Biometrical Journal **0** (0) 0, zzz–zzz / DOI: 10.1002/

Assessing agreement with multiple raters on correlated kappa statistics: Web supplement

Hongyuan Cao[*,1], **Pranab K. Sen** [2], **Anne F. Peery** [3], and **Evan S. Dellon** [3]

[1] Department of Statistics, University of Missouri-Columbia, Columbia, MO 65211, U.S.A.
[2] Department of Biostatistics, University of North Carolina-Chapel Hill, Chapel Hill, NC 27514, U.S.A.
[3] Department of Medicine, University of North Carolina-Chapel Hill, Chapel Hill, NC 27514, U.S.A.

Abstract: This supplementary material includes the proof of Theorem 2.1. Our main tools are central limit theorem and multinomial distribution theories.

1 Proofs of Theorem 2.1

Proof. Define $Z_{ica} = n_{ica}/n - \pi_{ica}$ and $Z_{icb} = n_{icb}/n - \pi_{icb}$. These are mean 0 random variables with variances $\pi_{ica}(1-\pi_{ica})/n$ and $\pi_{icb}(1-\pi_{icb})/n$ respectively. We shall write p_o^a, p_o^b, p_e^a and p_e^b in terms of Z_{ica} and Z_{icb}.

Taking into account the fact that $\sum_{c=1}^2 n_{ica} = n$, we can write

$$1-p_o^a = \frac{n}{N(n-1)}\sum_{c=1}^{2}\sum_{i=1}^{N}\{-Z_{ica}^2 + (1-2\pi_{ica})Z_{ica} + \pi_{ica}(1-\pi_{ica})\} = A_a + B_a + C_a, \quad (1)$$

where

$$\begin{aligned} A_a &= \frac{1}{(n-1)N}\sum_{c=1}^{2}\sum_{i=1}^{N}[-nZ_{ica}^2 + \pi_{ica}(1-\pi_{ica})] = O_p(\frac{1}{n\sqrt{N}}), \\ B_a &= \frac{n}{N(n-1)}\sum_{c=1}^{2}\sum_{i=1}^{N} Z_{ica}(1-2\pi_{ica}), \\ C_a &= \frac{1}{N}\sum_{c=1}^{2}\sum_{i=1}^{N}\pi_{ica}(1-\pi_{ica}). \end{aligned}$$

Similar decomposition holds for $1-p_o^b$.

Since the subjects are independent, by central limit theorem (with Lyapunov Condition), there exists a positive σ_{1a}^2, such that

$$\sqrt{nN}B_a \to N(0,\sigma_{1a}^2),$$

where

$$\sigma_{1a}^2 = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\{\sum_{c=1}^{2}\pi_{ica}(1-\pi_{ica})(1-2\pi_{ica})^2 - \sum_{c\neq c'}\pi_{ica}\pi_{ic'a}(1-2\pi_{ica})(1-2\pi_{ic'a})\}.$$

*Corresponding author: e-mail: caohong@missouri.edu, Phone: +1-573-884-8568, Fax: +1-573-884-5524

Thus we have

$$U_a = \sqrt{nN}(1 - p_o^a - C_a) \to N(0, \sigma_{1a}^2).$$

Similarly

$$U_b = \sqrt{nN}(1 - p_o^b - C_b) \to N(0, \sigma_{1b}^2),$$

where

$$\sigma_{1b}^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \{\sum_{c=1}^{2} \pi_{icb}(1-\pi_{icb})(1-2\pi_{icb})^2 - \sum_{c\neq c'} \pi_{icb}\pi_{ic'b}(1-2\pi_{icb})(1-2\pi_{ic'b})\}.$$

For $1 - p_e^a$, we have

$$\begin{aligned} 1 - p_e^a &= 1 - \sum_{c=1}^{2} [\frac{1}{N}\sum_{i=1}^{N}(Z_{ica} + \pi_{ica})]^2 \\ &= 1 - \sum_{c=1}^{2} [\bar{Z}_{ca}^2 + 2\bar{\pi}_{ca}\bar{Z}_{ca} + \bar{\pi}_{ca}^2] \\ &= \sum_{c=1}^{2} [-\bar{Z}_{ca}^2 - 2\bar{\pi}_{ca}\bar{Z}_{ca} + \bar{\pi}_{ca}(1-\bar{\pi}_{ca})], \end{aligned}$$

where $\bar{Z}_{ca} = 1/N \sum_{i=1}^{N} Z_{ica}$ and $\bar{\pi}_{ca} = 1/N \sum_{i=1}^{N} \pi_{ica}$ for $c = 1, 2$. Note that

$$\bar{Z}_{ca}^2 = O_p(\frac{1}{nN})$$

and

$$\sqrt{nN}\bar{Z}_{ca} \to N(0, \sigma_{\pi a}^2),$$

where $\sigma_{\pi a}^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \pi_{ica}(1-\pi_{ica})$.
So asymptotically we have

$$V_a = \sqrt{nN}[1 - p_e^a - \sum_{c=1}^{2} \bar{\pi}_{ca}(1-\bar{\pi}_{ca})] \to N(0, \sigma_{2a}^2),$$

where

$$\sigma_{2a}^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \{\sum_{c=1}^{2} 4\bar{\pi}_{ca}^2 \pi_{ica}(1-\pi_{ica}) - \sum_{c\neq c'} 4\bar{\pi}_{ca}\bar{\pi}_{c'a}\pi_{ica}\pi_{ic'a}.$$

Similarly,

$$V_b = \sqrt{nN}[1 - p_e^b - \sum_{c=1}^{2} \bar{\pi}_{cb}(1-\bar{\pi}_{cb})] \to N(0, \sigma_{2b}^2),$$

where

$$\sigma_{2b}^2 = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \{\sum_{c=1}^{2} 4\bar{\pi}_{cb}^2 \pi_{icb}(1-\pi_{icb}) - \sum_{c\neq c'} 4\bar{\pi}_{cb}\bar{\pi}_{c'b}\pi_{icb}\pi_{ic'b}.$$

Recall

$$C_a^* = \sum_{c=1}^{2} \bar{\pi}_{ca}(1-\bar{\pi}_{ca}) \quad \text{and} \quad C_b^* = \sum_{c=1}^{2} \bar{\pi}_{cb}(1-\bar{\pi}_{cb}). \tag{2}$$

Note that

$$\begin{aligned}\sqrt{nN}\{\frac{1-p_o^a}{1-p_e^a}-\frac{C_a}{C_a^*}\} &= \frac{C_a^*}{1-p_e^a}\frac{\sqrt{nN}}{C_a^{*2}}[C_a^*(1-p_o^a)-C_a(1-p_e^a)]\\ &= \frac{C_a^*U_a-C_aV_a}{C_a^{*2}}\{1+O_p(\frac{1}{\sqrt{nN}})\}\end{aligned}$$

and

$$\begin{aligned}\sqrt{nN}\{\frac{1-p_o^b}{1-p_e^b}-\frac{C_b}{C_b^*}\} &= \frac{C_b^*}{1-p_e^b}\frac{\sqrt{nN}}{C_b^{*2}}[C_b^*(1-p_o^b)-C_b(1-p_e^b)]\\ &= \frac{C_b^*U_b-C_bV_b}{C_b^{*2}}\{1+O_p(\frac{1}{\sqrt{nN}})\}\end{aligned}$$

We need to get the asymptotic distribution of

$$\sqrt{nN}\{\frac{1-p_o^a}{1-p_e^a}-\frac{C_a}{C_a^*}-[\frac{1-p_o^b}{1-p_o^b}-\frac{C_b}{C_b^*}]\}=[\frac{C_a^*U_a-C_aV_a}{C_a^{*2}}-\frac{C_b^*U_b-C_bV_b}{C_b^{*2}}]+O_p\{(nN)^{-1/2}\}. \quad (3)$$

Recall that

$$U_a=\sqrt{nN}\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{2}Z_{ica}(1-2\pi_{ica})+O_p(\frac{1}{\sqrt{n}})$$

and

$$V_a=\sqrt{nN}\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{2}(-2\bar{\pi}_{ca}Z_{ica})+O_p(\frac{1}{\sqrt{nN}})$$

Therefore, we have

$$\begin{aligned}Var(C_a^*U_a-C_aV_a) &= C_a^{*2}Var(U_a)+C_a^2Var(V_a)-2C_a^*C_aCov(U_a,V_a)\\ &= 4C_a^{*2}\frac{1}{N}\sum_{i=1}^{N}\pi_{i1a}(1-\pi_{i1a})(1-2\pi_{i1a})^2\\ &+ 4C_a^2(1-2\bar{\pi}_{1a})^2\frac{1}{N}\sum_{i=1}^{N}\pi_{i1a}(1-\pi_{i1a})\\ &- 8C_aC_a^*(1-2\bar{\pi}_{1a})\frac{1}{N}\sum_{i=1}^{N}\pi_{i1a}(1-\pi_{i1a})(1-2\pi_{i1a})+O(\frac{1}{n}).\end{aligned}$$

By the same token,

$$\begin{aligned}Var(C_b^*U_b-C_bV_b) &= C_b^{*2}Var(U_b)+C_b^2Var(V_b)-2C_b^*C_bCov(U_b,V_b)\\ &= 4C_b^{*2}\frac{1}{N}\sum_{i=1}^{N}\pi_{i1b}(1-\pi_{i1b})(1-2\pi_{i1b})^2\\ &+ 4C_b^2(1-2\bar{\pi}_{1b})^2\frac{1}{N}\sum_{i=1}^{N}\pi_{i1b}(1-\pi_{i1b})\\ &- 8C_bC_b^*(1-2\bar{\pi}_{1b})\frac{1}{N}\sum_{i=1}^{N}\pi_{i1b}(1-\pi_{i1b})(1-2\pi_{i1b})+O(\frac{1}{n}).\end{aligned}$$

We make use of the fact that $Z_{i2a} = -Z_{i1a}, Z_{i2b} = -Z_{i1b}$,

$$\begin{aligned} E(Z_{i1a}Z_{i1b}) &= E(Z_{i2a}Z_{i2b}) = \frac{1}{n}(\theta_{i11}\theta_{i22} - \theta_{i12}\theta_{i21}), \\ E(Z_{i2a}Z_{i1b}) &= E(Z_{i1a}Z_{i2b}) = -\frac{1}{n}(\theta_{i11}\theta_{i22} - \theta_{i12}\theta_{i21}), \end{aligned}$$

and obtain

$$\begin{aligned} & Cov(C_a^*U_a - C_aV_a, C_b^*U_b - C_bV_b) \\ = \quad & C_a^*C_b^*\frac{1}{N}\sum_{i=1}^{N} 4(1-2\pi_{i1a})(1-2\pi_{i1b})[\theta_{i11} - \pi_{i1a}\pi_{i1b}] \\ - \quad & C_a^*C_b 4(1-2\bar{\pi}_{1b})\frac{1}{N}\sum_{i=1}^{N}(1-2\pi_{i1a})[\theta_{i11} - \pi_{i1a}\pi_{i1b}] \\ - \quad & C_aC_b^* 4(1-2\bar{\pi}_{1a})\frac{1}{N}\sum_{i=1}^{N}(1-2\pi_{i1b})[\theta_{i11} - \pi_{i1a}\pi_{i1b}] \\ + \quad & C_aC_b 4(1-2\bar{\pi}_{1a})(1-2\bar{\pi}_{1b})\frac{1}{N}\sum_{i=1}^{N}[\theta_{i11} - \pi_{i1a}\pi_{i1b}] + O(\frac{1}{n}). \end{aligned}$$

Combining above results, we have

$$\begin{aligned} V \quad = \quad & Var(\frac{C_a^*U_a - C_aV_a}{C_a^{*2}} - \frac{C_b^*U_b - C_bV_b}{C_b^{*2}}) \\ = \quad & \frac{1}{C_a^{*4}}\{4C_a^{*2}\frac{1}{N}\sum_{i=1}^{N}\pi_{i1a}(1-\pi_{i1a})(1-2\pi_{i1a})^2 + 4C_a^2(1-2\bar{\pi}_{1a})^2\frac{1}{N}\sum_{i=1}^{N}\pi_{i1a}(1-\pi_{i1a}) \\ - \quad & 8C_a^*C_a(1-2\bar{\pi}_{1a})\frac{1}{N}\sum_{i=1}^{N}\pi_{i1a}(1-\pi_{i1a})(1-2\pi_{i1a})\} \\ + \quad & \frac{1}{C_b^{*4}}\{4C_b^{*2}\frac{1}{N}\sum_{i=1}^{N}\pi_{i1b}(1-\pi_{i1b})(1-2\pi_{i1b})^2 + 4C_b^2(1-2\bar{\pi}_{1b})^2\frac{1}{N}\sum_{i=1}^{N}\pi_{i1b}(1-\pi_{i1b}) \\ - \quad & 8C_b^*C_b(1-2\bar{\pi}_{1b})\frac{1}{N}\sum_{i=1}^{N}\pi_{i1b}(1-\pi_{i1b})(1-2\pi_{i1b})\} \\ - \quad & \frac{1}{C_a^{*2}C_b^{*2}}\{8C_a^*C_b^*\frac{1}{N}\sum_{i=1}^{N}(1-2\pi_{i1a})(1-2\pi_{i1b})[\theta_{i11} - \pi_{i1a}\pi_{i1b}] \\ - \quad & 8C_a^*C_b(1-2\bar{\pi}_{1b})\frac{1}{N}\sum_{i=1}^{N}(1-2\pi_{i1a})[\theta_{i11} - \pi_{i1a}\pi_{i1b}] \\ - \quad & 8C_aC_b^*(1-2\bar{\pi}_{1a})\frac{1}{N}\sum_{i=1}^{N}(1-2\pi_{i1b})[\theta_{i11} - \pi_{i1a}\pi_{i1b}] \\ + \quad & 8C_aC_b\frac{1}{N}\sum_{i=1}^{N}(1-2\bar{\pi}_{1a})(1-2\bar{\pi}_{1b})[\theta_{i11} - \pi_{i1a}\pi_{i1b}]\} + O(\frac{1}{n}) \end{aligned}$$

Therefore

$$\Sigma = \lim_{n \to \infty, N \to \infty} V. \tag{4}$$

Combining (3) and (4), we complete the proof of Theorem 2.1. □