

# 可靠度資料分析

# Reliability Data Analysis

許舒涵 (Shu-han Hsu)

成功大學 資訊工程系

**Lecture 3 –Reliability Basics**

# Class Presentation (After Midterm)

- Start thinking about a topic to pick
  - Submit topic title by 3/11/2024 on Moodle
- Apply what you learn to actual applications
- Each person presents 20min presentation on topic of choice
  - Can be application of reliability your research
  - Or, pick topic of interest related to reliability
- Research current state of reliability
- Identify future research / potential solutions to improve reliability
- Presentation format:
  - 15 min talk
  - 5 min Q & A
  - Either English or Chinese



# Example Topics

- Software reliability
- Reliability in artificial intelligence
  - Ex. Explainable AI
- Reliability in electronic design automation/circuit testing
- Predictive maintenance
- Control charts
- Data sampling in clinical trials
- Testing in your research area
- Etc.....



# Presentation Example Outline

| Story Arc                    |
|------------------------------|
| Problem statement            |
|                              |
| Why important                |
|                              |
| Background                   |
|                              |
| Issues with current practice |
|                              |
| Potential solutions          |
|                              |
| Recommendations              |
|                              |
| Future work                  |

# Example Part 1 of 3

From researcher in device radiation reliability:

| Story Arc         | Example of material discussed in presentation   |
|-------------------|---|
| Problem statement | Evaluating impact of new device geometries on sensitivity to ionizing radiation   |
| Why important     | Traditional scaling techniques no longer practical - forcing designers to develop new device geometries (example FinFET, 3D, etc).                |
| Background        | Characterization of ionizing radiation on traditional device geometry has been well documented, but sensitivity of new geometries presents a risk |

Note: Topics related to reliability may not necessarily have the word “reliability” in the subject

# Example Part 2 of 3

| Story Arc                    | Example of material discussed in presentation  |
|------------------------------|--|
| Issues with current practice | Traditional approach was to run radiation tests after devices have been fabricated and document sensitivity. Additional system level features could be implemented to detect and correct single bit errors. If new geometries are too sensitive, then system level mitigations may not be enough to mitigate device level failure rates. |
|                              |  |
| Potential solutions          | In order to rapidly evaluate potential device geometries for radiation, these options were evaluated:<br>[1] development of special test structures<br>[2] creation of tool to model radiation performance based on geometry   |

# Example Part 3 of 3

| Story Arc       | Example of material discussed in presentation  |
|-----------------|--|
| Recommendations | Based on cost and schedule impacts, it was determined that development of test structures offered the lowest risk of implementation and provides the most realistic results to provide rapid feedback to the design team |
|                 |  |
| Future work     | Development of model tool should be investigated, and test structure results could provide validation of model accuracy  |

# Grading Rubric

Presenter Name:

| Element                       | Value   | score | Comments |
|-------------------------------|---------|-------|----------|
| Clarity of presentation       | 1 to 10 |       |          |
| Technical content             | 1 to 10 |       |          |
| Delivery                      | 1 to 10 |       |          |
| Duration (15 min + 5 min Q&A) | 1 to 5  |       |          |
| Fielding questions            | 1 to 5  |       |          |
| Total                         | 40      |       |          |



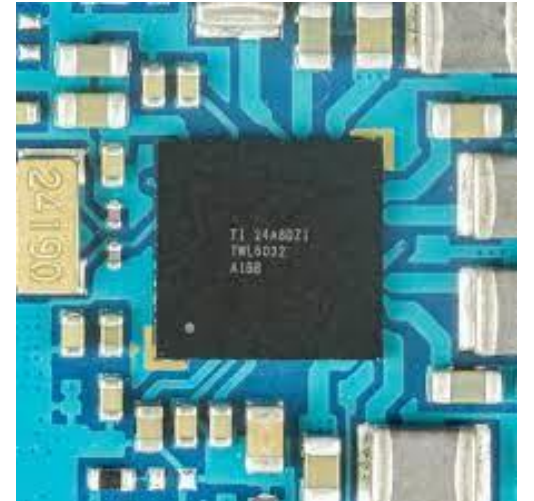
# Where to find literature...

- Example: Semiconductor Reliability
  - Conferences:
    - IEEE International Reliability Physics Symposium (IRPS)
    - European Symposium on Reliability of Electron Devices Failure Physics and Analysis (ESREF)
    - IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)
  - Journals:
    - IEEE Transactions on Very Large Scale Integration (VLSI) Systems
    - Can also be found in other IEEE Transactions related journals in semiconductors, devices, and circuits
    - Microelectronics Reliability



# Where to find literature...

- Example: EDA/Circuit Reliability Testing
  - Conferences:
    - International Test Conference (ITC)
    - IEEE VLSI Test Symposium (VTS)
    - ACM/IEEE Design Automation Conference (DAC)
    - Asia and South Pacific Design Automation Conference (ASP-DAC)
    - Design, Automation and Test in Europe Conference (DATE)
    - IEEE European Test Symposium (ETS)
    - IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)



# Where to find literature...

- Example: Reliability (More Statistics-oriented)
  - Journals:
    - IEEE Transactions on Reliability



# Last week

- Reviewed probability concepts for single random variables
- What if there are more than one random variable?

# Joint Random Variables

- Many problems in statistics and probability lead to models involving more than one random variable simultaneously. In these cases, we want to jointly model the behavior of two or more distributions.
- In this review, we will consider two cases:
  - Joint Discrete Random Variables
  - Joint Continuous Random Variables

# Joint Discrete Random Variables

- For two discrete rv's  $X : S_X \rightarrow RX$  and  $Y : S_Y \rightarrow RY$ , we define the joint probability mass function to be a function  $p : R_X \times R_Y \rightarrow [0, 1]$  such that:

$$p(x, y) = P(X = x \text{ and } Y = y).$$

- For any set  $A$  consisting of pairs  $(x, y)$  define:

$$P((X, Y) \in A) = \sum_{(x,y) \in A} p(x, y).$$

# Joint Discrete Random Variables

- We define the joint cumulative distribution function  $F : R_X \times R_Y \rightarrow [0, 1]$  such that:

$$F(x, y) = P(X \leq x \text{ and } Y \leq y) = \sum_{s \leq x, t \leq y} p(s, t)$$

joint probability mass function

- Also we define the marginal probability functions of X and Y and denote  $p_X$ , and, respectively,  $p_Y$ :

$$p_X(x) = \sum_{y \in R_Y} p(x, y)$$
$$p_Y(y) = \sum_{x \in R_X} p(x, y)$$

# Example

$$p_{X,Y}(x, y) = P(\{X = x\} \cap \{Y = y\}) = P(X = x, Y = y)$$

Sums to 1  $\rightarrow \sum_{x,y} p_{x,y}(x, y) = 1$  joint probability mass function

$$p_X(x) = \sum_y p_{X,Y}(x, y)$$

$$p_Y(y) = \sum_x p_{X,Y}(x, y)$$

|                  |   | $p_{X,Y}(x, y)$ |                |                |                |
|------------------|---|-----------------|----------------|----------------|----------------|
| $y \backslash x$ | 3 | 0               | $\frac{1}{10}$ | $\frac{1}{10}$ | $\frac{2}{10}$ |
|                  | 2 | $\frac{1}{10}$  | $\frac{3}{10}$ | $\frac{2}{10}$ | $\frac{6}{10}$ |
|                  | 1 | $\frac{2}{10}$  | 0              | 0              | $\frac{2}{10}$ |
|                  |   | $\frac{3}{10}$  | $\frac{4}{10}$ | $\frac{3}{10}$ | $p_X(x)$       |



# Joint Discrete Random Variables

- Similar to conditional probabilities, given the distribution of the rv  $X$  we can update/revise its distribution if partial information is available from other distributions. Therefore, given two rv's  $X$  and  $Y$  and their joint pmf we can define the conditional distribution of  $X$  given  $Y$  and denote  $X | Y$  as follows:

$$p(x|y) = P(X = x|Y = y) = \frac{p(x, y)}{p_Y(y)}$$

joint probability mass function  
marginal probability function

# Joint Continuous Random Variables

- For two continuous rv's  $X : S_X \rightarrow \mathbb{R}$  and  $Y : S_Y \rightarrow \mathbb{R}$ , we define the joint probability density function a function  $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  such that:

$$\begin{aligned} f(x, y) &\geq 0 \\ \int_{\mathbb{R} \times \mathbb{R}} f(x, y) dx dy &= 1 \end{aligned}$$

- For any set  $A$  consisting of pairs  $(x, y)$  define:

$$P((X, Y) \in A) = \iint_A f(x, y) dx dy$$

# Joint Continuous Random Variables

- We define the joint cumulative distribution function  $F : R_X \times R_Y \rightarrow [0, 1]$  such that:

$$F(x, y) = P(X \leq x \text{ and } Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy$$

- Also we define the marginal probability functions of X and Y and denote  $f_X$ , and, respectively,  $f_Y$ :

$$f_X(x) = \int_{R_Y} f(x, y) dy$$

$$f_Y(y) = \int_{R_X} f(x, y) dx$$

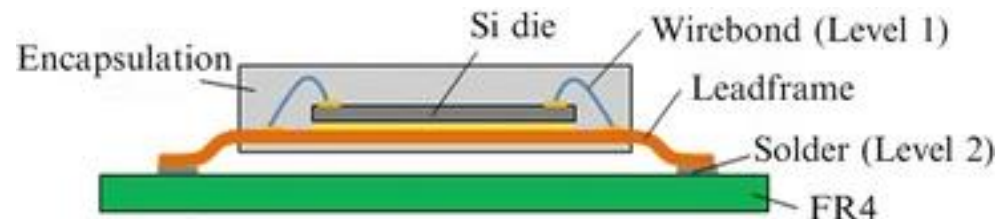
# Example

- You are estimating the failure time for the interconnect and solder joint to fail in product testing. Let  $X$  be the time it takes the interconnect to fail (in minutes) and  $Y$  be the time it takes the solder joint to fail, and suppose  $X$  and  $Y$  have the joint pdf

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-(x+2y)} & x, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- What is the probability that both the interconnect and solder joint fail in the next minute?

Example of interconnect (wire) and solder joint in an electronic product



# Joint Continuous Random Variables

- For two continuous rv's  $X$  and  $Y$  and their joint pdf's we can define the conditional distribution of  $X$  given  $Y$  and denote  $X | Y$  as follows:

$$f(x|y) = \frac{f(x, y)}{f_Y(y)}$$

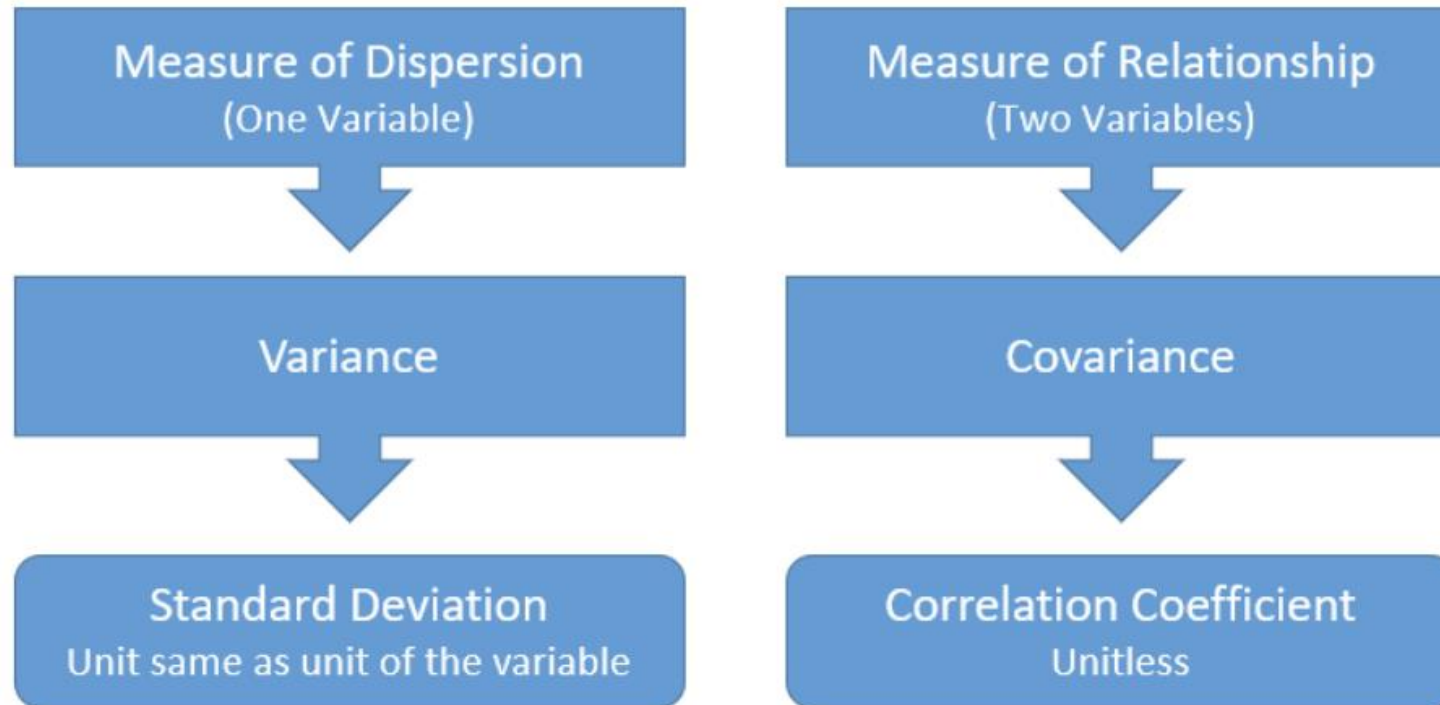
# Covariance, Correlation, and Independence

- Given the joint distributions of  $X$  and  $Y$ , we define their covariance:  
$$\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = E[XY] - E[X]E[Y]$$
- and we define their correlation:

$$\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{V(X)V(Y)}}$$

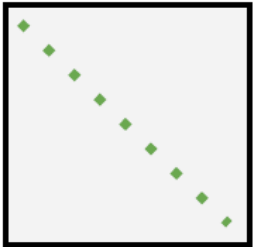
- Two rv's are independent if  $\text{cov}(X, Y) = 0$  or  $\text{cor}(X, Y) = 0$ . Also we can define independence in terms of the marginal and joint distributions (see next slide).

# Variance vs. Covariance

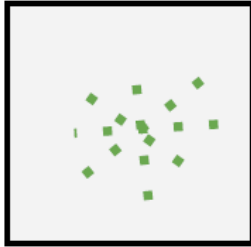


# Covariance vs. Correlation

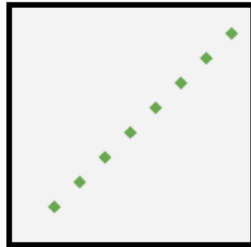
## COVARIANCE



Large Negative  
Covariance

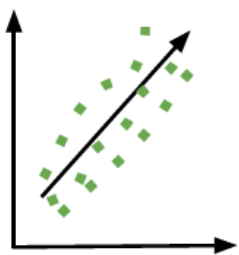


Nearly Zero  
Covariance

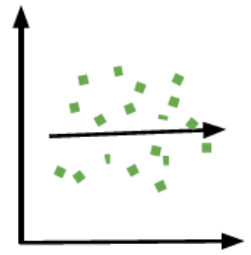


Large Positive  
Covariance

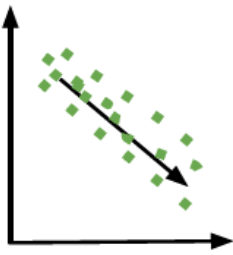
## CORRELATION



Positive  
Correlation



Zero  
Correlation



Negative  
Correlation

| Covariance   | Correlation   |
|--|---|
| Covariance is a measure of how much two random variables vary together | Correlation is a statistical measure that indicates how strongly two variables are related. |
| involve the relationship between two variables or data sets            | involve the relationship between multiple variables as well                                 |
| Lie between -infinity and +infinity                                    | Lie between -1 and +1   |
| Measure of correlation   | Scaled version of covariance  |
| provide direction of relationship                                      | provide direction and strength of relationship  |
| dependent on scale of variable   | independent on scale of variable  |
| have dimensions  | dimensionless   |



# Covariance, Correlation, and Independence

- Two discrete rv's  $X$  and  $Y$  with joint pmf  $p(x, y)$  and marginal pmf's  $p_X(x)$  and  $p_Y(y)$  are independent if:

$$p(x, y) = p_X(x)p_Y(y)$$

- Two continuous rv's  $X$  and  $Y$  with joint pdf  $f(x, y)$  and marginal pdf's  $f_X(x)$  and  $f_Y(y)$  are independent if:

$$f(x, y) = f_X(x)f_Y(y)$$

- Based on the definition above, two rv's with pmf's or pdf's  $f_X$  and  $f_Y$  are independent if  $f(x | y) = f(x)$  (that is, the conditional pmf or pdf of  $X | Y$  does not depend on  $y$ ).

Moving to more reliability specific concepts....



# Reliability Function

- Suppose  $n_o$  identical components are tested under their designed operating conditions. Furthermore, let us assume that during the interval  $(t - \Delta t, t)$ , we observed  $n_f(t)$  failed components, and  $n_s(t)$  surviving components such that  $[n_f(t) + n_s(t) = n_o]$
- Reliability at time  $t$ ,  $R(t)$ , is defined as the cumulative probability function of success, then;

$$R(t) = \frac{n_s(t)}{n_s(t) + n_f(t)}$$

- If  $T$  is a random variable denoting the time to failure, then the reliability function at time  $t$  can be expressed as;

$$R(t) = P(T > t)$$

Often called the **SURVIVAL FUNCTION!**

# Reliability, CDF, & PDF

- The cumulative distribution function CDF of failure,  $F(t)$  is the complement of  $R(t)$ , that is;

$$R(t) + F(t) = 1$$

- If  $T$  has a probability density function (p.d.f),  $f(t)$ , then,

$$R(t) = 1 - F(t) = 1 - \int_0^t f(v)dv$$

- Taking the derivative of the above equation with respect to  $t$ , we get,

$$\frac{dR(t)}{dt} = -f(t)$$

# Mean Time to Failure

- One of the key measures of a system's reliability is the MTTF.
  - Note that MTTF is usually used when the system is nonrepairable. For repairable systems, the failure time between two successive failures is usually referred to as MTBF (Between)
- Consider  $n$  identical nonrepairable systems and their time to failure are given by  $t_1, t_2, \dots, t_n$ . Then the mean time to failure is given as,

$$\widehat{MTTF} = \frac{1}{n} \sum_{i=1}^n t_i$$

# Mean Time to Failure

- If  $t$  is a random variable representing time to failure, then the Mean-Time-To-Failure, MTTF can be defined as follows:

$$MTTF = E[T] = \int_0^{\infty} t f(t) dt$$

$$\begin{aligned}\mu &= \int_0^{\infty} x f(x) dx \\ \mu &= E[X]\end{aligned}$$

- Another measure that is often used to describe the distribution of the time to failure is its variance  $\sigma^2$

$$\begin{aligned}\sigma^2 &= \int_0^{\infty} (MTTF)^2 f(t) dt \\ \sigma^2 &= \int_0^{\infty} t^2 f(t) dt - (MTTF)^2\end{aligned}$$

$$\sigma^2 = \int_0^{\infty} (x - \mu)^2 f(x) dx$$

# Mean Time to Failure

- If  $t$  is a random variable representing time to failure, then

$$MTTF = \int_0^{\infty} t f(t) dt$$

- Recall that  $f(t) = \frac{dF(t)}{dt} = -\frac{dR(t)}{dt}$ , thus

$$MTTF = \int_0^{\infty} t \left( -\frac{dR(t)}{dt} \right) dt = - \int_0^{\infty} t dR(t) = -tR(t) \Big|_0^{\infty} + \int_0^{\infty} R(t) dt$$

- Since  $R(\infty) = 0$  and  $R(0) = 1$ , we have,

$$MTTF = \int_0^{\infty} R(t) dt$$

Integration by parts

$$\int u dv = uv - \int v du$$

# Failure Rate

- The probability of failure of a component in a given interval of time  $[t_1, t_2]$  in terms of its reliability function is given as

$$\int_{t_1}^{t_2} f(t)dt = R(t_1) - R(t_2)$$

- Failure Rate in the interval  $[t_1, t_2]$  is defined as the probability that a failure per unit time occurs in the interval given that no failure occurred prior to  $t_1$ , in other words;

$$\frac{R(t_1) - R(t_2)}{(t_2 - t_1)R(t_1)}$$

This is conditional probability given  $R(t_1)$



- By replacing  $t_1$  and  $t_2$  with  $t$  and  $t + \Delta t$ , we can rewrite the above equation as follows,

$$\frac{R(t) - R(t + \Delta t)}{\Delta t R(t)}$$



# Failure Rate

- Note: There might be other definitions of failure rate in literature. Please check definition before usage.
  - Ex. In industry, it is common to use the definition:
    - Failure rate is defined in failures per hour

# Hazard Rate

- The Hazard Function is defined as the limit of the failure rate as  $\Delta t$  approaches zero, i.e., it is the **instantaneous Failure Rate**.

$$\begin{aligned}\lambda(t) = h(t) &= \lim_{\Delta t \rightarrow 0} \frac{R(t) - R(t + \Delta t)}{\Delta t R(t)} \\ &= \lim_{\Delta t \rightarrow 0} \frac{-[R(t + \Delta t) - R(t)]}{\Delta t} \frac{1}{R(t)} = \frac{-dR(t)}{dt} \frac{1}{R(t)} = \frac{f(t)}{R(t)}\end{aligned}$$

- The cumulative hazard function  $H(t)$  is the condition probability of failure over a period  $t$ . It is also the total number of failures during the time interval 0 to  $t$

$$L(t) = H(t) = \int_0^t h(v) dv$$

# Hazard Rate

- **A particular hazard rate function will uniquely determine a reliability function.**

- A.k.a. Can use the hazard rate function to figure out the reliability function
- To see this, let:

$$\lambda(t) = \frac{-dR(t)}{dt} \frac{1}{R(t)} \quad \text{or} \quad \lambda(t)dt = \frac{-dR(t)}{R(t)}$$

$$\text{Integrating, } \int_0^t \lambda(t')dt' = \int_1^{R(t)} \frac{-dR(t')}{R(t')},$$

where  $R(0) = 1$  establishes the lower limit in the integral

Then

$$-\int_0^t \lambda(t')dt' = \ln R(t)$$

Or

$$R(t) = \exp\left[-\int_0^t \lambda(t')dt'\right] \rightarrow$$

Can be used to derive the reliability function from a known hazard rate function

# Hazard Rate

- Commonly used form of hazard rate:

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

- Also called the instantaneous (conditional) failure rate
  - The failure rate is sometimes called a "conditional failure rate" since the denominator  $1 - F(t)$  (i.e., the population survivors) converts the expression into a conditional rate, given survival past time  $t$ .

# Average Hazard Rate

- Another useful function is the Average Hazard Rate, denoted as  $AHR(t_1, t_2)$ , and defined as,

$$AHR(t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \lambda(t') dt' = \frac{\ln R(t_1) - \ln R(t_2)}{t_2 - t_1}$$

- If  $t_1 = 0$  and  $t_2 = t$ , then  $AHR$  can be written as follows:

$$AHR(t) = \frac{-\ln R(t)}{t}$$

# Example

- An engineer approximates the reliability of a software system as:

$$R(t) = \begin{cases} (1 - \frac{t}{t_0})^2, & 0 \leq t < t_0 \\ 0, & t \geq t_0 \end{cases}$$

- a) Determine the hazard rate
- b) Does the hazard rate increase or decrease with time?
- c) Determine the MTTF.

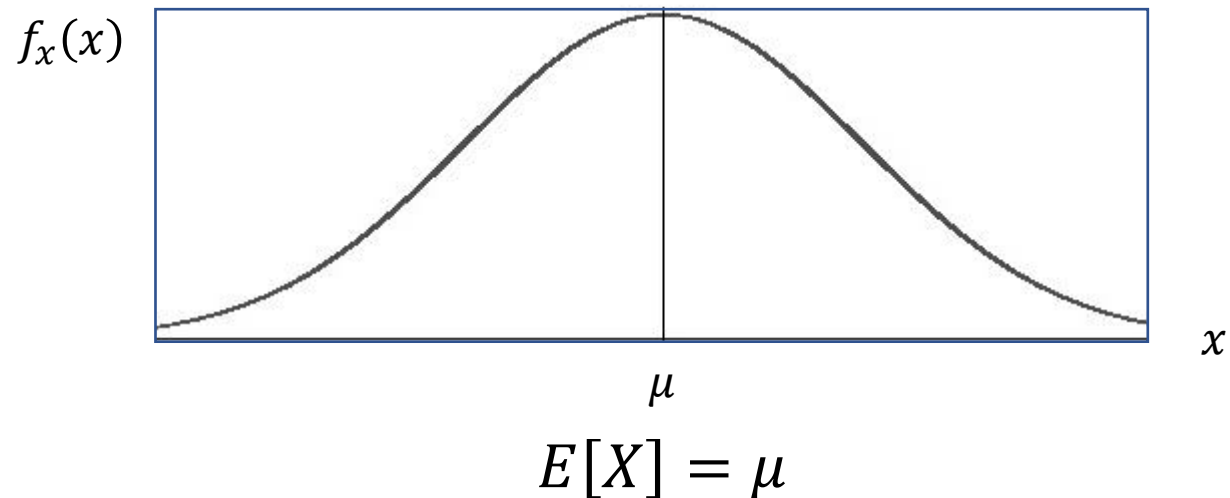
# Normal distribution

- We call a continuous random variable normal or Gaussian if it has pdf:

$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

where  $\mu$  and  $\sigma$  are parameters

- (Note: we need  $\sigma > 0$  for this to be a valid PMF)



# Normal distribution

- The CDF of a normal random variable  $X$  is:

$$P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} du$$

- There are no closed form expressions for this function. To calculate probabilities for normal random variables, must use tables, software functions, etc.
- For convenience, these are typically stated in terms of a **standard normal random variable**  $Y \sim \text{Normal}(0,1)$  which has CDF

$$\Phi(y) = P(Y \leq y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{u^2}{2}} du$$



# Linear transformations

- A linear transformation of a normal random variable is still a normal random variable.

**Theorem:**

If  $X \sim N(\mu_X, \sigma_X^2)$ , and  $Y = aX + b$ , where  $a, b \in \mathbb{R}$ , then  $Y \sim N(\mu_Y, \sigma_Y^2)$  where

$$\mu_Y = a\mu_X + b, \quad \sigma_Y^2 = a^2\sigma_X^2.$$

Proof:

We can write

$$X = \sigma_X Z + \mu_X \quad \text{where } Z \sim N(0, 1).$$

Thus,

$$\begin{aligned} Y &= aX + b \\ &= a(\sigma_X Z + \mu_X) + b \\ &= (a\sigma_X)Z + (a\mu_X + b). \end{aligned}$$

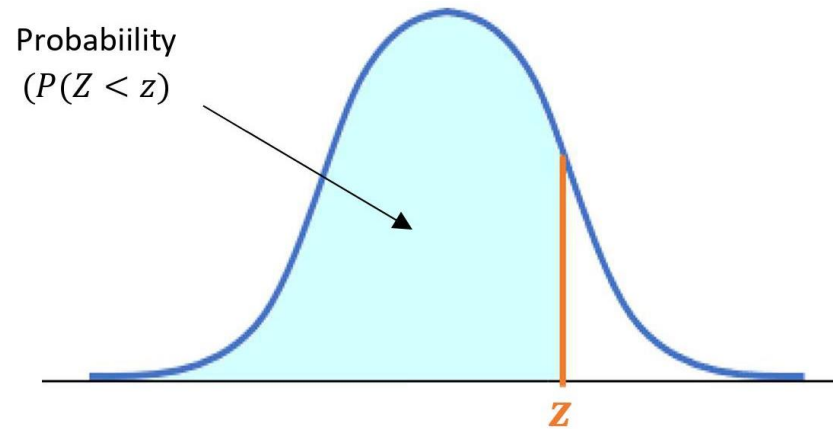
Therefore,

$$Y \sim N(a\mu_X + b, a^2\sigma_X^2).$$

**So, what does this all mean?**

# Normal distribution

- Any normal distribution can be converted to a standard normal distribution to find the CDF



If mean and standard deviation are known, use the Z-Table to find the CDF

**Z-Table**

$$Z = \frac{x - \mu}{\sigma}$$

| z   | 0.00   | 0.01   | 0.02   | 0.03   | 0.04   | 0.05   | 0.06   | 0.07   | 0.08   | 0.09   |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5754 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |

# Note: Z-Score Tables come in different formats!

- **First type (don't recommend):**

- Yields probability or area starting at the mean and going to the right of the mean up to the needed z-score.
- Usually labeled "**cumulative from mean**"
- This table basically works with half of the area under the normal curve, and the user must take this into consideration and make adjustments when using this table. This type of table lists positive z-scores only.

# Note: Z-Score Tables come in different formats!

- **Second type (use this):**

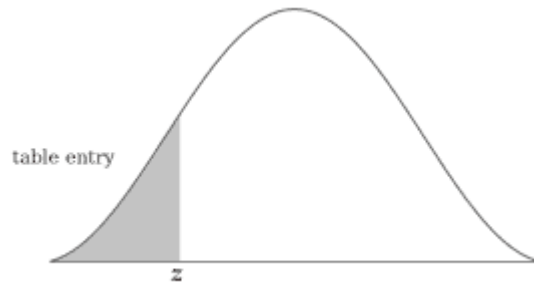
- Yields probability or area starting from negative infinity (the farthest left) and going to the right up to the needed z-score
- Usually labeled "**cumulative from the left**"
- This table works with the entire area under the normal curve, and requires less adjustments than the first option. This table lists both positive and negative z-scores.

# Note: Z-Score Tables come in different formats!

- **Second type (use this):**

- A.k.a. Need two tables, one positive and one negative.

Negative Z score table



| $z$  | 0      | 0.01   | 0.02   | 0.03   | 0.04   | 0.05   | 0.06   | 0.07   | 0.08   | 0.09   |
|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| -0   | .50000 | .49601 | .49202 | .48803 | .48405 | .48006 | .47608 | .47210 | .46812 | .46414 |
| -0.1 | .46017 | .45620 | .45224 | .44828 | .44433 | .44034 | .43640 | .43251 | .42858 | .42465 |
| -0.2 | .42074 | .41683 | .41294 | .40905 | .40517 | .40129 | .39743 | .39358 | .38974 | .38591 |
| -0.3 | .38209 | .37828 | .37448 | .37070 | .36693 | .36317 | .35942 | .35569 | .35197 | .34827 |
| -0.4 | .34458 | .34090 | .33724 | .33360 | .32997 | .32636 | .32276 | .31918 | .31561 | .31207 |
| -0.5 | .30854 | .30503 | .30153 | .29806 | .29460 | .29116 | .28774 | .28434 | .28096 | .27760 |
| -0.6 | .27425 | .27093 | .26763 | .26435 | .26109 | .25785 | .25463 | .25143 | .24825 | .24510 |
| -0.7 | .24196 | .23885 | .23576 | .23270 | .22965 | .22663 | .22363 | .22065 | .21770 | .21476 |

- Use the negative Z score table to find values on the left of the mean
- Corresponding values which are less than the mean are marked with a negative score in the z-table and represent the area under the bell curve to the left of  $z$ .

# Note: Z-Score Tables come in different formats!

- **Second type (use this):**

- A.k.a. Need two tables, one positive and one negative.

Positive Z score table

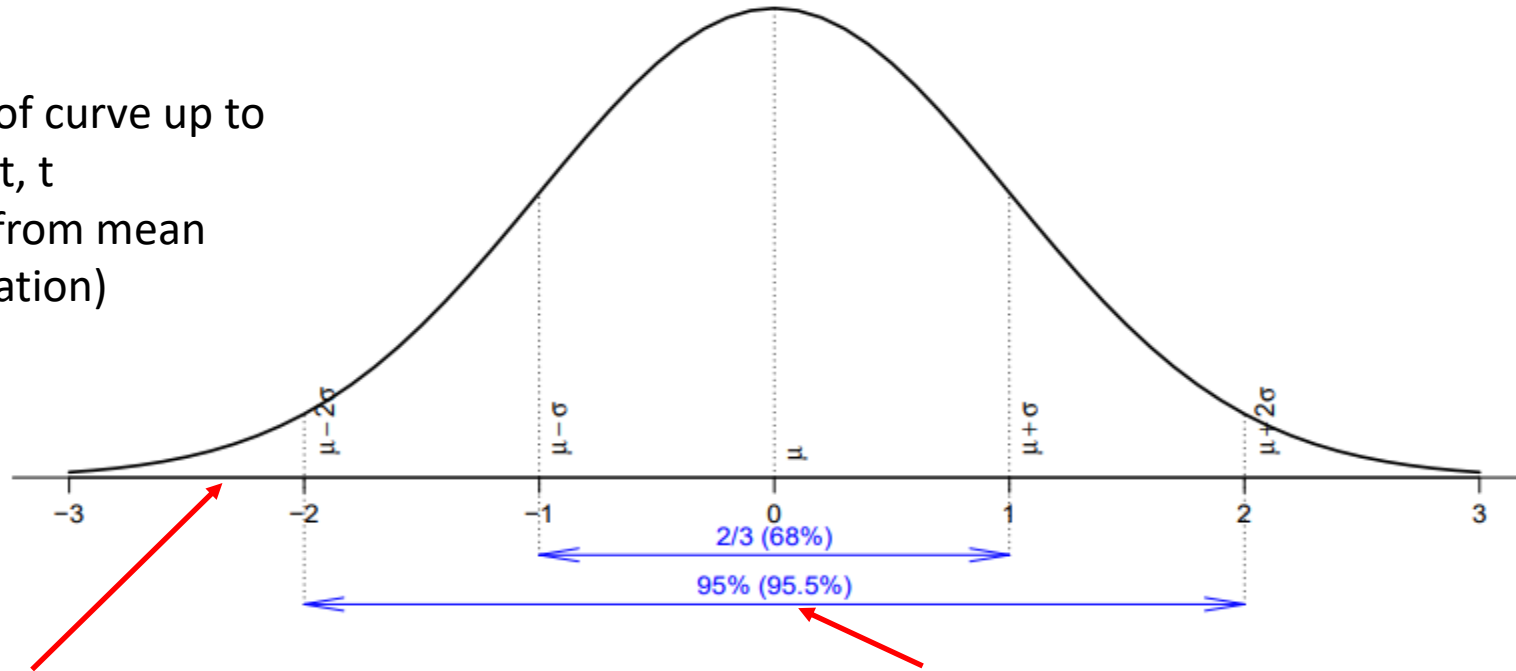


| $z$  | 0      | 0.01   | 0.02   | 0.03   | 0.04   | 0.05   | 0.06   | 0.07   | 0.08   | 0.09   |
|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| +0   | .50000 | .50399 | .50798 | .51197 | .51595 | .51994 | .52392 | .52790 | .53188 | .53586 |
| +0.1 | .53983 | .54380 | .54776 | .55172 | .55567 | .55966 | .56360 | .56749 | .57142 | .57535 |
| +0.2 | .57926 | .58317 | .58706 | .59095 | .59483 | .59871 | .60257 | .60642 | .61026 | .61409 |
| +0.3 | .61791 | .62172 | .62552 | .62930 | .63307 | .63683 | .64058 | .64431 | .64803 | .65173 |
| +0.4 | .65542 | .65910 | .66276 | .66640 | .67003 | .67364 | .67724 | .68082 | .68439 | .68793 |
| +0.5 | .69146 | .69497 | .69847 | .70194 | .70540 | .70884 | .71226 | .71566 | .71904 | .72240 |
| +0.6 | .72575 | .72907 | .73237 | .73565 | .73891 | .74215 | .74537 | .74857 | .75175 | .75490 |
| +0.7 | .75804 | .76115 | .76424 | .76730 | .77035 | .77337 | .77637 | .77935 | .78230 | .78524 |
| +0.8 | .78814 | .79103 | .79389 | .79673 | .79955 | .80234 | .80511 | .80785 | .81057 | .81327 |
| +0.9 | .81594 | .81859 | .82121 | .82381 | .82639 | .82894 | .83147 | .83398 | .83646 | .83891 |
| +1   | .84134 | .84375 | .84614 | .84849 | .85083 | .85314 | .85543 | .85769 | .85993 | .86214 |

- Use the positive Z score table to find values on the right of the mean.
- Corresponding values which are greater than the mean are marked with a positive score in the z-table and represent the area under the bell curve to the left of  $z$ .

# Two Ways of Looking at Normal Probability Density Function

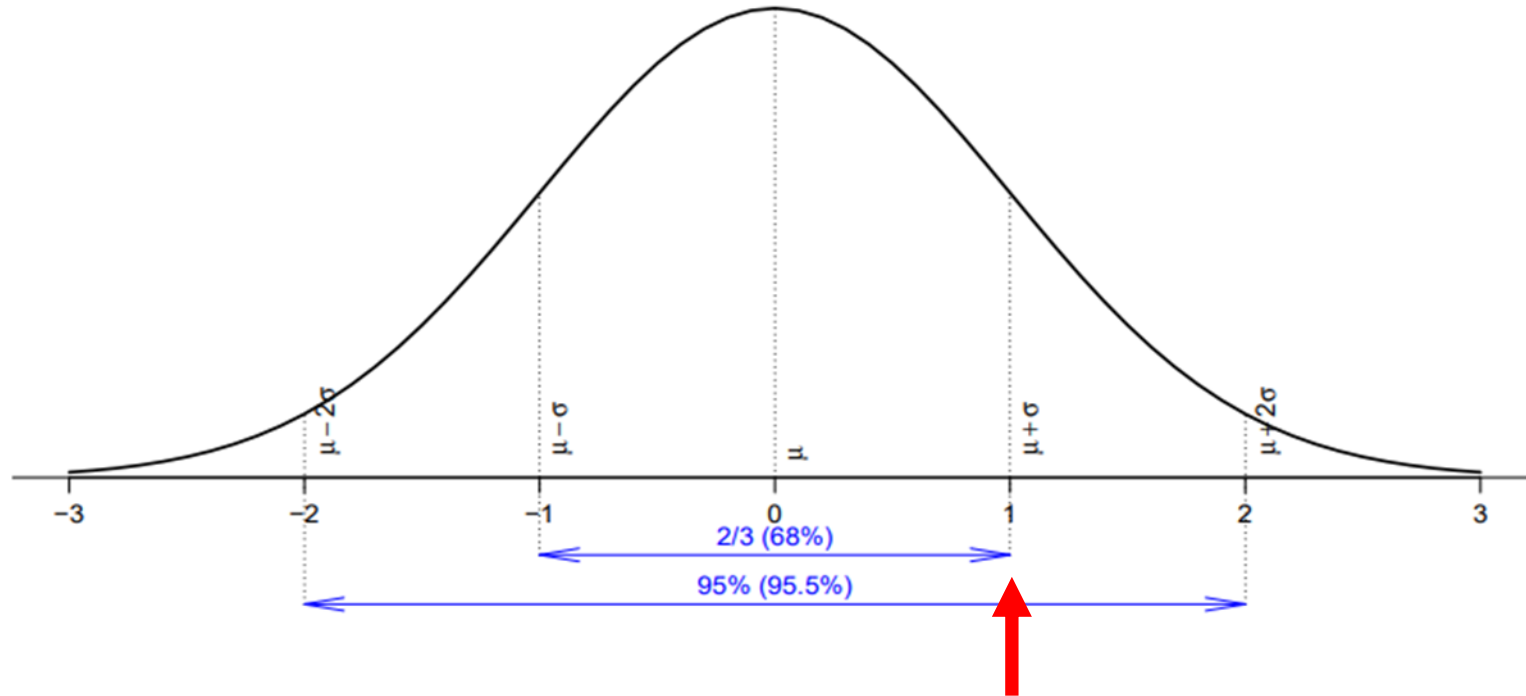
1. From left side of curve up to time of interest,  $t$
2. How far away from mean (standard deviation)



X axis can be anything.  
For this discussion it is time,  
specifically time to failure

Some applications such as  
the dimension of an object,  
one calculates area centered  
around the mean

# Z: Number of Standard Deviations



$$Z = \frac{x - \mu}{\sigma}$$

total number of standard deviations of a data point that lies below or above the mean.

$$Z = \frac{(\mu + \sigma) - \mu}{\sigma} = 1$$

Z = 1 is one standard deviation above the mean

(Note: Z = -1 is one standard deviation below the mean)

0.833 Area to the left of Z = 1



# Z-score vs Standard Deviation

- Both describe how far away from the mean the data is.
  - Z-score:  $z = \frac{x - \mu}{\sigma}$
  - Standard deviation:  $\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$ 
    - $x_i$  = every point in the dataset
    - $N$  = number of datapoints
- Difference is that standard deviation has units but Z score doesn't.
- Z score is helpful when comparing the values of two different samples that carry similar type of values, but are in different units.
- Example:
  - To compare the heights of 14 year old boys from two different schools, but school 1 reports the heights in feet & school 2 reports them in cm. Getting a Z score can help compare the heights without giving an erroneous result.

# Example

- You are running a test on 1,000 components. It is known that the product failure times follow a normal distribution with a mean time to failure of 10,000 hours.
- After 15,000 hours you have logged 695 failures.
- What must the standard deviation be of this population of components?

# Area Under the Curve

$$Z = \frac{x - \mu}{\sigma}$$

$\Phi(Z)$

| Z   | 0.00   | 0.01   | 0.02   | 0.03   | 0.04   | 0.05   | 0.06   | 0.07   |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8079 |

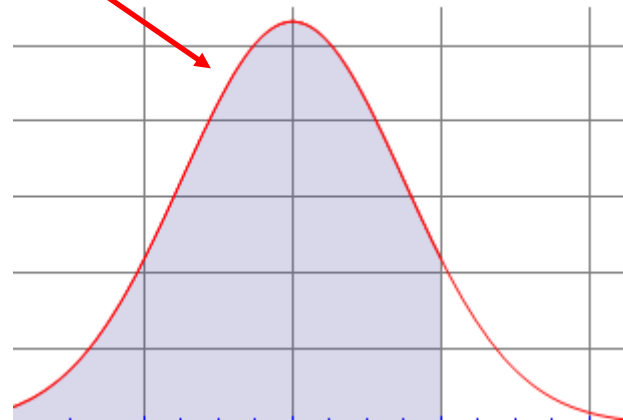
After looking up on table:

$$Z = \frac{x - \mu}{\sigma} = 0.51$$

$$\frac{15,000 - 10,000}{\sigma} = 0.51$$

$$\sigma = 9,803$$

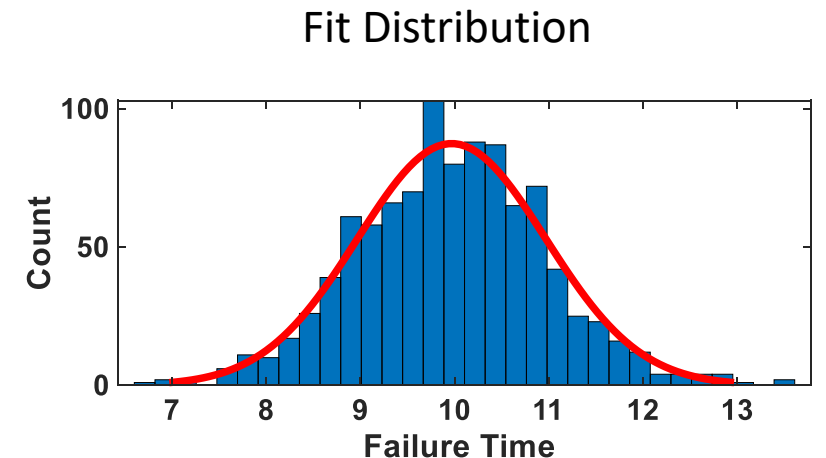
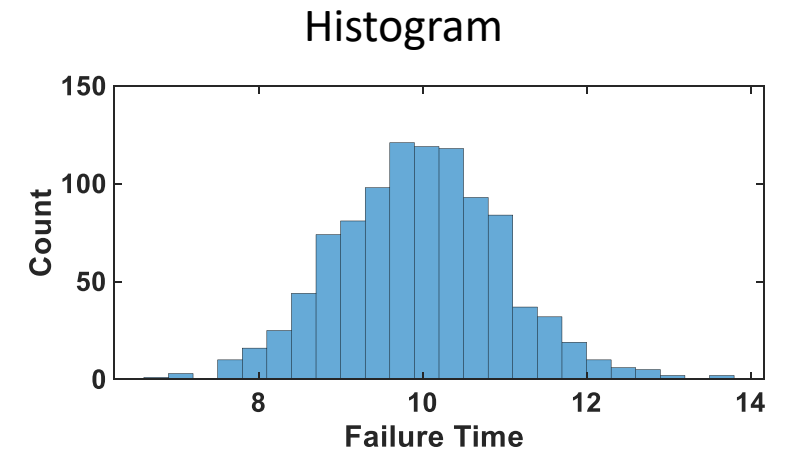
$$\Phi\left(\frac{x-\mu}{\sigma}\right) = 0.695$$



Putting it all together...

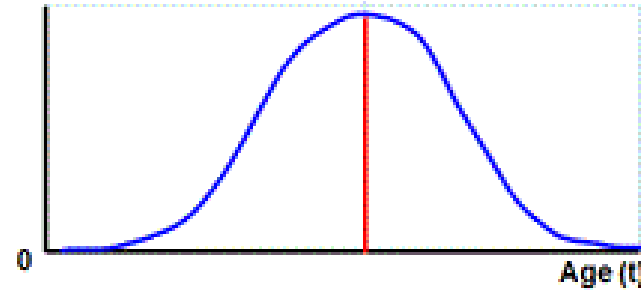
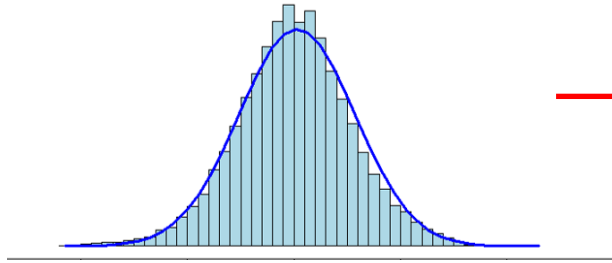
# Probability / Statistics

- Start with Histogram of Failures
  - Failure times dictate the functions
  - Create mathematical representations to fit reality
- Mathematical representations
  - Probability Density Function (PDF)
  - Cumulative Distribution Function (CDF)
  - Reliability Function (R)

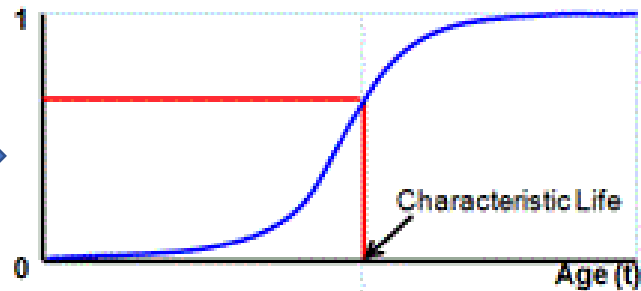


# Histogram -> PDF -> CDF -> Rel

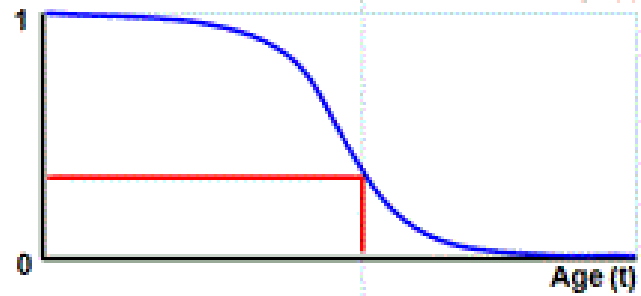
(Height in histogram is the number of failures in each time interval)



**Probability  
Density Function**



**Cumulative Failure  
Distribution  $F(t)$**

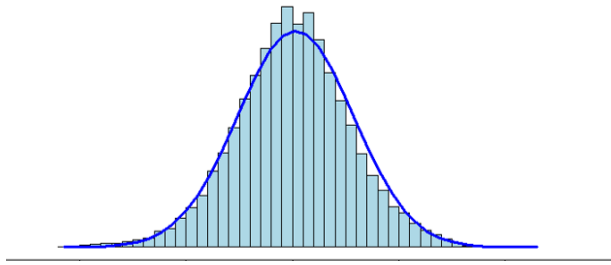


**Reliability  
Function  $R(t)$**

- CDF is lookup table for area under PDF
- When moving time further to right for CDF, more area under PDF is accumulated, increasing CDF value



# Raw Data to Normal PDF



Mean

$$\hat{\mu} = \bar{x} \equiv \frac{1}{n} \sum_{i=1}^n x_i$$

Variance

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

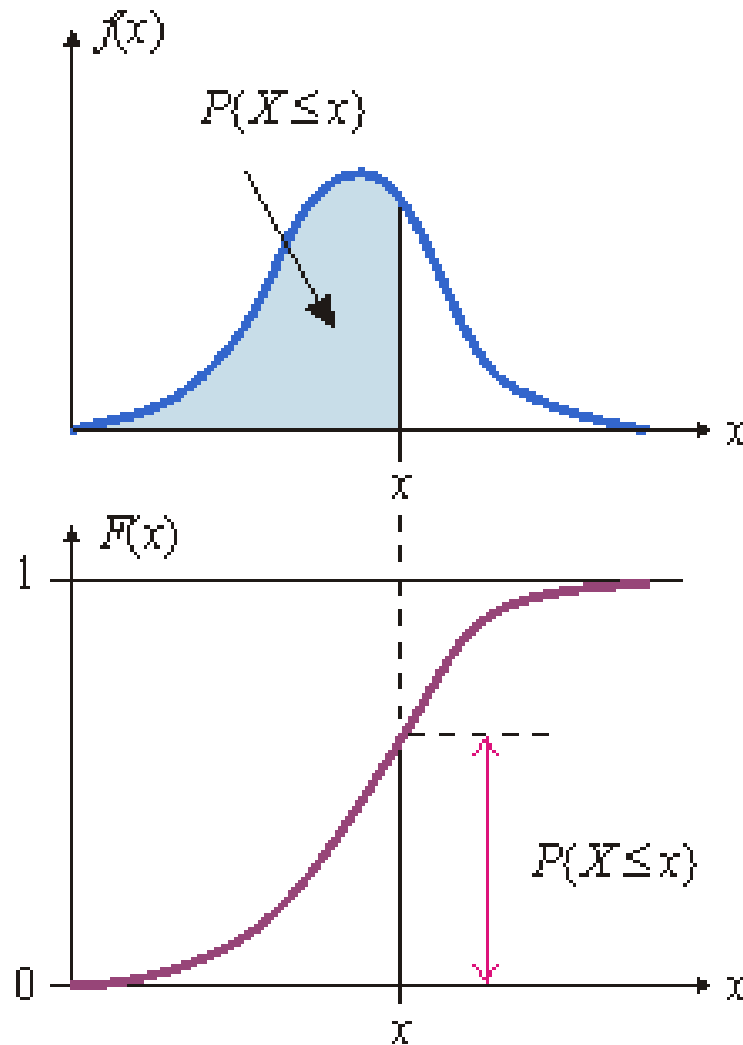
$$\textit{Standard Deviation} = \sqrt{\textit{variance}} = \sigma$$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{(x - \mu)^2}{2\sigma^2} \right], \quad -\infty < \mu < \infty, \quad \sigma^2 > 0.$$

*$f(x)$  is generic*

*$f(t)$  is applicable to time to failure*

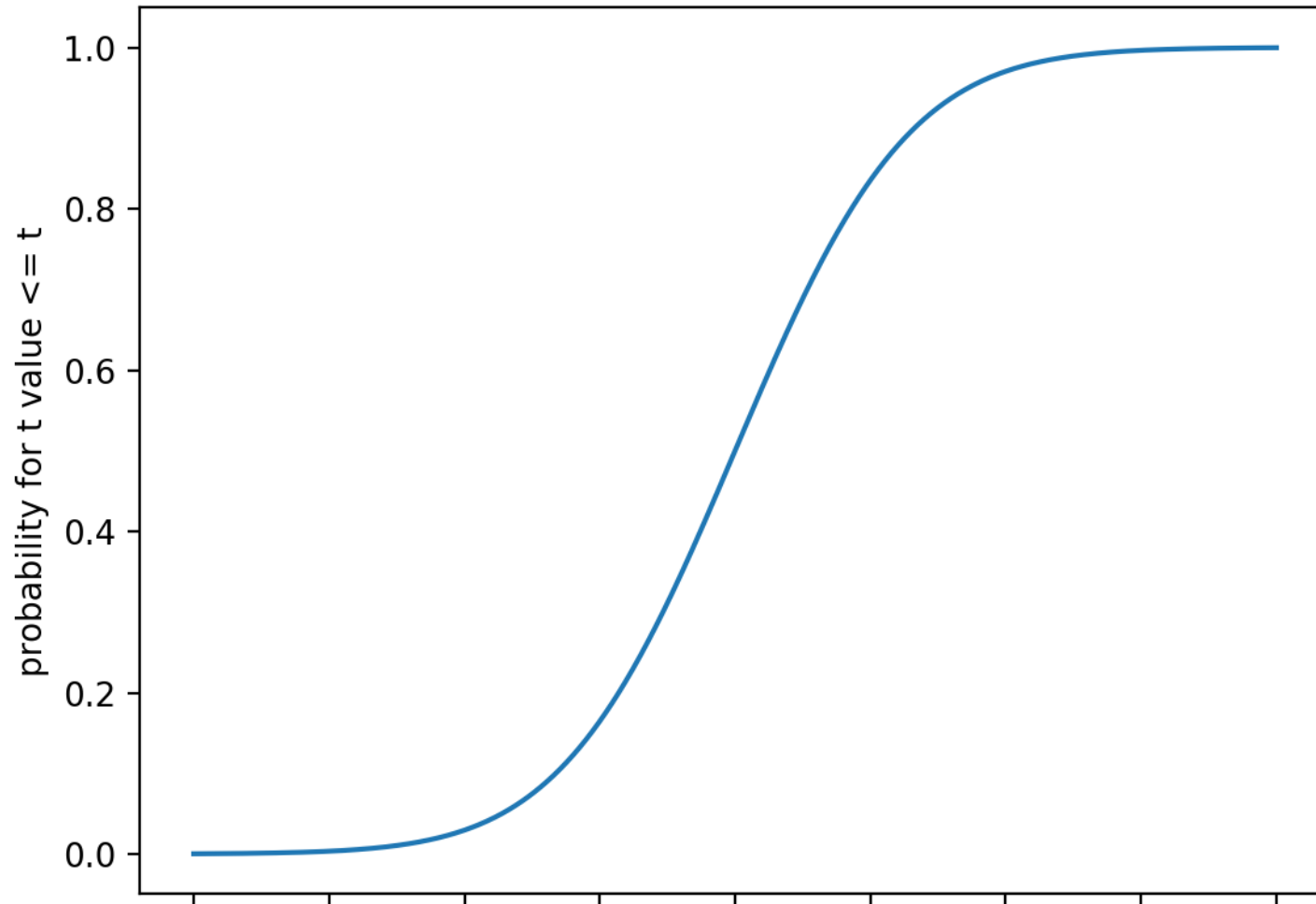
# PDF and CDF Relationship



The CDF acts as a  
Look up table for  
area of the PDF for  
each value of  $x$



# Cumulative Distribution Function

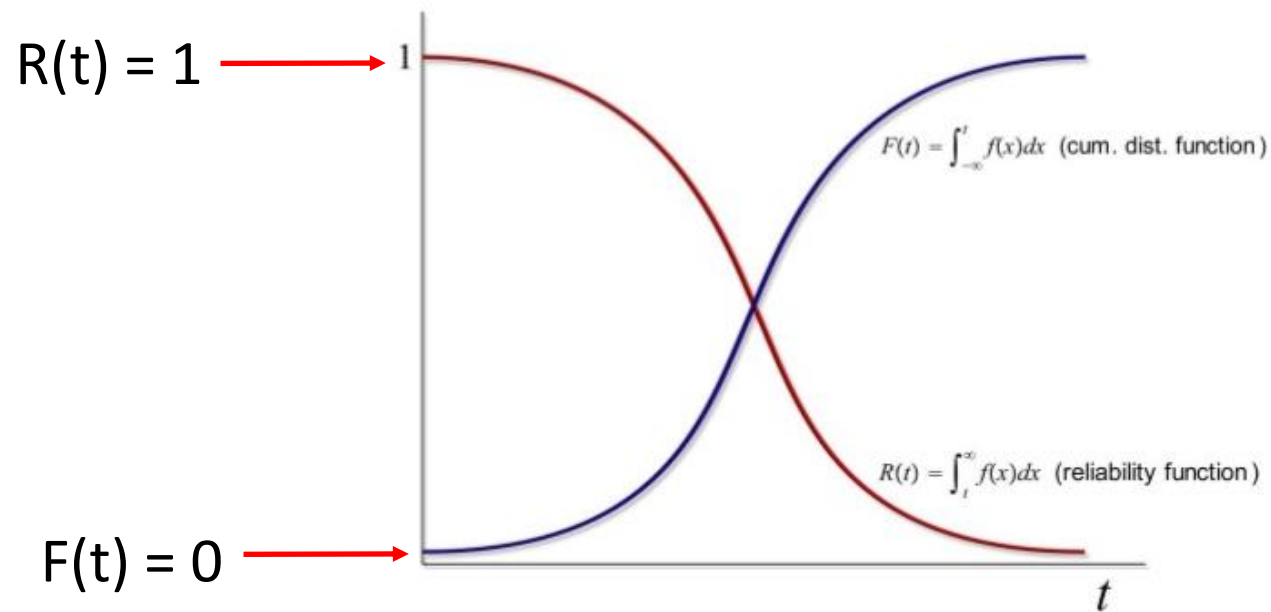


At time 0: probability of failure = 0

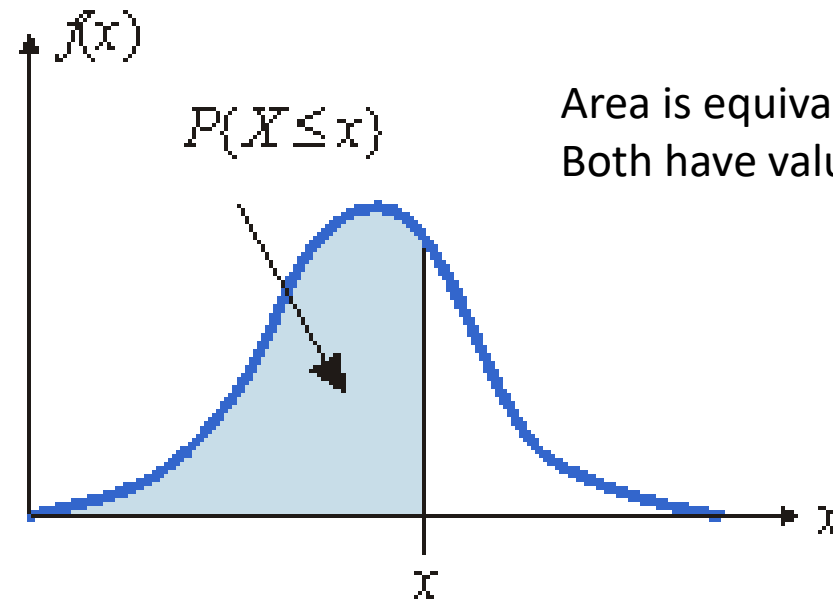
As time approaches infinity: probability of failure = 1

# CDF and Reliability

$$R(t) = 1 - F(t)$$



# Area and Probability

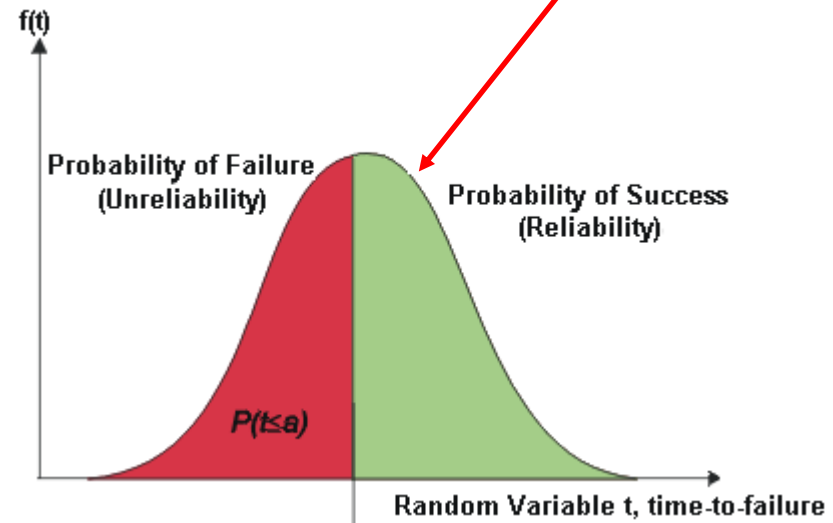


Area is equivalent to Probability  
Both have values from 0 -> 1

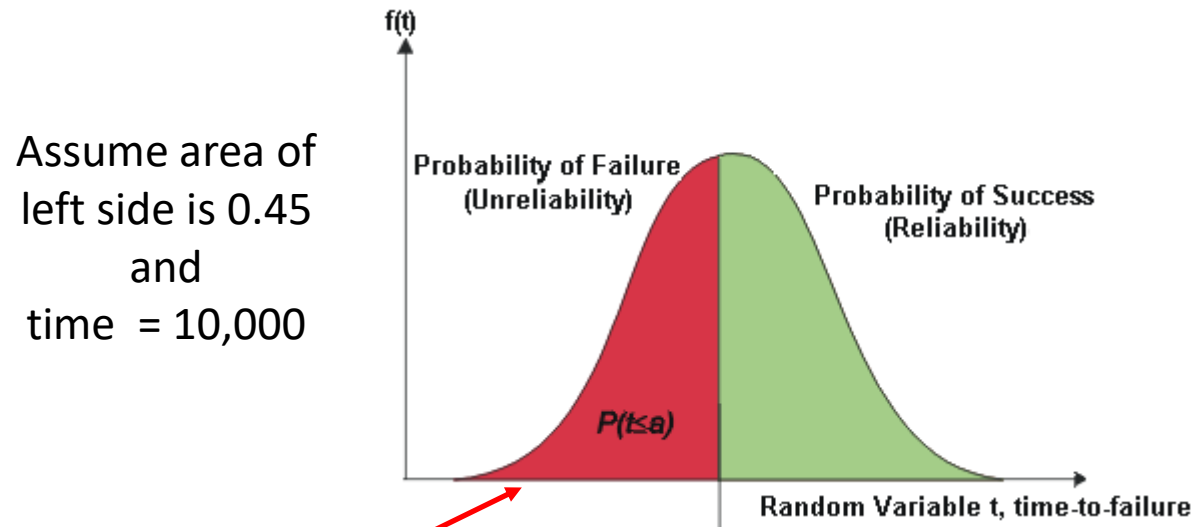
For this class,  
We are interested  
In time from 0 to  $x$

# Reliability Function

$$R(t) = 1 - F(t) = \int_t^{\infty} f(x) dx$$



# PDF and Reliability

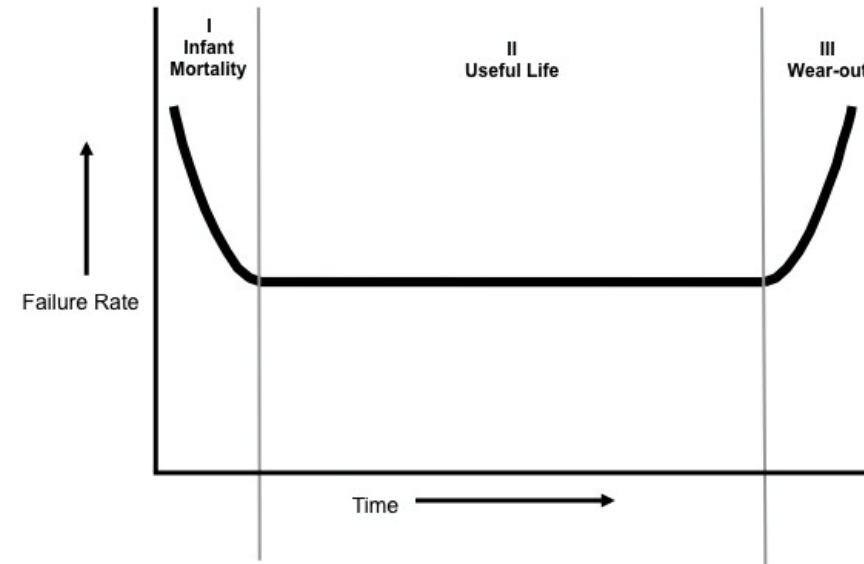


There are two ways to look at this area

[A] Probability of failure of 1 device is 0.45

[B] Test starts with 100 devices, after 10,000 hours – 45 will fail

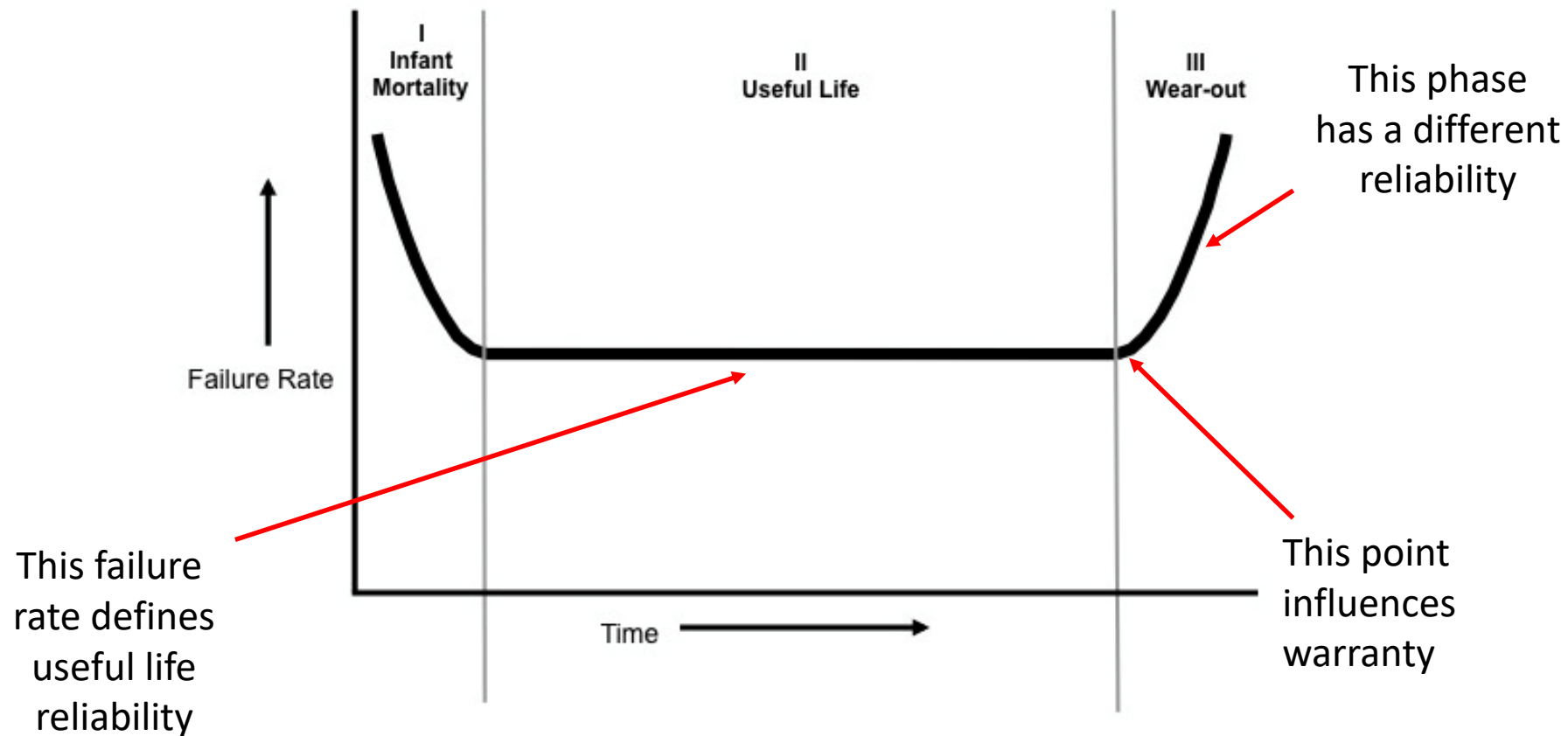
# Bathtub Curve



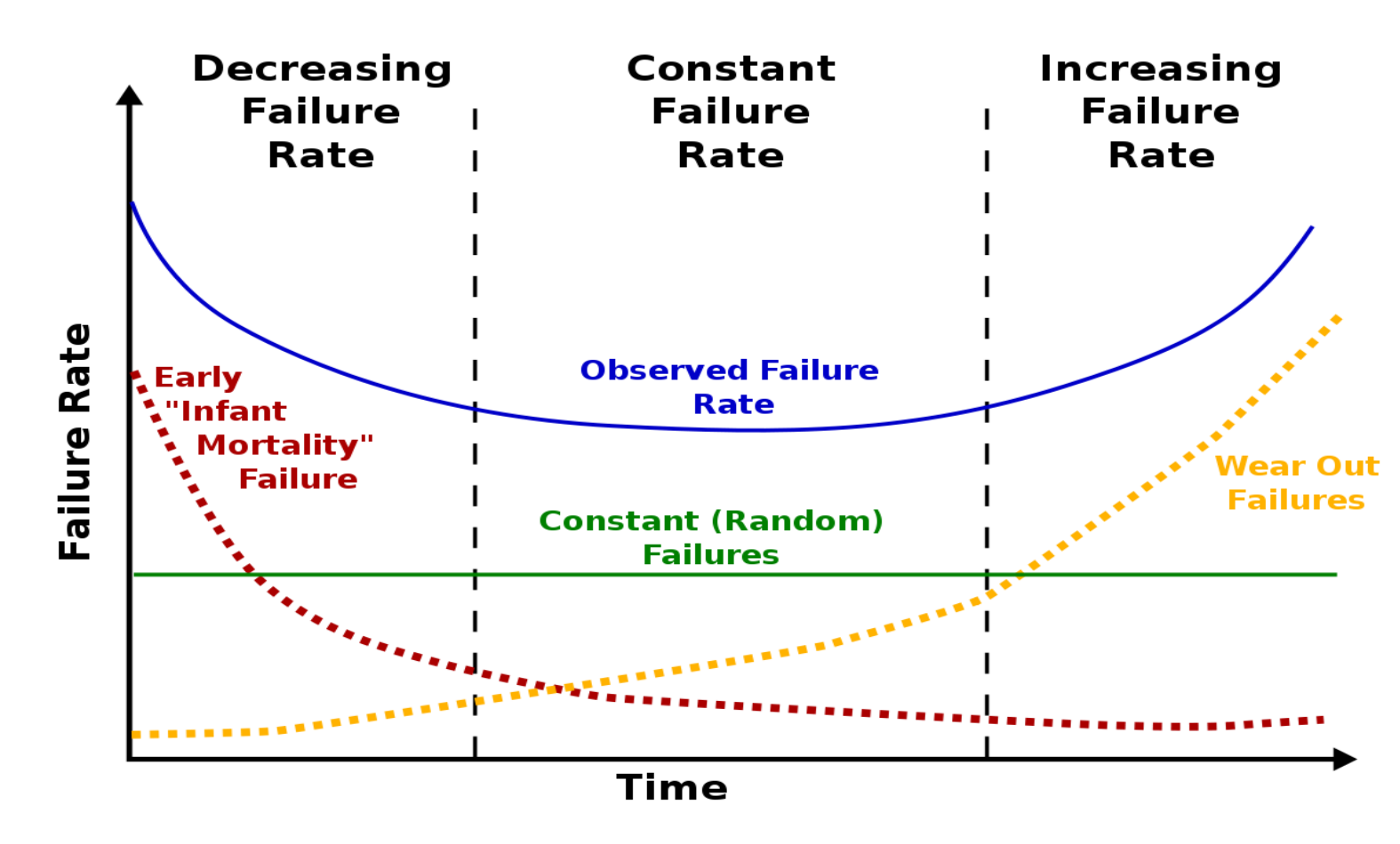
← Bathtub

# Bathtub Curve

Consider a population of identical components and take a large sample  $N$ .



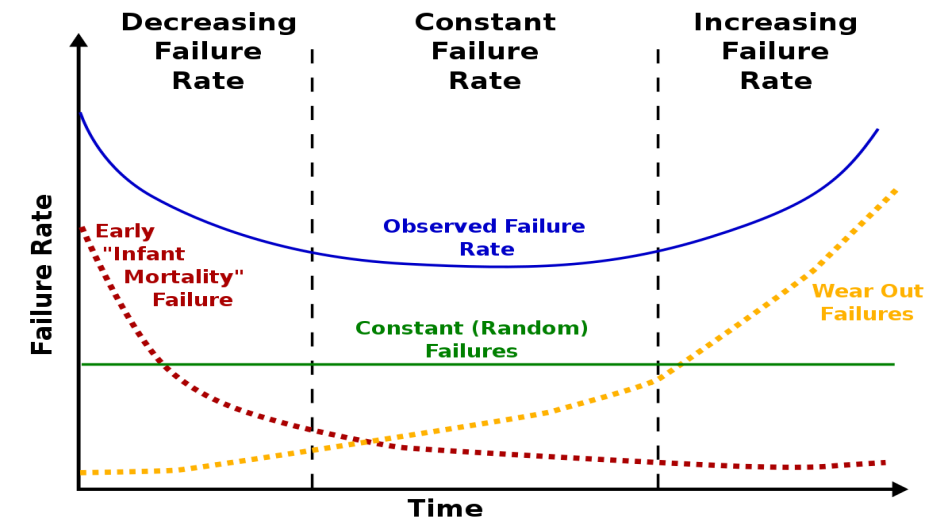
# Bathtub Curve





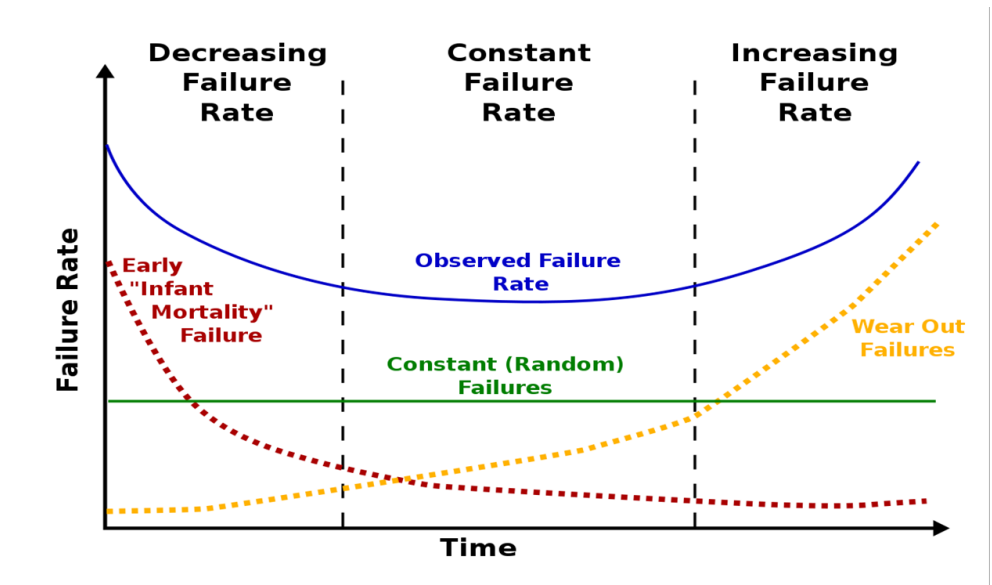
# Bathtub Curve–Decreasing Failure Rate

- Units experience a high failure rate at the beginning of the operation time due to manufacturing defects or poor installations and begins to decrease over time (decreasing failure rate, DFR), i.e., the time between failures begin to increase. This phase is known as infant mortality.
  - This phase is typically caused by manufacturing defects, cracks, poor workmanship, quality control, defective parts, contamination.
  - This phase can be reduced through burn-in testing where the unit is subjected to slightly more severe conditions than those encountered under normal operation.



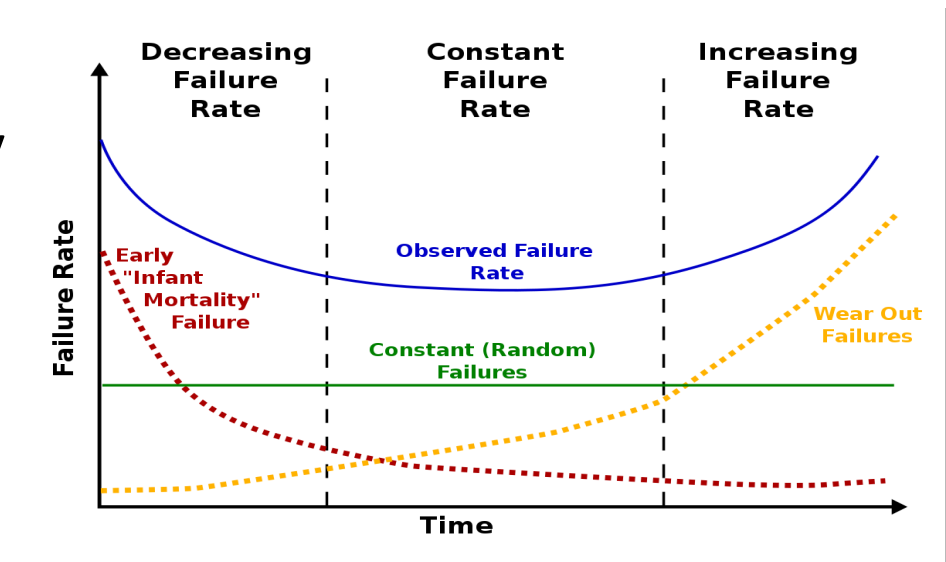
# Bathtub Curve–Constant Failure Rate

- The failure rate begins to level for a period of time which is characterized by a constant failure rate (CFR). In this region, failures are random and do not follow a predictable pattern.
  - This phase is typically referred to as the “useful life”.
  - Failure may be caused by random loads, human error, chance events, “Act of God”-type events
  - This phase can be reduced by redundancy or excess strength.



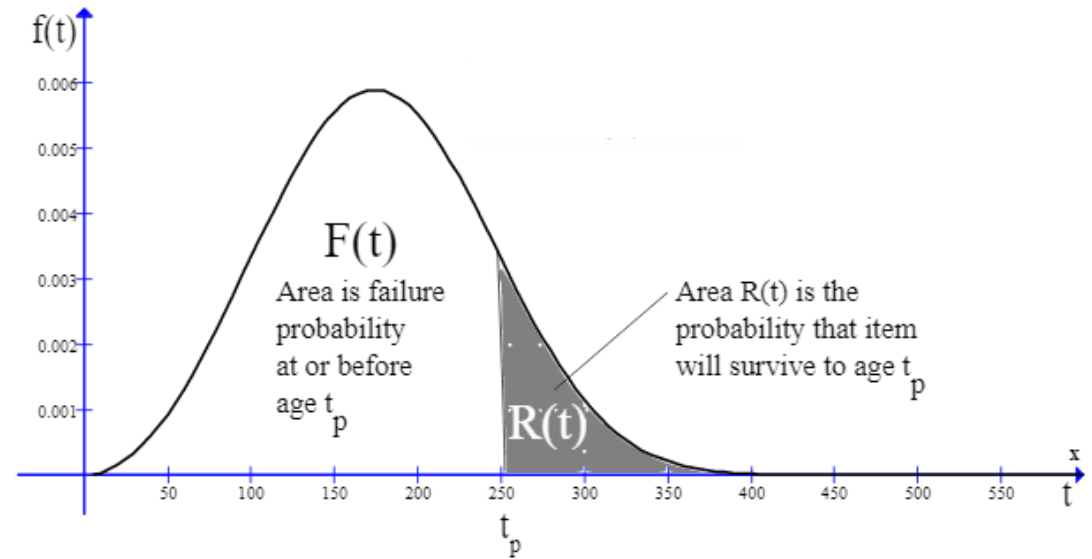
# Bathtub Curve–Increasing Failure Rate

- The third region also known as the wear-out phase is characterized by an increasing failure rate (IFR). Failures in this phase are no longer characterized by being random and are mostly due to aging and wear
  - Typical causes of failure in this phase are fatigue due to cyclic loading, wear, corrosion.
  - This phase can be reduced through derating, preventive maintenance, parts replacement, condition monitoring using sensor technology
    - Note: Derating is the practice of limiting thermal, electrical, and mechanical “stresses” to levels below the manufacturer's specified ratings, to improve reliability.



# Summary

- Reliability function  $R(t)$  : 1-CDF
  - Also called survival function
  - $R(t) + F(t) = 1$
  - $\frac{dR(t)}{dt} = -f(t)$
- Mean time to failure:
  - $MTTF = E[T] = \int_0^{\infty} tf(t)dt$
  - average amount of time a non-repairable asset operates before it fails
  - part's average lifespan
- Mean time between failure:
  - For repairable systems, the failure time between two successive failures is usually referred to as MTBF (Between)



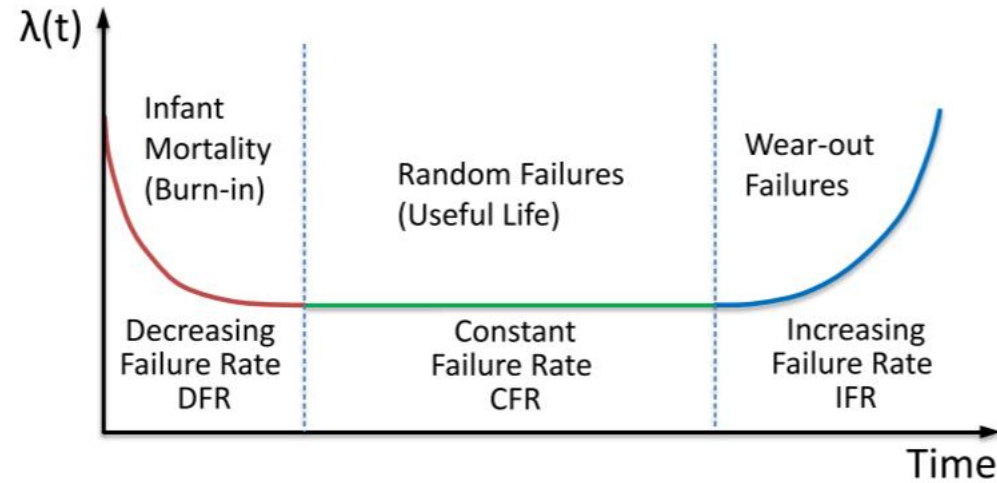
<https://www.livingreliability.com/en/posts/real-meaning-of-the-six-rcm-curves/>

# Summary

- Failure rate:  $f(t) = -\frac{dR(t)}{dt}$ 
  - May have other definitions, ex. number of failures per hour
- Hazard rate:  $h(t) = \frac{f(t)}{1-F(t)} = \frac{f(t)}{R(t)}$ 
  - instantaneous(conditional) failure rate, may also be written as  $\lambda(t)$
  - More mathematical definition:  $\lambda(t) = \frac{-dR(t)}{dt} \frac{1}{R(t)}$
- Can determine reliability function from hazard rate:
  - $R(t) = \exp[-\int_0^t \lambda(t')dt']$

# Summary

## Bathtub Curve



|             | Characterized By | Caused By   | Reduced By  |
|-------------|------------------|---|---|
| Burn-in     | DFR              | Manufacturing defects: Welding flaws, defective parts, poor quality/workmanship, contamination. | Burn-in testing, screening, acceptance testing, quality control |
| Useful Life | CFR              | Environment, random loads, human error, chance events   | Redundancy, excess strength                                     |
| Wear-out    | IFR              | Fatigue, corrosion, aging, cyclic loading   | Derating, part replacement, preventive maintenance              |