The natural logarithm*

Ragnar Nymoen

January 28, 2013

1 Logarithms and rates of change

We often make use of the approximation

$$\ln\left(1+\frac{\Delta Y}{Y}\right)=\ln\left(\frac{Y+\Delta Y}{Y}\right)=\ln\left(Y+\Delta Y\right)-\ln Y$$
$$\approx\frac{\Delta Y}{Y},\quad \frac{\Delta Y}{Y}\text{ is small.}$$

We will see how this approximation works, and how good it is.

When we consider marginal changes we have the precise relationship, i.e., from the derivative of a function

$$\frac{d\ln Y}{dY}=\frac{1}{Y},\text{ and hence } d\ln Y=\frac{dY}{Y}.$$

To get to the approximation we need the following properties of the logarithmic function: $\ln X$:

$$\ln X=\begin{cases}=\left(X-1\right), & X=1\\ <\left(X-1\right), & X\neq 1\end{cases}$$

The relationships holds because the function $f\left(X\right)=\left(X-1\right)$ is a straight line with slope coefficient $\frac{d(X-1)}{dX}=1$, while the function $g\left(X\right)=\ln X$ is concave, with a slope coefficient that decreases with increasing X:

$$\frac{d\ln X}{dX}=\frac{1}{X}\begin{cases}<1, X>1\\ >1, X<1.\end{cases}$$

This result is illustrated in figure 1.

If we let $X=\frac{Y+\Delta Y}{Y}=1+\frac{\Delta Y}{Y}$, we get the relationship between logarithms and rates of change:

$$\ln\left(1+\frac{\Delta Y}{Y}\right)=\ln\left(\frac{Y+\Delta Y}{Y}\right)=\ln\left(Y+\Delta Y\right)-\ln Y\leq\frac{\Delta Y}{Y}.$$

Changes in logarithms will therefore never give larger values than the exact rates of change. But how good is the approximation, and why does it become poorer when

*This note is a translation of Appendix 3.A in **?**.

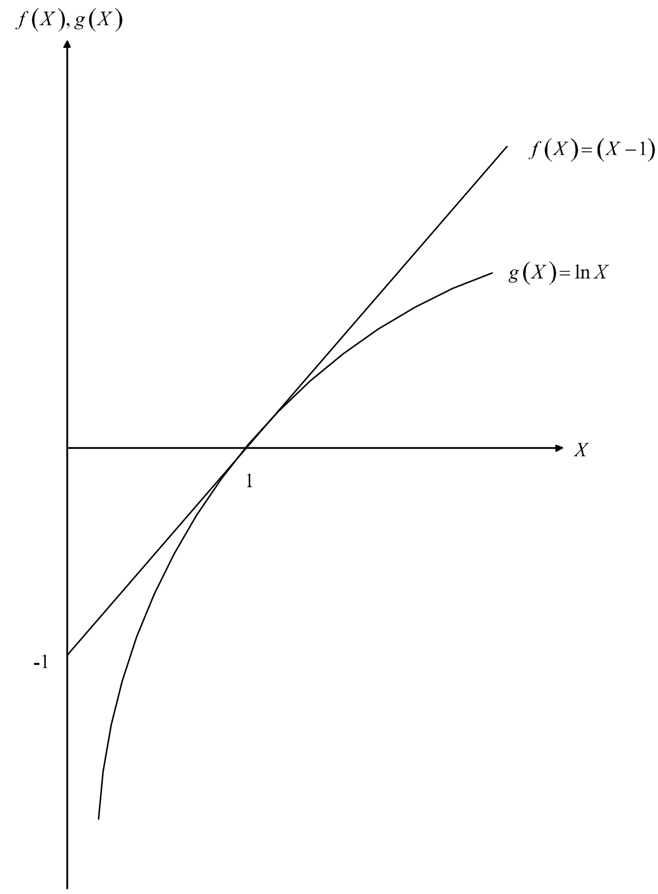

Figure 1: The log-function as a an approximation.

the rate of changes becomes larger? We can investigate this by taking a second order Taylor expansion of $\ln X$ around the value $X_0 = 1$:

$$\begin{aligned}
\ln X &\approx \ln X_0 + \frac{1}{X_0}(X - X_0) - \left(\frac{1}{X_0^2}\right)\frac{(X - X_0)^2}{2} \\
&= \ln 1 + 1(X - 1) - \frac{1}{2}(X - 1)^2 \\
\ln X &\approx (X - 1) - \frac{1}{2}(X - 1)^2 . \qquad (1)
\end{aligned}$$

If we substitute $X = \frac{Y+\Delta Y}{Y}$, as above, we can see how close the approximation is, depending on the size of the rate of change:

$$\ln\left(\frac{Y + \Delta Y}{Y}\right) = \ln(Y + \Delta Y) - \ln Y \approx \frac{\Delta Y}{Y} - \frac{1}{2}\left(\frac{\Delta Y}{Y}\right)^2 .$$

Clearly the approximation is best for smallish rates of change, but it also works well for a 10 % change in Y, as the table shows:

$\frac{\Delta Y}{Y}$	0.01	0.05	0.1	0.2	0.5	1
$\ln\left(1 + \frac{\Delta Y}{Y}\right)$	0.00995	0.04879	0.09531	0.18232	0.405 47	0.693 15

.

2 The standard error of a log-linear model estimated by OLS

Assume that we have estimated a model for $\ln Y_i$, so that we can write:

$$\ln Y_i = \ln \hat{Y}_i + \hat{\varepsilon}_i,$$

where $\hat{Y}_i = e^{\hat{\beta}_0} X_i^{\hat{\beta}_1}$ in the case of a single explanatory variable. The relationship for the variable Y_i becomes:

$$Y_i = \hat{Y}_i e^{\hat{\varepsilon}_i}.$$

From (1), and by dropping the second order term for simplicity, we have:

$$\ln Y_i - \ln \hat{Y}_i = \hat{\varepsilon}_i \approx \frac{Y_i - \hat{Y}_i}{Y_i}, \tag{2}$$

Equation (2) shows that the residual $\hat{\varepsilon}_i$ has an interpretation as a relative prediction error. Hence $100\hat{\varepsilon}_i \approx 100\left(\frac{Y_i - \hat{Y}_i}{Y_i}\right)$ can be interpreted as percentage prediction error. The standard errors of the regression:

$$\begin{aligned} 100\hat{\sigma}_\varepsilon &= 100\sqrt{\frac{1}{n-2}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2} \\ &\approx 100\sqrt{\frac{1}{n-2}\sum_{i=1}^{n}\left(\frac{Y_i - \hat{Y}_i}{Y_i}\right)^2} \end{aligned}$$

is the percentage unexplained standard deviation in the dependent variable. For example, if we have $\hat{\sigma}_\varepsilon = 0.01$, it means that 1 % of the standard deviation of the dependent variable is unexplained by the model we have estimated.

This interpretation is independent of the number of explanatory variables. If we have k variables, we replace $n-2$ by $n-k-1$.