Forecasting from Time Trend & Time Series

1. **Trend Line Fitting**

   No underlying behavioral econometric model or theory is involved

Trend Types	Model	Forecasts
Straight Line	$Y_t = b_1 + b_2 t + u_t$	$\hat{Y}_t = \hat{b}_1 + \hat{b}_2 t$
Quadratic	$Y_t = b_0 + b_1 t + b_2 t^2 + u_t$	$\hat{Y}_t = \hat{b}_0 + \hat{b}_1 t + \hat{b}_2 t^2$
Cubic	$Y_t = b_0 + b_1 t + b_2 t^2 + b_3 t^3 + u_t$	$\hat{Y}_t = \hat{b}_0 + \hat{b}_1 t + \hat{b}_2 t^2 + \hat{b}_3 t^3$
Linear-Log	$Y_t = b_1 + b_2 \ln(t) + u_t$	$\hat{Y}_t = \hat{b}_1 + \hat{b}_2 \ln(t)$
Reciprocal	$Y_t = b_1 + \frac{b_2}{t} + u_t$	$\hat{Y}_t = \hat{b}_1 + \frac{\hat{b}_2}{t}$
Log-linear	$\ln Y_t = b_1 + b_2 t + u_t;\ Y_t > 0$	$\hat{Y}_t = \exp\left(\hat{b}_1 + \hat{b}_2 t + \frac{\hat{o}^2}{2}\right)$
Double-log	$\ln Y_t = b_1 + b_2 \ln(t) + u_t;\ Y_t > 0$	$\hat{Y}_t = \exp(\hat{b}_1) t\wedge(\hat{b}_2) \exp\left(\frac{\hat{o}^2}{2}\right)$
Logistic	$\ln\left[\frac{Y_t}{1 - Y_t}\right] = b_1 + b_2 t + u_t;\ 0 < Y_t < 1$	$\hat{Y}_t = \frac{1}{1 + \exp\left(-\left[\hat{b}_1 + \hat{b}_2 t + \frac{\hat{o}^2}{2}\right]\right)}$

- If the independent variable is in logarithmic form, then forecasts are biased:

  $\exp(\ln Y_t) = Y_t = \exp(b_1 + b_2 \ln(t) + u_t = \exp(b_1) t\wedge b_2 \exp(u_t)$

  Taking the expected value of both sides:

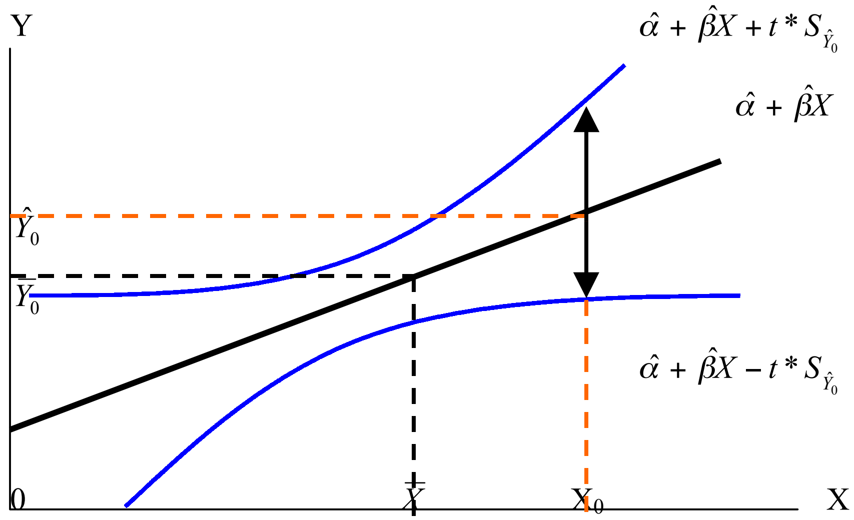
  $E[Y_t] = \exp(b_1) t\wedge b_2 E[\exp(u_t)] \neq \exp(b_1) t\wedge b_2$, because $E[u_t] = 0$ does not imply $E[\exp(u_t)] = 1$. However, it is possible to estimate $E[\exp(u_t)]$ by using the fact that $E[\exp(u_t)] = \exp\left(\frac{o^2}{2}\right)$ of which the estimate is $\exp\left(\frac{\hat{o}^2}{2}\right)$.

2. **Confidence Interval for $\hat{Y}$**

   $E[\hat{Y} | X = X_0] = E[\hat{\alpha}] + E[\hat{\beta}] X_0 = \alpha + \beta X_0 = E[Y | X = X_0]$ However, since $\hat{\alpha}$ *and* $\hat{\beta}$ are estimated with imprecision, $\hat{Y}_0$ is also subject to error.

$$S_{\hat{Y}_0}^2 = \hat{\sigma}^2\left[\frac{1}{n} + \frac{(X_0 - \overline{X})^2}{S_X^2}\right] \rightarrow \text{C.I. for } \hat{Y}_0 = \hat{Y}_0 \pm t^* S_{\hat{Y}_0}\text{, where t* is the critical t.}$$

The farther away X_0 is from $\overline{X}$, the larger $S_{\hat{Y}_0}$ is and the wider the C.I. is for $\hat{Y}_0$.



Confidence Interval for the Point Forecast

$$S^2 = \hat{\sigma}^2\left[1 + \frac{1}{n} + \frac{(X_0 - \overline{X})^2}{S_X^2}\right] > S_{\hat{Y}_0}^2 \quad (\because \hat{u}_0 = Y_0 - \hat{Y}_0) \rightarrow \text{C.I. for } \hat{u}_0 = \hat{u}_0 \pm t^* S$$

Proof)

$$\sigma^2 = \sigma_{Y_0}^2 + \sigma_{\hat{Y}_0}^2 - 2\sigma_{Y_0,\hat{Y}_0} = \sigma_{Y_0}^2 + \sigma_{\hat{Y}_0}^2\left[\frac{1}{n} + \frac{(X_0 - \overline{X})^2}{S_X^2}\right] + 0 = \sigma^2\left[1 + \frac{1}{n} + \frac{(X_0 - \overline{X})^2}{S_X^2}\right],$$

where $\sigma^2 = \sigma_{Y_0}^2 \quad \because Y_0 = \alpha + \beta X_0 + u_0 \quad and \quad \sigma_{Y_0,\hat{Y}_0} = 0 \ \because \sigma_{U_0,\hat{U}_0} = 0$

As $n \rightarrow \infty$, $S^2 = \hat{\sigma}^2[1 + 0 + 0] = \hat{\sigma}^2 \rightarrow$ Large sample C.I. for $\hat{Y}_0 = \hat{Y}_0 + 2\hat{\sigma}$.

3. **Deterndng**

a. Simple Time-Detrending

   i) Fit a long-term trend for the dependent variable, remove the trend, and obtain the residual $\hat{u}_t$.

   ii) Relate the residuals to short-run variables.

b. Multicollinearity Detrending (Stepwise Regression)

Given $Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t$, if Y_t and X_{2t}, X_{1t} and X_{2t} trend together, and if b_1 is the parameter of our interest:

i) Regress $Y_t = \mu X_{2t} + \upsilon_t$, and $X_{1t} = mX_{2t} + v_t$ respectively and obtain fitted values $\hat{Y}_t = \hat{\mu} X_{2t}$ and $\hat{X}_{1t} = \hat{m} X_{2t}$.

ii) Let $Y^* = Y_t - \hat{Y}_t$ and $X^* = X_{1t} - \hat{X}_{1t}$. Now, Y* is detrended and free of any influence of X_{2t} and X* is detrended and free of any influence of X_{2t}.

iii) Regressing Y* on X* will give parameter estimates free of time trend between Y_t and X_{2t}, X_{1t} and X_{2t}.

4. Smoothing Time Series

a. Moving Average Smoothing: Interested only in underlying trend, moving average smoothing can reduce short-term volatility of the series.

$$Y_t = \frac{1}{m} \sum (X_t + X_{t-1} + \cdots + X_{t-m+1})$$

b. Exponential Smoothing: A new series is obtained as a weighted average of present and past values with geometrically declining weights.

$$Y_t = \lambda [X_t + (1-\lambda)X_{t-1} + (1-\lambda)^2 X_{t-2} + \cdots]$$

$$Y_{t-1} = \lambda [X_{t-1} + (1-\lambda)X_{t-2} + (1-\lambda)^2 X_{t-3} + \cdots]$$

$$Y_t = \lambda X_t + (1-\lambda)Y_{t-1}$$

5. Econometric Forecasting

a. No Lagged Dependent Variables but Serially Correlated Errors

Let $u_t = \rho u_{t-1} + \varepsilon_t \rightarrow \hat{u}_{t+1} = \hat{\rho} u_t, \quad \hat{u}_{t+2} = \hat{\rho}^2 u_t, \quad \hat{u}_{t+h} = \hat{\rho}^h u_t$

Then, $Y_{t+h} = \beta_1 + \beta_2 X_{t+h,2} + \beta_3 X_{t+h,3} \cdots + \beta_k X_{t+h,k} + \hat{\rho}^h u_t$

In the general case of a q-th order AR error structure,

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_q u_{t-q} + \varepsilon_t$$

A one-period-ahead forecast error is estimated as $\hat{u}_{t+1} = \hat{\rho}_1 u_t + \hat{\rho}_2 u_{t-1} + \cdots + \hat{\rho}_q u_{t-q+1}$

Then, $\hat{Y}_{t+1} = \hat{\beta}_1 + \hat{\beta}_2 X_{t+1,2} + \hat{\beta}_3 X_{t+1,3} + \cdots + \hat{\beta}_k X_{t+1,k} + \hat{u}_{t+1}$

Subsequent forecast can be generated in a similar way.

b. Lagged Dependent Variables and Serially Correlated Errors.

$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + \beta_1 X_{t1} + \cdots \beta_k X_{tk} + u_t$, where

$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_q u_{t-q} + \varepsilon_t$

For given values of $X_{t+1,1}, X_{t+1,2}, \cdots X_{t+1,k}$, a one-period-ahead forecast is given by

$$\hat{Y}_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 Y_t + \hat{\alpha}_2 Y_{t-1} + \cdots + \hat{\alpha}_p Y_{t-p+1} + \hat{\beta}_1 X_{t+1,2} + \hat{\beta}_2 X_{t+1,2} + \hat{\beta}_3 X_{t+1,3} + \cdots + \hat{\beta}_k X_{t+1,k} + \hat{u}_{t+1}$$

where $\hat{u}_{t+1} = \hat{\rho}_1 u_t + \hat{\rho}_2 u_{t-1} + \cdots + \hat{\rho}_q u_{t-q+1}$

6. Time Series Forecasting

A time series often consists of i) deterministic trend term, ii) seasonal term, and iii) stochastic term $Y_t = T_t + S_t + u_t$ or $Y_t = T_t S_t u_t$, where $T_t = \alpha + \beta\, t$.

a. ARMA(*p,q*) or ARIMA(*p,d,q*) Model

$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_p Y_{t-p} + u_t - \beta_1 u_{t-1} - \beta_2 u_{t-2} - \cdots - \beta_q u_{t-q}$

If the series had to be differenced *d* times to be stationary, it is said to be integrated of order "d", and the resulting model is called ARIMA(*p,d,q*).

b. Autocorrelation Function (ACF), Partial AF (PACF) & Correlogram

According to Box-Jenkins, $\rho(s) = \rho_{u_t,u_{t-s}} = \dfrac{\sigma_{u_t,u_{t-s}}}{\sigma_{u_t}^2} = \dfrac{E[u_t, u_{t-s}]}{E[u_t^2]}$ and the resulting plot (Correlogram) are used to identify the specification of *p,d,q*.

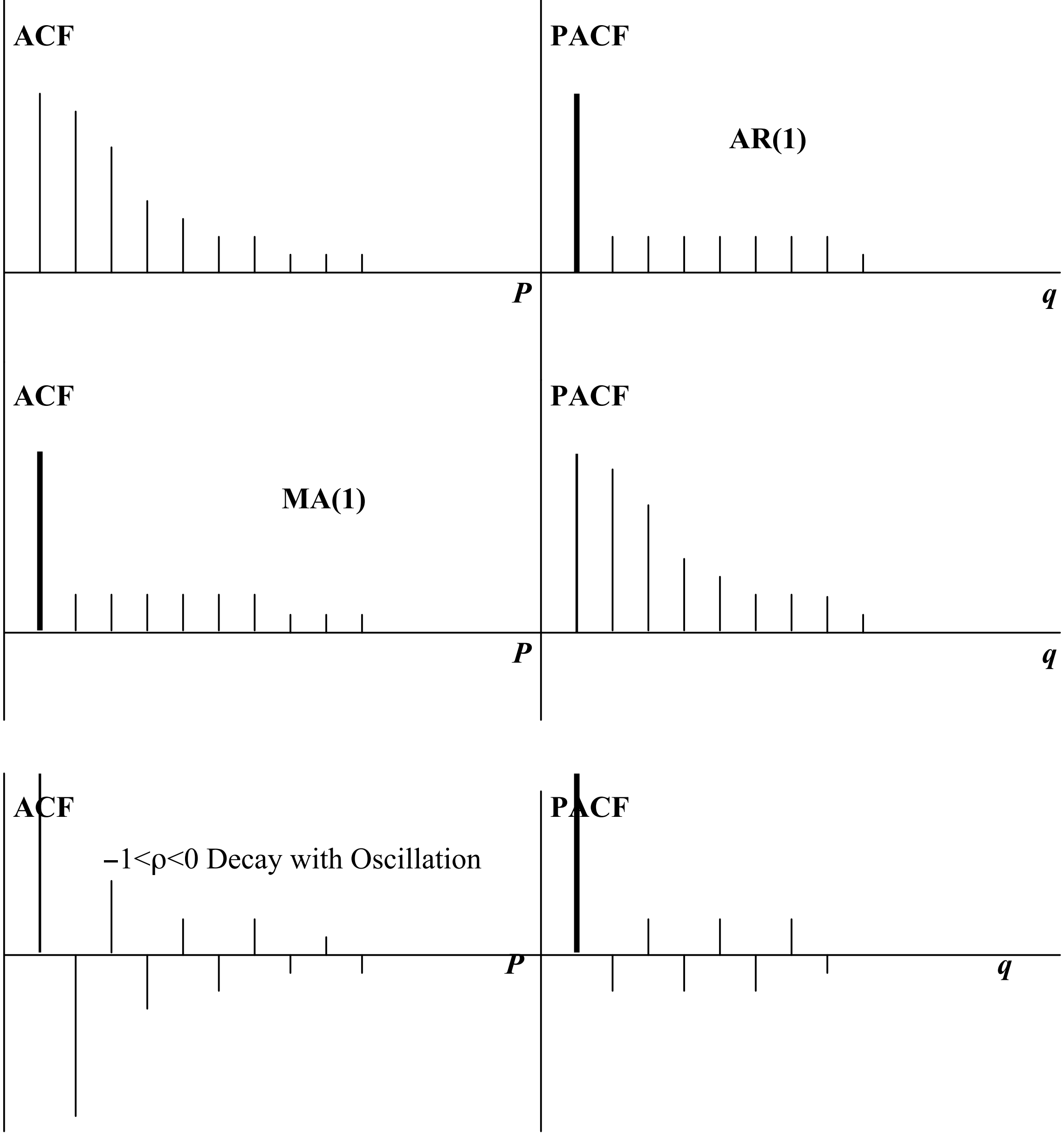
i) For large values of the lag (k), the theoretical ACF of AR(*p*) is approximately $A\rho^k$, where $|\rho| < 1$. The correlogram of sample ACF dies out (decays) with dampening pattern. The correlogram of sample PACF has spikes up to lag *p* and then cuts off immediately and remains near zero thereafter. If $-1 < \rho < 0$, the ACF will die out with oscillation.

ii) For MA(q), the correlogram of sample ACF cuts off immediately after lag *q* and remains near zero thereafter. The correlogram of sample PACF dies out (decays) with dampening pattern.

c. Nonstationarity due to seasonal effects is handled by deseasonalizing the series. This will be evident from seasonal spikes at regular intervals in ACF.

i) Remove the seasonal component in a monthly data series by differencing $Y_t - Y_{t-12}$.

ii) Alternatively, use dummy variable, and obtain the residuals free of seasonal effects.



d.. Diagnostics for correct model specification can be done by out-of-sample testing. (R)MSE and AIC are commonly used summary statistics. If the model fits the data well, then the residuals should be white-noise. This can be checked by Box-Pierce Q-statistic. $Q = n\sum_{k=1}^{m} \hat{\rho}_k^2 \sim \chi^2 \; w/m \; df, \quad where \;\; \hat{\rho}_k = \dfrac{\sum_{t=k+1}^{n} e_t e_{t-k}}{\sum_{t=1}^{n} e_t^2}$

- **SAS ARIMA Code**

```
filename ip 'c:\jeff\intl paper\ip.prn';
data ip;
infile ip;
input  quarter rit  delta1  delta  chgdiv chgrm invcons;
rit=rit-lag(rit);         /*or rit(1) in i var=rit(1) step*/
proc arima data=ip;
i var=rit nlag=30;   /*nlag must be w/in 25% of the data series.*/
run;

e p=(1,2,4) q=1;      /*p=(n) (m) for seasonal adjustment*/
run;
f lead=12;
run;
```