stationarity

March 3, 2020

1 Stationarity

The file covers the main features of a stationary time series.
An accompanying Rmd file is provided.

1.1 Conditions

There are three conditions for a time series to be stationary.

- Constant Mean
- Constant Variance
- Constant Autocorrelation

1.1.1 Constant Mean

All subpopulations of X_t have the same mean for each time t i.e. the mean does not depend on time.

$$E[X_t] = \mu$$

1.1.2 Constant Variance

Variance of X_t does not depend on time and the variance is finite.

$$Var[X_t] = \sigma^2 < \infty$$

1.1.3 Constant Autocorrelation

The correlation of X_2 and X_1 only depends on $t_2 - t_1$. That is, the correlation between the data points depends only on how far a part the obervations are in time, not where the observations are located in time.

$$Corr(X_t, X_{t+h}) = \rho_h$$

where h is the difference between two points in time.

1.1.4 Checking Assumptions in Practice

For constant mean and variance, plot the time series and visually assess the validity of the assumption. For the constant autocorrelation, split the time series into subpopulations and plots the ACFs for the subpopulations, then visually assess the validity of the assumption.

1.2 Parameter Estimation

A single realization can be used to estimate the mean, variance, and autocorrelation of a stationary time series when the autocorrelation approach zero as the lag increases.

Parameters can be estimated in the following order due to dependence:

- mean
- sample autocorrelations
- variance
- confidence interval for the mean

Let X_t be a discrete stationary time series for the estimate procedures.

1.2.1 Estimation of the Mean

Then $\bar{x}$ is an unbiased estimate of the mean of the time series. Thus, the mean will be estimated as

$$\hat{Mean}(X_t) = \bar{x} = \frac{\sum^N x_i}{N}$$

1.2.2 Estimation of the Autocorrelations

Recall the definition of ρ_k:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{E[(X_t - \mu)(X_{t+k} - \mu)]}{\sigma_x^2}$$

The autocorrelations can be estimated with the **sample autocorrelations**:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\frac{1}{N} \sum^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\frac{1}{N} \sum^N (x_t - \bar{x})^2}$$

Note that the amount of data used to estimate the sample autocorrelations decreases as $k \to N$.

1.2.3 Estimation of the Variance

The variance of the can be estimated by

$$\hat{Var}(X_t) = \frac{\hat{\sigma}^2}{N} \sum_{k=-(N-1)}^{N-1} \left(1 - \frac{|k|}{N}\right) \hat{\rho}_k$$

simplifying the summation by noting that $\rho_k = \rho_{-k}$,

$$= \frac{\hat{\sigma}^2}{N} \left(1 + 2 \sum_{k=1}^{N-1} \left(1 - \frac{|k|}{N} \right) \hat{\rho}_k \right)$$

where $\hat{\rho}_k$ represents the kth autocorrelation estimate and σ^2 is estimated with all the data as shown below.

$$\hat{\sigma}^2 = \frac{1}{N} \sum^{N} (x - \bar{x})^2$$

1.2.4   Estimation of a Confidence for the Mean

The confidence interval for the mean follows the common form of

$$CI : (Mean\ Estimate) \pm (Multiplier)(Variance\ Estimate)$$

Substituting for the estimation, we have

$$CI : \bar{x} \pm t_{1-\frac{\alpha}{2}} \sqrt{ \frac{\hat{\sigma}^2}{N} \left(1 + 2 \sum_{k=1}^{N-1} \left(1 - \frac{|k|}{N} \right) \hat{\rho}_k \right) }$$

This confidence interval is **interpreted as** "We are $X\%$ confident that the mean is contained in the interval $[CI\ Lower, CI\ Upper]$"

1.3   References

[1] W. Woodward and B. Salder, "Stationary", SMU, 2019