THEORETICAL AUTOCORRELATIONS

$$\rho_j = \frac{Cov(y_t, y_{t-j})}{Var(y_t)} = \frac{E(y_t - E(y_t))(y_{t-j} - E(y_t))}{E(y_t - E(y_t))^2}$$

$$j = 1, 2, \cdots$$

$\rho_0 = 1$ and $Cov(y_t, y_{t-j})$ is often denoted by γ_j
while $Var(y_t)$ if often denoted by γ_0. Note that
$\gamma_j = \gamma_{-j}$ and $\rho_j = \rho_{-j}$ and because of this symmetry
the theoretical autocorrelation function and the sample autocorrelation function (below) only need be examined over the positive lags
$j = 1, 2 \cdots$.

SAMPLE AUTOCORRELATIONS

$$r_j = \frac{\sum_{t=j+1}^{T} (y_t - \overline{y})(y_{t-j} - \overline{y})}{\sum_{t=1}^{T} (y_t - \overline{y})^2}, \; j = 1, 2, \cdots.$$

The r_j are consistent estimators of the theoretical autocorrelation coefficients ρ_j. Under the assumption that y_t follows a white noise process the standard errors of these r_j are approximately equal to $1/\sqrt{T}$. Thus, under the null hypothesis that y_t follows a white noise process, roughly 95% of the r_j should fall within the range of $\pm 1.96/\sqrt{T}$. If more than 5% of the r_j fall outside of this range, then most likely y_t does not follow a white noise process.

THEORETICAL PARTIAL AUTOCORRELATIONS

$$\phi_{jj} = \frac{Cov(y_t, y_{t-j} \mid y_{t-1}, \cdots, y_{t-j+1})}{Var(y_t \mid y_{t-1}, \cdots, y_{t-j+1})}$$

$$= \frac{E(y_t - E(y_t \mid y_{t-1}, \cdots, y_{t-j+1}))(y_{t-j} - E(y_t \mid y_{t-1}, \cdots, y_{t-j+1}))}{E(y_t - E(y_t \mid y_{t-1}, \cdots, y_{t-j+1}))^2}$$

$$j = 1, 2, \cdots$$

= the correlation between y_t and y_{t-j} after netting out the effects the intervening values $y_t, \cdots, y_{t-j+1}$ have on both of them

SAMPLE PARTIAL AUTOCORRELATIONS

$$\hat{\phi}_{jj}$$

are calculated using the formulas for the theoretical autocorrelations for a given ARMA(p,q) model (see my ACF_PACF_Table.doc Word document for the formulas) but replacing all of the theoretical autocorrelations (ρ_j) with the above sample autocorrelations (r_j) and all of the unknown Box-Jenkins coefficients (ϕ_i, θ_i) with their corresponding estimates ($\hat{\phi}_i, \hat{\theta}_i$) obtained by the method of moments or some other method. The $\hat{\phi}_{jj}$ are consistent estimators of the theoretical partial autocorrelations, ϕ_{jj} . Under the assumption that y_t follows a white noise process the standard errors of these $\hat{\phi}_{jj}$ are approximately equal to $1/\sqrt{T}$. Thus, under the null hypothesis that y_t follows a white noise process, roughly 95% of the $\hat{\phi}_{jj}$ should fall within the range of $\pm 1.96/\sqrt{T}$. If more than 5% of the $\hat{\phi}_{jj}$ fall outside of this range, then most likely y_t does not follow a white noise process.

GOODNESS-OF-FIT MEASURES

1. **AIC (Akaike Information Criterion)**

$$AIC = -2L(\sum \hat{a}_t^2) + 2K$$

where K = p + q + 1, L($\sum \hat{a}_t^2$) = the log of the likelihood function of the Box-Jenkins ARMA(p,q) model, $\hat{a}_t$ = the residual at time t for the Box-Jenkins model and the log likelihood function, $L(\sum \hat{a}_t^2)$, is a monotonically decreasing function of the sum of squared residuals, $\sum \hat{a}_t^2$. In other words, the smaller $\sum \hat{a}_t^2$ is, the larger L($\sum \hat{a}_t^2$) is and vice versa.

2. **SBC (Schwartz Bayesian Criterion)**

$$SBC = -2L(\sum \hat{a}_t^2) + K\ln(n) \quad SBC = -2L(\sum \hat{a}_t^2) + K\ln(n)$$

where n is the number of residuals computed for the model.

In terms of choosing a Box-Jenkins model, the smaller these goodness-of-fit measures, the better. That is, we prefer the Box-Jenkins model that has the **smallest** AIC and SBC measures. Notice that, as you add coefficients to the Box-Jenkins model, (ϕ_i, θ_i), the fit of the model, as measured by the sum of squared residuals, $\sum \hat{a}_t^2$, **always** decreases and, therefore, adding coefficients **always** increases the log likelihood, $L(\sum \hat{a}_t^2)$, of the Box-Jenkins model. To offset the tendency for adding coefficients to a model just to improve its fit, the above goodness-of-fit (information) criteria each include a "penalty" term. (For the AIC criterion the penalty term is +2K while for the SBC measure the penalty term is +Kln(T). Thus, with these criteria, as one adds coefficients to the Box-Jenkins model, the improvement in fit coming from reduction in the sum of squared residuals will eventually be offset by the penalty term moving in the opposite direction. The goodness-of-fit criteria are then intended to keep us from building large order Box-Jenkins models just to improve the fit just to find that such large order models don't forecast very well. Shibata (1976) has shown that, for a finite-order AR process, the AIC criterion asymptotically overestimates the order with positive probability. Thus, an estimator of the AR order (p) based on AIC will **not** be consistent. (By consistent we mean that, as the sample size goes to infinity, the correct order of an AR(p) Box-Jenkins model will be correctly chosen with probability one.) In contrast, the SBC criterion is consistent in choosing the correct order of an AR(p) model. Often these two criteria choose the same Box-Jenkins model as being the best model. However, when there is a difference in choice, the AIC measure invariably implies a Box-Jenkins model of bigger order (K = p + q + 1) than the order of the model implied by the SBC criterion. In other words, the SBC criterion tends to pick the more parsimonious model when there is a "split" decision arising from using these criteria. Personally, I prefer to rely on the SBC criterion in the case of "split" decisions.

A TEST FOR WHITE NOISE RESIDUALS

(and thus the Box-Jenkins model's "completeness")

H_0: Residuals of Estimated Box-Jenkins model are white noise (i.e. uncorrelated at all lags). Other things held constant, the estimated Box-Jenkins model is adequate.

H_1: Residuals of Estimated Box-Jenkins model are **not** white noise. In this case, a better model can be found by adding more parameters to the model.

The chi-square test used to test for white noise residuals is calculated as

$$\chi_m^2 = n(n+2)\sum_{j=1}^{m} \frac{r_j^2(\hat{a}_t)}{(n-j)}$$

Where

$$r_j(\hat{a}_t) = \frac{\sum_{t=1}^{n-j} \hat{a}_t \hat{a}_{t+j}}{\sum_{t=1}^{n} \hat{a}_t^2}$$

n = number of residuals, and $\hat{a}_t$ is the time t residual of the Box-Jenkins model. This statistic was suggested by Ljung and Box (1978) and is called the **Ljung-Box chi-square statistic** for testing for white noise residuals. The null hypothesis above is accepted if the observed chi-square statistic is small (i.e. has a probability value greater than 0.05) and is rejected if the chi-square statistic is "large" (i.e. has a probability value less than 0.05). As far as the choice of the number of lags, m, to use, I would suggest m = 12 for quarterly data and m = 24 for monthly data to increase the power of the test given the frequency with which the data is observed.

CONSTRUCTION OF THE P-Q BOX

In this class we will be constructing a "P-Q Box" of the form

		Q		
		0	1	2
	0	.	.	.
P	1	.	.	
	2	.		

where **.** represents the following numbers in each cell: AIC, SBC, χ_m^2, and the p-value of the Ljung-Box chi-square statistic, χ_m^2. These cells represent the most prevalent Box-Jenkins models that apply to **non-seasonal** economic time series data, namely, the ARMA(0,0), AR(1), AR(2), MA(1), MA(2), and ARMA(1,1) models. Using the sample ACF and sample PACF of the data one can often narrow down the choice between these cell (models) but not always with certainty. Thus, the P-Q Box can often help confirm which Box-Jenkins model is best for the data. The model with the lowest AIC and SBC measures and having white noise residuals is the model that the P-Q Box statistics suggest. Hopefully, after looking at the sample ACF and sample PACF and the P-Q Box results one can come to a tentative choice for the p and q orders of the Box-Jenkins model.

OVERFITTING EXERCISE

To confirm the choice of model suggested by the sample ACF, sample PACF, and the P-Q Box, one should conduct an **overfitting exercise**. That is, you should fit two additional Box-Jenkins models, one having **one more** autoregressive coefficient and one having **one more** moving average coefficient and then examining (individually) the statistical significance of the extra coefficient in each model. For example, if your tentative choice is p = 1 and q = 0 (an AR(1) model), you should examine the AR2 coefficient in an AR(2) model and determine whether this "overfitting" coefficient is statistically significant or not. If it is not statistically significant (i.e. the p-value is > 0.05), you can "fall" back to your original choice. The other overfitting model for the AR(1) model is the ARMA(1,1) model. So when you fit it, the overfitting parameter is the MA1 parameter. If it is not statistically significant, then you can "fall" back to your original "almost final" choice again and make it your "final" choice for forecasting purposes. Of course, if either of the overfitting parameters is statistically significant, you need to continue the model building process.