Statistica Sinica **18**(2008): Supplement, S??∼S??

HYBRID PERMUTATION TEST WITH APPLICATION TO SURFACE SHAPE ANALYSIS

CHUNXIAO ZHOU AND YONGMEI MICHELLE WANG

University of Illinois at Urbana-Champaign

Supplementary Material

Appendix 1.

Derivation of the Third and Fourth Moments for Mean Difference Test Statistics

The third moment: $U = U^{(1,1,1)} \bigcup U^{(1,2)} \bigcup U^{(3)}$

$$\begin{aligned} a_{(1,1,1)} &= \left(\binom{n_1}{3}\frac{1}{n_1^3} - \binom{n_2}{3}\frac{1}{n_2^3} - \binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^2 n_2} + \binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1 n_2^2}\right)\Big/\binom{n_1+n_2}{3} \\ &= \frac{2(n_2-n_1)}{(n_1^2 n_2^2 (n_1+n_2-1)(n_1+n_2-2))}; \end{aligned}$$

$$\begin{aligned} a_{(1,2)} &= \left(\binom{n_1}{2}\frac{1}{n_1^3} - \binom{n_2}{2}\frac{1}{n_2^3} - \binom{n_1}{1}\binom{n_2}{1}\frac{1/2}{n_1^2 n_2} + \binom{n_1}{1}\binom{n_2}{1}\frac{1/2}{n_1 n_2^2}\right)\Big/\binom{n_1+n_2}{2} \\ &= \frac{(n_1-n_2)}{(n_1^2 n_2^2 (n_1+n_2-1))}; \end{aligned}$$

$$a_{(3)} = \left(\binom{n_1}{1}\frac{1}{n_1^3} - \binom{n_2}{1}\frac{1}{n_2^3}\right)\Big/\binom{n_1+n_2}{3} = \frac{(n_1-n_2)}{(n_1^2 n_2^2)};$$

The fourth moment: $U = U^{(1,1,1,1)} \bigcup U^{(1,1,2)} \bigcup U^{(1,3)} \bigcup U^{(2,2)} \bigcup U^{(4)}$

$$\begin{aligned} a_{(1,1,1,1)} = \Bigg(&\binom{n_1}{4}\frac{1}{n_1^4} - \binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^3 n_2} + \binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^2 n_2^2} \\ &-\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1 n_2^3} + \binom{n_2}{4}\frac{1}{n_2^4}\Bigg)\Big/\binom{n_1+n_2}{4}; \end{aligned}$$

$$\begin{aligned} a_{(1,1,2)} = \Bigg(&\binom{n_1}{3}\frac{1}{n_1^4} - \frac{2}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^3 n_2} + \frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^2 n_2^2} \\ &+\frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^2 n_2^2} - \frac{2}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1 n_2^3} + \binom{n_2}{3}\frac{1}{n_2^4}\Bigg)\Big/\binom{n_1+n_2}{3}; \end{aligned}$$

$$a_{(1,3)} = \left(\binom{n_1}{2}\frac{1}{n_1^4} - \frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^3 n_2} - \frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1 n_2^3} + \binom{n_2}{2}\frac{1}{n_2^4}\right)\Big/\binom{n_1+n_2}{2};$$

$$a_{(2,2)} = \left(\binom{n_1}{2}\frac{1}{n_1^4} + \binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^2 n_2^2} + \binom{n_2}{2}\frac{1}{n_2^4}\right) \Big/ \binom{n_1+n_2}{2};$$

$$a_{(4)} = \left(\binom{n_1}{1}\frac{1}{n_1^4} + \binom{n_2}{1}\frac{1}{n_2^4}\right) \Big/ \binom{n_1+n_2}{1};$$

Appendix 2.

Derivation of the Third and Fourth Moments for Modified Hotelling's T^2 Test Statistics

The third moment: $U = U^{(1,1,1,1,1,1)} \bigcup U^{(1,1,1,1,2)} \bigcup U^{(1,1,1,3)} \bigcup U^{(1,1,2,2)} \bigcup U^{(1,1,4)} \bigcup U^{(1,2,3)} \bigcup U^{(2,2,2)} \bigcup U^{(1,5)} \bigcup U^{(2,4)} \bigcup U^{(3,3)} \bigcup U^{(6)}$;

$$\begin{aligned} a_{(1,1,1,1,1,1)} = \Bigg(& \binom{n_1}{6}\frac{1}{n_1^6} - \binom{n_1}{5}\binom{n_2}{1}\frac{1}{n_1^5 n_2^1} + \binom{n_1}{4}\binom{n_2}{2}\frac{1}{n_1^4 n_2^2} - \binom{n_1}{3}\binom{n_2}{3}\frac{1}{n_1^3 n_2^3} \\ & + \binom{n_1}{2}\binom{n_2}{4}\frac{1}{n_1^2 n_2^4} - \binom{n_1}{1}\binom{n_2}{5}\frac{1}{n_1^1 n_2^5} + \binom{n_2}{6}\frac{1}{n_2^6}\Bigg) \Big/ \binom{n_1+n_2}{6}; \end{aligned}$$

$$\begin{aligned} a_{(1,1,1,1,2)} = \Bigg(& \binom{n_1}{5}\frac{1}{n_1^6} - \frac{4}{5}\binom{n_1}{4}\binom{n_2}{1}\frac{1}{n_1^5 n_2^1} + \frac{1}{5}\binom{n_1}{4}\binom{n_2}{1}\frac{1}{n_1^4 n_2^2} + \frac{3}{5}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^4 n_2^2} \\ & - \frac{2}{5}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^3 n_2^3} - \frac{2}{5}\binom{n_1}{2}\binom{n_2}{3}\frac{1}{n_1^3 n_2^3} + \frac{3}{5}\binom{n_1}{2}\binom{n_2}{3}\frac{1}{n_1^2 n_2^4} \\ & + \frac{1}{5}\binom{n_1}{1}\binom{n_2}{4}\frac{1}{n_1^2 n_2^4} - \frac{4}{5}\binom{n_1}{1}\binom{n_2}{4}\frac{1}{n_1^1 n_2^5} + \binom{n_2}{5}\frac{1}{n_2^6}\Bigg) \Big/ \binom{n_1+n_2}{5}; \end{aligned}$$

$$\begin{aligned} a_{(1,1,1,3)} = \Bigg(& \binom{n_1}{4}\frac{1}{n_1^6} - \frac{3}{4}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^5 n_2^1} - \frac{1}{4}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^3 n_2^3} + \frac{1}{2}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^4 n_2^2} \\ & + \frac{1}{2}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^2 n_2^4} - \frac{1}{4}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^3 n_2^3} - \frac{3}{4}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^1 n_2^5} \\ & + \binom{n_2}{4}\frac{1}{n_2^6}\Bigg) \Big/ \binom{n_1+n_2}{4}; \end{aligned}$$

$$\begin{aligned} a_{(1,1,2,2)} = \Bigg(& \binom{n_1}{4}\frac{1}{n_1^6} - \frac{1}{2}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^5 n_2^1} + \frac{1}{2}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^4 n_2^2} + \frac{1}{6}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^4 n_2^2} \\ & - \frac{2}{3}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^3 n_2^3} + \frac{1}{6}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^2 n_2^4} + \frac{1}{2}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^2 n_2^4} \\ & - \frac{1}{2}\binom{n_1}{1}\binom{n_3}{3}\frac{1}{n_1^1 n_2^5} + \binom{n_2}{4}\frac{1}{n_2^6}\Bigg) \Big/ \binom{n_1+n_2}{4}; \end{aligned}$$

$$\begin{aligned} a_{(1,1,4)} = \Bigg(& \binom{n_1}{3}\frac{1}{n_1^6} - \frac{2}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^5 n_2^1} + \frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^2 n_2^4} \\ & + \frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^4 n_2^2} - \frac{2}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^1 n_2^5} + \binom{n_2}{3}\frac{1}{n_2^6}\Bigg) \Big/ \binom{n_1+n_2}{3}; \end{aligned}$$

$$
\begin{aligned}
a_{(1,2,3)} = \Bigg(&\binom{n_1}{3}\frac{1}{n_1^6} - \frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^5 n_2^1} + \frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^4 n_2^2} - \frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^3 n_2^3} \\
&- \frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^3 n_2^3} + \frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^2 n_2^4} - \frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^1 n_2^5} \\
&+ \binom{n_2}{3}\frac{1}{n_2^6}\Bigg) \Bigg/ \binom{n_1+n_2}{3};
\end{aligned}
$$

$$
a_{(2,2,2)} = \Bigg(\binom{n_1}{3}\frac{1}{n_1^6} - \binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^4 n_2^2} + \binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^2 n_2^4} + \binom{n_2}{3}\frac{1}{n_2^6}\Bigg) \Bigg/ \binom{n_1+n_2}{3};
$$

$$
a_{(1,5)} = \Bigg(\binom{n_1}{2}\frac{1}{n_1^6} - \frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^5 n_2^1} - \frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^1 n_2^5} + \binom{n_2}{2}\frac{1}{n_2^6}\Bigg) \Bigg/ \binom{n_1+n_2}{2};
$$

$$
a_{(2,4)} = \Bigg(\binom{n_1}{2}\frac{1}{n_1^6} + \frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^4 n_2^2} + \frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^2 n_2^4} + \binom{n_2}{2}\frac{1}{n_2^6}\Bigg) \Bigg/ \binom{n_1+n_2}{2};
$$

$$
a_{(3,3)} = \Bigg(\binom{n_1}{2}\frac{1}{n_1^6} - \binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^3 n_2^3} + \binom{n_2}{2}\frac{1}{n_2^6}\Bigg) \Bigg/ \binom{n_1+n_2}{2};
$$

$$
a_{(6)} = \Bigg(\binom{n_1}{1}\frac{1}{n_1^6} + \binom{n_2}{2}\frac{1}{n_2^6}\Bigg) \Bigg/ \binom{n_1+n_2}{1};
$$

The fourth moment: $U = U^{(1,1,1,1,1,1,1,1)} \bigcup U^{(1,1,1,1,1,1,2)} \bigcup U^{(1,1,1,1,1,3)} \bigcup U^{(1,1,1,1,2,2)} \bigcup U^{(1,1,1,1,4)} \bigcup U^{(1,1,1,2,3)} \bigcup U^{(1,1,2,2,2)} \bigcup U^{(2,2,2,2)} \bigcup U^{(1,1,1,5)} \bigcup U^{(1,1,2,4)} \bigcup U^{(1,1,3,3)} \bigcup U^{(1,2,2,3)} \bigcup U^{(1,1,6)} \bigcup U^{(1,2,5)} \bigcup U^{(1,3,4)} \bigcup U^{(2,2,4)} \bigcup U^{(2,3,3)} \bigcup U^{(1,7)} \bigcup U^{(2,6)} \bigcup U^{(3,5)} \bigcup U^{(4,4)} \bigcup U^{(8)}$

$$
\begin{aligned}
a_{(1,1,1,1,1,1,1,1)} = \Bigg(&\binom{n_1}{8}\frac{1}{n_1^8} - \binom{n_1}{7}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} + \binom{n_1}{6}\binom{n_2}{2}\frac{1}{n_1^6 n_2^2} - \binom{n_1}{5}\binom{n_2}{3}\frac{1}{n_1^5 n_2^3} \\
&+ \binom{n_1}{4}\binom{n_2}{4}\frac{1}{n_1^4 n_2^4} - \binom{n_1}{3}\binom{n_2}{5}\frac{1}{n_1^3 n_2^5} + \binom{n_1}{2}\binom{n_2}{6}\frac{1}{n_1^2 n_2^6} \\
&- \binom{n_1}{1}\binom{n_2}{7}\frac{1}{n_1^1 n_2^7} + \binom{n_2}{8}\frac{1}{n_2^8}\Bigg) \Bigg/ \binom{n_1+n_2}{8};
\end{aligned}
$$

$$
\begin{aligned}
a_{(1,1,1,1,1,1,2)} = \Bigg(&\binom{n_1}{7}\frac{1}{n_1^8} - \frac{6}{7}\binom{n_1}{6}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} + \frac{1}{7}\binom{n_1}{6}\binom{n_2}{1}\frac{1}{n_1^6 n_2^2} \\
&+ \frac{5}{7}\binom{n_1}{5}\binom{n_2}{2}\frac{1}{n_1^6 n_2^2} - \frac{2}{7}\binom{n_1}{5}\binom{n_2}{2}\frac{1}{n_1^5 n_2^3} - \frac{4}{7}\binom{n_1}{4}\binom{n_2}{3}\frac{1}{n_1^5 n_2^3} \\
&+ \frac{3}{7}\binom{n_1}{4}\binom{n_2}{3}\frac{1}{n_1^4 n_2^4} + \frac{3}{7}\binom{n_1}{3}\binom{n_2}{4}\frac{1}{n_1^4 n_2^4} - \frac{4}{7}\binom{n_1}{3}\binom{n_2}{4}\frac{1}{n_1^3 n_2^5} \\
&- \frac{2}{7}\binom{n_1}{2}\binom{n_2}{5}\frac{1}{n_1^3 n_2^5} + \frac{5}{7}\binom{n_1}{2}\binom{n_2}{5}\frac{1}{n_1^2 n_2^6} + \frac{1}{7}\binom{n_1}{1}\binom{n_2}{6}\frac{1}{n_1^2 n_2^6} \\
&- \frac{6}{7}\binom{n_1}{1}\binom{n_2}{6}\frac{1}{n_1^1 n_2^7} + \binom{n_2}{7}\frac{1}{n_2^8}\Bigg) \Bigg/ \binom{n_1+n_2}{7};
\end{aligned}
$$

$$a_{(1,1,1,1,1,3)} = \Big(\binom{n_1}{6}\frac{1}{n_1^8} - \frac{5}{6}\binom{n_1}{5}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} - \frac{1}{6}\binom{n_1}{5}\binom{n_2}{1}\frac{1}{n_1^5 n_2^3} + \frac{4}{6}\binom{n_1}{4}\binom{n_2}{2}\frac{1}{n_1^6 n_2^2}$$
$$+\frac{2}{6}\binom{n_1}{4}\binom{n_2}{2}\frac{1}{n_1^4 n_2^4} - \frac{3}{6}\binom{n_1}{3}\binom{n_2}{3}\frac{1}{n_1^5 n_2^3} - \frac{3}{6}\binom{n_1}{3}\binom{n_2}{3}\frac{1}{n_1^3 n_2^5}$$
$$+\frac{2}{6}\binom{n_1}{2}\binom{n_2}{4}\frac{1}{n_1^4 n_2^4} + \frac{4}{6}\binom{n_1}{2}\binom{n_2}{4}\frac{1}{n_1^2 n_2^6} - \frac{1}{6}\binom{n_1}{1}\binom{n_2}{5}\frac{1}{n_1^3 n_2^5}$$
$$-\frac{5}{6}\binom{n_1}{1}\binom{n_2}{5}\frac{1}{n_1^1 n_2^7} + \binom{n_2}{6}\frac{1}{n_2^8}\Big)\Big/\binom{n_1+n_2}{6};$$

$$a_{(1,1,1,1,2,2)} = \Big(\binom{n_1}{6}\frac{1}{n_1^8} - \frac{4}{6}\binom{n_1}{5}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} + \frac{2}{6}\binom{n_1}{5}\binom{n_2}{1}\frac{1}{n_1^6 n_2^2}$$
$$+\frac{6}{15}\binom{n_1}{4}\binom{n_2}{2}\frac{1}{n_1^6 n_2^2} - \frac{8}{15}\binom{n_1}{4}\binom{n_2}{2}\frac{1}{n_1^5 n_3^3} + \frac{1}{15}\binom{n_1}{4}\binom{n_2}{2}\frac{1}{n_1^4 n_2^4}$$
$$-\frac{4}{20}\binom{n_1}{3}\binom{n_2}{3}\frac{1}{n_1^5 n_2^3} + \frac{12}{20}\binom{n_1}{3}\binom{n_2}{3}\frac{1}{n_1^4 n_2^4} - \frac{4}{20}\binom{n_1}{3}\binom{n_2}{3}\frac{1}{n_1^3 n_2^5}$$
$$+\frac{1}{15}\binom{n_1}{2}\binom{n_2}{4}\frac{1}{n_1^4 n_2^4} - \frac{8}{15}\binom{n_1}{2}\binom{n_2}{4}\frac{1}{n_1^3 n_2^5} + \frac{6}{15}\binom{n_1}{2}\binom{n_2}{4}\frac{1}{n_1^2 n_2^6}$$
$$+\frac{2}{6}\binom{n_1}{1}\binom{n_2}{5}\frac{1}{n_1^2 n_2^6} - \frac{4}{6}\binom{n_1}{1}\binom{n_2}{5}\frac{1}{n_1^1 n_2^7} + \binom{n_2}{6}\frac{1}{n_2^8}\Big)\Big/\binom{n_1+n_2}{6};$$

$$a_{(1,1,1,1,4)} = \Big(\binom{n_1}{5}\frac{1}{n_1^8} - \frac{4}{5}\binom{n_1}{4}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} + \frac{1}{5}\binom{n_1}{4}\binom{n_2}{1}\frac{1}{n_1^4 n_2^4}$$
$$+\frac{3}{5}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^6 n_2^2} - \frac{2}{5}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^3 n_2^5} - \frac{2}{5}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^5 n_2^3}$$
$$+\frac{3}{5}\binom{n_1}{2}\binom{n_2}{3}\frac{1}{n_1^2 n_2^6} + \frac{1}{5}\binom{n_1}{1}\binom{n_2}{4}\frac{1}{n_1^4 n_2^4} - \frac{4}{5}\binom{n_1}{1}\binom{n_2}{4}\frac{1}{n_1^1 n_2^7}$$
$$+\binom{n_2}{5}\frac{1}{n_2^8}\Big)\Big/\binom{n_1+n_2}{5};$$

$$a_{(1,1,1,2,3)} = \Big(\binom{n_1}{5}\frac{1}{n_1^8} - \frac{3}{5}\binom{n_1}{4}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} + \frac{1}{5}\binom{n_1}{4}\binom{n_2}{1}\frac{1}{n_1^6 n_2^2}$$
$$-\frac{1}{5}\binom{n_1}{4}\binom{n_2}{1}\frac{1}{n_1^5 n_2^3} + \frac{3}{10}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^6 n_2^2} - \frac{3}{10}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^5 n_3^3}$$
$$+\frac{3}{10}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^4 n_2^4} - \frac{1}{10}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^3 n_2^5} - \frac{1}{10}\binom{n_1}{2}\binom{n_2}{3}\frac{1}{n_1^5 n_2^3}$$
$$+\frac{3}{10}\binom{n_1}{2}\binom{n_2}{3}\frac{1}{n_1^4 n_2^4} - \frac{3}{10}\binom{n_1}{2}\binom{n_2}{3}\frac{1}{n_1^3 n_2^5} + \frac{3}{10}\binom{n_1}{2}\binom{n_2}{3}\frac{1}{n_1^2 n_2^6}$$
$$-\frac{1}{5}\binom{n_1}{1}\binom{n_2}{4}\frac{1}{n_1^3 n_2^5} - \frac{1}{5}\binom{n_1}{1}\binom{n_2}{4}\frac{1}{n_1^3 n_2^5} + \frac{1}{5}\binom{n_1}{1}\binom{n_2}{4}\frac{1}{n_1^2 n_2^6}$$
$$-\frac{3}{5}\binom{n_1}{1}\binom{n_2}{4}\frac{1}{n_1^1 n_2^7} + \binom{n_2}{5}\frac{1}{n_2^8}\Big)\Big/\binom{n_1+n_2}{5};$$

$$
\begin{aligned}
a_{(1,1,2,2,2)} = &\Bigg(\binom{n_1}{5}\frac{1}{n_1^8} - \frac{2}{5}\binom{n_1}{4}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} + \frac{3}{5}\binom{n_1}{4}\binom{n_2}{1}\frac{1}{n_1^6 n_2^2} \\
&+\frac{1}{10}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^6 n_2^2} - \frac{6}{10}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^5 n_2^3} + \frac{3}{10}\binom{n_1}{3}\binom{n_2}{2}\frac{1}{n_1^4 n_2^4} \\
&+\frac{3}{10}\binom{n_1}{2}\binom{n_2}{3}\frac{1}{n_1^4 n_2^4} - \frac{6}{10}\binom{n_1}{2}\binom{n_2}{3}\frac{1}{n_1^3 n_2^5} + \frac{1}{10}\binom{n_1}{2}\binom{n_2}{3}\frac{1}{n_1^2 n_2^6} \\
&+\frac{3}{5}\binom{n_1}{1}\binom{n_2}{4}\frac{1}{n_1^2 n_2^6} - \frac{2}{5}\binom{n_1}{1}\binom{n_2}{4}\frac{1}{n_1^1 n_2^7} + \binom{n_2}{5}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{5};
\end{aligned}
$$

$$
\begin{aligned}
a_{(2,2,2,2)} = &\Bigg(\binom{n_1}{4}\frac{1}{n_1^8} + \binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^6 n_2^2} + \binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^4 n_2^4} \\
&+\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^2 n_2^6} + \binom{n_2}{4}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{4};
\end{aligned}
$$

$$
\begin{aligned}
a_{(1,1,1,5)} = &\Bigg(\binom{n_1}{4}\frac{1}{n_1^8} - \frac{3}{4}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} - \frac{1}{4}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^3 n_2^5} \\
&+\frac{1}{2}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^6 n_2^2} + \frac{1}{2}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^2 n_2^6} - \frac{1}{4}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^5 n_2^3} \\
&-\frac{3}{4}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^1 n_2^7} + \binom{n_2}{4}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{4};
\end{aligned}
$$

$$
\begin{aligned}
a_{(1,1,2,4)} = &\Bigg(\binom{n_1}{4}\frac{1}{n_1^8} - \frac{1}{2}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} + \frac{1}{4}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^6 n_2^2} \\
&+\frac{1}{4}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^4 n_2^4} + \frac{1}{6}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^6 n_2^2} - \frac{1}{3}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^5 n_2^3} \\
&-\frac{1}{3}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^3 n_2^5} + \frac{1}{6}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^2 n_2^6} + \frac{1}{4}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^4 n_2^4} \\
&+\frac{1}{4}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^2 n_2^6} - \frac{1}{2}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^1 n_2^7} + \binom{n_2}{4}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{4};
\end{aligned}
$$

$$
\begin{aligned}
a_{(1,1,3,3)} = &\Bigg(\binom{n_1}{4}\frac{1}{n_1^8} - \frac{1}{2}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} - \frac{1}{2}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^5 n_2^3} + \frac{1}{6}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^6 n_3^2} \\
&+\frac{2}{3}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^4 n_3^4} + \frac{1}{6}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^2 n_2^6} - \frac{1}{2}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^3 n_2^5} \\
&-\frac{1}{2}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^1 n_2^7} + \binom{n_2}{4}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{4};
\end{aligned}
$$

$$
\begin{aligned}
a_{(1,2,2,3)} = &\Bigg(\binom{n_1}{4}\frac{1}{n_1^8} - \frac{1}{4}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^7 n_2^1} + \frac{1}{2}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^6 n_2^2} \\
&-\frac{1}{4}\binom{n_1}{3}\binom{n_2}{1}\frac{1}{n_1^5 n_2^3} - \frac{1}{3}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^5 n_3^3} + \frac{1}{3}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^4 n_3^4}
\end{aligned}
$$

$$-\frac{1}{3}\binom{n_1}{2}\binom{n_2}{2}\frac{1}{n_1^3n_2^5}-\frac{1}{4}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^3n_2^5}+\frac{1}{2}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^2n_2^6}$$
$$-\frac{1}{4}\binom{n_1}{1}\binom{n_2}{3}\frac{1}{n_1^1n_2^7}+\binom{n_2}{4}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{4};$$

$$a_{(1,1,6)}=\Bigg(\binom{n_1}{3}\frac{1}{n_1^8}-\frac{2}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^7n_2^1}+\frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^2n_2^6}+\frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^6n_2^2}$$
$$-\frac{2}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^1n_2^7}+\binom{n_2}{3}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{3};$$

$$a_{(1,2,5)}=\Bigg(\binom{n_1}{3}\frac{1}{n_1^8}-\frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^7n_2^1}+\frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^6n_2^2}-\frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^3n_2^5}$$
$$-\frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^5n_2^3}+\frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^2n_2^6}-\frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^1n_3^7}$$
$$+\binom{n_2}{3}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{3};$$

$$a_{(1,3,4)}=\Bigg(\binom{n_1}{3}\frac{1}{n_1^8}-\frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^7n_2^1}-\frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^5n_2^3}+\frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^4n_2^4}$$
$$+\frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^4n_2^4}-\frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^3n_2^5}-\frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^1n_3^7}$$
$$+\binom{n_2}{3}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{3};$$

$$a_{(2,2,4)}=\Bigg(\binom{n_1}{3}\frac{1}{n_1^8}+\frac{2}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^6n_2^2}+\frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^4n_2^4}+\frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^4n_2^4}$$
$$+\frac{2}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^2n_3^6}+\binom{n_2}{2}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{3};$$

$$a_{(2,3,3)}=\Bigg(\binom{n_1}{3}\frac{1}{n_1^8}+\frac{1}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^6n_2^2}-\frac{2}{3}\binom{n_1}{2}\binom{n_2}{1}\frac{1}{n_1^5n_2^3}-\frac{2}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^3n_2^5}$$
$$+\frac{1}{3}\binom{n_1}{1}\binom{n_2}{2}\frac{1}{n_1^2n_2^6}+\binom{n_2}{2}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{3};$$

$$a_{(1,7)}=\Bigg(\binom{n_1}{2}\frac{1}{n_1^8}-\frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^7n_2^1}-\frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^1n_2^7}+\binom{n_2}{2}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{2};$$

$$a_{(2,6)}=\Bigg(\binom{n_1}{2}\frac{1}{n_1^8}+\frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^6n_2^2}+\frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^2n_2^6}+\binom{n_2}{2}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{2};$$

$$a_{(3,5)}=\Bigg(\binom{n_1}{2}\frac{1}{n_1^8}-\frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^5n_2^3}-\frac{1}{2}\binom{n_1}{1}\binom{n_2}{1}\frac{1}{n_1^3n_2^5}+\binom{n_2}{2}\frac{1}{n_2^8}\Bigg)\Bigg/\binom{n_1+n_2}{2};$$

$$a_{(4,4)} = \left(\binom{n_1}{2} \frac{1}{n_1^8} + \binom{n_1}{1} \binom{n_2}{1} \frac{1}{n_1^4 n_2^4} + \binom{n_2}{2} \frac{1}{n_2^8} \right) \Big/ \binom{n_1+n_2}{2};$$

$$a_{(8)} = \left(\binom{n_1}{2} \frac{1}{n_1^8} + \binom{n_2}{1} \frac{1}{n_2^8} \right) \Big/ \binom{n_1+n_2}{2};$$

Appendix 3.

Matlab Code (Simulated Data for Mean Difference Test Statistics)

Usage: >> ComparePerm_MeanDiff

```
%%%%%%%%%%%%%ComparePerm_MeanDiff.m%%%%%%%%%%%

% Compare the efficiency and accuracy of our hybrid pemutation
% with random permutation and exact permutation. We generate six different
% simulated data sets to evaluate our hybrid permutation approach.

% Test Statistic: mean difference

% Author: Chunxiao Zhou
% Email: czhou4@uiuc.edu
% Revision Date: 2007/12/10

%   References:
%       Chunxiao Zhou and Yongmei Michelle Wang (2008)
%       Hybrid Permutation Test With Application To Surface Shape Analysis,
%       Statistica Sinica.

% set number of random permutation
num_rp = 20000;

%------- Case 1: Normal Distribution (balanced) --------------------
% initialize sample size
n1 = 10; n2 = 10; n = n1+n2;
% initialize moments coefficient
[Coef_3rd, Coef_4th] = MomentsCoef_MeanDiff(n, n1);
% generate random two-group samples
Data_A1 = normrnd(0,1,1,n1); Data_B1 = normrnd(1,0.5,1,n2); Data1 =
[Data_A1 Data_B1];
% exact permutaion
[Pval_ep1, t_ep1, mean_diff1] = ExactPerm_MeanDiff(Data1, n, n1);
% random permutation
[Pval_rp1, t_rp1] = RandPerm_MeanDiff(Data1, n, n1, num_rp);
% hybrid permutation
[Pval_hp1, fitcdf1, t_hp1]=HybridPerm_MeanDiff(Data1, n,
```

```
n1,Coef_3rd,Coef_4th,mean_diff1);
% display comparison results;
disp('For case 1, the p-value estimations');
disp(sprintf('exact permutation  Pval_ep1=%f',Pval_ep1));
disp(sprintf('random permutation Pval_rp1=%f',Pval_rp1));
disp(sprintf('hybrid permutation Pval_hp1=%f',Pval_hp1));
disp('For case 1, the computation costs (in seconds)');
disp(sprintf('exact permutation  t_ep1=%f',t_ep1));
disp(sprintf('random permutation t_rp1=%f',t_rp1));
disp(sprintf('hybrid permutation t_hp1=%f',t_hp1));
disp(' ');
% display fitted Pearson distribution and exact permutation distribution
[Stat1, Ind1] = sort(mean_diff1); x1 = Stat1; y1 =fitcdf1(Ind1);
figure, mycdfplot(mean_diff1);hold on;
plot(Stat1,fitcdf1(Ind1),'--b','LineWidth',3);
legend('true','fitted',2); hold off;

%------- Case 2: Normal Distribution (unbalanced) ------------------
% initialize sample size
n1 = 6; n2 = 18; n = n1+n2;
% initialize moments coefficient
[Coef_3rd, Coef_4th] = MomentsCoef_MeanDiff(n, n1);
% generate random two-group samples
Data_A2 = normrnd(0,1,1,n1); Data_B2 = normrnd(1,0.5,1,n2); Data2 =
[Data_A2 Data_B2];
% exact permutaion
[Pval_ep2, t_ep2, mean_diff2] = ExactPerm_MeanDiff(Data2, n, n1);
% random permutation
[Pval_rp2, t_rp2] = RandPerm_MeanDiff(Data2, n, n1, num_rp);
% hybrid permutation
[Pval_hp2, fitcdf2, t_hp2]=HybridPerm_MeanDiff(Data2, n,
n1,Coef_3rd,Coef_4th,mean_diff2);
% display comparison results;
disp('For case 2, the p-value estimations');
disp(sprintf('exact permutation  Pval_ep2=%f',Pval_ep2));
disp(sprintf('random permutation Pval_rp2=%f',Pval_rp2));
disp(sprintf('hybrid permutation Pval_hp2=%f',Pval_hp2));
disp('For case 2, the computation costs (in seconds)');
disp(sprintf('exact permutation  t_ep2=%f',t_ep2));
disp(sprintf('random permutation t_rp2=%f',t_rp2));
disp(sprintf('hybrid permutation t_hp2=%f',t_hp2));
disp(' ');
% display fitted Pearson distribution and exact permutation distribution
[Stat2, Ind2] = sort(mean_diff2); x2 = Stat2; y2 =fitcdf2(Ind2);
figure, mycdfplot(mean_diff2);hold on;
```

```
plot(Stat2,fitcdf2(Ind2),'--b','LineWidth',3);
legend('true','fitted',2); hold off;

%------- Case 3: Gama Distribution (balanced) ----------------------
% initialize sample size
n1 = 10; n2 = 10; n = n1+n2;
% initialize moments coefficient
[Coef_3rd, Coef_4th] = MomentsCoef_MeanDiff(n, n1);
% generate random two-group samples
Data_A3 = gamrnd(3,3,1,n1); Data_B3 = gamrnd(3,2,1,n2); Data3 =
[Data_A3 Data_B3];
% exact permutaion
[Pval_ep3, t_ep3, mean_diff3] = ExactPerm_MeanDiff(Data3, n, n1);
% random permutation
[Pval_rp3, t_rp3] = RandPerm_MeanDiff(Data3, n, n1, num_rp);
% hybrid permutation
[Pval_hp3, fitcdf3, t_hp3]=HybridPerm_MeanDiff(Data3, n,
n1,Coef_3rd,Coef_4th,mean_diff3);
% display comparison results;
disp('For case 3, the p-value estimations');
disp(sprintf('exact permutation  Pval_ep3=%f',Pval_ep3));
disp(sprintf('random permutation Pval_rp3=%f',Pval_rp3));
disp(sprintf('hybrid permutation Pval_hp3=%f',Pval_hp3));
disp('For case 3, the computation costs (in seconds)');
disp(sprintf('exact permutation  t_ep3=%f',t_ep3));
disp(sprintf('random permutation t_rp3=%f',t_rp3));
disp(sprintf('hybrid permutation t_hp3=%f',t_hp3));
disp(' ');
% display fitted Pearson distribution and exact permutation distribution
[Stat3, Ind3] = sort(mean_diff3); x3 = Stat3; y3 =fitcdf3(Ind3);
figure, mycdfplot(mean_diff3);hold on;
plot(Stat3,fitcdf3(Ind3),'--b','LineWidth',3);
legend('true','fitted',2); hold off;

%------- Case 4: Gama Distribution (unbalanced) --------------------
% initialize sample size
n1 = 6; n2 = 18; n = n1+n2;
% initialize moments coefficient
[Coef_3rd, Coef_4th] = MomentsCoef_MeanDiff(n, n1);
% generate random two-group samples
Data_A4 = gamrnd(3,2,1,n1); Data_B4 = gamrnd(3,1,1,n2); Data4 =
[Data_A4 Data_B4];
% exact permutaion
[Pval_ep4, t_ep4, mean_diff4] = ExactPerm_MeanDiff(Data4, n, n1);
% random permutation
```

```
[Pval_rp4, t_rp4] = RandPerm_MeanDiff(Data4, n, n1, num_rp);
% hybrid permutation
[Pval_hp4, fitcdf4, t_hp4]=HybridPerm_MeanDiff(Data4, n,
n1,Coef_3rd,Coef_4th,mean_diff4);
% display comparison results;
disp('For case 4, the p-value estimations');
disp(sprintf('exact permutation  Pval_ep4=%f',Pval_ep4));
disp(sprintf('random permutation Pval_rp4=%f',Pval_rp4));
disp(sprintf('hybrid permutation Pval_hp4=%f',Pval_hp4));
disp('For case 4, the computation costs (in seconds)');
disp(sprintf('exact permutation  t_ep4=%f',t_ep4));
disp(sprintf('random permutation t_rp4=%f',t_rp4));
disp(sprintf('hybrid permutation t_hp4=%f',t_hp4));
disp(' ');
% display fitted Pearson distribution and exact permutation distribution
[Stat4, Ind4] = sort(mean_diff4); x4 = Stat4; y4 =fitcdf4(Ind4);
figure, mycdfplot(mean_diff4);hold on;
plot(Stat4,fitcdf4(Ind4),'--b','LineWidth',3);
legend('true','fitted',2); hold off;

%------- Case 5: Beta Distribution (balanced) ----------------------
% initialize sample size
n1 = 10; n2 = 10; n = n1+n2;
% initialize moments coefficient
[Coef_3rd, Coef_4th] = MomentsCoef_MeanDiff(n, n1);
% generate random two-group samples
Data_A5 = betarnd(0.8,0.8,1,n1); Data_B5 = betarnd(0.1,0.1,1,n2);
Data5 = [Data_A5 Data_B5];
% exact permutaion
[Pval_ep5, t_ep5, mean_diff5] = ExactPerm_MeanDiff(Data5, n, n1);
% random permutation
[Pval_rp5, t_rp5] = RandPerm_MeanDiff(Data5, n, n1, num_rp);
% hybrid permutation
[Pval_hp5, fitcdf5, t_hp5]=HybridPerm_MeanDiff(Data5, n,
n1,Coef_3rd,Coef_4th,mean_diff5);
% display comparison results;
disp('For case 5, the p-value estimations');
disp(sprintf('exact permutation  Pval_ep5=%f',Pval_ep5));
disp(sprintf('random permutation Pval_rp5=%f',Pval_rp5));
disp(sprintf('hybrid permutation Pval_hp5=%f',Pval_hp5));
disp('For case 5, the computation costs (in seconds)');
disp(sprintf('exact permutation  t_ep5=%f',t_ep5));
disp(sprintf('random permutation t_rp5=%f',t_rp5));
disp(sprintf('hybrid permutation t_hp5=%f',t_hp5));
disp(' ');
```

```
% display fitted Pearson distribution and exact permutation distribution
[Stat5, Ind5] = sort(mean_diff5); x5 = Stat5; y5 =fitcdf5(Ind5);
figure, mycdfplot(mean_diff5);hold on;
plot(Stat5,fitcdf5(Ind5),'--b','LineWidth',3);
legend('true','fitted',2); hold off;

%------- Case 6: Beta Distribution (unbalanced) --------------------
% initialize sample size
n1 = 6; n2 = 18; n = n1+n2;
% initialize moments coefficient
[Coef_3rd, Coef_4th] = MomentsCoef_MeanDiff(n, n1);
% generate random two-group samples
Data_A6 = betarnd(0.8,0.8,1,n1); Data_B6 = betarnd(0.1,0.1,1,n2);
Data6 = [Data_A6 Data_B6];
% exact permutaion
[Pval_ep6, t_ep6, mean_diff6] = ExactPerm_MeanDiff(Data6, n, n1);
% random permutation
[Pval_rp6, t_rp6] = RandPerm_MeanDiff(Data6, n, n1, num_rp);
% hybrid permutation
[Pval_hp6, fitcdf6, t_hp6]=HybridPerm_MeanDiff(Data6, n,
n1,Coef_3rd,Coef_4th,mean_diff6);
% display comparison results;
disp('For case 6, the p-value estimations');
disp(sprintf('exact permutation  Pval_ep6=%f',Pval_ep6));
disp(sprintf('random permutation Pval_rp6=%f',Pval_rp6));
disp(sprintf('hybrid permutation Pval_hp6=%f',Pval_hp6));
disp('For case 6, the computation costs (in seconds)');
disp(sprintf('exact permutation  t_ep6=%f',t_ep6));
disp(sprintf('random permutation t_rp6=%f',t_rp6));
disp(sprintf('hybrid permutation t_hp6=%f',t_hp6));
disp(' ');
% display CDFs of fitted Pearson distribution and exact permutation distribution
[Stat6, Ind6] = sort(mean_diff6); x6 = Stat6; y6 =fitcdf6(Ind6);
figure, mycdfplot(mean_diff6);hold on;
plot(Stat6,fitcdf6(Ind6),'--b','LineWidth',3);
legend('true','fitted',2); hold off;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [Coef_3rd, Coef_4th] = MomentsCoef_MeanDiff(n, n1)
% [Coef_3rd, Coef_4th] = MomentsCoef_MeanDiff(n, n1)
%-----------------Initialize moments coefficients for hybrid permutation----
% Input-------
%     n: number of samples for both two groups;
```

```
%     n1: sample size for group A;
% Output-------
%     Coef_3rd: 3rd moment coefficients for hybrid permutation using
%               mean difference test statistics;
%     Coef_4th: 4th moment coefficients for hybrid permutation using
%               mean difference test statistics;
% Author: Chunxiao Zhou Email: czhou4@uiuc.edu

n2 = n-n1;
% Calculate 3rd moment coefficients (it only needs to be done once for repeated
% tests with the same sample size)
a_1_1_1 = 2*(n2-n1)/(n1^2)/(n2^2)/(n1+n2-1)/(n1+n2-2); a_1_2 =
(n1-n2)/(n1^2)/(n2^2)/(n1+n2-1); a_3 = (n2-n1)/(n1^2)/(n2^2);
Coef_3rd = [a_1_1_1, a_1_2, a_3];

%---------------------------------------------------------------------------
% Calculate 4th moment coefficients (it only needs to be done once for repeated
% tests with the same sample size)
a_1_1_1_1 =
nchoosek(n1,4)/(n1^4)-nchoosek(n1,3)*nchoosek(n2,1)/(n1^3)/n2+...
    nchoosek(n1,2)*nchoosek(n2,2)/(n1^2)/(n2^2)-nchoosek(n1,1)*...
    nchoosek(n2,3)/(n1)/(n2^3)+nchoosek(n2,4)/(n2^4);
a_1_1_1_1 = a_1_1_1_1/nchoosek(n,4); a_1_1_2 =
nchoosek(n1,3)*(1/(n1^4))-2/3*nchoosek(n1,2)*...
    nchoosek(n2,1)/(n1^3)/n2+1/3*nchoosek(n1,2)*nchoosek(n2,1)/(n1^2)/(n2^2)...
    +1/3*nchoosek(n1,1)*nchoosek(n2,2)/(n1^2)/(n2^2)-2/3*nchoosek(n1,1)*...
    nchoosek(n2,2)/(n1)/(n2^3)+nchoosek(n2,3)/(n2^4);
a_1_1_2 = a_1_1_2/nchoosek(n,3); a_1_3 =
nchoosek(n1,2)/(n1^4)-1/2*nchoosek(n1,1)*nchoosek(n2,1)/(n1^3)/n2-...
    1/2*nchoosek(n1,1)*nchoosek(n2,1)/(n1)/(n2^3)+nchoosek(n2,2)*(1/(n2^4));
a_1_3 = a_1_3/nchoosek(n,2); a_2_2 =
nchoosek(n1,2)/(n1^4)+nchoosek(n1,1)*nchoosek(n2,1)/(n1^2)/(n2^2)+...
    nchoosek(n2,2)/(n2^4);
a_2_2 = a_2_2/nchoosek(n,2); a_4 =
nchoosek(n1,1)/(n1^4)+nchoosek(n2,1)/(n2^4); a_4 =
a_4/nchoosek(n,1); Coef_4th = [a_1_1_1_1, a_1_1_2, a_1_3, a_2_2,
a_4];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [Pval_ep, t_ep, mean_diff]=ExactPerm_MeanDiff(Data, n, n1)
% [Pval_ep, t_ep, mean_diff]=ExactPerm_MeanDiff(Data, n, n1)
% Exact Permutation for mean difference test statistic
% Input-------
```

```
%     n: number of samples for both two groups;
%     n1: sample size for group A;
%     Data: univariate data sequences; first n1 items belong to group A,
%            the rest belongs to group B;
% Output-------
%     Pval_ep: P value;
%     t_ep: computation time;
%     mean_diff: mean difference test statistics for all permutations.
% Author: Chunxiao Zhou Email: czhou4@uiuc.edu

x = Data; tic; sum_x = sum(x); num_permutation = nchoosek(n,n1);
perm_all = nchoosek([1:n],n1); sum_a = sum(x(perm_all),2); sum_b =
sum_x-sum_a; mean_diff = sum_a/n1-sum_b/(n-n1); [Stat,IndStat] =
sort(mean_diff); I1 = find(IndStat==1); Pval_ep =
I1/num_permutation; t_ep = toc;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [Pval_rp, t_rp]=RandPerm_MeanDiff(Data, n, n1, num_rp)
% [Pval_rp, t_rp]=RandPerm_MeanDiff(Data, n, n1, num_rp)
% Random Permutation for mean difference test statistic
% Input-------
%     n: number of samples for both two groups;
%     n1: sample size for group A;
%     Data: univariate data sequences; first n1 items belong to group A,
%            the rest belongs to group B;
%     num_rp: number of random permutations;
% Output-------
%     Pval_rp: P value;
%     t_rp: computation time;
% Author: Chunxiao Zhou Email: czhou4@uiuc.edu

stat_rp = zeros(1, num_rp); stat_rp(1) =
mean(Data(1:n1))-mean(Data((n1+1):n)); tic; for i = 2:num_rp
    perm_ind = randperm(n);
    stat_rp(i) = mean(Data(perm_ind(1:n1)))-mean(Data(perm_ind(n1+1:n)));
end [stat_rp_all,IndStat] = sort(stat_rp); I2 = find(IndStat==1);
Pval_rp = I2/num_rp; t_rp = toc;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [Pval_hp, fitpdf, t_hp]=HybridPerm_MeanDiff(Data, n,
n1,Coef_3rd,Coef_4th,mean_diff)
```

```
% [Pval_hp, fitpdf, t_hp]=HybridPerm_MeanDiff(Data, n, n1,Coef_3rd,Coef_4th,mean_diff)
% Hybrid Permutation for mean difference test statistic
% Input-------
%     n: number of samples for both two groups;
%     n1: sample size for group A;
%     Data: univariate data sequences; first n1 items belong to group A,
%             the rest belongs to group B;
%     Coef_3rd: coefficient for the third moment calculation;
%     Coef_4th: coefficient for the fourth moment calculation;
%     mean_diff: mean difference test statistics for all permutations.
% Output-------
%     Pval_hp: P value;
%     fitpdf: pdf values of fitted Pearson distribution at values of mean_diff;
%     t_hp: computation time;
% Author: Chunxiao Zhou Email: czhou4@uiuc.edu

x = Data; n2 = n-n1;
% get original mean difference test statistic
initialStat = mean(x(1:n1))-mean(x((n1+1):n));
%---------------------------calculate summation terms---------------------
tic; sum_x = sum(x); sum_x2 = sum(x.^2); sum_x3 = sum(x.^3); sum_x4
= sum(x.^4); sum_x1_1 = sum_x^2-sum_x2; sum_x1_2 =
3*(sum_x2*sum_x-sum_x3); sum_x1_3 = 4*(sum_x*sum_x3-sum_x4);
sum_x2_2 = 3*(sum_x2^2-sum_x4); sum_x1_1_1 =
sum_x1_1*sum_x-sum_x1_2*2/3; sum_x1_1_2 =
6*(sum_x1_1*sum_x2-sum_x1_3/2); sum_x1_1_1_1 =
sum_x^4-sum_x4-sum_x2_2-sum_x1_3-sum_x1_1_2;

Theoretic_1st_moment_x = 0; Theoretic_2nd_moment_x =
(1/n1+1/n2)*(std(x)^2); Theoretic_3rd_moment_x =
Coef_3rd*[sum_x1_1_1,sum_x1_2,sum_x3]'; Theoretic_4th_moment_x =
Coef_4th*[sum_x1_1_1_1,sum_x1_1_2,sum_x1_3,sum_x2_2,sum_x4]';

mu = Theoretic_1st_moment_x; sigma = sqrt(Theoretic_2nd_moment_x);
skew = Theoretic_3rd_moment_x/(sigma^3); kurt =
Theoretic_4th_moment_x/(sigma^4); t_hp = toc;
% Pearson distribution seriers fitting;
[Pval_hp,type,fitpdf,t_fitting] =
pearsfitting(mu,sigma,skew,kurt,initialStat,mean_diff); t_hp =
t_hp+t_fitting;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [pvalue,type,fitcdf,t_fitting] =
```

```
pearsfitting(mu,sigma,skew,kurt,statvalue,allstat)
% [pvalue,type,fitcdf,t_fitting] = pearsfitting(mu,sigma,skew,kurt,statvalue,allstat)
%revision from the matlab function pearsrnd;
%Approximate the empirical distribution by Pearson distribution system.
% Input-------
%     mu: mean of empirical distribution;
%     sigma: standard deviation of empirical distribution;
%     skew: skewness of  empirical distribution;
%     kurt: kurtosis of  empirical distribution;
%     statvalue: a value of distribution variable;
%     allstat: a sequence of values of distribution variable;
% Output-------
%     pvalue: p-value for statvalue;
%     fitpdf: pdf values of allstat;
%     type: type of Pearson distribution;
%     t_fitting: computation cost of p-value for statvalue;

%   The seven types of distribution in the the Pearson series correspond to the
%   following distributions:
%      Type 0: Normal distribution
%      Type 1: Four-parameter beta
%      Type 2: Symmetric four-parameter beta
%      Type 3: Three-parameter gamma
%      Type 4: Not related to any standard distribution.  Density proportional
%              to (1+((x-a)/b)^2)^(-c) * exp(-d*arctan((x-a)/b)).
%      Type 5: Inverse gamma location-scale
%      Type 6: F location-scale
%      Type 7: Student's t location-scale

%   PEARSRND uses transformations of various standard random variates for types
%   0-III and types V-VII, and a rejection algorithm for type IV.

%   References:
%      [1] Johnson, N.L., S. Kotz, and N. Balakrishnan (1994)
%           Continuous Univariate Distributions, Volume 1,  Wiley-Interscience.
%      [2] Devroye, L. (1986)
%           Non-Uniform Random Variate Generation, Springer-Verlag.

tic if nargin < 4
    error('stats:pearsrnd:TooFewInputs','Requires at least four input arguments.');
elseif ~(isscalar(mu) && isscalar(sigma) && isscalar(skew) &&
isscalar(kurt))
    error('stats:pearsrnd:NonScalarInputs','MU, SIGMA, SKEW, and KURT must be scalars.');
end
```

```
outClass = superiorfloat(mu,sigma,skew,kurt);

beta1 = skew.^2; beta2 = kurt;

% Return NaN for illegal parameter values.
if (sigma < 0) || (beta2 <= beta1 + 1)
    pvalue = NaN;
    type = NaN;
    return
end statvalue = (statvalue-mu)./sigma; allstat=(allstat-mu)./sigma;
% Classify the distribution and find the roots of c0 + c1*x + c2*x^2
c0 = (4*beta2 - 3*beta1); % ./ (10*beta2 - 12*beta1 - 18);
c1 = skew .* (beta2 + 3); % ./ (10*beta2 - 12*beta1 - 18);
c2 = (2*beta2 - 3*beta1 - 6); % ./ (10*beta2 - 12*beta1 - 18);
if c1 == 0 % symmetric dist'ns
    if beta2 == 3
        type = 0;
    else
        if beta2 < 3
            type = 2;
        elseif beta2 > 3
            type = 7;
        end
        a1 = -sqrt(abs(c0./c2));
        a2 = -a1; % symmetric roots
    end
elseif c2 == 0 % kurt = 3 + 1.5*skew^2
    type = 3;
    a1 = -c0 ./ c1; % single root
else
    kappa = c1.^2 ./ (4*c0.*c2);
    if kappa < 0
        type = 1;
    elseif kappa < 1-eps
        type = 4;
    elseif kappa <= 1+eps
        type = 5;
    else
        type = 6;
    end
    % Solve the quadratic for general roots a1 and a2 and sort by their real parts
    tmp = -(c1 + sign(c1).*sqrt(c1.^2 - 4*c0.*c2)) ./ 2;
    a1 = tmp ./ c2;
    a2 = c0 ./ tmp;
    if (real(a1) > real(a2)), tmp = a1; a1 = a2; a2 = tmp; end
```

```
end

denom = (10*beta2 - 12*beta1 - 18); if abs(denom) > sqrt(realmin)
    c0 = c0 ./ denom;
    c1 = c1 ./ denom;
    c2 = c2 ./ denom;
    coefs = [c0 c1 c2];
else
    type = 1; % this should have happened already anyway
    % beta2 = 1.8 + 1.2*beta1, and c0, c1, and c2 -> Inf.  But a1 and a2 are
    % still finite.
end

% generate standard (zero mean, unit variance) values
switch type

case 0
    % normal: standard support (-Inf,Inf)
pvalue = normcdf(statvalue,0,1); t_fitting = toc; fitcdf =
normcdf(allstat);

case 1
    % four-parameter beta: standard support (a1,a2)
    if abs(denom) > sqrt(realmin)
        m1 = (c1 + a1) ./ (c2 .* (a2 - a1));
        m2 = -(c1 + a2) ./ (c2 .* (a2 - a1));
    else
        % c1 and c2 -> Inf, but c1/c2 has finite limit
        m1 = c1 ./ (c2 .* (a2 - a1));
        m2 = -c1 ./ (c2 .* (a2 - a1));
    end
    pvalue = betacdf((statvalue-a1)./(a2-a1),m1+1,m2+1);
    t_fitting = toc;
    fitcdf = betacdf((allstat-a1)./(a2-a1),m1+1,m2+1);
case 2
    % symmetric four-parameter beta: standard support (-a1,a1)
    m = (c1 + a1) ./ (c2 .* 2*abs(a1));
    pvalue = betacdf((statvalue-a1)/2./abs(a1),m+1,m+1);
    t_fitting = toc;
    fitcdf = betacdf((allstat-a1)/2./abs(a1),m+1,m+1);
case 3
    % three-parameter gamma: standard support (a1,Inf) or (-Inf,a1)
    m = (c0./c1 - c1) ./ c1;
    pvalue = gamcdf((statvalue-a1)./c1,m+1,1);
    t_fitting = toc;
```

```
    fitcdf = gamcdf((allstat-a1)./c1,m+1,1);
case 4
    % Pearson IV is not a transformation of a standard distribution: density
    % proportional to (1+((x-lambda)/a)^2)^(-m) * exp(-nu*arctan((x-lambda)/a)),
    % standard support (-Inf,Inf)
    m = 1 ./ (2*c2);
    nu = 2.*c1.*(1 - m) ./ sqrt((4.*c0.*c2 - c1.^2));
    b = 2*(m-1);
    a = sqrt(b.^2 .* (b-1) ./ (b.^2 + nu.^2)); % gives unit variance
    lambda = a.*nu ./ b; % gives zero mean
    pvalue = pearson4cdf(statvalue,m,nu,a,lambda);
    t_fitting = toc;
    fitcdf = pearson4cdf(allstat,m,nu,a,lambda);
case 5
    % inverse gamma location-scale: standard support (-C1,Inf) or (-Inf,-C1)
    C1 = c1 ./ (2*c2);
    pvalue = gamcdf(-((c1 - C1) ./ c2) ./(statvalue+C1),(1./c2 - 1),1);
    t_fitting = toc;
    fitcdf = gamcdf(-((c1 - C1) ./ c2) ./(allstat+C1),(1./c2 - 1),1);
case 6
    % F location-scale: standard support (a2,Inf) or (-Inf,a1)
    m1 = (a1 + c1) ./ (c2.*(a2 - a1));
    m2 = -(a2 + c1) ./ (c2.*(a2 - a1));
    % a1 and a2 have the same sign, and they've been sorted so a1 < a2
    if a2 < 0
        nu1 = 2*(m2 + 1);
        nu2 = -2*(m1 + m2 + 1);
        pvalue = fcdf((statvalue-a2)./(a2 - a1)./(nu1./nu2),nu1,nu2);
        t_fitting = toc;
        fitcdf = fcdf((allstat-a2)./(a2 - a1)./(nu1./nu2),nu1,nu2);
    else % 0 < a1
        nu1 = 2*(m1 + 1);
        nu2 = -2*(m1 + m2 + 1);
        pvalue = fcdf((statvalue-a1)./(a1 - a2)./(nu1./nu2),nu1,nu2);
        t_fitting = toc;
        fitcdf = fcdf((allstat-a1)./(a1 - a2)./(nu1./nu2),nu1,nu2);
end case 7
    % t location-scale: standard support (-Inf,Inf)
    nu = 1./c2 - 1;
    pvalue = tcdf(statvalue./sqrt(c0 ./ (1-c2)),nu);
    t_fitting = toc;
    fitcdf = tcdf(allstat./sqrt(c0 ./ (1-c2)),nu);
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function pvalue = pearson4cdf(r,m,nu,a,lambda)
% pvalue = pearson4cdf(r,m,nu,a,lambda)
%   Pearson distribution series cumulative distribution function (cdf).
%
%   References:
%      [1] Devroye, L. (1986). Non-Uniform Random Variate Generation,
%          Springer-Verlag.  Also available in PDF format on-line at
%          http://cgm.cs.mcgill.ca/~luc/rnbookindex.html.
%      [2] Heinrich, J. (2004). A Guide to the Pearson Type IV Distribution,
%          CDF/MEMO/STATISTICS/PUBLIC/6820, available on-line at
%          http://www-cdf.fnal.gov/publications/cdf6820_pearson4.pdf.

pvalue=zeros(length(r),1); for i = 1:length(r)
 if r(i)>(lambda+a*sqrt(3))
     nu = -nu;
     lambda = -lambda;
     r(i) = -r(i);
 else
  if r(i)>(lambda-a*sqrt(3))
     pvalue(i) = 1/(1-exp(-(nu+2i*m)*pi)) - i*a*pearson4pdf(r(i),m,nu,a,lambda)/...
  (i*nu-2*m+2)*(1+((r(i)-lambda)/a).^2)*hypergeom([1,2-2*m],2-m+i*nu/2,(1+i*(r(i)-lambda)/a)/2);
  else
     pvalue(i) = a/(2*m-1)*(i-(r(i)-lambda)/a)*pearson4pdf(r(i),m,nu,a,lambda)*...
     hypergeom([1,m+i*nu/2],2*m,2/(1-i*(r(i)-lambda)/a));
     if r(i)>(lambda+a*sqrt(3))
        pvalue(i) = 1-pvalue(i);
     end
  end
 end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function f = pearson4pdf(r,m,nu,a,lambda)
% function f = pearson4pdf(r,m,nu,a,lambda)
%   Pearson distribution series probability density function (pdf).

%   References:
%      [1] Heinrich, J. (2004). A Guide to the Pearson Type IV Distribution,
%          CDF/MEMO/STATISTICS/PUBLIC/6820, available on-line at
%          http://www-cdf.fnal.gov/publications/cdf6820_pearson4.pdf.
```

```
K = (1./HypGeo(m,nu/2)).*exp(gammaln(m) - gammaln(m-.5)) ./
(sqrt(pi)*a); f =
K*((1+((r-lambda)/a).^2).^m).*exp(-nu*atan((r-lambda)/a));

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [handleCDF] = mycdfplot(x)
%Revised from the matlab function CDFPLOT
%MYCDFPLOT Display an empirical cumulative distribution function.
%   MYCDFPLOT(X) plots an empirical cumulative distribution function (CDF)
%   of the observations in the data sample vector X. X may be a row or
%   column vector, and represents a random sample of observations from
%   some underlying distribution.
%
%   H = MYCDFPLOT(X) plots F(x), the empirical (or sample) CDF versus the
%   observations in X. The empirical CDF, F(x), is defined as follows:
%
%   F(x) = (Number of observations <= x)/(Total number of observations)
%
%   for all values in the sample vector X. If X contains missing data
%   indicated by NaN's (IEEE arithmetic representation for
%   Not-a-Number), the missing observations will be ignored.
%
%   H is the handle of the empirical CDF curve (a Handle Graphics 'line'
%   object).

[yy,xx,n,emsg,eid] = cdfcalc(x); if ~isempty(eid)
   error(sprintf('stats:mycdfplot:%s',eid),emsg);
end
% Create vectors for plotting
k = length(xx); n = reshape(repmat(1:k, 2, 1), 2*k, 1); xCDF    =
[-Inf; xx(n); Inf]; yCDF    = [0; 0; yy(1+n)];
% Now plot the sample (empirical) CDF staircase.
hCDF = plot(xCDF , yCDF,'r','LineWidth',3); if (nargout>0),
handleCDF=hCDF; end
```

Departments of Statistics, Electrical & Computer Engineering, Psychology, Bioengineering, Beckman Institute, University of Illinois at Urbana-Champaign, IL, U.S.A.

E-mail: czhou4@uiuc.edu

Departments of Statistics, Electrical & Computer Engineering, Psychology, Bioengineering, Beckman Institute, University of Illinois at Urbana-Champaign, IL, U.S.A.

E-mail: ymw@uiuc.edu

(Received April 2007; accepted March 2008))