Tim B.

1 Standard approach

This document explores a simple strategy for improving the robustness of randomization tests. Let $T_1, \ldots, T_{B-1}$ be the $B-1$ resampled test statistics. Let T^* be the statistic computed on the raw data. Let $\text{rank}(T^*; B)$ be the rank of T^* in the set $\{T^*, T_1, \ldots, T_{B-1}\}$, i.e.

$$\text{rank}(T^*; B) = |\{T \in \{T^*, T_1, \ldots, T_{B-1}\} : T^* \geq T| = 1 + \sum_{i=1}^{B-1} \mathbb{I}(T^* \geq T_i).$$

Under the null hypothesis, $(T^*, T_1, \ldots, T_{B-1})$ is exchangeable. Therefore,

$$\text{rank}(T^*; B) = \text{Unif}(\{1, \ldots, B\}).$$

We can construct a valid p-value as follows:

$$p = \text{rank}(T^*; B)/B.$$

Letting $T_{(1)}, \ldots, T_{(B)}$ denote the order statistics of $T_1, \ldots, T_B$, we also can compute the rank of T^* as

$$\text{rank}(T^*; B) = 1 + \sum_{i=1}^{B-1} \mathbb{I}(T^* \geq T_{(i)}),$$

i.e., we can order the T_is before computing the rank.

2 A simple strategy for improving robustness

Denote $R := \text{rank}(T^*; B)$. Let $f : \{1, \ldots, B\} \to \mathbb{R}$ be a function. We consider the transformed random variable $W := f(R)$. Let $\mathcal{X} := \{f(r)\}_{r=1}^{B}$. For $w \in \mathcal{X}$, we have that

$$\mathbb{P}\left(W = w\right) = \mathbb{P}\left(R \in f^{-1}(w)\right) = |f^{-1}(w)|,$$

where $f^{-1}(w) = \{r \in \{1, \ldots, B\} : f(r) = w\}$ is the preimage of w. The latter equality follows because R is uniformly distributed on $\{1, \ldots, B\}$. If f

is injective, then $P(W = w) = 1/B$ for all w, implying that W is uniformly distributed over $\mathcal{X}$. Therefore,

Example. Let $a_1, \ldots, a_{B-1} > 0$ be weights. Denote $a = [a_1, \ldots, a_{B-1}]^T \in \mathbb{R}^{B-1}$. Define the *weighted* rank W as follows:

$$W = 1 + \sum_{i=1}^{B-1} a_i \mathbb{I}\left(T^* \geq T_{(i)}\right).$$

Clearly, W is not (in general) uniformly distributed on $\{1, \ldots, B\}$, and so W/B is not a valid p-value. However, we can compute the distribution of W. Define the function $f : \{1, \ldots, B\} \to \mathbb{R}$ by

$$f(j) = 1 + \sum_{i=1}^{j-1} a_i.$$

Observe that $W = f(R)$, where

$$R = \sum_{i=1}^{B-1} \mathbb{I}(T^* \geq T_{(i)})$$

is the (unweighted) rank of T^*. Hence, R is uniformly distributed over $\{1, \ldots, B\}$, and

$$\mathbb{P}(W = f(j)) = 1/B$$

for $j \in \{1, \ldots, B\}$.