Statistics 191: Introduction to Applied Statistics

Correlated Errors, Whitening

Jonathan Taylor
Department of Statistics
Stanford University

February 22, 2010

Correlated Erorrs

Statistics 191: Introduction to Applied Statistics

Jonathan Taylor
Department of Statistics
Stanford University

Topics

- Autocorrelation.
- Whitening.

Correlated errors

Autocorrelation

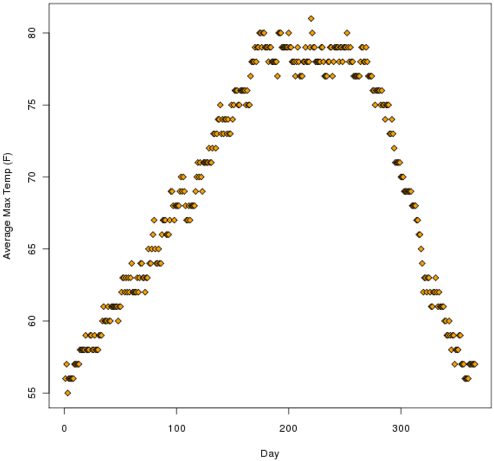
- In the random effects model, outcomes within groups were correlated.
- Other regression applications also have correlated outcomes (i.e. errors).
- Common examples: time series data.
- Why worry? Can lead to underestimates of SE $\rightarrow$ inflated t's $\rightarrow$ false positives.

Correlated errors

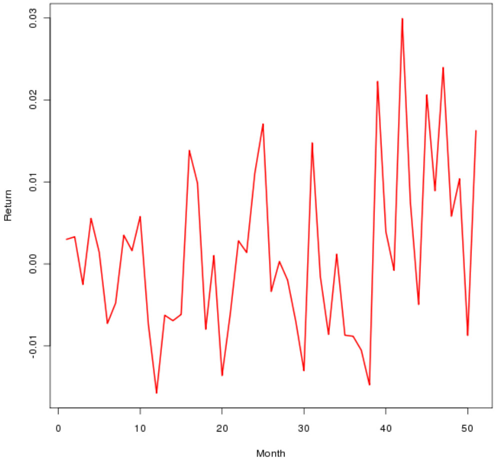
Autocorrelation

- Suppose we plot Palo Alto's daily average temperature – clearly we would see a pattern in the data.
- Sometimes, this pattern can be attributed to a deterministic phenomenon (i.e. predictable seasonal fluctuations).
- Other times, "patterns" are due to correlations in the noise, maybe small time fluctuations in the stock market, economy, etc.
- Example: financial time series, monthly bond return.
- Example: residuals regressing consumer expenditure on money stock.

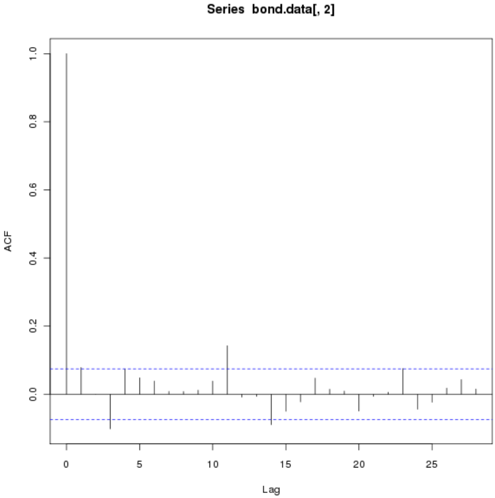
Average Max Temp in Palo Alto



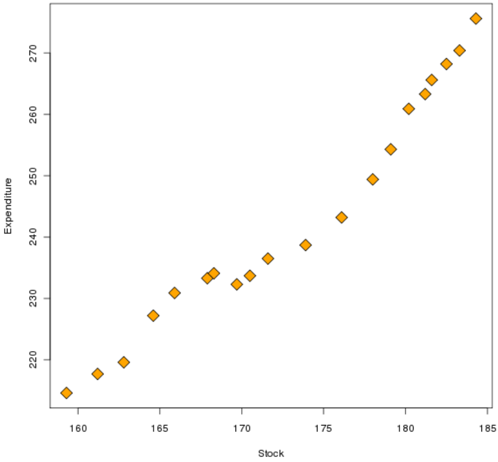
Monthly bond return



Monthly bond return, ACF

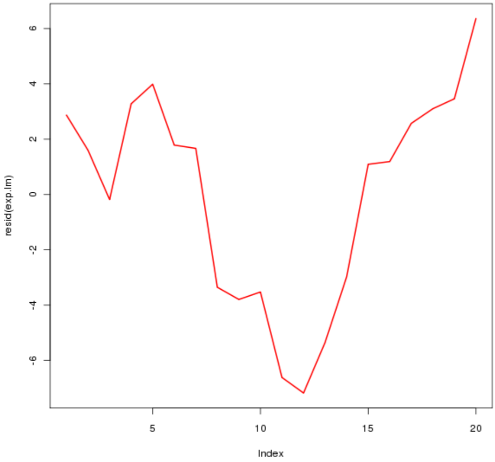


Expenditure vs. stock



Expenditure vs. stock: residuals

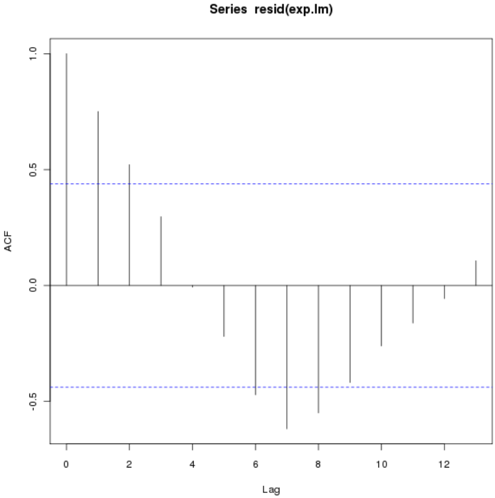
Jonathan Taylor
Department of Statistics
Stanford University



ACF of residuals

Statistics 191: Introduction to Applied Statistics

Jonathan Taylor
Department of Statistics
Stanford University



Correlated errors

AR(1) noise

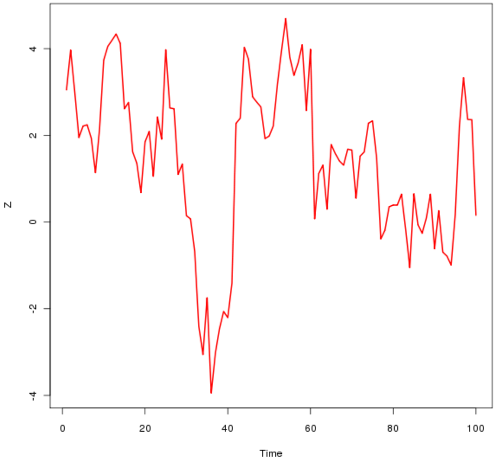
- Suppose that, instead of being independent, the errors in our model were

$$\varepsilon_t = \rho \cdot \varepsilon_{t-1} + \omega_t, \qquad -1 < \rho < 1$$

  with $\omega_t \sim N(0, \sigma^2)$ independent.
- If ρ is close to 1, then errors are very correlated, $\rho = 0$ is independence.
- This is "Auto-Regressive Order (1)" noise (AR(1)). Many other models of correlation exist: ARMA, ARIMA, ARCH, GARCH, etc.

AR(1) noise, $\rho = 0.9$

Jonathan Taylor
Department of Statistics
Stanford University



Correlated errors

Autocorrelation function

- For a "stationary" time series $(Z_t)_{1 \leq t \leq \infty}$ define

$$ACF(t) = \text{Cor}(Z_s, Z_{s+t}).$$

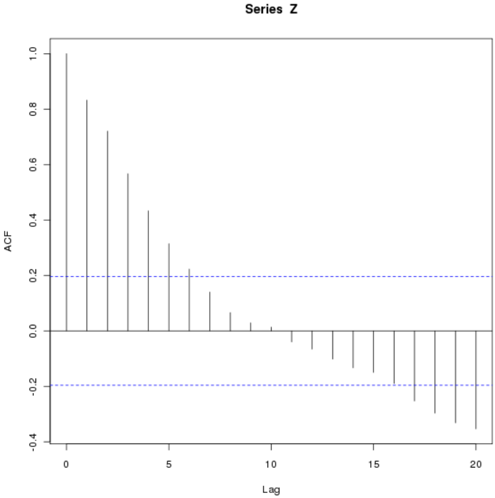
- Stationary means that correlation above does not depend on s.
- For AR(1) model,

$$ACF(t) = \rho^t.$$

- For a sample $(Z_1, \dots, Z_n)$ from a stationary time series

$$\widehat{ACF}(t) = \frac{\sum_{j=1}^{n-t}(Z_j - \overline{Z})(Z_{t+j} - \overline{Z})}{\sum_{j=1}^{n}(Z_j - \overline{Z})^2}.$$

ACF of AR(1) noise, $\rho = 0.9$



Correlated errors

Jonathan Taylor
Department of Statistics
Stanford University

Effects on inference

- So far, we have just mentioned that things *may* be correlated, but not thought about how it affects inference.
- Suppose we are in the "one sample problem" setting and we observe

$$W_i = Z_i + \mu, \qquad 1 \leq i \leq n$$

with the Z_i's from an $AR(1)$ time series. It is easy to see that

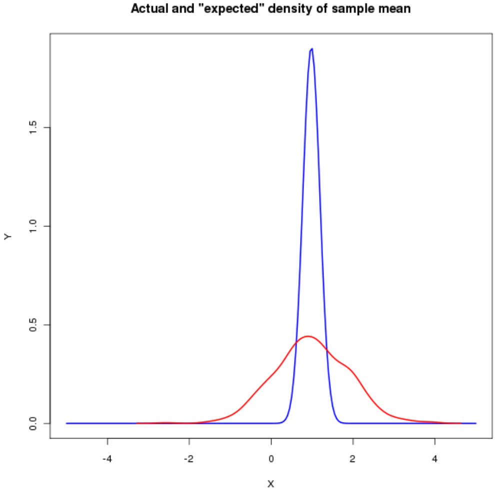
$$E(\overline{W}) = \mu$$

BUT, generally

$$\text{Var}(\overline{W}) > \frac{\text{Var}(Z_1)}{n}$$

how much bigger depends on ρ.

Misleading inference ignoring autocorrelation



Regression with auto-correlated errors

Model

item Observations:

$$Y_t = \beta_0 + \sum_{j=1}^{p} X_{tj}\beta_j + \varepsilon_t, \qquad 1 \leq t \leq n$$

Errors:

$$\varepsilon_t = \rho \cdot \varepsilon_{t-1} + \omega_t, \qquad -1 < \rho < 1$$

Question: who do we determine if autocorrelation is present?Question: what do we do to correct for autocorrelation?

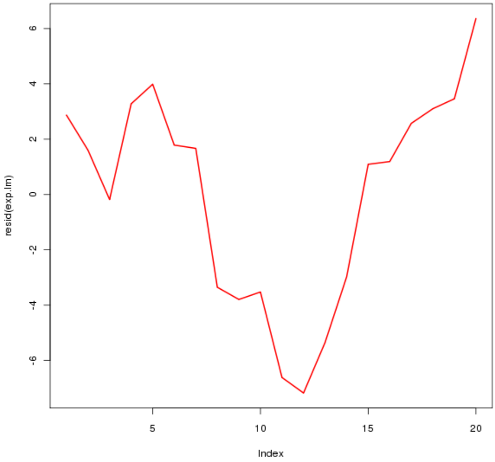
Regression with auto-correlated errors

Graphical checks for autocorrelation

- A plot of residuals vs. time is helpful.
- Residuals clustered above and below 0 line can indicate autocorrelation.
- Example: regressing consumer expenditure on money stock.

Expenditure vs. stock: residuals

Jonathan Taylor
Department of Statistics
Stanford University



Regression with auto-correlated errors

Jonathan Taylor
Department of Statistics
Stanford University

Durbin-Watson test

- In regression setting, if noise is AR(1), a simple estimate of ρ is obtained by (essentially) regressing e_t onto e_{t-1}

$$\widehat{\rho} = \frac{\sum_{t=2}^n (e_t e_{t-1})}{\sum_{t=1}^n e_t^2}.$$

- To formally test $H_0 : \rho = 0$ (i.e. whether residuals are independent vs. they are AR(1)), use Durbin-Watson test, based on

$$d = 2(1 - \widehat{\rho}).$$

Regression with auto-correlated errors

Jonathan Taylor
Department of Statistics
Stanford University

Correcting for AR(1)

- Suppose we know ρ, if we "pre-whiten" the data and regressors

$$\tilde{Y}_{t+1} = Y_{t+1} - \rho Y_t, t > 1$$
$$\tilde{X}_{(t+1)j} = X_{(t+1)j} - \rho X_{tj}, i > 1$$

  for $1 \leq t \leq n-1$. This model satisfies "usual" assumptions, i.e. the errors

$$\tilde{\varepsilon}_t = \varepsilon_{t+1} - \rho \cdot \varepsilon_t$$

  are independent $N(0, \sigma^2)$.
- For coefficients in new model $\tilde{\beta}$, $\beta_0 = \tilde{\beta}_0/(1-\rho)$, $\beta_j = \tilde{\beta}_j$.
- Problem: in general, we don't know ρ.

Regression with auto-correlated errors

Jonathan Taylor
Department of Statistics
Stanford University

Two-stage regression

- Step 1: Fit linear model to unwhitened data (OLS: ordinary least squares, i.e. no pre-whitening).
- Step 2: Estimate ρ with $\widehat{\rho}$.
- Step 3: Pre-whiten data using $\widehat{\rho}$ – refit the model.

Other models of covariance

- Suppose we model covariance of ε's differently, i.e. $ARMA(p,q)$.
- As long as we can estimate parameters of covariance, from residuals of the OLS fit, we can use this two-stage procedure.
- This is very similar to weighted least squares.

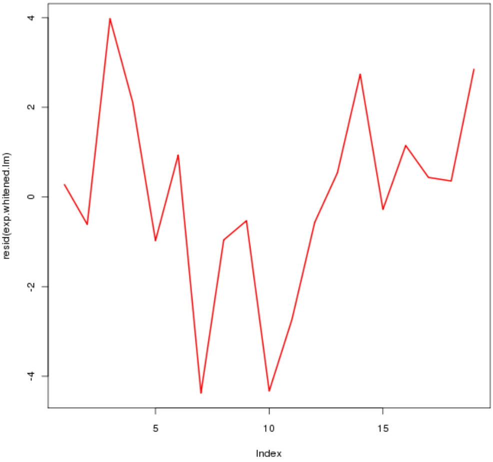
Regression with auto-correlated errors

Jonathan Taylor
Department of Statistics
Stanford University

Interpreting results of two-stage fit

- Basically, interpretation is unchanged, but the exact degrees of freedom in the error is not exactly clear.
- Common argument: "this works for large degrees of freedom, so we better hope we have enough degrees of freedom so this point is not important."
- Can treat t-statistics as Z-statistics, F's as χ^2, appealing to asymptotics:
    - t_ν, with ν large is like $N(0,1)$;
    - $F_{j,\nu}$, with ν large is like χ^2_j/j.

Expenditure vs. stock: whitened residuals



Expenditure vs. stock: ACF of whitened residuals

Jonathan Taylor
Department of Statistics
Stanford University