

What is the 'lognormal distribution' is ... and how to use it

If reliability engineering isn't making you money, then you are doing it wrong



Workbook

Presented by

Dr Chris Jackson

Reliability engineer



Copyright © 2025 Christopher Jackson

All rights reserved. No part of this publication may be reproduced, distributed or transmitted in any form or by any means, including photocopying or other electronic or mechanical methods without the prior written permission of the author (Christopher Jackson), except as permitted under Section 107 or 108 of the 1976 United States Copyright Act and certain other non-commercial uses permitted by copyright law.

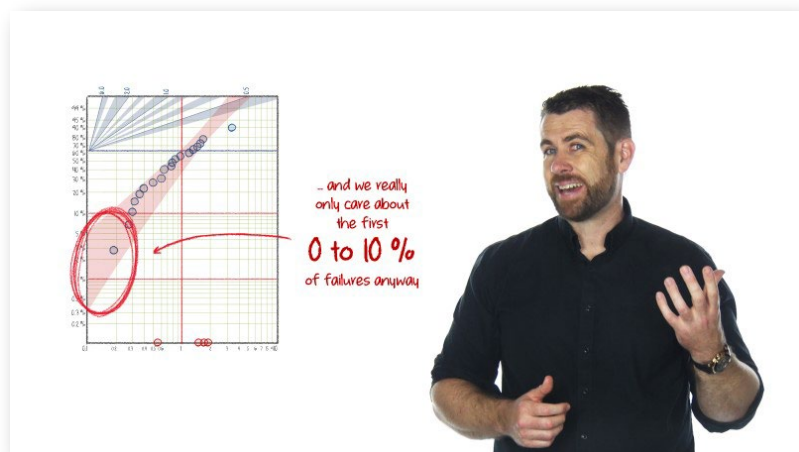
For permission requests, write to Christopher Jackson using the details on the following pages.

Limit of Liability/Disclaimer of Warranty: While the author has used his best efforts in preparing this document, he makes no representations or warranties with respect to the accuracy or completeness of the contents of this document and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. The author shall not be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Who is your teacher? ... Dr Chris Jackson

Dr Jackson holds a Ph.D. and MSc in Reliability Engineering from the University of Maryland's Center of Risk and Reliability. He also holds a BE in Mechanical Engineering, a Masters of Military Science from the Australian National University (ANU) and has substantial experience in the field of leadership, management, risk, reliability and maintainability. He is a Certified Reliability Engineer (CRE) through the American Society of Quality (ASQ) and a Chartered Professional Engineer (CPEng) through Engineers Australia.

Dr Jackson is the cofounder of www.is4.com online learning, was the inaugural director of the University of California, Los Angeles (UCLA's) Center of Reliability and Resilience Engineering and the founder of its Center for the Safety and Reliability of Autonomous Systems (SARAS). He has had an extensive career as a Senior Reliability Engineer in the Australian Defence Force, and is the Director of Acuitas Reliability Pty Ltd.



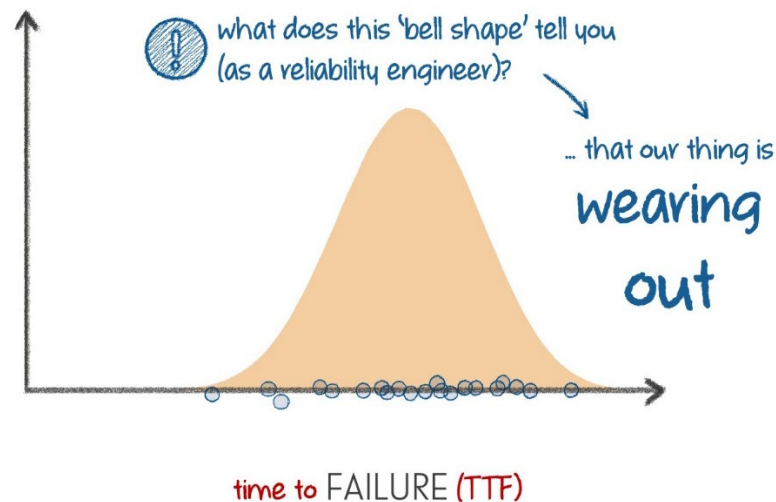
He has supported many complex and material systems develop their reliability performance and assisted in providing reliability management frameworks that support business outcomes (that is, make more money). He has been a reliability management consultant to many companies across industries ranging from medical devices to small satellites. He has also led initiatives in complementary fields such as systems engineering and performance-based management frameworks (PBMFs). He is the author of two reliability engineering books, co-author of another, and has authored several journal articles and conference papers.

Table of Contents

... the CENTRAL LIMIT THEOREM	5
and the LOGNORMAL DISTRIBUTION.....	8
... the normal distribution does addition – this one does multiplication	8
PROBABILITY DENSITY FUNCTION (PDF)	11
CUMULATIVE DISTRIBUTION FUNCTION (CDF)	11
RELIABILITY FUNCTION	12
HAZARD RATE.....	12
... where do we use LOGNORMAL DISTRIBUTIONS?	13
... and the FINAL WORD	18

The **normal distribution PDF curve** is a 'perfectly symmetrical' bell shape. This shape isn't there because it looks 'pretty.' This shape is actually based on an underlying process that appears in lots of probability and statistics scenarios, random processes and the natural world.

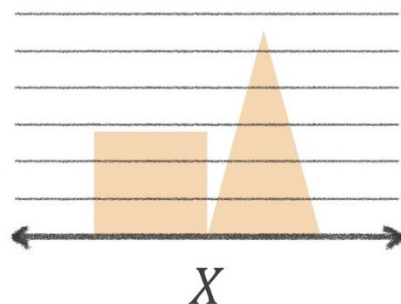
But first, remember what it means when a **TTF random variable** is described by a **PDF curve** with a 'bell shape?'



So ... where does this nice shape come from?

... the CENTRAL LIMIT THEOREM

Let's make up a **random variable**, which has a clearly 'absurd' **PDF curve**. Let's make this absurd **PDF** look like it has a square on the left and a triangle on the right.

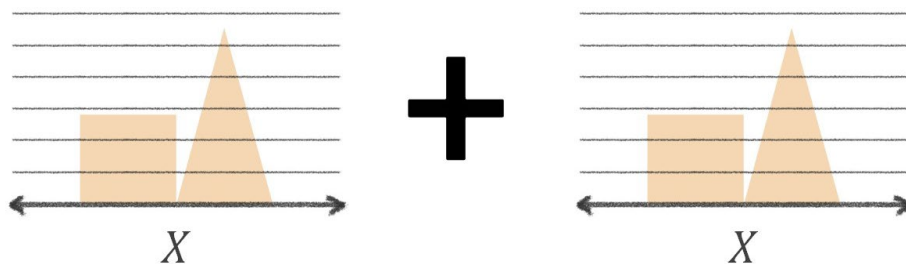


Hopefully you can agree that this **random variable** is not likely to exist! But now let's add this **random variable** to itself. That is ...

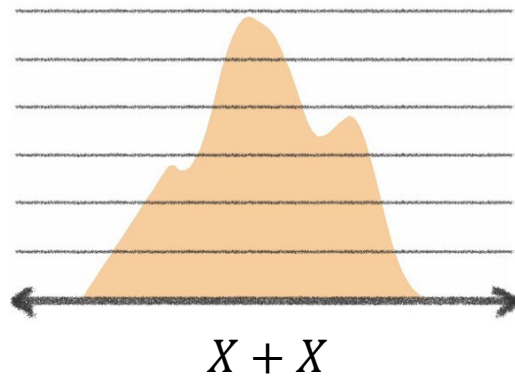
$$X + X$$

What does the **PDF** of the sum of these two **random variables** look like?

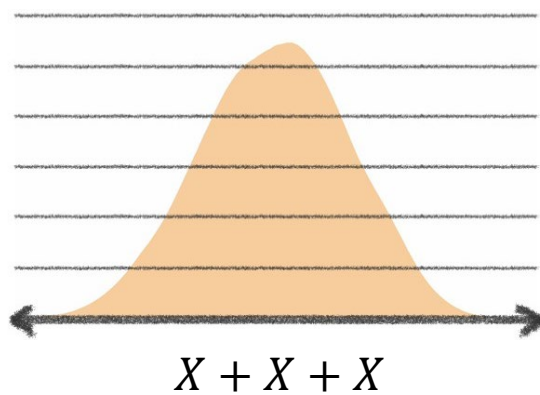
So we are trying to ...



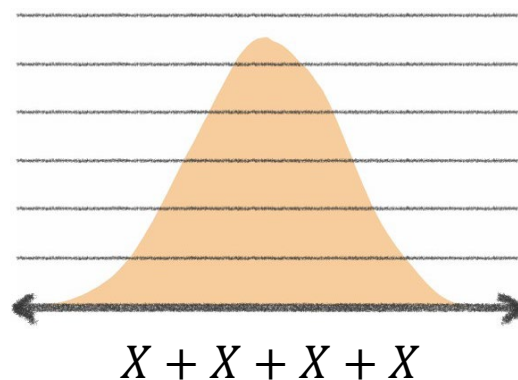
It turns out the PDF of this new **random variable** ($X + X$) looks like ...



You can already see that the shape of this new **PDF curve** is substantially different. It is smoother, and perhaps more like a bell shape! But let's add another X and see what happens ...

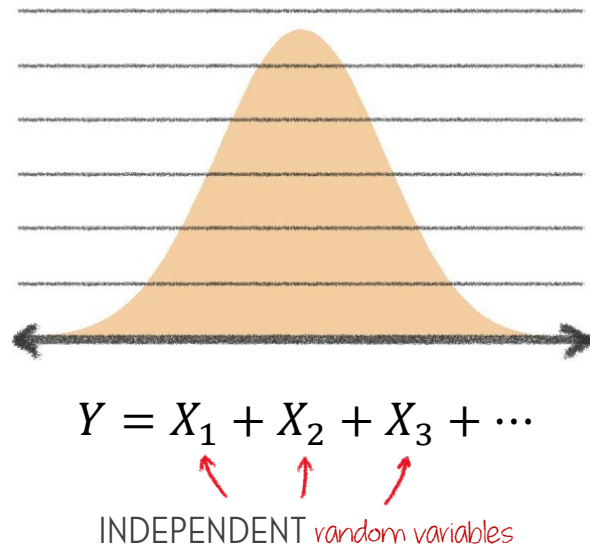


And let's add another X ...



It should be really obvious that the **PDF** of this **random variable** that is the sum of lots of 'X's' has a **PDF curve** that looks like a bell shape. And this is where the **normal distribution** comes from. If there a **random process** is based on the sum of lots of random processes, the output random variable will probably be described by a **normal distribution** with its bell-shaped **PDF curve**.

In other words, **random variable 'Y'** is a **normally distributed random variable** if ...



This phenomenon is known as the **Central Limit Theorem**.

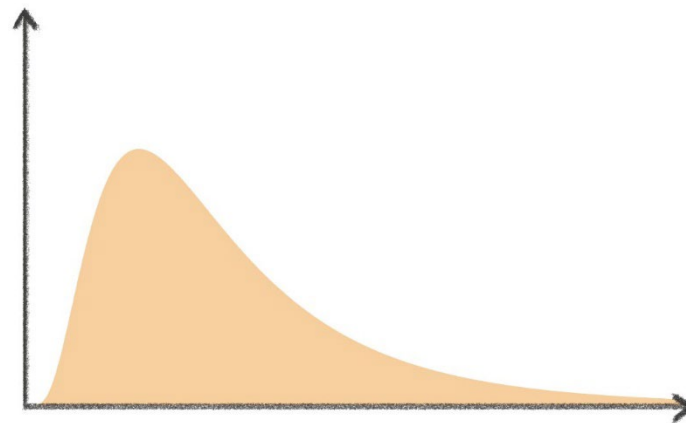
the CENTRAL LIMIT THEOREM states that any sum of (independently individually distributed) random variables will be approximately NORMALLY DISTRIBUTED.

So ... the **normal distribution** IS the 'bell curve.'

and the LOGNORMAL DISTRIBUTION

The **lognormal distribution** sounds like the **normal distribution**. It is obviously different, but still related. **Lognormal distribution PDF curves** are not all similar. They have different shapes based on different parameters.

That said, all **lognormal distribution curves** have a couple of common traits. They all have a 'left skew' with a right tail (as illustrated below). The nature of this skewness and tail length varies across different **lognormal distributions**.

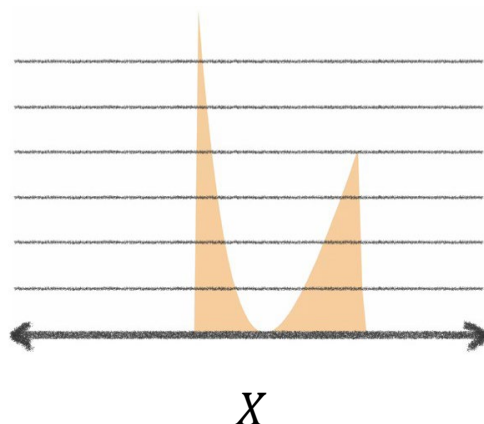


Because there is still a 'hill' or a 'mound' (even if it doesn't look like a perfect 'bell'), the **lognormal distribution** models **TTF** of products, systems or services that wear out.

So ... where does lognormal PDF shape come from?

... the normal distribution does addition – this one does multiplication

Let's make up another **random variable**, which has a clearly 'absurd' **PDF curve**.

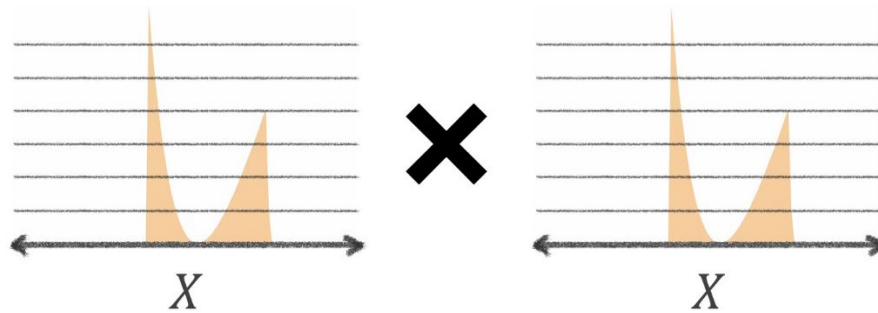


Now let's multiply this random variable with itself. That is ...

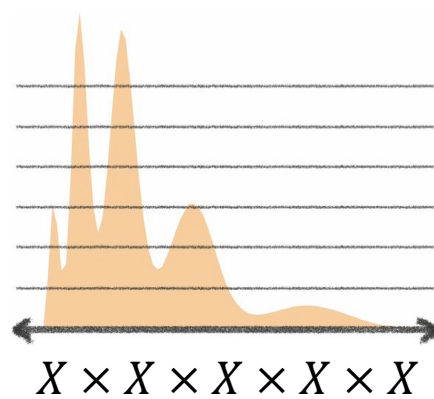
$$X \times X$$

What does the **PDF** of the product of these two **random variables** look like?

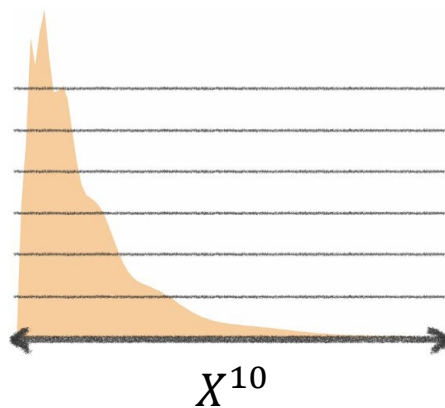
So we are trying to ...



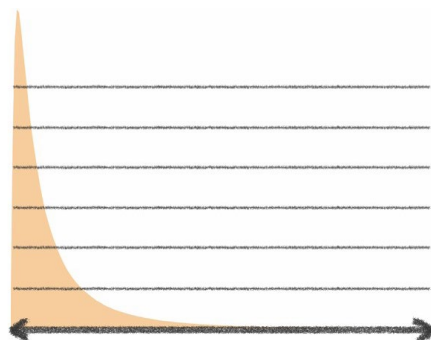
Let's multiply it a few times and look at the resulting **PDF** ...



Let's multiply the random variable 10 times ...



Let's multiply the random variable 20 times ...



$$X^{20}$$

You can see how the **PDF** of the new **random variable**, which is the product of lots of other random variables, is now quite smooth with that typical **lognormal PDF curve** shape: left skew with right tail. So what is happening here?

It comes back to **logarithms** and **exponents**. Look at the following **equation** ...

$$100 \times 1\,000 = 100\,000$$

Another way of writing this **equation** where we use **exponents** with 'bases' of 10 is ...

$$10^2 \times 10^3 = 10^5$$

Now if we look at all the **exponents** (**logarithms** to base 10), we can see that the sum of the **exponents** on the left-hand side (2 + 3) is equal to the sum of the **exponents** on the right-hand side (5).

$$2 + 3 = 5$$

So how is this relevant to the **lognormal distribution**? Well, remember that we are looking at the product of lots of different independent **random variables**. For example ...

$$T_1 \times T_2 \times T_3 \times T_4 \times T_5 = X$$

We know that because T is a **random variable**, then X is too. So let's use some actual figures:

$$16.58 \times 18.72 \times 20.36 \times 20.93 \times 23.13 = 3\,059\,236$$

Now let's take the **natural logarithms** of each value. For example, $\ln 16.58 = 2.808$. So therefore ...

$$2.808 + 2.930 + 3.014 + 3.041 + 3.141 = 14.934$$

So in general ...

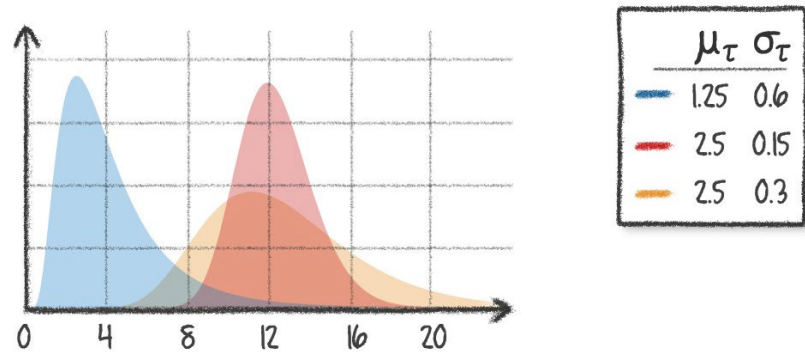
$$\ln T_1 + \ln T_2 + \ln T_3 + \ln T_4 + \ln T_5 = \ln X$$

... this random variable is
NORMALLY DISTRIBUTED

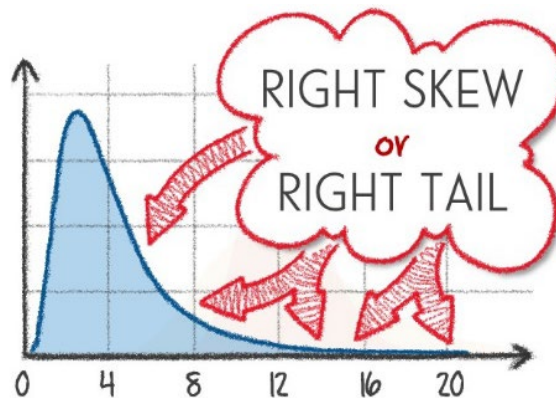
When this occurs, the **random variable** X is described by the **lognormal distribution**. So the **Central Limit Theorem** helps us understand what happens when we **ADD** lots of **random variables** **AND** what happens when we **MULTIPLY** lots of **random variables**.

PROBABILITY DENSITY FUNCTION (PDF)

The equations and formulae for the **PDF** and other functions and statistics based on the **Lognormal distribution** (including the **mean**) are at the end of this chapter. The most important thing to understand when dealing with **PDFs** is knowing what they represent in terms of failure. Three example **lognormal distribution PDF curves** are illustrated below.

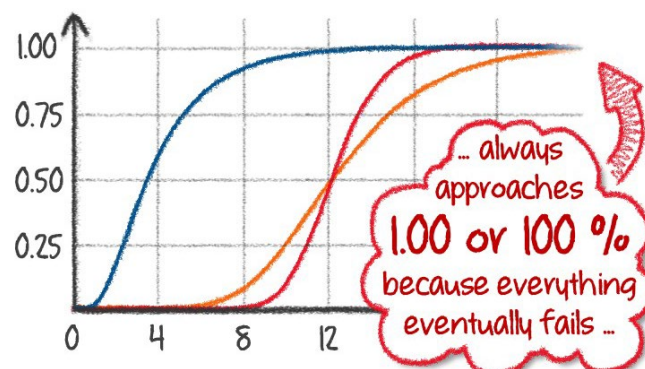


Lognormal distribution PDF curves have a particular shape ...



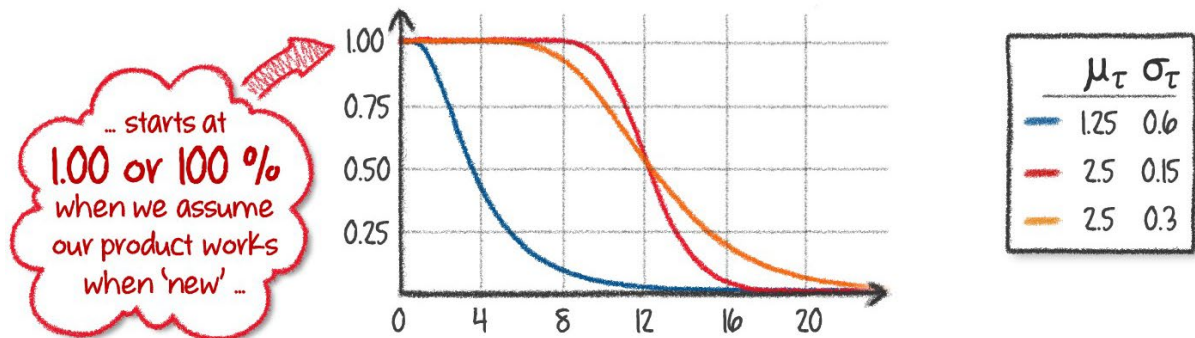
CUMULATIVE DISTRIBUTION FUNCTION (CDF)

Three example **Lognormal distribution CDF curves** are illustrated below. Each **CDF curve** corresponds to one of the **PDF curves** above. You can see here that the **Weibull distribution CDF** (like all other **CDFs**) starts at 0 and approaches 1 (or 100 %) as **TTF** increases.



RELIABILITY FUNCTION

Three example **Lognormal distribution reliability curves** are illustrated below. Each **reliability curve** corresponds to one of the **PDF and CDF curves** above. You can see here that the **Lognormal distribution reliability curve** (like all other **reliability curves**) starts at 1 (or 100 %) and approaches 0 as **TTF** increases.



HAZARD RATE

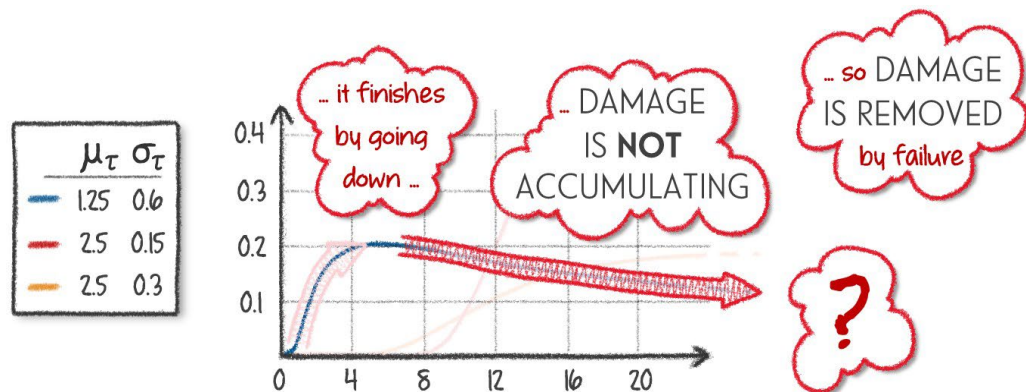
One way of looking at the **hazard rate** is by seeing it as the rate at which working (non-failed) systems are **failing**.

$$\frac{f(t)}{R(t)}$$

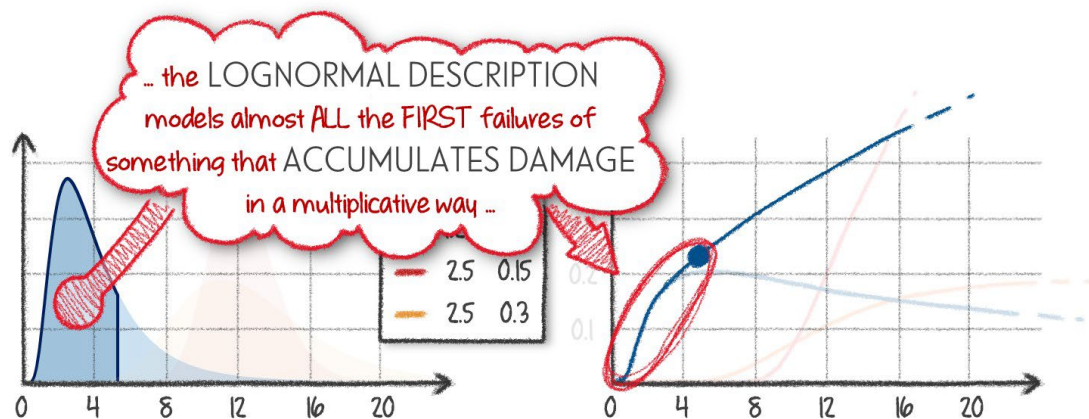
... the RATE at which a 'thing' that is still working - FAILS

PROBABILITY DENSITY FUNCTION (PDF)
RELIABILITY FUNCTION

Three example **lognormal distribution hazard rate curves** are illustrated below. Each **hazard rate curve** corresponds to one of the **PDF, CDF and reliability curves** above. But with the lognormal distribution, the hazard rate is unusual ...



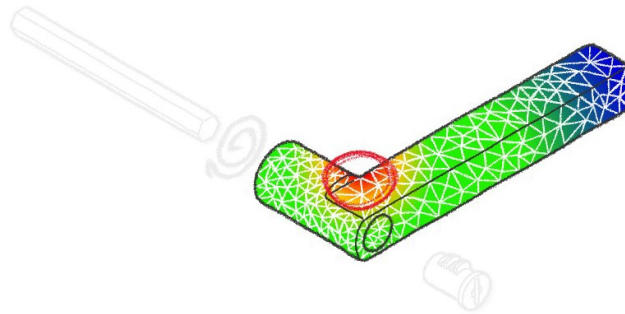
This is not possible for a **failure mechanism**. If a **failure mechanism** is based on the accumulation of damage, it never gets to a stage where the accumulation of damage stops.



... where do we use LOGNORMAL DISTRIBUTIONS?

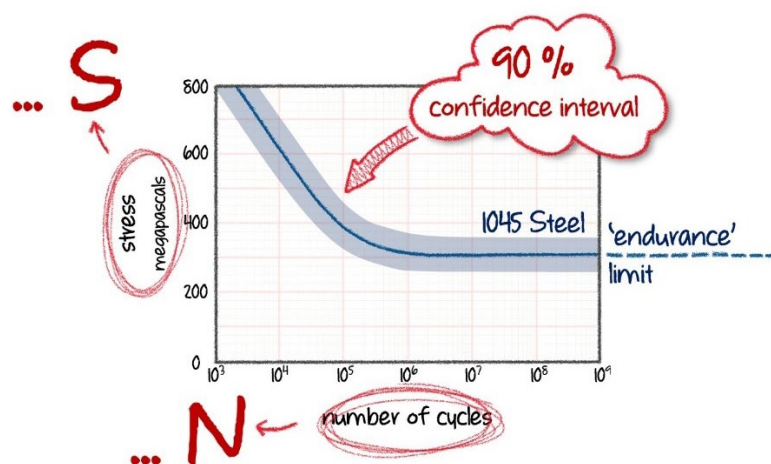
Let's look at the handle of a smart lock. We are planning on manufacturing it out of 1045 steel, which can be prone to fatigue failure. The door handle will routinely be opened and shut, so there will be a cyclic stress applied to it.

We conduct **Finite Element Analysis (FEA)** on the preliminary design and identify a region of stress concentration on the inside of the handle. And the local stress the handle will experience when the handle is opened with a 'high' level of force is 500 MPa. This will be used at the (conservatively high) assumed stress for each door handle use.



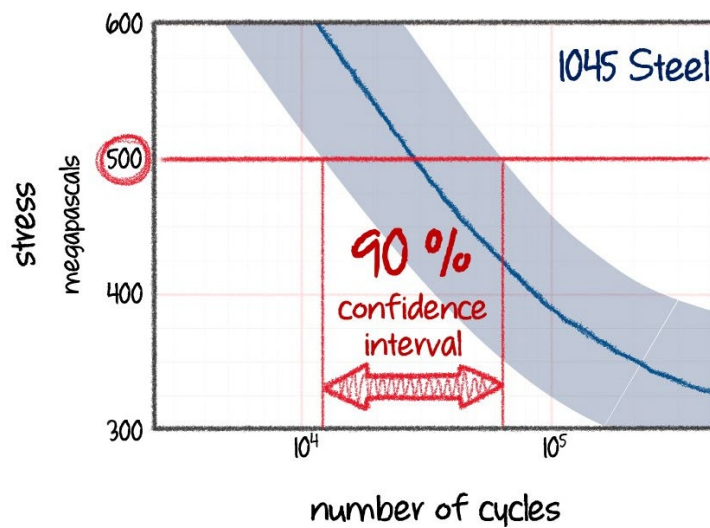
We know (from experience) that the **lognormal distribution** does a good job of modelling fatigue failure. As a fatigue crack grows bigger, it concentrates more stress at the crack tip. So the crack grows faster in a 'multiplicative' way.

We can obtain what we call an 'S-N' chart for 1045 steel which describes its fatigue characteristics.

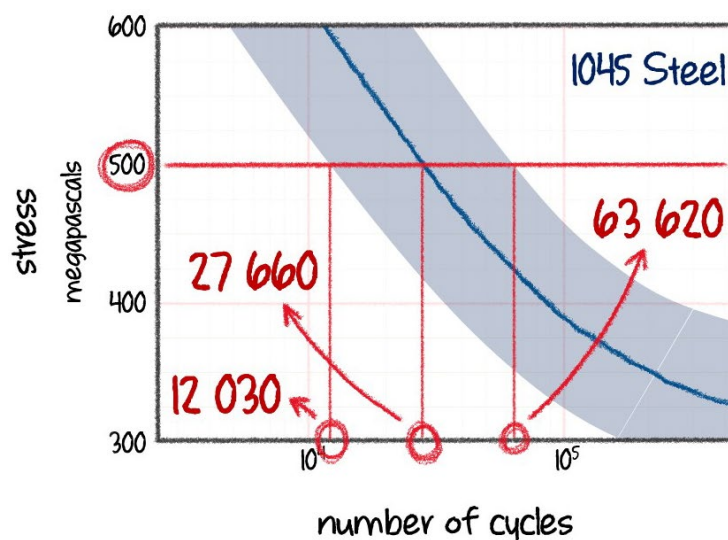


You can see that the horizontal scale of the 'S-N' chart above is **logarithmic**. And because fatigue cycles to failure is described by the **lognormal distribution**, the *logarithm* of fatigue cycles to failure should be described by the **normal distribution**.

Now let's zoom in on 500 MPa. And this allows us to identify the 90 % **confidence interval** of cycles to failure - within which we expect our 1045 steel to **fail** due to fatigue.



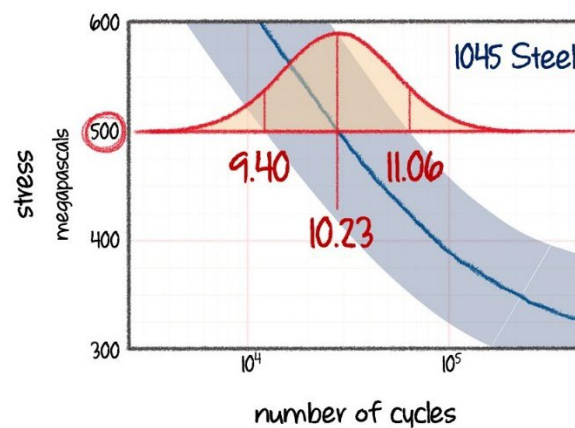
We can identify the boundaries and mid point of this interval using this chart.



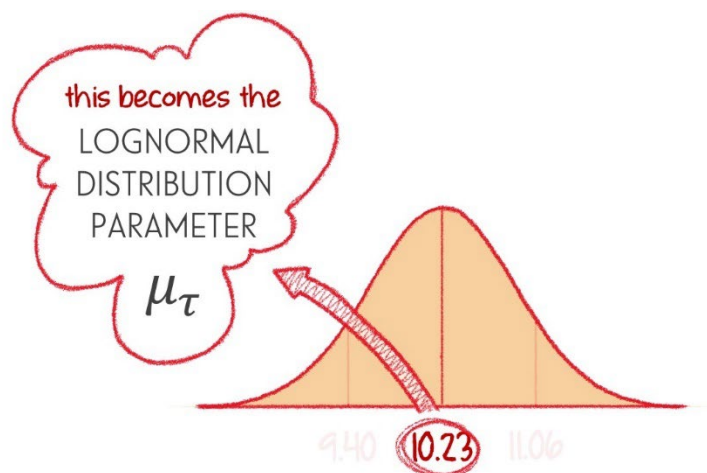
What we are going to do now is use the **normal distribution** (noting that this chart uses a **logarithmic scale**). So we can't use the cycle numbers above - we need to use the **logarithms** of these numbers

	NATURAL LOGARITHM - ln
12 030	9.40
27 660	10.23
63 620	11.06

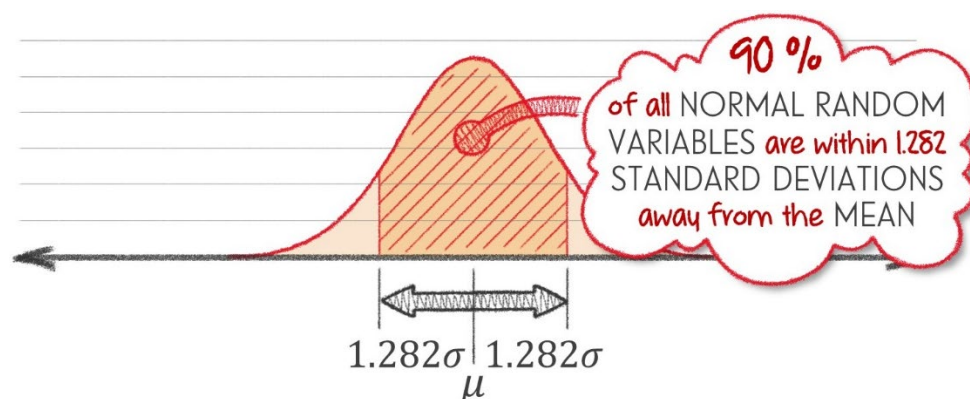
So let's superimpose a bell curve (**normal distribution**) over our **confidence interval**.



So the middle of this superimposed **normal distribution PDF curve** gives us the first parameter of our **lognormal distribution**.

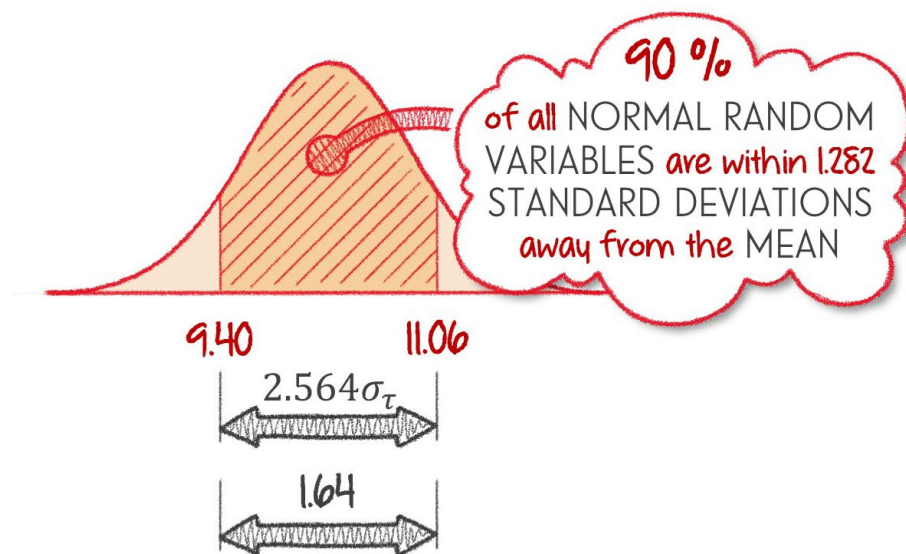


Now remember from the [normal distribution lesson](#), if we are interested in the 'middle' 90 % of **normally distributed random variables** ...



This means that the 90 % **confidence bounds** are each 1.282 **standard deviations** from the **mean**.

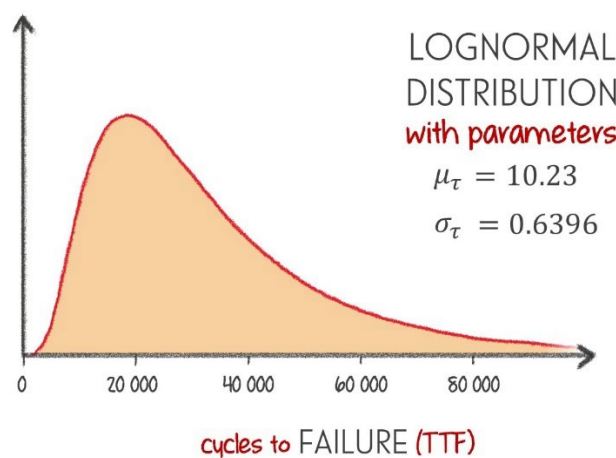
For our **normal distribution PDF curve** superimposed onto the 90 % **confidence interval** ...



This allows us to calculate the second **lognormal distribution** parameter, σ_τ ...

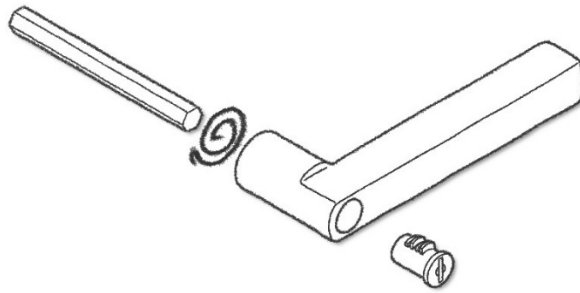
$$\sigma_\tau = \frac{1.64}{2.564} = 0.6396$$

This allows us to model the cycles to failure for our door handle if it is based on fatigue as follows ...



EXERCISE

But we don't just want to admire (measure) a problem!



What sort of 'things' could you do to the design or manufacture to reduce the likelihood of fatigue failure?

... and the FINAL WORD

The **lognormal distribution** models the TTF of products, systems and services that wear out – when *the rate of wear out accelerates in a multiplicative way*. Fatigue is a classic example of a **failure mechanism** that is really well modelled by a lognormal distribution.

The **lognormal distribution** has also empirically been found to model **Corrective Maintenance (CM) time**. Many repairs or **CM** are quick. A few repairs take a lot longer. This creates a **PDF curve** that has a left skew with a right tail.

But ... the next **probability distribution** is perhaps the 'best' yet!

LOGNORMAL DISTRIBUTION

Equations and Formulae

Parameters

μ_τ – mean of the logarithms of the random variable values

σ_τ – standard deviation of the logarithms of the random variable values

Probability Density Function (PDF)

$$f(t) = \frac{1}{t\sigma_\tau\sqrt{2\pi}} e^{-\left(\frac{\ln t - \mu_\tau}{\sigma_\tau}\right)^2} \dots \text{X} = \text{LOGNORM.DIST}([t],[\mu_\tau],[\sigma_\tau],\text{FALSE})$$

Cumulative Distribution Function (CDF)

$$F(t) = \int_0^t \frac{1}{t'\sigma_\tau\sqrt{2\pi}} e^{-\left(\frac{\ln t' - \mu_\tau}{\sigma_\tau}\right)^2} dt' \dots \text{X} = \text{LOGNORM.DIST}([t],[\mu_\tau],[\sigma_\tau],\text{TRUE})$$

Reliability Function

$$R(t) = 1 - \int_0^t \frac{1}{t'\sigma_\tau\sqrt{2\pi}} e^{-\left(\frac{\ln t' - \mu_\tau}{\sigma_\tau}\right)^2} dt' = \int_t^\infty \frac{1}{t'\sigma_\tau\sqrt{2\pi}} e^{-\left(\frac{\ln t' - \mu_\tau}{\sigma_\tau}\right)^2} dt' \\ \dots \text{X} = 1 - \text{LOGNORM.DIST}([t],[\mu_\tau],[\sigma_\tau],\text{TRUE})$$

Hazard rate

$$\lambda(t) = \frac{\frac{1}{t} e^{-\left(\frac{\ln t - \mu_\tau}{\sigma_\tau}\right)^2}}{\int_t^\infty \frac{1}{t'} e^{-\left(\frac{\ln t' - \mu_\tau}{\sigma_\tau}\right)^2} dt'} \\ \dots \text{X} = \text{LOGNORM.DIST}([t],[\mu_\tau],[\sigma_\tau],\text{FALSE}) / (1 - \text{LOGNORM.DIST}([t],[\mu_\tau],[\sigma_\tau],\text{TRUE}))$$

Mean

$$\mu = e^{\left(\mu_\tau + \frac{\sigma_\tau^2}{2}\right)} \dots \text{X} = \text{EXP}([\mu_\tau] + ([\sigma_\tau]^2)/2)$$