Ch3. TRENDS

Time Series Analysis
3.1 Deterministic Versus Stochastic Trends

The simulated random walk in Exhibit 2.1 shows a “upward trend”. However, it is caused by a strong correlation between the series at nearby time points. The true process is $Y_t = Y_{t-1} + e_t$ or $Y_t = \sum_{i=1}^{t} e_t$, which has zero mean and increasing variance over time. We call such a trend a stochastic trend.

By comparison, the average monthly temperature series in Exhibit 1.7 shows a cyclical trend. We may model it by $Y_t = \mu_t + X_t$, where $\mu_t$ is a deterministic function with period 12 ($\mu_t = \mu_{t-12}$), and $X_t$ is a zero mean unobserved variation (or noise). We say that this model has a deterministic trend.
3.2 Estimation of a Constant Mean

Consider a model with a constant mean function

\[ Y_t = \mu + X_t, \quad E(X_t) = 0. \]

The most common estimate of \( \mu \) is the sample mean

\[ \overline{Y} = \frac{1}{n} \sum_{t=1}^{n} Y_t. \]

Then \( E(\overline{Y}) = \mu \). So \( \overline{Y} \) is a unbiased estimator of \( \mu \). If \( \{Y_t\} \) is stationary with acf \( \rho_k \), Ex 2.17 shows that

\[
\text{Var}\overline{Y} = \frac{\gamma_0}{n} \left[ 1 + 2 \sum_{k=1}^{n-1} \left( 1 - \frac{k}{n} \right) \rho_k \right]. \tag{1}
\]

For many stationary processes, the acf decays quickly that

\[ \sum_{k=0}^{\infty} |\rho_k| < \infty. \]

Under this assumption, we have

\[
\text{Var}\overline{Y} \approx \frac{\gamma_0}{n} \left[ \sum_{k=-\infty}^{\infty} \rho_k \right] \quad \text{for large } n.
\]
So for many stationary processes \( \{Y_t\} \), the variance of \( \overline{Y} \) is inversely proportional to \( n \). Increasing \( n \) will improve the estimation.

**Ex.** We use (1) to investigate \( \overline{Y} \):

1. If \( \{X_t\} \) is white noise, then \( \rho_k = 0 \) for \( k > 0 \). So \( \text{Var} \overline{Y} = \frac{\gamma_0}{n} \).

2. If \( \rho_1 \neq 0 \) but \( \rho_k = 0 \) for \( k > 1 \), then \( \text{Var} \overline{Y} \to \frac{\gamma_0}{n} (1 + 2\rho_1) \) as \( n \to \infty \).

3. If \( \{Y_t\} \) is a MA(1) model: \( Y_t = e_t + \theta e_{t-1} \), then \( \rho_1 = \frac{\theta}{1 + \theta^2} \) and \( \rho_k = 0 \) for \( k > 1 \). We have

\[
\text{Var} \overline{Y} = \frac{\gamma_0}{n} \left[ 1 + \frac{2\theta}{1 + \theta^2} \left( \frac{n - 1}{n} \right) \right] \approx \frac{(1 + \theta)^2 \gamma_0}{1 + \theta^2} \frac{n}{n}.
\]

- When \( -1 < \theta < 0 \), say \( \theta = -0.5 \), \( \text{Var} \overline{Y} \approx \frac{0.2 \gamma_0}{n} \). Comparing with the white noise, the negative correlation \( \rho_1 \) improves the estimation significantly.

- When \( 0 < \theta < 1 \), say \( \theta = +0.5 \), \( \text{Var} \overline{Y} \approx \frac{1.8 \gamma_0}{n} \). The positive correlation \( \rho_1 \) reduces the accuracy of estimation.
Ex. For an order one autoregressive process

\[ Y_t = \phi Y_{t-1} + e_t, \quad -1 < \phi < 1, \]

we can show that \( \rho_k = \phi^{|k|} \) for all \( k \). Then

\[
\text{Var} \bar{Y} \approx \frac{\gamma_0}{n} \left[ \sum_{k=-\infty}^{\infty} \phi^{|k|} \right] = \frac{(1 + \phi)}{(1 - \phi)} \frac{\gamma_0}{n}.
\]

We plot in R the function \( \frac{(1 + \phi)}{(1 - \phi)} \) for \(-1 < \phi < 1\):

```r
phi=seq(-.95,.95,0.05)
factor=(1+phi)/(1-phi)
plot(y=factor,x=phi,type='l')
abline(h=1,col='red')
```

When \( \phi \to 1^- \), the variance of \( \bar{Y} \) dramatically increases. In fact, \( \phi = 1 \) is the nonstationary random walk process \( Y_t = Y_{t-1} + e_t \). The variance of \( \bar{Y} \) can be very different for a nonstationary process.
3.3 Regression Methods

We use classical regression analysis to estimate the parameters in the nonconstant mean trend models \( Y_t = \mu_t + X_t \), where \( \mu_t \) is deterministic and \( \mathbb{E}(X_t) = 0 \). These include linear, quadratic, seasonal means, and cosine trends.

3.3.1 Linear and Quadratic Trends: The deterministic trend linear model is the most widely used model:

\[
\mu_t = \beta_0 + \beta_1 t
\]

we call \( \beta_0 \) the **intercept**, and \( \beta_1 \) the **slope**, resp. The classical least squares regression is to find \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \) to minimize

\[
Q(\beta_0, \beta_1) = \sum_{t=1}^{\infty} [Y_t - (\beta_0 + \beta_1 t)]^2.
\]

Use partial derivatives to solve that (formulas are not required):

\[
\hat{\beta}_1 = \frac{\sum_{t=1}^{n}(Y_t - \overline{Y})(t - \overline{t})}{\sum_{t=1}^{n}(t - \overline{t})^2}, \quad \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{t}, \quad \text{where} \quad \overline{t} = \frac{n + 1}{2}.
\]
**Exhibit. SCL in the Capital Asset Pricing Model (CAPM)**

Security characteristic line

From Wikipedia, the free encyclopedia

**Security characteristic line** (SCL) is a regression line,[1] plotting performance of a particular security or portfolio against that of the market portfolio at every point in time. The SCL is plotted on a graph where the Y-axis is the excess return on a security over the risk-free return and the X-axis is the excess return of the market in general. The slope of the SCL is the security's beta, and the intercept is its alpha.[2]

Content: [hide]
1 Formula
2 See also
3 References
4 External links

**Formula** [edit]

\[ SCL : R_{i,t} - R_f = \alpha_i + \beta_i (R_{M,t} - R_f) + \epsilon_{i,t} \]

where:

- \( \alpha_i \) is called the asset's alpha (abnormal return)
- \( \beta_i \) is a non-diversifiable or systematic risk
- \( \epsilon_{i,t} \) is the non-systematic or diversifiable, non-market or idiosyncratic risk
- \( R_{M,t} \) is a market risk
- \( R_f \) is a risk-free rate

**Positive abnormal return (\( \alpha \))**: Above-average returns that cannot be explained as compensation for added risk

**Negative abnormal returns (\( \alpha \))**: Below-average returns that cannot be explained by below-market risk
The function `lm` fits a linear model (a regression model) with its first argument being a formula, i.e., an expression including a tilde sign (`∼`), the left-hand side of which is the response variable and the right-hand side are the covariates or explanatory variables (separated by plus signs if there are two or more covariates). The intercept term can be removed by adding the term “-1” on the right-hand side of the tilde sign. The covariates must be separated with a plus sign (`+`).
The function `abline` is a low-level graphics function. If a fitted simple regression model is passed to it, it adds the fitted straight line to an existing graph. Any straight line of the form $y = \beta_0 + \beta_1 x$ can be superimposed on the graph by running the command

```
abline(a=beta0, b=beta1)
```
**Exhibit.** Suppose we want to fit a quadratic time trend model to the rwalk series. We need to create a new covariate that contains the square of the time indices. The quadratic variable may be created before invoking the `lm` function. Or it may be created on the fly when invoking the `lm` function. The latter approach is illustrated here.

```r
model1a=lm(rwalk~time(rwalk)+I(time(rwalk)^2))
summary(model1a)
```

The expression `time(rwalk)^2` is enclosed within the `I` function which instructs R to create a new variable by executing the command passed into the `I` function. Without the `I` function, R will use formula convention to interpret the term.

In the summary, all significant covariates are marked with asterisks (*): more asterisks means higher significance, that is, smaller p-value, as explained in the line labeled as Signif. codes. We see that the quadratic term is not significant so that it is not needed.
3.3.2 Cyclical or Seasonal Trends

Suppose that a seasonal trend, such as the average monthly temperature data in Exhibit 1.7, can be modeled as

\[ Y_t = \mu_t + X_t, \quad E(X_t) = 0 \text{ for all } t. \]

One suitable model may be a **seasonal mean model**, where there are 12 constants (parameters), \( \beta_1, \beta_2, \cdots, \beta_{12} \), giving the expected average temperature for each of the 12 months, such that

\[
\mu_t = \begin{cases} 
\beta_1 & t = 1, 13, 25, \cdots \\
\beta_2 & t = 2, 14, 26, \cdots \\
& \vdots \\
\beta_{12} & t = 12, 24, 36, \cdots 
\end{cases}
\]
As an example of this model consider the average monthly temperature data shown in Exhibit 1.7 on page 6. To fit such a model, we need to set up indicator variables (sometimes called dummy variables) that indicate the month to which each of the data points pertains. The procedure for doing this will depend on the particular statistical software that you use. We also need to note that the model as stated does not contain an intercept term, and the software will need to know this also. Alternatively, we could use an intercept and leave out any one of the $\beta$’s in Equation (3.3.3).

Exhibit 3.3 displays the results of fitting the seasonal means model to the temperature data. Here the $t$-values and $Pr(>|t|)$-values reported are of little interest since they relate to testing the null hypotheses that the $\beta$’s are zero—not an interesting hypothesis in this case.

Exhibit 3.4 shows how the results change when we fit a model with an intercept term. The software omits the January coefficient in this case. Now the February coefficient is interpreted as the difference between February and January average temperatures, the March coefficient is the difference between March and January average temperatures, and so forth. Once more, the $t$-values and $Pr(>|t|)$ ($p$-values) are testing hypotheses of little interest in this case. Notice that the Intercept coefficient plus the February coefficient here equals the February coefficient displayed in Exhibit 3.3.

cap3.R:

- Exhibit 3.3 Regression Results for the Seasonal Means Model
- Exhibit 3.4 Results for Seasonal Means Model with an Intercept
R: The expression `season(tempdub)` outputs the monthly index of `tempdub` as a factor, and saves it into the object `month`. A **factor** is a kind of data structure for handling qualitative (nominal) data that do not have a natural ordering like numbers do. However, for purposes of summary and graphics, the user may supply the `levels` argument to indicate an ordering among the factor values.

See TS-ch3.R

See TS-ch3.R

The model `milk.m2` is a combination of linear model with seasonal mean model. We may decompose a time series into a sum or product of trend, seasonal, and stochastic components by “decompose” command.

See TS-ch3.R
3.3.3 Cosine Trends:
In some cases, seasonal trends can be modeled economically with cosine curves (that incorporate the smooth change expected from one time period to the next while still preserving the seasonality.) The simplest such model is \( Y_t = \mu_t + X_t \) with

\[
\mu_t = \beta_0 + \beta \cos(2\pi ft + \Phi) = \beta_0 + \beta_1 \cos(2\pi ft) + \beta_2 \sin(2\pi ft)
\]

where \( \beta_1 = \beta \cos \Phi \) and \( \beta_2 = -\beta \sin \Phi \). We call \( \beta (> 0) \) the amplitude, \( f \) the frequency, and \( \Phi \) the phase of the curve. \( 1/f \) is called the period of the cosine wave.
In any practical example, we must be careful how we measure time, as our choice of time measurement will affect the values of the frequencies of interest. For example, if we have monthly data but use 1, 2, 3,... as our time scale, then 1/12 would be the most interesting frequency, with a corresponding period of 12 months. However, if we measure time by year and fractional year, say 1980 for January, 1980.08333 for February of 1980, and so forth, then a frequency of 1 corresponds to an annual or 12 month periodicity.

Exhibit 3.5 is an example of fitting a cosine curve at the fundamental frequency to the average monthly temperature series.

- Exhibit 3.5 Cosine Trend Model for Temperature Series (chap3.R)

In this output, time is measured in years, with 1964 as the starting value and a frequency of 1 per year. A graph of the time series values together with the fitted cosine curve is shown in Exhibit 3.6. The trend fits the data quite well with the exception of most of the January values, where the observations are lower than the model would predict.

- Exhibit 3.6 Cosine Trend for the Temperature Series (chap3.R)

The harmonic function will return a matrix of the values of \( \cos(2\pi t) \) and \( \sin(2\pi t) \).

See TS-ch3.R for some basic commands of matrix
3.4 Reliability and Efficiency of Regression Estimates

Reliability and efficiency of estimates lie on the variances of the estimates. Let us analyze the reliability of estimates for some deterministic trends.

Assumption: the series is represented as $Y_t = \mu_t + X_t$, where

1. $\mu_t$ is a deterministic trend of the kind linear, seasonal means, or cosine curves, and
2. $\{X_t\}$ is a zero-mean stationary process with autocovariance and autocorrelation functions $\gamma_k$ and $\rho_k$, respectively.

Parameters in a linear model are estimated according to ordinary least squares regression.
3.4.1 The seasonal means
For \( N \) complete years of monthly data, the estimate for the mean for the \( j \)th month is

\[
\hat{\beta}_j = \frac{1}{N} \sum_{i=0}^{N-1} Y_{j+12i}
\]

By (1),

\[
\text{Var} \hat{\beta}_j = \frac{\gamma_0}{N} \left[ 1 + 2 \sum_{k=1}^{N-1} \left( 1 - \frac{k}{N} \rho_{12k} \right) \right], \quad \text{for } j = 1, 2, \ldots, 12.
\]

(2)

Only the seasonal autocorrelations, \( \rho_{12}, \rho_{24}, \ldots \), enter into the above equation. If \( \{X_t\} \) is white noise, then \( \text{Var} \hat{\beta}_j = \frac{\gamma_0}{N} \).
3.4.2 The cosine trends $\mu_t = \beta_0 + \beta_1 \cos(2\pi ft) + \beta_2 \sin(2\pi ft)$

For any frequency of the form $f = m/n$, where $m$ is an integer satisfying $1 \leq m < n/2$, the estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ are

$$
\hat{\beta}_1 = \frac{2}{n} \sum_{t=1}^{n} \left[ \cos \left( \frac{2\pi mt}{n} \right) Y_t \right], \quad \hat{\beta}_2 = \frac{2}{n} \sum_{t=1}^{n} \left[ \sin \left( \frac{2\pi mt}{n} \right) Y_t \right],
$$

(3)

Because these are linear functions of $\{Y_t\}$, their variances can be computed by (1).

1. If $\{X_t\}$ is white noise, we get $\text{Var}\hat{\beta}_1 = \text{Var}\hat{\beta}_2 = 2\gamma_0/n$.

2. If $m/n = 1/12$ and $\rho_k = 0$ for $k > 1$, then

$$
\text{Var}\hat{\beta}_1 = \frac{2\gamma_0}{n} \left[ 1 + \frac{4\rho_1}{n} \sum_{t=1}^{n-1} \cos \left( \frac{\pi t}{6} \right) \cos \left( \frac{\pi t + 1}{6} \right) \right]
$$

When $n \to \infty$, we have (see (3.4.5) in the textbook)

$$
\text{Var}\hat{\beta}_1 \to \frac{2\gamma_0}{n} \left( 1 + 2\rho_1 \cos \left( \frac{\pi}{6} \right) \right) \approx \frac{2\gamma_0}{n} \left( 1 + 1.732\rho_1 \right).
$$

$\text{Var}\hat{\beta}_2$ has the same asymptotic behavior.
3.4.3 The seasonal mean trends v.s. the cosine trends
Seasonal means and cosine trends are often competing models for a cyclical trend.

Consider the two estimates for the trend in $\mu_1$ (January).

- For the seasonal means, the variance is given by (2).
- With the cosine trend model, the estimate is

$$\hat{\mu}_1 = \hat{\beta}_0 + \hat{\beta}_1 \cos \left( \frac{2\pi}{12} \right) + \hat{\beta}_2 \sin \left( \frac{2\pi}{12} \right).$$

The terms $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_2$ are uncorrelated by (3). So

$$\text{Var}\hat{\mu}_1 = \text{Var}\hat{\beta}_0 + \text{Var}\hat{\beta}_1 \left[ \cos \left( \frac{2\pi}{12} \right) \right]^2 + \text{Var}\hat{\beta}_2 \left[ \sin \left( \frac{2\pi}{12} \right) \right]^2.$$ (4)
Consider two scenarios:

1. Assume that \( \{X_t\} \) is white noise. The seasonal means model gives \( \text{Var} \hat{\mu}_1 = \gamma_0 / N \). The cosine model has \( \text{Var} \hat{\beta}_1 = \text{Var} \hat{\beta}_2 = 2\gamma_0 / n \) and so

\[
\text{Var} \hat{\mu}_1 = \frac{\gamma_0}{n} \left\{ 1 + 2 \left[ \cos \left( \frac{\pi}{6} \right) \right]^2 + 2 \left[ \sin \left( \frac{\pi}{6} \right) \right]^2 \right\} = \frac{3\gamma_0}{n}.
\]

Since \( N = 12n \), the ratio of the standard deviation in the cosine model to that in the seasonal means model is

\[
\sqrt{\frac{3\gamma_0}{n}} = \sqrt{\frac{3N}{n}} = 0.5.
\]
2. Assume that \( \{X_t\} \) has \( \rho_k = 0 \) for \( k > 1 \). The seasonal means model still has \( \text{Var} \hat{\mu}_1 = \gamma_0 / N \). The cosine trend model for large \( n \) has

\[
\text{Var} \hat{\beta}_0 \approx \frac{\gamma_0}{n} (1 + 2\rho_1), \quad \text{Var} \hat{\beta}_1 \approx \text{Var} \hat{\beta}_2 \approx \frac{2\gamma_0}{n} \left( 1 + 2\rho_1 \cos \left( \frac{\pi}{6} \right) \right).
\]

So (4) implies that

\[
\text{Var} \hat{\mu}_1 = \frac{\gamma_0}{n} \left\{ 1 + 2\rho_1 + 2 \left[ 1 + 2\rho_1 \cos \left( \frac{\pi}{6} \right) \right] \right\} = \frac{\gamma_0}{n} \left\{ 3 + 2\rho_1 \left[ 1 + 2 \cos \left( \frac{\pi}{6} \right) \right] \right\}.
\]

If \( \rho_1 = -0.4 \), then we get \( 0.814\gamma_0 / n \), and the ratio of the standard deviation in the cosine case to the standard deviation in the seasonal means case is

\[
\sqrt{\left[ \frac{(0.814\gamma_0)/n}{\gamma_0/N} \right]} = \sqrt{\frac{0.814N}{n}} = 0.26, \text{ a big reduction.}
\]
3.4.4 The linear time trends $\mu_t = \beta_0 + \beta_1 t$

The least squares estimate of the slope is

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{n} (t - \bar{t}) Y_t}{\sum_{t=1}^{n} (t - \bar{t})^2}.$$ 

Since $\hat{\beta}_1$ is a linear combination of $Y$'s, we have

$$\text{Var}\hat{\beta}_1 = \frac{12\gamma_0}{n(n^2 - 1)} \left[ 1 + \frac{24}{n(n^2 - 1)} \sum_{s=2}^{n} \sum_{t=1}^{s-1} (t - \bar{t})(s - \bar{t}) \rho_{s-t} \right]$$

For the case $\rho_k = 0$ for $k > 1$, we get

$$\text{Var}\hat{\beta}_1 = \frac{12\gamma_0}{n(n^2 - 1)} \left[ 1 + 2\rho_1 \left( 1 - \frac{3}{n} \right) \right] \approx \frac{12\gamma_0 (1 + 2\rho_1)}{n(n^2 - 1)} \text{ for large } n.$$ 

In this situation, the variance increases as $\rho_1$ increases. The case $\rho_1 = 0$ is the white noise case, where

$$\text{Var}\hat{\beta}_1 = \frac{12\gamma_0}{n(n^2 - 1)}. \quad (5)$$
3.4.5 Best linear unbiased estimates (BLUE, or GLS)

We turn now to comparing the least squares estimates with the so-called best linear unbiased estimates (BLUE) or the generalized least squares (GLS) estimates. If the stochastic component \( \{X_t\} \) is not white noise, estimates of the unknown parameters in the trend function may be made; they are linear functions of the data, are unbiased, and have the smallest variances among all such estimates—the so-called BLUE or GLS estimates. These estimates and their variances can be expressed fairly explicitly by using certain matrices and their inverses. (Details may be found in Draper and Smith (1981).) However, constructing these estimates requires complete knowledge of the covariance function of the stochastic component, a function that is unknown in virtually all real applications. It is possible to iteratively estimate the covariance function for \( \{X_t\} \) based on a preliminary estimate of the trend. The trend is then estimated again using the estimated covariance function for \( \{X_t\} \) and thus iterated to an approximate BLUE for the trend. This method will not be pursued here, however.

Fortunately, there are some results based on large sample sizes that support the use of the simpler least squares estimates for the types of trends that we have considered. In particular, we have the following result (see Fuller (1996), pp. 476–480, for more details): We assume that the trend is either a polynomial in time, a trigonometric polynomial, seasonal means, or a linear combination of these. Then, for a very general stationary stochastic component \( \{X_t\} \), the least squares estimates for the trend have the same variance as the best linear unbiased estimates for large sample sizes.
3.5 Interpreting Regression Output

In many statistical softwares, some properties of the regression output depend heavily on the regression assumption that \( \{X_t\} \) is white noise, and some depend on the further assumption that \( \{X_t\} \) is approximately normally distributed. These assumptions are false for many time series!

Therefore, we shall be very careful in interpreting the regression outputs for time series.
Consider the deterministic trend model $Y_t = \mu_t + X_t$. Let $\hat{\mu}_t$ denote the estimated trend. We estimate $X_t$ by $Y_t - \hat{\mu}_t$, and the standard deviation $\sqrt{\gamma_0}$ of $X_t$ by the residual standard deviation

$$s = \sqrt{\frac{1}{n-p} \sum_{t=1}^{n} (Y_t - \hat{\mu}_t)^2}$$

where $p$ is the number of parameters estimated in $\mu_t$ and $n - p$ is the degree of freedom for $s$. The value of $s$ (dependent on the unit) gives an absolute measure of the goodness of fit of the estimated trend.

A unitless measure of the goodness of fit is the $R^2$. (the coefficient of determination, or multiple R-squared).

- $R^2$ is the square of the sample correlation coefficient between the observed series and the estimated trend.
- $R^2$ is also the fraction of the variation in the series that is explained by the estimated trend.
According to Exhibit 3.7, about 81% of the variation in the random walk series is explained by the linear time trend. The adjusted R-squared value is a small adjustment to $R^2$ that yields an approximately unbiased estimate based on the number of parameters estimated in the trend. It is useful for comparing models with different numbers of parameters. Various formulas for computing $R^2$ may be found in any book on regression, such as Draper and Smith (1981). The standard deviations of the coefficients labeled Std. Error on the output need to be interpreted carefully. They are appropriate only when the stochastic component is white noise—the usual regression assumption.
3.6 Residual Analysis

In the model

\[ Y_t = \mu_t + X_t, \]

the unobserved stochastic component \( \{X_t\} \) can be predicted by the residual

\[ \hat{X}_t = Y_t - \hat{\mu}_t. \]

We may investigate the standardized residual \( \hat{X}_t/s \). (most statistics software will produce standardized residuals using a more complicated standard error in the denominator that takes into account the specific regression model being fit.)
The expression `rstudent(model3)` returns the (externally) studentized residuals from the fitted model. [whereas `rstandard(model3)` gives the (internally) standardized residuals.]
There are no apparent patterns relating to different months of the year.

- Exhibit 3.9 Residuals versus Time with Seasonal Plotting Symbols (chap3.R)
We look at the standardized residuals versus the corresponding trend estimate (fitted value). No dramatic patterns are found.

- Exhibit 3.10 Standardized Residuals versus Fitted Values for the Temperature Seasonal Means Model (chap3.R)
The normality of the standardized residuals may be checked by histogram and QQ plot.

Gross nonnormality can be assessed by plotting a histogram of the residuals or standardized residuals.
Normality can be checked more carefully by plotting the so-called normal scores or **quantile-quantile (QQ) plot**. Such a plot displays the quantiles of the data versus the theoretical quantiles of a normal distribution. The straight-line pattern here supports the assumption of a normally distributed stochastic component in this model.

```
qqnorm(rstudent(model3))
```

A reference straight line can be superimposed on the Q-Q normal score plot by

```
qqline(rstudent(model3))
```
Exhibit 3.12  Q-Q Plot: Standardized Residuals of Seasonal Means Model

```
> win.graph(width=2.5,height=2.5,pointsize=8)
> qqnorm(rstudent(model3))
```
An excellent test of normality is known as the Shapiro-Wilk test. It essentially calculates the correlation between the residuals and the corresponding normal quantiles.

\[ \text{shapiro.test}(\text{rstudent}(\text{model3})) \]

The **runs test** tests the independence of a sequence of random variables by checking whether there are too many or too few runs above (or below) the median.

\[ \text{runs}(\text{rstudent}(\text{model3})) \]
The Sample Autocorrelation Function

Another very important diagnostic tool for examining dependence is the sample autocorrelation function.

**Def.** If \( \{ Y_t \} \) is a sequence of data, we define the *sample autocorrelation function* \( r_k \) at lag \( k \) as

\[
r_k = \frac{\sum_{t=k+1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2}, \quad \text{for} \quad k = 1, 2, \ldots
\]

(7)

A plot of \( r_k \) versus lag \( k \) is often called a correlogram.

When \( \{ Y_t \} \) is a stationary process, \( r_k \) gives an estimate of \( \rho_k \).
In our present context, we are interested in discovering possible dependence in the stochastic component; therefore the sample autocorrelation function for the standardized residuals is of interest. Exhibit 3.13 displays the sample autocorrelation for the standardized residuals from the seasonal means model of the temperature series. All values are within the horizontal dashed lines, which are placed at zero plus and minus two approximate standard errors of the sample autocorrelations, namely $\pm 2/\sqrt{n}$. The values of $r_k$ are, of course, estimates of $\rho_k$. As such, they have their own sampling distributions, standard errors, and other properties. For now we shall use $r_k$ as a descriptive tool and defer discussion of those topics until Chapters 6 and 8. According to Exhibit 3.13, for $k = 1, 2, ..., 21$, none of the hypotheses $\rho_k = 0$ can be rejected at the usual significance levels, and it is reasonable to infer that the stochastic component of the series is white noise.

```r
win.graph(width=4.875, height=3, pointsize=8)
acf(rstudent(model3))
```

Exhibit 3.13 Sample Autocorrelation of Residuals of Seasonal Means Model (chap3.R)
We now consider the standardized residuals from fitting a straight line to the rwalk dataset.

Exhibit 3.14 Residuals from Straight Line Fit of the Random Walk

```
> plot(y=rstudent(model1), x=as.vector(time(rwalk)),
     ylab='Standardized Residuals', xlab='Time', type='o')
```

In this plot, the residuals “hang together” too much for white noise—the plot is too smooth. Furthermore, there seems to be more variation in the last third of the series than in the first two-thirds. Exhibit 3.15 shows a similar effect with larger residuals associated with larger fitted values.
Exhibit 3.15 Residuals versus Fitted Values from Straight Line Fit

```r
> win.graph(width=4.875, height=3, pointsize=8)
> plot(y=rstudent(model1), x=fitted(model1),
      ylab='Standardized Residuals', xlab='Fitted Trend Line Values',
      type='p')
```

- Exhibit 3.15 Residuals versus Fitted Values from Straight Line Fit (chap3.R)
The sample autocorrelation function of the standardized residuals, shown in Exhibit 3.16, confirms the smoothness of the time series plot that we observed in Exhibit 3.14. The lag 1 and lag 2 autocorrelations exceed two standard errors above zero and the lag 5 and lag 6 autocorrelations more than two standard errors below zero. This is not what we expect from a white noise process.

Exhibit 3.16  Sample Autocorrelation of Residuals from Straight Line Model

> acf(rstudent(model1))
Let us turn to the annual rainfall in Los Angeles (Exhibit 1.1). We found no evidence of dependence in that series, but we now look for evidence against normality. Exhibit 3.17 displays the normal quantile-quantile plot for that series. We see considerable curvature in the plot.

Exhibit 3.17  Quantile-Quantile Plot of Los Angeles Annual Rainfall Series

```
> win.graph(width=2.5, height=2.5, pointsize=8)
> qqnorm(larain); qqline(larain)
```

Exhibit 3.17 Quantile-Quantile Plot of Los Angeles Annual Rainfall Series (chap3.R)

(a) Display and interpret the time series plot for these data.

(b) Now construct a time series plot that uses separate plotting symbols for the various months. Does your interpretation change from that in part (a)?

(c) Use least squares to fit a seasonal-means trend to this time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

(d) Construct and interpret the time series plot of the standardized residuals from part (c). Be sure to use proper plotting symbols to check on seasonality in the standardized residuals.

(e) Use least squares to fit a seasonal-means plus quadratic time trend to the beer sales time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

See TS-ch3.R
(f) Construct and interpret the time series plot of the standardized residuals from part (e). Again use proper plotting symbols to check for any remaining seasonality in the residuals.

(g) Perform a runs test on the standardized residuals from part (e) and interpret the results.

(h) Calculate and interpret the sample autocorrelations for the standardized residuals.

(i) Investigate the normality of the standardized residuals (error terms). Consider histograms and normal probability plots. Interpret the plots.

See TS-ch3.R