6. Quantitative Comparisons: Results of Finite Sample Size Simulation Studies

The investigations of the last chapter were based on asymptotic arguments. It remains to show that the results of the comparisons are transferable to the finite sample size. Moreover, the properties of the methods themselves have been examined so far only asymptotically, and the estimation of variance is also based on asymptotic results.

In the first section of this chapter we investigate properties of the consistent estimation methods (ML Estimation, PML Estimation and Filling) for the finite sample size. The aim is to demonstrate that these methods work well with respect to unbiasedness and validity of confidence intervals. As we have to expect small deviations of the bias from 0 and of the coverage probability of the confidence intervals from their nominal level, we include Complete Case Analysis in our examination, so that we can regard the deviations as acceptable, if they are not larger than those of Complete Case Analysis.

The comparison of the power to detect an effect $\beta_1 \neq 0$ with these methods is the topic of the second section. Here the major aim is to establish the gain in power in comparison with Complete Case Analysis.

The third section considers the finite properties of Conditional Probability Imputation. Here our interests concern the consequences of asymptotic bias and underestimation of the asymptotic variance for the finite sample size situation.

All examinations for finite sample sizes are based on simulation studies. We restrict our investigation to the case of two dichotomous covariates and we use the same parametrization as in the last chapter. We will also refer to some of the special parameter constellations D1-D6 used there. For each parameter constellation considered we create 2500 data sets of size $n$, and apply the methods to them, i.e. results for different methods but the same parameter constellation are based on an identical sequence of data sets. For the investigations of an estimate $\hat{\beta}_i$ the following statistics were computed and appear in the representations of the results:

- **bias**: The average difference between the estimates and the true parameter.
- **cvg**: The relative frequency to find the true parameter within the 95% confidence interval. This is an approximation to the true coverage probability of the confidence interval. The limits of the confidence interval are computed as $\hat{\beta}_i \pm 1.96 \sqrt{\text{Var}(\hat{\beta}_i)}$ and the variance is estimated as described for each method in Chapter 4.
- **pow**: The relative frequency to find 0 outside of the 95% confidence interval for $\beta_1$. This is an approximation to the power of the associated test to reject the null hypothesis $\beta_1 = 0$.

It is well known that ML estimates for the parameters of a logistic regression model need not to exist, if in some of the strata defined by the combination of the categories of the covariates the outcome variable shows no variation. Hence we cannot exclude that for some of our data sets our estimation procedure fails. This may have also other reasons, e.g. in the PML Estimation procedure the computation of the estimates of the nuisance parameters can result in a division by 0. We try to choose the sample size $n$ always so large that the number of failures of the estimation procedure is less than 25 (i.e., 1% of all data sets); exceptions are indicated in the tables. The reported relative frequencies are always based on the total number of data sets with success of the estimation procedure.

W. Vach, *Logistic Regression with Missing Values in the Covariates*  
© Springer-Verlag New York, Inc. 1994
Quantitative Comparisons: Results of Finite Sample Size Simulation Studies

The choice of 2500 repetitions in the simulation study implies that the standard error of the estimated coverage probability of the 95% confidence interval is less than 0.005 and that the standard error of the estimated power is less than 0.01.

For the computation of the estimates we always used the scoring variant of the Newton-Raphson method. No EM algorithm was used. The initial value for the regression parameters is 0.0. For the computation of the ML estimates the initial value of $\pi_{klj}$ is $\frac{1}{K}$. Convergence was assumed, if the maximal absolute difference between two consecutive parameter values was less than $10^{-8}$. The maximal number of iterations was 25.

6.1 Finite Behavior of ML Estimation, PML Estimation, Filling and Complete Case Analysis

We start our investigation for the parameter constellation of the design $D_1$. In Table 6.1 estimates of the bias of $\hat{\beta}_1$ and the coverage probability of the 95% confidence interval for $\beta_1$ are shown for the following five methods: ML Estimation, the two variants of PML Estimation, Filling and Complete Case Analysis. The results for the first four methods are very similar; differences only concern the third decimal digit. Contrary, the bias of Complete Case Analysis is often distinctly larger. Although the coverage probabilities are only estimated, we observe very few values less than 0.95, so that we can conclude that the confidence intervals are slightly conservative.

The investigation of the parameter constellations $D_2$ and $D_3$ in Tables 6.2 and 6.3 allow similar observations; only for the design $D_2$ with correlated covariates we cannot regard the confidence intervals as conservative.

The fourth investigation is based on the parameter constellation of the design $D_6$ (Table 6.4. As we have here an MDXY missing value mechanism, the variant PMLX of PML Estimation is not applicable, and the estimates of Complete Case Analysis are asymptotically biased. We can observe this bias also in the finite sample size situation; the bias of the other methods is distinctly smaller. The results for ML Estimation, PML Estimation and Filling are again very similar, but Filling shows an increased bias for some constellation.

One may argue that the finite behavior of the methods may break down for parameter constellations with extreme differences in the missing rates. Hence we investigate at least one such constellation; we chose an extreme MDY mechanism with missing rates of 0.1 and 0.9, and the other parameters are chosen such that they cover a constellation with minimal asymptotic relative efficiency with respect to the estimation of $\beta_1$ between ML Estimation and Filling (cf. Table 5.7). We have to choose a sample size of $n = 600$ in order get enough executable data sets for our study. Again the violations are of similar magnitude for all methods, but note that the bias of Filling exceeds for some constellations the bias of Complete Case Analysis.