MUNI FSS Katedra psychologie

Přednáška 3 (samostudium): Chyba měření a intervaly spolehlivosti

28. 9. 2021 | PSYn4790 | **Psychometrika: Měření v psychologii**
Katedra psychologie, Fakulta sociálních studií MU

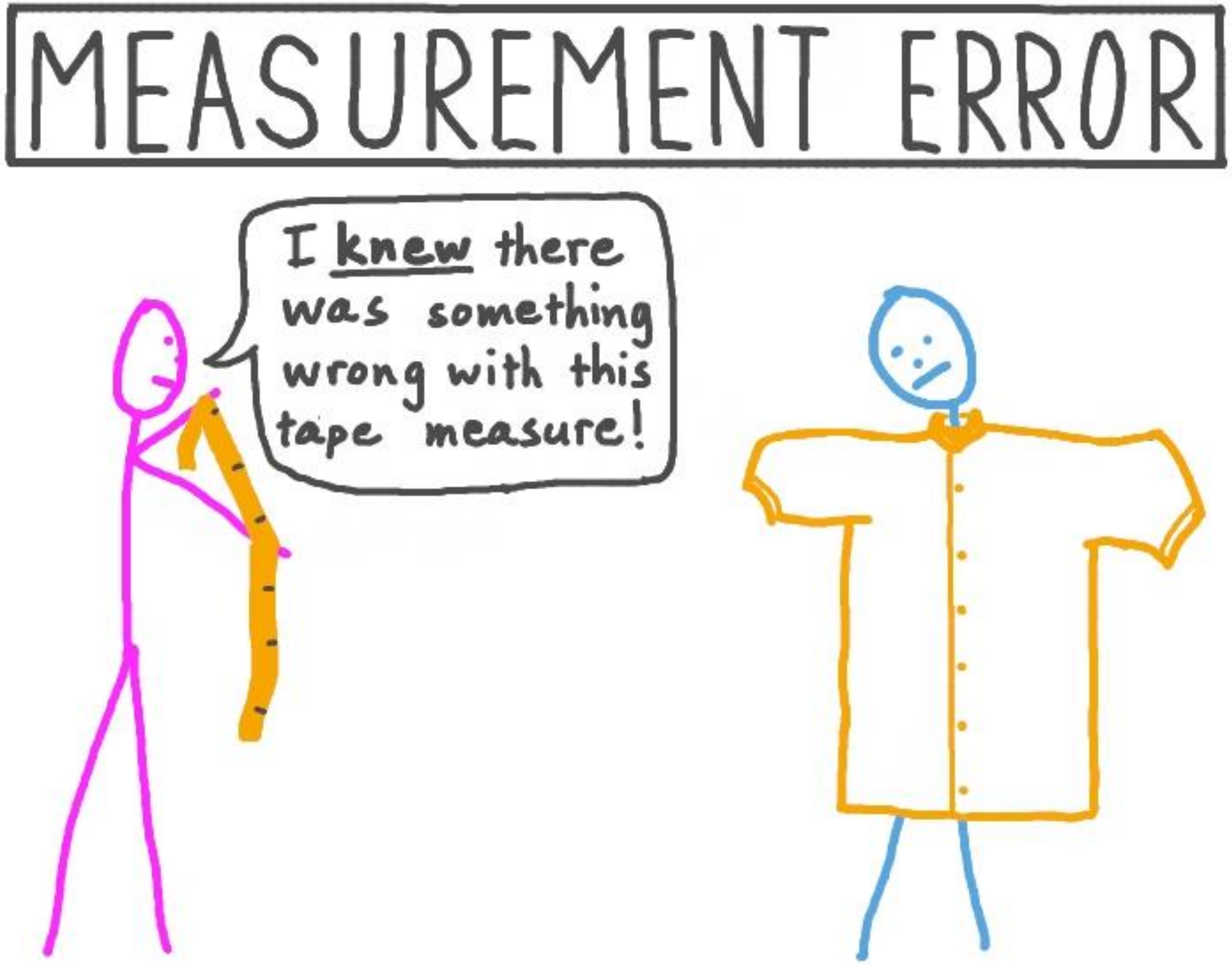
Hynek Cígler

Chyba měření a intervaly spolehlivosti

Opakování:
standardní chyba měření
standardní chyba predikce
standardní chyba rozdílu

Statisticky významný rozdíl

Klinicky významný rozdíl



https://www.nagwa.com/en/videos/138104137874/

Otázky spojené s chybou měření

Respondentovi naměřím výšku 178 cm.
Jaké otázky si mohu položit?

- Kolik měří právě teď?
- Kolik bude měřit příště?
- Kolik mu můžu naměřit příště, pokud se jeho výška nezmění?
- Kolik mu musím naměřit příště, abych mohl konstatovat, že se jeho výška změnila?

Kromě toho naměřím i jeho hmotnost 65 kg.
Jaké další otázky si mohu položit?

- Je „vyšší než těžší“?
- Je „vyšší než těžší“ oproti jiným respondentům?

Chyba měření

Standardní chyba měření: směrodatná odchylka pozorovaných hodnot okolo skutečné úrovně atributu

Příklad:

- https://www.zoology.ubc.ca/~whitlock/Kingfisher/SamplingNormal.htm
- https://www.zoology.ubc.ca/~whitlock/Kingfisher/CLT.htm

Chyba měření a CI

Rozložení naměřených hodnot je normálně rozložené a definované svým M a SD.

Proto, když konstruujeme CI, musíme vědět:

- Okolo čeho? Jaký je průměr rozložení?
- Jak nepřesné? Jaká je směrodatná odchylka rozložení (SE?)

Tři klíčové vzorce
(z nichž lze vše odvodit)

1. Základní teorém CTT:

$$X = \tau + e$$

- X – pozorované, τ – pravé skóre a e – chyba.

2. Reliabilita $r_{xx'}$ je podíl vysvětleného rozptylu:

$$r_{xx'} = \frac{\sigma_\tau^2}{\sigma_x^2} = \frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma_e^2} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$$

- Symbol sigma (σ^2) označuje rozptyl.

3. Rozptyl součtu dvou náhodných proměnných A+B má rozptyl:

$$\sigma_{A\pm B}^2 = \sigma_A^2 + \sigma_B^2 + 2\sigma_{AB} = \sigma_A^2 + \sigma_B^2 \pm 2r_{AB}\sigma_A\sigma_B$$

- $\sigma_{AB} = \mathrm{cov}(A, B)$ – kovariance, r_{AB} – jejich korelace (grafická ilustrace)
- Protože $r_{\tau e} = 0$, pak z 1 a 3 vyplývá $\sigma_x^2 = \sigma_\tau^2 + \sigma_e^2$.

Standardní chyba měření

Když rovnici $r_{xx'} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$ vyřešíme pro σ_e, získáme vzorec standardní chyby měření:

$$SE = \sigma_e = \sigma_x \sqrt{1 - r_{xx'}}$$

Standardní chyba měření

Když rovnici $r_{xx'} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$ vyřešíme pro σ_e, získáme vzorec standardní chyby měření:

$$SE = \sigma_e = \sigma_x\sqrt{1 - r_{xx'}}$$

- reliabilita – podíl vysvětleného rozptylu

Standardní chyba měření

Když rovnici $r_{xx'} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$ vyřešíme pro σ_e, získáme vzorec standardní chyby měření:

$$SE = \sigma_e = \sigma_x\sqrt{1 - r_{xx'}}$$

- reliabilita – podíl vysvětleného rozptylu
- „nereliabilita“ – podíl nevysvětleného rozptylu

Standardní chyba měření

Když rovnici $r_{xx'} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$ vyřešíme pro σ_e, získáme vzorec standardní chyby měření:

$$SE = \sigma_e = \sigma_x \sqrt{1 - r_{xx'}}$$

- reliabilita – podíl vysvětleného rozptylu
- „nereliabilita“ – podíl nevysvětleného rozptylu
- převod z rozptylu na směrodatnou odchylku
  - podíl směrodatné odchylky pravého skóru, která je „způsobena“ chybou

Standardní chyba měření

Když rovnici $r_{xx'} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$ vyřešíme pro σ_e, získáme vzorec standardní chyby měření:

$$SE = \sigma_e = \sigma_x\sqrt{1 - r_{xx'}}$$

- reliabilita – podíl vysvětleného rozptylu
- „nereliabilita“ – podíl nevysvětleného rozptylu
- převod z rozptylu na směrodatnou odchylku
  - podíl směrodatné odchylky pravého skóru, která je „způsobena“ chybou
- převod z podílu (z-skóre) přímo na škálu směrodatné odchylky

Standardní chyba měření

Když rovnici $r_{xx'} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$ vyřešíme pro σ_e, získáme vzorec standardní chyby měření:

$$SE = \sigma_e = \sigma_x\sqrt{1 - r_{xx'}}$$

- reliabilita – podíl vysvětleného rozptylu
- „nereliabilita“ – podíl nevysvětleného rozptylu
- převod z rozptylu na směrodatnou odchylku
  - podíl směrodatné odchylky pravého skóru, která je „způsobena“ chybou
- převod z podílu (z-skóre) přímo na škálu směrodatné odchylky
- směrodatná odchylka chyby měření

Standardní chyba měření

Když rovnici $r_{xx'} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$ vyřešíme pro σ_e, získáme vzorec standardní chyby měření:

$$SE = \sigma_e = \sigma_x\sqrt{1 - r_{xx'}}$$

Ve vzorci je $r_{xx'}$ vysvětlený rozptyl; viz koeficient determinace (PSYb1170).

- Tedy rozptyl měření vysvětlený pravým skórem. Na rozdíl od koeficientu determinace tam není mocnina, protože reliabilita je už přímo „umocněná“.
- $r_{x\tau} = \sqrt{r_{xx'}}$ a tedy $r_{x\tau}^2 = r_{xx'}$

Středová hodnota

Chyba se nepohybuje kolem pozorovaného, ale kolem pravého skóre.

Jaká je nejpravděpodobnější hodnota pravého skóre při určitém pozorovaném skóre x?

O trochu blíže k průměru (protože pravé skóry mají menší rozptyl než pozorované skóry).

Regresní model CTT:

$$\mathrm{E}(T|x) = r_{xx'}x + (1 - r_{xx'})M_x$$

- $\mathrm{E}(T|x)$: očekávané (expected), nejpravděpodobnější pravé skóre.
- $r_{xx'}$: reliabilita; „směrnice“.
- M_x : průměrné skóre; $(1 - r_{xx'})M_x$ je „průsečík“.
- Čím větší reliabilita, tím větší vliv pozorovaného skóre a menší vliv průměru (a naopak).

Směrodatná odchylka pravého skóre: $\sigma_\tau = \sqrt{r_{xx'}}\sigma_x$

Chyba měření (v CTT)

Takto spočítanou chybu měření mohu použít pro konstrukci intervalu spolehlivosti.

$$CI_i = \mathrm{E}(X) \pm z_i \sigma_e$$

- $\mathrm{E}(X)$ = očekávaná hodnota, okolo které interval konstruuji.
- σ_e = chyba měření
- z_i = kvantil normálního rozdělení

Kvantily normálního rozdělení:

- 95% CI: $z_{95\%} \cong 1{,}96$
- 90% CI: $z_{90\%} \cong 1{,}64$
- 80% CI: $z_{80\%} \cong 1{,}28$
- 68% CI: $z_{68\%} \cong 1{,}00$

Shrnutí: Důležité prvky práce s SE

Co je očekávanou hodnotou, okolo které interval konstruuji?

- Pozorované skóre?
- Odhad pravého skóre?
- Nula (pro rozdíl dvou skórů)?

Jak spočítám chybu pro daný účel/diagnostickou otázku?

Jaký odhad reliability nejlépe použiju pro daný účel?

Scénář 1: Standardní chyba měření

Pokud jsme naměřili pozorované skóre X, jaké jiné alternativní X jsme mohli rovněž naměřit?

Slouží pro popis chyby měření a intervalu spolehlivosti jednoho jediného měření.

Velikost chyby:

$$\sigma_e = \sigma_x \sqrt{1 - r_{xx'}}$$

Středová hodnota: odhad pravého skóre

$$\mathrm{E}(T|x) = r_{xx'}x + (1 - r_{xx'})M_x$$

Scénář 2: Chyba odhadu pravého skóre

Pokud jsme naměřili pozorované skóre X, jaká je chyba odhadu pravého skóre τ?

Vzorec je stejný, jen namísto SD pozorovaného skóre použijeme odhad SD pravého skóre:

Velikost chyby:

$$\sigma_{e(\tau)} = \sigma_\tau \sqrt{1 - r_{xx'}} = \sigma_x \sqrt{r_{xx'}} \sqrt{1 - r_{xx'}}$$

Středová hodnota:

$$\mathrm{E}(T|x) = r_{xx'} x + (1 - r_{xx'}) M_x$$

Někteří autoři tento postup doporučují, ale potíž s interpretací.

- Zajímá nás chyba na škále použité při konstrukci norem. Zpravidla tedy nepoužitelné.
- Nicméně např. WISC-5[UK] – pro standardizaci na IQ použil právě σ_τ
  - Standardizace $IQ = 15\frac{(X - M_x)}{\sigma_x \sqrt{r_{xx'}}} + 100$ namísto běžného $IQ = 15\frac{(X - M_x)}{\sigma_x} + 100$

Scénář 3: Standardní chyba predikce

Naměřil jsem X. V jakém rozsahu bude ležet příští měření, pokud se úroveň atributu nezmění?

- *„Zlepšil se klient v terapii?“ „Je účinný výukový program?“*

Velikost chyby:

$$\sigma_{pred} = \sigma_x \sqrt{1 - r_{xx'}^2}$$

- $r_{xx'}^2$ - druhá mocnina (test-retest) reliability.
- jde o úpravu $\sigma_{pred} = \sqrt{\sigma_e^2 + \sigma_{e(\tau)}^2}$, tedy rozdíl chyby odhadu pravého skóru a chyby měření

Středová hodnota = očekávaný skór při retestu: odhad pravého skóre:

$$\mathrm{E}(T|x) = r_{xx'}x + (1 - r_{xx'})M_x$$

Scénář 4: Statisticky významný rozdíl

Standardní chyba rozdílu. Rozdíl dvou nezávislých testů jedné osoby; případně rozdíl dvou osob.

Jaká je očekávaná odlišnost v měření dvěma testy?

- „Dosáhla vyššího skóru Anežka nebo Bedřich?“ „Je Cyril vyšší nebo těžší?“
- Musí být ve stejných jednotkách.

Velikost chyby:

$$\sigma_{e(A-B)} = \sqrt{\sigma_{e(A)}^2 + \sigma_{e(B)}^2} = \sigma_{ab}\sqrt{2 - r_{aa'} - r_{bb'}}$$

- Pokud jde o měření jediným testem (dvěma testy se stejnou reliabilitou), lze zjednodušit:

$$\sigma_{e(A-B)} = \sqrt{2}SE = \sigma_x\sqrt{2}\sqrt{1 - r_{xx'}}$$

Středová hodnota:

- Jde o rozdíl a očekávaný rozdíl je zpravidla žádný rozdíl, **proto zpravidla 0**.
- To není úplně pravda; pokud $r_{aa'} \neq r_{bb'}$, pak je střední hodnotou $E(\tau_A' - \tau_B') = \sqrt{r_{AA'}}(A - M) - \sqrt{r_{BB'}}(B - M)$, ale výsledek bude velmi podobný. Zanedbejte.

Scénář 5: Klinicky významný rozdíl

Liší se dva skóry téhož respondenta více či méně než u „běžných“ respondentů?

- To, že se skóry liší, neznamená, že se liší více, než bychom čekali u náhodně vybraného člověka.
- Klinické hypotézy: *„Rozkolísaný profil schopností...“*, *„Je rozdíl ‚klinicky‘ významný?“* atd.

Příklad:

- **Statisticky významný rozdíl:** *„Člověk má vyšší váhu než výšku (ve standardních jednotkách, např. IQ skórech)“.*
- **Klinicky významný rozdíl:** *„Člověk má vyšší váhu, než by odpovídalo jeho výšce, je tedy obézní.“*

Scénář 5: Klinicky významný rozdíl

Více postupů. Nejjednodušší používá pouze korelaci a je zcela shodný s postupem pro chybu predikce.

Odhad chyby:

$$\sigma_{A-B} = \sigma_{AB}\sqrt{1 - r_{AB}^2}$$

- r_{AB} je korelace testů A a B, σ_{AB} je směrodatná odchylka obou testů (musí být shodná)

Středová hodnota:

$$\mathrm{E}(B|A) = r_{AB}A + (1 - r_{AB})M_{AB}$$

Scénář 6: Více měření

Lze testovat, zda má klient celkově „rozkolísaný profil".

- Např.: *„Liší se subtesty ve WAIS-III od celkového IQ více, než bychom čekali?"*
- Analogie F-testu u lineární regrese s více prediktory.

Poskytují jen některé diagnostické metody, není pravidlem.

Technicky vzato není ideální interpretovat „profil", pokud test celkového rozdílu není signifikantní na zvolené *p*-hladině.

Ruční výpočet je příliš náročný.

Studijní zdroje

Viz interaktivní osnova.

Pracovní verze Diagnostické kalkulačky (bez záruky).