

Linear Fit				
TCPI = $-41295.44 + 20.854732 \cdot \text{Year}$				
Summary of Fit				
RSquare		0.961187		
RSquare Adj		0.96024		
Root Mean Square Error		53.25903		
Mean of Response		393.1699		
Observations (or Sum Wgts)		43		
Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F-Ratio
Model	1	2880039.2	2880039	1015.341
Error	41	116297.5	2837	Prob > F
C. Total	42	2996336.7		< .0001*
Parameter Estimates				
Team	Estimate	Std Error	t Ratio	Prob> t
Intercept	-41295.44	1308.338	-31.56	< .0001*
Year	20.854732	0.654483	31.86	< .0001*

Figure 13.4.2

The estimate of the slope, 20.8547, tells us that on average the TCPI is increasing at a rate of 20.8547 per year. Given how well the line fits the data ($R^2 \approx 0.9612$), the trend line is a good descriptor of the data.

The trend line can also be used for short-term prediction. Suppose you wanted to estimate the TCPI for 2021. If the data are not available, the trend model can be used.

$$\text{Estimated TCPI} = -41295.44 + 20.8547(2021) = 851.9087$$

(Prediction of the TCPI for 2021)

One of the problems with this prediction, as with all predictions in this chapter, is that the accuracy of the prediction is unknown. It might be very close to the true value or it could be very inaccurate. If there were some knowledge about the accuracy of the prediction it would be more useful. In later sections, we will return to this topic and study inferential methods.

13.4 Exercises

Basic Concepts

1. Why is the mean not a reasonable descriptor for nonstationary time series data?
2. What is a linear time trend?
3. What is the independent variable in a linear trend model?
4. Is there a difference between the way the best fit line is determined for time series data and the way it is determined for other types of data?
5. Identify a problem with predictions that are made using a time trend model.

Exercises

6. Consider the following table containing the Consumer Price Index (CPI) for all urban consumers in the United States from 1990 to 2010. The index is based on 1982–84 prices.

Consumer Price Index			
Year	Consumer Price Index (CPI)	Year	Consumer Price Index (CPI)
1990	130.7	2001	177.1
1991	136.2	2002	179.9
1992	140.3	2003	184.0
1993	144.5	2004	188.9
1994	148.2	2005	195.3
1995	152.4	2006	201.6
1996	156.9	2007	207.34
1997	160.5	2008	215.30
1998	163.0	2009	214.54
1999	166.6	2010	218.06
2000	172.2		

Source: Bureau of Labor Statistics

- Looking at the data in the table, do you believe the trend line will slope upward or downward?
 - Suppose we are interested in constructing a linear trend model for these data. Identify the independent and dependent variables for this model.
 - Write the general equation for the time trend model in terms of year and CPI.
 - Using statistical software, the following least squares model was determined.

$$\text{Estimated CPI} = -8647.4245 + 4.4107(\text{Year})$$
 Use this model to predict the price level in 2015.
 - Can we determine the accuracy of this prediction? Explain.
7. Consider the following monthly sales data for an up-and-coming technology company.

Sales Data	
Month	Sales (Thousands of Dollars)
1	321
2	542
3	540
4	581
5	641
6	700
7	698
8	710
9	799
10	821
11	833
12	850

- Identify the independent and dependent variables for the linear time trend model.
- Using statistical software, the following summary output was produced.

Data

This data set can be found on stat.hawkeslearning.com under **Discovering Business Statistics, Second Edition > Data Sets > Consumer Price Index.**

SUMMARY OUTPUT				
<i>Regression Statistics</i>				
Multiple R		0.949341195		
R Square		0.901248704		
Adjusted R Square		0.891373575		
Standard Error		51.20789475		
Observations		12		
ANOVA				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Regression	1	239318.1818	239318.1818	91.26449427
Residual	10	26222.48485	2622.248485	
Total	11	265540.6667		
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	403.7575758	31.51628057	12.81107949	1.57569E-07
Month	40.90909091	4.282219283	9.553245222	2.41268E-06

Write the estimated regression equation.

- What is the mean square error for this model? The standard error?
 - Using this model, predict the company's sales for the 13th month.
 - What percent of the variation in sales is explained by the linear time trend model? Does this model seem to accurately fit the data?
8. Consider the following table containing unemployment rates for North Carolina and South Carolina in 2000 through 2010.

Unemployment Rates 2000–2010		
Year	Unemployment Rate (%)	
	North Carolina	South Carolina
2000	3.7	3.6
2001	5.6	5.2
2002	6.6	6.0
2003	6.5	6.7
2004	5.5	6.8
2005	5.3	6.8
2006	4.8	6.4
2007	4.7	5.6
2008	6.2	6.8
2009	10.8	11.3
2010	10.6	11.2

Source: Bureau of Labor Statistics

- Using statistical software, estimate the following linear time trend model:

$$\text{N.C. Unemployment Rate} = \beta_0 + \beta_1 (\text{Year}) + \varepsilon_i.$$

Write the estimated regression equation using the least squares estimates for β_0 and β_1 .

- Using statistical software, estimate the following linear time trend model:

$$\text{S.C. Unemployment Rate} = \beta_0 + \beta_1 (\text{Year}) + \varepsilon_i.$$

Write the estimated regression equation using the least squares estimates for β_0 and β_1 .

- Use the equations in parts **a.** and **b.** to estimate the unemployment rates for North and South Carolina in the year 2013.

- d. What is the coefficient of determination for the regression model in part **a**?
- e. What is the coefficient of determination for the regression model in part **b**?
- f. Do you think that these regression models are reliable in predicting future unemployment rates? Of the two models, which seems to fit the data better?

13.5 Inference Concerning the Slope

Since β_1 specifies the rate of change between x and y , in most linear models the parameter of interest is β_1 . Two inferential techniques are useful in evaluating the estimate of β_1 . Confidence intervals, similar in structure to those used for means and proportions, will be developed. In addition, a hypothesis testing procedure will be presented to test whether β_1 is equal to some particular value.

The Confidence Interval for β_1

Developing a confidence interval for β_1 requires thinking about the estimate b_1 as a random variable. Each random sample from the population will produce different data and hence different estimates of b_0 and b_1 . The confidence interval will serve two purposes: to place bounds on the location of β_1 and to provide information about the quality of the point estimate, b_1 . The form of the confidence interval is familiar.

$$\text{Sample estimate of parameter} \pm \left(\begin{array}{l} \text{A certain number of standard} \\ \text{deviation units depending on} \\ \text{the desired confidence} \end{array} \right) \cdot \left(\begin{array}{l} \text{The standard} \\ \text{deviation of the} \\ \text{sample estimate} \end{array} \right)$$

The sample estimate of β_1 is b_1 . The variance of b_1 is given by

$$\sigma_{b_1}^2 = \frac{\sigma_\varepsilon^2}{\sum(x_i - \bar{x})^2}$$

but like all population measurements, $\sigma_{b_1}^2$ usually has to be estimated from the data. Notice that the denominator of the expression above is equal to the variance of x , multiplied by the sample size, n . This indicates that the variance of b_1 is reduced if the variance of the error terms decreases, the sample size increases, or the variance of x increases.

The sample estimate of the variance of b_1 is given by

$$s_{b_1}^2 = \frac{s_\varepsilon^2}{\sum(x_i - \bar{x})^2}.$$

The only difference in the computation of $\sigma_{b_1}^2$ and $s_{b_1}^2$ is the replacement of the population variance of the error terms, σ_ε^2 , with the corresponding sample statistic, s_ε^2 . The standard deviation (standard error) of the sample estimate b_1 is

$$s_{b_1} = \sqrt{\frac{s_\varepsilon^2}{\sum(x_i - \bar{x})^2}}.$$

Formula

100(1- α)% Confidence Interval for β_1

The 100(1 - α)% confidence interval for β_1 is given by

$$b_1 \pm t_{\alpha/2, df} s_{b_1},$$

where $t_{\alpha/2, df}$ is the critical value for a t -distribution with $n - 2$ degrees of freedom.