

Discrete Distributions and How to Use them for Fleet Planning and Spare Parts Analysis

Accendo Reliability Webinar 25th October 2021



Workbook

Presented by

Dr Chris Jackson

Reliability engineer



Copyright © 2021 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.

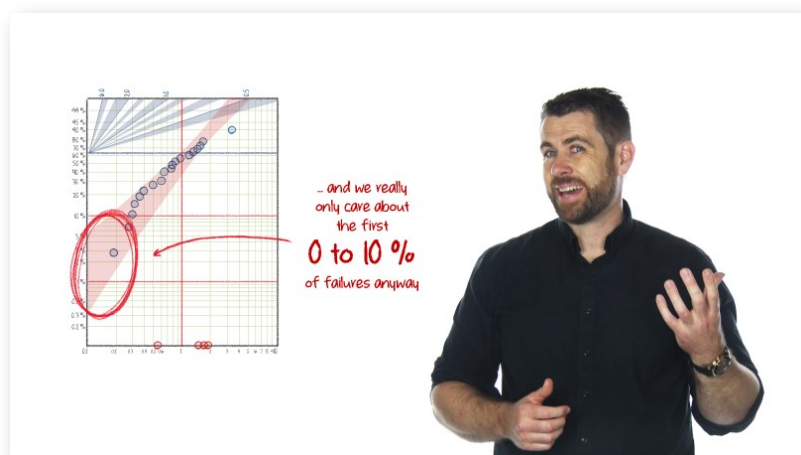
For permission requests, write to Christopher Jackson using the details on the following pages.

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.

A lot of reliability engineering focuses on what we call '**continuous random variables**.' These are **random variables** that represent some sort of quantity that can have an infinite number of possible values over some interval. The most common **continuous random variable** we use in **reliability engineering** is **Time to Failure (TTF)**. But that is not the only **random variable** we are interested in.

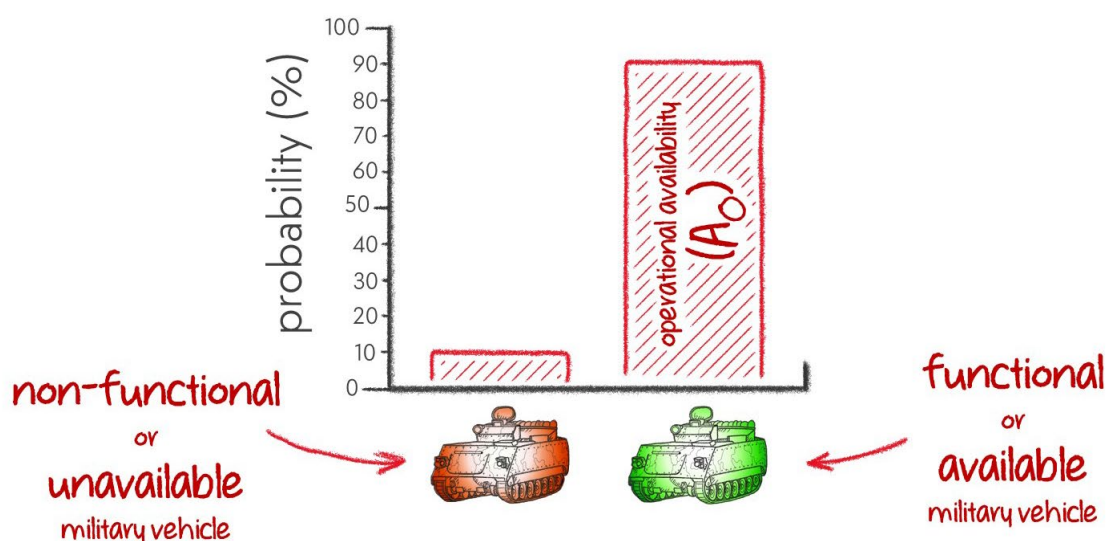
Sometimes we want to 'count' the number of things that are working. This could be the number of '**available**' items you have in a fleet of equipment. Or the number of '**functional**' items at the end of a **reliability demonstration test**. Or how many spares you need over a calendar year. In scenarios like this, we need to move away from **continuous random variables** and start focusing on what we call '**discrete random variables**.'

... the BINOMIAL DISTRIBUTION

Let's say you have a fleet of 'things.' Like Armoured Personnel Carriers (APCs).

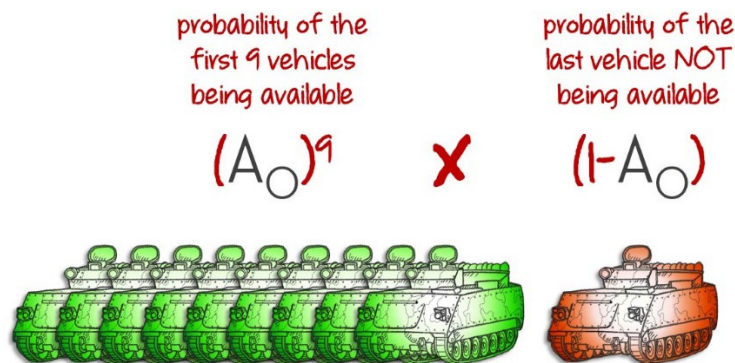


This means that if we were to select an APC at random, there is a 90 % chance that it is working. Another way of looking at it is ...

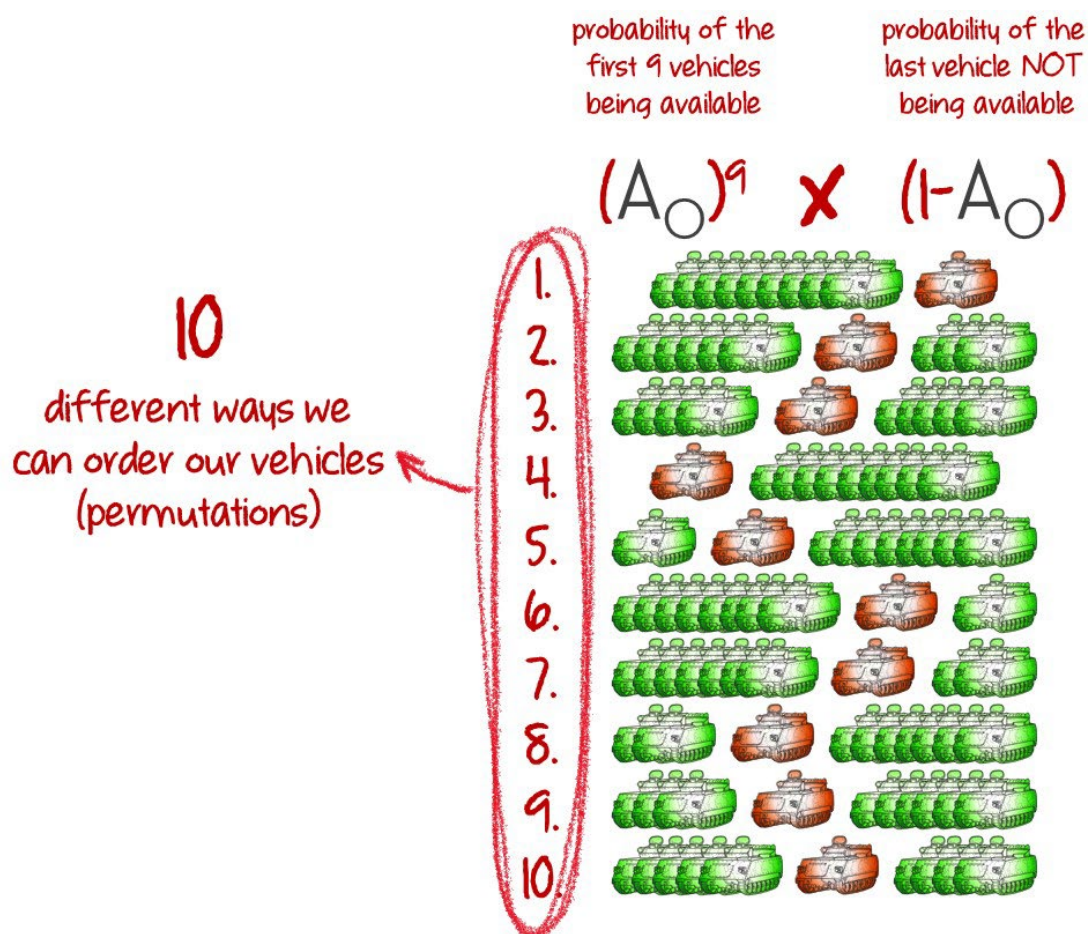


We typically need a certain fraction of our fleet or number of 'things' to be working to do our job. For example, a mechanised infantry company might have a total of ten APCs and can only launch a mission if nine are **available**. We obviously are interested in knowing the **probability** of having nine APCs **available** so we can launch our mission.

The first step in working this out is just focusing on finding the probability of EXACTLY nine APCs being **available**. So let's imagine a scenario where we happen to select (at random) APCs from our fleet of ten one at a time AND the first nine we selected happened to be **available** AND the last or tenth APC being **unavailable**.



This is ONLY one scenario of MANY POSSIBLE scenarios where we select exactly nine **available** APCs and one **unavailable** APC. The MANY POSSIBLE scenarios are ...



So for our fleet of ten APCs, the **probability** of EXACTLY nine APCs being **available** is ...

$$\begin{aligned} & 10 \times (A_o)^9 \times (1 - A_o) \\ & = 10 \times (0.9)^9 \times (1 - 0.9) \\ & = 10 \times (0.9)^9 \times 0.1 = 38.74 \% \end{aligned}$$

When we follow this logic when trying to understand the **probability** of any particular number of APCs being **available** (let's call it ' x '), we essentially get ...

(the number of ways we can order these ' x ' available vehicles)

\times

(the probability of the FIRST ' x ' vehicles being available)

\times

(the probability of the LAST ' $n - x$ ' vehicles being UN-functional)

This can be summarized in the following **equation** that defines what we call **the Probability Mass Function (PMF)**. We often look at scenarios like this as a number of '**trials**' that have that can either be '**successful**' or '**unsuccessful**.' For our APCs above, each 'trial' is the act of walking up to an APC at random and identifying if it is **available** or **unavailable**.

$$f(x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

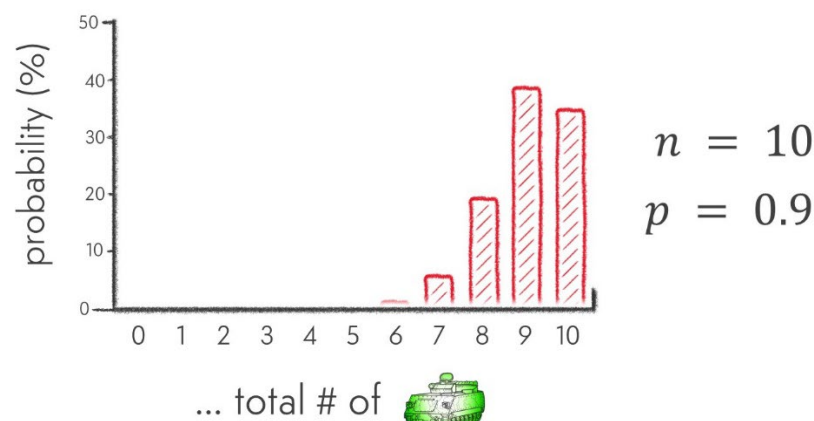
the probability of the first ' x ' things being successful

the number of ways we can order ' x successful' things out of a total of ' n ' things

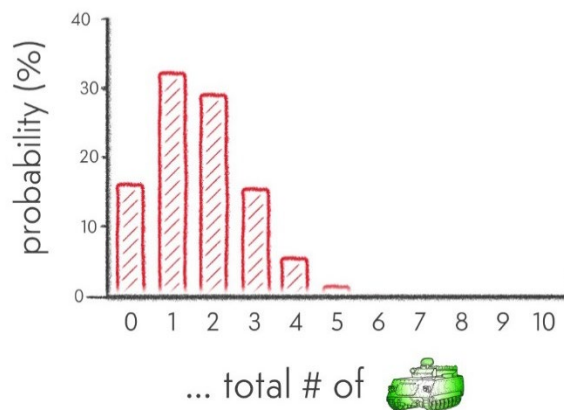
the probability of the last ' $n - x$ ' things being UN-successful

... remembering that $\binom{n}{x} = \frac{n!}{x!(n-x)!}$.

The **PMF** simply gives us the **probability** that **random variable**, X (the number of 'successful trials') is equal to a specific value, x .




And for a 'probability of success' which describes a low **operational availability** of 16.67% ...



$$n = 10$$

$$p = \frac{1}{6}$$

This is called the **binomial distribution**. And the Microsoft Excel function that give us the **PMF** is ...

 = binom.dist
([x], [n], [p], FALSE)

The '**number of available APCs**' is an example of a **discrete random variable** as the values can only be one of a set of possible values.

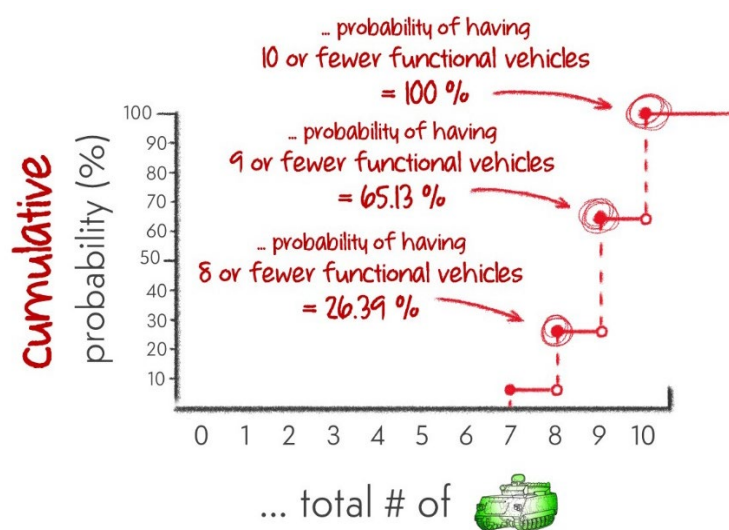


A **PMF** (for a **discrete random variable**) is not the same as a **PDF** (for a **continuous random variable**). Why? Well, a **discrete random variable** does have a theoretical chance of assuming a particular value. For example, we worked out that there is 38.74 % chance of nine APCs being **available**. But the probability that the **Time to Failure (TTF** – a **continuous random variable**) is EXACTLY nine years (for example) is infinitesimally small. Even if your maintenance records state that the system failed at exactly nine years, in practice it will be 9.000012 ... (followed by an infinite number of digits) that could be very close to nine, but not exactly nine.


But while the **continuous random variable PDF** might be different to the **discrete random variable PMF**, the **Cumulative Distribution Functions (CDFs)** measure the same thing. But ... the **discrete random variable CDF** goes up in 'steps' where the **continuous random variable CDF** is smooth. This is because **discrete random variables** can only be one of a set of allowable values. But because they are **CDFs** (and describe the probability that the **random variable value** is less than or equal to a specified value), they both start at 0 and finish at 1 or 100 %. In the case of the **binomial distribution**, the **equation** for the **CDF** is ...

$$F(x) = \Pr(X \leq x) = \sum_{x'=0}^x f(x') = \sum_{x'=0}^x \binom{n}{x'} p^{x'} (1-p)^{n-x'}$$

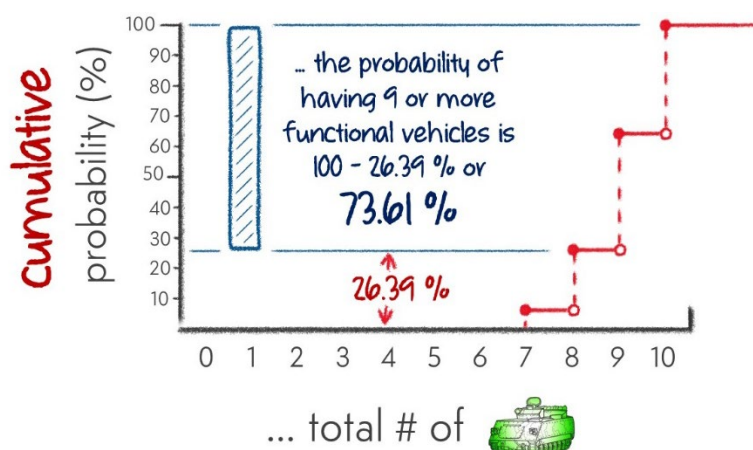
The **CDF** for the number of **available APCs** is illustrated below.



The Microsoft Excel function for the **binomial distribution CDF** is illustrated below. The only difference between this **function** and the function for the **PMF** is that the last argument is 'TRUE.'


= binom.dist
([x], [n], [p], TRUE)

And ... we can manipulate the **CDF** to help us answer certain questions.




EXERCISE

From the chart on the previous page, we can tell that there is a 0.736099 or 73.6099 % probability that if we select 10 APCs, 9 or more of them will be available.

But what happens if we want to be 100 % certain that we will have 9 or more functional APCs out of a number we select?
... how many do we need to select?

We can use the Microsoft Excel function above to help us. For example, we can use the following Microsoft Excel function to calculate that there is a 0.910438 or 91.0438 % probability that if we select 11 APCs, 9 or more will be available:

 = 1 - binom.dist
(8, 11, 0.9, TRUE)

The reason we see '8' in the function is that this part of the function is used to calculate the probability of having 8 or fewer APCs functional when we select 11. You can see that this is subtracted from 1 to give us the probability we are after. We can keep doing this for any number of selected APCs ...

Number of APCs selected	Probability of 9 or more being functional
10	0.736099
11	0.910438
12	0.9743625
13	0.9935398
14	0.9985259
15	0.9996894
16	0.9999387
17	0.9999885
18	0.9999980
19	0.9999996
20	0.9999999
21	0.99999999
22	0.999999999
23	0.9999999998
24	0.99999999997
25	0.999999999995
26	0.999999999999
27	0.9999999999999
28	0.99999999999999

EXERCISE

The problem is that no matter how many APCs we select, there is always a finite chance that all of them are unavailable. This chance gets smaller and smaller as we select more and more. But it never goes away.

All we can do is specify a **CONFIDENCE** of having 9 or more functional APCs. For example, if we want to be 95 % confident that we will have 9 or more functional APCs, we need to select at least 12 APCs.

Number of APCs selected	Probability of 9 or more being functional
11	0.910438
12	0.9743625
13	0.9935398

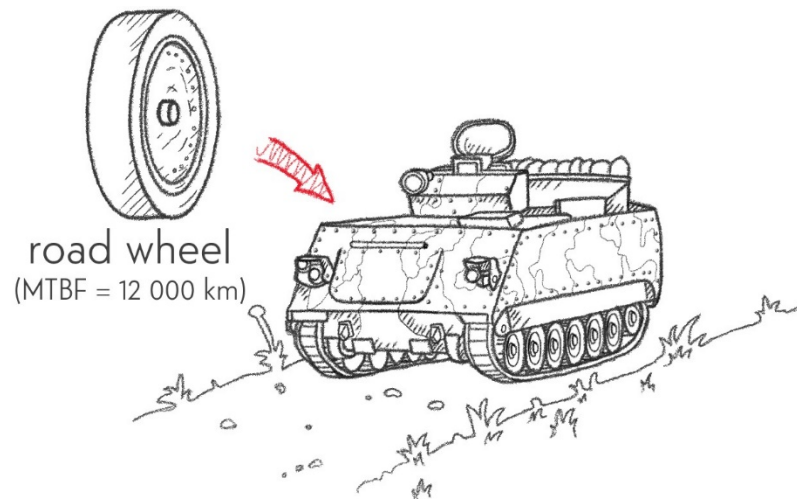
All we can do is specify a **CONFIDENCE** of having 9 or more functional APCs. For example, if we want to be 99 % confident that we will have 9 or more functional APCs, we need to select at least 13 APCs.

Number of APCs selected	Probability of 9 or more being functional
12	0.9743625
13	0.9935398
14	0.9985259

So you if you aren't comfortable specifying a confidence yet
... then you need to **START FINDING A WAY!**

... the POISSON DISTRIBUTION

APCs have 'road wheels.' These are wheels that have solid rubber tires that travel within and along the tracks of the APC. They do eventually wear out and need to be replaced.



So let's say that we know this APC (which has 12 road wheels) needs to complete a 2 000 km journey. How many SPARE road wheels do we need? Another way of looking at this is that the total distance the road wheels travel is 24 000 km.

Given the road wheel has an **MTBF** of 12 000 km, we would EXPECT 2 of them to fail (if they fail at a constant **hazard rate**). But this doesn't mean it isn't possible for 3, 4, 5 or more road wheels to fail. Remember – **failure is a random process**.

If we assume a constant **hazard rate** (which we should rarely do), we can work out that if **random variable** X is the number of events that will occur in a specific interval, then the **PMF** for this **random variable** is ...

$$f(x) = \Pr(X = x) = e^{-\mu} \frac{\mu^x}{x!}$$

... where μ is the expected number of events for that interval.

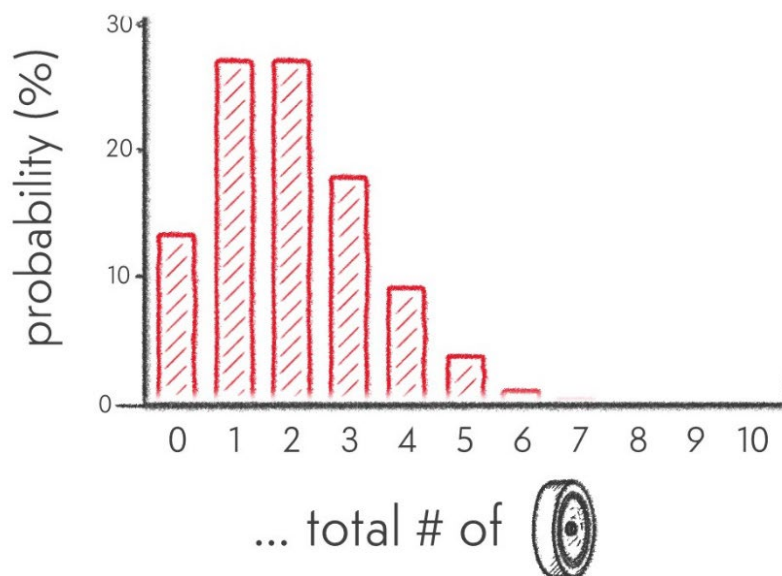
So why can we assume the constant **hazard rate** ... sometimes? Well, for this APC, the road wheels will be of all different ages and stages of deterioration. This might mean some are 'close to the end' and others might be 'brand new.' So because these road wheels are not all starting in the same condition, and IF their ages are random, then we can expect them (as a group) to fail with a constant **hazard rate**.

The **PMF equation** above describes what we call the **Poisson distribution**. The **Poisson distribution** models the number of events that occur at a constant rate in specified interval.


The Microsoft Excel function for the **Poisson distribution PMF** is ...

 `= poisson.dist
([x], [μ], FALSE)`

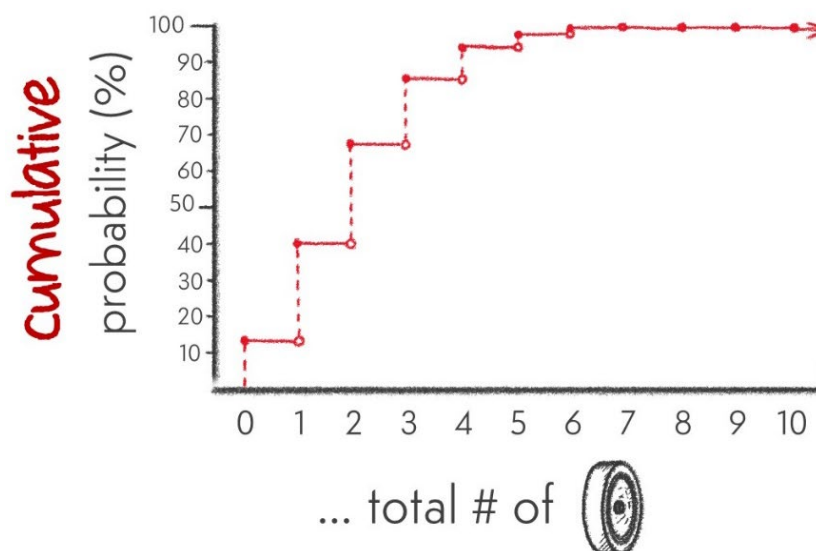
So for our single APC that is expecting 2 road wheels to fail over its 2 000 km journey, the **PMF** is illustrated below ...



The Microsoft Excel function for the **Poisson distribution CDF** is ...

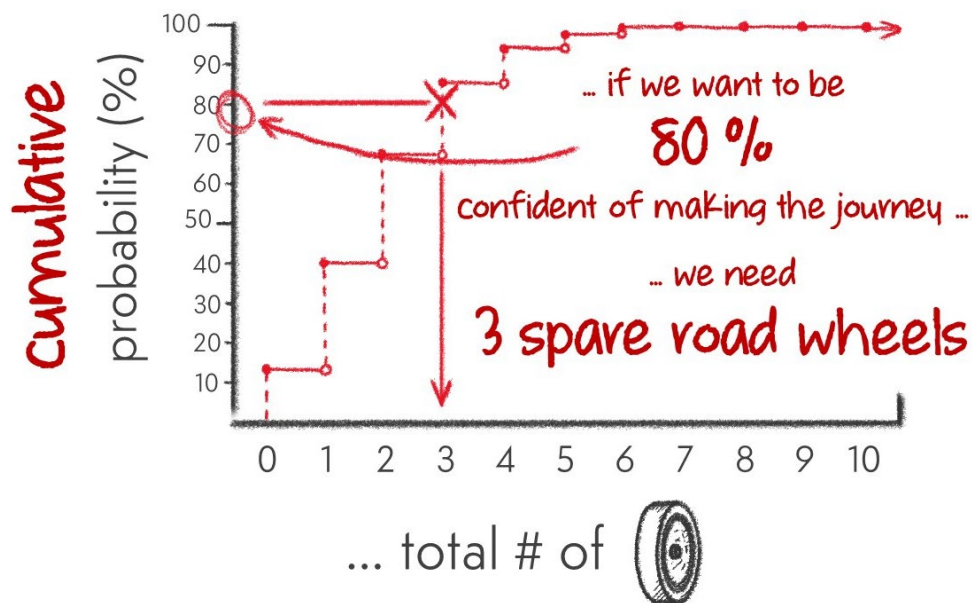
 `= poisson.dist
([x], [μ], TRUE)`

... and the corresponding **CDF** is illustrated below ...

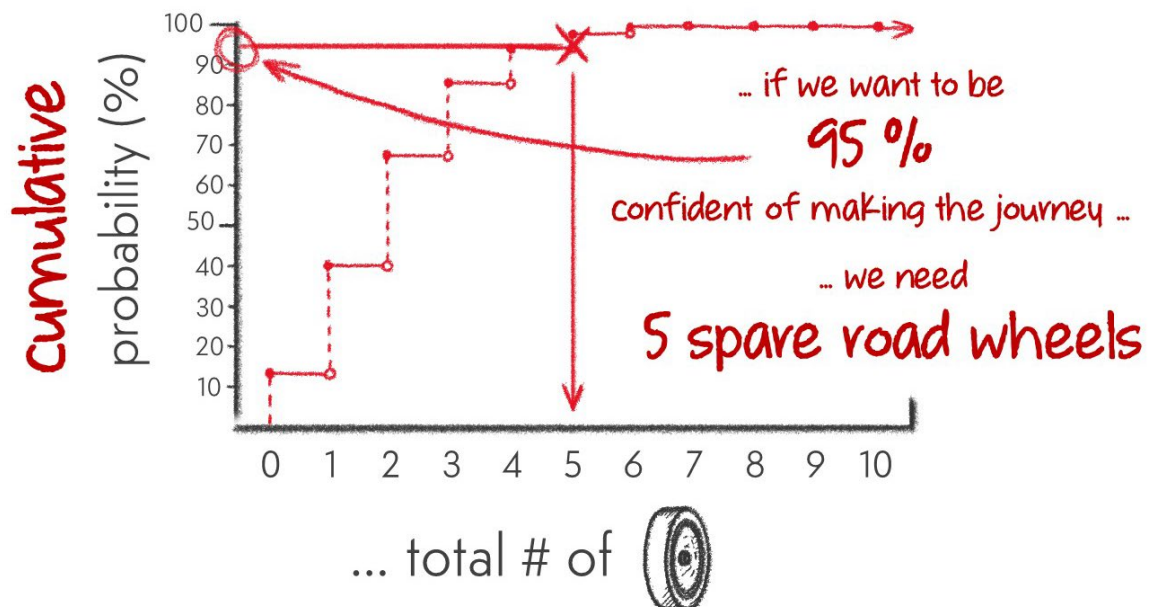


You might remember that when we were trying to work out how many APCs we needed to select to guarantee that at least nine would be **available** ... we couldn't select enough! It is not possible. There was always a finite chance of most of our fleet being **unavailable**. And a similar scenario applies for the road wheels. We cannot have enough spare road wheels to be 100 % certain that we can travel 2 000 km without running out. Again, we need to deal with confidence and risk.

We can use our **CDF** to work out that we need to have at least 3 spare road wheels to be at least 80 % confident that we complete our 2 000 km journey.

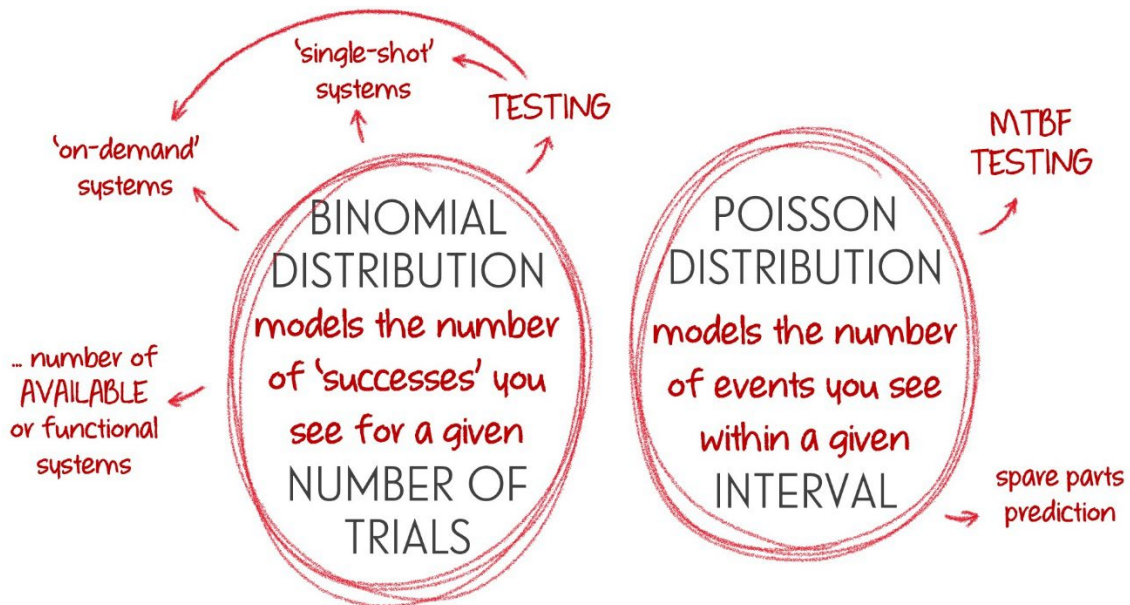


And likewise if we want to be 95 % **confident** that we complete out 2 000 km journey ...



... so know your DISCRETE DISTRIBUTIONS

... and when to use them. There are more **discrete distributions** than just the **binomial** and **Poisson distributions**. But they are very important for **reliability engineers**.



They can help you work out really important things like how many systems are functional or how many spare parts you need. And they are very important for testing ... especially when tests involved 'numbers' of **failures** observed.