Nonparametric Estimation of Reference Intervals by Simple and Bootstrap-based Procedures, Kristian Linnet
(Laboratory of Clinical Biochemistry, Psychiatric University Hospital, Skovagervej 2, DK-8240 Risskov, Denmark; fax 45-89773549, e-mail linnet@post7.tele.dk)

In recent years, increasing interest has arisen in nonparametric estimation of reference intervals. The IFCC recommendation focuses on the nonparametric procedure, and the NCCLS guideline on reference interval estimation deals exclusively with the nonparametric approach (1, 2). The mentioned reports are based on the simple nonparametric approach, taking as a basis the sorted sample values. In addition to this basic approach, modern computer-based procedures have been introduced, which have made it possible to attain slightly increased precision for the nonparametric approach by applying resampling methods, weighted percentile estimation, or smoothing techniques (3, 4). In the present report, both the simple nonparametric reference interval estimation procedure and the resampling (bootstrap) principle were studied using simulations based on distribution types that should be relevant for clinical chemistry, i.e., gaussian and skewed distributions.

According to the procedure recommended by the IFCC and NCCLS, the observations are ranked according to size, and the 2.5 and 97.5 percentiles are obtained as the 0.025 (n + 1) and 0.975 (n + 1) ordered observations (1, 2). If the estimated rank values are not integers, then linear interpolation is carried out. In the statistical literature, various modifications of the computation procedure have been considered (5–7). Here the traditional one used in clinical chemistry as outlined above (called method I) is compared with an alternative (called method II): p/100 × n + 0.5, where p indicates the percentile (6). For the 2.5 and 97.5 percentiles, method II yields the 0.025n + 0.5 and 0.975n + 0.5 ordered values, respectively. In the following, the above-mentioned calculation principles are referred to as “simple” procedures (IS or IIS) as opposed to “bootstrap” modifications described below (IB or IIB).

The bootstrap principle consists of repeated random resampling of the original observations with replacement, which is performed by a computer (8, 9). Each of the original observations is assigned the same probability of being resampled, i.e., 1/n. For each set of n resampled values, percentile estimates are computed as usual. After repetition of the procedure a large number of times, e.g., 50–500 times, the bootstrap estimates are obtained as the means of these percentile estimates. In the present study, 100 replications were carried out.

A gaussian population distribution was considered as a basis (Fig. 1, top left panel). A mean (μ) of zero and a standard deviation (σ) of 1 were selected as parameters, i.e., corresponding to the standard gaussian distribution. The true 2.5 and 97.5 percentile values for a gaussian distribution are μ ± 1.96 × σ, respectively, i.e., −1.96 and +1.96 for the standard gaussian distribution.

The skewed distribution was generated on the basis of standard gaussian distributed values subjected to the inverse of the Manly exponential transformation [y = [log (1 + ky)]/k] with a selected parameter producing a coefficient of skewness of 1.5 (Fig. 1, top right panel) (10, 11). This degree of skewness may, for example, be observed for reference value distributions of serum concentrations of enzymes (11).

The root mean squared error (RMSE) of percentile estimates represents an overall measure of bias and precision for a given procedure and allows a ranking of the studied procedures. The bottom left panel of Fig. 1 displays a comparison of the RMSE values of simple and bootstrap modifications of methods I and II with regard to the upper percentile for the gaussian distribution (same as for the lower percentile). The IIS procedure clearly outperforms the IS method, and the bootstrap version of method II has the lowest RMSE of all procedures for all sample sizes. The same ranking of the procedures is valid for estimation of the upper percentile of the skewed distribution (Fig. 1, bottom right panel). In general, the differences between the procedures are most pronounced at low to moderate sample sizes. At a sample size of 40, the RMSE of the bootstrap version of method II is 30% (gaussian distribution) or 42% (skewed distribution) lower than that of the IS procedure. The differences are 8% and 7% at a sample size of 500.

A valid statistical estimation method should provide a realistic estimate of the uncertainty associated with the procedure, e.g., expressed as a 90% confidence interval. For the simple estimation procedures (IS and IIS), a 90% confidence interval for the reference interval limits may be derived from the sorted sample values when n is at least 120 (1). Simulations showed that for both the gaussian and the skewed distributions, the actual coverage...
was 91–93%, i.e., in reasonable agreement with the expected value of 90%. For the bootstrap modifications, 90% confidence intervals can also be estimated at lower sample sizes on the basis of the distribution of individual bootstrap percentile estimates, from which the SE is estimated. At the low sample size of 40, a too-low coverage of ~70% was provided, but at higher sample sizes, the coverage was in fairly good agreement with the expected value of 90%. The low coverage at n = 40 relies in a tendency of the bootstrap procedure to underestimate the real uncertainty at small sample sizes.

The general, approximate relationship between SE and sample size is that of a square root relationship: to halve the SE, a fourfold increase in sample size is required. Table 1 displays the relationship between the expected SE and sample size for the IIS and IIB procedures with regard to percentile estimation for gaussian (columns 2 and 3) and skewed distributions (columns 4 and 5), in the latter case with regard to the upper limit. The SE of the percentile estimate is presented here as a percentage of the width of the reference interval to make the relationship generally applicable. The approximate 90% confidence interval is obtained as ± 1.65 SE around the percentile estimate. For a sample size of 120, the 90% confidence interval corresponds to approximately ± 10% around the percentile in case of a gaussian distribution, and approximately ± 25% for the skewed distribution (the percentage of the width of the reference interval). The bootstrap version provides SEs that are 4–12.5% lower than those of the simple version, corresponding to sample

---

**Fig. 1.** Gaussian distribution (top left panel) and skewed distribution with coefficient of skewness of 1.5 (top right panel), with the 2.5 and 97.5 percentile limits demarcated, and RMSEs of the estimated 97.5 percentile as a function of sample size for both distributions (bottom panels). (Bottom left panel), RMSEs for the gaussian distribution; (bottom right panel), RMSEs for the skewed distribution. Based on 5000–10 000 simulation runs.
size savings of 8–24%. In relation to parametric estimation (x̄ ± 1.96 SD), procedure IIB has efficiencies of 48.7% and 57.5% for the gaussian and the skewed distributions, respectively (evaluated at n = 100).

A theoretical treatment of nonparametric percentile estimation shows that the detailed percentile computation formula depends on the type of distribution being considered (7). For a gaussian distribution, the expression (p/100) (n + 0.2) + 0.4 has been recommended for the percentile p (7). This corresponds to the (0.025n + 0.4) and (0.975n + 0.6) ordered values for the 2.5 and 97.5 percentile, respectively. These expressions are very close to the IIS formula, which in some preliminary simulations turned out to perform slightly better and thus was brought into focus in the present study. The above-mentioned expressions have actually been considered for use in clinical chemistry, but they have not received much attention (12, 13).

Overall, procedure IIB provided the lowest RMSE of percentiles for both the gaussian and the skewed distributions at all studied sample sizes. This was also confirmed for other types of distributions that might be of relevance, e.g., the log-normal distribution, the skew model of Box and Cox, and symmetric distributions with kurtosis deviating from that of the gaussian distribution (11). However, at small sample sizes, the estimated confidence interval has a low coverage. Thus, it is advisable to apply the bootstrap procedure mainly at sample sizes exceeding 100. In addition, the high general degree of uncertainty also suggests that a sample size of at least 100 should be considered for nonparametric reference interval estimation.

The bootstrap procedure is related to the weighted percentile method suggested by Harrell and Davis (14). Percentiles are estimated as a weighted average of all possible percentiles, which may reduce the RMSE by ~10–15% for a sample size of 119 (3, 14).

The present study shows that irrespective of what type of nonparametric procedure is used, nonparametric reference interval estimation at small-to-moderate sample sizes is associated with a large degree of uncertainty. A minimum sample size of 120 for nonparametric reference interval estimation has been suggested previously with reference to the lower limit for specification of the 90% confidence intervals of the percentile limits on basis of the sorted sample values. At this sample size, the width of the 90% confidence interval is likely to be ~20% of the length of the reference interval, given a symmetric distribution, but for skewed distributions, the percentage is larger. For sample sizes exceeding 100, the bootstrap procedure, preferably in the IIB version, can be recommended, and the improvement in efficiency is likely to correspond to sample size savings of 10–15%. The bootstrap procedure for reference interval estimation is available in the RefVal program (IB version) distributed by Solberg (15), and in the CBstat program (IB and IIB versions), which is a Windows program distributed (free) by the author (16).

References

Estimation of Serum Apolipoprotein B by a Modified Homogeneous Assay for HDL-Cholesterol, Maureen L. Sampson, Gyorgy Csako, and Alan T. Remaley* (Clinical Pathology Department, Clinical Center, National Institutes of Health, Bldg. 10, Rm. 2C-407, Bethesda, MD 20892-1508; * author for correspondence: fax 301-402-1885, e-mail aremaley@nih.gov)

Serum lipoprotein analysis frequently is used in assessing the risk for coronary artery disease and for monitoring cholesterol-lowering therapy (1, 2). A recent improvement in the analysis of lipoproteins is the development of homogeneous assays for HDL-cholesterol (HDL-C) (3, 4) that are easier to perform because they do not require the

<p>| Table 1. Relationship between sample size and expected SE of percentile estimates obtained by simple (IIS) and bootstrap (IIB) procedures. |</p>
<table>
<thead>
<tr>
<th>Sample size</th>
<th>Gaussian</th>
<th>Skewed distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>IIS</td>
<td>IIB</td>
<td>IIS</td>
</tr>
<tr>
<td>40</td>
<td>9.5</td>
<td>8.6</td>
</tr>
<tr>
<td>120</td>
<td>5.9</td>
<td>5.4</td>
</tr>
<tr>
<td>250</td>
<td>4.2</td>
<td>3.9</td>
</tr>
<tr>
<td>500</td>
<td>3.0</td>
<td>2.8</td>
</tr>
</tbody>
</table>
| *For the skewed distribution, the SE refers to the upper percentile estimate.