

Question on Lack of Fit test in Simple Linear Regression.

Source: <http://sci.tech-archive.net/Archive/sci.stat.math/2007-09/msg00394.html>

- *From:* isabellesup@xxxxxxxxxxxx
 - *Date:* Sat, 22 Sep 2007 08:46:33 -0700
-

Dear Forum,

The lack of fit test (for simple linear regression) tests the following hypotheses:

$H_0: E\{Y\} = \beta_0 + \beta_1 X$

$H_a: E\{Y\} \text{ not equal to } \beta_0 + \beta_1 X$

Basically, the test works as follows:

1/Write a "full" model $Y_{ij} = \mu_j + \epsilon_{ij}$
compute $SSE(\text{Full}) = \sum(i, \sum(j, \{Y_{ij} - \bar{Y}_{.j}\}^2))$

2/Write a "reduced" model $Y_{ij} = \beta_0 + \beta_1 X_j + \epsilon_{ij}$
compute $SSE(\text{Reduced}) = \sum(i, \sum(j, \{Y_{ij} - \beta_0 - \beta_1 X_j\}^2))$

3/Set up a F test that compares $SSE(\text{Full})$ with $SSE(\text{Reduced})$

This test requires repeat observations at one or more X levels.

I am trying to answer the following question:

Is there any advantage in having an equal number of replications at each of the X levels? Is there any disadvantage?

At this point, I cannot find a reason why an equal number of replications would be an advantage.

Many thanks for your input.

.