CHAPTER 3

TRENDS

In a general time series, the mean function is a totally arbitrary function of time. In a stationary time series, the mean function must be constant in time. Frequently we need to take the middle ground and consider mean functions that are relatively simple (but not constant) functions of time. These trends are considered in this chapter.

3.1 Deterministic Versus Stochastic Trends

"Trends" can be quite elusive. The same time series may be viewed quite differently by different analysts. The simulated random walk shown in Exhibit 2.1 might be considered to display a general upward trend. However, we know that the random walk process has zero mean for all time. The perceived trend is just an artifact of the strong positive correlation between the series values at nearby time points and the increasing variance in the process as time goes by. A second and third simulation of exactly the same process might well show completely different "trends." We ask you to produce some additional simulations in the exercises. Some authors have described such trends as **stochastic trends** (see Box, Jenkins, and Reinsel, 1994), although there is no generally accepted definition of a stochastic trend.

The average monthly temperature series plotted in Exhibit 1.7 on page 6, shows a cyclical or seasonal trend, but here the reason for the trend is clear—the Northern Hemisphere's changing inclination toward the sun. In this case, a possible model might be $Y_t = \mu_t + X_t$, where μ_t is a deterministic function that is periodic with period 12; that is μ_t, should satisfy

$$\mu_t = \mu_{t-12} \qquad \text{for all } t$$

We might assume that X_t, the unobserved variation around μ_t, has zero mean for all t so that indeed μ_t is the mean function for the observed series Y_t. We could describe this model as having a **deterministic trend** as opposed to the stochastic trend considered earlier. In other situations we might hypothesize a deterministic trend that is linear in time (that is, $\mu_t = \beta_0 + \beta_1 t$) or perhaps a quadratic time trend, $\mu_t = \beta_0 + \beta_1 t + \beta_2 t^2$. Note that an implication of the model $Y_t = \mu_t + X_t$ with $E(X_t) = 0$ for all t is that the deterministic trend μ_t applies for all time. Thus, if $\mu_t = \beta_0 + \beta_1 t$, we are assuming that the *same* linear time trend applies forever. We should therefore have good reasons for assuming such a model—not just because the series looks somewhat linear over the time period observed.

In this chapter, we consider methods for modeling deterministic trends. Stochastic trends will be discussed in Chapter 5, and stochastic seasonal models will be discussed in Chapter 10. Many authors use the word trend only for a slowly changing mean function, such as a linear time trend, and use the term seasonal component for a mean function that varies cyclically. We do not find it useful to make such distinctions here.

3.2 Estimation of a Constant Mean

We first consider the simple situation where a constant mean function is assumed. Our model may then be written as

$$Y_t = \mu + X_t \tag{3.2.1}$$

where $E(X_t) = 0$ for all t. We wish to estimate μ with our observed time series $Y_1, Y_2, \ldots, Y_n$. The most common estimate of μ is the sample mean or average defined as

$$\bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t \tag{3.2.2}$$

Under the minimal assumptions of Equation (3.2.1), we see that $E(\bar{Y}) = \mu$; therefore $\bar{Y}$ is an unbiased estimate of μ. To investigate the precision of $\bar{Y}$ as an estimate of μ, we need to make further assumptions concerning X_t.

Suppose that $\{Y_t\}$, (or, equivalently, $\{X_t\}$ of Equation (3.2.1)) is a stationary time series with autocorrelation function ρ_k. Then, by Exercise 2.17, we have

$$\begin{aligned} Var(\bar{Y}) &= \frac{\gamma_0}{n}\left[\sum_{k=-n+1}^{n-1}\left(1-\frac{|k|}{n}\right)\rho_k\right] \\ &= \frac{\gamma_0}{n}\left[1+2\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\rho_k\right] \end{aligned} \tag{3.2.3}$$

Notice that the first factor, γ_0/n, is the process (population) variance divided by the sample size—a concept with which we are familiar in simpler random sampling contexts. If the series $\{X_t\}$ of Equation (3.2.1) is just white noise, then $\rho_k = 0$ for $k > 0$ and $Var(\bar{Y})$ reduces to simply γ_0/n.

In the (stationary) moving average model $Y_t = e_t - \tfrac{1}{2}e_{t-1}$, we find that $\rho_1 = -0.4$ and $\rho_k = 0$ for $k > 1$. In this case, we have

$$\begin{aligned} Var(\bar{Y}) &= \frac{\gamma_0}{n}\left[1+2\left(1-\frac{1}{n}\right)(-0.4)\right] \\ &= \frac{\gamma_0}{n}\left[1-0.8\left(\frac{n-1}{n}\right)\right] \end{aligned}$$

For values of n usually occurring in time series ($n > 50$, say), the factor $(n-1)/n$ will be close to 1, so that we have

$$Var(\bar{Y}) \approx 0.2\frac{\gamma_0}{n}$$

We see that the negative correlation at lag 1 has improved the estimation of the mean compared with the estimation obtained in the white noise (random sample) situation. Because the series tends to oscillate back and forth across the mean, the sample mean obtained is more precise.

On the other hand, if $\rho_k \geq 0$ for all $k \geq 1$, we see from Equation (3.2.3) that $Var(\bar{Y})$ will be larger than γ_0/n. Here the positive correlations make estimation of the mean *more* difficult than in the white noise case. In general, some correlations will be positive and some negative, and Equation (3.2.3) must be used to assess the total effect.

For many stationary processes, the autocorrelation function decays quickly enough with increasing lags that

$$\sum_{k=0}^{\infty} |\rho_k| < \infty \tag{3.2.4}$$

(The random cosine wave of Chapter 2 is an exception.)

Under assumption (3.2.4) and given a large sample size n, the following useful approximation follows from Equation (3.2.3) (See Anderson, 1971, p. 459, for example)

$$Var(\bar{Y}) \approx \frac{\gamma_0}{n}\left[\sum_{k=-\infty}^{\infty} \rho_k\right] \qquad \text{for large } n \tag{3.2.5}$$

Notice that to this approximation the variance is inversely proportional to the sample size n.

As an example, suppose that $\rho_k = \phi^{|k|}$ for all k, where ϕ is a number strictly between -1 and $+1$. Summing a geometric series yields

$$Var(\bar{Y}) \approx \frac{(1+\phi)}{(1-\phi)}\frac{\gamma_0}{n} \tag{3.2.6}$$

For a nonstationary process (but with a constant mean), the precision of the sample mean as an estimate of μ can be strikingly different. As a useful example, suppose that in Equation (3.2.1) $\{X_t\}$ is a random walk process as described in Chapter 2. Then directly from Equation (2.2.8) we have

$$\begin{aligned} Var(\bar{Y}) &= \frac{1}{n^2}Var\left[\sum_{i=1}^{n} Y_i\right] \\ &= \frac{1}{n^2}Var\left[\sum_{i=1}^{n}\sum_{j=1}^{i} e_j\right] \end{aligned}$$

$$= \frac{1}{n^2} Var(e_1 + 2e_2 + 3e_3 + \cdots + ne_n)$$

$$= \frac{\sigma_e^2}{n^2} \sum_{k=1}^{n} k^2$$

so that

$$Var(\bar{Y}) = \sigma_e^2 (2n+1) \frac{(n+1)}{6n} \tag{3.2.7}$$

Notice that in this special case the variance of our estimate of the mean actually *increases* as the sample size n increases. Clearly this is unacceptable, and we need to consider other estimation techniques for nonstationary series.

3.3 Regression Methods

The classical statistical method of regression analysis may be readily used to estimate the parameters of common nonconstant mean trend models. We shall consider the most useful ones: linear, quadratic, seasonal means, and cosine trends.

Linear and Quadratic Trends in Time

Consider the deterministic time trend expressed as

$$\mu_t = \beta_0 + \beta_1 t \tag{3.3.1}$$

where the *slope* and *intercept*, β_1 and β_0 respectively, are unknown parameters. The classical least squares (or regression) method is to choose as estimates of β_1 and β_0 values that minimize

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

The solution may be obtained in several ways, for example, by computing the partial derivatives with respect to both β's, setting the results equal to zero, and solving the resulting linear equations for the β's. Denoting the solutions by $\hat{\beta}_0$ and $\hat{\beta}_1$, we find that

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{n} (Y_t - \bar{Y})(t - \bar{t})}{\sum_{t=1}^{n} (t - \bar{t})^2} \tag{3.3.2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{t}$$

where $\bar{t} = (n+1)/2$ is the average of 1, 2,…, n. These formulas can be simplified somewhat, and various versions of the formulas are well-known. However, we assume that the computations will be done by statistical software and we will not pursue other expressions for $\hat{\beta}_0$ and $\hat{\beta}_1$ here.

Example

Consider the random walk process that was shown in Exhibit 2.1. Suppose we (mistakenly) treat this as a linear time trend and estimate the slope and intercept by least-squares regression. Using statistical software we obtain Exhibit 3.1.

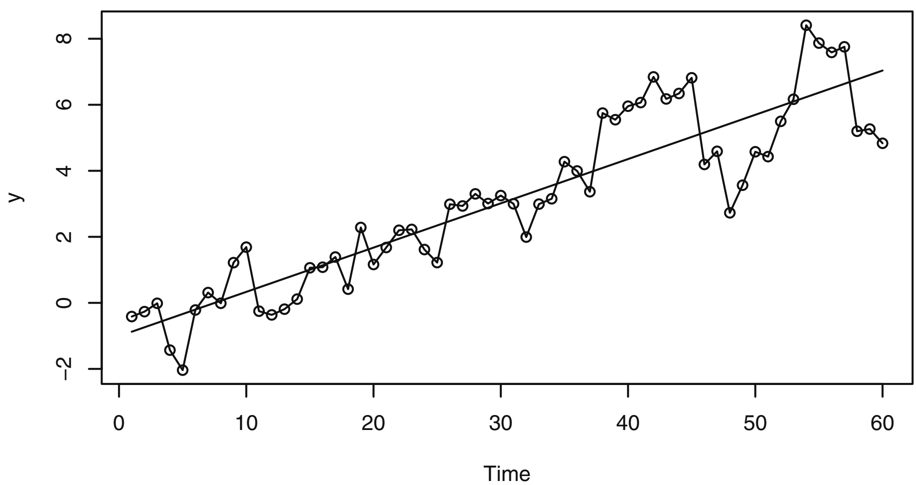
Exhibit 3.1 Least Squares Regression Estimates for Linear Time Trend

	Estimate	Std. Error	*t* value	*Pr*(>|*t*|)
Intercept	−1.008	0.2972	−3.39	0.00126
Time	0.1341	0.00848	15.82	< 0.0001

```
> data(rwalk)
> model1=lm(rwalk~time(rwalk))
> summary(model1)
```

So here the estimated slope and intercept are $\hat{\beta}_1 = 0.1341$ and $\hat{\beta}_0 = -1.008$, respectively. Exhibit 3.2 displays the random walk with the least squares regression trend line superimposed. We will interpret more of the regression output later in Section 3.5 on page 40 and see that fitting a line to these data is not appropriate.

Exhibit 3.2 Random Walk with Linear Time Trend




```
> win.graph(width=4.875, height=2.5,pointsize=8)
> plot(rwalk,type='o',ylab='y')
> abline(model1) # add the fitted least squares line from model1
```

Cyclical or Seasonal Trends

Consider now modeling and estimating seasonal trends, such as for the average monthly temperature data in Exhibit 1.7. Here we assume that the observed series can be represented as

$$Y_t = \mu_t + X_t$$

where $E(X_t) = 0$ for all t.

The most general assumption for μ_t with monthly seasonal data is that there are 12 constants (parameters), β_1, β_2,…, and β_{12}, giving the expected average temperature for each of the 12 months. We may write

$$\mu_t = \begin{cases} \beta_1 & \text{for } t = 1, 13, 25, \ldots \\ \beta_2 & \text{for } t = 2, 14, 26, \ldots \\ \vdots & \\ \beta_{12} & \text{for } t = 12, 24, 36, \ldots \end{cases} \tag{3.3.3}$$

This is sometimes called a **seasonal means** model.

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 β'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 β's are zero—not an interesting hypothesis in this case.

Exhibit 3.3 Regression Results for the Seasonal Means Model

	Estimate	Std. Error	t-value	$Pr(>\|t\|)$
January	16.608	0.987	16.8	< 0.0001
February	20.650	0.987	20.9	< 0.0001
March	32.475	0.987	32.9	< 0.0001
April	46.525	0.987	47.1	< 0.0001
May	58.092	0.987	58.9	< 0.0001
June	67.500	0.987	68.4	< 0.0001
July	71.717	0.987	72.7	< 0.0001

	Estimate	Std. Error	*t*-value	*Pr*(>|*t*|)
August	69.333	0.987	70.2	< 0.0001
September	61.025	0.987	61.8	< 0.0001
October	50.975	0.987	51.6	< 0.0001
November	36.650	0.987	37.1	< 0.0001
December	23.642	0.987	24.0	< 0.0001

```
> data(tempdub)
> month.=season(tempdub) # period added to improve table display
> model2=lm(tempdub~month.-1) # -1 removes the intercept term
> summary(model2)
```

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.

Exhibit 3.4 Results for Seasonal Means Model with an Intercept

	Estimate	Std. Error	*t*-value	*Pr*(>|*t*|)
Intercept	16.608	0.987	16.83	< 0.0001
February	4.042	1.396	2.90	0.00443
March	15.867	1.396	11.37	< 0.0001
April	29.917	1.396	21.43	< 0.0001
May	41.483	1.396	29.72	< 0.0001
June	50.892	1.396	36.46	< 0.0001
July	55.108	1.396	39.48	< 0.0001
August	52.725	1.396	37.78	< 0.0001
September	44.417	1.396	31.82	< 0.0001
October	34.367	1.396	24.62	< 0.0001
November	20.042	1.396	14.36	< 0.0001
December	7.033	1.396	5.04	< 0.0001

```
> model3=lm(tempdub~month.) # January is dropped automatically
> summary(model3)
```

Cosine Trends

The seasonal means model for monthly data consists of 12 independent parameters and does not take the shape of the seasonal trend into account at all. For example, the fact that the March and April means are quite similar (and different from the June and July means) is not reflected in the model. 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.

Consider the cosine curve with equation

$$\mu_t = \beta\cos(2\pi ft + \Phi) \tag{3.3.4}$$

We call β (> 0) the *amplitude*, f the *frequency*, and Φ the *phase* of the curve. As t varies, the curve oscillates between a maximum of β and a minimum of $-\beta$. Since the curve repeats itself exactly every $1/f$ time units, $1/f$ is called the period of the cosine wave. As noted in Chapter 2, Φ serves to set the arbitrary origin on the time axis. For monthly data with time indexed as 1, 2,…, the most important frequency is $f = 1/12$, because such a cosine wave will repeat itself every 12 months. We say that the *period* is 12.

Equation (3.3.4) is inconvenient for estimation because the parameters β and Φ do not enter the expression linearly. Fortunately, a trigonometric identity is available that reparameterizes (3.3.4) more conveniently, namely

$$\beta\cos(2\pi ft + \Phi) = \beta_1\cos(2\pi ft) + \beta_2\sin(2\pi ft) \tag{3.3.5}$$

where

$$\beta = \sqrt{\beta_1^2 + \beta_2^2}\,, \qquad \Phi = \operatorname{atan}(-\beta_2/\beta_1) \tag{3.3.6}$$

and, conversely,

$$\beta_1 = \beta\cos(\Phi), \qquad \beta_2 = \beta\sin(\Phi) \tag{3.3.7}$$

To estimate the parameters β_1 and β_2 with regression techniques, we simply use $\cos(2\pi ft)$ and $\sin(2\pi ft)$ as regressors or predictor variables.

The simplest such model for the trend would be expressed as

$$\mu_t = \beta_0 + \beta_1\cos(2\pi ft) + \beta_2\sin(2\pi ft) \tag{3.3.8}$$

Here the constant term, β_0, can be meaningfully thought of as a cosine with frequency zero.

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.

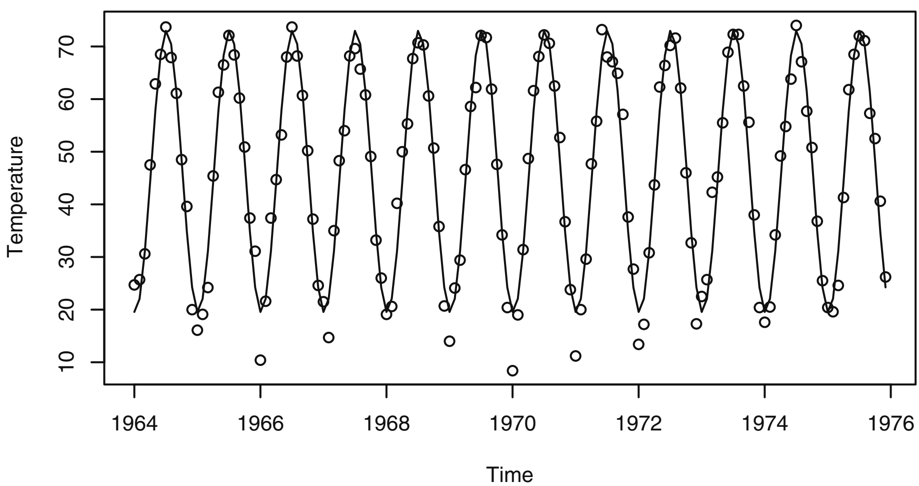
Exhibit 3.5 Cosine Trend Model for Temperature Series

Coefficient	Estimate	Std. Error	*t*-value	*Pr*(>|*t*|)
Intercept	46.2660	0.3088	149.82	< 0.0001
$\cos(2\pi t)$	–26.7079	0.4367	–61.15	< 0.0001
$\sin(2\pi t)$	–2.1697	0.4367	–4.97	<0.0001

```
> har.=harmonic(tempdub,1)
> model4=lm(tempdub~har.)
> summary(model4)
```

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




```
> win.graph(width=4.875, height=2.5,pointsize=8)
> plot(ts(fitted(model4),freq=12,start=c(1964,1)),
   ylab='Temperature',type='l',
> ylim=range(c(fitted(model4),tempdub))); points(tempdub)
> # ylim ensures that the y axis range fits the raw data and the
   fitted values
```

Additional cosine functions at other frequencies will frequently be used to model cyclical trends. For monthly series, the higher harmonic frequencies, such as 2/12 and 3/12, are especially pertinent and will sometimes improve the fit at the expense of adding more parameters to the model. In fact, it may be shown that any periodic trend with period 12 may be expressed exactly by the sum of six pairs of cosine-sine functions. These ideas are discussed in detail in Fourier analysis or spectral analysis. We pursue these ideas further in Chapters 13 and 14.

3.4 Reliability and Efficiency of Regression Estimates

We assume that the series is represented as $Y_t = \mu_t + X_t$, where μ_t is a deterministic trend of the kind considered above and $\{X_t\}$ is a zero-mean stationary process with autocovariance and autocorrelation functions γ_k and ρ_k, respectively. Ordinary regression estimates parameters in a linear model according to the criterion of least squares regardless of whether we are fitting linear time trends, seasonal means, cosine curves, or whatever.

We first consider the easiest case—the seasonal means. As mentioned earlier, the least squares estimates of the seasonal means are just seasonal averages; thus, if we have N (complete) years of monthly data, we can write the estimate for the mean for the jth season as

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

Since $\hat{\beta}_j$ is an average like $\overline{Y}$ but uses only every 12th observation, Equation (3.2.3) can be easily modified to give $Var(\hat{\beta}_j)$. We replace n by N (years) and ρ_k by ρ_{12k} to get

$$Var(\hat{\beta}_j) = \frac{\gamma_0}{N}\left[1 + 2\sum_{k=1}^{N-1}\left(1-\frac{k}{N}\right)\rho_{12k}\right] \qquad \text{for } j = 1, 2, \ldots, 12 \tag{3.4.1}$$

We notice that if $\{X_t\}$ is white noise, then $Var(\hat{\beta}_j)$ reduces to γ_0/N, as expected. Furthermore, if several ρ_k are nonzero but $\rho_{12k} = 0$, then we still have $Var(\hat{\beta}_j) = \gamma_0/N$. In any case, only the seasonal autocorrelations, ρ_{12}, ρ_{24}, ρ_{36},..., enter into Equation (3.4.1). Since N will rarely be very large (except perhaps for quarterly data), approximations like those shown in Equation (3.2.5) will usually not be useful.

We turn now to the cosine trends expressed as in Equation (3.3.8). For any frequency of the form $f = m/n$, where m is an integer satisfying $1 \le m < n/2$, explicit expressions are available for the estimates $\hat{\beta}_1$ and $\hat{\beta}_2$, the amplitudes of the cosine and sine:

$$\hat{\beta}_1 = \frac{2}{n}\sum_{t=1}^{n}\left[\cos\left(\frac{2\pi mt}{n}\right)Y_t\right], \qquad \hat{\beta}_2 = \frac{2}{n}\sum_{t=1}^{n}\left[\sin\left(\frac{2\pi mt}{n}\right)Y_t\right] \tag{3.4.2}$$

(These are effectively the correlations between the time series $\{Y_t\}$ and the cosine and sine waves with frequency m/n.)

Because these are linear functions of $\{Y_t\}$, we may evaluate their variances using Equation (2.2.6). We find

$$Var(\hat{\beta}_1) = \frac{2\gamma_0}{n}\left[1 + \frac{4}{n}\sum_{s=2}^{n}\sum_{t=1}^{s-1}\cos\left(\frac{2\pi mt}{n}\right)\cos\left(\frac{2\pi ms}{n}\right)\rho_{s-t}\right] \tag{3.4.3}$$

where we have used the fact that $\sum_{t=1}^{n}[\cos(2\pi mt/n)]^2 = n/2$. However, the double sum in Equation (3.4.3) does not, in general, reduce further. A similar expression holds for $Var(\hat{\beta}_2)$ if we replace the cosines by sines.

If $\{X_t\}$ is white noise, we get just $2\gamma_0/n$. If $\rho_1 \neq 0$, $\rho_k = 0$ for $k > 1$, and $m/n = 1/12$, then the variance reduces to

$$Var(\hat{\beta}_1) = \frac{2\gamma_0}{n}\left[1 + \frac{4\rho_1}{n}\sum_{t=1}^{n-1}\cos\left(\frac{\pi t}{6}\right)\cos\left(\frac{\pi t + 1}{6}\right)\right] \tag{3.4.4}$$

To illustrate the effect of the cosine terms, we have calculated some representative values:

n	$Var(\hat{\beta}_1)$	
25	$\left(\frac{2\gamma_0}{n}\right)(1 + 1.71\rho_1)$	
50	$\left(\frac{2\gamma_0}{n}\right)(1 + 1.75\rho_1)$	
500	$\left(\frac{2\gamma_0}{n}\right)(1 + 1.73\rho_1)$	
∞	$\left(\frac{2\gamma_0}{n}\right)\left(1 + 2\rho_1\cos\left(\frac{\pi}{6}\right)\right) = \left(\frac{2\gamma_0}{n}\right)(1 + 1.732\rho_1)$	(3.4.5)

If $\rho_1 = -0.4$, then the large sample multiplier in Equation (3.4.5) is $1+1.732(-0.4) = 0.307$ and the variance is reduced by about 70% when compared with the white noise case.

In some circumstances, seasonal means and cosine trends could be considered as competing models for a cyclical trend. If the simple cosine model is an adequate model, how much do we lose if we use the less parsimonious seasonal means model? To approach this problem, we must first consider how to compare the models. The parameters themselves are not directly comparable, but we can compare the estimates of the trend at comparable time points.

Consider the two estimates for the trend in January; that is, μ_1. With seasonal means, this estimate is just the January average, which has variance given by Equation (3.4.1). With the cosine trend model, the corresponding estimate is

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

To compute the variance of this estimate, we need one more fact: With this model, the estimates $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\beta}_2$ are uncorrelated.† This follows from the orthogonality relationships of the cosines and sines involved. See Bloomfield (1976) or Fuller (1996) for more details. For the cosine model, then, we have

$$Var(\hat{\mu}_1) = Var(\hat{\beta}_0) + Var(\hat{\beta}_1)\left[\cos\left(\frac{2\pi}{12}\right)\right]^2 + Var(\hat{\beta}_2)\left[\sin\left(\frac{2\pi}{12}\right)\right]^2 \tag{3.4.6}$$

For our first comparison, assume that the stochastic component is white noise. Then the variance of our estimate in the seasonal means model is just γ_0/N. For the cosine model, we use Equation (3.4.6), and Equation (3.4.4) and its sine equivalent, to obtain

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

$$= 3\frac{\gamma_0}{n}$$

since $(\cos\theta)^2 + (\sin\theta)^2 = 1$. Thus the ratio of the standard deviation in the cosine model to that in the seasonal means model is

$$\sqrt{\frac{3\gamma_0/n}{\gamma_0/N}} = \sqrt{\frac{3N}{n}}$$

In particular, for the monthly temperature series, we have $n = 144$ and $N = 12$; thus, the ratio is

$$\sqrt{\frac{3(12)}{144}} = 0.5$$

Thus, in the cosine model, we estimate the January effect with a standard deviation that is only half as large as it would be if we estimated with a seasonal means model—a substantial gain. (Of course, this assumes that the cosine trend plus white noise model is the correct model.)

Suppose now that the stochastic component is such that $\rho_1 \neq 0$ but $\rho_k = 0$ for $k > 1$. With a seasonal means model, the variance of the estimated January effect will be unchanged (see Equation (3.4.1) on page 36). For the cosine trend model, if we have a reasonably large sample size, we may use Equation (3.4.5), an identical expression for $Var(\hat{\beta}_2)$, and Equation (3.2.3) on page 28 for $Var(\hat{\beta}_0)$ to obtain

† This assumes that 1/12 is a "Fourier frequency"; that is, it is of the form m/n. Otherwise, these estimates are only approximately uncorrelated.

$$Var(\hat{\mu}_1) = \frac{\gamma_0}{n}\left\{1 + 2\rho_1 + 2\left[1 + 2\rho_1 \cos\left(\frac{2\pi}{12}\right)\right]\right\}$$

$$= \frac{\gamma_0}{n}\left\{3 + 2\rho_1\left[1 + 2\cos\left(\frac{\pi}{6}\right)\right]\right\}$$

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

$$\sqrt{\left[\frac{(0.814\gamma_0)/n}{\gamma_0/N}\right]} = \sqrt{\frac{0.814N}{n}}$$

If we take $n = 144$ and $N = 12$, the ratio is

$$\sqrt{\frac{0.814(12)}{144}} = 0.26$$

a very substantial reduction indeed!

We now turn to linear time trends. For these trends, an alternative formula to Equation (3.3.2) on page 30 for $\hat{\beta}_1$ is more convenient. It can be shown that the least squares estimate of the slope may be written

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{n}(t-\bar{t})Y_t}{\sum_{t=1}^{n}(t-\bar{t})^2} \tag{3.4.7}$$

Since the estimate is a linear combination of Y-values, some progress can be made in evaluating its variance. We have

$$Var(\hat{\beta}_1) = \frac{12\gamma_0}{n(n^2-1)}\left[1 + \frac{24}{n(n^2-1)}\sum_{s=2}^{n}\sum_{t=1}^{s-1}(t-\bar{t})(s-\bar{t})\rho_{s-t}\right] \tag{3.4.8}$$

where we have used $\sum_{t=1}^{n}(t-\bar{t})^2 = n(n^2-1)/12$. Again the double sum does not in general reduce.

To illustrate the effect of Equation (3.4.8), consider again the case where $\rho_1 \neq 0$ but $\rho_k = 0$ for $k > 1$. Then, after some algebraic manipulation, again involving the sum of consecutive integers and their squares, Equation (3.4.8) can be reduced to

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

For large n, we can neglect the $3/n$ term and use

$$Var(\hat{\beta}_1) = \frac{12\gamma_0(1+2\rho_1)}{n(n^2-1)} \tag{3.4.9}$$

If $\rho_1 = -0.4$, then $1 + 2\rho_1 = 0.2$, and then the variance of $\hat{\beta}_1$ is only 20% of what it would be if $\{X_t\}$ were white noise. Of course, if $\rho_1 > 0$, then the variance would be larger than for the white noise case.

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. These methods are pursued further in Chapter 11.

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.

Although the simple least squares estimates may be asymptotically efficient, it does not follow that the estimated standard deviations of the coefficients as printed out by all regression routines are correct. We shall elaborate on this point in the next section. We also caution the reader that the result above is restricted to certain kinds of trends and cannot, in general, be extended to regression on arbitrary predictor variables, such as other time series. For example, Fuller (1996, pp. 518–522) shows that if $Y_t = \beta Z_t + X_t$, where $\{X_t\}$ has a simple stochastic structure but $\{Z_t\}$ is also a stationary series, then the least squares estimate of β can be very inefficient and biased even for large samples.

3.5 Interpreting Regression Output

We have already noted that the standard regression routines calculate least squares estimates of the unknown regression coefficients—the betas. As such, the estimates are reasonable under minimal assumptions on the stochastic component $\{X_t\}$. However, some of the properties of the regression output depend heavily on the usual regression assumption that $\{X_t\}$ is white noise, and some depend on the further assumption that $\{X_t\}$ is approximately normally distributed. We begin with the items that depend least on the assumptions.

Consider the regression output shown in Exhibit 3.7. We shall write $\hat{\mu}_t$ for the estimated trend regardless of the assumed parametric form for μ_t. For example, for the linear time trend, we have $\mu_t = \beta_0 + \beta_1 t$. For each t, the unobserved stochastic component

X_t can be estimated (predicted) by $Y_t - \hat{\mu}_t$. If the $\{X_t\}$ process has constant variance, then we can estimate the standard deviation of X_t, namely $\sqrt{\gamma_0}$, by the **residual standard deviation**

$$s = \sqrt{\frac{1}{n-p}\sum_{t=1}^{n}(Y_t - \hat{\mu}_t)^2} \tag{3.5.1}$$

where p is the number of parameters estimated in μ_t and $n - p$ is the so-called *degrees of freedom* for s. The value of s gives an absolute measure of the goodness of fit of the estimated trend—the smaller the value of s, the better the fit. However, a value of s of, say, 60.74 is somewhat difficult to interpret.

A unitless measure of the goodness of fit of the trend is the value of R^2, also called the **coefficient of determination** or multiple R-squared. One interpretation of R^2 is that it is the square of the sample correlation coefficient between the observed series and the estimated trend. It is also the fraction of the variation in the series that is explained by the estimated trend. Exhibit 3.7 is a more complete regression output when fitting the straight line to the random walk data. This extends what we saw in Exhibit 3.1 on page 31.

Exhibit 3.7 Regression Output for Linear Trend Fit of Random Walk

	Estimate	Std. Error	*t*-value	*Pr*(>|*t*|)
Intercept	−1.007888	0.297245	−3.39	0.00126
Time	0.134087	0.008475	15.82	< 0.0001

Residual standard error	1.137	with 58 degrees of freedom
Multiple *R*-Squared	0.812	
Adjusted *R*-squared	0.809	
***F*-statistic**	250.3	with 1 and 58 df; p-value < 0.0001

```
> model1=lm(rwalk~time(rwalk))
> summary(model1)
```

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.

For example, in Exhibit 3.7 the value 1.137 is obtained from the square root of the value given by Equation (3.4.8) when $\rho_k = 0$ for $k > 0$ and with γ_0 estimated by s^2, that is, to within rounding,

$$0.008475 = \sqrt{\frac{12(1.137)^2}{60(60^2 - 1)}}$$

The important point is that these standard deviations assume a white noise stochastic component that will rarely be true for time series.

The t-values or t-ratios shown in Exhibit 3.7 are just the estimated regression coefficients, each divided by their respective standard errors. If the stochastic component is normally distributed white noise, then these ratios provide appropriate test statistics for checking the significance of the regression coefficients. In each case, the null hypothesis is that the corresponding unknown regression coefficient is zero. The significance levels and p-values are determined from the t-distribution with $n - p$ degrees of freedom.

3.6 Residual Analysis

As we have already noted, the unobserved stochastic component $\{X_t\}$ can be estimated, or predicted, by the **residual**

$$\hat{X}_t = Y_t - \hat{\mu}_t \tag{3.6.1}$$

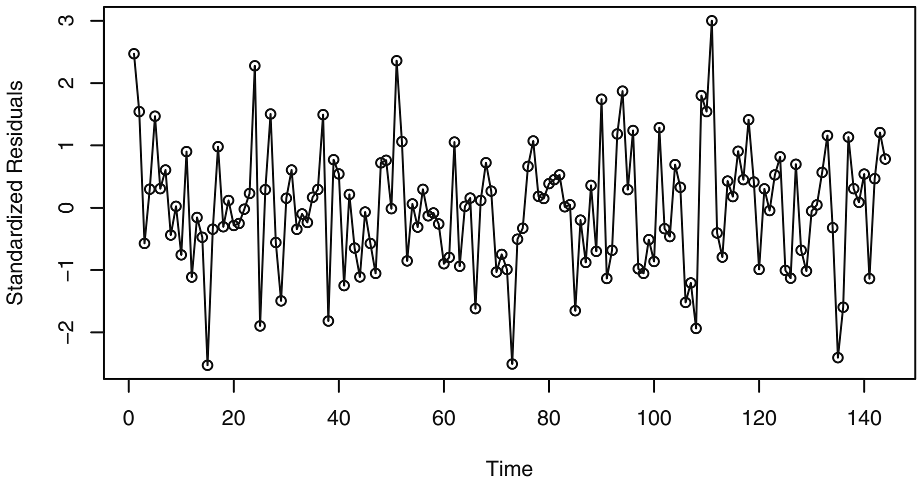
Predicted is really a better term. We reserve the term estimate for the guess of an unknown parameter and the term predictor for an estimate of an unobserved random variable. We call $\hat{X}_t$ the residual corresponding to the tth observation. If the trend model is reasonably correct, then the residuals should behave roughly like the true stochastic component, and various assumptions about the stochastic component can be assessed by looking at the residuals. If the stochastic component is white noise, then the residuals should behave roughly like independent (normal) random variables with zero mean and standard deviation s. Since a least squares fit of any trend containing a constant term automatically produces residuals with a zero mean, we might consider standardizing the residuals as $\hat{X}_t / s$. However, 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.

With the residuals or standardized residuals in hand, the next step is to examine various residual plots. We first look at the plot of the residuals over time. If the data are possibly seasonal, we should use plotting symbols as we did in Exhibit 1.9 on page 7, so that residuals associated with the same season can be identified easily.

We will use the monthly average temperature series which we fitted with seasonal means as our first example to illustrate some of the ideas of residual analysis. Exhibit 1.7 on page 6 shows the time series plot of that series. Exhibit 3.8 shows a time series plot for the standardized residuals of the monthly temperature data fitted by seasonal means. If the stochastic component is white noise and the trend is adequately modeled, we would expect such a plot to suggest a rectangular scatter with no discernible trends whatsoever. There are no striking departures from randomness apparent in this display.

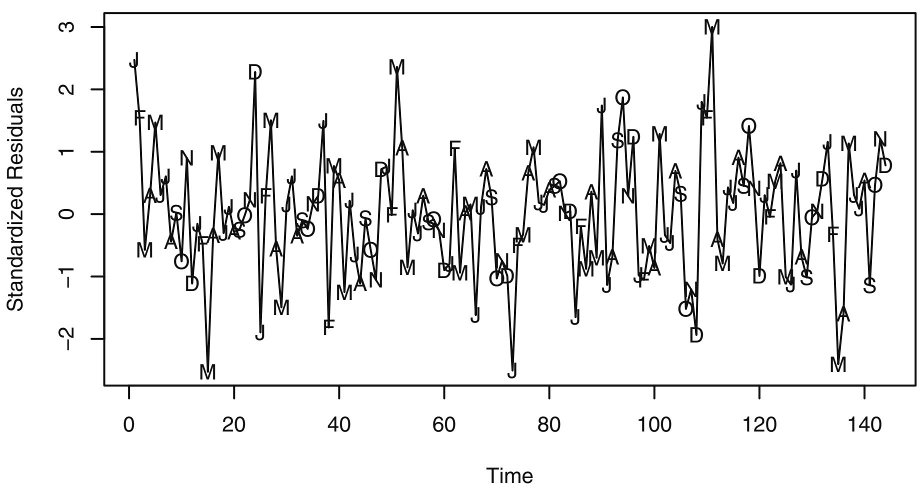
Exhibit 3.9 repeats the time series plot but now with seasonal plotting symbols. Again there are no apparent patterns relating to different months of the year.

Exhibit 3.8 Residuals versus Time for Temperature Seasonal Means




```
> plot(y=rstudent(model3),x=as.vector(time(tempdub)),
   xlab='Time',ylab='Standardized Residuals',type='o')
```

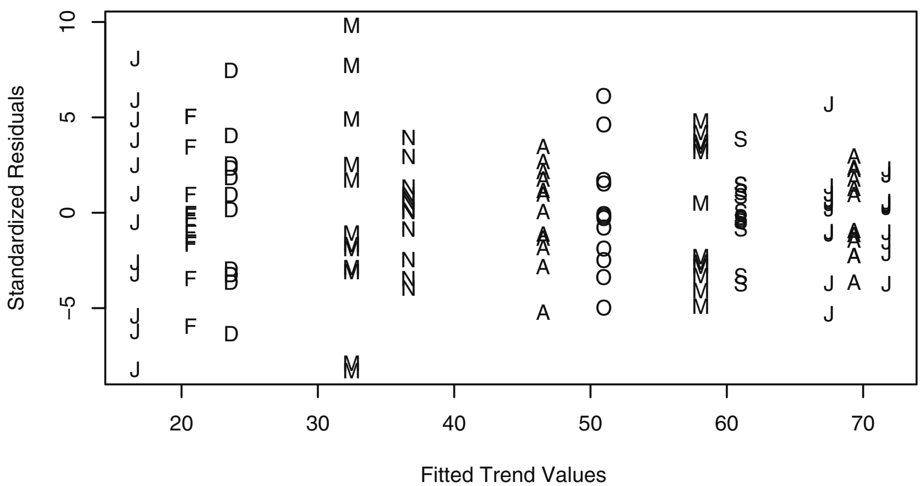
Exhibit 3.9 Residuals versus Time with Seasonal Plotting Symbols




```
> plot(y=rstudent(model3),x=as.vector(time(tempdub)),xlab='Time',
> ylab='Standardized Residuals',type='l')
> points(y=rstudent(model3),x=as.vector(time(tempdub)),
   pch=as.vector(season(tempdub)))
```

Next we look at the standardized residuals versus the corresponding trend estimate, or fitted value, as in Exhibit 3.10. Once more we are looking for patterns. Are small residuals associated with small fitted trend values and large residuals with large fitted trend values? Is there less variation for residuals associated with certain sized fitted trend values or more variation with other fitted trend values? There is somewhat more variation for the March residuals and less for November, but Exhibit 3.10 certainly does not indicate any dramatic patterns that would cause us to doubt the seasonal means model.

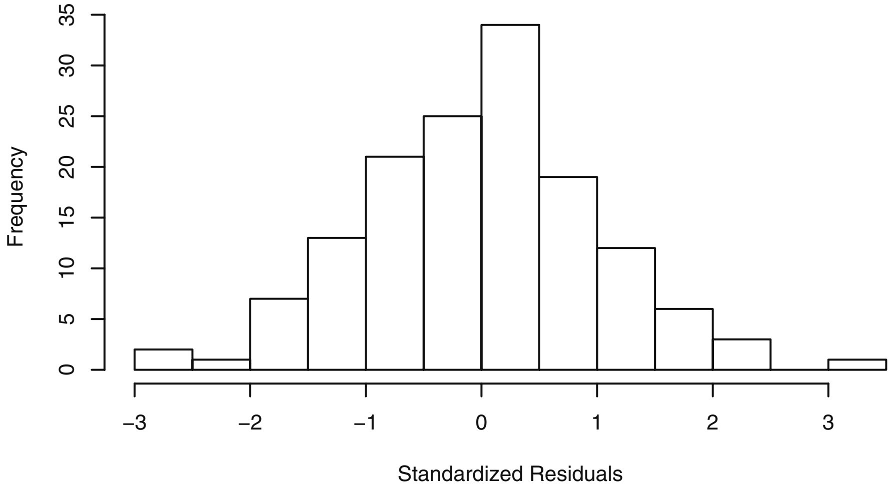
Exhibit 3.10 Standardized Residuals versus Fitted Values for the Temperature Seasonal Means Model




```
> plot(y=rstudent(model3),x=as.vector(fitted(model3)),
   xlab='Fitted Trend Values',
> ylab='Standardized Residuals',type='n')
> points(y=rstudent(model3),x=as.vector(fitted(model3)),
   pch=as.vector(season(tempdub)))
```

Gross nonnormality can be assessed by plotting a histogram of the residuals or standardized residuals. Exhibit 3.11 displays a frequency histogram of the standardized residuals from the seasonal means model for the temperature series. The plot is somewhat symmetric and tails off at both the high and low ends as a normal distribution does.

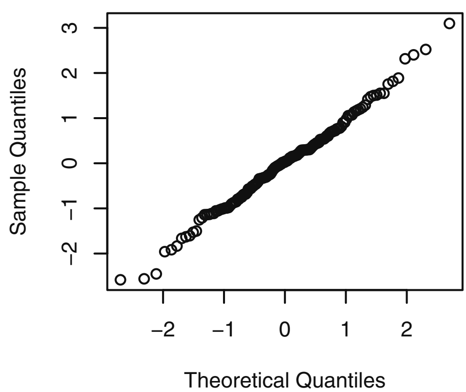
Exhibit 3.11 Histogram of Standardized Residuals from Seasonal Means Model




```
> hist(rstudent(model3),xlab='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. With normally distributed data, the QQ plot looks approximately like a straight line. Exhibit 3.12 shows the QQ normal scores plot for the standardized residuals from the seasonal means model for the temperature series. The straight-line pattern here supports the assumption of a normally distributed stochastic component in this model.

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. The lower this correlation, the more evidence we have against normality. Applying that test to these residuals gives a test statistic of $W = 0.9929$ with a p-value of 0.6954. We cannot reject the null hypothesis that the stochastic component of this model is normally distributed.

Independence in the stochastic component can be tested in several ways. The **runs test** examines the residuals in sequence to look for patterns—patterns that would give evidence against independence. Runs above or below their median are counted. A small number of runs would indicate that neighboring residuals are positively dependent and tend to "hang together" over time. On the other hand, too many runs would indicate that the residuals oscillate back and forth across their median. Then neighboring residuals are negatively dependent. So either too few or too many runs lead us to reject independence. Performing a runs test‡ on these residuals produces the following values: observed runs = 65, expected runs = 72.875, which leads to a p-value of 0.216 and we cannot reject independence of the stochastic component in this seasonal means model.

The Sample Autocorrelation Function

Another very important diagnostic tool for examining dependence is the sample autocorrelation function. Consider any sequence of data $Y_1, Y_2, \ldots, Y_n$—whether residuals, standardized residuals, original data, or some transformation of data. Tentatively assuming stationarity, we would like to estimate the autocorrelation function ρ_k for a variety of lags $k = 1, 2, \ldots$. The obvious way to do this is to compute the sample correlation between the pairs k units apart in time. That is, among (Y_1, Y_{1+k}), (Y_2, Y_{2+k}), $(Y_3, Y_{3+k}), \ldots$, and (Y_{n-k}, Y_n). However, we modify this slightly, taking into account that we are assuming stationarity, which implies a common mean and variance for the series. With this in mind, we define the **sample autocorrelation function**, r_k, at lag k as

$$r_k = \frac{\sum_{t=k+1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2} \qquad \text{for } k = 1, 2, \ldots \tag{3.6.2}$$

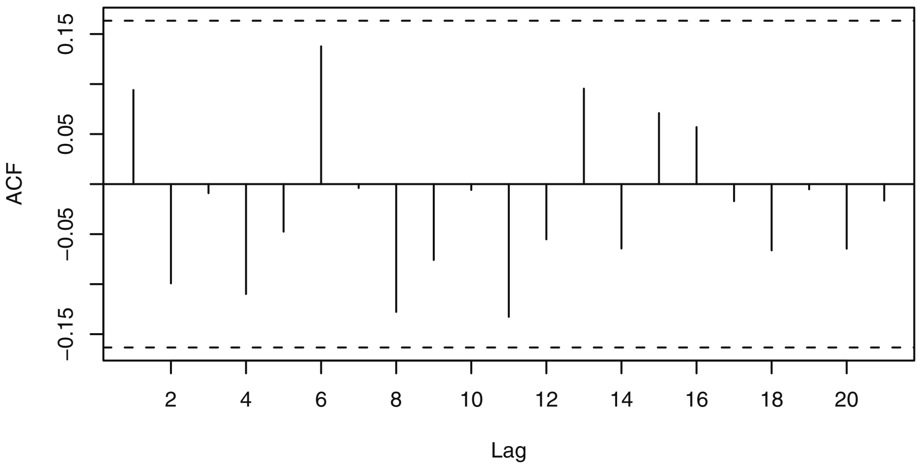
Notice that we used the "grand mean," $\bar{Y}$, in all places and have also divided by the "grand sum of squares" rather than the product of the two separate standard deviations used in the ordinary correlation coefficient. We also note that the denominator is a sum of n squared terms while the numerator contains only $n - k$ cross products. For a variety of reasons, this has become the standard definition for the sample autocorrelation function. A plot of r_k versus lag k is often called a **correlogram**.

† Royston, P. (1982) "An Extension of Shapiro and Wilk's *W* Test for Normality to Large Samples." *Applied Statistics*, **31**, 115–124.

‡ R code: `runs(rstudent(model3))`

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 ρ_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.

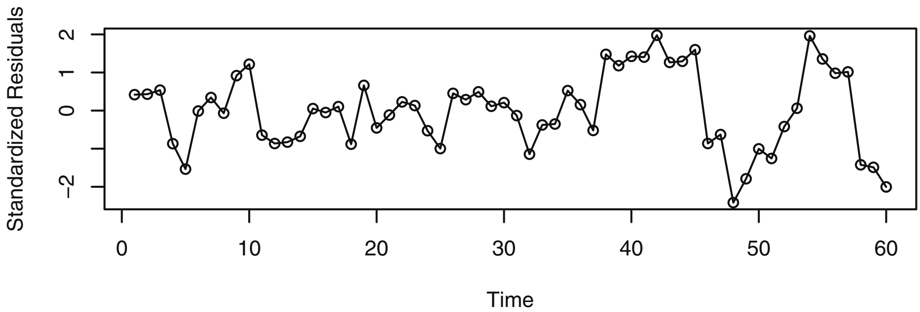
Exhibit 3.13 Sample Autocorrelation of Residuals of Seasonal Means Model




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

As a second example consider the standardized residuals from fitting a straight line to the random walk time series. Recall Exhibit 3.2 on page 31, which shows the data and fitted line. A time series plot of the standardized residuals is shown in Exhibit 3.14.

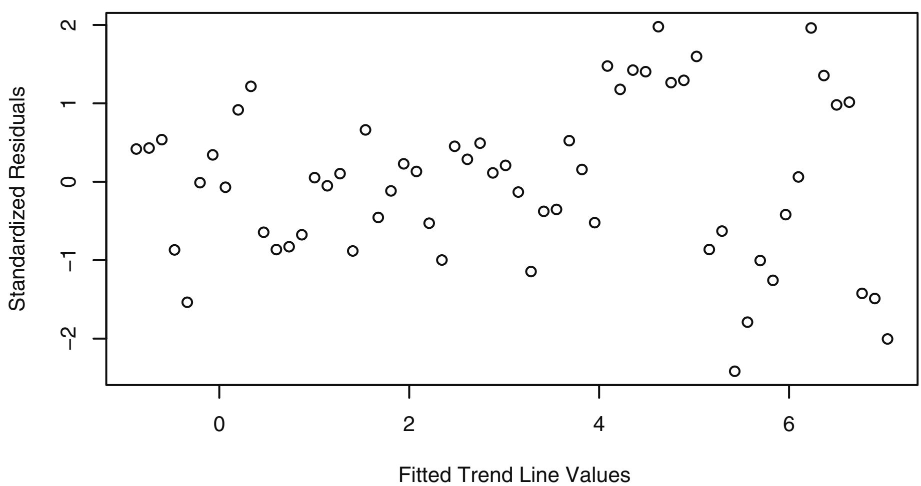
Exhibit 3.14 Residuals from Straight Line Fit of the Random Walk




```
> 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.

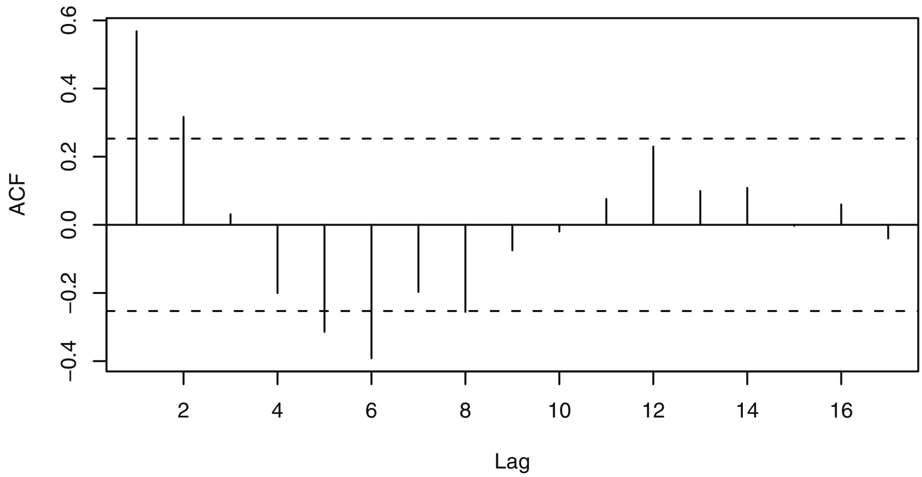
Exhibit 3.15 Residuals versus Fitted Values from Straight Line Fit




```
> 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')
```

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.

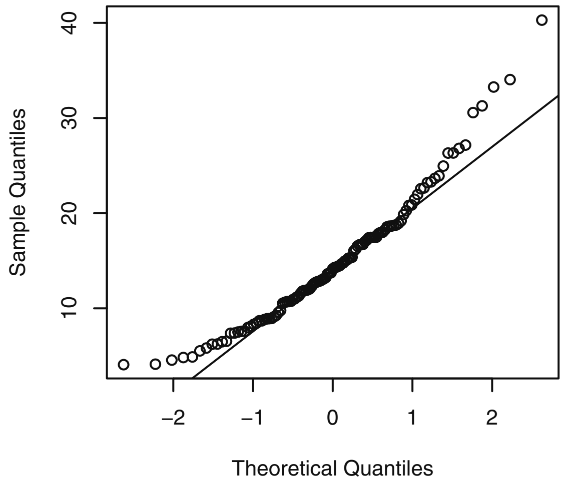
Exhibit 3.16 Sample Autocorrelation of Residuals from Straight Line Model




```
> acf(rstudent(model1))
```

Finally, we return to the annual rainfall in Los Angeles shown in Exhibit 1.1 on page 2. 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. A line passing through the first and third normal quartiles helps point out the departure from a straight line 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)
```

3.7 Summary

This chapter is concerned with describing, modeling, and estimating deterministic trends in time series. The simplest deterministic “trend” is a constant-mean function. Methods of estimating a constant mean were given but, more importantly, assessment of the accuracy of the estimates under various conditions was considered. Regression methods were then pursued to estimate trends that are linear or quadratic in time. Methods for modeling cyclical or seasonal trends came next, and the reliability and efficiency of all of these regression methods were investigated. The final section began our study of residual analysis to investigate the quality of the fitted model. This section also introduced the important sample autocorrelation function, which we will revisit throughout the remainder of the book.

Exercises

3.1 Verify Equation (3.3.2) on page 30, for the least squares estimates of β_0 and of β_1 when the model $Y_t = \beta_0 + \beta_1 t + X_t$ is considered.

3.2 Suppose $Y_t = \mu + e_t - e_{t-1}$. Find $Var(\bar{Y})$. Note any unusual results. In particular, compare your answer to what would have been obtained if $Y_t = \mu + e_t$. (Hint: You may avoid Equation (3.2.3) on page 28 by first doing some algebraic simplification on $\sum_{t=1}^{n}(e_t - e_{t-1})$.)

3.3 Suppose $Y_t = \mu + e_t + e_{t-1}$. Find $Var(\bar{Y})$. Compare your answer to what would have been obtained if $Y_t = \mu + e_t$. Describe the effect that the autocorrelation in $\{Y_t\}$ has on $Var(\bar{Y})$.

3.4 The data file hours contains monthly values of the average hours worked per week in the U.S. manufacturing sector for July 1982 through June 1987.

(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)?

3.5 The data file wages contains monthly values of the average hourly wages (in dollars) for workers in the U.S. apparel and textile products industry for July 1981 through June 1987.

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

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

(c) Construct and interpret the time series plot of the standardized residuals from part (b).

(d) Use least squares to fit a quadratic time trend to the wages time series. Interpret the regression output. Save the standardized residuals from the fit for further analysis.

(e) Construct and interpret the time series plot of the standardized residuals from part (d).

3.6 The data file beersales contains monthly U.S. beer sales (in millions of barrels) for the period January 1975 through December 1990.

(a) Display and interpret the plot 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.

(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.

3.7 The data file winnebago contains monthly unit sales of recreational vehicles from Winnebago, Inc., from November 1966 through February 1972.

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

(b) Use least squares to fit a line to these data. Interpret the regression output. Plot the standardized residuals from the fit as a time series. Interpret the plot.

(c) Now take natural logarithms of the monthly sales figures and display and interpret the time series plot of the transformed values.

(d) Use least squares to fit a line to the logged data. Display and interpret the time series plot of the standardized residuals from this fit.

(e) Now use least squares to fit a seasonal-means plus linear time trend to the logged sales time series and save the standardized residuals for further analysis. Check the statistical significance of each of the regression coefficients in the model.

(f) Display the time series plot of the standardized residuals obtained in part (e). Interpret the plot.

3.8 The data file retail lists total U.K. (United Kingdom) retail sales (in billions of pounds) from January 1986 through March 2007. The data are not "seasonally adjusted," and year 2000 = 100 is the base year.

(a) Display and interpret the time series plot for these data. Be sure to use plotting symbols that permit you to look for seasonality.

(b) Use least squares to fit a seasonal-means plus linear time trend to this time series. Interpret the regression output and save the standardized residuals from the fit for further analysis.

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

3.9 The data file prescrip gives monthly U.S. prescription costs for the months August 1986 to March 1992. These data are from the State of New Jersey's Prescription Drug Program and are the cost per prescription claim.

(a) Display and interpret the time series plot for these data. Use plotting symbols that permit you to look for seasonality.

(b) Calculate and plot the sequence of month-to-month percentage changes in the prescription costs. Again, use plotting symbols that permit you to look for seasonality.

(c) Use least squares to fit a cosine trend with fundamental frequency 1/12 to the percentage change series. Interpret the regression output. Save the standardized residuals.

(d) Plot the sequence of standardized residuals to investigate the adequacy of the cosine trend model. Interpret the plot.

3.10 (Continuation of Exercise 3.4) Consider the hours time series again.

(a) Use least squares to fit a quadratic trend to these data. Interpret the regression output and save the standardized residuals for further analysis.

(b) Display a sequence plot of the standardized residuals and interpret. Use monthly plotting symbols so that possible seasonality may be readily identified.

(c) Perform the Runs test of the standardized residuals and interpret the results.

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

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

3.11 (Continuation of Exercise 3.5) Return to the wages series.

(a) Consider the residuals from a least squares fit of a quadratic time trend.

(b) Perform a runs test on the standardized residuals and interpret the results.

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

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

3.12 (Continuation of Exercise 3.6) Consider the time series in the data file beersales.

(a) Obtain the residuals from the least squares fit of the seasonal-means plus quadratic time trend model.

(b) Perform a runs test on the standardized residuals and interpret the results.

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

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

3.13 (Continuation of Exercise 3.7) Return to the winnebago time series.

(a) Calculate the least squares residuals from a seasonal-means plus linear time trend model on the logarithms of the sales time series.

(b) Perform a runs test on the standardized residuals and interpret the results.

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

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

3.14 (Continuation of Exercise 3.8) The data file retail contains U.K. monthly retail sales figures.

(a) Obtain the least squares residuals from a seasonal-means plus linear time trend model.

(b) Perform a runs test on the standardized residuals and interpret the results.

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

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

3.15 (Continuation of Exercise 3.9) Consider again the prescrip time series.

(a) Save the standardized residuals from a least squares fit of a cosine trend with fundamental frequency 1/12 to the percentage change time series.

(b) Perform a runs test on the standardized residuals and interpret the results.

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

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

3.16 Suppose that a stationary time series, $\{Y_t\}$, has an autocorrelation function of the form $\rho_k = \phi^k$ for $k > 0$, where ϕ is a constant in the range $(-1,+1)$.

(a) Show that $Var(\bar{Y}) = \dfrac{\gamma_0}{n}\left[\dfrac{1+\phi}{1-\phi} - \dfrac{2\phi}{n}\dfrac{(1-\phi^n)}{(1-\phi)^2}\right]$.

(Hint: Use Equation (3.2.3) on page 28, the finite geometric sum

$\sum_{k=0}^{n} \phi^k = \dfrac{1-\phi^{n+1}}{1-\phi}$, and the related sum $\sum_{k=0}^{n} k\phi^{k-1} = \dfrac{d}{d\phi}\left[\sum_{k=0}^{n} \phi^k\right]$.)

(b) If n is large, argue that $Var(\bar{Y}) \approx \dfrac{\gamma_0}{n}\left[\dfrac{1+\phi}{1-\phi}\right]$.

(c) Plot $(1+\phi)/(1-\phi)$ for ϕ over the range -1 to $+1$. Interpret the plot in terms of the precision in estimating the process mean.

3.17 Verify Equation (3.2.6) on page 29. (Hint: You will need the fact that

$\sum_{k=0}^{\infty} \phi^k = \dfrac{1}{1-\phi}$ for $-1 < \phi < +1$.)

3.18 Verify Equation (3.2.7) on page 30. (Hint: You will need the two sums

$\sum_{t=1}^{n} t = \dfrac{n(n+1)}{2}$ and $\sum_{t=1}^{n} t^2 = \dfrac{n(n+1)(2n+1)}{6}$.)