

13.2: Evils of Stepwise Regression

Almost all statistical software packages (including RR) permit a number of mechanical “search strategies” for finding IVs that make a statistically significant contribution to the prediction of the model-dependent variable. The most common of these is called **stepwise regression**, which may also be referred to as forward, backward (or maybe even upside down!) stepwise regression. Stepwise procedures do not require that the analyst think – you just have to designate a pool of possible IVs and let the package go to work, sifting through the IVs to identify those that (on the basis of your sample data) appear to be related to the model dependent variable. The stepwise procedures use sequential F-tests, sequentially adding variables that “improve the fit” of the mindless model until there are no more IVs that meet some threshold (usually $p < 0.05$) of statistical significance. These procedures are like mechanically wringing all of the explanation you can get for YY out of some pool of XX.

You should already recognize that these kinds of methods pose serious problems. First and foremost, this is an atheoretical approach to model building. But, what if you have no theory to start with – is a stepwise approach appropriate then? No, for several reasons. If any of the candidate XX variables are strongly correlated, the inclusion of the first one will “use up” some of the explanation of the second, because of the way OLS calculates partial regression coefficients. For that reason, once one of the variables is mechanically selected, the other will tend to be excluded because it will have less to contribute to YY. Perhaps more damning, stepwise approaches are highly susceptible to inclusion of spuriously related variables. Recall that we are using samples, drawn from the larger population, and that samples are subject to random variation. If the step-wise process uses the classical 0.05 cut-off for inclusion of a variable, that means that one time in twenty (in the long run) we will include a variable that meets the criterion only by random chance.²³ Recall that the classical hypothesis test requires that we specify our hypothesis in advance; step-wise processes simply rummage around within a set of potential IVs to find those that fit.

There have been notable cases in which mechanical model building has resulted in seriously problematic “findings” that have very costly implications for society. One is recounted in the PBS Frontline episode called “Currents of Fear”.²⁴ [The program was written, produced and directed by Jon Palfreman, and it was first broadcast on June 13, 1995. The full transcript can be found [here](#). The story concerns whether electromagnetic fields (EMFs) from technologies including high-voltage power lines cause cancer in people who are exposed. The problem was that “cancer clusters” could be identified that were proximate to the power lines, but no laboratory experiments could find a connection. However, concerned citizens and activists persisted in believing there was a causal relationship. In that context, the Swedish government sponsored a very ambitious study to settle the question. Here is the text of the discussion from the Frontline program:

... in 1992, a landmark study appeared from Sweden. A huge investigation, it enrolled everyone living within 300 meters of Sweden’s high-voltage transmission line system over a 25-year period. They went far beyond all previous studies in their efforts to measure magnetic fields, calculating the fields that the children were exposed to at the time of their cancer diagnosis and before. This study reported an apparently clear association between magnetic field exposure and childhood leukemia, with a risk ratio for the most highly exposed of nearly 4.

The Swedish government announced it was investigating new policy options, including whether to move children away from schools near power lines. Surely, here was the proof that power lines were dangerous, the proof that even the physicists and biological naysayers would have to accept. But three years after the study was published, the Swedish research no longer looks so unassailable. This is a copy of the original contractor’s report, which reveals the remarkable thoroughness of the Swedish team. Unlike the published article, which just summarizes part of the data, the report shows everything they did in great detail, all the things they measured and all the comparisons they made.

When scientists saw how many things they had measured – nearly 800 risk ratios are in the report – they began accusing the Swedes of falling into one of the most fundamental

| *errors in epidemiology, sometimes called the multiple comparisons fallacy.*

So, according to the Frontline report, the Swedish EMF study regressed the incidence of nearly 800 possible cancers onto the proximity of its citizens to high-voltage power lines. In some cases, there appeared to be a positive relationship. These they reported. In other cases, there was no relationship, and in some the relationship was negative - which would seem to imply (if you were so silly as to do so) that living near the high voltage lines actually protected people from cancer. But only the positive relationships were included in the reports, leading to a false impression that the study had confirmed that proximity to high-voltage lines causes cancer. Embarrassing to the study authors, to put it mildly.

This page titled [13.2: Evils of Stepwise Regression](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Jenkins-Smith et al. \(University of Oklahoma Libraries\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.