Case 1 : One-way ANOVA

Model: $y_{j(i)} = \underset{\substack{\uparrow \\ \text{ground} \\ \text{mean}}}{\mu} + \underset{\substack{\uparrow \\ \text{subject} \\ \text{effect}}}{S_i} + \underset{\substack{\uparrow \\ \text{error}}}{\varepsilon_{j(i)}}$, $\quad S_i \overset{iid}{\sim} N(0, \sigma_S^2) \perp\!\!\!\perp \varepsilon_{j(i)} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$

subject $i = 1, \dots, N$

rater $j = 1, \dots, M$

$j(i)$ = raters are nested within the subject

Reliability : Use

single scores / average scores

* Single scores

raters are exchageable

$$\rho_{yy'} = Cor(y_{j(i)}, y_{j'(i)})$$

$$= \frac{\sigma_S^2}{\sigma_S^2 + \sigma_\varepsilon^2} = ICC(1)$$

* average scores

$$\rho_{\bar{y}\bar{y}'} = Cor(\bar{y}_{\cdot(i)}, \bar{y}_{\cdot'(i)})$$

$$= \frac{\sigma_S^2}{\sigma_S^2 + \sigma_\varepsilon^2/M} = ICC(k = M)$$

Illustration:

* single score

	$R_1(1)$	$R_2(1)$	$R_3(1)$	$R_1(2)$	$R_2(2)$	$R_3(2)$	...
S_1							
S_2							
S_3							
⋮							

	R_j	$R_{j'}$
S_1		
S_2		
S_3		
	⋮	⋮

$= Cor(y_{\bar{j}(x)}, y_{\bar{j}'(x)})$

× 1000 times

- only use the partial information
- repeat this procedure many times and then take average

* average scores

$$\bar{y}_{\cdot(i)} = \frac{1}{m}\sum_{j=1}^{m} y_{j(i)}$$

$$\bar{y}'_{\cdot(i)} = \bar{y}_{\cdot(i)}' = ?$$

another group of raters?

$\boxed{\text{Ideally}}$

Imaging repeated measurement (Raters are changed)

	$R_{1(1)}$	$R_{2(1)}$	$R_{3(1)}$	$R_{1(2)}$	$R_{2(2)}$	$R_{3(2)}$	...
S_1	x	x	x				
S_2				x	x	x	
S_3							x x x
⋮							

ave

	$R'_{1(1)}$	$R'_{2(1)}$	$R'_{3(1)}$	$R'_{1(2)}$	$R'_{2(2)}$	$R'_{3(2)}$	...
S_1	x	x	x				
S_2				x	x	x	
S_3							x x x
⋮							

ave

	$\bar{R}_{\cdot}$	$\bar{R}'_{\cdot}$
S_1		
S_2		
S_3		
	⋮	⋮

$$= Cor(\bar{y}_{\cdot(i)}, \bar{y}'_{\cdot(i)})$$

× 1000 times

ave

ave

$\boxed{\text{In practice}}$, we only have a single assessment

	$R_{1(1)}$	$R_{2(1)}$	$R_{3(1)}$	$R_{1(2)}$	$R_{2(2)}$	$R_{3(2)}$	...
S_1	1	2	3				
S_2				1	2	3	
S_3							1 2 3
⋮							

=

	$R_{1(1)}$	$R_{2(1)}$	$R_{3(1)}$	$R_{1(2)}$	$R_{2(2)}$	$R_{3(2)}$	...
S_1	x	x	x				
S_2				x	x	x	
S_3							x x x
⋮							

☆ sampling with replacement

Same measurements (Raters are the same)

Case 2: Two-way Random Effect ANOVA

Model: $y_{ij} = \mu + S_i + R_j + \varepsilon_{ij}$

$$\perp\!\!\!\perp \begin{cases} S_i \overset{iid}{\sim} N(0, \sigma_S^2) \\ R_j \overset{iid}{\sim} N(0, \sigma_R^2) \\ \varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \end{cases}$$

subjects × raters

Reliability = Use × Decision

- Use: single scores (1), average score (k)
- Decision: agreement (A), consistence (C)

* single score

$$\rho_A = Cor(y_{ij}, y_{ij'})$$

$$= \frac{\sigma_S^2}{\sigma_S^2 + \sigma_R^2 + \sigma_\varepsilon^2} = ICC(A, 1)$$

$$\rho_C = Cor(y_{ij}, y_{ij'} \mid j, j')$$

$$= \frac{\sigma_S^2}{\sigma_S^2 + \sigma_\varepsilon^2} = ICC(C, 1)$$

$Cor(y_{ij} - R_j, y_{ij'} - R_{j'})$
[Winer's (1971) adjustment for anchor point]

* average score

$$\rho_{Ak} = Cor(\bar{y}_{i\cdot}, \bar{y}_{i\cdot'})$$

$$= \frac{\sigma_S^2}{\sigma_S^2 + \sigma_R^2/M + \sigma_\varepsilon^2/M} = ICC(A, k=M)$$

$$\rho_{ck} = \mathrm{Cor}\left(\bar{y}_{i\cdot}, \bar{y}_{i\cdot'} \mid \forall j, \forall j'\right)$$

$$= \frac{\sigma_\delta^2}{\sigma_\delta^2 + \sigma_\varepsilon^2/M} = ICC(C, k=M)$$

$\equiv$

Cronbach's alpha

Illustration

* single scores, agreement

	R_1	R_2	R_3	...
S_1				
S_2				
S_3				
⋮	⋮			⋱

	R_J	$R_{J'}$
S_1		
S_2		
S_3		
⋮	⋮	⋮

$$\| \quad \mathrm{Cor}(y_{\pi J}, y_{\pi J'})$$

× 1000 times

* single scores, consistence

	R_1	R_2	R_3	...
S_1				
S_2				
S_3				
⋮	⋮			⋱

	R_J	$R_{J'}$
S_1		
S_2		
S_3		
⋮	⋮	⋮

$$\| \quad \mathrm{Cor}(y_{\pi J}, y_{\pi J'} \mid J, J')$$

× 1000 times

* average scores, agreement

	R_1	R_2	R_3	...
S_1	✓✓ ×	✓ ×	×	...
S_2	×	✓✓✓ ×	×	...
S_3	✓ ×	✓ ×	✓ ×	...
⋮	⋮			⋱

	R_1	R_2	R_3	...
S_1	×	✓ ×	✓✓ ×	...
S_2	✓✓ ×	×	✓ ×	...
S_3	✓✓ ×	✓ ×	×	...
⋮	⋮			⋱

ave

ave

X 1000 times

	R_J	$R_{J'}$
S_1	②	④
S_2	②	④
S_3	②	④
⋮	⋮	⋮

$= Cor(\bar{y}_{\bar{s}\bullet}, \bar{y}_{\bar{s}\bullet'})$

X 1000 times

$= Cor\left(\bar{y}_{\bar{s}\bullet}, \bar{y}_{\bar{s}\bullet'} \mid \forall J, \forall J'\right)$

ave

ave

* average scores, consistency

	R_1	R_2	R_3	...
S_1	✓✓ ×	✓ ×	×	...
S_2	✓✓ ×	✓ ×	×	...
S_3	✓✓ ×	✓ ×	×	...
⋮	⋮			⋱

	R_1	R_2	R_3	...
S_1	×	✓✓ ×	✓ ×	...
S_2	×	✓✓ ×	✓ ×	...
S_3	>	✓✓ ×	✓ ×	...
⋮	⋮			⋱

Case 3: Two-way Mixed Effect ANOVA

Model: $y_{ij} = \mu + S_i + R_j + \varepsilon_{ij}$

- $S_i \overset{iid}{\sim} N(0, \sigma_S^2)$
- $\sum_{j=1}^{M} R_j = 0$, $\theta_R^2 \triangleq \frac{\sum R_j^2}{M-1}$
- $\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$

Reliability: Same ICCs used in Case 2. But...

$$Cov(y_{ij}, y_{ij'}) = ?$$

$$Var(y_{ij}) \begin{cases} \sigma_S^2 + \sigma_\varepsilon^2 \\ \sigma_S^2 + \theta_R^2 + \sigma_\varepsilon^2 \end{cases} ??$$

r.v. for ....

$$Var(Y_{Ij}) \text{ vs. } Var(Y_{IJ})$$

random sample for subjects

random sample for rater

$\boxed{\text{sample}}$

* Argue: $ICC_2(A,1) = ICC_3(A,1)$

$ICC_2(C,1) = ICC_3(C,1)$

Illustration: the same procedure in Case 2

Liljequist et. al., 2019

Model	Expected Mean Squares	Population ICC	Sample ICC
Model 1 $x_{ij} = \mu + r_i + v_{ij}$ ↑ ↑ random	$MSBS \approx k \cdot \sigma_r^2 + \sigma_v^2$ $MSBM \approx \sigma_v^2$ $MSWS \approx \sigma_v^2$ $MSWM \approx \sigma_r^2 + \sigma_v^2$ $MSE \approx \sigma_v^2$	$\rho_1 = \dfrac{\sigma_r^2}{\sigma_r^2 + \sigma_v^2}$ →	$ICC(1) = \dfrac{MSBS - MSWS}{MSBS + (k-1)MSWS}$
Model 2 $x_{ij} = \mu + r_i + c_j + v_{ij}$ ↑ ↑ ↑ random	$MSBS \approx k \cdot \sigma_r^2 + \sigma_v^2$ $MSBM \approx n \cdot \sigma_c^2 + \sigma_v^2$ $MSWS \approx \sigma_c^2 + \sigma_v^2$ $MSWM \approx \sigma_r^2 + \sigma_v^2$ $MSE \approx \sigma_v^2$	Agreement $\rho_{2A} = \dfrac{\sigma_r^2}{\sigma_r^2 + \sigma_c^2 + \sigma_v^2}$ → Consistency $\rho_{2C} = \dfrac{\sigma_r^2}{\sigma_r^2 + \sigma_v^2}$ →	$ICC(A,1) = \dfrac{MSBS - MSE}{MSBS + (k-1)MSE + \left(\dfrac{k}{n}\right)(MSBM - MSE)}$ $ICC(C,1) = \dfrac{MSBS - MSE}{MSBS + (k-1)MSE}$
Model 3 fixed ↓ $x_{ij} = \mu + r_i + c_j + v_{ij}$ ↑ ↑ random	$MSBS \approx k \cdot \sigma_r^2 + \sigma_v^2$ $MSBM \approx n \cdot \theta_c^2 + \sigma_v^2$ $MSWS \approx \theta_c^2 + \sigma_v^2$ $MSWM \approx \sigma_r^2 + \sigma_v^2$ $MSE \approx \sigma_v^2$ $\theta_C^2 = \dfrac{\sum (c_j - \bar{c})^2}{k-1}$	Agreement $\rho_{3A} = \dfrac{\sigma_r^2}{\sigma_r^2 + \theta_c^2 + \sigma_v^2}$ → Consistency $\rho_{3C} = \dfrac{\sigma_r^2}{\sigma_r^2 + \sigma_v^2}$ →	$ICC(A,1) = \dfrac{MSBS - MSE}{MSBS + (k-1)MSE + \left(\dfrac{k}{n}\right)(MSBM - MSE)}$ $ICC(C,1) = \dfrac{MSBS - MSE}{MSBS + (k-1)MSE}$

Fig 1. Survey of models, mean square relations and ICC formulas.

https://doi.org/10.1371/journal.pone.0219854.g001