

HOW DO I DO MEAN TIME BETWEEN FAILURE (MTBF) TESTING?



Workbook

Presented by

Dr Chris Jackson

Reliability engineer

Copyright © 2023 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

chris.jackson@acuitas.com

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 Acuitas Reliability (www.acuitas.com) and 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.

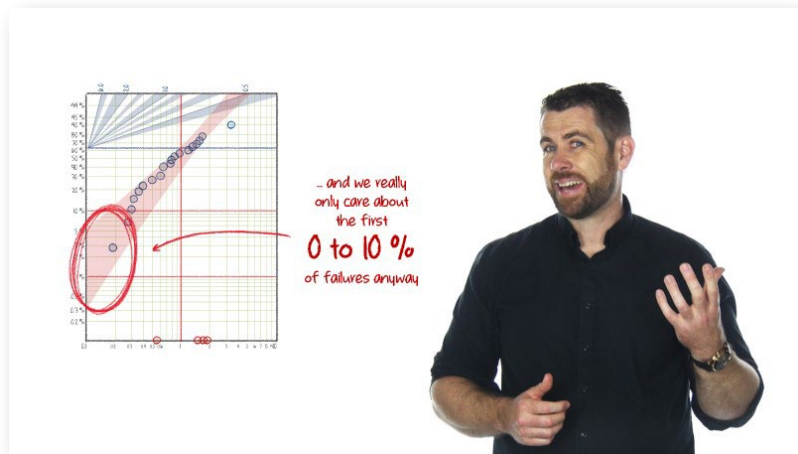


Table of Contents

MEAN TIME BETWEEN FAILURE (MTBF) TESTING WITH CONFIDENCE.....	5
... let's start with the MTBF.....	7
... so where do we start?	11
... but these tests might not be EASY TO PASS.....	13

MEAN TIME BETWEEN FAILURE (MTBF) testing with confidence



Why do we **test**? To generate **CONFIDENCE** about something. If we were already **confident**, we wouldn't need to **test**. Of course, there are **tests** imposed on us by customers and regulators. But ideally, we want to already be **confident** in the reliability or quality of our product, system or service to the extent that we aren't frightened to pass a **test** someone else tells us we need to do.

So what are we trying to be **confident** about? A **DECISION**.

Every **test** should only be conducted to inform a **decision**. If the outcomes of your **test** won't change anything ... **DON'T DO IT!** You are wasting time and money ... and perhaps generating **confidence** when you shouldn't.

Your **decision** could be defining the warranty period of your product, identifying the 'weakest part' of your manufacturing line that you need to focus on improving, coming up with availability targets for individual components and so on. If you can't make these **decisions** with the **information** you have at hand, then a **test** could generate enough information to give you confidence about what you decided to do.

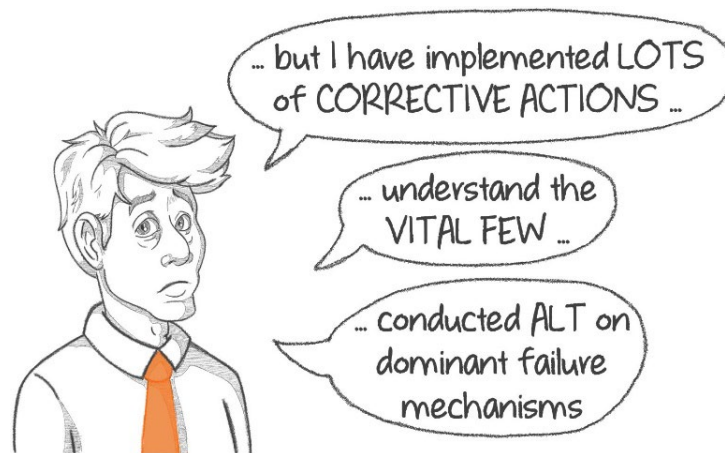
But of course, not everyone has the same **decisions** to make. So **confidence** is all about the perspective of the **decision-maker**.

Reliability and quality testing is almost a rite of passage. Or at least perceived to be something like that. In practice, **reliability testing** should always be about supporting a **decision**. Which means we work out what **decisions** we are going to make before we even consider if testing needs to be a thing.

For the start up company in the example on the previous page, they were trying to generate the **WRONG TYPE OF CONFIDENCE**. They hadn't taken **quality** or **reliability** seriously. They wanted to generate some level of very **personal confidence** that what they had done so far was OK. They weren't

genuinely interested in making the wearable smart device waterproof. With the pressure to succeed and move on quickly, launching as soon as possible was a more desirable outcome than a reliable and high-quality product. So the **decision-makers** needed a '**security blanket**' in the form of passing some sort of **credentialed test**. This allowed them to say they had 'done something' to facilitate moving on as quickly as possible.

When we are planning an *MTBF* test, we are measuring. Not improving. What does this mean?



Demonstration Testing is an approach that is used in certain scenarios and industries where assessments regarding **reliability performance** are exclusively based on a highly constrained test framework.

*Demonstration Testing is used to verify that your product, system or service
PROBABLY meets a reliability specification.*

The word 'probably' is key here. Because statistical analysis always involves at least some level of uncertainty on any inferred conclusions, we can never use a **demonstration test** to be able to 'truly demonstrate' that **reliability** exceeds a **specification**. And while there are lots of different types of **tests** that in principle can help us understand **performance** (including many that are faster and cheaper), **demonstration** is based on several 'largely unwritten' premises about what we can and can't do statistically. This means that **demonstration testing** involves the following constraints and characteristics:

1. RDT tends to focus either on MTBF (which isn't reliability) or ON-DEMAND RELIABILITY only
2. RDT focusing on MTBF assumes that the object of the test has a CONSTANT HAZARD RATE, which means it is neither wearing in or wearing out
3. ... which means that any IDENTIFIED FAILURE TRENDS NEED TO BE IGNORED, which rules out approaches like Weibull probability plotting where we try and identify a Weibull distribution that best describes the data

4 ... which further means that we have to
IGNORE INFORMATION ASSOCIATED WITH ANY APPARENT FAILURE TRENDS

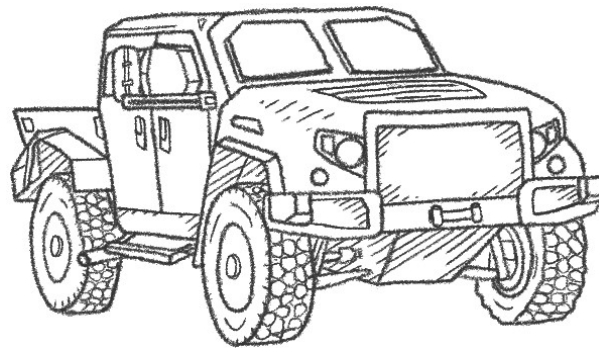
5. we need to IGNORE ANY PREVIOUS INFORMATION
that might have been gathered through observation or previous tests

6. a CONFIDENCE LEVEL needs to be specified so we can work
out what 'probably' meeting specifications mean

... let's start with the MTBF

The *MTBF* can be useful for logistics engineering and asset management (where we work out how many spare parts we need, frequency of failure for repairs and downtime, or the required fleet size,) but it has limited 'pure' **reliability applications**.

Let's look at an example military vehicle. Military customers typically specify '**reliability**' in terms of an *MTBF requirement*.

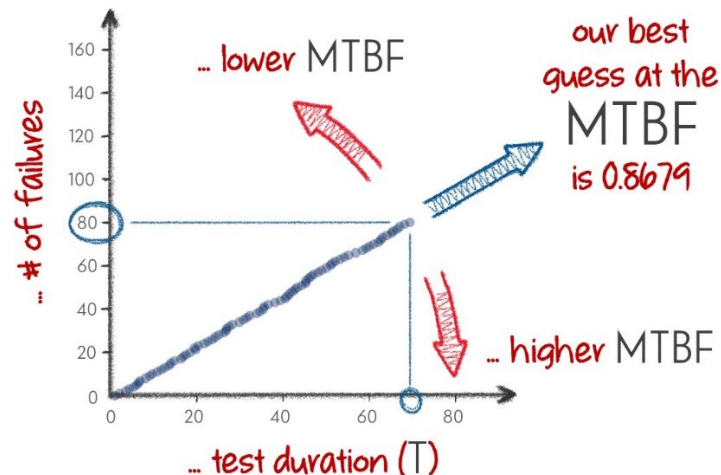


For example ...

... the vehicle *SHALL* have an MTBF of no less than 1 000 km demonstrated with a
90 % level of confidence ...

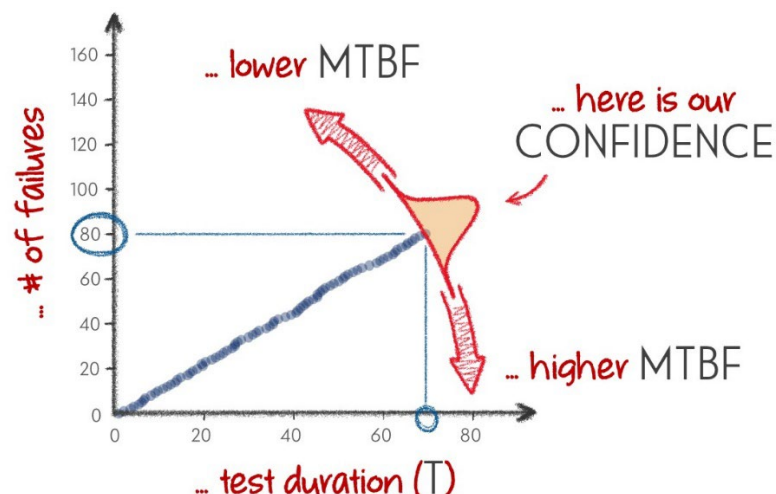
Demonstration testing consists of several prototypes being used in accordance with an agreed use case or mission profile. **Failures** are observed and recorded until we reach criteria that 'satisfies' the **requirement**.

A typical way of representing a **demonstration test plan**, and tracking progress, is with a **cumulative failure chart**. An example is illustrated below that has tracked a test that involved 80 failures.

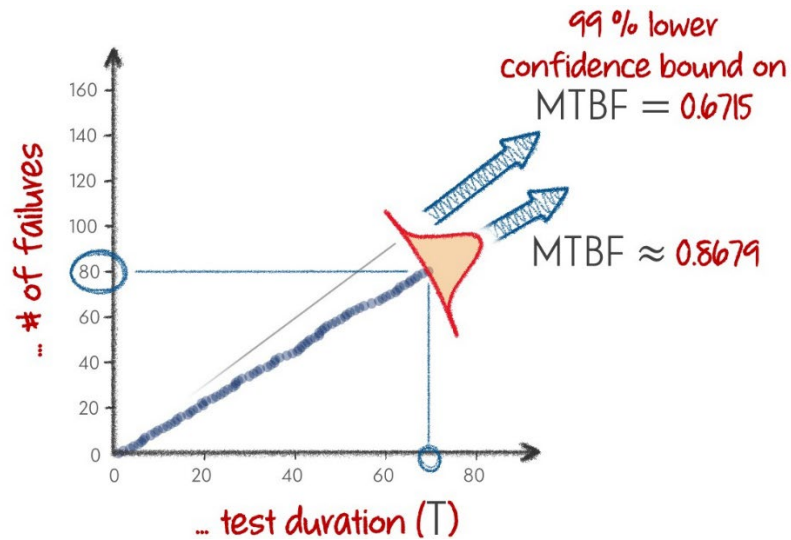


The data tends to follow a straight line (assuming system **reliability characteristics** never change) with a gradient that is equal to the inverse of the *MTBF*. In the example above, this slope tells us that our 'best guess' of the *MTBF* is 0.8679 usage units per failure. If the **cumulative failure line** has a steeper slope, it means **failures** are occurring more often. Which means we have a lower *MTBF* (bad). Conversely, if the **cumulative failure line** isn't as steep, it means **failures** are occurring less often. Which means we have a higher *MTBF* (good).

But as with any data analysis, we need to start talking about **confidence**. The confidence we have in hypothetical *MTBF* values are illustrated by the **density curve** below.

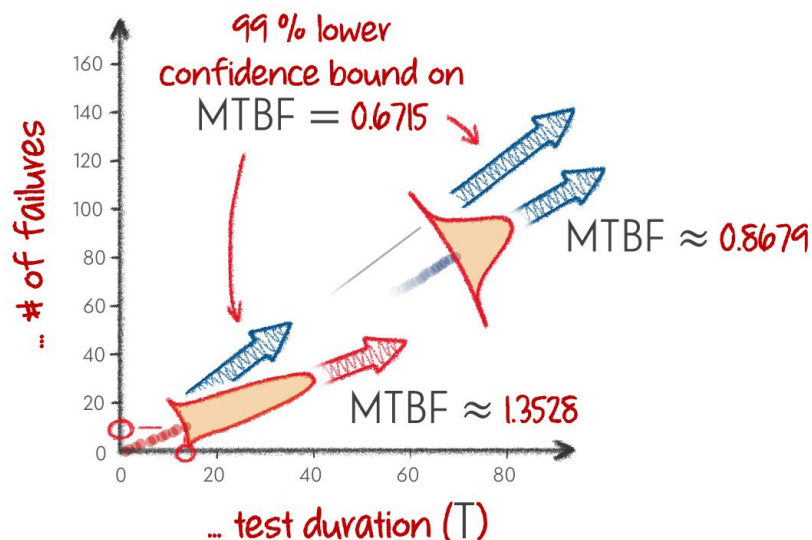


So even though we have a 'best guess' of *MTBF*, most **specifications** are interested in a value that they are reasonably **confident** is met. In other words, the **lower confidence bound**. For the example test results on the previous page, the 99 % **lower confidence bound** is shown in the following illustration.



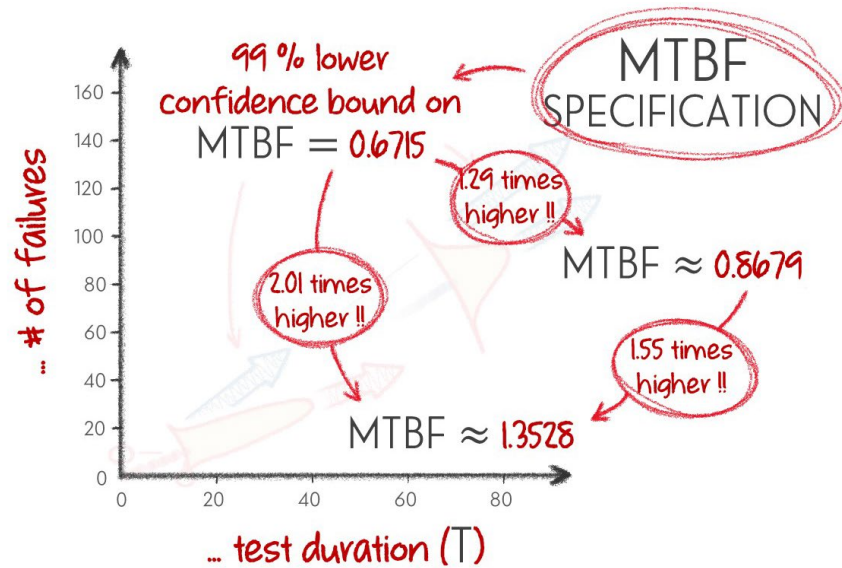
As you would expect, the **MTBF 99 % lower confidence bound** (0.6715 usage units per failure) is lower than the **MTBF 'best guess'** (0.8679 usage units per failure).

But 80 **failures** represents a LOT of information. More information than most **demonstration tests** are going to provide. So let's look at another **demonstration test** that was a lot shorter and only involved 10 failures.



You can see that even though the shorter test also involves the same **MTBF 99 % lower confidence bound**, the item being tested needed a very high **ACTUAL MTBF** to take into account less information (and more uncertainty).

The effects can be considerable.



So to 'demonstrate' the same *MTBF* with 99 % **confidence**, items involved in **shorter duration tests** need to be more reliable.

This is one of the big problems with **demonstration tests**. So we have some more additions to the first six **demonstration test** characteristics and constraints we looked at five pages ago.

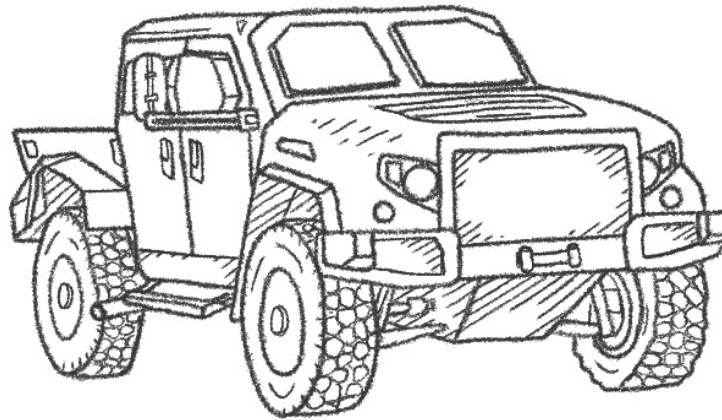
7. items under test need to have
RELIABILITY CHARACTERISTICS THAT EXCEED THE SPECIFICATION
to successfully demonstrate reliability

8. DECREASING RDT DURATION means you need to either
INCREASE THE RELIABILITY OF THE ITEM or
DECREASE YOUR REQUIRED CONFIDENCE LEVELS

... so where do we start?

We can't just test until we are happy. We need test facilities to be booked in advance, operators identified and all sorts of other things included in the budget. So how do we do this?

Let's start with the following requirement ...



... the military vehicle shall have an MTBF of no less than 1 000 km demonstrated with an 80 % level of confidence ...

So our **required MTBF** is ...

$$MTBF_{Req} = 1\,000\text{ km}$$

The first thing we do is define **decision risk**. If the specification requires an 80 % **level of confidence**, this means we are willing to accept a 20 % chance that an unreliable item is incorrectly accepted. This is called the **consumer's risk** ... which is represented with the lowercase Greek letter β (... note that we only have a certain number of Greek letters so we sometimes need to recycle them!)

$$\beta = 20\%$$

The next thing we do is focus on a test which involves zero allowable failures. What need to come up with a **test duration** (T) where by this point, if our item has NOT FAILED, we pass the test.

We can rearrange the **Poisson distribution PMF** equation (outside of this course!) to calculate the **test duration** ...

$$T = \frac{MTBF_{Req} \times F_{\chi^2}^{-1}(1 - \beta; 2r + 2)}{2}$$

... where $\chi^2_{p;2r+2}$ is the chi-squared random variable with a **CDF** of p and $2r + 2$ degrees of freedom.

The Microsoft Excel function that evaluates this **equation** is ...



$$=[MTBF_{REQ}] * CHISQ.INV(1-[\beta], (2*[r])+2)/2$$

So for zero allowable failures, we need a **test duration** of 1609.4 km to demonstrate with 80 % confidence that the **MTBF** exceeds 1 000 km.

EXERCISE

... get your **MTBF TEST PLANNER (EXCEL)** OPEN.

... let's create a test plan
where we demonstrate
1 000 km MTBF
with
80 % CONFIDENCE



Navigate to the tab '**3. MTBF Test Planner**' and enter the following data in **STEP 1**.

STEP 1: ENTER MTBF SPECIFICATION

MTBF requirement ($MTBF_{REQ}$):
Required Confidence Level (CL):

1,000
80%

... based on your inputs above, the specification that these test plans will verify is ...

... the item **SHALL have an MTBF of NO LESS THAN 1,000 demonstrated with 80.00% CONFIDENCE** ...

Then scroll down to **STEP 2**. This allows you to examine different test durations based on different allowable numbers of failures.

STEP 2: EXAMINE POTENTIAL TEST PLANS

Enter hypothetical 'allowable numbers of failures' below to identify the corresponding test durations (in terms of usage units). These will all be potential test plans to verify your specification described above.

Allowable number of failures (r)
0
1
2
3
4
5
6

Test Duration (T)
1,609
2,994
4,279
5,515
6,721
7,906
9,075

... but these tests might not be EASY TO PASS

But remember how we need to have a 'more' **reliable** item to have a reasonable chance of passing a short **demonstration test**? ... and a test that involves zero allowable **failures** is the shortest **demonstration test** possible?

The factor by which the *MTBF* of the item under test exceeds the **specification** is called the **Discrimination Ratio (DR)**. For example, if the **specified MTBF** is 1 000 hours, then a product with a *DR* of 2.0 is one that has a **MTBF design goal** of 2 000 hours ... *specifically to pass the demonstration test!* We can then use this 'more reliable' hypothetical device to understand the probability of it *failing* our **demonstration test**. We call this the **Producer's Risk**.

If we are paying a contractor to manufacture a military vehicle, we need to work out the chance it has of passing our **demonstration test**. And this is where that *DR* matters.

Let's assume that $DR = 2$. That means our contractor is aiming to provide us with a military vehicle that has an *MTBF* of 2 000 km (which is twice as large as our **required MTBF**). So if we have a test duration of 1609.4 km, and we aren't allowed any failures, how do we work out the probability of it passing the **demonstration test**?

EXERCISE



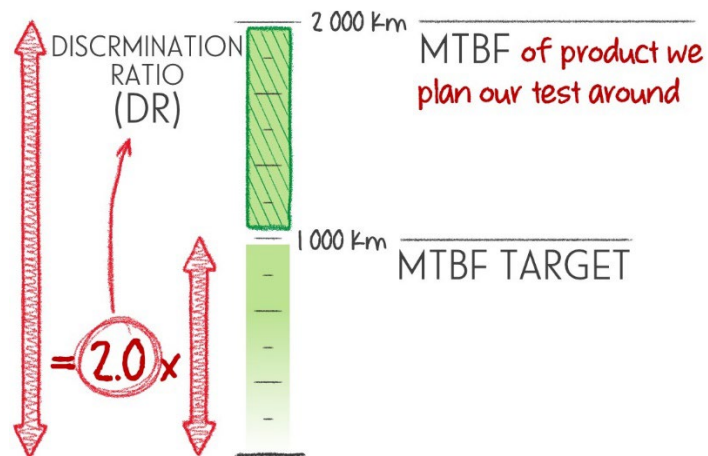
What is **PRODUCER's RISK** and why is it important here?

EXERCISE

... let's plan
our test using a
**DISCRIMINATION
RATIO of 2.0**



What does a **DISCRIMINATION RATIO** of 2.0 mean?



Now we need to use our **Poisson distribution** to understand the probability of our 'more reliable' failing any test in order to quantify the **Producer's Risk**. And the reason we can do this is because we are assuming a constant **hazard rate** (!!!). We will expect the following number of **failures** during our **demonstration** (duration 1609.4 km) for a vehicle with an *MTBF* of 2 000 km ...

$$\mu = \frac{T}{DR \times MTBF_{Req}} = \frac{1609.4}{2000} = 0.804719$$

So the probability of passing the **demonstration testing** is the probability of observing zero (or fewer) failures, which is simply the **Poisson distribution CDF**.

$$\text{probability of passing} = F_{Poisson}(0; \mu)$$

Which means that we can work out the probability of this vehicle also failing the **demonstration test** ...

$$\text{probability of failing} = 1 - F_{Poisson}(0; \mu)$$

We call the probability of failing the **producer's risk**, which we represent with the lowercase Greek letter α . So for an **demonstration test** that is 1609.4 km long, with zero allowable failures, and a vehicle that has an *MTBF* of 2 000 km, there is a 55.28 % probability of it failing the test. This is too high!

The Microsoft Excel function that gives us the **producer's risk** for r allowable failures is ...



**=1-POISSON.DIST([r],
[T]/([DR]*[MTBF_{REQ}]),TRUE)**

This allows us to assume an allowable number of **failures** (r), calculate the corresponding **RDT duration** (T), and use the **DR** to work out the **producer's risk** (α).

EXERCISE

Complete the following table for an MTBF requirement of 1 000 km
(using the Excel equation below) ...



=1-POISSON.DIST([r],
[T]/([DR]*[MTBF_{REQ}]),TRUE)

Discrimination Ratio (<i>DR</i>)	Actual MTBF ($MTBF_{REQ} \times DR$)	Probability of PASSING a test with zero allowable failures over 1 609.4 km	Probability of FAILING a test with zero allowable failures over 1 609.4 km
1.0	1 000 km		
1.5			
2.0	2000 km	44.7 %	55.3 %
3.0			



... this is called the
PRODUCER's RISK (α)

So if the **producer's risk** (α) is too high, we increase the number of allowable failures and repeat the process. This gives us the following values

... # of allowable failures (r)	... test duration (T)	... producer's risk (α)
0	1 609.4	55.3%
1	2 994.3	44.1%
2	4 279.0	36.1%
3	5 515.0	29.9%
4	6 721.0	24.9%
5	7 906.0	20.8%
6	9 075.4	17.4%
7	10 232.5	14.6%
8	11 379.8	12.2%
9	12 518.8	10.3%

So when do we stop? As a rule, we keep increasing **test duration** until the **producer's risk** is smaller than or equal to the **consumer's risk**. Of course, this is all premised on the contractor (producer) being able to produce something that has an *MTBF* that significantly exceeds the **requirement** (by a factor of *DR*). In this case, this occurs when we have six allowable failures with a **test duration** of 9 075.4 km. So this becomes the planned **demonstration test duration** (T).

There are ALL SORTS OF ISSUES with **demonstration testing** that start with discarding lot's of existing information, dismissing any apparent information (failure characteristics), assuming a constant **hazard rate**, and relying on a contractor (producer) to make something that exceeds **requirements**. So you should always approach **demonstration testing** cautiously, knowing that many of its assumptions will cost you time and money.

EXERCISE

... get your MTBF TEST PLANNER (EXCEL) OPEN.



Navigate to the tab '3. MTBF Test Planner' and enter the following data in STEP 3.

STEP 3: (OPTIONAL) EXAMINE CHANCES OF PASSING THE TEST

Enter assumed an assumed Discrimination Ratio and producer's risk below to identify the shortest test plan that provides a 'reliable' item a reasonable chance of passing that test.

Discrimination Ratio (DR):
Producer's risk (α):

2.0
20.0%

Based on your entries above, the following statement applies to the test plan below:

... if the item has an ACTUAL MTBF of 2,000 then it will have AT LEAST a 80.00% chance of passing the test plan described below ...

Allowable number of failures (x)
6

Test Duration (T)
9,075

You can see that it recommends the test duration we calculated before (it is rounded to the nearest integer). To ensure that we have a test plan that meets the Producer's Risk requirements, enter this duration into STEP 4. Note that because the value in STEP 3 was 'rounded down,' you might need to enter 9 076 in order to get the 6 allowable failures identified above.

STEP 4: SPECIFY TEST DURATION

Based on the information above ... what is the MAXIMUM duration of your test?

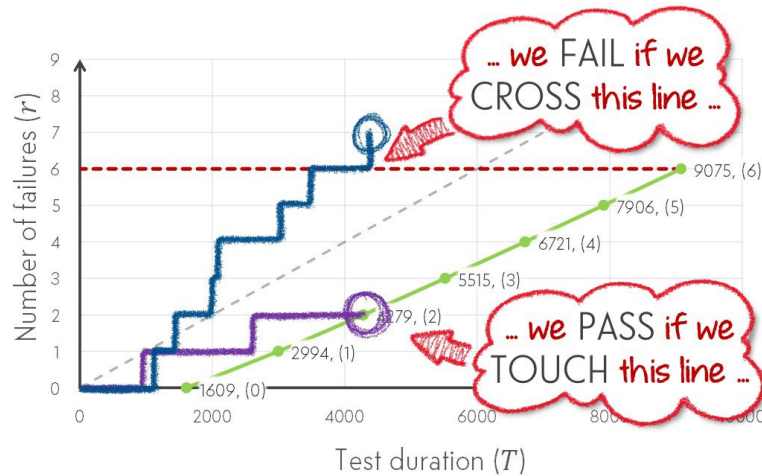
9,076

FINAL TEST PLAN ...

Allowable number of failures (x)	Number of trials or tests required (n)
6	9,075

EXERCISE

You will now get the following test plan on the tab
'4. Test Plan and Progress Chart.'



This test plan ensures that if any product 'passes' the test, it will meet the MTBF requirements (1 000 km with 80 % confidence) outlined previously. It also ensures that a product with an actual MTBF of 2 000 km (Discrimination Ratio of 2.0) will have an 80 % probability of passing.



Navigate to the tab '3. MTBF Test Planner' and enter the following data in STEP 5.

STEP 5: SPECIFY EARLY TERMINATION CRITERIA (OPTIONAL)

How CONFIDENT do you want to be that the system DOES NOT meet the requirement to terminate the test early?
(leave empty if you do not want early termination)

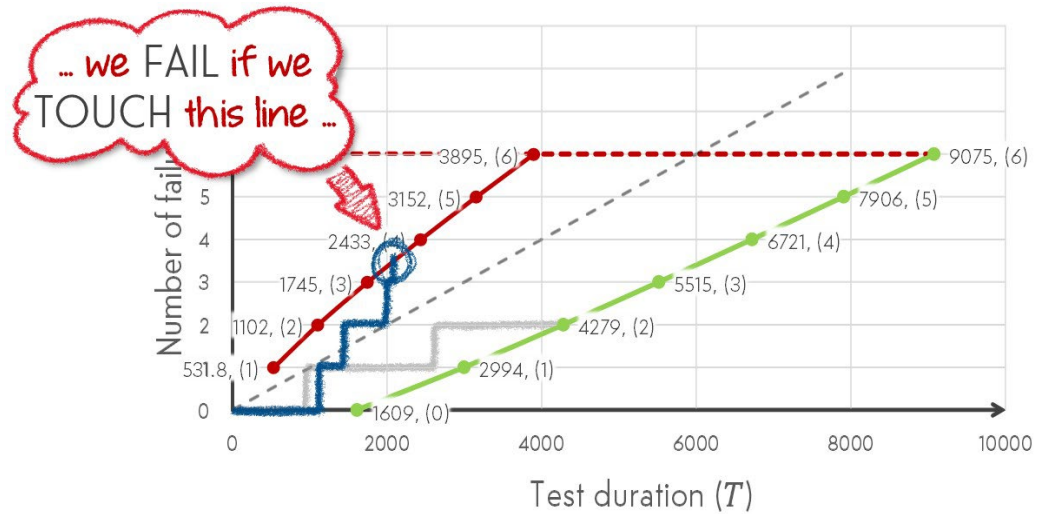
Early termination criteria (confidence of NOT meeting the requirement)

80.00%

... the test will be TERMINATED if we are 80.00% CONFIDENT that the item WILL NOT meet the requirement described above ...

EXERCISE

You will now get the following test plan on the tab
'4. Test Plan and Progress Chart.'



This new demonstration test plan now allows 'early termination' if the product doesn't perform well early. This saves time and money (including in terms of the number of prototypes or units under test).

