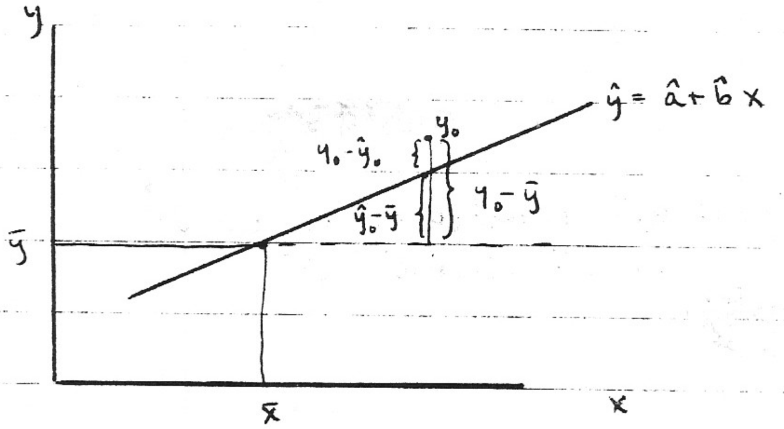
Measurement of Explanatory Power

Having "fitted" a line through the scatter diagram, next we ask, how well does it fit the data?



Consider

$$y_t - \bar{y} = y_t - \hat{y}_t + \hat{y}_t - \bar{y}$$

Note $y_t = \hat{y}_t + \hat{u}_t$ so $\hat{u}_t = y_t - \hat{y}_t$. Then

$$\Sigma (y_t - \bar{y})^2 = \Sigma \hat{u}_t^2 + \Sigma (\hat{y}_t - \bar{y})^2 + 2 \Sigma \hat{u}_t (\hat{y}_t - \bar{y})$$

But last term is

$$2 \left[\Sigma \hat{u}_t (\hat{a} + \hat{b} x_t - a - b\bar{x} - \bar{u}) \right] = 2 \left[(\hat{a} - a) \Sigma \hat{u}_t - b\bar{x} \Sigma \hat{u}_t - \bar{u} \Sigma \hat{u}_t + \hat{b} \Sigma \hat{u}_t x_t \right]$$

$$= 0$$

since $\Sigma \hat{u}_t = 0$ and $\Sigma \hat{u}_t x_t = 0$.

Hence

$$\Sigma(y_t-\bar{y})^2 = \Sigma(\hat{y}_t-\bar{y})^2 + \Sigma\,\hat{u}_t^2$$

$$\underset{\text{total sum of squares}}{TSS} = \underset{\text{explained sum of sq.}}{ESS} + \underset{\text{residual sum of sq.}}{RSS}$$

We then use the <u>coefficient of determination</u> :

$$R^2 = \frac{ESS}{TSS} \equiv 1 - \frac{RSS}{TSS}$$

Provided we include a constant, $0 \le R^2 \le 1$.

Note that if $R^2 = 0$, $\Sigma(\hat{y}-\bar{y})^2 = 0$ which implies $\hat{y}_t = \bar{y}$ $\forall t$. $\Rightarrow$

$$\hat{a} + \hat{b}x_t = \hat{a} + \hat{b}\bar{x} \Rightarrow \hat{b}(x_t - \bar{x}) = 0 \Rightarrow \hat{b} = 0$$

since x_t varies.

<u>Relation between Correlation Coefficient & R^2</u>

Consider

$$\hat{\rho}_{y\hat{y}} = \frac{\Sigma(y_t-\bar{y})(\hat{y}_t-\bar{y})}{\left[\Sigma(y_t-\bar{y})^2\,\Sigma(\hat{y}_t-\bar{y})^2\right]^{\frac{1}{2}}}$$

But $$\Sigma(y_t-\bar{y})(\hat{y}_t-\bar{y}) = \Sigma(\hat{y}_t+\hat{u}_t-\bar{y})(\hat{y}_t-\bar{y})$$

$$= \Sigma(\hat{y}_t-\bar{y})^2 + \Sigma\,\hat{u}_t(\hat{y}_t-\bar{y})$$

$$= \Sigma(\hat{y}_t-\bar{y})^2 + \Sigma\,\hat{u}_t\hat{y}_t - \bar{y}\,\Sigma\,\hat{u}_t = \Sigma(\hat{y}_t-\bar{y})^2$$

then

$$\hat{\rho}_{y\hat{y}} = \frac{\Sigma(\hat{y}_t-\bar{y})^2}{\left(\Sigma(y_t-\bar{y})^2\,\Sigma(\hat{y}_t-\bar{y})^2\right)^{\frac{1}{2}}} = \frac{RSS}{(RSS\cdot TSS)^{\frac{1}{2}}}$$

or

$$\hat{\rho}^2_{y\hat{y}} = R^2$$

Gauss-Markov Theorem

$\boxed{\hat{b} \text{ is "BLUE"}}$

We showed that

$$\hat{b} = \Sigma w_t y_t \qquad \text{where} \quad w_t = \frac{x_t - \bar{x}}{\Sigma (x_t - \bar{x})^2}$$

and

$$\sigma^2_{\hat{b}} = \sigma^2_u \Sigma w_t^2 = \frac{\sigma^2_u}{\Sigma (x_t - \bar{x})^2}$$

Now consider an arbitrary linear estimator $\tilde{b}$:

$$\tilde{b} = \Sigma (w_t + v_t) y_t = \hat{b} + \Sigma v_t y_t$$

$$= \hat{b} + \Sigma v_t (a + b x_t + u_t)$$

$$= \hat{b} + a \Sigma v_t + b \Sigma v_t x_t + \Sigma v_t u_t$$

If $\tilde{b}$ is unbiased, $E(\tilde{b}) = b$. Since we know $E(\hat{b}) = b$, it follows from

$$E(\tilde{b}) = E(\hat{b}) + a \Sigma v_t + b \Sigma v_t x_t + \Sigma v_t E(u_t)$$

the unbiasedness of $\tilde{b}$ implies

$$\sum v_t = 0 \quad \text{and} \quad \sum v_t x_t = 0$$

Since we established $\hat{b} = \sum w_t u_t + b$,

$$\tilde{b} = b + \sum (w_t + v_t) u_t$$

Then

$$\sigma^2_{\tilde{b}} = \sigma^2_u \sum (w_t + v_t)^2$$

Note that cross terms vanish:

$$\sum w_t v_t = \frac{\sum (x_t - \bar{x}) v_t}{\sum (x_t - \bar{x})^2} = \frac{\sum x_t v_t - \bar{x} \sum v_t}{\sum (x_t - \bar{x})^2} = 0$$

Therefore

$$\sigma^2_{\tilde{b}} = \sigma^2_u \sum (w_t^2 + v_t^2)$$

$$\sigma^2_{\hat{b}} = \sigma^2_u \sum w_t^2$$

so $\quad \sigma^2_{\tilde{b}} \geq \sigma^2_{\hat{b}}$ and $=$ only if $v_t \equiv 0$.

$\hat{b}$ is the best linear unbiased estimator of b qed.
(efficient - smallest variance)

Hypothesis Testing

Assumption: $u \sim N(0, \sigma_u^2)$

Then

$$\hat{a} \sim N(a, \sigma_{\hat{a}}^2)$$

$$\hat{b} \sim N(b, \sigma_{\hat{b}}^2)$$

where $\sigma_{\hat{a}}^2$ and $\sigma_{\hat{b}}^2$ were derived earlier.

In standardized form

$$\frac{\hat{a}-a}{\sigma_{\hat{a}}} \sim N(0,1) \quad ; \quad \frac{\hat{b}-b}{\sigma_{\hat{b}}} \sim N(0,1)$$

Then we make the probability statement

$$P\left(-1.96 \leq \frac{\hat{b}-b}{\sigma_{\hat{b}}} \leq 1.96\right) = 0.95$$

$\Rightarrow$

$$P(\hat{b} - 1.96\,\sigma_{\hat{b}} \leq b \leq \hat{b} + 1.96\,\sigma_{\hat{b}}) = .95$$

We are 95% confident that b lies in the interval

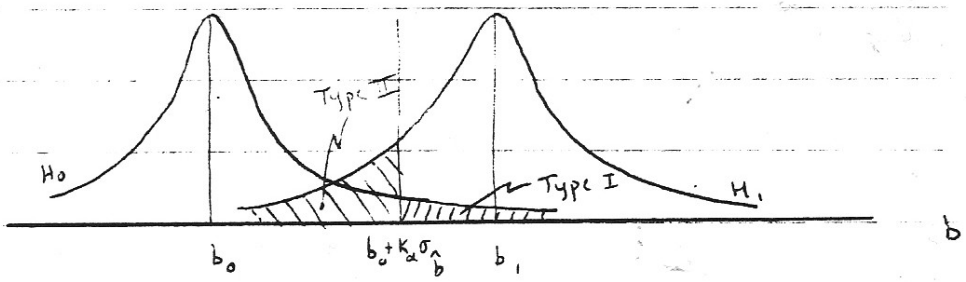
Then the confidence interval $(\hat{b} \pm 1.96\,\sigma_{\hat{b}})$ contains the true (nonstochastic) value b with probability .95.

The higher $\sigma_{\hat{b}}$ the wider the interval. This is an example of interval estimation as opposed to point estimation.

$$P(\hat{b} - K_\alpha \sigma_{\hat{b}} \le b \le \hat{b} + K_\alpha \sigma_{\hat{b}}) = 1-\alpha$$

α = significancl level or Type I error. This is the prob of rejecting H_0 when H_0 is true. Also is Type II error β, accepting H_0 when H_0 is false.

Type I and II Errors



1. As $\alpha \uparrow$, confidence interval widens. Then Type I error falls but Type II error rises.

2. Given Type I error α, Type II error increases as the H_1 moves closer to H_0. It's harder to distinguish H_0 from H_1 if H_0 and H_1 are very similar.

→ Specify a hypothesis to test: the "null" hp, to be compared with the "alternative" hp

1. Two-Tailed Tests

$H_0 : \quad b = b_0$

$H_1 : \quad b \neq b_0$

→ Choose a test statistic. Here $z = (\hat{b} - b_0)/\sigma_{\hat{b}} \sim N(0,1)$

→ Choose α and construct $\hat{b} \pm K_\alpha \sigma_{\hat{b}}$ and check whethe it includes hypothesized value b_0 under H_0.

$\hat{b} - K_\alpha \sigma_{\hat{b}} \leq b_0 \leq \hat{b} + K_\alpha \sigma_{\hat{b}}$

→ Reject H_0 if abs. value of z is greater than critical value. That is, reject if: $\hat{b} > b_0 + K_\alpha \sigma_{\hat{b}}$

or:

$$\hat{b} < b_0 - K_\alpha \sigma_{\hat{b}}$$

A special case is $H_0 : b = 0$ vs $H_1 : b \neq 0$. This just sets $b_0 = 0$; same procedure.

2. One-Tailed Test

$H_0 : \quad b = b_0$

$H_1 : \quad b > b_0$

$$P\left(\frac{\hat{b} - b_0}{\sigma_{\hat{b}}} < K_\alpha\right) = 1 - \alpha$$

or $\quad P(\hat{b} < b_0 + K_\alpha \sigma_{\hat{b}}) = 1-\alpha.$

$\hat{b} > b_0 + K_\alpha \sigma_{\hat{b}}$ reject H_0

$\hat{b} < b_0 + K_\alpha \sigma_{\hat{b}}$ accept H_0

$H_0: \; b = b_0$
$H_1: \; b < b_0$

$$P(\hat{b} - b_0 > K_\alpha \sigma_{\hat{b}}) = 1-\alpha$$

so $\quad \hat{b} < b_0 + K_\alpha \sigma_{\hat{b}}$ reject H_0
$\hat{b} > b_0 + K_\alpha \sigma_{\hat{b}}$ accept H_0

Problem:

$\underline{\sigma_u^2 \text{ not known}}$

Then we have $\hat{\sigma}_{\hat{b}}$ and $\hat{\sigma}_{\hat{a}}$. Then from elementary dist theory $(\hat{b}-b)/\hat{\sigma}_{\hat{b}} \sim t_{(T-2)}$ where $T-2$ = degrees of freedom. Then we use critical value of t rather than $N(0,1)$.

The reason is that $\hat{\sigma}_{\hat{b}}$ is distributed χ^2 (as sums of square of normal distributions are χ^2 distributed: $\hat{\sigma}_{\hat{b}} = \frac{\Sigma \hat{u}_t}{T-2}$) and the ratio of a normal to a χ^2 is distributed student t.

Rule of Thumb for $H_0: b=0$ vs $H_1: b \neq 0$.

$$\hat{b} \pm t \cdot \hat{\sigma}_{\hat{b}} \quad \text{contain zero ?}$$

Reject H_0 $\quad \hat{b} + t\,\hat{\sigma}_{\hat{b}} < 0 \qquad \hat{b} < 0$

$\hat{b} - t\,\hat{\sigma}_{\hat{b}} > 0 \qquad \hat{b} > 0$

Compactly

Reject H_0 if $\left| \dfrac{\hat{b}}{\hat{\sigma}_{\hat{b}}} \right| > t$

The test $H_0: b=0$ amounts to test whether or not the model is meaningful. Indeed, under H_0 $y_t = a + u_t$ so that x_t has no bearing on y_t.
A similar information is provided by the case when $R^2 = 0$, as a measure of goodness of fit. The R^2 is simply an indicator but we may want to test the hp $H_0: R^2 = 0$ against the alternative $H_1: R^2 \neq 0$.

Recall that: $R^2 = \dfrac{\Sigma(\hat{y}_t - \bar{y})^2}{\Sigma(y_t - \bar{y})^2}$ and $1 - R^2 = \dfrac{\Sigma \hat{u}_t^2}{\Sigma(y_t - \bar{y})^2}$

Consider the statistic: $F = \dfrac{(T-2)R^2}{1-R^2} = \dfrac{\Sigma(\hat{y}_t - \bar{y})^2}{\dfrac{\Sigma \hat{u}_t^2}{T-2}}$

The ratio of two χ^2 distributions is distributed F, Hence $F \sim F_{(T-2,\,T-1)}$