

Odpovědi

1. Teoreticky budou proměnnými všechny (tj. budou ve vymezené skupině nabývat alespoň 2 hodnot) kromě povolání, které by u studentů prezenčního studia mělo být konstantou. Pokud některé z nich budeme měřit/kódovat velmi hrubě (např. lateralita P n. L), může se stát, že budou ve vymezené skupině také konstantami. Když se přidají studenti kombinovaného studia, povolání přestane být konstantou. Nad rámec dichotomie konstanta vs. proměnná můžeme ještě uvažovat o tom, že v uvažované populaci bude značně omezená variabilita některých proměnných, např. rok narození a věk.

2. *d e i l m*

Komentář: Všimněte si formulace otázky. Úroveň měření není charakteristikou toho, co měříme, ale toho JAK měříme. Vyjadřuje naše přesvědčení o vztahu mezi hodnotami na škále, na které proměnnou měříme, a podobou (kvalitou/kvantitou) jevu, který měříme. Proto také v odpovědi na otázku, jako je tato, obvykle zahrnujeme i vysvětlení, proč považujeme škálu za nominální (či jinou).

3. *f g h* (letopočty mají nulu stanovenou arbitrárně), *n*. Zvláštní případ je *k*, dichotomie mají jen jeden interval, který je sám se sebou totožný. Proto je můžeme považovat za intervalovou škálu. Neuvažujeme o nich však jako o poměrových (byť nula často není arbitrární).

4. *b c j*

5. *a +* všechny výše označené za intervalové, u nichž vznikly nějaké zásadní pochyby ohledně rovnosti intervalů.

6. ano

7. poměrová

8. poměrové (intervalovost zůstane zachována, přibude smysluplná 0 – nulový rozdíl)

9. ano, poměrová.

10.

id	gr_level	group	pretest	posttest
1	1	1	60	107
...				
5	1	1	115	122
31	1	2	60	90
...				
35	1	2	120	121
6	2	1	65	118
...				
10	2	1	110	122
36	2	2	80	99
...				
40	2	2	120	123
...				

11.1 Proměnné jsou všechny charakteristiky kromě oblíbené zmrzliny. Oblíbená zmrzlina je na této populaci konstantní, nazveme ji tedy konstantou. Ano, je to trochu překvapivé, ale je možné, že v dané oblasti je nějaký fenomenální lokální výrobce vanilkové zmrzliny. Proměnná musí ve vymezené populaci nabývat alespoň dvou hodnot. Velmi důležité je zde to „ve vymezené populaci“. Oblíbená zmrzlina by tedy v jiné populaci (lokaci, věku...) velmi pravděpodobně byla také proměnnou!

11.2 Smysl má samozřejmě pouze otázka a). Barvu očí je možné měřit pouze na nominální škále.

U ordinální škály by mělo smysl klást si otázky a) i b), u intervalové přibude c) a u poměrové by měly smysl všechny otázky. *(Zde lze namítnout, že by bylo možné měřit barvu pomocí jí odpovídající vlnové délky světla. V tom případě by se mohlo jednat až o škálu poměrovou. V psychologické praxi se však s tímto měřením pravděpodobně nesetkáme. Navíc hnědá je nespektrální barva, nemá tedy příslušnou vlnovou délku.)*

11.3 Věk budeme měřit na poměrové škále. Věk dvou dětí je možno porovnávat, je možné zjišťovat rozdíl v jejich věku a dokonce je možné se ptát, kolikrát je jedno starší než druhé. (Škály, u kterých se lze ptát na velikost rozdílu mezi hodnotami, nazýváme metrické. Patří mezi ně intervalové a poměrové.

11.4 Znamku z chování – na ordinální škále ji měřit můžeme (neboť lze porovnat, která známka je vyšší a která nižší), avšak na intervalové už ne (rozdíl mezi známkou 1 a 2 může být naprosto odlišný od rozdílu mezi známkami 2 a 3, zvláště pak u chování).

11.5 Váhu budeme měřit na poměrové škále. Měřit ji na ordinální škále je také možné – změnu na nižší úroveň měření lze provést.

11.6 Lateralita. Dichotomická (= alternativní) znamená, že nabývá ve vymezené populaci pouze dvou hodnot. Na nominální škále měříme ještě vesnici a barvu očí – ta by hypoteticky mohla být také dichotomickou, avšak je velice pravděpodobné, že v uvažované populaci (která je širší než náš desetičlenný vzorek) se nachází i další barvy očí. Tedy barva očí může nabývat i další hodnoty, a proto není dichotomickou.

11.7 Vesnice a barva očí. Polytomická proměnná znamená, že nabývá více než dvou hodnot ve vymezené populaci, jde tedy o protiklad k dichotomické proměnné.

11.8 Spojitá je váha a věk, neboť nejsou omezeny jen na některé dovolené hodnoty, ale mohou nabývat libovolných desetinných hodnot.

Komentář: Pozor, nejde o to, jaké hodnoty se v datech vyskytly (těch je vždy konečně mnoho, což je technicky blízko definice diskrétnosti). Ani nejde o to, zda se v datech vyskytují desetinná čísla. Jde o to, zda jsou definovány/smysluplné i hodnoty mezi hodnotami, které se v datech vyskytly. U diskrétních proměnných (jako např. u četnosti výskytu, či hodnotících škál u dotazníkových položek) definovány nejsou.

11.9 Datová matice obsahuje v každém řádku údaje týkající se jednoho objektu a v každém sloupci obsahuje data pro jednu proměnnou. Za datovou matici ji tedy považovat můžeme. Dalšími úpravami by mohlo být vynechání sloupce s oblíbenou zmrzlinou (který zde není proměnnou) a kódování nečíselných hodnot pomocí čísel, jak je častým zvykem. (Tedy například místo modré barvy očí psát 1 a místo hnědé psát 2. Samozřejmě, v tom případě je nezbytné si poznamenat, co které číslo znamená!)

12.1

cislo	pohlavi	auto	tele	dige
1	1	1	0	1
2	2	1	9	0
3	1	0	1	9
4	1	1	1	0
5	2	0	1	1
6	2	1	1	1

12.2

cislo	pohlavi	zarizeni
1	1	1
2	1	2
3	1	99
4	2	99
5	2	2
6	2	3

12.3 b(cd) f h j l (Ne všechny termíny použité v otázkách existují, či jsou součástí obsahu PSY117.)

Komentář: Používání něčeho kódované 0=nepoužívá, 1=používá je proměnnou *dichotomickou*. Ta je zajímavá tím, že o ní můžeme uvažovat jako ordinální, intervalové i poměrové, podle toho, kde se ve svém uvažování zastavíme. Hodnoty lze podle míry používání něčeho seřadit (bez ohledu na kódování, takže je splněná podmínka pro pořadovou. Interval mezi hodnotami je zde jen jeden. Díky tomu zde nemohou být intervaly různě široké, a tak intervalovosti nestojí nic v cestě. A pokud je nepoužívání kódováno jako 0, pak jde o smysluplnou absolutní nulu – umožňující uvažovat o škále jako poměrové. Takto je dichotomická škála proměnnou, kterou lze popsat jak slovem poměrová, tak kategorická.

12.4 a f h

12.5 Při kódování M=muž a Z=žena: a f h l. Při kódování muž=0, žena=1 viz 12.3.

16.1

cislo	pohlavi	shlednut	minpokuta	maxpokuta
1	1	0	5000	5000
1	1	1	5000	5000
2	1	0	-9	-9
2	1	1	1000	-9
3	2	0	500	1500
3	2	1	500	1500

alternativně (vhodné napří pro párové testy v SPSS, Statistica)

cislo	pohlavi	minpokpre	minpokpo	maxpokpre	maxpokpo
1	1	5000	5000	5000	5000
2	1	-9	1000	-9	-9
3	2	500	500	1500	1500

Komentář: Tento úkol již není tak triviální jako předchozí, protože zahrnuje opakované měření a navíc otevřené otázky, na něž různí respondenti odpovídají různě. Je tedy nutné se nad kódováním zamyslet, aby bylo zachováno co největší množství informace.

Je otázka, jak správně definovat a kódovat poslední proměnnou, protože jsme díky ne zcela vhodnému zadání získali rozporné odpovědi – jeden respondent uvedl jednoznačnou výši pokuty, druhý vymezil spodní hranici pokuty a třetí vymezil její spodní i horní hranici. Pokud bychom chtěli zachovat co nejvíce platných hodnot, pak by asi bylo nejlepším řešením zakódovat odpovědi do více proměnných – dolní mez (minpok), horní mez (maxpok). Problém tohoto řešení je však ten, že nevíme, co přesně myslel svou odpovědí první respondent – zda minimální výši pokuty (jak v tabulce předpokládáme), nebo její průměrnou výši, nebo ještě něco jiného. Jestliže bychom tedy chtěli být precizní, bylo by lepší chápat jeho odpovědi jako chybějící (i když jako jediný odpověděl tak, jak jsme pravděpodobně předpokládali – tedy uvedením jednoho konkrétního čísla). Poučení:

při formulaci otázek v dotaznících je třeba být nanejvýš opatrný a přesný, protože jinak si nemůžeme být jisti, jak si je respondenti vyloží a na co vlastně odpovídají. Takto získaná data mohou být bezcenná, nebo nás přinejmenším připravit o informace od řady cenných respondentů.

16.2 d e g

16.3 b(cd) f h j l

18. poměrová

19. intervalová (porovnáváme dva intervaly pětibodový a šestibodový a jejich srovnání je podmíněno existencí jednotky, což je vlastností intervalové škály). Na poměrové škále bychom toto srovnání mohli učinit v jediné situaci, a to tehdy, kdyby výchozí hodnota Petra i Pavla v prvním měření byla stejná.

21. ano (nelze mít částečné partnery, snad jen s výjimkou Hannibala Lectera), poměrová

22.1 například jedno z řešení může být:

rodina	delka_vz	z_data	z_mista	m_data	m_mista
1	4	9	12	3	6
2	5	7	9	2	9
3	8	7	3	8	12
4	11	6	6	7	11
5	14	8	4	6	6
6	18	11	8	4	10

Je možné dáta zapísať aj vo forme: ID, rodina, pohlavie, dátumy, miesta

22.2 d, f, g

23

ID	skupina	pohlavi	p_kog	pos_hry	hrani1	Hrani2
1	1	1	...		pořád	-999
2	1	2			7	7
3	1	2			nehraji	0
4	1	1			15	15
....						
10	2	2			5-15, to je různé	-999
...						

24

a) ordinálna

b) nominálna

c) intervalová

d) ordinálna (pretože je možné povedať, že vyššie skóre je lepšie ako nižšie skóre, a nie je tu garantované, že rozdiely medzi napr. 2 a 3 reprezentujú rovnaký rozdiel vedomostí ako rozdiely medzi 3 a 4)

e) nominálna

25

	nominálna	poradová	intervalová	pomerová	diskrétna	spojitá
Počet televízorov				x	x	

v domácnostiach						
Výška dvadsaťročných mužov				x		x
Typy áut predávaných v autobazároch	x				x	
Umiestnenie tenistov v ATP		x		x	x	x
Vierovyznanie obyvateľov Košíc	x				x	
Reakčný čas na svetelný podnet				x		x
Preferované spôsoby trávenia víkendu	x				x	
Počet bodov získaných v teste				x		x
Školská známka		x			x	

26.1

ID	F1otazka	F1odhad	F2otazka	F2odhad
1	1	1	2	1
2	3	2	5	5
3	5	5	6	4
4	6	99	4	3
5	2	2	3	2
6	3	3	2	99
...				
60	4	5	2	5

26.2 Kvůli poslední kategorii a, e, k. Kdybychom ji ignorovali, pak b, e, h, k.

27 ...premeníme ich tak, že si zoradíme všetky hodnoty od najnižšej po najvyšiu a priradíme im hodnoty od 1 po....

28

ID	skupina	testdopo	testodpo
1	1		
2	2		
3	3		
4	1		
5	2		
...			

29

	nominálna	poradová	intervalová	pomerová
Je známe poradie hodnôt		x	x	x
Je možné kvantifikovať rozdiely medzi každou hodnotou			x	x
Je možné pričítať a odčítať hodnoty			x	x
Je možné násobiť a deliť hodnoty				x
Je možné určiť početnosti či frekvencie	x	x	x	x
Má „pravú nulu“				x

30.

ID	pohlavie	vek	skusenost	rola	skóre
1	1	25	1	1	17
2	1	22	1	1	13
3	1	45	2	3	25
4	1	43	2	4	21
5	1	-99	2	3	22
6	1	40	2	2	23
...					
12	2	22	1	1	16

30.1 Je to možné a to v tom prípade, ak si povieme, že byť neverný je viac ako byť podvedený a menej ako byť neverný aj podvedený.

31.

- 1.1 The entire student body of your college or university would be considered a population under any circumstances in which you want to generalize *only* to the student body of your college or university and no further.
- 1.3 The students of your college or university are a nonrandom sample of U.S. students, for example, because all U.S. students do not have an equal chance of being included in the sample.
- 1.5 Independent variables: (a) First grade students who attended Kindergarten versus those who did not. (b) Seniors, Masters, Submasters, and Juniors as categories of marathon runners. Dependent variables: (a) Social-adjustment scores assigned by first-grade teachers. (b) Time to run 26 miles, 385 yards.
- 1.7 Continuous variables: (a) Length of gestation. (b) Typing speed in words/minute. (c) Level of serotonin in a particular subcortical nucleus.
- 1.9 The planners of a marathon race would like to know the average times of Senior, Master, Submaster, and Junior runners so as to facilitate planning for handling the finish line.
- 1.11 Categorical data: (a) The number of Brown University students in an October, 1984, referendum voting For and the number voting Against the university's stockpiling suicide pills in case of nuclear disaster. (b) The number of students in a small Midwestern college who are white, African-American, Hispanic-American, Asian, Native American, Alaskan Native, or Other. (c) One year after an experimental program to treat alcoholism, the number of participants who are "still on the wagon", "drinking without having sought additional treatment", or "again under treatment".
- 1.13 Children's scores in an inner-city elementary school could be reported numerically (a measurement variable), or the children could be categorized as Bluebirds ($X > 90$), Robins ($X = 70-90$), or Cardinals ($X < 70$).
- 1.15 For adults of a given height and sex, weight is a ratio scale of body weight, but it is *at best* an ordinal scale of physical health.
- 1.17 Speed is probably a much better index of motivation than of learning.
- 1.19
 - a. The final grade point averages for low-achieving students taking courses that interested them could be compared with the averages of low-achieving students taking courses that don't interest them.
 - b. The quality of communication could be compared for happily versus unhappily married couples.