Cochran Formula for Regression Graphs

Marianna Bolla

March 21, 2020

In Fig. 1, we illustrate Cochran's formula with three variables within a regression graph. Let X_1, X_2 be on equal standing, connected with a dashed line in the same chain component, whereas their joint parent X_3 is in a chain component of their past (see Fig. 1a). First we predict X_1 only with X_3, with regression coefficient $\beta_{1|3}$. After constructing a Markov equivalent DAG, then say, there is a directed edge $X_2 \to X_1$ (replacing the former dashed line), see Fig. 1b. Along this DAG, we first predict X_2 with X_3, then X_1 with X_3 (old parent) and X_2 (new parent). But in this case the direct and indirect effect of X_3 is added together as $\beta_{1|3.2} + \beta_{1|2.3}\beta_{2|3}$, where $\beta_{1|3.2}$ is the partial regression coefficient of X_3 when regressing X_1, given also X_2 as regressor for X_1. By the Cochran's formula (see also Wermuth–Sadeghi):

$$\beta_{1|3} = \beta_{1|3.2} + \beta_{1|2.3}\beta_{2|3}, \tag{1}$$

so we get the same result if our variables are Gaussians, or we confine ourselves to the second moments. The above equivalence extends to several variables, actually the seminal paper (1923) of S. Wright (geneticist at the Harvard University) discusses it with a more complicated notation.

We can also see that $\beta_{1|3} = \beta_{1|3.2}$ if and only if either $\beta_{1|2.3} = 0$ or $\beta_{2|3} = 0$ in accord with the Theorem we discussed about the Simpson paradox (when it can be avoided).

From Equation (1) we can see that in the special case when $X_3 \to X_2 \to X_1$ form a Markov chain, i.e., $X_1 \perp\!\!\!\perp X_3 | X_2$ (see Fig. 1c), $\beta_{1|3.2} = 0$ and $\beta_{1|2.3} = \beta_{1|2}$. Therefore, $\beta_{1|3} = \beta_{1|2}\beta_{2|3}$.

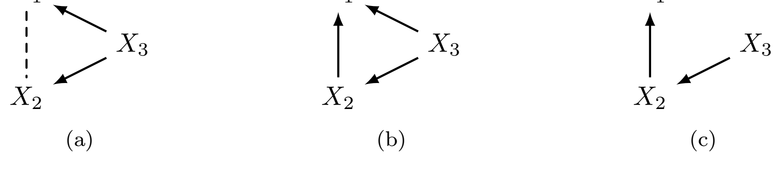

Figure 1: Examples for Cochran's formula