

CVIČENÍ Z ANALÝZY DAT SOCIOLOGICKÉHO VÝZKUMU

4. ÚKOL

Jméno a příjmení:

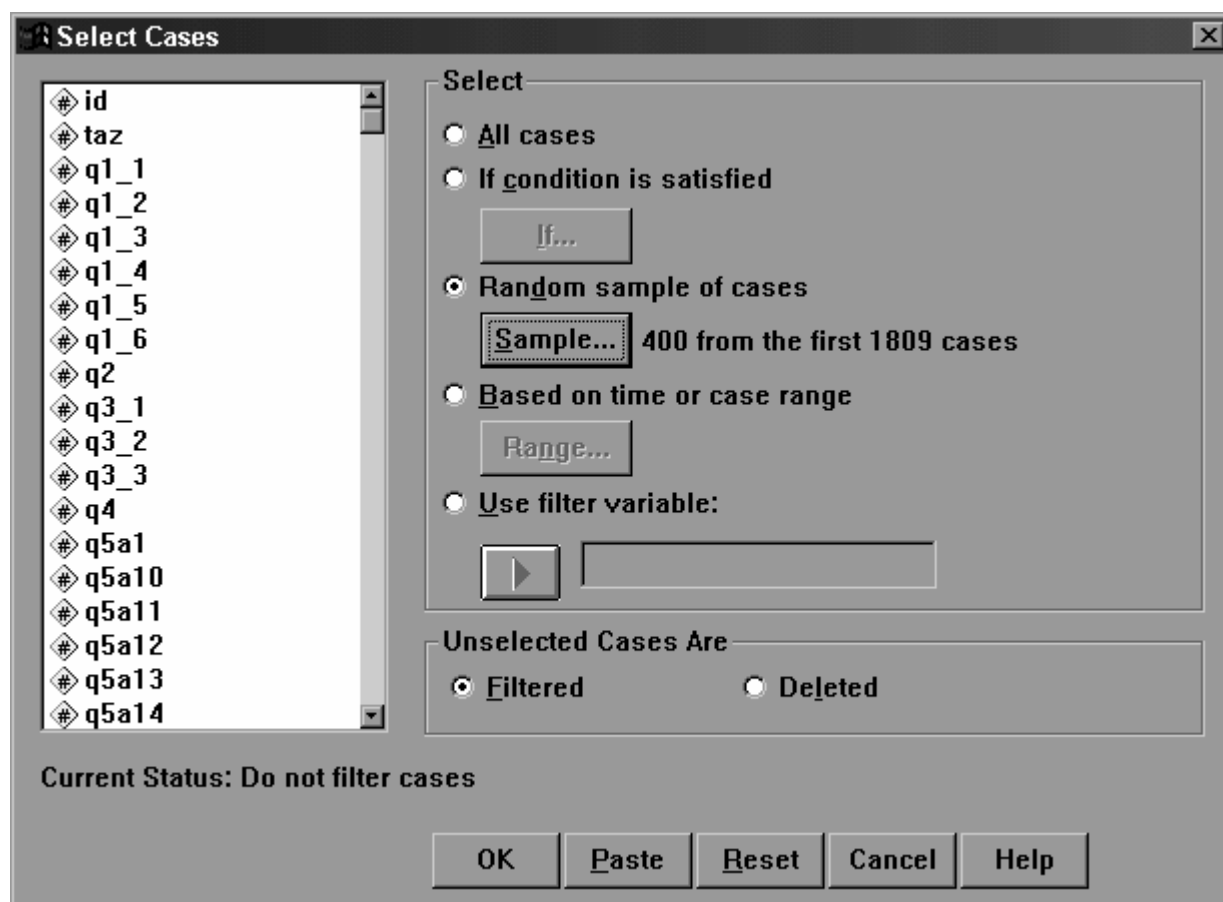
Imatrik. ročník:

C1: Udělejte 15 náhodných výběrů ze souboru EVS-CR99.sav tak, abyste v každém výběru náhodně vybrali vždy přesně 400 případů z celkových 1809. Sledujte, jak se mění průměr a směrodatná odchylka proměnné q54b v těchto výběrech a srovnajte tyto hodnoty s celkovým průměrem. Zobrazte hodnoty těchto průměrů zaokrouhlené na jedno desetinné místo do grafu a zjistěte, zdali tato distribuce má charakteristiky normálního rozložení. Budou výsledky vašich řešení identické s řešením vašich kolegů, nebo se budou lišit? Zdůvodněte svou odpověď.

Řešení:

Vždy 15 krát uděláme výběr souboru a pak spočítáme průměr a směrodatnou odchylku.

1. Výběr případů



2. Průměry v podsouborech:

Statistics

Q54B Nezaměstnaní nabízené zaměstnání:

N	Valid	378
	Missing	5
Mean		4,61
Std. Deviation		2,77

Tabulka 15 náhodných výběrů

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
N	376	379	397	380	397	393	386	377	385	392	380	398	395	396	394
Průměr	4,7	4,7	4,6	5	5	4,73	4,8	4,7	4,6	4,7	4,7	5	4,7	4,7	4,7
Std.															
Odchylka	2,6	2,7	2,7	2,7	2,6	2,8	2,6	2,7	2,7	2,8	2,7	2,7	2,6	2,6	2,7

Každý student bude mít těchto patnáct hodnot průměrů odlišných, neboť SPSS to vybírá náhodně. Pokud bychom udělali rozdělovací funkci těchto patnácti průměrů, tj. nakreslili graf, kde by na ose x byly jednotlivé hodnoty průměru a na ose y kolikrát nám vyšly, měli bychom získat přibližně normální rozložení.

To vychází z centrálního limitního teorému, který říká, že když provedeme mnoho výběrů o určité velikosti založených na pravděpodobnostním výběru (což jsme udělali, i když jich nebylo mnoho), tak rozložení výběrových průměrů bude přibližně normální a celkový průměr z těchto průměrů bude blízký průměru v populaci. Celkový průměr ze všech těchto výběrových průměrů je 4,76.

Pro získání průměru pro celý soubor musím zadat v Data – Select cases – All cases – Ok
Pak nechám spočítat průměr.

Statistics

Q54B Nezaměstnaní nabízené zaměstnání:

N	Valid	1874
	Missing	34
Mean		4,69
Std. Deviation		2,71

Průměr celého souboru byl 4,7.

C2: Stanovte interval spolehlivosti průměru odpovědi na otázku q54f z výzkumu EVS-ČR1999, zdali by u nás mělo vzrůstat soukromé vlastnictví obchodu a průmyslu, nebo zda by se mělo v co nejširší míře zachovat vlastnictví státní. Interval spolehlivosti stanovte se spolehlivostí 95 a 99 %. V odpovědi nad získaným výsledkem srovnajte českou situaci se situací v některých evropských zemích. K tomu použijte tabulku 1. Do ní dopočítejte intervaly spolehlivosti pro vámi zvolených pět zemí.

Řešení:

Descriptives

		Statistic	Std. Error	
Q54F Rozšiřovat/ zachovat by se mělo obchod. a prům.:	Mean	4,71	,06	
	95% Confidence Interval for Mean	Lower Bound	4,60	
		Upper Bound	4,83	
	5% Trimmed Mean	4,62		
	Median	5,00		
	Variance	6,317		
	Std. Deviation	2,51		
	Minimum	1		
	Maximum	10		
	Range	9		
	Interquartile Range	3,00		
	Skewness	,376	,057	
	Kurtosis	-,678	,114	

Descriptives

		Statistic	Std. Error	
Q54F Rozšiřovat/ zachovat by se mělo obchod. a prům.:	Mean	4,71	,06	
	99% Confidence Interval for Mean	Lower Bound	4,56	
		Upper Bound	4,86	
	5% Trimmed Mean	4,62		
	Median	5,00		
	Variance	6,317		
	Std. Deviation	2,51		
	Minimum	1		
	Maximum	10		
	Range	9		
	Interquartile Range	3,00		
	Skewness	,376	,057	
	Kurtosis	-,678	,114	

Interval spolehlivosti se počítá pomocí vzorečku $C.I._{95\%} = X \pm 1,96 \cdot s / \sqrt{N}$ přičemž s je směrodatná odchylka. Někdy se 1,96 zaokrouhluje na 2.

Tabulka 1: Odpovědi na otázku q54f v různých evropských zemích

Země	Průměr	Std. Dev.	N	95% Interval spolehlivosti	
Francie	4,03	2,23	1498	<3,92	4,14>
SRN	4,05	2,14	1836	<3,95	4,15>
Rakousko	3,41	2,08	1301	<3,30	3,52>
Itálie	4,11	2,21	1801	<4,01	4,21>
Nizozemsko	4,42	1,88	968	<4,30	4,54>
Finsko	4,22	2,14	956	<4,08	4,36>

Estonsko	5,74	2,44	902	<5,58	5,90>
Litevsko	4,94	2,85	840	<4,75	5,13>
Polsko	5,78	2,80	1017	<5,6	5,96>
Rumunsko	4,38	3,23	1032	< 4,18	4,6>
Chorvatsko	4,12	2,98	985	< 3,93	4,31>
Bělorusko	5,24	2,69	859	<5,06	5,42>
Ukrajina	5,60	2,99	1036	<5,42	5,78>
Rusko	6,11	2,77	2196	<5,99	6,23>
ČR	4,71	2,51	1851	<4,60	4,82>

Názor české populace na vlastnictví obchodu a průmyslu se nachází někde uprostřed mezi soukromým a státním vlastnictvím. To je průměrné i z hlediska dalších evropských zemí, z nichž nejvíce pro soukromé vlastnictví je Rakousko (průměr 3,41) a nejvíce pro státní vlastnictví Rusko (6,11).

C3: Stanovte interval spolehlivosti pro odpovědi na otázku q7_10 z výzkumu EVS-ČR1999, zdali by lidé nechtěli mít za sousedy lidi s nemocí AIDS. Kromě toho ve své odpovědi srovnajte český postoj k lidem s AIDS s ostatními evropskými zeměmi (viz tab. 2).

Tab. 2: Postoje k lidem s AIDS v různých evropských zemích

			V61 people with AIDS		Total	
			0 not mentioned	1 mentioned		
COUNTRY country code	9901	France	Count	1476	139	1615
			Row %	91,4%	8,6%	100,0%
	9902	UK	Count	759	241	1000
			Row %	75,9%	24,1%	100,0%
	9903	Germany	Count	1602	213	1815
			Row %	88,3%	11,7%	100,0%
	9905	Austria	Count	1241	281	1522
			Row %	81,5%	18,5%	100,0%
	9906	Italy	Count	1379	621	2000
			Row %	69,0%	31,1%	100,0%
	9907	Spain	Count	948	252	1200
			Row %	79,0%	21,0%	100,0%
	9909	Netherl	Count	922	81	1003
			Row %	91,9%	8,1%	100,0%
	9910	Belgium	Count	1688	224	1912
			Row %	88,3%	11,7%	100,0%
	9911	Denmark	Count	960	59	1019
			Row %	94,2%	5,8%	100,0%
	9913	Sweden	Count	948	67	1015
			Row %	93,4%	6,6%	100,0%
	9916	Northern Ireland	Count	681	319	1000
			Row %	68,1%	31,9%	100,0%
	9917	Ireland	Count	763	242	1005
			Row %	75,9%	24,1%	100,0%
	9918	Estonia	Count	589	416	1005
			Row %	58,6%	41,4%	100,0%
	9919	Latvia	Count	719	294	1013
			Row %	71,0%	29,0%	100,0%
	9920	Lithuani	Count	459	559	1018
			Row %	45,1%	54,9%	100,0%
	9921	Poland	Count	592	503	1095
			Row %	54,1%	45,9%	100,0%
	9922	Czechia	Count	1511	397	1908
			Row %	79,2%	20,8%	100,0%
	9923	Slovakia	Count	738	593	1331
			Row %	55,4%	44,6%	100,0%
	9924	Hungary	Count	233	767	1000
			Row %	23,3%	76,7%	100,0%
	9925	Romania	Count	606	540	1146
			Row %	52,9%	47,1%	100,0%
	9926	Bulgaria	Count	494	506	1000
			Row %	49,4%	50,6%	100,0%
	9928	Greece	Count	835	307	1142
			Row %	73,1%	26,9%	100,0%
	9929	Russia	Count	1168	1332	2500
			Row %	46,7%	53,3%	100,0%
	9934	Slovenia	Count	672	334	1006
			Row %	66,8%	33,2%	100,0%
	99031	SRN	Count	810	112	922
			Row %	87,9%	12,1%	100,0%
	99032	DDR	Count	792	101	893
			Row %	88,7%	11,3%	100,0%

Řešení: Nejprve si pomocí příkazu Recode změňme kódování této dichotomické proměnné na 0 a 100.

Q7_10 Nemocní AIDS

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 nezmínil/a	1515	79,4	79,4	79,4
1 nechce	393	20,6	20,6	100,0
Total	1908	100,0	100,0	

Q7_10 Nemocní AIDS

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0 nezmínil/a	1515	79,4	79,4	79,4
100	393	20,6	20,6	100,0
Total	1908	100,0	100,0	

Pak v *explore* spočítáme průměr a jeho interval spolehlivosti pro jistotu spolehlivosti 95 %.

Descriptives

		Statistic	Std. Error	
Q7_10 Nemocní AIDS	Mean	20,58	,93	
	95% Confidence Interval for Mean	Lower Bound	18,77	
		Upper Bound	22,40	
	5% Trimmed Mean	17,32		
	Median	,00		
	Variance	1635,537		
	Std. Deviation	40,44		
	Minimum	0		
	Maximum	100		
	Range	100		
	Interquartile Range	,00		
	Skewness	1,456	,056	
	Kurtosis	,121	,112	

Interval spolehlivosti pro otázku q7_10 je <18,8;22,4>. Tj. procento lidí, jimž by vadilo mít za souseda nemocného AIDS se v populaci bude pohybovat v tomto rozmezí. Musíme dát pozor na to, co nám tento průměr říká, zde jsme kód 100 měli pro odpověď zmínil, že by mu lidé s Aids vadili, tudíž interval spolehlivosti, který nám vyšel, je pro tuto odpověď.

79,4 % (1515) respondentů zvolilo odpověď 0 = „nezmínil/a“, což znamená, že většině nevadí mít za sousedy nemocné AIDS. 20,6% (393) respondentům by vadilo mít nemocné s AIDS za sousedy v mezinárodním srovnání se tak blížíme Velké Británii a Španělsku, ale v porovnání s Francií, Holandskem a především Dánskem jsme stále málo tolerantní vůči této skupině.

C4: Stanovte intervaly spolehlivosti pro jednotlivé odpovědi na otázku q17_4, zdali pracovat je povinností vůči společnosti. Využijte k tomu příslušný script nebo excelovskou tabulku..

Řešení:

Q17_4 Pracovat je povinnost

		Frequenc y	Percent	Valid Percent	Cumulative Percent
Valid	1 rozhodně souhlasí	358	18,8	19,0	19,0
	2 souhlasí	831	43,6	44,0	62,9
	3 ani souhlas ani nesouhlas	368	19,3	19,5	82,4
	4 nesouhlasí	278	14,6	14,7	97,1
	5 rozhodně souhlasí	55	2,9	2,9	100,0
	Total	1889	99,0	100,0	
Missing	-2 neodpověděl/a	5	,2		
	-1 neví	14	,7		
	Total	19	1,0		
Total		1908	100,0		

Četnostní tabulka s intervaly spolehlivosti proměnné Q17_4 Pracovat je povinnost

Hodnoty	Statistiky		
	Rel. četnost platných hodnot	Dolní mez ^a	Horní mez ^a
Platné			
1 rozhodně souhlasí	18,96%	17,19%	20,73%
2 souhlasí	43,98%	41,74%	46,22%
3 ani souhlas ani nesouhlas	19,45%	17,67%	21,24%
4 nesouhlasí	14,71%	13,11%	16,31%
5 rozhodně souhlasí	2,89%	2,14%	3,65%
Celkem	100,00%		

a. 95%ní interval spolehlivosti. K výpočtu je použita asymptotická metoda, která předpokládá, že celkový počet pozorování je větší než 30 a v každé kategorii se vyskytuje alespoň 5 případů.

Microsoft Excel - C5-3.xls

Soubor Úpravy Zobrazit Vložit Formát Nástroje Data Okno Nápověda

Times New Roman CE 11 B I U

E14 = C14-D14

	A	B	C	D	E	F	G	H	I
1									
2	Q17_4 Pracovat je povinnost								
3			Frequency	Percent	Valid Percent	Cumulative Percent			
4	Valid	1 rozhodně souhlasí	358	18,8	19,0	19,0			
5		2 souhlasí	831	43,6	44,0	62,9			
6		3 ani souhlas ani nesc	368	19,3	19,5	82,4			
7		4 nesouhlasí	278	14,6	14,7	97,1			
8		5 rozhodně souhlasí	55	2,9	2,9	100,0			
9		Total	1889	99,0	100,0				
10									
11	Q17_4 Pracovat je povinnost								
12				std. Error	CI	CI			
13		Freq.	Valid %	95%	dolní	dolní			
14	1 rozhodně souhlasí	358	19,0	1,81	17,2	20,8			
15	2 souhlasí	831	44,0	2,28	41,7	46,3			
16	3. Ani, ani	368	19,5	1,82	17,7	21,3			
17	4 nesouhlasí	278	14,7	1,63	13,1	16,3			
18	5 rozhodně nesouhlasí	55	2,9	0,77	2,1	3,7			
19	Total	1889	100,0						
20									
21									

List1 / List2 / List3 / List4 / List5 / List6 /

Připraven 123