

Supplemental Data

Wolfe P, Murphy J, McGinley J, Zhu Z, Jiang W, Gottschall EB, Thompson HJ. (2004) Using nuclear morphometry to discriminate the tumorigenic potential of cells: a comparison of statistical methods. *Cancer Epidemiol Biomarkers Prev.* Jun;13(6):976-88

Bootstrap Algorithm

When we compare univariate means, we assume $y_{ij} = \mu_i + \varepsilon_{ij}$, where the ε_{ij} are independent, have constant variance, and are normally distributed. In the bootstrap procedure we assume the ε_{ij} come from an unspecified distribution, G , and calculate the multiplicity adjustments from the empirical distribution, \hat{G} , as if it were the true distribution.

1. \hat{G} is estimated by pooling the residuals $\hat{\varepsilon}_{ij} = y_{ij} - \hat{\mu}_i = y_{ij} - \bar{y}_i$.
2. The bootstrap data are generated by resampling with replacement from the pooled residuals:

$$\tilde{y}_{ij} = \hat{\mu}_i + \hat{\varepsilon}_{ij}$$
3. P -values $\tilde{p}_1, \dots, \tilde{p}_k$ are computed from the bootstrap sample under the complete null hypothesis (In our example, $k=39$, the t -tests for each morphometric feature.)
4. Repeat steps 2 and 3 N times. The single-step adjusted p -value for the i th comparison is the proportion of samples in which $\tilde{p}_j \leq p_i$, where p_i is the p -value from the comparison in the original data.

The step-down algorithm is more complex:

1. P -values $\tilde{p}_1, \dots, \tilde{p}_k$ are computed from the bootstrap sample under the complete null hypothesis (as in step 3 above) and assigned the rank of the corresponding p -values in the original data $\tilde{p}_{r_1}, \dots, \tilde{p}_{r_k}$. The sequence $\{Westfall, 1993 611 /id\}$ is fixed throughout the simulation.
2. Define the successive minima:

$$\begin{aligned} \tilde{q}_k &= \tilde{p}_{r_k} \\ \tilde{q}_{k-1} &= \min(\tilde{q}_k, \tilde{p}_{r_{k-1}}) \\ \tilde{q}_{k-2} &= \min(\tilde{q}_{k-1}, \tilde{p}_{r_{k-2}}) \\ &\vdots \\ \tilde{q}_1 &= \min(\tilde{q}_2, \tilde{p}_{r_1}) \end{aligned}$$
3. Repeat steps 1 and 2 N times. $\tilde{p}_{(i)}^{(N)}$ = the proportion of samples where $\tilde{q}_i \leq p_{(i)}$
4. Monotonicity is enforced by successive maximization:

$$\begin{aligned} \tilde{p}_{(1)}^{(N)} &= \tilde{p}_{(1)}^{(N)} \\ \tilde{p}_{(2)}^{(N)} &= \max(\tilde{p}_{(1)}^{(N)}, \tilde{p}_{(2)}^{(N)}) \\ &\vdots \\ \tilde{p}_{(k)}^{(N)} &= \max(\tilde{p}_{(k-1)}^{(N)}, \tilde{p}_{(k)}^{(N)}) \end{aligned}$$

The estimates $\tilde{p}_{(j)}^{(N)}$ are reasonable approximations for $\tilde{p}_{(j)}$ for N sufficiently large. Westfall and Young recommend $N \geq 10000$.

Westfall PH, Young SS. Resampling-based multiple testing. New York: John Wiley & Sons, Inc.; 1993.