Contents:

Simultaneous Inference

Situations in which these arise in regression:

1. Joint inference on β_0 and β_1
2. Simultaneous inference on the mean μ
3. Simultaneous prediction intervals

IDEA:

Want to make inference on more than one parameter, at the same time without inflating the error rate.

Bonferroni's Joint confidence intervals:

- Very general method, can be used in a wide range of situations
- $(1-\alpha)$ confidence interval Bonferroni's method is a lower bound of the true level
- Conservative

IDEA:

Regular confidence interval is of form:

$$\text{Estimate} \pm t(\alpha/2, df_{error}) \text{ Std. Error}$$

If g comparisons are to be made then Bonferroni's method gives the following confidence interval:

$$\text{Estimate} \pm t(\alpha/2\mathbf{g}, df_{error}) \text{ Std. Error}$$

With a family confidence coefficient of $(1-\alpha)$

Other procedures like Working Hotelling procedure:

Estimate ± WStd. Error

Where, $W^2=2F(\alpha,\mathbf{2}, df_{error})$

And Scheffe's technique can also be used for g comparisons.

Estimate ± S Std. Error

Where, $S^2=gF(\alpha,\mathbf{g}, df_{error})$

Regression Through Origin:

Sometimes prior information on the data gives that if X=0 then Y=0

Example: X = units of sale
Y = tip or commission

In situations like this we want to fit a regression model through the origin.

$$Yi = \beta_1 X_i + \varepsilon_i$$

$$E(Y) = \beta_1 X_i$$

We want to minimize $\Sigma (Y_i - \beta_1 X_i)^2$

$$\hat{\beta}_1 * = \frac{\sum X_i Y_i}{\sum X_i^2} \text{ and } e_i^* = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_1^* X_i$$

Words of Caution:

- In regression through origin the residuals do not add up to zero.
- When data form a very curvilinear pattern SSTP may be less than SSE.

<u>In general best to use intercept even if it is close to 0.</u>

Measurement Errors:

In Y:

If the errors are:

- Independent
- Unbiased
- Uncorrelated

They are absorbed in $\varepsilon_{i.}$

In X:

More difficult problem, since the assumptions in regression, is no measurement errors in the X's.

Inverse Predictions:

Predicting x given y.

Situation:

Suppose we have two measuring instruments, one is quite accurate but is expensive and slow, the other is fast and less expensive but also not very accurate.

If the measurements are highly correlated a measurement for the more expensive machine can be predicted fairly well using the less expensive one.

So here X= more expensive, Y = less expensive

We could regress X using Y as a predictor. This is termed Inverse regression since Y is used as the predictor. Now the terms X and Y are arbitrary and the question is why not just change things around and call Y as X and vice versa.

The problem is our assumption that the X's are being measure without error.

So if we have our dependent variable not random that makes life considerably difficult, since we really have to bend the theory to justify results.

One approach is to regress Y using X as the predictor the usual way and then algebraically solve for a new value of x.

For example $X_{h(new)} = (Y_{h(new)} - b_0)/b_1$

And the prediction interval will be given by

$$\hat{X}_{h(new)} \pm t(1-\alpha/2)s\{predX\}$$

with s^2{predX} $= \frac{MSE}{b_1^2}[1 + \frac{1}{n} + \frac{(\hat{X}_{h(new)} - \bar{X})^2}{S_{xx}}]$

Examples: Selling price (Y) based on cost (X) of firms in a trade association. Suppose selling price of a firm in the association is known how do we estimate the cost?