

Log-linear Model

$$\ln(\text{wage}) = \beta_1 + \beta_2 S_i + \epsilon_i$$

β_2 = % change in wage from a 1 unit increase in S_i (schooling)

$$\hat{\ln(\text{wage})} = -7884 + .1038 S_i \quad n = 1000$$

(.0063)
(SE)

so, 1 year of additional schooling is expected to increase wage by 10.38%

95% CI

$$.1038 \pm 1.96(.0063)$$

$$(9.1\%, 11.6\%)$$

The increase in wage for a person earning \$10/hr.

$$\frac{\Delta \hat{E}(\text{wage})}{\Delta S_i} = \beta_2 \cdot \text{wage} = .1038 \cdot \$10 \\ = \$1.038 / \text{hr.}$$

A 95% CI

$$\text{Var}(\hat{\beta}_2 \text{ wage}) = \text{wage}^2 \cdot \text{Var}(\hat{\beta}_2)$$

$$\hat{SE} = \text{wage} \cdot \hat{SE} = 10 \cdot (.0063) \\ = .063$$

$$1.0\overset{37}{\cancel{4}} \pm 1.96(.063)$$

$$(.91, 1.16)$$

log-normal distribution

let $w = e^y$, then $y = \ln(w)$

if $\ln(w) \sim N(\mu, \sigma^2)$ then w is said to have a log-normal distribution.

(It's natural log is normal)

In general $E[\tilde{g}(y)] \neq g E(y)$

$$\therefore E(w) = E(e^y) \neq e^{E(y)}$$

However, $E(w) = e^{(\mu + \sigma^2)/2}$

$$\text{Var}(w) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$$

Special Case : Growth Model.

$$\ln(y) = \beta_1 + \beta_2 t + \epsilon$$

t is a time-trend i.e.,

$$t = 1, 2, 3, \dots, n$$

Let g = rate of growth in y for each year (constant).

y_0 is initial level of y .

$$y_1 = y_0(1+g) = y_0 + y_0 g$$

$$y_2 = y_1(1+g) = y_0(1+g)(1+g) = y_0(1+g)^2$$

$$y_3 = y_2(1+g) = y_0(1+g)^3$$

$$\vdots$$

$$y_t = y_0(1+g)^t$$

Take Log's

$$\begin{aligned} \ln(y_t) &= \ln(y_0(1+g)^t) \\ &= \ln(y_0) + \ln(1+g) * t \end{aligned}$$

$$\ln(y_0) - \text{constant} = \beta_1$$

$$\ln(1+g) - \text{constant} = \beta_2$$

Add error. ϵ_t

$$\ln(y_t) = \beta_1 + \beta_2 t + \epsilon_t$$

LS:

$$\hat{\beta}_2 = \ln(\hat{1+g})$$

Note: when g is "small" $\ln(1+g) \approx g$

In a model with an explanatory variable.

$s=0$ Wage₀ and return to another year = r

$$\text{wage}_1 = \text{wage}_0 (1+r)$$

$$s_i=1 \quad \text{wage}_2 = \text{wage}_1 (1+r) = \text{wage}_0 (1+r)^2$$

$$s_i=2$$

$$s_i=s_i \quad \text{wage}_i = \text{wage}_0 (1+r)^{s_i}$$

$$\ln(\text{wage}_i) = \ln(\text{wage}_0) + \ln(1+r) * s_i$$

$$= \beta_1 + \beta_2 \text{wage} + \epsilon_i$$

$$\hat{\beta}_2 \approx \hat{r}$$

In terms of linear regression,

$$\ln(y) = \beta_1 + \beta_2 x_i + e_i \quad e_i | y_i \sim N(0, \sigma^2)$$

Then

$$\begin{aligned} E(y_i) &= E(e^{\beta_1 + \beta_2 x_i + e_i}) \\ &= E[e^{\beta_1 + \beta_2 x_i} \cdot e^{e_i}] \\ &= e^{\beta_1 + \beta_2 x_i} \cdot E(e^{e_i}) = e^{\beta_1 + \beta_2 x_i} \cdot e^{\sigma^2/2} \\ &= e^{\beta_1 + \beta_2 x_i + \sigma^2/2} \\ &= e \end{aligned}$$

to estimate this, replace σ^2 with a consistent estimate, $\hat{\sigma}^2$.

$$\begin{aligned} E(\hat{y}_i) &= e^{\hat{\beta}_1 + \hat{\beta}_2 x_i + \hat{\sigma}^2/2} \\ &= \hat{y}_{\text{mon}} \cdot e^{\hat{\sigma}^2/2} \end{aligned}$$

$$\hat{y}_{\text{mon}} = \exp \{ \hat{\ln}(y) \}$$

Predicted value from the regression

For 10 years schooling

$$\begin{aligned}\widehat{\ln(\text{wage})} &= .7884 + .1038 \cdot 10 \\ &= .7884 + 1.038 = 1.826\end{aligned}$$

$$\widehat{\text{wage}_{nn}} = \exp\{1.826\} = 6.209$$

$$\widehat{\delta}^2 = .2402$$

$$\begin{aligned}\widehat{\text{wage}}_{\text{corrected}} &= 6.209 \cdot \exp\{.2402/2\} \\ &= 6.209 \cdot 1.27 \doteq 7\end{aligned}$$

Uncorrected prediction \$6.209
corrected " \$7.00

MLR

$$y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i \\ i=1, 2, \dots, n$$

$$E(\epsilon_i | X_{i,j}) = 0 \quad \forall i, j$$

$$V(\epsilon_i | X_{i,j}) = \sigma^2 \quad \forall i, j$$

$$\text{Cov}(\epsilon_i, \epsilon_j) = 0 \quad \forall i \neq j$$

Values of X_{ij} are not exact
average comb of others.

Interpreting coeffs.

$$\partial E(y_i | x_i) = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_k x_{ik}.$$

$$\frac{\partial E(y_i | x_i)}{\partial x_{ij}} = \beta_j$$

It is the effect of a 1 unit increase
in variable j on the average value
of y , holding all other variables in
the model constant.