

Analysis of TAutoCorr.R

Oliver Tarrant

Is the temperature of one year significantly correlated with that of the following year?

Method

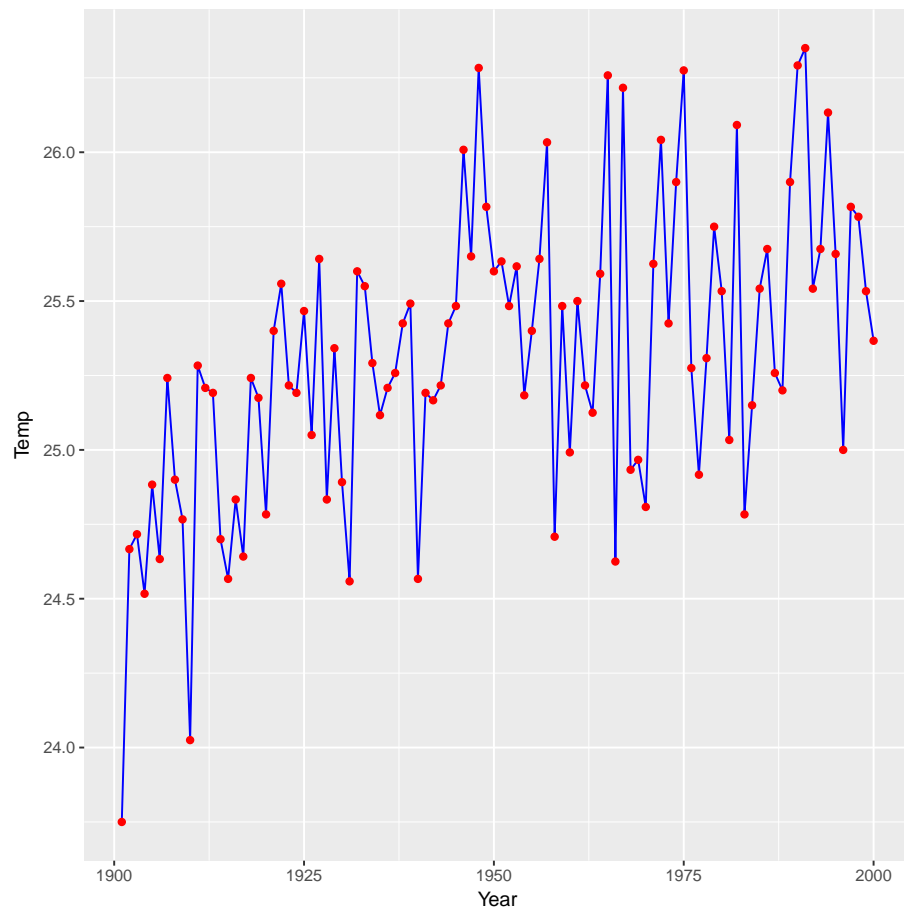
Using the TAutoCorr.R script, I have analysed the data from the "KeyWest-AnnualMeanTemperatures" to look if there is a significant correlation between the mean temperatures in consecutive years.

This has been done by calculating a correlation coefficient for each of the n-1 pairs of consecutive years. To see if this value has been produced by chance or if the correlation coefficient is statistically significant, this value is compared with 10000 correlation coefficients generated by looking for a correlation in randomly ordered sequences of the years. By looking at the ratio of the coefficients of these randomly generated sequences that are greater than the correlation coefficient from the original data, I have an estimate for the p value for the correlation.

Results

```
source("TAutoCorrCode.R")

## Loading objects:
##   ats
## 'data.frame': 100 obs. of  2 variables:
##  $ Year: int   1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 ...
##  $ Temp: num   23.8 24.7 24.7 24.5 24.9 ...
## [1] "The range of temperatures is: "
## [1] 23.75 26.35
## [1] "The range of years is: "
## [1] 1901 2000
## [1] "The mean temperatures is: "
## [1] 25.31475
## [1] "The median temperatures is: "
## [1] 25.2875
```



```
## [1] "The correlation coefficient is: "  
## [1] 0.3261697  
## [1] "The approximate p-value is: "  
## [1] 7e-04
```

Interpretation

For my investigation my null hypothesis is that the correlation is not significant (approximately 0), and my alternative hypothesis is that the data is positively correlated. By analysing my results I get a positive correlation coefficient of 0.326169651060742 . By approximating the corresponding p-value for this statistic I get a p-value of 7e-04 . This falls outside my 95 percent confidence interval for the null hypothesis. Thus at a 5 percent level of significance I will reject the null hypothesis in favour of the alternative. I.e the chances of my data being produced if the null hypothesis is true is less than 5 percent. Thus my test suggests that there is a positive correlation between the consecutive yearly mean temperatures in Florida.