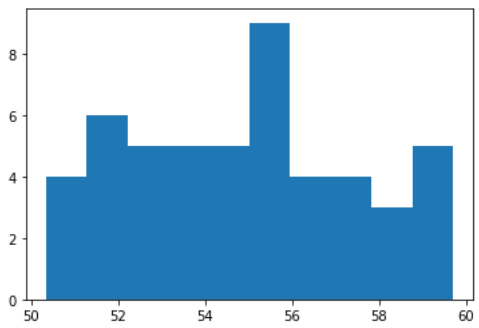
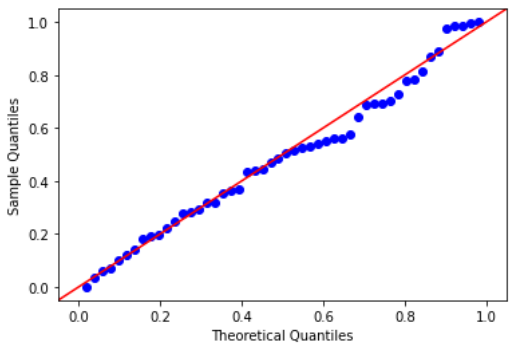
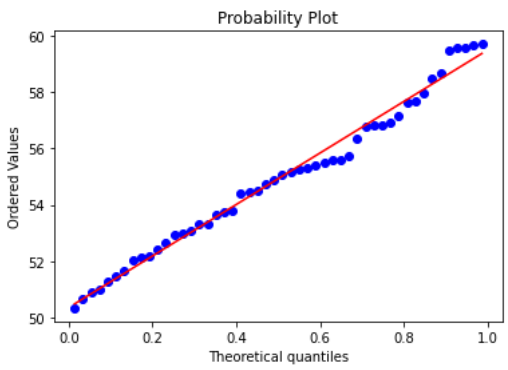
2.5 Guessing the Distribution of Dataset 2

Sunday, 02 October 2022 8:22

<table>
<tr><td></td><td>Uniform Distribution</td></tr>
<tr><td></td><td> 

Peaks on either side, but it's not symmetric in nature, in the sense there is no stepwise decrease on either side.
So, it's a non-Gaussian, somewhat symmetric.
</td></tr>
<tr><td></td><td>
<table>
<tr><td>count</td><td>50.000000</td></tr>
<tr><td>mean</td><td>54.925080</td></tr>
<tr><td>std</td><td>2.659768</td></tr>
<tr><td>min</td><td>50.340227</td></tr>
<tr><td>25%</td><td>52.965425</td></tr>
<tr><td>50%</td><td>54.981739</td></tr>
<tr><td>75%</td><td>56.831619</td></tr>
<tr><td>max</td><td>59.695597</td></tr>
</table>
{('Variance Observed', 7.07), ('Mean Observed', 54.93),
('Skew Observed', 0.18), ('Kurt Observed', -0.87)}

Between the minimum and 25%, 25% to 50%, 50% to 75%, and 75% to maximum, the intervals are around 2.
$\frac{min+max}{2} = \frac{50+60}{2} = \frac{110}{2} = 55 = mean$. We can assume that this comes from uniform which is between 50 and 60 .
Variance or the standard deviation is very small and it is symmetric.
From these numbers, we can conclude that it's a uniform distribution. Because between the quantiles, the split is almost the same, and the standard deviation is very small.
Kurtosis is negative.
Generally for normal distributions the Kurtosis would be around 3. So it cannot be a normal.
Skewness is very low. So, it's not skewed to either side. It is centred.
</td></tr>
<tr><td></td><td>Q-Q plot w.r.t Uniform
 

P-P plot w.r.t Uniform
 

From both plots, we can see that the observed points are very close to the red line.
So, it's a uniform distribution.</td></tr>
<tr><td></td><td>NULL HYPOTHESIS: The given data follows Uniform distribution.

ALTERNATE HYPOTHESIS: The given data does not follow Uniform distribution</td></tr>
<tr><td></td><td>

Each observation is observed frequency.
Expected frequency for each observation $= mean = 54.925080$
</td></tr>
<tr><td></td><td>Calculated chi square statistic = 6.31
p-value = 0.99

p-value > α
We accept the null.
</td></tr>
</table>

	$df = k - p - 1 = 50 - 0 - 1 = 49$ $k = 50$: number of observations $p = 0$: for uniform distribution
	Tabulated Chi Square value $= 66.34$ Tabulated value $>$ Calculated value We accept the null.
Case examples	• Throwing a die. The probability of getting each side is 16 . It's a uniform distribution. • Fuel Efficiency. If you put one litre petrol/diesel, every time you would get a different mileage. It could be 50, 52, 47. In case you're a rash driver, it could be low, say 40, 42.