Supplement to: Nonparametric MANOVA in meaningful effects[1]

Dennis Dobler*, Sarah Friedrich† and Markus Pauly‡

March 15, 2019

Abstract

In this supplemental material to the author's paper 'Nonparametric MANOVA in meaningful effects', we provide a computationally convenient representation of the wild bootstrapped estimator, additional simulation results, and explicit formulas for the asymptotic covariances.

Keywords: Covariance Heteroscedasticity; Multivariate Data; Multivariate Ordinal Data; Multiple Samples; Rank-based Methods; Wild Bootstrap

* Vrije Universiteit Amsterdam, Department of Mathematics, Netherlands
email: d.dobler@vu.nl

† University of Copenhagen, Section of Biostatistics, Danmark
email: safr@sund.ku.dk

‡ Technical University of Dortmund, Germany
email: markus.pauly@tu-dortmund.de

[1] Authors are in alphabetical order.

8 Computationally convenient representation of $\mathbf{p}^\star$

We derive the simplified representation for each entry of the wild bootstrapped effects vector:

$$\begin{aligned}
p_{ij}^\star &= \int G_j^\star(x)\mathrm{d}\widehat{F}_{ij}(x) - \int F_{ij}^\star(x)\mathrm{d}\widehat{G}_j(x) \\
&= \int \Big(\frac{1}{a}\sum_{\ell=1}^{a} F_{\ell j}^\star\Big)\mathrm{d}\widehat{F}_{ij} - \int F_{ij}^\star \mathrm{d}\Big(\frac{1}{a}\sum_{\ell=1}^{a}\widehat{F}_{\ell j}\Big) \\
&= \frac{1}{a}\sum_{\ell\neq i}\Big[\int F_{\ell j}^\star \mathrm{d}\widehat{F}_{ij} - \int F_{ij}^\star \mathrm{d}\widehat{F}_{\ell j}\Big] \\
&= \frac{1}{a}\sum_{\ell\neq i}\Big[\frac{1}{n_\ell}\sum_{k_\ell=1}^{n_\ell} D_{\ell k_\ell}\int (c(x-X_{\ell j k_\ell}) - \widehat{F}_{\ell j}(x))\mathrm{d}\widehat{F}_{ij}(x) \\
&\quad - \frac{1}{n_i}\sum_{k_i=1}^{n_i} D_{i k_i}\int (c(x-X_{ijk_i}) - \widehat{F}_{ij}(x))\mathrm{d}\widehat{F}_{\ell j}(x)\Big] \\
&= \frac{1}{a}\sum_{\ell\neq i}\Big[\frac{1}{n_\ell}\sum_{k_\ell=1}^{n_\ell} D_{\ell k_\ell}\Big\{1 - \frac{\text{placement of } X_{\ell j k_\ell} \text{ within } X_{ij1},\dots,X_{ijn_i}}{n_i} - \widehat{w}_{\ell i j}\Big\} \\
&\quad - \frac{1}{n_i}\sum_{k_i=1}^{n_i} D_{i k_i}\Big\{1 - \frac{\text{placement of } X_{ijk_i} \text{ within } X_{\ell j1},\dots,X_{\ell jn_\ell}}{n_\ell} - \widehat{w}_{i\ell j}\Big\}\Big] \\
&= \frac{1}{a}\sum_{\ell\neq i}\Big[\frac{1}{n_\ell}\sum_{k_\ell=1}^{n_\ell} D_{\ell k_\ell}\Big\{\widehat{w}_{i\ell j} - \frac{R_{\ell j k_\ell}^{(i\ell)} - R_{\ell j k_\ell}^{(\ell)}}{n_i}\Big\} - \frac{1}{n_i}\sum_{k_i=1}^{n_i} D_{ik_i}\Big\{\widehat{w}_{\ell i j} - \frac{R_{ijk_i}^{(\ell i)} - R_{ijk_i}^{(i)}}{n_\ell}\Big\}\Big]
\end{aligned}$$

In the last equality we used that $\widehat{w}_{i\ell j} = 1 - \widehat{w}_{\ell i j}$ and that the placement of $X_{\ell j k_\ell}$ within $X_{ij1},\dots,X_{ijn_i}$ equals $R_{\ell j k_\ell}^{(i\ell)} - R_{\ell j k_\ell}^{(\ell)}$. If we now insert $\widehat{w}_{i\ell j} = \frac{\bar{R}_{\ell j\cdot}^{(i\ell)} - (n_\ell+1)/2}{n_i}$, we obtain

$$\begin{aligned}
p_{ij}^\star = \frac{1}{a}\sum_{\ell\neq i}\Big[&\frac{1}{n_\ell}\sum_{k_\ell=1}^{n_\ell} D_{\ell k_\ell}\Big\{\frac{1}{n_i}\Big(\bar{R}_{\ell j\cdot}^{(i\ell)} - \frac{n_\ell+1}{2}\Big) - \frac{1}{n_i}(R_{\ell j k_\ell}^{(i\ell)} - R_{\ell j k_\ell}^{(\ell)})\Big\} \\
&- \frac{1}{n_i}\sum_{k_i=1}^{n_i} D_{ik_i}\Big\{\frac{1}{n_\ell}\Big(\bar{R}_{ij\cdot}^{(\ell i)} - \frac{n_i+1}{2}\Big) - \frac{1}{n_\ell}(R_{ijk_i}^{(\ell i)} - R_{ijk_i}^{(i)})\Big\}\Big]
\end{aligned}$$

The big advantage of this representation is that the involved mid-ranks do not need to be recalculated in every resampling iteration, and this saves many computer resources.

9 Additional simulation results

9.1 The Brunner-Munzel-Puri method

In case of the multivariate Behrens-Fisher problem with $a = 2$ we additionally compared our results to the ANOVA-type testing procedure of Brunner et al. (2002). This approach is based on approximating the unknown limiting distribution by an F-distribution with estimated degrees of freedom and thus does not provide an asymptotically valid procedure. The simulation scenarios are identical to the ones in Section 5 for continuous and heteroscedastic data in the one-way layout. The results are presented in Tables 5 and 6, respectively. For comparison, we also kept the results of the two bootstrap procedures. In the homoscedastic setting it keep the F-approximation (BMP) proposed by Brunner et al. (2002) leads to similar results as the wild bootstrap approach but with a more conservative behaviour for covariance setting S2 while less liberal behaviour for covariance setting S1. In the heteroscedastic setting, it leads to even more conservative results than the wild bootstrap for most scenarios. Here, the group-wise bootstrap remains the method of choice. Finally, in the ordinal data setting the BMP method shows good type-I error control for all settings except for covariance setting S2 with $d = 8$ dimensions – here it shows an extremely conservative behavior and the wild bootstrap remains the best method.

Table 5: Type-I-error results for the homoscedastic setting with normal and lognormal distributed data with $d = 4$ and $d = 8$ dimensions, varying sample sizes and different covariance settings.

				wild bootstrap			group-wise bootstrap			BMP		
	distr	Cov	$\boldsymbol{n}$	$m = 1$	2	5	1	2	5	1	2	5
$d = 4$	normal	S1	(10, 10)	5.9	5.5	5.6	7.3	6.4	5.9	3.9	4.4	4.8
			(10, 20)	5.6	5.5	5.2	6.6	6.3	5.5	4.0	4.4	4.7
		S2	(10, 10)	5.1	5.3	5.6	7.0	6.5	5.9	3.3	4.1	4.7
			(10, 20)	5.2	5.0	5.1	6.8	6.3	5.3	3.8	4.0	4.6
	lognormal	S1	(10, 10)	7.2	6.3	5.8	7.3	6.7	5.8	4.9	5.4	4.7
			(10, 20)	6.8	6.2	5.6	7.4	6.3	5.5	5.2	5.6	4.9
		S2	(10, 10)	6.5	6.1	5.6	7.2	6.6	5.7	4.5	5.4	4.6
			(10, 20)	6.6	5.9	5.5	7.0	6.5	5.7	4.6	5.1	4.6
$d = 8$	normal	S1	(10, 10)	6.6	5.7	5.2	7.6	6.2	5.5	4.5	4.4	4.7
			(10, 20)	6.1	5.9	5.1	6.9	6.5	5.4	4.2	4.9	4.8
		S2	(10, 10)	3.7	3.9	4.2	6.8	6.1	5.3	2.0	2.6	4.5
			(10, 20)	3.8	3.9	4.6	6.5	5.7	5.1	2.1	2.7	4.5
	lognormal	S1	(10, 10)	7.8	5.9	5.4	8.3	6.2	5.5	6.0	5.1	5.2
			(10, 20)	7.7	6.3	5.3	8.1	6.4	5.5	6.1	5.4	4.9
		S2	(10, 10)	5.4	4.6	4.7	7.8	6.0	5.5	3.2	3.5	4.8
			(10, 20)	5.3	5.2	4.6	7.1	6.2	5.1	3.3	4.1	4.6

Table 6: Type-I-error in % for the heteroscedastic setting.

	σ_i^2	$\boldsymbol{n}$	wild $m=1$	2	5	groupwise 1	2	5	BMP 1	2	5
$d=4$	(1, 2)	(10, 10)	0.6	1.3	3.1	6.3	5.2	5.5	0.2	0.6	4.1
		(10, 20)	0.7	1.7	3.5	5.6	5.2	5.1	0.2	1.1	3.9
		(20, 10)	1.3	2.0	3.3	6.5	5.4	5.2	0.4	1.3	3.6
	(1, 1)	(10, 10)	0.5	1.2	3.2	6	5.3	5.3	0.1	0.7	4.0
		(10, 20)	0.8	1.7	3.7	5.9	5.8	5.4	0.3	1.3	3.8
	(1.2, 1)	(10, 10)	0.6	1.2	3.2	6.2	5.3	5.5	0.1	0.7	4.1
		(10, 20)	0.9	1.9	3.7	6.2	5.9	5.3	0.3	1.3	4.0
		(20, 10)	0.9	1.9	3.5	5.9	5.6	5.2	0.3	1.2	3.9
$d=8$	(1, 2)	(10, 10)	0.1	0.4	2.5	4.8	4.7	4.6	< 0.01	0.1	3.7
		(10, 20)	0.1	0.9	3.2	4.3	4.3	5.3	< 0.01	0.3	3.6
		(20, 10)	0.2	0.9	3.4	4.3	4.8	5.0	0.1	0.3	3.4
	(1, 1)	(10, 10)	< 0.01	0.4	2.5	5.1	4.9	4.6	< 0.01	0.2	3.5
		(10, 20)	0.1	1.1	3.2	4.7	4.4	5.4	< 0.01	0.4	3.3
	(1.2, 1)	(10, 10)	< 0.01	0.3	2.3	4.8	4.9	4.8	< 0.01	0.2	3.4
		(10, 20)	0.2	1.1	3.4	4.9	4.4	5.4	< 0.01	0.4	3.4
		(20, 10)	0.1	0.7	3.4	4.3	4.8	5.6	< 0.01	0.4	3.6

9.2 A one-way layout with three groups

In addition to the simulation results presented in Section 5.1 we furthermore considered a one-way layout with $a = 3$ groups. Covariance settings and distributions were the same as in Section 5.1, i. e., we considered the following covariance settings:

$$\text{Setting 1:} \qquad \mathbf{V}_1 = \mathbf{I}_d + 0.5(\mathbf{J}_d - \mathbf{I}_d) = \mathbf{V}_2 = \mathbf{V}_3,$$
$$\text{Setting 2:} \qquad \mathbf{V}_1 = \left((0.6)^{|r-s|}\right)_{r,s=1}^{d} = \mathbf{V}_2 = \mathbf{V}_3.$$

However, we restricted the simulations to balanced designs starting with $\boldsymbol{n} = (10, 10, 10)'$ here. Sample sizes were again increase by a factor $m \in \{1, 2, 5\}$.

For the heteroscedastic setting, we again considered normally distributed random vectors with $\boldsymbol{\sigma}^2 = (\sigma_1^2, \sigma_2^2, \sigma_3^2)' \in \{(1, 2, 3)', (1, 1, 1)'\}$. Ordinal data was simulated as in Section 5.1. The results presented in Tables 8 – 10 are very similar to the ones obtained for $a = 2$ groups. In particular, the wild bootstrap behaves preferably in the homoscedastic setting and for ordinal data, but is very conservative in the heteroscedastic setting, where the group-wise bootstrap performs better.

Table 7: Type-I-error in % for ordinal data with different sample sizes and different covariance structures.

		$\boldsymbol{n}$	wild $m=1$	2	5	groupwise 1	2	5	BMP 1	2	5
$d=4$	S1	(10, 10)	6.3	5.5	5.0	8.1	6.6	5.4	4.5	4.5	5.2
		(10, 20)	6.7	5.5	5.1	7.7	6.2	5.6	5.1	4.4	4.9
	S2	(10, 10)	5.9	5.1	5.2	7.8	6.8	5.5	4.1	4.1	5.1
		(10, 20)	6.3	5.2	5.2	7.8	6.4	5.6	4.5	4.2	4.7
$d=8$	S1	(10, 10)	6.4	5.6	5.9	7.8	6.4	6.1	4.4	4.8	5.4
		(10, 20)	7.2	5.3	5.2	8.2	5.9	5.5	4.9	4.5	5.2
	S2	(10, 10)	3.6	3.8	5.1	7.6	6.2	6.2	1.8	2.9	4.6
		(10, 20)	4.5	3.8	4.4	7.4	5.6	4.8	2.6	2.7	4.9

Table 8: Type-I-error results in % for the homoscedastic setting with $a = 3$ groups, normal and lognormal distributed data with $d = 4$ and $d = 8$ dimensions and different covariance settings.

	distr	Cov	wild $m=1$	2	5	groupwise 1	2	5
$d=4$	normal	S1	5.9	5.3	5.2	7.0	6.3	5.4
		S2	5.9	5.0	5.3	7.3	6.4	5.8
	lognormal	S1	7.9	5.9	5.7	7.9	6.3	5.7
		S2	6.9	5.4	5.9	7.4	6.1	6.0
$d=8$	normal	S1	6.4	5.4	5.3	7.7	6.2	5.6
		S2	3.1	3.8	4.3	6.8	6.1	4.9
	lognormal	S1	8.5	6.3	5.5	8.5	6.6	5.5
		S2	5.1	5.4	4.9	7.6	6.7	5.3

9.3 Runtime comparisons

To further investigate the advantages and disadvantages of the two proposed bootstrap approaches, we compared their computational cost in yet another simulation study. Here, we considered three different distributions in a 2×2 design:

1. A normal distribution $N(0, \mathbf{I}_d)$
2. A normal distribution with an autoregressive covariance structure $V = \left((0.6)^{|r-s|}\right)_{r,s=1}^{d}$
3. A multinomial distribution with success probability $1/d$

The distributions were the same for all groups. Furthermore, we considered different sample sizes $n = n_1 = n_2 = n_3 = n_4$ and dimensions d. Runtimes are computed in seconds for 5000 bootstrap runs averaged over 500 simulation runs. It turned out that the wild bootstrap is

Table 9: Type-I-error results in % for the heteroscedastic setting with $d = 4$ and $d = 8$ dimensions.

	σ_i^2	wild			groupwise		
		$m = 1$	2	5	1	2	5
$d = 4$	(1, 2, 3)	0.8	1.8	3.4	6.2	5.9	5.3
	(1, 1, 1)	0.6	1.7	3.6	5.8	6.1	5.6
$d = 8$	(1, 2, 3)	< 0.01	0.5	2.8	4.4	4.0	5.1
	(1, 1, 1)	0.1	0.5	3.0	4.3	4.2	4.9

Table 10: Type-I-error rates in % for ordinal data with different covariance structures.

	Cov	$m = 1$	2	5	1	2	5
$d = 4$	S1	7.0	5.4	5.7	8.6	6.1	6.1
	S2	6.1	5.0	5.7	7.9	6.7	6.3
$d = 8$	S1	7.1	5.7	5.7	8.5	6.6	5.8
	S2	3.9	3.8	4.8	7.8	6.2	5.5

between 1.2 and 5.7 times faster than the group-wise bootstrap in the considered scenarios, see Table 11.

In addition, we also compared the runtimes for the simulation settings in Section 5.1. Here we restricted our simulations to $m = 1$, $\mathbf{n} = (10, 20)'$ in the one-way and $\mathbf{n} = (10, 20, 20, 50)'$ in the two-way case, covariance setting S1 as well as normally distributed data for the homoscedastic setting and $\boldsymbol{\sigma}^2 = (1, 2)$ in the heteroscedastic setting. The results are displayed in Table 12. The results are similar to the ones seen in Table 11, with the wild bootstrap always outperforming the group-wise bootstrap.

10 Covariances

In this section we derive the covariance matrix $\boldsymbol{\Sigma}$ of the multivariate limit normal distribution in Theorem 1. The exact representation may not be strictly necessary for the practical purposes in this paper because a covariance estimator is not required due to the wild bootstrap asymptotics as described in Theorem 3. But the covariances below will give some insights into the asymptotically independent components of $\sqrt{N}(\widehat{\mathbf{p}} - \mathbf{p})$ and what kind of studentization may be applied if one wishes to test sub-hypotheses. Furthermore, it is important to see that the limit distribution is not degenerate. Therefore, let $\widehat{\mathbf{w}}$ be the vector consisting of all

$$\widehat{w}_{\ell ij} = \int \widehat{F}_{\ell j} \mathrm{d}\widehat{F}_{ij} = \frac{1}{n_\ell}\frac{1}{n_i}\sum_{k=1}^{n_i}\sum_{r=1}^{n_\ell} c(X_{ijk} - X_{\ell jr}); \quad \ell, i = 1, \ldots, a, \ j = 1, \ldots, d.$$

This estimator is consistent for the vector, say, $\mathbf{w}$ consisting of the different $w_{\ell ij}$. As an intermediate result, we are interested in the asymptotic covariance matrix of the $\sqrt{N}(\widehat{\mathbf{w}} - \mathbf{w})$ vector, i.e. in the limits $\sigma_{\ell ij,\ell' i' j'}$ of

$$N \cdot cov\Big(\int \widehat{F}_{\ell j} \mathrm{d}\widehat{F}_{ij}, \int \widehat{F}_{\ell' j'} \mathrm{d}\widehat{F}_{i' j'}\Big).$$

To this end, we consider an asymptotically linear development which is due to the functional delta-method: Let $\psi : (f, g) \mapsto \int f \mathrm{d}g$ again denote the Wilcoxon functional; cf. Section 3.9.4.1 in van der Vaart and Wellner (1996). As $N \to \infty$ and $\lim n_i/N \to \lambda_i$, $\lim n_\ell/N \to \lambda_\ell$ (according to Condition 1),

$$\begin{aligned}
\sqrt{N}\Big(\int \widehat{F}_{\ell j} \mathrm{d}\widehat{F}_{ij} - \int F_{\ell j} \mathrm{d}F_{ij}\Big) &= \sqrt{N}(\psi(\widehat{F}_{\ell j}, \widehat{F}_{ij}) - \psi(F_{\ell j}, F_{ij}))\\
&= \sqrt{N}\psi'_{(F_{\ell j}, F_{ij})}(\widehat{F}_{\ell j} - F_{\ell j}, \widehat{F}_{ij} - F_{ij}) + o_p(1)\\
&= \int \sqrt{N}(\widehat{F}_{\ell j} - F_{\ell j})\mathrm{d}F_{ij} + \int \sqrt{N} F_{\ell j} \mathrm{d}(\widehat{F}_{\ell j} - F_{\ell j}) + o_p(1)\\
&= \sqrt{N}\Big[- \int F_{ij} \mathrm{d}\widehat{F}_{\ell j} + \int F_{\ell j} \mathrm{d}\widehat{F}_{ij} + \int F_{ij} \mathrm{d}F_{\ell j} - \int F_{\ell j} \mathrm{d}F_{ij}\Big] + o_p(1)\\
&= \sqrt{N}\Big[- \frac{1}{n_\ell}\sum_{r=1}^{n_\ell} F_{ij}(X_{\ell j r}) + \frac{1}{n_i}\sum_{k=1}^{n_i} F_{\ell j}(X_{ijk}) + \int F_{ij} \mathrm{d}F_{\ell j} - \int F_{\ell j} \mathrm{d}F_{ij}\Big] + o_p(1).
\end{aligned}$$

Thus, we know that $\sigma_{\ell ij,\ell' i' j'}$ is the limit of

$$\begin{aligned}
N \cdot cov\Big(&- \frac{1}{n_\ell}\sum_{r=1}^{n_\ell} F_{ij}(X_{\ell j r}) + \frac{1}{n_i}\sum_{k=1}^{n_i} F_{\ell j}(X_{ijk}),\\
&\qquad - \frac{1}{n_{\ell'}}\sum_{r'=1}^{n_{\ell'}} F_{i' j'}(X_{\ell' j' r'}) + \frac{1}{n_{i'}}\sum_{k'=1}^{n_{i'}} F_{\ell' j'}(X_{i' j' k'})\Big)\\
&= \delta_{\ell\ell'}\frac{N}{n_\ell} cov(F_{ij}(X_{\ell j1}), F_{i' j'}(X_{\ell j' 1})) - \delta_{\ell i'}\frac{N}{n_\ell} cov(F_{ij}(X_{\ell j1}), F_{\ell' j'}(X_{\ell j' 1}))\\
&\quad + \delta_{ii'}\frac{N}{n_i} cov(F_{\ell j}(X_{ij1}), F_{\ell' j'}(X_{ij' 1})) - \delta_{i\ell'}\frac{N}{n_i} cov(F_{\ell j}(X_{ij1}), F_{i' j'}(X_{ij' 1})),
\end{aligned}$$

where $\delta_{ii'} = 1\{i = i'\}$ is Kronecker's delta. We continue by calculating any of the above covariances, but we need to distinguish between two cases:
Equal coordinates $\underline{j = j'}$:

$$\begin{aligned}
cov(F_{ij}(X_{\ell j1}), F_{i' j}(X_{\ell j1})) &= \int F_{ij}(u) F_{i' j}(u) \mathrm{d}F_{\ell j}(u) - \int F_{ij}(u) \mathrm{d}F_{\ell j}(u) \int F_{i' j}(u) \mathrm{d}F_{\ell j}(u)\\
&= \tau_{ii'\ell j} - w_{i\ell j} w_{i'\ell j}.
\end{aligned}$$

Unequal coordinates $\underline{j \neq j'}$: Denote by $F_{\ell j j'}$ the joint normalized distribution function of $X_{\ell j1}$ and $X_{\ell j' 1}$.

$$\begin{aligned}
cov(F_{ij}(X_{\ell j1}), F_{i' j'}(X_{\ell j' 1})) &= \int F_{ij}(u) F_{i' j'}(v) \mathrm{d}F_{\ell j j'}(u, v) - \int F_{ij}(u) \mathrm{d}F_{\ell j}(u) \int F_{i' j'}(u) \mathrm{d}F_{\ell j'}(u)\\
&= \rho_{ii'\ell j j'} - w_{i\ell j} w_{i'\ell j'}.
\end{aligned}$$

Recall that $w_{iij} = \frac{1}{2}$. To sum up, we have the following asymptotic covariances (symmetric cases not listed):

$$\begin{cases} 0 & \{i,\ell\} \cap \{i',\ell'\} = \emptyset \text{ or } i = \ell = i' = \ell' \\ \frac{N}{n_\ell}(\tau_{ii\ell j} - w_{i\ell j}^2) + \frac{N}{n_i}(\tau_{\ell\ell ij} - w_{\ell ij}^2) & j = j', i = i' \neq \ell = \ell' \\ -\frac{N}{n_i}(\tau_{i\ell' ij} - w_{iij}w_{\ell' ij}) + \frac{N}{n_i}(\tau_{i\ell' ij} - w_{iij}w_{\ell' ij}) & j = j', i = i' = \ell \neq \ell' \\ \frac{N}{n_\ell}(\tau_{i\ell\ell j} - w_{i\ell j}w_{\ell\ell j}) - \frac{N}{n_\ell}(\tau_{i\ell\ell j} - w_{i\ell j}w_{\ell\ell j}) & j = j', i \neq i' = \ell = \ell' \\ -\frac{N}{n_\ell}(\tau_{ii\ell j} - w_{i\ell j}^2) - \frac{N}{n_i}(\tau_{\ell\ell ij} - w_{\ell ij}^2) & j = j', i = \ell' \neq i' = \ell \\ \frac{N}{n_i}(\tau_{\ell\ell' ij} - w_{\ell ij}w_{\ell' ij}) & j = j', i = i' \neq \ell \neq \ell' \neq i \\ -\frac{N}{n_i}(\tau_{\ell i' ij} - w_{\ell ij}w_{i' ij}) & j = j', i = \ell' \neq i' \neq \ell \neq i \\ \frac{N}{n_\ell}(\rho_{ii\ell jj'} - w_{i\ell j}w_{i\ell j'}) + \frac{N}{n_i}(\rho_{\ell\ell ijj'} - w_{\ell ij}w_{\ell ij'}) & j \neq j', i = i' \neq \ell = \ell' \\ -\frac{N}{n_i}(\rho_{i\ell' ijj'} - w_{iij}w_{\ell' ij'}) + \frac{N}{n_i}(\rho_{i\ell' ijj'} - w_{iij}w_{\ell' ij'}) & j \neq j', i = i' = \ell \neq \ell' \\ \frac{N}{n_\ell}(\rho_{i\ell\ell jj'} - w_{i\ell j}w_{\ell\ell j'}) - \frac{N}{n_\ell}(\rho_{i\ell\ell jj'} - w_{i\ell j}w_{\ell\ell j'}) & j \neq j', i \neq i' = \ell = \ell' \\ -\frac{N}{n_\ell}(\rho_{ii\ell jj'} - w_{i\ell j}w_{i\ell j'}) - \frac{N}{n_i}(\rho_{\ell\ell ijj'} - w_{\ell ij}w_{\ell ij'}) & j \neq j', i = \ell' \neq i' = \ell \\ \frac{N}{n_i}(\rho_{\ell\ell' ijj'} - w_{\ell ij}w_{\ell' ij'}) & j \neq j', i = i' \neq \ell \neq \ell' \neq i \\ -\frac{N}{n_i}(\rho_{\ell i' ijj'} - w_{\ell ij}w_{i' ij'}) & j \neq j', i = \ell' \neq i' \neq \ell \neq i \end{cases}$$

$$= \begin{cases} 0 & \{i,\ell\} \cap \{i',\ell'\} = \emptyset \text{ or } i = \ell = i' = \ell' \\ & \text{or } i = i' = \ell \neq \ell' \text{ or } i \neq i' = \ell = \ell' \\ \frac{N}{n_\ell}(\tau_{ii\ell j} - w_{i\ell j}^2) + \frac{N}{n_i}(\tau_{\ell\ell ij} - w_{\ell ij}^2) & j = j', i = i' \neq \ell = \ell' \\ -\frac{N}{n_\ell}(\tau_{ii\ell j} - w_{i\ell j}^2) - \frac{N}{n_i}(\tau_{\ell\ell ij} - w_{\ell ij}^2) & j = j', i = \ell' \neq i' = \ell \\ \frac{N}{n_i}(\tau_{\ell\ell' ij} - w_{\ell ij}w_{\ell' ij}) & j = j', i = i' \neq \ell \neq \ell' \neq i \\ -\frac{N}{n_i}(\tau_{\ell i' ij} - w_{\ell ij}w_{i' ij}) & j = j', i = \ell' \neq i' \neq \ell \neq i \\ \frac{N}{n_\ell}(\rho_{ii\ell jj'} - w_{i\ell j}w_{i\ell j'}) + \frac{N}{n_i}(\rho_{\ell\ell ijj'} - w_{\ell ij}w_{\ell ij'}) & j \neq j', i = i' \neq \ell = \ell' \\ -\frac{N}{n_\ell}(\rho_{ii\ell jj'} - w_{i\ell j}w_{i\ell j'}) - \frac{N}{n_i}(\rho_{\ell\ell ijj'} - w_{\ell ij}w_{\ell ij'}) & j \neq j', i = \ell' \neq i' = \ell \\ \frac{N}{n_i}(\rho_{\ell\ell' ijj'} - w_{\ell ij}w_{\ell' ij'}) & j \neq j', i = i' \neq \ell \neq \ell' \neq i \\ -\frac{N}{n_i}(\rho_{\ell i' ijj'} - w_{\ell ij}w_{i' ij'}) & j \neq j', i = \ell' \neq i' \neq \ell \neq i \end{cases}$$

In order to present the above covariances in a more compact matrix notation, we introduce the following matrices: Denote by $\mathbf{0}_{p\times q} \in \mathbb{R}^{p\times q}$ the $(p \times q)$-matrix of zeros, by $\mathbf{0}_r \in \mathbb{R}^r$ the r-dimensional column vector of zeros, by $\boldsymbol{\tau}_{ii'\ell\cdot} = diag(\tau_{ii'\ell 1}, \ldots, \tau_{ii'\ell d}) \in \mathbb{R}^{d\times d}$ the $(d \times d)$-diagonal matrices of τ's, by

$$\boldsymbol{\rho}_{ii'\ell\cdot\cdot} = \begin{pmatrix} 0 & \rho_{ii'\ell 12} & \rho_{ii'\ell 13} & \cdots & \rho_{ii'\ell 1d} \\ \rho_{ii'\ell 21} & 0 & \rho_{ii'\ell 23} & \cdots & \rho_{ii'\ell 2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{ii'\ell(d-1)1} & \rho_{ii'\ell(d-1)2} & \rho_{ii'\ell(d-1)3} & \cdots & \rho_{ii'\ell(d-1)d} \\ \rho_{ii'\ell d1} & \rho_{ii'\ell d2} & \rho_{ii'\ell d3} & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{d\times d}$$

the $(d \times d)$-matrices of ρ's with zeros along the diagonal entries, and the vector of treatment effects between groups i and i' by $\mathbf{w}_{ii'\cdot} = (w_{ii'1}, w_{ii'2}, \ldots, w_{ii'd})^T \in \mathbb{R}^d$. Recall that, in the whole w-vector, we first first the ℓ-value, then the i-value, so that we first go through the component index j. With the above notation, we thus obtain the following first block of the covariance matrix in which $\ell = \ell' = 1$ which, for general indices ℓ and ℓ', we denote by $\boldsymbol{\Sigma}_{\ell\ell'} \in \mathbb{R}^{da \times da}$:

$$
\begin{aligned}
\boldsymbol{\Sigma}_{11} = \frac{N}{n_1} &\begin{pmatrix} \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \boldsymbol{\tau}_{221\cdot} & \boldsymbol{\tau}_{231\cdot} & \ldots & \boldsymbol{\tau}_{2a1\cdot} \\ \mathbf{0}_{d\times d} & \boldsymbol{\tau}_{321\cdot} & \boldsymbol{\tau}_{331\cdot} & \ldots & \boldsymbol{\tau}_{3a1\cdot} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \boldsymbol{\tau}_{a21\cdot} & \boldsymbol{\tau}_{a31\cdot} & \ldots & \boldsymbol{\tau}_{aa1\cdot} \end{pmatrix} + \begin{pmatrix} \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \frac{N}{n_2}\boldsymbol{\tau}_{112\cdot} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \frac{N}{n_3}\boldsymbol{\tau}_{113\cdot} & \ldots & \mathbf{0}_{d\times d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \frac{N}{n_a}\boldsymbol{\tau}_{11a\cdot} \end{pmatrix} \\
+ \frac{N}{n_1} &\begin{pmatrix} \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \boldsymbol{\rho}_{221\cdot\cdot} & \boldsymbol{\rho}_{231\cdot\cdot} & \ldots & \boldsymbol{\rho}_{2a1\cdot\cdot} \\ \mathbf{0}_{d\times d} & \boldsymbol{\rho}_{321\cdot\cdot} & \boldsymbol{\rho}_{331\cdot\cdot} & \ldots & \boldsymbol{\rho}_{3a1\cdot\cdot} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \boldsymbol{\rho}_{a21\cdot\cdot} & \boldsymbol{\rho}_{a31\cdot\cdot} & \ldots & \boldsymbol{\rho}_{aa1\cdot\cdot} \end{pmatrix} + \begin{pmatrix} \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \frac{N}{n_2}\boldsymbol{\rho}_{112\cdot\cdot} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \frac{N}{n_3}\boldsymbol{\rho}_{113\cdot\cdot} & \ldots & \mathbf{0}_{d\times d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \frac{N}{n_a}\boldsymbol{\rho}_{11a\cdot\cdot} \end{pmatrix} \\
- \frac{N}{n_1} &\begin{pmatrix} \mathbf{0}_d \\ \mathbf{w}_{21\cdot} \\ \mathbf{w}_{31\cdot} \\ \vdots \\ \mathbf{w}_{a1\cdot} \end{pmatrix} \begin{pmatrix} \mathbf{0}_d \\ \mathbf{w}_{21\cdot} \\ \mathbf{w}_{31\cdot} \\ \vdots \\ \mathbf{w}_{a1\cdot} \end{pmatrix}^T - \begin{pmatrix} \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \frac{N}{n_2}\mathbf{w}_{12\cdot}\mathbf{w}_{12\cdot}^T & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \frac{N}{n_3}\mathbf{w}_{13\cdot}\mathbf{w}_{13\cdot}^T & \ldots & \mathbf{0}_{d\times d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \frac{N}{n_a}\mathbf{w}_{1a\cdot}\mathbf{w}_{1a\cdot}^T \end{pmatrix}
\end{aligned}
$$

Note that the other $\boldsymbol{\Sigma}_{\ell\ell}$-matrices have a similar structure but with the $\mathbf{0}_{d\times d}$-matrices in the ℓth block row and block column and with all 1's replaced by ℓ's. In the same way,

$$
\begin{aligned}
\boldsymbol{\Sigma}_{21} = -\frac{N}{n_2} &\begin{pmatrix} \mathbf{0}_{d\times d} & \boldsymbol{\tau}_{112\cdot} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \boldsymbol{\tau}_{132\cdot} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \boldsymbol{\tau}_{142\cdot} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \boldsymbol{\tau}_{1a2\cdot} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \end{pmatrix} - \frac{N}{n_1} \begin{pmatrix} \mathbf{0}_{d\times d} & \boldsymbol{\tau}_{221\cdot} & \boldsymbol{\tau}_{231\cdot} & \ldots & \boldsymbol{\tau}_{2a1\cdot} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \end{pmatrix} \\
- \frac{N}{n_2} &\begin{pmatrix} \mathbf{0}_{d\times d} & \boldsymbol{\rho}_{112\cdot\cdot} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \boldsymbol{\rho}_{132\cdot\cdot} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \boldsymbol{\rho}_{142\cdot\cdot} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \boldsymbol{\rho}_{1a2\cdot\cdot} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \end{pmatrix} - \frac{N}{n_1} \begin{pmatrix} \mathbf{0}_{d\times d} & \boldsymbol{\rho}_{221\cdot\cdot} & \boldsymbol{\rho}_{231\cdot\cdot} & \ldots & \boldsymbol{\rho}_{2a1\cdot\cdot} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \ldots & \mathbf{0}_{d\times d} \end{pmatrix}
\end{aligned}
$$

$$+\begin{pmatrix} \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \dots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \dots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \frac{N}{n_3}\boldsymbol{\tau}_{213\cdot} & \mathbf{0}_{d\times d} & \dots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \frac{N}{n_4}\boldsymbol{\tau}_{214\cdot} & \dots & \mathbf{0}_{d\times d} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \dots & \frac{N}{n_a}\boldsymbol{\tau}_{21a\cdot} \end{pmatrix}$$

$$+\begin{pmatrix} \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \dots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \dots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \frac{N}{n_3}\boldsymbol{\rho}_{213\cdot\cdot} & \mathbf{0}_{d\times d} & \dots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \frac{N}{n_4}\boldsymbol{\rho}_{214\cdot\cdot} & \dots & \mathbf{0}_{d\times d} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \dots & \frac{N}{n_a}\boldsymbol{\rho}_{21a\cdot\cdot} \end{pmatrix}$$

$$+\frac{N}{n_2}\left(\mathbf{0}_{da\times d} \quad \begin{pmatrix} \mathbf{w}_{12\cdot}^T \\ \mathbf{0}_d^T \\ \mathbf{w}_{32\cdot}^T \\ \mathbf{w}_{42\cdot}^T \\ \vdots \\ \mathbf{w}_{a2\cdot}^T \end{pmatrix} \otimes \mathbf{w}_{12\cdot} \quad \mathbf{0}_{da\times d(a-2)}\right)$$

$$+\frac{N}{n_1}\begin{pmatrix} \begin{pmatrix}\mathbf{0}_d^T & \mathbf{w}_{21\cdot}^T & \mathbf{w}_{31\cdot}^T & \dots & \mathbf{w}_{a1\cdot}^T\end{pmatrix} \otimes \mathbf{w}_{21\cdot} \\ \mathbf{0}_{d(a-1)\times da} \end{pmatrix}$$

$$-\begin{pmatrix} \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \dots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \dots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \frac{N}{n_3}\mathbf{w}_{23\cdot}\mathbf{w}_{13\cdot}^T & \mathbf{0}_{d\times d} & \dots & \mathbf{0}_{d\times d} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \frac{N}{n_4}\mathbf{w}_{24\cdot}\mathbf{w}_{14\cdot}^T & \dots & \mathbf{0}_{d\times d} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d} & \dots & \frac{N}{n_a}\mathbf{w}_{2a\cdot}\mathbf{w}_{1a\cdot}^T \end{pmatrix}$$

The representation of the general block matrix $\mathbf{\Sigma}_{\ell\ell'}$ with $\ell \neq \ell'$ is similarly obtained, where the zero-rows have to shifted to the row block number ℓ and the zero-column to the column block number ℓ'. Furthermore, the repeating 1's and 2's in the above representation need to be replaced with ℓ''s and ℓ's, respectively.

Since each $\widehat{p}_{ij}$ is the mean of $\widehat{w}_{1ij}, \widehat{w}_{2ij}, \dots, \widehat{w}_{aij}$, we conclude that the limit covariance matrix of $\sqrt{N}(\widehat{\mathbf{p}} - \mathbf{p})$ is given by

$$\mathbf{\Sigma} = \lim_{N\to\infty} \frac{1}{a^2} \sum_{\ell=1}^{a} \sum_{\ell'=1}^{a} \mathbf{\Sigma}_{\ell\ell'}.$$

□

References

Brunner, E., Munzel, U., and Puri, M. L. (2002). The multivariate nonparametric Behrens–Fisher problem. *Journal of Statistical Planning and Inference*, 108(1):37–53.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and empirical processes*. Springer, New York.

Table 11: Run times for 5000 bootstrap runs in seconds, averaged over 500 simulations

Setting	Sample size N	d	wild bootstrap	group-wise bootstrap
$N(0, \mathbf{I}_d)$	25	1	0.18	0.22
		2	0.37	0.59
		4	0.32	0.78
		8	0.23	0.74
	50	1	0.17	0.23
		2	0.21	0.42
		4	0.22	0.53
		8	0.22	0.97
	100	1	0.20	0.28
		2	0.30	0.61
		4	0.22	0.65
		8	0.30	1.73
$N(0, AR(0.6))$	25	1	0.17	0.23
		2	0.30	0.51
		4	0.48	1.05
		8	0.38	1.60
	50	1	0.31	0.42
		2	0.24	0.46
		4	0.36	0.89
		8	0.41	1.55
	100	1	0.38	0.55
		2	0.38	0.93
		4	0.21	0.73
		8	0.51	2.25
Multinomial$(N, 1/d)$	25	1	0.19	0.23
		2	0.21	0.35
		4	0.26	0.58
		8	0.57	1.76
	50	1	0.18	0.22
		2	0.30	0.58
		4	0.29	0.84
		8	0.22	0.81
	100	1	0.18	0.23
		2	0.48	0.89
		4	0.26	0.78
		8	0.47	1.98

Table 12: Run times for 5000 bootstrap runs in seconds, averaged over 500 simulations

	Setting	d	wild bootstrap	group-wise bootstrap
one-way	homoscedastic	4	0.47	1.15
	data	8	0.47	1.69
	heteroscedastic	4	0.28	0.71
	data	8	0.29	1.09
	ordinal	4	0.27	0.69
	data	8	0.24	0.60
two-way	homoscedastic	4	0.74	2.47
	heteroscedastic	4	0.46	1.33
	ordinal	4	0.44	1.26