Outline Solutions to Tutorial Sheet 3

1. Not true! The vast majority of people have both eyes the same colour so if the left eye is brown the right eye will almost certainly also be brown. So the probability that both eyes are brown is also about 0.6.
   More formally, the events are *not independent* so
   P(left eye brown **and** right eye brown) ≈
   P(left eye brown)×P(right eye brown **given** left eye brown) = 0.6×1 = 0.6

2. From the tree diagram:

		Joint Probability
P($MachineA \cap Defective$)	= 0.45×0.08	= **0.036**
P($MachineA \cap NotDefective$)	= 0.45×0.92	= **0.414**
P($MachineB \cap Defective$)	= 0.55×0.10	= **0.055**
P($MachineB \cap NotDefective$)	= 0.55×0.90	= **0.495**
	TOTAL	= 1.000

   Prob(item from Machine A **given** it's defective)

$$= P(A \mid Defective) = \frac{P(A \cap Defective)}{P(Defective)} = \frac{0.036}{0.036+0.055} = 0.396$$

3. Let X be the number of defective bulbs found.
   (i) $P(X=0) = \frac{6}{9}\times\frac{5}{8}\times\frac{4}{7} = \frac{120}{504} = 0.238$
   (ii) $P(X=1) = \left(\frac{3}{9}\times\frac{6}{8}\times\frac{5}{7}\right) + \left(\frac{6}{9}\times\frac{3}{8}\times\frac{5}{7}\right) + \left(\frac{6}{9}\times\frac{5}{8}\times\frac{3}{7}\right) = 3\times 0.1786 = 0.536$
   (iii) $P(X \geq 1) = 1 - P(X=0) = 1 - 0.238 = 0.762$

4. It helps to draw a tree diagram first. Then:

$$P(\text{Has disorder} \mid + \text{result}) = \frac{P(D\cap +)}{P(+)} = \frac{0.05\times 0.95}{(0.05\times 0.95)+(0.95\times 0.10)} = \frac{0.0475}{0.1425} = 0.333$$

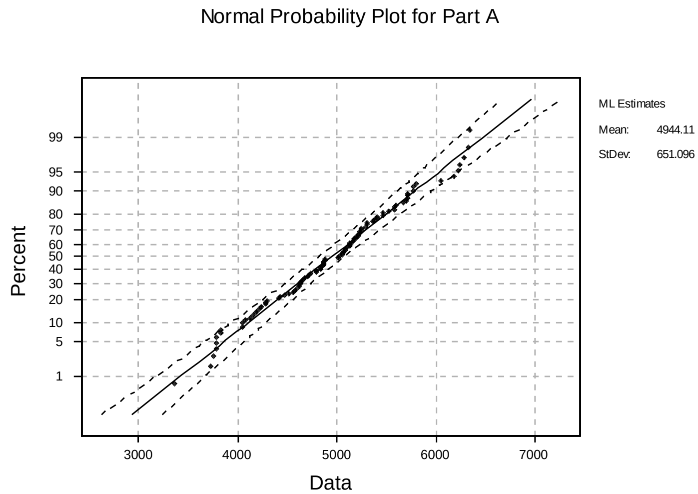
$$P(\text{No disorder} \mid - \text{result}) = \frac{P(NotD\cap -)}{P(-)} = \frac{0.95\times 0.90}{(0.05\times 0.05)+(0.95\times 0.90)} = \frac{0.855}{0.8575} = 0.997$$

   If you have a negative result, you are almost certainly free of the disorder.

If you have a positive result, you are still probably disorder free though your chances of having it are increased. Without the test you have a 5% chance of having it, with a positive test result you have a 33% chance of having it.

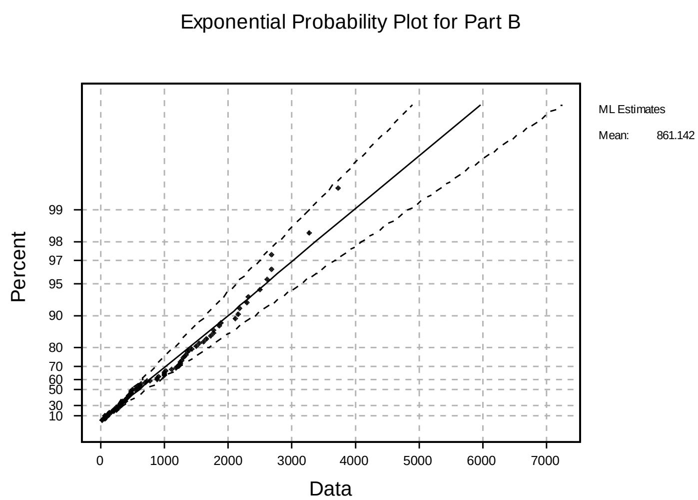
5. **Part A**

Both the normal distribution and the lognormal distribution are good models. In modelling generally, we would use the simplest model which is acceptable. In this case the *normal distribution.*

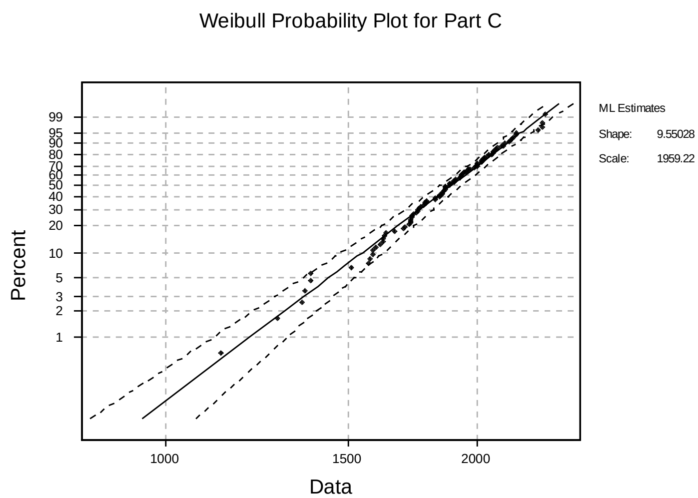


Part B

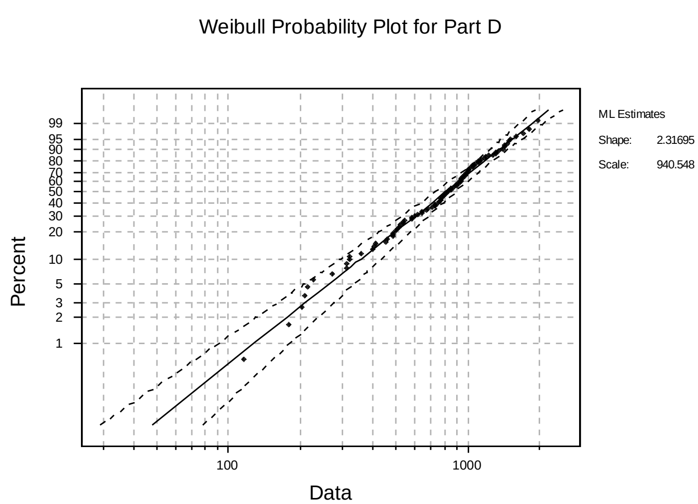
The *exponential distribution* can be used. Note that the *Weibull* plot is also linear. As we will see in a later section, the exponential distribution is a special case of the Weibull so this is as expected.



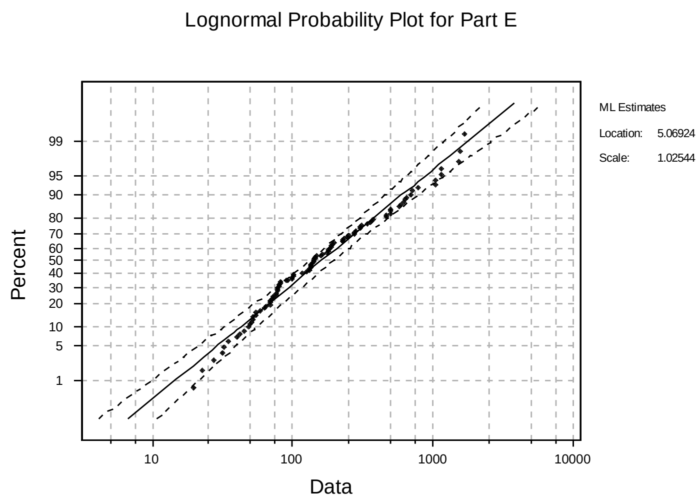
Part C
The *Weibull distribution* does best in the tails of the ID plot and all points are within limits in the Weibull plot below.



Part D
ID plots of all the data show a clear outlier which needs investigating. Looking at the data, this is item 89 recorded as 22250.0 Why is it so extreme? If it's a *correct* value, we would want to know why this item lasted so long – which would probably be the most interesting piece of information in the data. If it's *incorrect*, the value should be corrected or removed. Without any other information, it's best to just remove it. (Actually, it was a data entry error. 22250.0 was entered instead of 2250.0). Plots with the outlier removed show that the *Weibull distribution* is an excellent fit.

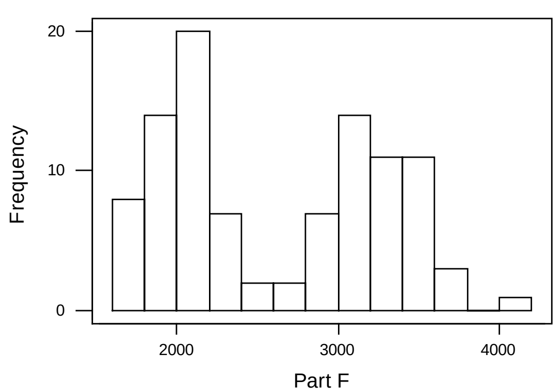


Part E
The *lognormal distribution* is a good model.

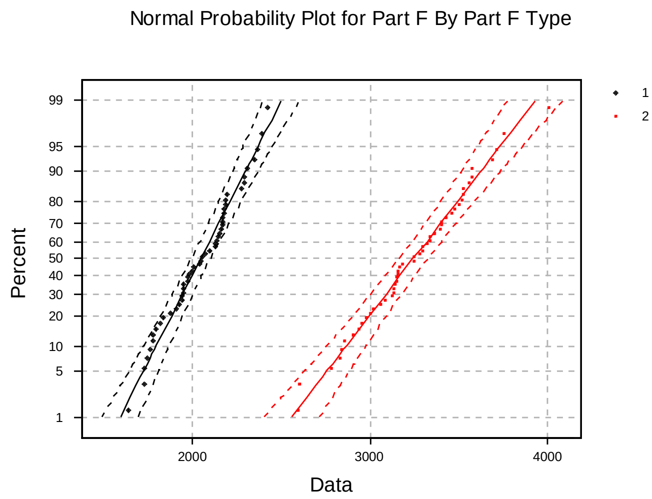


Part F

ID plots (except the exponential) are S-shaped which suggests the distribution has more than one peak. A histogram would help here. In fact it's *bimodal*.



Also on the worksheet in column C43 is an indicator for *Part F Type*. This shows that the first 50 items in column 6 are of one type and the second 50 are of another type. These should be analysed separately. Splitting (stratifying) by type suggests that the original distribution may be considered as a mixture of normal distributions. The probability plots are given below.



6. (i) P(lifetime < 6000) = 0.95 approximately

(ii) 90% of items have lifetimes less than 5800 hours approximately which is the same as saying that 10% of items will have lifetimes exceeding this amount.

7. (i) P(lifetime > 3000) = 1 – p(lifetime < 3000) = 1 – 0.97 approximately. Thus this item has 3% reliability at 3000 hours.

(ii) The median will have 50% of lifetimes below it. From the plot this is roughly 550 hours.

Using Minitab, the mean of the sample of 100 lifetimes is 861 hours. From the plot just over 60% (or a proportion of 0.60) will fail before this time. The distribution is very positively skew and, as expected, the mean is larger than the median.

8. P(1500 < lifetime < 2000) = P(lifetime < 2000) – P(lifetime < 1500)
From the plot this is approximately 0.70 – 0.075 = 0.625

9. The reliability at 1000 hours = P(lifetime > 1000) ≈ 1 – 0.96 = 0.04
Looking at the data in column 5 of the worksheet, it can be seen that 7 of the 100 items had lifetimes of at least 1000 hours – a relative frequency of 7/100 = 0.07. This is similar to the estimate obtained from the probability plot. However, because the estimate from the plot uses more of the information in the data, it is likely to be closer to the true population value.