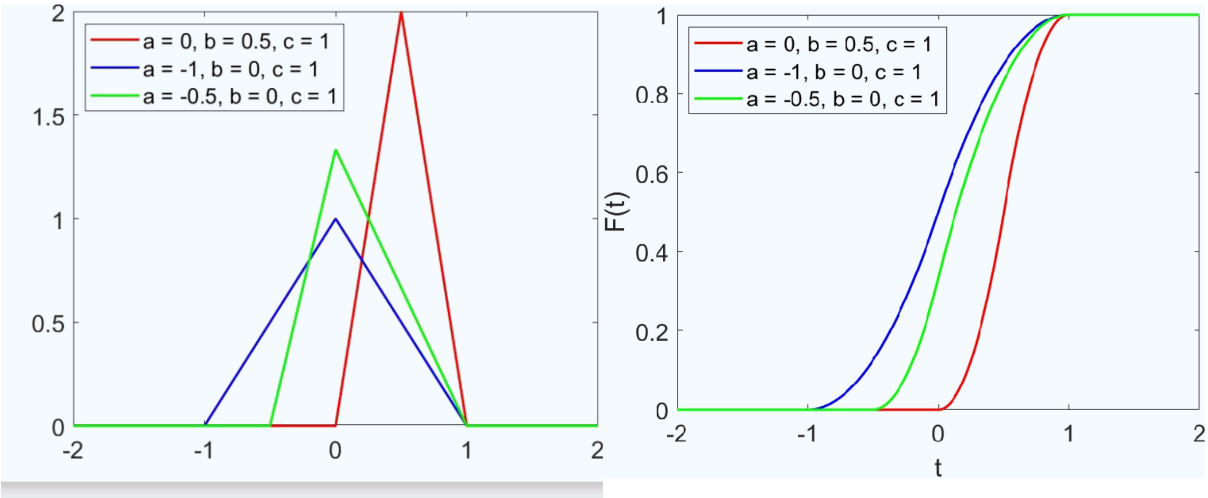

Figure 5: Plots of triangular PDFs [left] and plots of triangular CDFs [right].

2.6.3. Gaussian Distribution

Also known as the the **Normal Distribution** is the most commonly used probability distribution function. The PDF is shown in **Eq. 18** (**Note no close form CDF exists**).

$$f(t) = \frac{1}{\sigma(2\pi)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) \tag{18}$$

Where:

- $\boldsymbol{\mu}$: Mean (Location parameter).
- $\boldsymbol{\sigma}$: Standard Deviation (Scaling parameter).

The PDF and CDF for a Gaussian probability distribution are shown graphically in **Figure 6**.

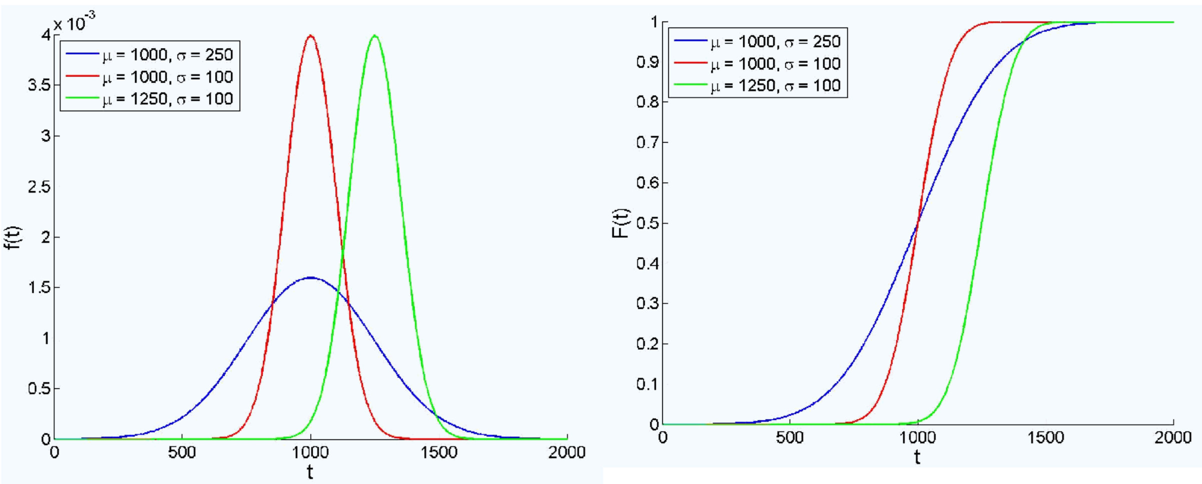

Figure 6: Plots of Gaussian PDFs [left] and plots of Gaussian CDFs [right].

The Gaussian distribution whilst being **symmetrical** also exhibits these key properties:

- **68.26%** of data is within 1 standard deviation of the mean$(\boldsymbol{\sigma})$.
- **95.44%** of data is within 2 standard deviations of the mean $(\boldsymbol{2\sigma})$.
- **99.74%** of data is within 3 standard deviations of the mean $(\boldsymbol{3\sigma})$.

2.6.4. Log Normal Distribution

A more versatile version of the Gaussian distribution that is better suited at modelling reliability data. The PDf is shown in **Eq. 19** (**Note no CDF exists**).

$$f(t) = \begin{cases} \frac{1}{t\sigma(2\pi)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{19}$$

Where:
- $\boldsymbol{\mu}$: Mean (Location parameter).
- $\boldsymbol{\sigma}$: Standard Deviation (Scaling parameter).

The PDF and CDF for a log normal probability distribution are shown graphically in **Figure 7**.

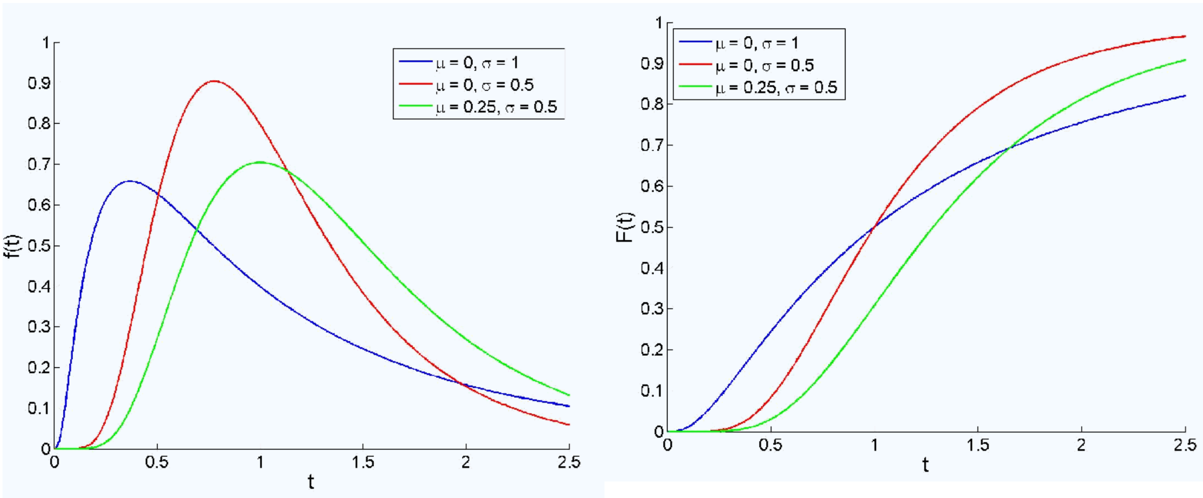

Figure 7: Plots of log normal PDFs [left] and plots of log normal CDFs [right].

2.6.5. Exponential Distribution

These distributions feature a **constant hazard rate** which is useful to model some processes. The PDf and CDF are shown mathematically in **Eq. 20**.

$$f(t) = \begin{cases} \lambda \exp(-\lambda t) & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{20.1}$$

$$F(t) = \begin{cases} 0 & t < 0 \\ 1 - \exp(-\lambda t) & t \geq 0 \end{cases} \tag{20.2}$$

Where:
- $\boldsymbol{\lambda}$: Scaling parameter (Also is the constant hazard rate)

It is important to note that $1/\lambda$ is the mean time to failure (MTTF). The PDF and CDF for an exponential distribution are shown graphically in **Figure 8**.

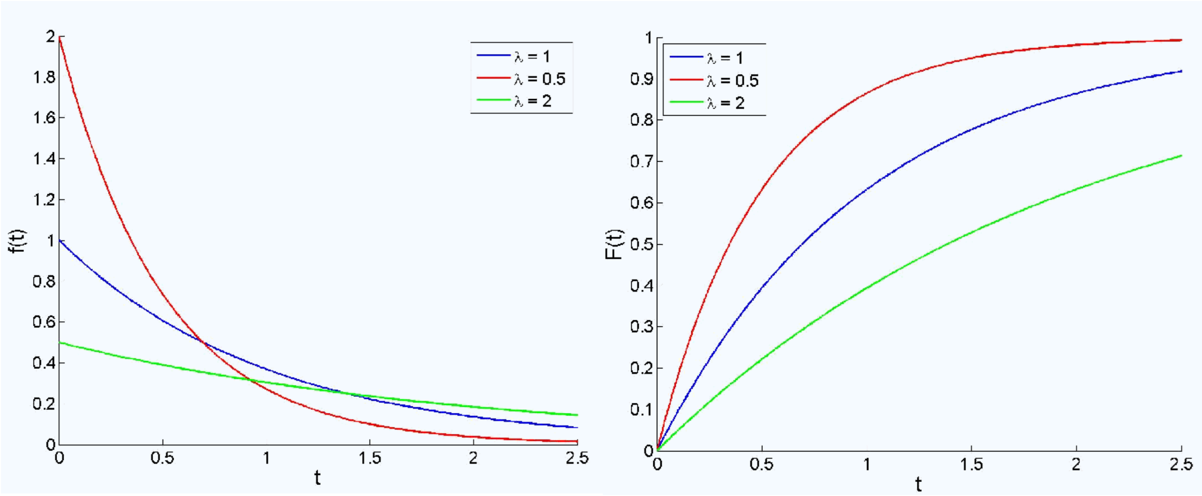

Figure 8: Plots of exponential PDFs [left] and plots of exponential CDFs [right].

2.6.6. Weibull Distribution

Is one of the most extensible and useful distributions out there, and can be used to model a lot of different distributions. The PDF, CDF and hazard rate are shown in **Eq. 21**.

$$f(t) = \begin{cases} \frac{\beta}{\eta^\beta} t^{\beta-1} \exp\left(-\left(\frac{t}{\eta}\right)^\beta\right) & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{21.1}$$

$$F(t) = 1 - \exp\left(-\left(\frac{t}{\eta}\right)^\beta\right) \qquad h(t) = \frac{\beta}{\eta^\beta} t^{\beta-1} \tag{21.2}$$

Where:

- $\boldsymbol{\beta}$: Shape parameter
- $\boldsymbol{\eta}$: Scaling parameter (characteristic life)

η is also the point at which 63.2% of the population have failed. Weibull distributions are so versatile as the β parameter changes the shape into different distributions:

- $\beta = 1$: Constant hazard function (exponential dist)
- $\beta < 1$: Decreasing hazard rate
- $\beta > 1$: Increasing hazard rate
- $\beta = 3.5$: Normal distribution

The PDF and CDF for various Weibull distributions are shown graphically in **Figure 9**.

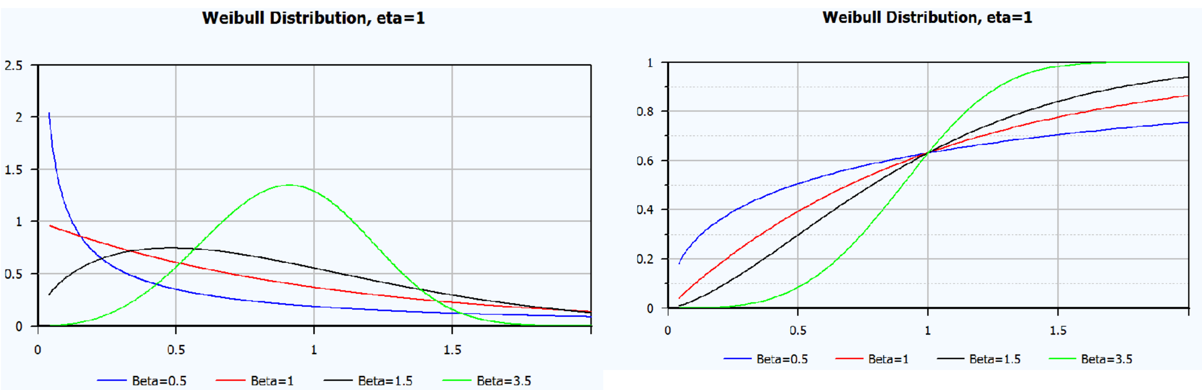

Figure 9: Plots of Weibull PDFs [left] and plots of Weibull CDFs [right] for various βs.

2.6.7. Three Parameter Weibull Distribution

Introduces a new parameter γ which is used to switch on the probability, its useful if the failures only start after a set time. The PDF, CDF and hazard rate are shown in **Eq. 22**.

$$f(t) = \begin{cases} \frac{\beta}{\eta^\beta}(t-\gamma)^{\beta-1}\exp\left(-\left(\frac{t-\gamma}{\eta}\right)^\beta\right) & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{22.1}$$

$$F(t) = 1 - \exp\left(-\left(\frac{t-\gamma}{\eta}\right)^\beta\right) \qquad h(t) = \frac{\beta}{\eta^\beta}(t-\gamma)^{\beta-1} \tag{22.2}$$

Where:

- $\boldsymbol{\beta}$: Shape parameter
- $\boldsymbol{\eta}$: Scaling parameter (characteristic life)
- $\boldsymbol{\gamma}$: Location parameter (failure free time)

2.6.8. Other Distribution Functions (Non-Examinable)

Like was stated previously, there are an infin ite number of PDFs as the only criteria is for the area under the curve to sum to 1. Some other common distributions and their purposes are mentioned below:

- **Rayleigh Distribution**: Similar to exponential but with a linearly increasing hazard rate.
- **Gamma Distribution**: Similar to Weibull in that it can model a wide number of distributions by varying the parameters.
- **Beta Distribution**: A complex distribution which uses multiple gamma distributions to ensure that the life is limited to a set interval.
- **Inverse Gamma Distribution**
- **Log-logistic Distribution**
- **Birnbaum-Saunders Distribution**

2.7. Discrete Distributions

Whereas continuous distributions can model the probability over time, discrete distributions model the probability per an n number of events, some common discrete distributions are shown below.

2.7.1. Binomial Distribution

Used where the outcome of each discrete event is either pass or fail, the PDF for a binomial distribution function is defined by **Eq. 23**.

$$f(x) = \binom{n}{x} p^x q^{n-x} \tag{23.1}$$

$$\frac{n!}{x!(n-x)!} \equiv \binom{n}{x} \tag{23.2}$$

Where:

- $\boldsymbol{x}$: The number of passes

- $\boldsymbol{n}$: Total number of trails
- $\boldsymbol{p}$: Probability of success
- $\boldsymbol{q}$: Probability of failure
- $\binom{\boldsymbol{n}}{\boldsymbol{x}}$: Binomial coefficient

Note that the **binomial coefficient** is a parameter that will appear often and is read as `"n choose x"`. Usefully, it also represents the **number of possible combinations of $\boldsymbol{n}$ from $\boldsymbol{x}$**.

2.7.2. Other Discrete Distributions

Some other commonly used discrete distributions are:

- **Poisson's Distribution**: Represents an event occurring at a constant rate and can approximate the binomial distribution.
- **Hypergeometric Distribution** Models the probability if there are no replacements.

3. Lecture 3

3.1. Parameter Estimation

This is the process of estimating the key parameters within a given PDF (μ, σ for Gaussian λ for exponential etc) by using the current set of data. for all parameter estimation techniques, the following must be true:

- **Unbiased:**: The estimator should not consistently under or overestimate the true value of the parameter.
- **Consistent**: The estimator should converge to the true value as the sample size increases.
- **Efficient**: The estimator should be consistent with a standard deviation in that estimate smaller than any other estimator for the same population.
- **Sufficient**: The estimator should use all of the information about the parameter that the data sample possesses.

All parameter estimation models depend on the quality of the data used. The three most common methods are the **method of moments**, the **maximum likelihood method** and the **least squares method**.

3.2. Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) is a parameter estimation processed based on the most likely estimate of the data, and roughly consists of the following steps:

1. Define the formulation of the PDF.
2. Define the parameter to estimate.
3. Define the natural log of the PDF
4. Define the log likelihood function.
5. Define it's derivative(s).
6. Equate the derivative(s) to zero and solve for the parameter(s).

3.3. How to Do MLE

For MLE we assume we already know which PDF form we are using (Gaussian, exponential, Weibull etc). The PDF as well as the parameters are written in the form shown in **Eq. 24**, known as the **density function**.

$$f(t; \theta_1, \theta_2, ..., \theta_m) \equiv f(t : \theta) \tag{24}$$

Where:

- $\boldsymbol{\theta}$: PDF parameters $(\theta_1, \theta_2, ..., \theta_m)$

For example, θ_1 and θ_2 for a Gaussian distribution would be μ and σ. Following on from **Eq. 24**, the likelihood is mathematically defined in **Eq. 25**.

$$L(\theta) = \prod_{i=1}^{n} f(t_i : \theta) \tag{25}$$

Where $L(\theta)$ is the **likelihood** for a given PDF parameter. Effectively, the likelihood itself will hel to fit the correct PDF parameters to the given form of the data. The definition of

$L(\theta)$ has its issues however, mainly that the product of the probabilities reduce very quickly and so instead the log-likelihood is used, defined in **Eq. 26**.

$$l(\theta) = \log L(\theta) \tag{26}$$

Where $l(\theta)$ is the log-likelihood and gets around the issue of very small probabilities. The maximum likelihood estimators of the PDF parameters are written in the notation shown in **Eq. 27**.

$$\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m \tag{27}$$

The values of **Eq. 27** can be obtained using an optimization algorithm, or can be obtained via analytical equations using the criteria shown in **Eq. 28**.

$$\frac{\partial l}{\partial \theta_j} = 0 \qquad j = 1, 2, ..., m \tag{28}$$

3.3.1. MLE for a Gaussian Distribution

Recall that the form that the PDF takes for a Gaussian distribution is given in **Eq. 18**. Currently, μ & σ are unknown for a given set of data, MLE can be used to obtain $\hat{\mu}$ & $\hat{\sigma}$. For a given observation t_i, the log likelihood of the Guassian PDF is given in **Eq. 29**.

$$\ln(f(t_i)) = -\ln(\sigma) - \frac{1}{2}\ln(2\pi) - \frac{1}{2}\left(\frac{t_i - \mu}{\sigma}\right)^2 \tag{29}$$

Moving on from **Eq. 29**, the MLE parameters are defined and the equation is re-written in **Eq. 30**.

$$l(\mu, \sigma) = \ln(L(\mu, \sigma)) = \ln\left(\prod_{i=1}^{n} f(t_i : \mu, \sigma)\right) = \sum_{i=1}^{n} \ln(f(t_i : \mu, \sigma)) \tag{30}$$

Note that in **Eq. 30**, using the log likelihood here is advantageous as it converts the products into much more manageable sums. Applying the sums in **Eq. 30** to **Eq. 29** yields **Eq. 31**.

$$\ln(f(t_i)) = -n\ln(\sigma) - \frac{n}{2}\ln(2\pi) - \sum_{i=1}^{n}\frac{1}{2}\left(\frac{t_i - \mu}{\sigma}\right)^2 \tag{31}$$

Taking the partial derivative of the log likelihood function with respect to the PDF parameters yields **Eq. 32**, note that when these equations are set to zero and rearranged, the MLE parameters are obtained.

$$\frac{\partial l}{\partial \mu} = -\frac{1}{\sigma^2}\sum_{i=1}^{n}(\mu - t_i)^2 \quad \rightarrow \quad \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} t_i \tag{32.1}$$

$$\frac{\partial l}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(t_i - \mu)^2 \quad \rightarrow \quad \hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(t_i - \mu)^2} \tag{32.2}$$

Note that the MLE params for a Guassian distribution are the mean and standard deviation.

3.3.2. MLE for an Exponential Distribution

Recall that the PDF for a exponential distribution is defined in **Eq. 20**. Currently λ is unknown for a set of data, MLE can find $\hat{\lambda}$. Taking logs of the PDF, for a given observation t_i, the log likelihood is given in **Eq. 33**.

$$\ln(f(t_i)) = \ln(\lambda \exp(-\lambda t_i)) = \ln(\lambda) - t_i \tag{33}$$

Taking the product of the probability of all of the events yields **Eq. 34**. This equation is then differentiated with respect to λ, set to zero and rearranged for the MLE parameter.

$$l(\lambda) = n \exp(\lambda) - \lambda \sum_{i=1}^{n} t_i \tag{34.1}$$

$$\frac{\partial l}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} t_i \quad \rightarrow \quad \hat{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i} \tag{34.2}$$

Note that $\hat{\lambda}$ is the $1/\mu$ and is also the reaction that $1/\lambda$ is the mean time to failure (MTTF).

3.4. Parameter Confidence

Now that the MLE parameters have been defined, the confidence in the value of these parameters must also be quantified. It is clear that **as the data set size increases**, the **confidence in the calculated MLE parameter should increase**. To quantify parameter confidence, a Fisher information matrix is used, shown in **Eq. 35**.

$$\text{The } \ I_{ij} \text{ component} \rightarrow I_{ij} = E\left[-\frac{\partial^2 l(t:\theta)}{\partial\theta_i \partial\theta_j}\right] \tag{35}$$

Calculating and constructing the Fisher information matrix and then inverting it yields the **covariance** matrix which allows for the calculation of the variance of a specific MLE parameter and the covariance between two MLE parameters, this is shown in

$$I^{-1} = \begin{bmatrix} \mathrm{Var}(\theta_1) & \mathrm{Cov}(\theta_1,\theta_2) & ... & \mathrm{Cov}(\theta_1,\theta_k) \\ \mathrm{Cov}(\theta_2,\theta_1) & \mathrm{Var}(\theta_2) & ... & \mathrm{Cov}(\theta_2,\theta_k) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(\theta_k,\theta_1) & \mathrm{Cov}(\theta_k,\theta_2) & ... & \mathrm{Var}(\theta_k) \end{bmatrix}_{k\times k} \tag{36}$$

3.4.1. Parameter Confidence for Guassian Distribution

To form the Fisher information matrix for a Guassian distribution, the second partial derivative as well as the mixed partial derivative must be calculated, this is shown in **Eq. 37**. Note that E in the Fisher matrix **represents the expected value**, allowing the moments of a normal to be applied

$$\text{Moments of a Guassian Dist:} \begin{cases} E[t_i - \mu] = 0 \\ E[\mu - t_i] = 0 \\ E\left[(t_i - \mu)^2\right] = \sigma^2 \\ E\left[(\mu - t_i)^2\right] = -\sigma^2 \end{cases} \tag{37.1}$$

$$E\left[-\frac{\partial^2 l}{\partial \mu^2}\right] = E\left[-\frac{\partial}{\partial \mu}\left(-\frac{1}{\sigma^2}\sum_{i=1}^{n}(\mu - t_i)\right)\right] = \frac{\boldsymbol{n}}{\boldsymbol{\hat{\sigma}}} \tag{37.2}$$

$$E\left[-\frac{\partial^2 l}{\partial \sigma^2}\right] = E\left[-\frac{\partial}{\partial \sigma}\left(-\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(t_i - \mu)^2\right)\right] = E\left[-\frac{n}{\sigma^2} + \frac{3}{\sigma^4}\sum_{i=1}^{n}(t_i - \mu)^2\right] \tag{37.3}$$

$$= -\frac{n}{\sigma^2} + \frac{3}{\sigma^4}(\sigma^2) = \frac{\boldsymbol{2n}}{\boldsymbol{\hat{\sigma}^2}} \tag{37.4}$$

$$E\left[-\frac{\partial^2 l}{\partial \mu \partial \sigma}\right] = \left[-\frac{\partial}{\partial \sigma}\left(\frac{\partial}{\partial \mu}\right)\right] = E\left[\frac{2}{\sigma^3}\sum_{i=1}^{n}(\mu - t_i)\right] = 0 \tag{37.5}$$

Now the Fisher information matrix can be constructed, noting that both of the covariance terms are zero, this is shown in **Eq. 38** alongside the inverted matrix.

$$I_{2\times 2} = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{2n}{\sigma^2} \end{bmatrix}_{2\times 2} \quad \rightarrow \quad I_{2\times 2}^{-1} = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{\sigma^2}{2n} \end{bmatrix}_{2\times 2} \tag{38.1}$$

$$\therefore \mathrm{Var}(\hat{\mu}) = \frac{\hat{\sigma}^2}{n} \quad \mathrm{Var}(\hat{\sigma}) = \frac{\hat{\sigma}^2}{2n} \tag{38.2}$$

This means that each of the MLE parameters themselves have a normal distribution associated with themselves where for $\hat{\mu} : \mu = \hat{\mu}, \sigma = \sqrt{\mathrm{Var}(\hat{\mu})}$ and for $\hat{\sigma} : \mu = \hat{\sigma}, \sigma = \sqrt{\mathrm{Var}(\hat{\sigma})}$. The confidence in a given MLE parameter is therefore written in the form show in **Eq. 39**.

$$P\left(\theta_l \leq \hat{\theta}, \leq \theta_u\right) = \gamma \tag{39}$$

Where θ_l & θ_u are the upper and lower bounds of the MLE parameter and γ is the confidence level. For example if $\gamma = 0.95$, then 95% of the time, θ is within the upper and lower bounds. The upper and lower bounds are calculated using the formula shown in.

$$\theta_l = \hat{\theta} - Z_{\alpha/2}\sqrt{\mathrm{Var}\left(\hat{\theta}\right)} \tag{40.1}$$

$$\theta_u = \hat{\theta} + Z_{\alpha/2}\sqrt{\mathrm{Var}\left(\hat{\theta}\right)} \tag{40.2}$$

Where $\alpha = 1 - \gamma$ and Z is the the standard normal statistic (negative of the inverse CDF of a normal distribution with mean of 0 and a standard deviation of 1 for the probability $\alpha/2$).

4. Lecture 4

4.1. Multivariate Models

In many scenarios, more than one factor will effect the failure and therefore the reliability and most of the time, these factors are not independent from one another. For univariate data, the CDF can be used to calculate the probability in the manner shown in **Eq. 41**.

$$P(a < x \leq b) = F(b) - F(a) \tag{41}$$

For multivariate models, multiple occurrence are observed at the same time. The probability two variables fall between two bounds is shown in **Eq. 42**.

$$P(a_1 < x < b_1, a_2 < y < b_2) = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) \tag{42}$$

4.1.1. CDF of a Multivariate Model

In a similar method to univariate model, the CDF of a given multivariate PDF is the integral between minus and positive infinity of that PDF with respect to each variable, this is mathematically written in **Eq. 43**

$$f(x_1, x_2, ..., x_n) \tag{43.1}$$

$$F(a_1, a_2, ..., a_n) = \int_{-\infty}^{a_n} ... \int_{-\infty}^{a_2} \int_{-\infty}^{a_1} f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n \tag{43.2}$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n = 1 \tag{43.3}$$

4.2. Multivariate Normal Distribution

A multivariate normal distribution has a different vectorized form from the equation shown in **Eq. 18**, the multivariate form is shown in **Eq. 44**.

$$f(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^p \ |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})\right) \tag{44}$$

Where:

- $\boldsymbol{x}$: Vector of variables.
- $\boldsymbol{p}$: Number of variables
- $\boldsymbol{\Sigma}$: Covariance matrix between variables
- $\boldsymbol{\mu}$: Vector of means for each variable

Note that when **Eq. 44** has $p = 1$, the equation reduces to the univariate case. Some example plots of bivariate normal distributions are shown in **Figure 10**.

4.2.0.1. Fitting Data to Multivariate Normal Distributions

The MLE process can be applied to obtain the MLE parameters for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, these are both shown in **Eq. 45** note that $\hat{\boldsymbol{\Sigma}}_{ij}$ is the same as the sample covariance matrix.

$$\hat{\boldsymbol{\mu}} = \left[\frac{1}{n}\sum_{i=1}^{n} x_{1_i}, \frac{1}{n}\sum_{i=1}^{n} x_{2_i}, ..., \frac{1}{n}\sum_{i=1}^{n} x_{k_i}\right] \quad \hat{\boldsymbol{\Sigma}}_{ij} = (\boldsymbol{x_i} - \mu_i)^T (\boldsymbol{x_j} - \mu_j)/n \tag{45}$$

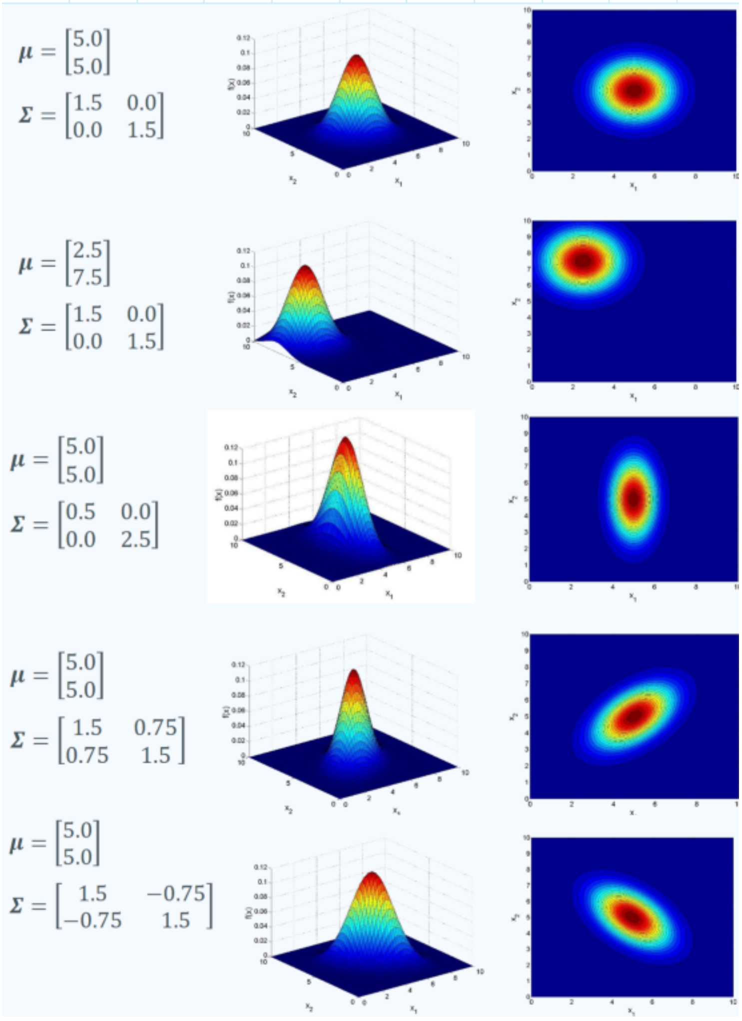

Figure 10: Various bivariate normal distributions with variations in the $\boldsymbol{\mu}$ vector and $\boldsymbol{\Sigma}$ matrix

4.3. Joint Distribution Functions

Joint distributions can be created for **two variables that are independent from one another** given that they satisfy a few conditions, these conditions are shown in **Eq. 46**

$$F(-\infty, -\infty, ..., -\infty) = 0 \qquad F(\infty, \infty, ..., \infty) = 0 \tag{46.1}$$

$$\text{If } a < b \text{ and } c < d \text{ then } F(a, c) < F(b, d) \tag{46.2}$$

If the conditions in **Eq. 46** are satisfied, then the n number of joinable distributions, the PDF and CDF can be written in the form shown in **Eq. 47**.

$$F(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} F_i(x) \tag{47.1}$$

$$f(x_1, x_2, ..., x_n) = \frac{\partial^n f}{\partial x_1 \partial x_2 ... \partial x_{\emptyset}} \tag{47.2}$$

4.3.1. Bivariate Exponential Distribution

Assuming that two distributions are statistically independent from one another, they can be joined in the manner defined in the previous section. The CDF for the bivariate exponential distribution is shown in **Eq. 48**.

$$F(x) = 1 - \exp(-\lambda x) \tag{48.1}$$

$$F(x_1, x_2) = F_1(x)F_2(x) = \begin{cases} (1 - \exp(-\lambda_1 x_1))(1 - \exp(-\lambda_2 x_2)) & x_{1,2} \geq 0 \\ 0 & x_{1,2} < 0 \end{cases} \tag{48.2}$$

The resulting PDF can be calculating by applying **Eq. 47** to **Eq. 48**, this yields **Eq. 49**.

$$f(x_1, x_2) = \frac{\partial^2 F}{\partial x_1 \partial x_2} \tag{49.1}$$

$$f(x_1, x_2) = \begin{cases} \lambda_1 \lambda_2 \exp(-\lambda_1 x_1 - \lambda_2 x_2) & x_{1,2} \geq 0 \\ 0 & x_{1,2} < 0 \end{cases} \tag{49.2}$$

An example bivariate exponential distribution with $\lambda_1 = 5, \lambda_2 = 2.5$ is shown in **Figure 11**

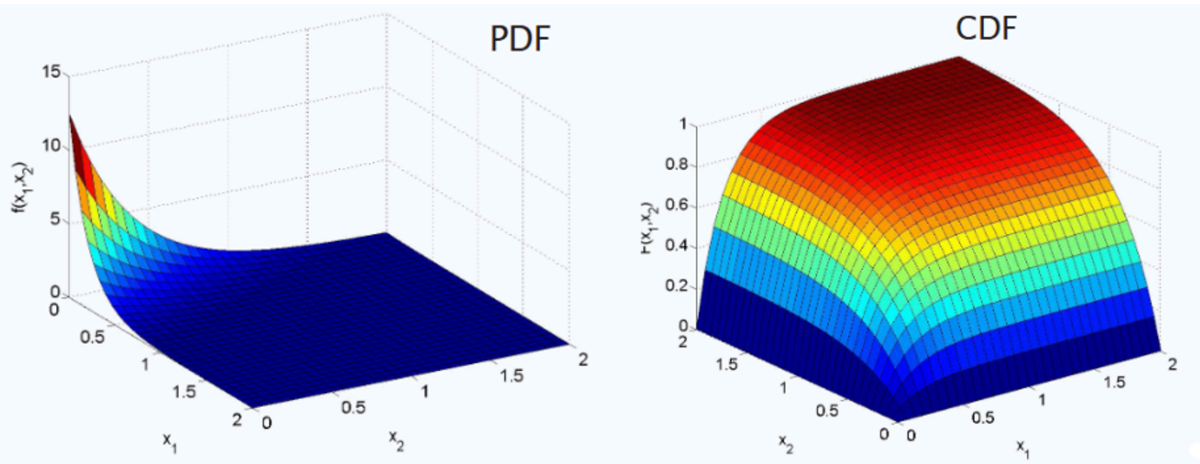

Figure 11: A bivariate exponential distributions PDF [Left] and CDF [Right]