

The presence of missing values complicates statistical analyses. In design of experiments, missing values are particularly problematic when constructing optimal designs, as it is not known which values are missing at the design stage. When data are missing at random it is possible to incorporate this information into the optimality criterion that is used to find designs; Imhof, Song and Wong (2002) develop such a framework. However, when data are not missing at random this framework can lead to inefficient designs. We investigate and address the specific challenges that not missing at random values present when finding optimal designs for linear regression models. We show that the optimality criteria will depend on model parameters that traditionally do not affect the design, such as regression coefficients and the residual variance. We also develop a framework that improves efficiency of designs over those found assuming values are missing at random.