CURVEFIT

Eleven models can be selected to fit times series and produce forecasts, forecast errors, and confidence limits. In all of the models, the observed series is some function of time.

Notation

The following notation is used throughout this chapter unless otherwise stated:

Y_t	Observed series; $t = 1, \ldots, n$
$E(Y_t)$	Expected value of Y_t
$\hat{Y}_t$	Predicted value for Y_t

Models

CURVEFIT allows the user to specify a model with or without a constant term designated by β_0. If this constant term is excluded, simply set it zero or one depending upon whether it appears in an additive or multiplicative manner in the models listed below.

(1) Linear	$E(Y_t) = \beta_0 + \beta_1 t$
(2) Logarithmic	$E(Y_t) = \beta_0 + \beta_1 \ln(t)$
(3) Inverse	$E(Y_t) = \beta_0 + \beta_1 / t$
(4) Quadratic	$E(Y_t) = \beta_0 + \beta_1 t + \beta_2 t^2$
(5) Cubic	$E(Y_t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$
(6) Compound	$E(Y_t) = \beta_0 \beta_1^t$
(7) Power	$E(Y_t) = \beta_0 t^{\beta_1}$
(8) S	$E(Y_t) = \exp(\beta_0 + \beta_1 / t)$

(9) Growth	$E(Y_t) = \exp(\beta_0 + \beta_1 t)$
(10) Exponential	$E(Y_t) = \beta_0 e^{\beta_1 t}$
(11) Logistic	$E(Y_t) = \left(\frac{1}{u} + \beta_0 \beta_1^t \right)^{-1}$

Assumption

We assume that nonlinear models (6) to (11) can be expressed in linear model form by logarithmic transformation. So, for models (6) to (10),

$$\ln(Y_t) = \ln(E(Y_t)) + \varepsilon_t$$

and for model (11),

$$\ln\left(\frac{1}{Y_t} - \frac{1}{u} \right) = \ln\left(\frac{1}{E(Y_t)} - \frac{1}{u} \right) + \varepsilon_t$$

with ε_t, $t = 1, \ldots, n$ being independently identically distributed $N(0, \sigma^2)$.

Application of Regression

Each of the models is expressed in linear form and computational techniques described in the REGRESSION procedure are applied. The dependent variable and independent variables for each model are listed as follows:

Model	Dependent Variable	Independent Variables	Coefficients
(1)	Y	t	β_0, β_1
(2)	Y	$\ln(t)$	β_0, β_1
(3)	Y	$1/t$	β_0, β_1
(4)	Y	t, t^2	$\beta_0, \beta_1, \beta_2$
(5)	Y	t, t^2, t^3	$\beta_0, \beta_1, \beta_2, \beta_3$

Model	Dependent Variable	Independent Variables	Coefficients
(6)	$\ln(Y)$	t	β_0^*, β_1^*
(7)	$\ln(Y)$	$\ln(t)$	β_0^*, β_1
(8)	$\ln(Y)$	$1/t$	β_0, β_1
(9)	$\ln(Y)$	t	β_0, β_1
(10)	$\ln(Y)$	t	β_0^*, β_1
(11)	$\ln\left(\frac{1}{Y}-\frac{1}{u}\right)$	t	β_0^*, β_1^*

where $\beta_0^* = \ln(\beta_0)$ and $\beta_1^* = \ln(\beta_1)$.

The ANOVA table, coefficient estimates and their standard errors, *t*-values, and significance levels are computed as in the REGRESSION procedure. Note that for the nonlinear models (6) to (11), we have

$$se\left(\hat{\beta}_0\right) \approx \exp\left(\hat{\beta}_0^*\right) \times se\left(\hat{\beta}_0^*\right)$$

and

$$se\left(\hat{\beta}_1\right) \approx \exp\left(\hat{\beta}_1^*\right) \times se\left(\hat{\beta}_1^*\right).$$

Predicted Values and Confidence Intervals

The regression coefficients for models (1) to (5) are used to obtain the predicted values. For the transformed models, more computations are required to obtain the predicted values for the original models. The formulas are listed below:

(1) $\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 t$

(2) $\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 \ln(t)$

(3) $\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 / t$

(4) $\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2$

(5) $$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2 + \hat{\beta}_3 t^3$$

(6) $$\hat{Y}_t^* = \hat{\beta}_0^* + \hat{\beta}_1^* t$$

(7) $$\hat{Y}_t^* = \hat{\beta}_0^* + \hat{\beta}_1 \ln(t)$$

(8) $$\hat{Y}_t^* = \hat{\beta}_0 + \hat{\beta}_1 / t$$

(9) $$\hat{Y}_t^* = \hat{\beta}_0 + \hat{\beta}_1 t$$

(10) $$\hat{Y}_t^* = \hat{\beta}_0^* + \hat{\beta}_1 t$$

(11) $$\hat{Y}_t^* = \hat{\beta}_0^* + \hat{\beta}_1^* t$$

where $\hat{Y}_t^* = \ln\left(\hat{Y}_t\right)$ in models (5) to (10), and $\hat{Y}_t^* = \ln\left(\frac{1}{\hat{Y}_t} - \frac{1}{u}\right)$ in model (11).

The 95% prediction interval for an observation at time t is constructed as follows:

For models (1) to (5):

$$\hat{Y}_t \pm t_{0.025}\sqrt{MSE\left(1 + h_t + \frac{1}{n}\right)} \quad \text{if constant term is included}$$

$$\hat{Y}_t \pm t_{0.025}\sqrt{MSE(1 + h_t)} \quad \text{otherwise}$$

For models (6) to (10):

$$\exp\left(\hat{Y}_t^* \pm t_{0.025}\sqrt{MSE\left(1 + h_t + \frac{1}{n}\right)}\right)$$

and for model (11):

$$\frac{1}{\exp\left(\hat{Y}_t^* \pm t_{0.025}\sqrt{MSE\left(1 + h_t + \frac{1}{n}\right)}\right) + \frac{1}{u}}$$

where MSE is the mean square error obtained by fitting the linear model, $t_{0.025}$ is the 97.5 percentage point from Student t-distribution with MSE degrees of freedom, and h_t is the leverage (computational detail in the REGRESSION procedure).