proLékaře.cz | 20.5.2024

Analýza dat v neurologii

LXXIX. Koeficient vnitrotřídní korelace

L. Dušek, T. Pavlík, J. Jarkovský, K. Pokorová, J. Koptíková

Institut biostatistiky a analýz, LF MU, Brno

prof. RNDr. Ladislav Dušek, Ph.D.
Institut biostatistiky a analýz, LF MU, Brno
e-mail: dusek@iba.muni.cz

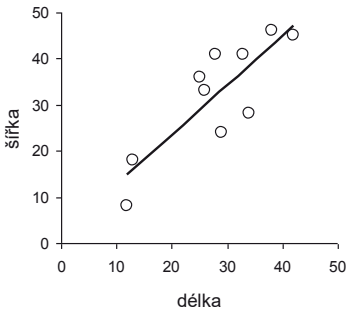
Tímto dílem seriálu otevíráme poslední část rozboru korelačních analýz, a to výkladem významu tzv. koeficientu vnitrotřídní korelace (intraclass correlation coefficient; ICC). Ačkoli tato statistika nese přímo v názvu slovo korelace, její význam je poněkud odlišný od klasických korelačních koeficientů. Korelací obecně myslíme míru síly a významnosti vztahu dvou nebo více kvantitativních (spojitých) proměnných. Analýza ICC je naopak základním nástrojem pro měření spolehlivosti (konzistence, stability) opakovaných měření či experimentů. ICC tedy nejčastěji vyjadřuje shodu několika opakovaných měření, jinými slovy jejich homogenitu či „reliabilitu".

Ačkoli zde jde o jiný cíl než u klasické korelační analýzy, od tématu korelace nijak neodbočíme. Základní princip posuzování spolehlivosti měření je velmi jednoduchý a skutečně směřuje ke korelační analýze. Pokud totiž v určitém experimentálním plánu opakujeme stejná měření, měli bychom logicky mezi po sobě jdoucími pokusy dosáhnout vysoké korelace. Tato však nemusí být absolutní z důvodu náhodné variability výsledků opakovaných měření. Odhad hodnoty ICC umožní posoudit míru vzájemné podobnosti takových měření a případně rozhodnout o výběru těch nejspolehlivějších.

Význam odhadu ICC je však ještě širší a nemusí se omezovat pouze na série opakovaných měření prováděných za stejných podmínek. ICC může obecně posoudit vzájemnou podobnost či homogenitu jakýchkoli kvantitativních hodnot spojité ná-

1a.
Modelová data pro klasickou korelační analýzu proměnných *X* a *Y* v párově uspořádaném experimentu.

Číslo hodnocené léze	*X* délka (mm)	*Y* šířka (mm)
1	13	18
2	29	24
3	33	41
4	42	45
5	12	8
6	26	33
7	34	28
8	38	46
9	25	36
10	28	41



V experimentu sledujeme korelaci mezi šířkou a délkou nádorových lézí při diagnostickém vyšetření. Jde o klasickou korelační analýzu, která pracuje se dvěma rozdílnými proměnnými (*X* a *Y*) a vychází přitom z párově uspořádaného experimentu; dvojice hodnot *X* a *Y* jsou tak vždy měřeny na témže subjektu.

1b.
Modelová data pro hodnocení vnitrotřídní korelace v datovém souboru rozděleném do skupin.

Datový soubor o *n* = 30 měřeních proměnné *X* (objem léze v mm³). Hodnoty *X* jsou rozdělené do tří skupin: jde o měření třemi různými metodami – hodnocena je jedna léze vždy 10× každou metodou.

skupina 1	skupina 2	skupina 3
14	17	22
9	15	19
15	19	21
17	18	18
18	29	35
20	36	34
10	14	18
22	50	48
17	22	24
13	15	18

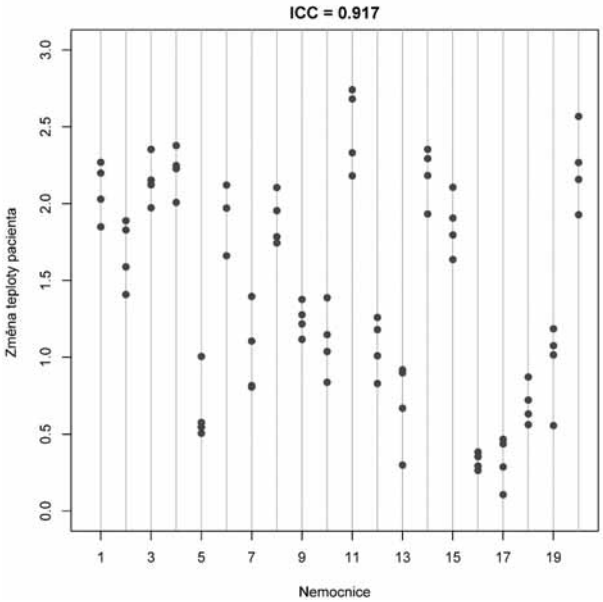
Experiment posuzuje konzistenci odhadů objemu nádorových lézí prováděných třemi různými zobrazovacími metodami. Na rozdíl od klasické korelační analýzy zde pracujeme pouze s jednou spojitou proměnnou *X*. Hodnoty ve skupinách jsou datové série z 10× nezávisle opakovaných měření jedné léze. Vnitrotřídní korelace zde umožní posoudit homogenitu (vzájemnou podobnost) hodnot uvnitř skupin. Podobné uspořádání by mohlo být i párové, v takovém případě by řádky představovaly např. 10 hodnocených různých lézí a skupiny by byly tvořeny opět třemi různými metodami nebo třemi různými hodnotiteli.

Příklad 1. Schematické znázornění primárních dat pro hodnocení vnitrotřídní korelace – srovnání rozdílů oproti klasické korelační analýze.

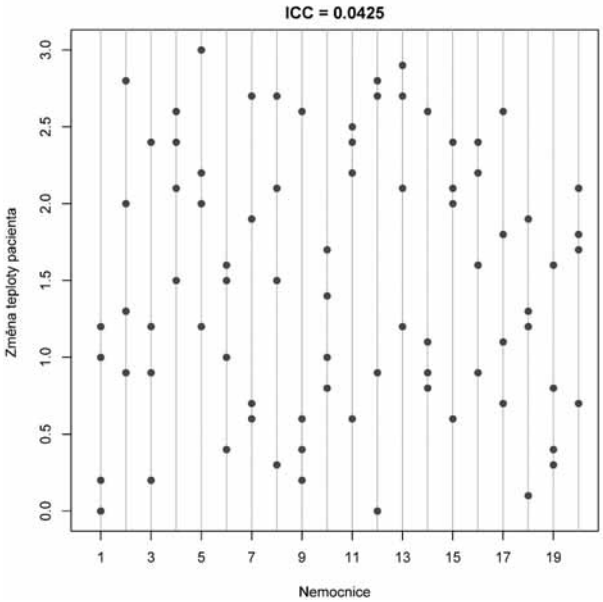
proLékaře.cz | 20.5.2024

Příklad zobrazuje dva datové soubory o *n* pacientech, u kterých sledujeme v klinické studii změnu tělesné teploty po určité terapii. Jak vyplývá z grafů, změna byla pozorována v rozsahu od 0 do 3 °C. Do klinické studie bylo zapojeno více nemocnic a každá zařadila 4 pacienty. Smyslem analýzy vnitrotřídní korelace je posoudit, zda jsou hodnoty v rámci nemocnic konzistentní, a tedy vzájemně si podobné. Vnitřní homogenita hodnot v rámci nemocnic by totiž mohla představovat riziko zkreslení anebo naopak by mohla odhalit postupy, které v určité nemocnici vedou ke konzistentní reakci pacientů na danou léčbu. Příklad 2a ukazuje výsledky vedoucí na velmi vysokou hodnotu ICC – je zde patrný silný vliv nemocnice jako třídícího faktoru a reakce pacientů v rámci nemocnic mají tendenci se vzájemně podobat. Naopak příklad 2b zobrazuje data s velmi vysokou vnitroskupinovou variabilitou, která překrývá rozdíly mezi nemocnicemi.

2a.
Datový soubor s vvsokou hodnotou ICC.



2b.
Datový soubor s nízkou hodnotou ICC.



Příklad 2. Schematické znázornění primárních dat vedoucích k vysoké a naopak nízké hodnotě koeficientu vnitrotřídní korelace (ICC).

hodné veličiny seskupených do určitých skupin.

Typický experiment vedoucí k výpočtu vnitrotřídní korelace sleduje kvantitativní znak měřený u jedinců (subjektů) uspořádaných do skupin. Výpočet ICC hodnotí, jak hodně se hodnoty náležející do stejné skupiny vzájemně podobají. Tuto situaci ukazuje příklad 1, ze kterého je také patrný principiální rozdíl mezi ICC a klasickou korelací. Běžné korelační koeficienty vycházejí z dat, kdy jsou korelované proměnné měřeny u stejných subjektů a jsou znázornitelné v jednoduchém korelačním bodovém grafu. Klasická korelace primárně vychází z párově uspořádaných souborů, kde jsou hodnoty korelovaných proměnných vždy přiřazeny stejnému jedinci. Korelace typicky vyjadřuje vztah dvou rozdílných spojitých náhodných veličin, např. výšky (*X*) a hmotnosti lidské postavy (*Y*). Naopak ICC pracuje s hodnotami jediné proměnné (*X*) uspořádanými do skupin v rámci jednoho datového souboru. Proto ICC bývá někdy v literatuře označován za koeficient skupinové korelace.

Pro výpočet ICC lze v učebnicích nalézt různé vztahy, avšak pro základní pochopení se nám jeví jako nejlepší výpočet pracující s variabilitou opakovaných měření proměnné *X*. Koeficient vnitrotřídní korelace charakterizuje podobnost hodnot měřené veličiny ve skupinách (třídách, sériích) dat v porovnání s variabilitou celého souboru, ve kterém jsou tyto skupiny spojeny do celku. Lze jej tedy jednoduše definovat jako poměr variability mezi těmito skupinami a variability v celém souboru:

$$\mathrm{ICC} = \frac{var(B)}{var(B) + var(W)}, \text{ kde}$$

var(B) znamená variabilitu mezi skupinami či datovými sériemi (between groups; B) a *var(W)* značí variabilitu uvnitř skupin (within groups; W).

Obě složky variability lze sečíst a jmenovatel tak vyjadřuje hodnotu celkové variability spojeného datového souboru. Výše uvedený vztah můžeme též přepsat pro populační hodnotu ICC (hodnotu v cílové populaci) takto:

$$\mathrm{ICC} = \frac{\sigma^2_\beta}{\sigma^2_\beta + \sigma^2_E}, \text{ kde}$$

σ^2_β je rozptyl hodnot způsobený vlivem roztřídění do skupin (meziskupinová variabilita) a σ^2_E je rozptyl hodnot uvnitř těchto skupin (vnitroskupinová variabilita).

Pouze pro úplnost tohoto výkladu zde přiblížíme další výpočet ICC pomocí nástrojů analýzy rozptylu (viz též příklad 5). Pro výpočet *var(B)* a *var(W)* z primárních dat ve formátu tabulky naměřených hodnot proměnné *X* o *r* řádcích (*r* vyjadřuje hodnocené skupiny) a *s* sloupcích (měření v rámci skupiny) využijeme vztahu:

$$var(B) = MS_R - MS_E \text{ a } var(W) = s \times MS_E, \text{ kde}$$

$$MS_R = \frac{S_R}{r-1}, MS_E = \frac{S_E}{(r-1)(s-1)} \text{ a}$$

$$S_R = \Sigma_{i=1}^{r} \Sigma_{j=1}^{s} (\overline{x}_i - \overline{x})^2, S_C = \Sigma_{i=1}^{r} \Sigma_{j=1}^{s} (\overline{x}_j - \overline{x})^2,$$

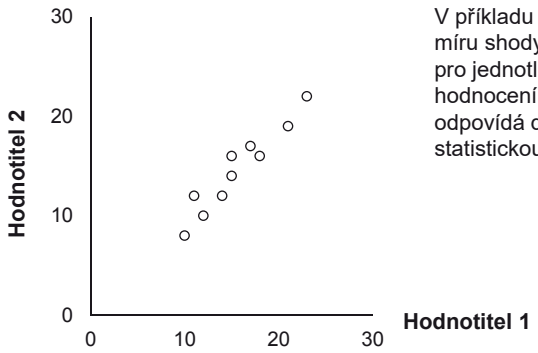
$$S_T = \Sigma_{i=1}^{r} \Sigma_{j=1}^{s} (x_{ij} - \overline{x})^2 \text{ a } S_E = S_T - S_R - S_C.$$

proLékaře.cz | 20.5.2024

Nejběžnějším případem výpočtu ICC je zjištění reliability (vzájemné spolehlivosti) hodnocení stejných subjektů (pacientů) dvěma či více hodnotiteli. Příklady níže popisují situaci s vysokým a nízkým ICC.

3a. Data vedoucí k vysoké hodnotě ICC.

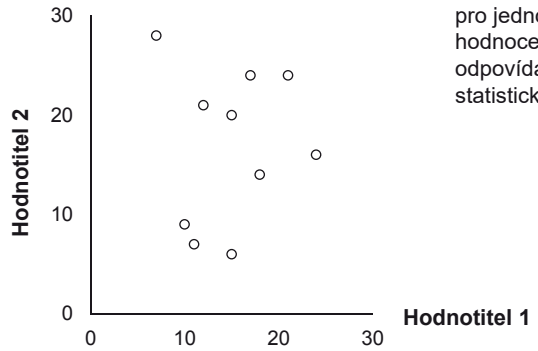
	Hodnotitel 1	Hodnotitel 2
Pacient 1	10	8
Pacient 2	12	10
Pacient 3	14	12
Pacient 4	15	16
Pacient 5	17	17
Pacient 6	21	19
Pacient 7	23	22
Pacient 8	15	14
Pacient 9	18	16
Pacient 10	11	12



V příkladu 3a pozorujeme vysokou míru shody obou hodnotitelů, ICC pro jednotlivá měření (vyjadřuje, zda hodnocení jednoho hodnotitele odpovídá druhému) je 0,957 se statistickou významností $p < 0{,}001$.

3b. Data vedoucí k nízké hodnotě ICC.

	Hodnotitel 1	Hodnotitel 2
Pacient 1	10	9
Pacient 2	12	21
Pacient 3	24	16
Pacient 4	15	6
Pacient 5	17	24
Pacient 6	21	24
Pacient 7	7	28
Pacient 8	15	20
Pacient 9	18	14
Pacient 10	11	7



V příkladu 3b pozorujeme nízkou míru shody obou hodnotitelů, ICC pro jednotlivé měření (vyjadřuje, zda hodnocení jednoho hodnotitele odpovídá druhému) je 0,046 se statistickou významností $p = 0{,}447$.

Příklad 3. Koeficient vnitrotřídní korelace (ICC) srovnávající dva hodnotitele.

Těmito výpočty ovšem nechceme čtenáře odradit od dalšího čtení. Pouze tak dokumentujeme, že ICC je možné kalkulovat jednoduchým dosazením primárních dat. To ovšem dnes již není nutné, výpočet ICC provede téměř každý statistický software. Při použití těchto nástrojů je však nutné znát podstatu výpočtu a volit kalkulaci odpovídající nastavení experimentu.

Výpočet ICC tedy provádíme přes odhady rozptylu měřené veličiny *X*, přičemž se rozlišuje rozptyl uvnitř jednotlivých tříd (sérií měření) a rozptyl celkový, kalkulovaný na spojeném souboru dat. ICC vyjadřuje homogenitu či nehomogenitu rozptylu *X* v jednotlivých třídách. Z výše uvedených vztahů vyplývají některé podstatné vlastnosti ICC, které shrnujeme v následujících bodech a dokumentujeme na příkladu 2:

- ICC nabývá hodnot od 0 do 1. Na rozdíl od klasické korelace tedy ICC nemůže nabývat záporných hodnot. Maximální teoreticky dosažitelná hodnota ICC je +1, které koeficient dosáhne při nulové variabilitě (rozptylu) hodnot *X* uvnitř skupin (všechna měření v rámci skupin jsou identická, mají stejnou hodnotu *X*). Datová sada vykazuje úplnou homogenitu hodnot uvnitř skupin a veškerá variabilita souboru je asociována s rozdíly mezi skupinami. Taková situace je samozřejmě extrémní a v praxi krajně nepravděpodobná, nicméně ilustruje fakt, že rostoucí hodnota ICC značí rostoucí podíl variability mezi hodnocenými skupinami na celkové variabilitě hodnot *X* v datovém souboru. Neboli s rostoucí hodnotou ICC roste konzistence (podobnost) měření v rámci skupin.
- Výpočet ICC vychází z analýzy rozptylu, nikoli ze středních hodnot rozdělení proměnných. Výpočet probíhá na základě rozkladu celkového rozptylu datového souboru na složku vnitroskupinovou (hodnotí rozdíly, variabilitu, mezi jedinci či subjekty měření) a meziskupinovou (variabilita mezi skupinami, které představují třídění souboru do kategorií vyšší úrovně).
- Hodnotu ICC lze vynásobit 100 a vyjádřit ji tak v %. Jde o % celkové variability datového souboru, které připadá na rozdíly mezi skupinami.
- Příklad 2a ilustruje experiment s vysokou hodnotou ICC, kde se hodnoty *X* v rámci skupin vzájemně podobají a podíl variability mezi skupinami na celkovém rozptylu hodnot v datovém souboru je vysoký. Příklad 2b naopak zobrazuje situaci s vysokou variabilitou hodnot *X* uvnitř skupin, které se od sebe v důsledku toho vzájemně nijak neodlišují. Hodnoty nejsou v rámci skupin homogenní a hodnota ICC je nízká.

Interpretaci hodnot ICC samozřejmě určují zejména design a cíl experimentu, tedy za jakým účelem jsou hodnoty roztříděny do skupin a proč rozdíly mezi skupinami sledujeme:

- Jedním z modelů je rozdělení datového souboru do experimentálních skupin, u kterých očekáváme vliv na variabilitu hodnot *X*. Při studiu charakteristik nemocí to mohou být např. geografická území, sociální kategorie obyvatel nebo typ domác-

proLékaře.cz | 20.5.2024

4a. Data vedoucí k vysoké hodnotě ICC.

	Hodnotitel 1	Hodnotitel 2	$\bar{x}_{i.}$
Pacient 1	10	8	**9**
Pacient 2	12	10	**11**
Pacient 3	14	12	**13**
Pacient 4	15	16	**15,5**
Pacient 5	17	17	**17**
Pacient 6	21	19	**20**
Pacient 7	23	22	**22,5**
Pacient 8	15	14	**14,5**
Pacient 9	18	16	**17**
Pacient 10	11	12	**11,5**
$\bar{x}_{.j}$	**15,6**	**14,6**	$\bar{x} = 15,1$

Tabulka má *r* = 10 řádků (počet hodnocených pacientů) a *s* = 2 sloupce (počet hodnotitelů).

Výpočet součtů čtverců odchylek:

$$S_R = \sum_{i=1}^{r}\sum_{j=1}^{s}(\bar{x}_{i.} - \bar{x})^2 = 315{,}8;\ S_C = \sum_{i=1}^{r}\sum_{j=1}^{s}\left(\bar{x}_{.j} - \bar{x}\right)^2 = 5;$$

$$S_T = \sum_{i=1}^{r}\sum_{j=1}^{s}\left(x_{ij} - \bar{x}\right)^2 = 327{,}8;\ S_E = S_T - S_R - S_C = 7.$$

Vztah pro řádkové průměrné čtverce:

$$MS_R = \frac{S_R}{r-1} = \frac{315{,}8}{9} = 35{,}0889.$$

Vztah pro průměrné čtverce uvnitř skupin:

$$MS_E = \frac{S_E}{(r-1)(s-1)} = \frac{7}{9 * 1} = 0{,}7778.$$

Hodnota vnitrotřídního korelačního koeficientu:

$$ICC = \frac{MS_R - MS_E}{MS_R + (s-1)MS_E} = \frac{35{,}0889 - 0{,}7778}{35{,}0889 + 1 * 0{,}7778} = 0{,}9566.$$

95% IS pro ICC: ⟨0,836, 0,989⟩, *p*-hodnota < 0,001 (tzn. na hladině významnosti 0,05 zamítáme nulovou hypotézu ICC = 0 ve prospěch alternativy ICC > 0).

4b. Data vedoucí k nízké hodnotě ICC.

	Hodnotitel 1	Hodnotitel 2	$\bar{x}_{i.}$
Pacient 1	10	9	**9,5**
Pacient 2	12	21	**16,5**
Pacient 3	24	16	**20**
Pacient 4	15	6	**10,5**
Pacient 5	17	24	**20,5**
Pacient 6	21	24	**22,5**
Pacient 7	7	28	**17,5**
Pacient 8	15	20	**17,5**
Pacient 9	18	14	**16**
Pacient 10	11	7	**9**
$\bar{x}_{.j}$	**15**	**16,9**	$\bar{x} = 15,95$

Tabulka má *r* = 10 řádků (počet hodnocených pacientů) a *s* = 2 sloupce (počet hodnotitelů).

Výpočet součtů čtverců odchylek:

$$S_R = \sum_{i=1}^{r}\sum_{j=1}^{s}(\bar{x}_{i.} - \bar{x})^2 = 409{,}45;\ S_C = \sum_{i=1}^{r}\sum_{j=1}^{s}\left(\bar{x}_{.j} - \bar{x}\right)^2 = 18{,}05;$$

$$S_T = \sum_{i=1}^{r}\sum_{j=1}^{s}\left(x_{ij} - \bar{x}\right)^2 = 800{,}95;\ S_E = S_T - S_R - S_C = 373{,}45.$$

Vztah pro řádkové průměrné čtverce:

$$MS_R = \frac{S_R}{r-1} = \frac{409{,}45}{9} = 45{,}4944.$$

Vztah pro průměrné čtverce uvnitř skupin:

$$MS_E = \frac{S_E}{(r-1)(s-1)} = \frac{373{,}45}{9 * 1} = 41{,}4944.$$

Hodnota vnitrotřídního korelačního koeficientu:

$$ICC = \frac{MS_R - MS_E}{MS_R + (s-1)MS_E} = \frac{45{,}4944 - 41{,}4944}{45{,}4944 + 1 * 41{,}4944} = 0{,}046.$$

95% IS pro ICC: ⟨−0,572, 0,631⟩, *p*-hodnota = 0,447 (tzn. na hladině významnosti 0,05 nezamítáme nulovou hypotézu ICC = 0).

Příklad 4. Modelová ukázka výpočtu hodnoty koeficientu vnitrotřídní korelace (ICC) doplněná odhadem 95% intervalu spolehlivosti a statistickými testy (data z příkladu 3).

ností. Jde o objektivně dané kategorie, v rámci kterých má smysl zkoumat stupeň homogenizace hodnot *X*.

- Jiný modelový příklad představuje experiment sledující vnitro- či meziskupinovou opakovatelnost měření. V takovém modelu generují skupiny hodnot *X* jednotliví pozorovatelé či experimentátoři, kteří hodnoty *X* opakovaně měří. ICC vyjadřuje vnitřní konzistenci jejich výsledků. ICC je využitelný jako míra opakovatelnosti současně pro více než dvě série měření (neomezeně ve výpočtu můžeme zahrnout 1, …, *s* pozorovatelů). Vysoká hodnota ICC blízká 1 zde znamená vysokou podobnost opakovaných měření jednotlivých pozorovatelů. Naopak hodnota ICC blízká nule ukazuje na situaci, kdy opakovaná měření získaná od téhož pozorovatele si nejsou podobná.

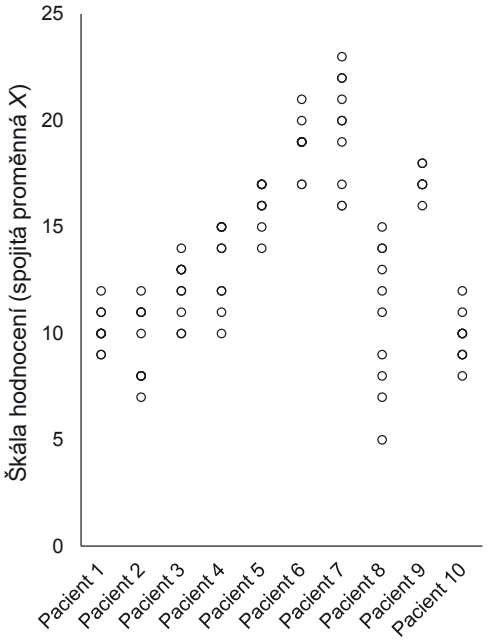
Vzorová data a hodnocení vnitrotřídní korelace dokumentují příklady 3–5. Z uvedených postupů vyplývá, že bodový odhad hodnoty ICC lze standardně doplnit odhadem intervalu spolehlivosti a též testem statistické významnosti. Nulovou hypotézou v těchto testech je předpoklad nulové hodnoty ICC, přičemž statistická významnost znamená hodnotu ICC prokazatelně rozdílnou od nuly. Avšak kromě těchto standardních statistických postupů se v literatuře velmi často hodnotí přímo absolutní hodnota ICC, zejména u testů konzistence v opakovaných měřeních. Hodnoty ICC menší než 0,5 značí nízkou konzistenci, ICC v intervalu 0,5–0,75 jsou označovány za mírnou až střední konzistenci, ICC nad 0,75 indikuje vysokou konzistenci a hodnoty koeficientu nad 0,9 značí excelentní, vysoce průkaznou konzistenci měření.

Příklady 2–5 záměrně ukazují různé modelové situace, pro které má smysl hodnotit vnitrotřídní korelaci. Snad se nám podařilo doložit, že možnosti aplikace ICC jsou velmi různorodé a v podstatě nelze uvést nějaké univerzální pravidlo pro interpretaci hodnot ICC. Vždy jde o ukazatel homogenity hodnot měřených v rámci určitých skupin. A průkaz této homogenity může mít jak pozitivní, tak negativní interpretaci dle smyslu a cíle experimentu. Např. prokážeme-li vysokou homogenitu a konzistenci měření různých pozorovatelů, je to jistě pozitivní zpráva pro danou testovanou metodu, neboť ta vede k opakovatelným, a tedy spolehlivým výsledkům. Avšak příklad 2 tohoto dílu seriálu ukazuje klinickou studii, kde lze průkaz homogenity považovat za indikaci možného rizika zkreslení. Pokud pacienty zařazené do multicentrické studie rozdělíme do skupin dle nemocnice, kde byli léčeni, pak nemocnice představují skupiny, tedy vyšší úroveň třídění souboru. Pokud se v měřeném znaku (léčebná odpověď, výskyt komplikací, doba hojení po operaci apod.) prokáže vysoká vzájemná podobnost pacientů v rámci nemocnic, může to znamenat, že vliv konkrétních center na výsledky studie není náhodný a může být zdrojem zkreslení. Příklad 2a ukazuje, že se hodnocená změna tělesné teploty pacientů v rámci konkrétních nemocnic vzájemně podobá a některé nemocnice čelí

proLékaře.cz | 20.5.2024

Na rozdíl od běžného Pearsonova korelačního koeficientu (spadajícího mezi tzv. mezitřídní korelační koeficienty, „interclass correlation coefficient“) není odhad ICC omezen pouze na dvojice čísel, ale je možné jej analyzovat pro větší počet hodnotitelů. V příkladu je zahrnuto 10 pacientů hodnocených prostřednictvím 10 hodnotitelů. Uspořádání je tedy párové, každý pacient je hodnocen všemi hodnotiteli. Výsledek hodnocení je znázorněn na obrázku, kde datová řada pro každého pacienta představuje skupinu sdružující hodnocení daného pacienta hodnotiteli; výsledek jednotlivých hodnocení znázorňují body v grafu.

5a. Grafické znázornění vstupních dat.



Výpočtem můžeme získat dvě varianty ICC:
i) pro reliabilitu hodnotitele samostatně (ICC = 0,896)
ii) průměrnou reliabilitu hodnotitelů (ICC = 0,989).

Druhá varianta je využitelná pouze v případě, kdy uvažujeme, že všichni hodnotitelé budou průměrováni do výsledné hodnoty (tedy *s* = 1); tato hodnota je vždy větší než ICC pro hodnotitele samostatně. Za běžných okolností nás však zejména zajímá, zda hodnocení jednoho hodnotitele odpovídá hodnocení ostatních, tedy varianta první.

Za tohoto předpokladu je výsledné ICC = 0,896 s $p < 0.001$ znamenající silnou korelaci mezi hodnotiteli.

Příklad 5b. Rozbor vstupních dat.

	Hodnotitel 1	Hodnotitel 2	Hodnotitel 3	Hodnotitel 4	Hodnotitel 5	Hodnotitel 6	Hodnotitel 7	Hodnotitel 8	Hodnotitel 9	Hodnotitel 10
Pacient 1	10	11	10	9	9	10	10	10	11	12
Pacient 2	12	11	11	11	10	8	8	8	8	7
Pacient 3	14	13	13	13	13	12	12	10	10	11
Pacient 4	15	15	15	15	14	14	12	12	11	10
Pacient 5	17	17	16	17	17	17	16	16	15	14
Pacient 6	21	20	19	19	19	19	19	19	17	17
Pacient 7	23	22	22	21	20	20	19	17	16	16
Pacient 8	15	14	14	13	12	11	9	8	7	5
Pacient 9	18	17	17	17	18	17	16	17	17	17
Pacient 10	11	12	10	10	10	10	9	9	9	8

i) Reliabilita hodnotitele samostatně

$$ICC = \frac{MS_R - MS_E}{MS_R + (s-1)MS_E} = \frac{153{,}7289 - 1{,}7585}{153{,}7289 + 9 * 1{,}7585} = 0{,}8963$$

95% IS pro ICC: ⟨0,789; 0,968⟩
p-hodnota* < 0,001

ii) Průměrná reliabilita hodnotitelů

$$ICC = \frac{MS_R - MS_E}{MS_R} = \frac{153{,}7289 - 1{,}7585}{153{,}7289} = 0{,}9886$$

95% IS pro ICC: ⟨0,974; 0,997⟩
p-hodnota* < 0,001

Vztah výpočtu k obecnému vztahu $ICC = \frac{var(B)}{var(W) + var(B)}$

1) Pro *var(B)* a *var(W)* platí:
var(B) = $MS_R - MS_E$ a *var(W)* = $s \times MS_E$

2) Po dosazení:

$$ICC = \frac{var(B)}{var(W) + var(B)} = \frac{MS_R - MS_E}{s \times MS_E + MS_R - MS_E} = \frac{MS_R - MS_E}{MS_R + (s-1)MS_E}$$

* F test

Příklad 5. Koeficient vnitrotřídní korelace (ICC) srovnávající více hodnotitelů.

proLékaře.cz | 20.5.2024

konzistentně vyššímu nárůstu jejích hodnot. Naopak výsledky studie zobrazené v příkladu 2b nejsou zatíženy rozdíly mezi nemocnicemi a variabilita jedinců v rámci nemocnic odpovídá žádoucímu náhodnému chování, nezávislému na konkrétním centru.

Na závěr tohoto dílu si dovolíme čtenářům doporučit velmi přínosné pojednání rozebírající význam různých analýz opakovatelnosti a spolehlivosti experimentálních měření, které v roce 2016 vydali autoři Koo a Li. Jistě zde zaujme i velmi recentní datum této práce, které svědčí o tom, že problematika měření spolehlivosti je v klinickém výzkumu stále aktuální a živá. Je to až s podivem, když uvážíme, že koncept analýzy rozptylu do skupin tříděných dat, a tedy i koncept výpočtu ICC, vznikl z pera R. A. Fishera v 50. letech minulého století (Fisher 1954) a nejvýznamnější primární práce k hodnocení spolehlivosti opakovaných měření byly publikovány v druhé polovině minulého století (Bartko 1966; Shrout a Fleiss 1979; Bland a Altman 1986). Přesto se k této problematice více či méně kriticky vrací řada současných odborníků na metodiku klinického výzkumu (např. Bruton et al 2000; Field 2005; Lee et al 2012). Jejich práce potvrzují, že spolehlivost a opakovatelnost sledování mají zejména v klinické praxi zásadní význam. K této problematice se tedy vrátíme ještě v dalších dílech seriálu.

Literatura

Bartko JJ. The intraclass correlation coefficient as a measure of reliability. Psychol Rep 1966; 19(1): 3–11. doi: 10.2466/pr0.1966.19.1.3.

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1(8476): 307–310.

Bruton A, Conway JH, Holgate ST. Reliability: what is it, and how is it measured? Physiotherapy 2000; 86(2): 94–99. doi: 10.1016/S0031-9406(05)61211-4.

Field AP. Intraclass correlation. In: Everitt BS, Howell DC. Encyclopedia of statistics in behavioral sciences. Chichester: Wiley 2005.

Fisher RA. Statistical methods for research workers. Edinburgh: Oliver and Boyd 1954.

Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Ciropr Med 2016; 15(2): 155–163.

Lee KM, Lee J, Chung CY et al. Pitfalls and important issues in testing reliability using intraclass correlation coefficients in orthopaedic research. Clin Orthop Surg 2012; 4(2): 149–155. doi: 10.4055/cios.2012.4.2.149.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979; 86(2): 420–428. doi: 10.1037//0033-2909.86.2.420.

Na webu csnn.eu naleznete aktualizované pokyny pro autory.