Otázky k tématu 8 – statistická indukce a intervalové odhady

1. Je k výběru prostého náhodného vzorku třeba, aby byly splněny následující podmínky (ano-ne)?

    1.1 Normální rozložení skórů v populaci

    1.2 Každý jedinec v populaci (jednotka zkoumání) musí mít stejnou pravděpodobnost, že bude vybrán do vzorku.

    1.3 Výběr kteréhokoli jedince (jednotky zkoumání) musí být zcela nezávislý na výběru kohokoli jiného.

2. U Wechslerových inteligenčních škál jsou skóry normálně rozložené s $\mu = 100$ a $\sigma = 15$. Představte si, že jsme otestovali náhodný vzorek 9 lidí, spočítali jejich průměrný skór a celou proceduru zopakovali 1000krát.

    2.1 Odhadněte směrodatnou odchylku těch 1000 průměrů.

    2.2 Zhruba jaká část (v %) těchto výběrových průměrů s n = 9 by byla vyšší než 105? A než 110?

    2.3. Zhruba jaká část (v %) těchto výběrových průměrů s n = 9 by byla mezi 95 a 105? A mezi 90 a 110?

    2.4 Budou tyto výběrové průměry normálně rozložené?

    2.5 Jaký je rozptyl tohoto rozložení výběrových průměrů?

    2.6 Kdyby $n = 225$ (místo 9), jaká by byla hodnota směrodatné chyby?

    2.7 Kdyby $n = 225$, jaká část (v %) výběrových průměrů by se pohybovala nejvýše jeden bod od 100, tj. mezi 99 a 101?

    2.8 Bude rozložení výběrových průměrů přibližně normální i tehdy, kdyby rozložení skórů v populaci nebylo normální?

3. Jsou následující páry ekvivalentními výrazy?

    3.1 směrodatná chyba m a směrodatná odchylka výběrového rozložení m

    3.2 σ^2/n a směrodatná chyba m.

    3.3 σ_m^2 a rozptyl výběrového rozložení m.

    3.4 populační rozptyl σ^2 a n krát σ_m^2

    3.5 průměr výběrového rozložení m a σ_m^2

    3.6 μ a $(\sum X)/n$

    3.7 m a $(\sum x)/n$

    3.8 s^2 a $\sum x^2/(n-1)$

4. Odpovězte na následující otázky

    4.1 Pokud budeme provádět mnoho různých studií na různá témata a vždy v nich budeme k odhadovaným statistikám tvořit 95% intervaly spolehlivosti, kolik z těchto intervalů asi obsahuje odhadovaný parametr?

4.2 V kterém z následujících případů by nárůst velikost vzorku způsobil největší zúžení intervalu spolehlivosti?

a) z 5 na 25
b) z 10 na 30
c) ze 40 na 60

4.3 Který typ odhadu lépe sděluje přesnost odhadu – bodový, nebo intervalový?

4.4 *z* se má k σ_m jako *t* se má k

a) σ
b) σ^2
c) *s*
d) s_m

4.5 Za jakých podmínek platí $\sigma = \sigma_m$?

4.6 Jaký je vztah mezi zvolenou hladinou spolehlivosti a šířkou intervalu spolehlivosti?

4.7 Který matematický teorém tvrdí, že výběrové rozložení průměru se s rostoucím *n* blíží normálnímu, bez ohledu na tvar rozložení proměnné v populaci?

4.8 Známe-li σ, platí, že m +/- $1{,}96\sigma_m$ tvoří 95% interval spolehlivosti pro jakékoli *n*?

4.9 Předpokládejme proměnnou s normálním rozložením a známou σ. Pokud vybereme náhodně 2 vzorky o $n = 100$ a na obou spočítáme 68% interval spolehlivosti, budou tyto dva intervaly shodné?

4.10 Jak velké musí být *n*, aby směrodatná odchylka výběrového rozložení průměru σ_m byla pouze 10% směrodatné odchylky *s* rozložení proměnné X?

5. Provádíme výzkum toho, zda a jak dlouhodobá hospitalizace škodí dětem ve vývoji. Jednou z výzkumných otázek je, zda nedochází ke zabrzdění vývoje intelektu. Pro tento účel jsme 30 dlouhodobě hospitalizovaným dětem v mladším školním věku rozdali inteligenční test s následujícími výsledky: m_{IQ}=98, s_{IQ}11.

5.1 Stanovte 95% interval spolehlivosti pro průměrnou hodnotu inteligence v populaci dlouhodobě hospitalizovaných dětí v mladším školním věku (μ_{IQ}).

5.2 Dále jsme zjistili, že délka hospitalizace (ve dnech) koreluje s IQ, r=-0,1. Stanovte 95% interval spolehlivosti pro korelaci mezi délkou hospitalizace a IQ (ρ) (viz Hendl 252, pozor na chybu, výběrové rozložení Fischerova Z je normální, nikoli *t*)

6. Výzkumník se zabývá rychlostí čtení u vysokoškolských studentů. U šesti studentů změřil počet přečtených slov za minutu: 200, 240, 300, 410, 450 a 600.

6.1 Spočítejte průměr a směrodatnou odchylku.

6.2 Jaké rozdělení použijete při konstrukci intervalu spolehlivosti, normální nebo t-rozdělení? Proč?

6.3 Zkonstruujte 95% interval spolehlivosti pro průměr.

6.4 Bude 99% interval spolehlivosti pro průměr širší nebo užší oproti 95% intervalu? Zkonstruujte i 99% interval spolehlivosti.

6.5 Výzkumník měl pocit, že pouze 6 zkoumaných osob je příliš málo. Rozhodl se předstírat, že zkoumal 18 osob, přitom zachoval průměr a směrodatnou odchylku vzorku. Jak se změní intervaly spolehlivosti pro průměr? Co vše bude ovlivněno? Vypočítejte 95% interval spolehlivosti.

7. Badatel se zabývá Stroopovým efektem. U každé pokusné osoby změřil dobu, za jakou pojmenují barvy 60 barevných obdélníčků, a dobu, za jakou pojmenují barvy 60 slov vytištěných barevným inkoustem, avšak tyto slova byly také názvy barev. Příslušné časy obou měření a jejich rozdíl pro 10 osob shrnuje tabulka:

barevné obdélníčky	barevná slova	rozdíl
17	38	21
15	58	43
18	35	17
20	39	19
18	33	15
20	32	12
20	45	25
19	52	33
17	31	14
21	29	8

7.1 Spočítejte 95% interval spolehlivosti pro průměr rozdílů časů dat z tabulky.

7.2 Uvedená tabulka byla pouze částí dat, celkem se výzkumu zúčastnilo 47 osob, průměrný rozdíl byl 16,362 sekund a směrodatná odchylka 7,470. Spočítejte 95% interval spolehlivosti pro průměr rozdílů časů celého výzkumu.

8. Několik studentů psychologie dospělo k podezření, že je výčepní v hospodě okrádá. Objednali si tedy 8 piv a změřili jejich objem: 0.51, 0.462, 0.491, 0.466, 0.461, 0.503, 0.495, 0.488 (v litrech). Zkonstruujte 95% interval spolehlivosti pro průměr.

9. Badatel zkoumá inteligenci u dětí v zájmových kroužcích. U dětí z modelářského kroužku naměřil následující hodnoty IQ: 128, 117, 139, 122, 101, 152, 120, 127, 136, 108.

9.1 Spočítejte 95% interval spolehlivosti pro průměr IQ za předpokladu, že populační směrodatná odchylka IQ je 15. Co tento předpoklad ovlivní?

9.2 Spočítejte 95% interval spolehlivosti pro průměr, pokud neznáte populační směrodatnou odchylku.

9.3 Spočítejte 95% interval spolehlivosti pro rozptyl[1].

9.4 Spočítejte 99% interval spolehlivosti pro rozptyl.

10. Národní centrum pro výuku statistiky (příklad je z amerického prostředí) zkoumá dobu, kterou vysokoškolští studenti potřebují k dosažení bakalářského titulu. Velikost vzorku je 101, průměr 5,15 let a směrodatná odchylka 1,68. Zkonstruujte 99% interval spolehlivosti pro populační směrodatnou odchylku.

11. Firma dodávající náhradní díly automobilů chce zjistit, jak dlouho si majitelé plánují svůj vůz ponechat. Náhodný vzorek 25 majitelů má průměrnou dobu 7,01 rok a směrodatnou odchylku 3,74. Spočítejte 95% interval spolehlivosti pro populační průměr a 95% interval spolehlivosti pro směrodatnou odchylku.

12. V roce 1961 byla zjištěna výška u 15 náhodně vybraných chlapců z populace všech desetiletých chlapců žijících v Československu. Určete 95% interval spolehlivosti pro střední výšku, pokud víte, že v roce 1951 byla zjištěna směrodatná odchylka s = 6,4 cm a je známo, že variabilita výšek postavy se v různých generacích příliš nemění.

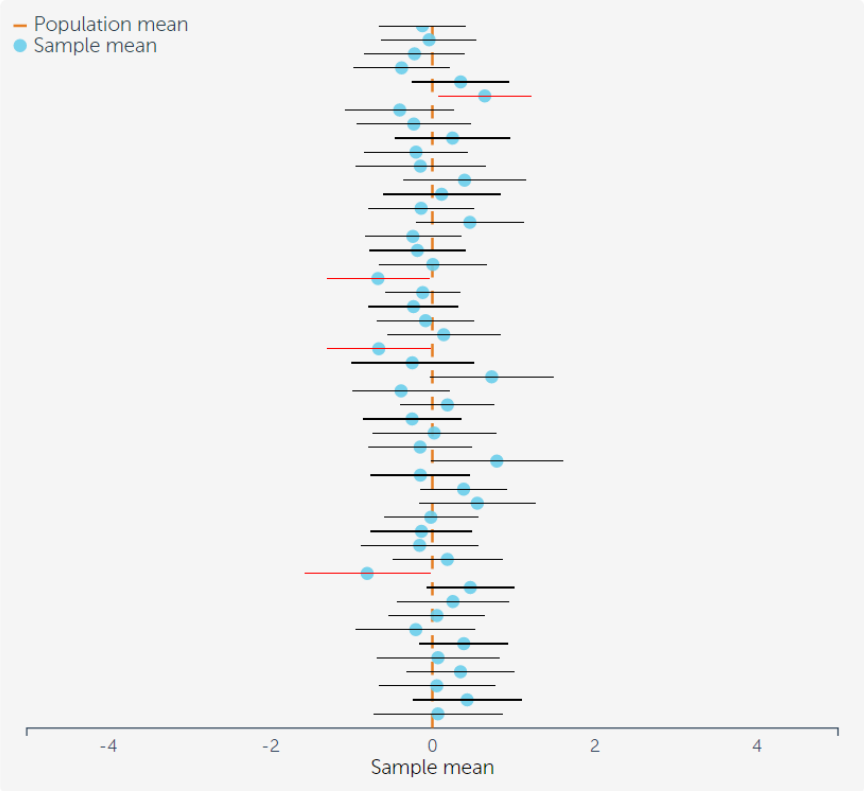
[1] Interval spolehlivosti pro rozptyl/sm. odchylku je jen pro zvídavé...

Naměřené výšky:
130, 140, 136, 141, 139, 133, 149, 151, 139, 136, 138, 142, 127, 139, 147

13. Sestavte 95% interval spolehlivosti pro korelační koeficient ρ, znáte-li výběrovou hodnotu korelace mezi čistým měsíčním příjmem na člena domácnosti a výdaji na sport, rekreaci a zábavu $r = 0,45$ vypočtenou z údajů získaných od 20 rodin.

14. Hypotéza pre jeden experiment znela, že priemerné skóre je rovné 20. Ak je 99% CI (18;24), co si budeme o hypotéze myslet?

15. Ktoré z týchto 95% CI pre rozdiel medzi priemermi reprezentujú signifikantný rozdiel na hladine 0,05?
a) (-4,6;-1,8)
b) (-0,2; 8,1)
c) (-5,1; 6,7)
d) (3; 10,9)

16. Ak 95%CI obsahuje 0, bude 0 obsahovať aj 99%CI?

17. Na webu R-psychologist (http://rpsychologist.com/d3/CI/) jsem využil simulaci počítání intervalů spolehlivosti v nekonečně mnoha replikacích nějakého výzkumu. Níže je screenshot 50 výsledků tohoto simulování. Jakou hladinu spolehlivosti jsem asi nastavil?