Chapter 6 Regression Analysis

Linear regression models are usually unweighted. That is, they give all past data equal weight. This is a serious mistake when the relationship between X and Y changes with time, the usual case with business data. Thus DISCREG improves forecast accuracy by giving recent data more weight in computing the regression coefficients. Another regression alternative is AUTOREG, in which the independent variable is a series of previous values of the dependent variable. This type of regression is useful in modeling data containing cycles.

6.1 Discounted regression (DISCREG)

The linear growth model (LIN) in the previous chapter automatically uses time (a series of consecutive integers) as the independent variable. This can easily be changed when the independent variable is causal in nature: simply replace the X range with new data.

One drawback to LIN is that it gives equal weight to all past data in forecasting. This is a serious mistake when the relationship between X and Y changes with time. DISCREG in Figure 6-1 often improves forecast accuracy by giving recent data more weight in computing the regression equation.[1] This has been the experience at Brookshire Cookware in Brookshire, Texas, which uses DISCREG to predict warranty claims based on production in the previous quarter.

In Figure 6-1, X data are production in the previous quarter, while Y data are warranty claims in the current quarter. All data are rounded to the nearest thousand units. For example, the first data pair shows that warranty claims in the quarter ending December, 1997, were 33,000, while production in the previous quarter (ending September, 1997) was 397,000 units. Warranty claims have grown with production volume, although growth slowed during 1999.

To model the relationship between claims and production, we could fit a straight line using ordinary least squares that gives equal weight to all data; unfortunately, this model pulls the regression line toward the early data (see Figure 6-1, which is the example tab in DISCREG). The reason is that the line is computed to minimize the sum of squared errors, defined as distances between the line and the actual data. The farther a data point is from the line, the more serious is the error. A better choice is the discounted regression line in Figure 6-1. The discounted line is also computed to minimize the sum of squared errors, but recent data are weighted more heavily. The result is to pull the line toward recent data and produce more accurate forecasts.

The worksheet in Figure 6-2 shows how the discounting procedure works. In cell C12, enter a discount factor, a fractional number which determines the weights in column E. The current discount factor is 0.70, resulting in weights that decline from 1.00 for the most recent data to 0.29 for the oldest. Actual data are multiplied by the weights in column E to produce discounted or weighted data (columns G - I) for input to the regression.

---

[1] *This model is based on Everette S. Gardner, Jr., "Forecasting the Failure of Component Parts in Computer Systems: A Case Study," International Journal of Forecasting, Vol. 9, No. 2 (August, 1993), pp. 245-253.*
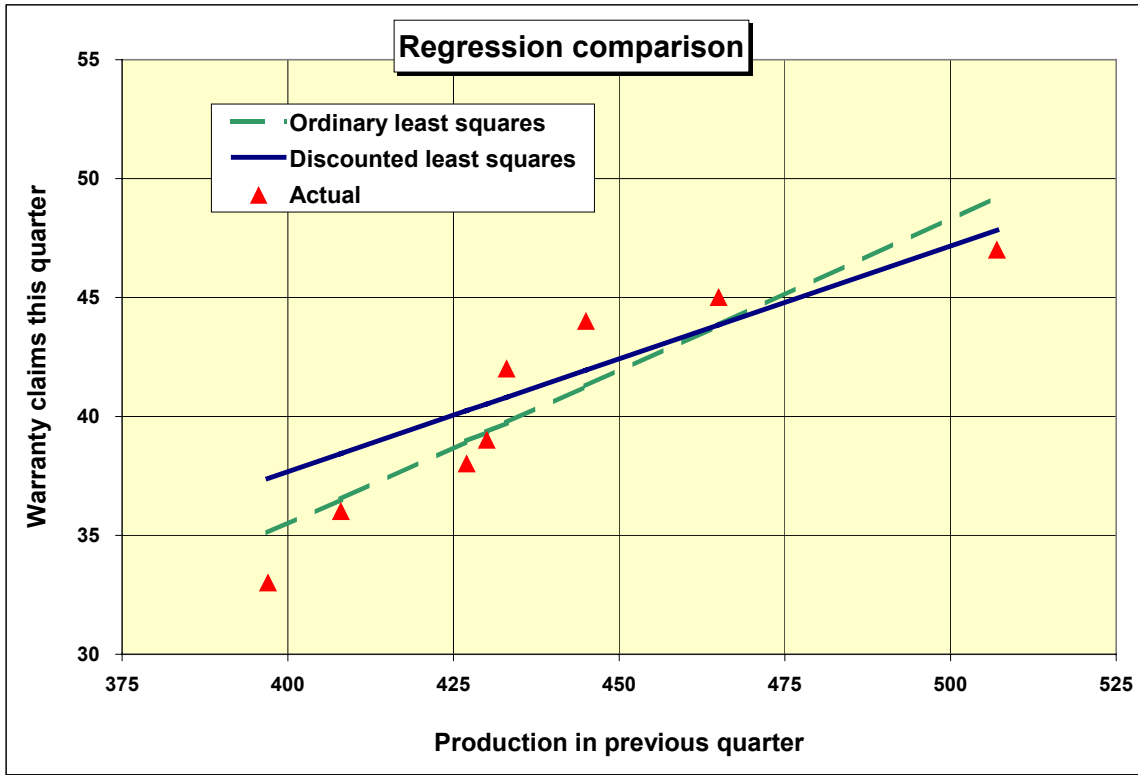
Figure 6-1

Figure 6-2

	A	B	C	D	E	F	G	H	I	J
1	DISCREG			Warranty claims vs. production						
2	**Discounted Regression**							Number of data		8
3	Enter data starting at row B18:D18.							RMSE		6.34
4	Enter discount factor in C12.									
5	Select Tools Data Analysis Regression.									
6	Select Y range from I18 through last cell containing a value.									
7	Select X range from *G18..H18* through last cell containing a value.									
8	Enter output range = O1.									
9	*Check box for "Constant is zero"*									
10	Click OK.									
11										
12	Discount factor		**0.70**							
13	Square root		0.84							
14										
15								**Regression range:**		
16			Claims	Production				**Discounted data**		
17		Date	Y	X	Weights	Intercept	X1	X2	Y	Est Y
18	1	Dec-97	33	397	0.29	1	0.29	113.93	9.47	37.39
19	2	Mar-98	36	408	0.34	1	0.34	139.94	12.35	38.44
20	3	Jun-98	38	427	0.41	1	0.41	175.05	15.58	40.24
21	4	Sep-98	39	430	0.49	1	0.49	210.70	19.11	40.52
22	5	Dec-98	42	433	0.59	1	0.59	253.59	24.60	40.81
23	6	Mar-99	44	445	0.70	1	0.70	311.50	30.80	41.95
24	7	Jun-99	45	465	0.84	1	0.84	389.05	37.65	43.85
25	8	Sep-99	47	507	1.00	1	1.00	507.00	47.00	47.83
26	9	Dec-99		520	#N/A	1	#N/A	#N/A	#N/A	49.06
27	10				#N/A	1	#N/A	#N/A	#N/A	#N/A
28	11				#N/A	1	#N/A	#N/A	#N/A	#N/A
29	12				#N/A	1	#N/A	#N/A	#N/A	#N/A
30	13				#N/A	1	#N/A	#N/A	#N/A	#N/A

When the discount factor is 1.0, all data get equal weight and the result is the same as an ordinary-least-squares regression. Thus you can use DISCREG as a general-purpose regression model.

As the discount factor is reduced, the weights on older data fall off substantially. How do you choose the best discount factor? This is a matter of judgment. Many professional forecasters avoid carrying the discounting idea too far and use moderate values of 0.70, 0.80, or 0.90. I chose 0.70 in this example since the changes in growth are substantial and this discount factor pulls the regression line closest to the last data point.

You can also experiment with a range of different factors and choose the one that gives the smallest MSE or RMSE. Note that DISCREG does not divide the data into warm-up and forecasting samples like the earlier models. The reason is that the weights increase according to a function in column E that cannot be interrupted. However, after the regression is run, you can add as many new X values as you like at the end of the input data in column D. Forecasts for each new X appear in column J.

Another application for DISCREG is to forecast the Y variable as a discounted function of time. For example, we could replace the X data in Figure 6-2 with a column of consecutive integers, the numbers 1-8. The result is a forecast based on a straight-line trend that gives more weight to recent data. Try this model and the forecasts are very similar to the model based on production data. However, management preferred to use the causal model since it allowed what-if analysis on production rate changes.

Forecasts based on time are very similar to the TREND exponential smoothing model in Chapter 4. DISCREG is actually simpler than TREND because it has only one parameter, the discount factor. TREND has three parameters: smoothing weights for level and trend and a trend modifier.

Pay special attention to the regression instructions at the top of Figure 6-2. The regression range is found in columns G – I. Data in this range are the original data multiplied by the weights in column D. The Y range is column I while the X range includes columns G and H. The regression has two X variables: the first is the column of weights and the second is the original X data times the weights. The column of weights is used to compute the regression intercept, which moves with time. Note that the "Constant is zero" box must be checked before executing the regression. This is because the model computes its own weighted intercept.

DISCREG does not contain a control chart for forecast errors. This would be misleading because early errors are not important. The model focuses on achieving a good fit to the most recent data.