SRM-1 Project

Graduation tests

Graduation tests

Data Csv = Graduation.csv

Q1} Calculating Crude rates

#Mortalitydata$CRUDE<-Crude\_rate

Q2} Using Gompertz law for Graduated rates

Since,

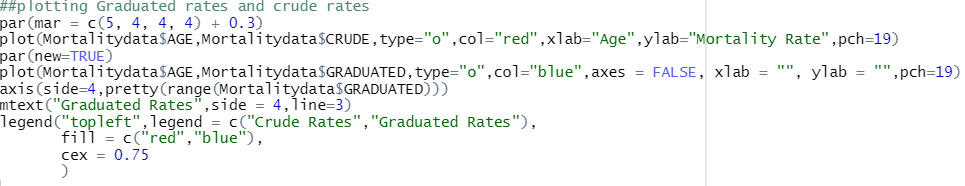
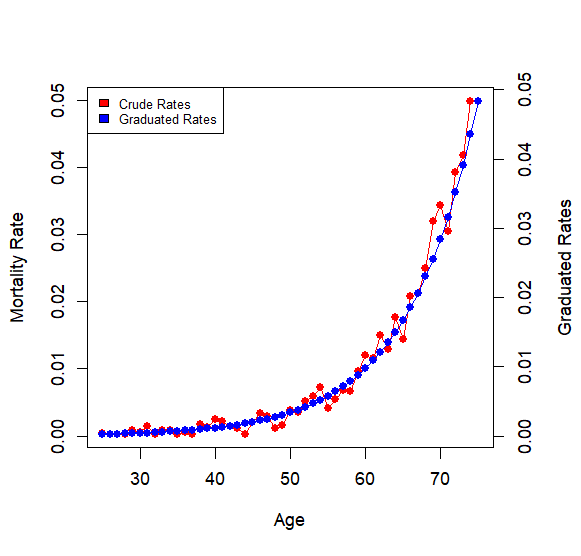
Taking log of both sides

,which is a linear equation for x

Using lm() function to put Gompertz law on data and coef() & as.numeric() for coefficients value

coef() to get the coefficients out of the model and as.numeric to remove headings

Now calculating Graduated rates:



Q3} Applying Third Difference to check Smoothness

Given: diff1 = function(x)x[-1]-x[-length(x)]

The values obtained were multiplied 10^6 to make them easy to compare

The Crude rates are much higher in magnitude as compared to Graduated rates and are very irregular because they are directly calculated from DEATHS and ETR, however the max value to Graduated went till 50 because a parametric formula is used with two parameters and the slight irregularity is because of rounding.

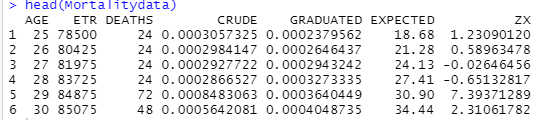
Q4} Calculating Values in EXPECTED AND ZX

Expected:-

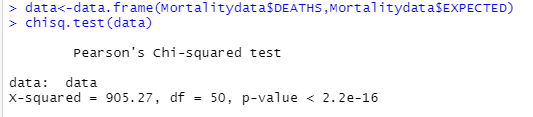
Mortalitydata$EXPECTED<-round(Mortalitydata$GRADUATED\*Mortalitydata$ETR,2)

ZX:-

Mortalitydata$ZX<-(Mortalitydata$DEATHS-Mortalitydata$EXPECTED)/sqrt(Mortalitydata$EXPECTED)



Using Chi-square test to check Goodness of fit of DEATHS & EXPECTED



degrees of freedom = 50

H0: EXPECTED and DEATHS are a good fit of one another

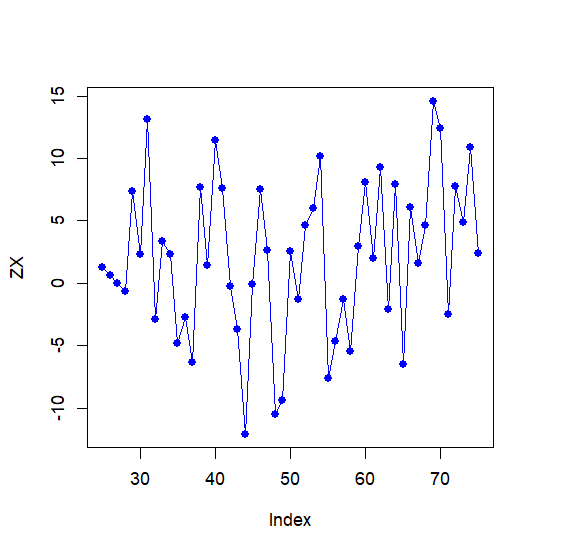
Since p-value is less than 0.05 we accept null Hypothesis.

Q5}

A] Standardized Deviation Test

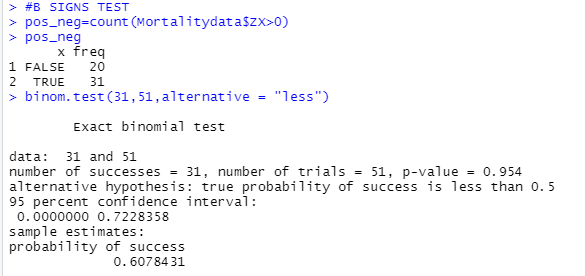
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (-inf,-3) | (-3,-2) | (-2,-1) | (-1,0) | (0,1) | (1,2) | (2,3) | (3,inf) |
| Obs | 10 | 4 | 2 | 4 | 1 | 4 | 6 | 20 |
| expected | 0.0501 | 1.07 | 6.936 | 17.391 | 17.391 | 6.936 | 1.07 | 0.0501 |

* Overall Shape : the values are positively skewed, so it doesn’t follow normal distribution
* Absolute Deviations : lots of values concentrated towards the tails
* Outliers : there a lot of outliers in the data



* Symmetry : No symmetry, positively skewed data
* Final conclusion about Null hypothesis: Reject the null hypothesis that the data is a good representation of the underlying mortality rates. We reject our null hypothesis at 5%.

B] Sign Test



p value = 0.954, Binomial(51,0.5)

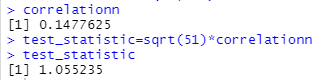
We accept the Null Hypothesis : No bias is present in data

C] Cumulative Deviation Test

H0:No overall bias present

The observed value of test statistic is 18.83024

For a test at 5% level of significance level we compare the value of test statistic with lower and upper 2.5% points, i.e., with +-1.96. Since 18.83 is greater than 1.96 we reject the Null Hypothesis and conclude the graduated rates are too low.

D] Serial Correlation test

H0: No grouping of signs

The value of test statistic is 1.055235

As we are testing only for positive correlation, we compare the value of test statistic with 1.6449, the upper point of 5%. We find there is insufficient evidence to reject the null hypothesis, there is no grouping of signs