# 1 Autocorrelation and Time Series Analysis

## 1.1 Introduction

- A ***time series*** is a **sequence** of random variables $\boldsymbol{Y_1, Y_2, Y_3, \ldots}$ observed at distinct points in **time**.

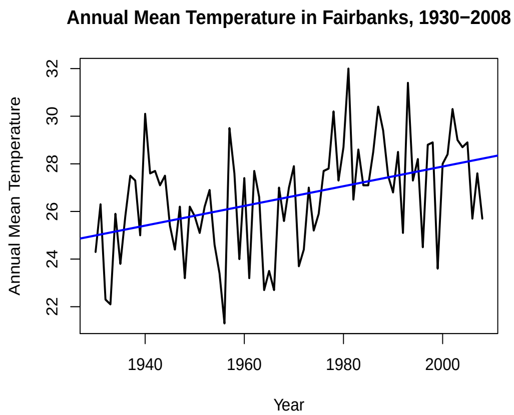
  We'll assume the time points are **regular** (equally spaced), and we'll denote those **time points** by $\boldsymbol{X_1, X_2, X_3, \ldots}$.

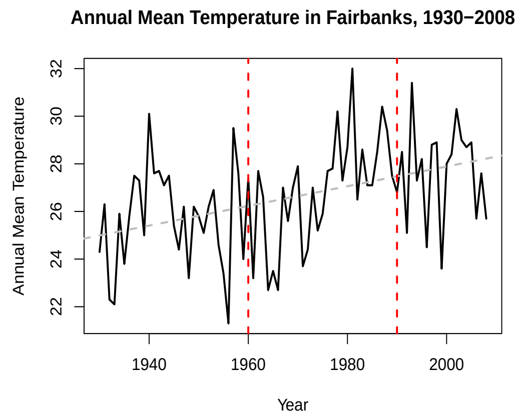
- The **random errors** (i.e. the $\epsilon$'s) in time series data often exhibit ***autocorrelation*** (also called ***serial correlation***), meaning observations made close together in time are correlated.

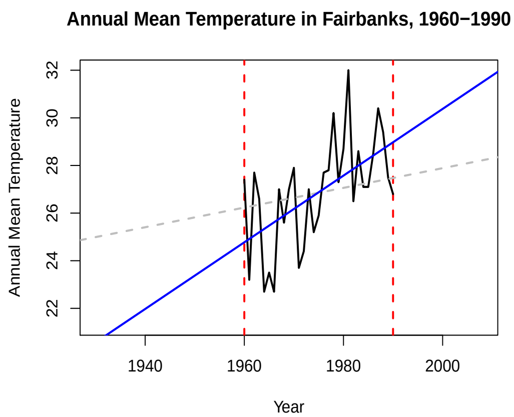
- For autocorrelated data, ordinary least squares regression *inference* procedures (e.g. the $t$ test or CI for $\beta_1$) that assume *independent* errors **aren't** appropriate.

  Some **consequences** of using such procedures when autocorrelation is present include:

  1. The **MSE** may **underestimate** the true error variance $\boldsymbol{\sigma^2}$.
  2. The **estimated standard error** $\boldsymbol{s\{b_k\}}$ of a coefficient $b_k$ may **underestimate** the true standard error $\boldsymbol{\sigma\{b_k\}}$.
  3. **Confidence intervals** and $\boldsymbol{t}$ **and** $\boldsymbol{F}$ **tests** for coefficients **may no longer be valid**. In particular, actual Type I error probabilities can be greater than the nominal level (e.g. $\alpha = 0.05$), and actual CI coverage probabilities may be smaller than the nominal level (e.g. 95%).




- However, if we're just interested in fitting a model to *estimate* the regression parameters, without carrying out hypothesis tests or constructing CIs, ordinary least squares is still valid with autocorrelated data, and the estimators are still unbiased.

## 1.2 Regression Model with First Order Autoregressive Errors

- Let $\boldsymbol{Y_t}$ denote the observation at time $\boldsymbol{X_t}$. One possible model for $Y_t$ is the regression model with ***first order autoregressive errors*** (or ***AR(1) errors***):

**Regression Model with First Order Autoregressive Errors (or AR(1) Errors)**:

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t \qquad (1)$$

with

$$\epsilon_t = \rho\epsilon_{t-1} + u_t \tag{2}$$

where

▷ $\boldsymbol{\rho}$ is a parameter satisfying $|\rho| < 1$.

▷ $\boldsymbol{u_t}$ is a $N(0, \sigma^2)$ random variable and:

1. The $u_t$'s are independent of each other.
2. $u_t$ is independent of $\epsilon_s$ for $s < t$.

- More complicated mean response models (e.g. polynomial models) can be used in place of $\beta_0 + \beta_1 X$ in (8) with the AR(1) errors.
- It can be shown that the $\epsilon_t$'s defined by (2) have the following properties:

1. The expected value of $\epsilon_t$ is zero, i.e.

$$\mathrm{E}(\epsilon_t) = 0$$

2. The variance and standard deviation of $\epsilon_t$ are

$$\sigma^2\{\epsilon_t\} = \frac{\sigma^2}{1-\rho^2} \qquad \text{and} \qquad \sigma\{\epsilon_t\} = \sqrt{\frac{\sigma^2}{1-\rho^2}}$$

3. The covariance between two successive errors $\epsilon_t$ and $\epsilon_{t-1}$ is

$$\sigma\{\epsilon_t, \epsilon_{t-1}\} = \rho \cdot \frac{\sigma^2}{1-\rho^2}$$

4. The correlation between two successive errors $\epsilon_t$ and $\epsilon_{t-1}$ is

$$\rho\{\epsilon_t, \epsilon_{t-1}\} = \frac{\sigma\{\epsilon_t, \epsilon_{t-1}\}}{\sigma\{\epsilon_t\}\sigma\{\epsilon_{t-1}\}} = \rho \tag{3}$$

5. The covariance between error terms $\epsilon_t$ and $\epsilon_{t-s}$ (i.e. error terms $s$ time units apart) is

$$\rho\{\epsilon_t, \epsilon_{t-s}\} = \rho^s \cdot \frac{\sigma^2}{1-\rho^2} \tag{4}$$

6. The correlation between error terms $\epsilon_t$ and $\epsilon_{t-s}$ (i.e. error terms $s$ time units apart) is

$$\rho\{\epsilon_t, \epsilon_{t-s}\} = \frac{\rho\{\epsilon_t, \epsilon_{t-s}\}}{\sigma\{\epsilon_t\}\sigma\{\epsilon_{t-s}\}} = \rho^s \tag{5}$$

- Note that (3) says that the **value** of $\boldsymbol{\rho}$ determines the degree of **autocorrelation** between successive errors $\boldsymbol{\epsilon_t}$ and $\boldsymbol{\epsilon_{t-1}}$. Values of $\rho$ close to **1.0** indicate a high degree of autocorrelation. Values close to **0.0** indicate a very little autocorrelation.

- Note that because $|\rho| < 1$, (5) says that the **correlation** between errors $\boldsymbol{s}$ **time units apart** goes to **zero** as $\boldsymbol{s}$ **increases**.

- The right sides of (4) and (5), as functions of the so-called ***lag s***, are called the ***autocovariance function*** and ***autocorrelation function***, respectively.

## 1.3 Durbin-Watson Test for Autocorrelation

- For the regression model with **AR(1)** errors, i.e.

$$Y_t \;=\; \beta_0 + \beta_1 X_t + \epsilon_t \tag{6}$$

with

$$\epsilon_t \;=\; \rho\epsilon_{t-1} + u_t \,,$$

**if** $\boldsymbol{\rho > 0}$, then **consecutive** errors $\boldsymbol{\epsilon_t}$ and $\boldsymbol{\epsilon_{t-1}}$ are **positively correlated**, meaning they tend to be similar in value.

On the other hand, **if** $\boldsymbol{\rho = 0}$, then the $\epsilon_t$'s are the same as the $u_t$'s, in which case they're **independent** and therefore **uncorrelated**.

- The ***Durbin-Watson Test*** for autocorrelation is a test of

$$H_0 : \rho \;=\; 0 \qquad \text{(there's no correlation between consecutive errors)}$$
$$H_a : \rho \;>\; 0 \qquad \text{(there's a correlation)}$$

The **test statistic** is

**Durbin Watson Test Statistic**:

$$D \;=\; \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

where $\boldsymbol{e_t}$ is the $\boldsymbol{t}$**th residual** obtained from fitting the regression function to the data using least squares, i.e.

$$e_t \;=\; Y_t - \hat{Y}_t$$

where $\boldsymbol{\hat{Y}_t}$ is the usual **fitted value** at $X_t$.

**Small** values of $D$ provide evidence against $H_0$ in favor of $H_a$ since $D$ will be small when consecutive residuals $e_t$ and $e_{t-1}$ are similar in value.

The exact distribution of $D$ when $H_0$ is true is unknown, but approximate critical values $d_u$ and $d_l$ have been determined, for given significance levels (e.g. $\alpha = 0.01$ or $\alpha = 0.05$), such that the decision rule is

Reject $H_0$ if $D < d_l$
Fail to reject $H_0$ if $D > d_u$

If $d_l \leq D \leq d_u$, the test is **inconclusive**.

- Note that the Durbin-Watson test can be used exactly as described with more complicated mean response models (e.g. polynomial regression models).

## 1.4 Remedial Measures

- When autocorrelation in the errors is present, ordinary least squares inference procedures can lead to erroneous conclusions.
- Some remedial measures for dealing with autocorrelated errors are:

1. Add more predictors to the model – there may be variables that are driving the apparent autocorrelations, and adding them to the model will moderate the degree of autocorrelation in the errors.
2. Use the transformed responses

$$Y_t' = Y_t - \rho Y_{t-1} \tag{7}$$

It's easy to show that if the original model is

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t \tag{8}$$

where $\epsilon_t = \rho\epsilon_{t-1} + u_t$, then the transformation (7) leads to the model

$$Y_t' = \beta_0' + \beta_1' X_t' + u_t \tag{9}$$

where

$$\begin{aligned} \beta_0' &= \beta_0(1-\rho) \\ \beta_1' &= \beta_1 \\ X_t' &= X_t - \rho X_{t-1} \end{aligned}$$

Then the errors $u_t$ in (9) are independent and, after estimating $\rho$ (see the textbook), ordinary least inference procedures can be used. Once the least squares estimates $b_0'$ and $b_1'$ of $\beta_0'$ and $\beta_1'$ are obtained, they can be transformed back to get estimates of $\beta_0$ and $\beta_1$ using

$$\begin{aligned} b_0 &= \frac{b_0'}{1-\rho} \\ b_1 &= b_1' \end{aligned} \tag{10}$$

In practice, an estimate $r$ of $\rho$ must be used in the transformation (7) and the back-transformation (10), using for example either the ***Cochrane-Orcutt*** or <u>***Hildreth-Lu***</u> procedures (see the book).

3. Use the <u>***first differences procedure***</u> by calculating the differences

$$Y_t' = Y_t - Y_{t-1}$$

It's easy to show that this leads to the model

$$Y_t' = \beta_1' X_t' + u_t \tag{11}$$

where

$$\begin{aligned} \beta_1' &= \beta_1 \\ X_t' &= X_t - X_{t-1} \end{aligned}$$

and since the errors $u_t$ in the model (11) are independent, ordinary least squares procedures can be used to estimate and draw inferences about $\beta_1'$. Once the estimate $b_1'$ of $\beta_1'$ is obtained, estimates of $\beta_0$ and $\beta_1$ in (8) can be obtained as

$$\begin{aligned} b_0 &= \bar{Y} - b_1'\bar{X} \\ b_1 &= b_1' \end{aligned}$$

(See the textbook).

## 1.5 Forecasting with Autocorrelated Errors

- When errors in a regression model are autocorrelated, an error $\epsilon_n$ can be used to predict the next response $Y_{n+1}$.
- Assuming the model (6), the $(n+1)$st response is

$$\begin{aligned} Y_{n+1} &= \beta_0 + \beta_1 X_{n+1} + \epsilon_{n+1} \\ &= \beta_0 + \beta_1 X_{n+1} + \rho\epsilon_n + u_{n+1} \end{aligned} \tag{12}$$

- The ***forecast*** for time period $n+1$ (i.e. the prediction for the value of $Y_{n+1}$) involves replacing $u_{n+1}$ in (12) by zero (its expected value), replacing $\beta_0$, $\beta_1$, and $\rho$ by estimates $b_0$, $b_1$, and $r$, and replacing $\epsilon_n$ by the residual $e_n$.

  Thus the **forecast**, denoted $\boldsymbol{F_{n+1}}$, is

  $$F_{n+1} = \hat{Y}_{n+1} + re_n$$

  where

  $$\hat{Y}_{n+1} = b_0 + b_1 X_{n+1}$$

  and the residual $e_n$ is

  $$\begin{aligned} e_n &= Y_n - \hat{Y}_n \\ &= Y_n - (b_0 + b_1 X_n) \end{aligned}$$

  with $b_0$ and $b_1$ estimated using the transformed responses (7) and the back-transformations (10), and $r$ an estimate of $\rho$ obtained using either the ***Cochrane-Orcutt*** or ***Hildreth-Lu*** procedure (see the textbook).

- Forecasts **more than one time period into the future** can also be made. For example, the forecast of $Y_{n+2}$ is

  $$F_{n+2} = \hat{Y}_{n+2} + r^2 e_n$$

  See the textbook.