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

\[
\hat{\sigma} = g\hat{\theta} - \sum_{j=1}^{g} \frac{(n-m_j)\hat{\theta}_{-j}}{n}
\]

\[
= \sum_{j=1}^{g} (\hat{\theta} - \hat{\theta}_{-j}) + \sum_{j=1}^{g} \frac{m_j\hat{\theta}_{-j}}{n}
\]  

(1)

\( \hat{\sigma}^2 \) 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 \( \hat{\sigma} = c \). In particular, if \( \theta \) is a vector of probabilities summing to 1 then \( \hat{\sigma} \) will also be. We now give an estimate \( \hat{\sigma}^2 \) for the variance \( \sigma^2 \) of \( \hat{\theta} \). Write \( h_j = n/m_j \). Define a pseudovalud \( \tau_j \) by

\[
\tau_j = h_j\hat{\theta} - (h_j - 1)\hat{\theta}_{-j}
\]  

(3)

Then

\[
\hat{\sigma}^2 = \frac{1}{g} \sum_{j=1}^{g} \frac{(\tau_j - \hat{\theta})^2}{h_j - 1}
\]  

(4)

If \( \theta \) is a vector quantity then similar formulae can be given for the covariance.

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

References