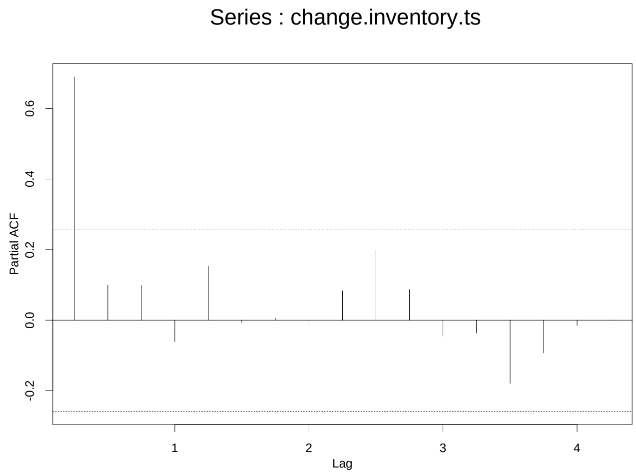
Statistics 451 Examination 1 Name ________________________

Spring 1999 **You must show all of your work**

When asked to explain something, provide an explanation that could be understood by someone who does not have formal training in statistical methods.

1. It is customary to plot the sample ACF and PACF functions showing spikes for each lag going above or below the zero line, as shown in the following figure.



This is because $\rho_k = 0$ and $\phi_{kk} = 0$ are often in mind as a null hypothesis values when using these functions to make decisions in model building.

(a) Give an expression for the standard error of $\widehat{\phi}_{kk}$.

(b) Briefly explain what happens to the standard error of $\widehat{\phi}_{kk}$ as the realization size n get larger.

(c) In order to help interpret a plot of a sample ACF or PACF, it is useful to plot warning limits such that it a spike protrudes through the limits, one can say that the estimate is "significantly different from 0." Briefly explain how such warning limits should be computed and plotted.

2. For the ARMA model $\phi_2(\mathrm{B})\dot{Z}_t = \theta_2(\mathrm{B})a_t$ written in compact form, write down the "unscrambled" model for Z_t, including a constant term that is related to the non-zero mean of the Z_t time series.

3. Probability plots are useful tools for data analysis.

   (a) Briefly explain the primary purpose for using a normal probability plot to examine the residuals from a regression analysis.

   (b) Briefly explain how simulation can be used to help interpret a normal probability plot.

   (c) Given that the central limit theorem tells us that sample statistics like means and regression coefficients have sampling distributions that can be approximated by normal distributions, why do we care about the normality of residuals?

4. Hypothesis testing is a widely used (and abused) statistical technique.

   (a) Briefly explain why the usual assumptions behind hypothesis testing do not hold during the process of model identification.

   (b) Briefly explain the proper role of hypothesis testing in model identification.

5. Consider the MA(1) model $Z_t = \theta_0 + \theta_1 a_{t-1} + a_t$ where $a_t \sim \text{nid}(0, \sigma_a^2)$.

   (a) Derive an expression for the mean of Z_t.

   (b) Derive an expression for the variance of Z_t.

   (c) Derive an expression for the autocovariance between Z_t and Z_{t-1}.

   (d) Why is the autocorrelation easier to interpret than the autocovariance?

6. For some purposes it is useful to decompose a time series in component parts known as trend, seasonal (or periodic), cyclical, and irregular (or residual).

   (a) Briefly explain the difference between a cyclical component and a periodic component of a time series.

   (b) Briefly explain why it is important to distinguish between the cyclical and periodic components of a time series.

7. It is easy to show that simple random sampling with replacement from a finite population leads to iid (independent and identically distributed) observations. Furthermore, if the population size is large relative to the sample size, sampling without replacement (as is commonly used in practice) provides a good approximation to sampling with replacement. Briefly explain why time series data, obtained by sampling from a process over time, cannot be expected to correspond to iid observations.

8. For each of the following differencing schemes, write down W_t as a function of past and present values of Z_t.

   (a) $W_t = (1 - \mathrm{B})^2 Z_t$

   (b) $W_t = (1 - \mathrm{B})(1 - \mathrm{B}^5)^1 Z_t$

9. Briefly explain the practical interpretation of the sample autocorrelation $\widehat{\rho}_1$.