1 A two-step adaptive permutation procedure

We describe an adaptive, two-step permutation testing procedure that we call the "checkpoint permutation test." The checkpoint permutation test is not new. However, we formalize the procedure, demonstrate its correctness, and precisely characterize its power-computation trade-off, which enables the principled selection of method hyperparameters in practice.

Algorithm 1: Left-tailed checkpoint permutation test

input : Number of null test statistics B_1 and B_2 to draw in the first and second rounds, respectively; threshold c.

output: Normalized rank p of the ground truth statistic among the null statistics.

Compute ground truth test statistic t and null statistics $y_1, \ldots, y_{B_1}$;
$\mathsf{r}_1 \leftarrow \text{rank}(t; \{y_1, \ldots, y_{B_1}\})$;
if $\mathsf{r}_1 \leq c$ **then**
    Compute fresh null statistics $z_1, \ldots, z_{B_2}$;
    $\mathsf{r}_2 \leftarrow \text{rank}(t; \{z_1, \ldots, z_{B_1}\})$;
    $\mathsf{p} \leftarrow \mathsf{r}_2/(B_2 + 1)$
else
    $\mathsf{p} \leftarrow \mathsf{r}_1/(B_1 + 1)$
end

We begin with the left-tailed checkpoint permutation test (Algorithm 1). Let $t \in \mathbb{R}$ be the ground truth test statistic. For $B_1 \in \mathbb{N}$, let $\mathcal{Y} = \{y_1, \ldots, y_{B_1}\}$ be the set of null test statistics drawn during the first round. We compute the rank r_1 of t among the y_is:

$$r_1 := \text{rank}(t; \mathcal{Y}) := |\{y \in \mathcal{Y} : y \leq t\}| + 1,$$

i.e., r_1 is the number of elements in the set $\mathcal{Y} \cup \{t\}$ that are less than or equal to t (including t itself). Note that r_1 takes values on $[B_1 + 1]$. Let $c \in [B_1 + 1]$ be a threshold. If $r_1 > c$, then we set the normalized rank p to $p = r_1/(B_1 + 1)$ and stop. Otherwise, we proceed to round two. For $B_2 \in \mathbb{N}$, let $\mathcal{Z} = \{z_1, \ldots, z_{B_2}\}$ be the set of null test statistics drawn during the second round. Let $r_2 := \text{rank}(t; \mathcal{Z})$ be the rank of t among the z_is. We set the normalized rank p to $p = r_2/(B_2 + 1)$ and stop, returning p to the user. Assume that there are no ties among elements within the set $\{t\} \cup \mathcal{Y} \cup \mathcal{Z}$.

(Otherwise, break ties by jittering.) Furthermore, assume that under the null hypothesis, the random variables in $\{t\} \cup \mathcal{Y} \cup \mathcal{Z}$ are exchangeable.

Proposition. The support $\mathcal{S}$ of p is $[B_1+1]/(B_1+1) \cup [B_2+1]/(B_2+1)$, where $0 < i \leq 1$ for all $i \in \mathcal{S}$. The distribution of p is

$$\mathbb{P}(P \qquad \leq \qquad x) \qquad =$$

Lemma. We map the above onto a combinatorial problem. Suppose that we have y_{tot} yellow balls, g_{tot} green balls, and one black ball that we can arrange in any single-file order. Assume that each of the $(y_{\text{tot}} + g_{\text{tot}} + 1)!$ possible arrangements of the balls is equally probable. Let $Y_l \in \{0, \ldots, y_{\text{tot}}\}$ (resp., $G_l \in \{0, \ldots, g_{\text{tot}}\}$) denote the number of yellow (resp., green) balls that falls to the left of the black ball. Our goal is to compute the probability that a given number of green balls falls to the left of the black ball given that a certain number of yellow balls falls to the left of the black ball.

First, we compute the total number of configurations such that $Y_l = y_l$ and $G_l = g_l$, i.e., such that y_l yellow balls and g_l green balls fall to the left of the black ball. There are $\binom{y_{\text{tot}}}{y_l}$ (resp., $\binom{g_{\text{tot}}}{g_l}$) ways to choose the set of yellow (resp., green) balls that falls to the left of the black ball. Among the $y_l + g_l$ balls that fall to the left of the black ball, there are $(y_l + g_l)!$ possible permutations. Likewise, there are $(y_{\text{tot}} - y_l + g_{\text{tot}} - g_l)!$ possible permutations among the balls that fall to the right of the black ball. Therefore, the total number of configurations such that $Y_l = y_l$ and $G_l = g_l$ is

$$\binom{y_{\text{tot}}}{y_l}\binom{g_{\text{tot}}}{g_l}(y_l + g_l)!(y_{\text{tot}} - y_l + g_{\text{tot}} - g_l)!.$$

Given that each arrangement is equally likely, we have that

$$\begin{aligned}\mathbb{P}(G_l = g_l, Y_l = y_l) &= \frac{\binom{y_{\text{tot}}}{y_l}\binom{g_{\text{tot}}}{g_l}(y_l + g_l)!(y_{\text{tot}} - y_l + g_{\text{tot}} - g_l)!}{(y_{\text{tot}} + g_{\text{tot}} + 1)!} \\ &= \binom{y_{\text{tot}}}{y_l}\binom{g_{\text{tot}}}{g_l} B(g_l + y_l + 1, g_{\text{tot}} - g_l + y_{\text{tot}} - y_l + 1),\end{aligned}$$

where B denotes the beta function. Next, for given $\tau \in \{0, \ldots, y_{\text{tot}}\}$, we have by the definition of conditional probability that

$$\mathbb{P}(G_l = g_l | Y_l \leq \tau) = \frac{\mathbb{P}(G_l = g_l, Y_l \leq \tau)}{\mathbb{P}(Y_l \leq y_l)}. \tag{1}$$

We can decompose the numerator of (1) as

$$\mathbb{P}(G_l = g_l, Y_l \leq \tau) = \sum_{y_l=0}^{\tau} \binom{y_{\text{tot}}}{y_l} \binom{g_{\text{tot}}}{g_l} B(g_l + y_l + 1, g_{\text{tot}} - g_l + y_{\text{tot}} - y_l + 1).$$

Meanwhile, because the rank of the black ball among the yellow balls is uniformly distributed marginally,

$$\mathbb{P}(Y_l \leq \tau) = \sum_{y_l=0}^{\tau} \mathbb{P}(Y_l = \tau) = \frac{\tau + 1}{y_{\text{tot}} + 1}.$$

□

Next, in the context of the original problem, the total number of yellow balls y_{tot} maps to the number of statistics resampled in the first round, B_1; the total number of green balls g_{tot} maps to the number of statistics sampled in the second round, B_2; the number of yellow balls to the left of the black ball maps to the rank of the original statistic among the first-round statistics minus 1, $r_1 = \text{rank}(t; \mathcal{Y}) - 1$; and the total number of green balls to the left of the black ball maps to the rank of the original statistics among the second-round statistics minus 1, $r_2 = \text{rank}(y; \mathcal{Z}) - 1$. Hence, by the derivation above,

$$\mathbb{P}(r_2 = k, r_1 \leq c) = \sum_{i=1}^{1}$$

Algorithm 2: Two-tailed checkpoint permutation test

input : B_1, B_2 and c.
output: Normalized rank p of ground truth statistic among null statistics.

Compute ground truth test statistic t and null statistics $y_1, \ldots, y_{B_1}$;
$\mathsf{rank}_1 \leftarrow \text{rank}(t; \{y_1, \ldots, y_{B_1}\})$;
if $\mathsf{rank}_1 \leq c$ *or* $\mathsf{rank}_1 \geq B_1 + 1 - c$ **then**
    Compute fresh null statistics $z_1, \ldots, z_{B_2}$;
    $\mathsf{rank}_2 \leftarrow \text{rank}(t; \{z_1, \ldots, z_{B_1}\})$;
    if $\mathsf{rank}_1 \leq c$ **then**
        `// left-tailed test in round 2`
        $\mathsf{p} \leftarrow \mathsf{rank}_2/(B_2 + 1)$;
    else
        `//` $\mathsf{rank}_1 \geq B_1 + 1 - c$`; right-tailed test in round 2`
        $\mathsf{p} \leftarrow (B_2 + 1 - \mathsf{rank}_2)/(B_2 + 1)$;
    end
else
    `//` $c < \mathsf{rank}_1 < B_1 + 1 - c$`; two-tailed test in round 1`
    $\mathsf{p} \leftarrow 2 \min\{\mathsf{rank}_1, B_1 + 1 - \mathsf{rank}_1\}/(B_1 + 1)$;
end