Economic Review (Otaru University of Commerce),
Vol.68, No.4, 1-7, March, 2018.

Supplement to the paper "Accurate distributions of Mallows' C_p and its unbiased modifications with applications to shrinkage estimation"

Haruhiko Ogasawara

This article supplements Ogasawara (2017).

Proof of Theorem 1

Since

$$\mathrm{C}_{pq} = (n - p_\Omega)\mathrm{tr}(\mathbf{U}_\Omega^{-1}\mathbf{U}) - nq + 2pq$$
$$= (n - p_\Omega)\mathrm{tr}(\mathbf{I}_{(q)} + \mathbf{U}_\Omega^{-1}\mathbf{U}_{p|\Omega}) - nq + 2pq$$

and $$\mathrm{E}(\mathbf{\Sigma}_0^{1/2}\mathbf{U}_\Omega^{-1}\mathbf{U}_{p|\Omega}\mathbf{\Sigma}_0^{-1/2} \mid \mathbf{\Lambda} = \mathbf{O}) = \frac{p_\Omega - p}{n - p_\Omega - q - 1}\mathbf{I}_{(q)} \qquad \text{(A.1)}$$

(see e.g. Siotani, Hayakawa & Fujikoshi, 1985, Equation (2.4.11)), we have

$$\mathrm{E}(\mathrm{C}_{pq} \mid \mathbf{\Lambda} = \mathbf{O}) = (n - p_\Omega)\left(1 + \frac{p_\Omega - p}{n - p_\Omega - q - 1}\right)q - nq + 2pq$$

$$= -p_\Omega q + 2pq + \frac{(n - p_\Omega)(p_\Omega - p)q}{n - p_\Omega - q - 1} = pq + \frac{q(q+1)(p_\Omega - p)}{n - p_\Omega - q - 1}.$$

When $\mathbf{\Lambda} = \mathbf{O}$, $\mathrm{E}(\mathrm{GD}_{pq}) = pq$, which gives from the above result

$$\mathrm{E}(\overline{\mathrm{C}}_{pq}) - \mathrm{E}(\mathrm{GD}_{pq}) = \mathrm{E}(\mathrm{C}_{pq}) - \frac{q(q+1)(p_\Omega - p)}{n - p_\Omega - q - 1} - pq = 0.$$

Proof of Theorem 2

The expectations in (3.4) are given by (2.2), (2.4), (2.5) and (2.6) for $\mathbf{\Lambda} = \mathbf{O}$. For the variances of (3.4), noting that under normality $\mathbf{U}_\Omega^{-1}$ and $\mathbf{U}_{p|\Omega}$ are independent, the following result will be used when X_i is independent of Y_j $(i, j = 1, 2)$:

$$\begin{aligned}
&\mathrm{cov}(X_1Y_1,\, X_2Y_2) = \mathrm{E}(X_1Y_1X_2Y_2) - \mathrm{E}(X_1Y_1)\mathrm{E}(X_2Y_2) \\
&= \mathrm{E}(X_1X_2)\mathrm{E}(Y_1Y_2) - \mathrm{E}(X_1)\mathrm{E}(Y_1)\mathrm{E}(X_2)\mathrm{E}(Y_2) \\
&= \{\mathrm{cov}(X_1, X_2) + \mathrm{E}(X_1)\mathrm{E}(X_2)\}\{\mathrm{cov}(Y_1,\, Y_2) + \mathrm{E}(Y_1)\mathrm{E}(Y_2)\} \\
&\quad - \mathrm{E}(X_1)\mathrm{E}(X_2)\mathrm{E}(Y_1)\mathrm{E}(Y_2) \\
&= \mathrm{cov}(X_1, X_2)\mathrm{cov}(Y_1,\, Y_2) + \mathrm{E}(X_1)\mathrm{E}(X_2)\mathrm{cov}(Y_1,\, Y_2) \\
&\quad + \mathrm{cov}(X_1, X_2)\mathrm{E}(Y_1)\mathrm{E}(Y_2).
\end{aligned} \tag{A.2}$$

When $\mathbf{\Lambda} = \mathbf{O}$, since $\mathbf{U}^*_{p|\Omega} \equiv \mathbf{\Sigma}_0^{-1/2}\mathbf{U}_{p|\Omega}\mathbf{\Sigma}_0^{-1/2}$ is Wishart-distributed with the covariance matrix $\mathbf{I}_{(n)}$ and $p_\Omega - p$ degrees of freedom, which is denoted by $\mathrm{W}(\mathbf{I}_{(n)}, p_\Omega - p)$, we have

$$\mathrm{cov}\{(\mathbf{U}^*_{p|\Omega})_{ij}, (\mathbf{U}^*_{p|\Omega})_{kl}\} = (p_\Omega - p)(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \ (i, j, k, l = 1, \ldots, q),$$

where $(\cdot)_{ij}$ indicates the (*i, j*)th element of a matrix and δ_{ik} is the Kronecker delta. On the other hand, $\mathbf{U}^{*-1}_\Omega \equiv \mathbf{\Sigma}_0^{1/2}\mathbf{U}^{-1}_\Omega\mathbf{\Sigma}_0^{1/2}$ is inverse-Wishart distributed as $\mathrm{W}^{-1}(\mathbf{I}_{(n)}, n - p_\Omega)$ and

$$\mathrm{cov}\{(\mathbf{U}^{*-1}_\Omega)_{ij}, (\mathbf{U}^{*-1}_\Omega)_{kl}\} = \frac{2\delta_{ij}\delta_{kl} + (n - p_\Omega - q - 1)(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})}{(n - p_\Omega - q)(n - p_\Omega - q - 1)^2(n - p_\Omega - q - 3)} \quad (i, j, k, l = 1, \ldots, q) \tag{A.3}$$

(see e.g. Siotani et al., 1985, Equation (2.4.12)).

From (A.2),

$$\begin{aligned}
&\mathrm{var}\{\mathrm{tr}(\mathbf{U}^{-1}_\Omega\mathbf{U}_{p|\Omega})\} = \mathrm{var}\{\mathrm{tr}(\mathbf{U}^{*-1}_\Omega\mathbf{U}^*_{p|\Omega})\} \\
&= (p_\Omega - p)^2\,\mathrm{var}\left\{\sum_{i=1}^q(\mathbf{U}^{*-1}_\Omega)_{ii}\right\} + (n - p_\Omega - q - 1)^{-2}\,\mathrm{var}\left\{\sum_{i=1}^q(\mathbf{U}^*_{p|\Omega})_{ii}\right\} \\
&\quad + \sum_{i,j,k,l=1}^q \mathrm{cov}\{(\mathbf{U}^{*-1}_\Omega)_{ij},\, (\mathbf{U}^{*-1}_\Omega)_{kl}\}\mathrm{cov}\{(\mathbf{U}^*_{p|\Omega})_{ij},\, (\mathbf{U}^*_{p|\Omega})_{kl}\},
\end{aligned}$$

where

$$\mathrm{cov}\{(\mathbf{U}^{*-1}_\Omega)_{ii}, (\mathbf{U}^{*-1}_\Omega)_{jj}\} = \frac{2 + 2(n - p_\Omega - q - 1)\delta_{ij}}{(n - p_\Omega - q)(n - p_\Omega - q - 1)^2(n - p_\Omega - q - 3)},$$

$$\mathrm{cov}\{(\mathbf{U}^*_{p|\Omega})_{ii}, (\mathbf{U}^*_{p|\Omega})_{jj}\} = 2(p_\Omega - p)\delta_{ij} \quad (i, j = 1, \ldots, q).$$

Consequently,

$$\begin{aligned}
&\mathrm{var}\{\mathrm{tr}(\mathbf{U}_\Omega^{-1}\mathbf{U}_{p|\Omega})\} \\
&= (p_\Omega - p)^2 \sum_{i,j=1}^{q} \frac{2+2(n-p_\Omega-q-1)\delta_{ij}}{(n-p_\Omega-q)(n-p_\Omega-q-1)^2(n-p_\Omega-q-3)} \\
&\quad + (n-p_\Omega-q-1)^{-2} \sum_{i,j=1}^{q} 2(p_\Omega-p)\delta_{ij} \\
&\quad + \sum_{i,j,k,l=1}^{q} (p_\Omega-p)(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}) \frac{2\delta_{ij}\delta_{kl}+(n-p_\Omega-q-1)(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})}{(n-p_\Omega-q)(n-p_\Omega-q-1)^2(n-p_\Omega-q-3)} \\
&= \frac{2(p_\Omega-p)^2\{q^2+q(n-p_\Omega-q-1)\}}{(n-p_\Omega-q)(n-p_\Omega-q-1)^2(n-p_\Omega-q-3)} + \frac{2(p_\Omega-p)q}{(n-p_\Omega-q-1)^2} \\
&\quad + \frac{(p_\Omega-p)\{4q+2(n-p_\Omega-q-1)(q^2+q)\}}{(n-p_\Omega-q)(n-p_\Omega-q-1)^2(n-p_\Omega-q-3)},
\end{aligned} \tag{A.4}$$

which gives the variances in (3.4). Equation (A.4) is partially justified in that when q = 1, (A.4) with (3.3) gives the well-known variance

$$\mathrm{var}(F) = \frac{2(n-p_\Omega)^2(n-p-2)}{(p_\Omega-p)(n-p_\Omega-2)^2(n-p_\Omega-4)} \tag{A.5}$$

of the central F distribution with $p_\Omega - p$ and $n - p_\Omega$ degrees of freedom.

Proof of Corollary 2

From (3.2) when q =1,

$$\begin{aligned}
\mathrm{C}_p &= (n-p_\Omega)\left(1+\frac{p_\Omega-p}{n-p_\Omega}F^*\right) - n + 2p = (p_\Omega-p)F^* + 2p - p_\Omega, \\
\overline{\mathrm{C}}_p &= \mathrm{C}_p - \frac{2(p_\Omega-p)}{n-p_\Omega-2} = (p_\Omega-p)F^* + 2p - p_\Omega - \frac{2(p_\Omega-p)}{n-p_\Omega-2}, \\
\mathrm{MC}_{pq} &= (p_\Omega-p)\frac{n-p_\Omega-2}{n-p_\Omega}F^* + 2p - p_\Omega \qquad \text{(A.6)} \\
&\left(= \overline{\mathrm{C}}_p + 2(p_\Omega-p)\left(\frac{1}{n-p_\Omega-2} - \frac{F^*}{n-p_\Omega}\right)\right),
\end{aligned}$$

which yield the results of Corollary 2.

Proof of Corollary 3

The properties of the noncentral F distribution are well documented (e.g., Johnson, Kotz & Balakrishnan, 1994, Chapter 30). The expectation when $n > p_\Omega + 2$ and variance when $n > p_\Omega + 4$ for the noncentral F distribution denoted by F^* in Corollary 2 are

$$\mathrm{E}(F^*) = \frac{(p_\Omega - p + \lambda)(n - p_\Omega)}{(p_\Omega - p)(n - p_\Omega - 2)},$$

$$\mathrm{var}(F^*) = 2\left(\frac{n - p_\Omega}{p_\Omega - p}\right)^2 \frac{(p_\Omega - p + \lambda)^2 + (p_\Omega - p + 2\lambda)(n - p_\Omega - 2)}{(n - p_\Omega - 2)^2 (n - p_\Omega - 4)}, \quad \text{(A.7)}$$

respectively. Then, when $\lambda = O(n)$,

$$\begin{aligned}\mathrm{E}(\mathrm{C}_p) &= (p_\Omega - p)\mathrm{E}(F^*) + 2p - p_\Omega \\ &= \frac{(p_\Omega - p + \lambda)(n - p_\Omega)}{n - p_\Omega - 2} + 2p - p_\Omega = \lambda + O(1),\end{aligned}$$

$$\begin{aligned}\mathrm{E}(\overline{\mathrm{C}}_p) &= (p_\Omega - p)\mathrm{E}(F^*) + 2p - p_\Omega - \frac{2(p_\Omega - p)}{n - p_\Omega - 2} \\ &= \frac{(p_\Omega - p + \lambda)(n - p_\Omega)}{n - p_\Omega - 2} + 2p - p_\Omega - \frac{2(p_\Omega - p)}{n - p_\Omega - 2} = \lambda + O(1),\end{aligned}$$

$$\begin{aligned}\mathrm{E}(\mathrm{MC}_{pq}) &= (p_\Omega - p)\frac{n - p_\Omega - 2}{n - p_\Omega}\mathrm{E}(F^*) + 2p - p_\Omega \\ &= (p_\Omega - p + \lambda) + (2p - p_\Omega) = p + \lambda = \lambda + O(1),\end{aligned}$$

which give (3.7).

Using (A.7),

$$\mathrm{var}(F^*) = \frac{2(\lambda^2 + 2\lambda n)}{(p_\Omega - p)^2 n} + O(1) = O(n). \quad \text{(A.8)}$$

follows. Equation (A.8) gives (3.8). From the unbiased property of MC_{pq} and the definitions of C_p and $\overline{\mathrm{C}}_p$, we have the results of (3.9) except its last inequality $\mathrm{MSE}(\overline{\mathrm{C}}_p) < \mathrm{MSE}(\mathrm{C}_p)$, which is given by

$$\{\mathrm{E}(\mathrm{C}_p)-(p+\lambda)\}^2=\left\{\frac{(p_\Omega-p+\lambda)(n-p_\Omega)}{n-p_\Omega-2}+p-\lambda-p_\Omega\right\}^2$$
$$=\frac{4(p_\Omega-p+\lambda)^2}{(n-p_\Omega-2)^2},$$
$$\{\mathrm{E}(\overline{\mathrm{C}}_p)-(p+\lambda)\}^2=\frac{1}{(n-p_\Omega-2)^2}\{-2(p-\lambda-p_\Omega)-2(p_\Omega-p))\}^2$$
$$=\frac{4\lambda^2}{(n-p_\Omega-2)^2}<\{\mathrm{E}(\mathrm{C}_p)-(p+\lambda)\}^2$$
(recall the assumption $p_\Omega > p$ in Section 1) and $\mathrm{var}(\mathrm{C}_p)=\mathrm{var}(\overline{\mathrm{C}}_p)$.

Proof of Lemma 1

Since $\mathrm{MSE}(d\hat{\theta})=(d-1)^2\theta_0^2+d^2\sigma_{\theta n}^2$, $\mathrm{MSE}(d\hat{\theta})$ is minimized when $d=d_{\min}=\theta_0^2/(\theta_0^2+\sigma_{\theta n}^2)=1/\{1+c_{\mathrm{V}}^2(\hat{\theta})\}$. The minimized MSE is
$$\theta_0^2-\frac{\theta_0^4}{\theta_0^2+\sigma_{\theta n}^2}=\frac{\sigma_{\theta n}^2}{1+c_{\mathrm{V}}^2(\hat{\theta})}=\frac{\mathrm{MSE}(\hat{\theta})}{1+c_{\mathrm{V}}^2(\hat{\theta})}.$$

Proof of Corollary 4

First, we obtain
$$\mathrm{MSE}(\mathrm{MC}_{pq})-\mathrm{MSE}(d_{\min\overline{\mathrm{C}}_{pq}}\overline{\mathrm{C}}_{pq})$$
$$=\left\{\left(\frac{n-p_\Omega-q-1}{n-p_\Omega}\right)^2-\frac{1}{1+\mathrm{var}(\overline{\mathrm{C}}_{pq})(pq)^{-2}}\right\}\mathrm{var}(\overline{\mathrm{C}}_{pq})$$
$$=\frac{(n-p_\Omega-q-1)^2\{(pq)^2+\mathrm{var}(\overline{\mathrm{C}}_{pq})\}-(n-p_\Omega)^2(pq)^2}{(n-p_\Omega)^2\{(pq)^2+\mathrm{var}(\overline{\mathrm{C}}_{pq})\}}\mathrm{var}(\overline{\mathrm{C}}_{pq}), \quad \text{(A.9)}$$
which can be positive or negative, as shown in the following examples. When $q = 1$, the numerator of the first factor on the right-hand side of the last equation of (A.9) is

$$
\begin{aligned}
&(n-p_\Omega-2)^2\{p^2+\mathrm{var}(\overline{\mathrm{C}}_p)\}-(n-p_\Omega)^2p^2\\
&=-4(n-p_\Omega)p^2+4p^2+\frac{2(p_\Omega-p)(n-p_\Omega)^2(n-p-2)}{n-p_\Omega-4}\\
&=-|O(n)|+|O(1)|+|O(n^2)|,
\end{aligned}
\tag{A.10}
$$

where for $\mathrm{var}(\overline{\mathrm{C}}_p)$, (3.2) and (A.5) are used.

When n is sufficiently large, (A.10) is positive, demonstrating that in this case, $\mathrm{MSE}(\mathrm{MC}_{pq})>\mathrm{MSE}(d_{\min\overline{\mathrm{C}}_{pq}}\overline{\mathrm{C}}_{pq})$. However, when n is relatively small, we define $n-p_\Omega=a>4$ (see a condition for (A.3)) and $p_\Omega-p=b>0$ (recall the assumption $p_\Omega>p$ in Section 1). Then, (A.10) becomes $-4ap^2+4p^2+2ba^2(a+b-2)/(a-4)$, which is negative when $p^2>ba^2(a+b-2)/\{2(a-1)(a-4)\}$. For instance, when a = 5 and b = 1, the last inequality holds when $p\geq 4$. From this result, we have the central inequality $\min\{\cdot\}\leq\max\{\cdot\}$ in (4.2). The remaining inequalities are given by the unbiased property of MC_{pq} and the definitions of C_{pq} and $\overline{\mathrm{C}}_{pq}$.

Proof of Theorem 4

From (A.6) and (A.7), we have

$$
\begin{aligned}
\mathrm{var}(\mathrm{MC}_{pq})&=(p_\Omega-p)^2\left(\frac{n-p_\Omega-2}{n-p_\Omega}\right)^2\mathrm{var}(F^*)\\
&=2\frac{(p_\Omega-p+\lambda)^2+(p_\Omega-p+2\lambda)(n-p_\Omega-2)}{n-p_\Omega-4}.
\end{aligned}
\tag{A.11}
$$

Substituting (A.11) for the first equation of (4.3) given by Lemma 1, the second equation of (4.3) follows.

Results associated with Theorem 4 when $\lambda=O(1)$ and $\lambda=0$

When $\lambda=O(1)$, from (A.11) we have

$$
\begin{aligned}
&\mathrm{var}(\mathrm{MC}_{pq})=2(p_\Omega-p+2\lambda)+O(n^{-1}),\\
&d^*_{\min\mathrm{MC}_{pq}}=\frac{(p+\lambda)^2}{(p+\lambda)^2+2(p_\Omega-p+2\lambda)}+O(n^{-1}).
\end{aligned}
\tag{A.12}
$$

Note that when $\lambda = 0$, (3.2) and (A.7) yield

$$\mathrm{var}(\mathrm{C}_p) = \mathrm{var}(\bar{\mathrm{C}}_p) = (p_\Omega - p)^2 \,\mathrm{var}(F_{p_\Omega - p, n - p_\Omega})$$

$$= \frac{2(p_\Omega - p)(n - p_\Omega)^2 (n - p - 2)}{(n - p_\Omega - 2)^2 (n - p_\Omega - 4)} = \left(\frac{n - p_\Omega}{n - p_\Omega - 2} \right)^2 \mathrm{var}(\mathrm{MC}_{pq}) \qquad \text{(A.13)}$$

$$> \mathrm{var}(\mathrm{MC}_{pq}) = \frac{2(p_\Omega - p)(n - p - 2)}{n - p_\Omega - 4},$$

$$\mathrm{var}(\mathrm{C}_p) = \mathrm{var}(\bar{\mathrm{C}}_p) = 2(p_\Omega - p) + O(n^{-1}),$$

$$\mathrm{var}(\mathrm{MC}_{pq}) = 2(p_\Omega - p) + O(n^{-1}),$$

$$d_{\min \mathrm{MC}_{pq}} = \frac{p^2}{p^2 + \mathrm{var}(\mathrm{MC}_{pq})} = \frac{p^2 (n - p_\Omega - 4)}{p^2 (n - p_\Omega - 4) + 2(p_\Omega - p)(n - p - 2)}$$

$$= \frac{p^2}{p^2 + 2(p_\Omega - p)} + O(n^{-1})$$

(see (4.1)). From (A.12) and (A.13), when $\lambda = O(1)$, it is seen that (A.12) is given from the last two sets of results of (A.13) by replacing $p_\Omega - p$ and p^2 with $p_\Omega - p + 2\lambda$ and $(p + \lambda)^2$, respectively. However, as described earlier, generally $\lambda = O(n)$, giving (A.8).

References

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions* Vol.2 (2nd ed.). New York: Wiley.

Ogasawara, H. (2017). Accurate distributions of Mallows' C$_p$ and its unbiased modifications with applications to shrinkage estimation. *Journal of Statistical Planning and Inference, 184*, 105-116.

Siotani, M., Hayakawa, T., & Fujikoshi, Y. (1985). *Modern multivariate statistical analysis: A graduate course and handbook*. Columbus, OH: American Sciences Press.