

Re: Selection of explanatory variables in a model

Source: <http://sci.tech-archive.net/Archive/sci.stat.math/2006-11/msg00602.html>

- *From:* "Beliavsky" <beliavsky@xxxxxxx>
 - *Date:* 28 Nov 2006 05:35:48 -0800
-

Cat wrote:

Dear colisters,

I have to build a logistic regression model and I've got a couple of questions:

– I heard / read (on the Internet) that the number of variables to put as explanatory characters must not be greater than $n/10$ or $n/20$ (n is the sample size). Does someone can provide me with a serious bibliographic reference about this ?

I think it is better to choose a subset using an information criterion such as AIC than to rely on such a rule of thumb. Two books on variable selection are

Subset Selection in Regression
Alan J. Miller

Regression and Time Series Model Selection
by Allan D. R. McQuarrie, Chih-Ling Tsai

.