Formulae for the weighted Jackknife

Nick Patterson

July 21 2005

This is just a short writeup of easy results on the weighted Jackknife, taken from [1]. We suppose we have n observations split into g groups. Group j has size m_j so that

$$\sum_{j=1}^{g} m_j = n$$

Suppose that $\hat{\theta}$ is an estimator of θ, on the whole data set and that $\hat{\theta}_{-j}$ is the corresponding estimator of θ after removing group j. We first define a Jackknifed estimate of θ:

$$\hat{\theta}_J = g\hat{\theta} - \sum_{j=1}^{g} \frac{(n-m_j)\hat{\theta}_{-j}}{n} \quad (1)$$

$$= \sum_{j=1}^{g} (\hat{\theta} - \hat{\theta}_{-j}) + \sum_{j=1}^{g} \frac{m_j \hat{\theta}_{-j}}{n} \quad (2)$$

This estimate will be in practice close to $\hat{\theta}$ but can be shown in some cases to reduce bias. As a sanity check, if $\hat{\theta}$ and $\hat{\theta}_{-j}$ all equal c then $\hat{\theta}_J = c$. In particular, if θ is a vector of probabilities summing to 1 then $\hat{\theta}_J$ will also be. We now give an estimate $\hat{\sigma}^2$ for the variance σ_J^2 of $\hat{\theta}_J$. Write $h_j = n/m_j$. Define a pseudovalue τ_j by

$$\tau_j = h_j \hat{\theta} - (h_j - 1)\hat{\theta}_{-j} \quad (3)$$

Then

$$\hat{\sigma}^2 = \frac{1}{g} \sum_{j=1}^{g} \frac{\left(\tau_j - \hat{\theta}_J\right)^2}{h_j - 1} \quad (4)$$

If θ is a vector quantity then similar formulae can be given for the covariance.

Technically $\hat{\sigma}^2$ is an estimate of the variance of $\hat{\theta}_J$ not of $\hat{\theta}$ and it would be preferable to use $\hat{\theta}_J$ as the basic estimator. This is often not done.

References

[1] F.M.T.A. Busing, E. Meijer, and R. van der Leeden. Delete-m jackknife for unequal m. *Statistics and Computing*, 9:3–8, 1999.