

Cvičení 1: Bodové zpracování četností

Příklad na bodové zpracování četností: U 20 studentů 1. ročníku byly zjišťovány známky z matematiky, angličtiny a údaje o pohlaví (viz skripta Popisná statistika, příklad 2.4). Příslušný datový soubor se jmenuje **znamky.sta**.

Úkol 1.: Do programu STATISTICA načtete datový soubor **znamky.sta**.

Úkol 2.: Znaky nazvěte X, Y, Z, vytvořte jim návěští (X - známka z matematiky, Y - známka z angličtiny, Z - pohlaví studenta) a popište, co znamenají jednotlivé varianty (u znaků X a Y: 1 - výborně, 2 - velmi dobře, 3 - dobře, 4 - neprospěl, u znaku Z: 0 - žena, 1 - muž). Soubor uložte.

Návod: Kurzor nastavíme na Prom1 – 2x klikneme myší – Jméno X – Dlouhé jméno známka z matematiky, Text. hodnoty – 1 výborně, 2 velmi dobře, 3 dobře, 4 neprospěl, OK.

U proměnné Y lze textové hodnoty okopírovat z proměnné X – v Editoru textových hodnot zvolíme Kopírovat z proměnné X.

Přepínání mezi číselnými hodnotami a jejich textovým popisem se děje pomocí tlačítka s ikonou štítku.

Úkol 3.: U znaků X a Y vypočtete absolutní četnosti, relativní četnosti a relativní kumulativní četnosti.

Návod: Statistika – Základní statistiky a tabulky – Tabulky četností – OK – Proměnné X, Y – OK – Výpočet.

Pro proměnnou X dostaneme tabulku:

Kategorie	Tabulka četností: X: známka z matematiky (znamky)			
	Četnost	Kumulativní četnost	Rel.četnost	Kumulativní rel.četnost
výborně	7	7	35,00000	35,0000
velmi dobře	3	10	15,00000	50,0000
dobře	2	12	10,00000	60,0000
neprospěl	8	20	40,00000	100,0000
ChD	0	20	0,00000	100,0000

Aby student uspěl, musí mít známku 1, 2 nebo 3. Ve sloupci Kumulativní četnost vidíme, že takových studentů bylo 12. Ve sloupci Kumulativní relativní četnost zjistíme, že úspěšných studentů z matematiky bylo 60 %.

Analogicky zjistíme požadované údaje pro úspěšné angličtináře:

Kategorie	Tabulka četností: Y: známka z angličtiny (znamky)			
	Četnost	Kumulativní četnost	Rel.četnost	Kumulativní rel.četnost
výborně	4	4	20,00000	20,0000
velmi dobře	4	8	20,00000	40,0000
dobře	7	15	35,00000	75,0000
neprospěl	5	20	25,00000	100,0000
ChD	0	20	0,00000	100,0000

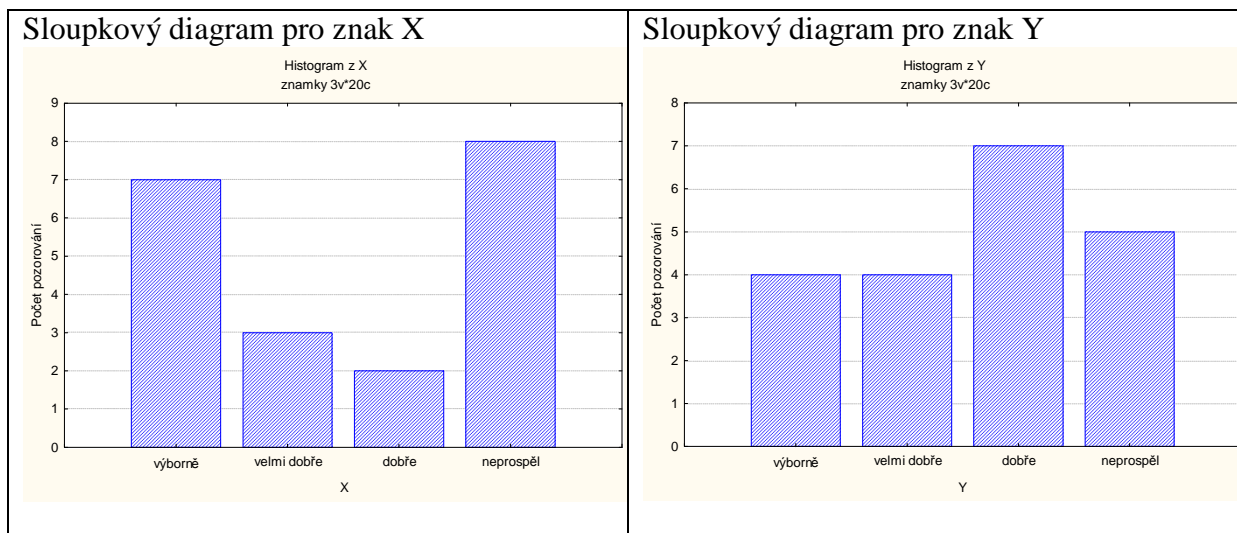
Úspěšných angličtinářů bylo 15, tj. 75 %.

Úkol 4.: Vytvořte

- sloupkový diagram absolutních četností znaků X a Y,
- polygon absolutních četností znaků X a Y
- graf četnostní funkce znaku X,
- graf empirické distribuční funkce znaku X.

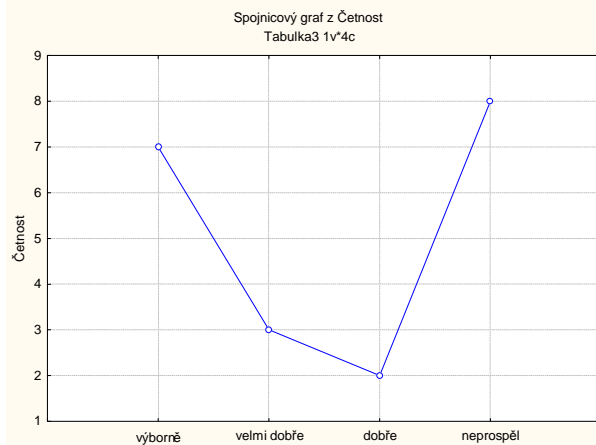
Návod:

ad a) Grafy – Histogramy – Proměnné X, Y – OK- vypneme Normální proložení – Detaily– zaškrtneme Mezery mezi sloupci - OK.

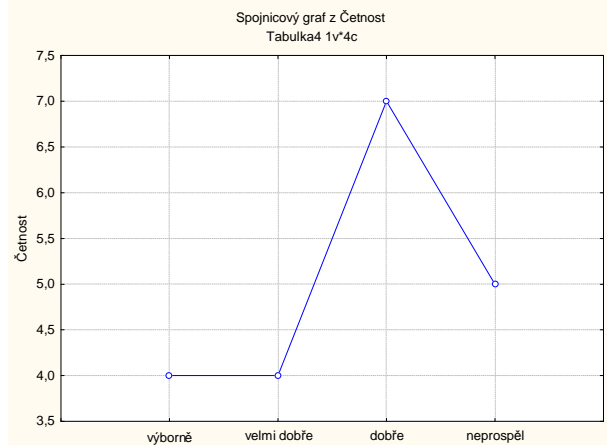


ad b) V pracovním sešitu vstoupíme do tabulky rozložení četností proměnné X resp. Y. Nastavíme se na řádek označený ChD. Pomocí Případy – Odstranit vymažeme tento řádek. Nastavíme se kurzorem na Četnost - klikneme pravým tlačítkem – Grafy bloku dat – Spojnicový graf: celé sloupce. Vykreslí se polygon absolutních četností.

Polygon absolutních četností pro znak X

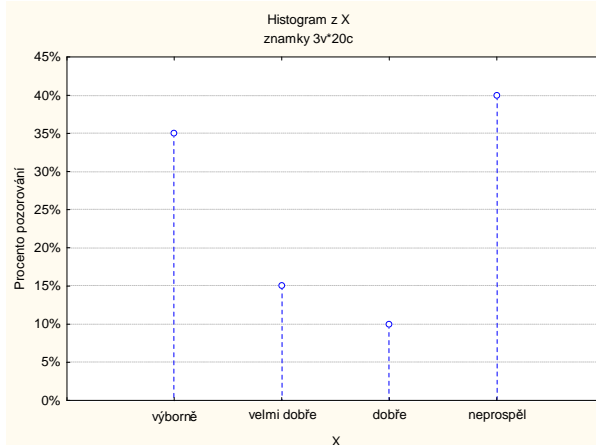


Polygon absolutních četností pro znak Y

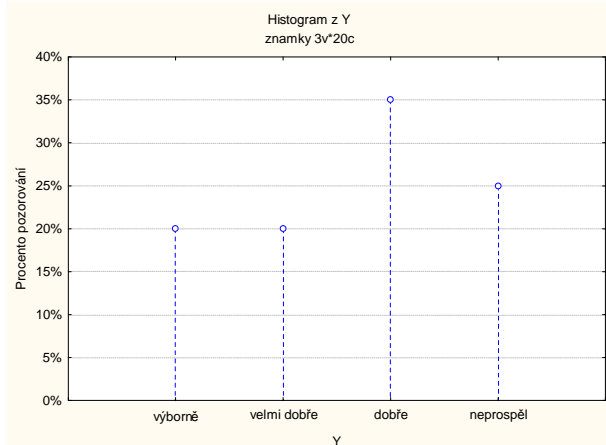


ad c) Při tvorbě histogramu vypneme Normální proložení, zadáme v Detailech Osa Y % - 2 x klikneme myší na pozadí grafu – vybereme Spojnice: Obecné – zaškrtneme Značky – vybereme Sloupce – Typ: Čáry, nastavíme čárkovanou čáru.

Graf četnostní funkce pro znak X

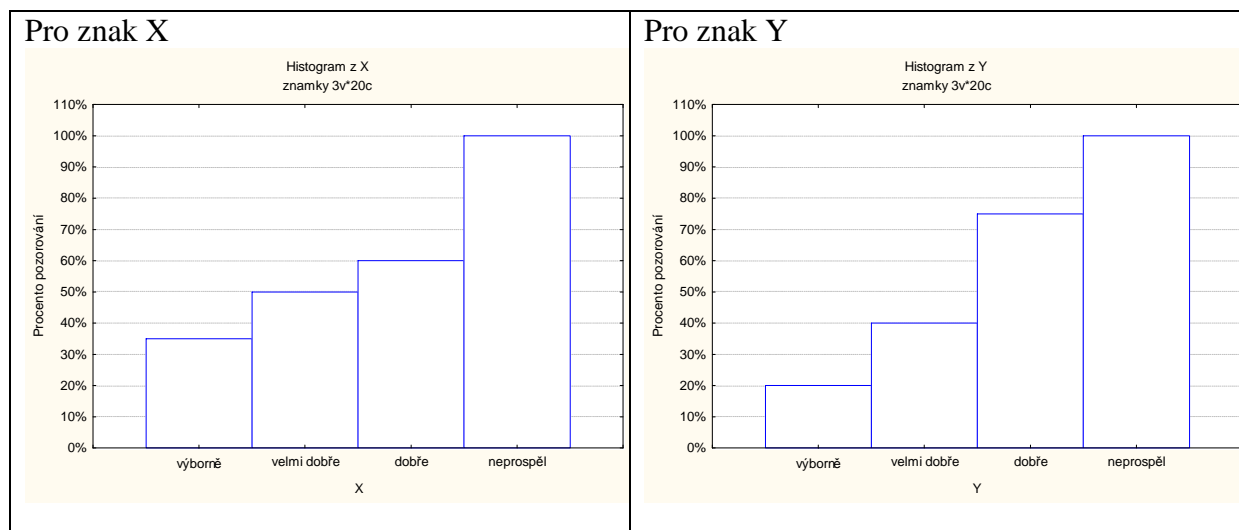


Pro znak Y



ad d) Při tvorbě histogramu vypneme Normální proložení, zadáme v Detailech volbu Zobrazovaný typ: Kumulativní, Osa Y % - 2x klikneme myší na pozadí grafu – vybereme Spojnice: Sloupce – Typ: Obdélníky.

Upozornění: V tomto grafu se objeví svislé čáry, které samozřejmě do grafu empirické distribuční funkce nepatří, avšak v systému STATISTICA je nelze jednoduše odstranit.



Úkol 5.: Vytvořte variační řady známek z matematiky a angličtiny pouze

- a) pro ženy,
- b) pro muže.

Návod:

ad a) Statistika – Základní statistiky a tabulky – Tabulky četností – OK – Proměnné X, Y – OK – vybereme Select Cases - zaškrtneme Zapnout filtr – do okénka některé, vybrané pomocí výrazu napíšeme $Z = 0$, OK, Výpočet.

Variační řada známek z matematiky pro ženy:

Kategorie	Tabulka četností: X: známka z matematiky (znamky) Zhrnout podmínku: Z=0			
	Četnost	Kumulativní četnost	Rel. četnost	Kumulativní rel. četnost
výborně	5	5	50,0000	50,0000
velmi dobře	2	7	20,0000	70,0000
dobře	1	8	10,0000	80,0000
neprospěl	2	10	20,0000	100,0000
ChD	0	10	0,0000	100,0000

Variační řada známek z angličtiny pro ženy:

Kategorie	Tabulka četností: Y: známka z angličtiny (znamky) Zhrnout podmínku: Z=0			
	Četnost	Kumulativní četnost	Rel. četnost	Kumulativní rel. četnost
výborně	4	4	40,0000	40,0000
velmi dobře	2	6	20,0000	60,0000
dobře	1	7	10,0000	70,0000
neprospěl	3	10	30,0000	100,0000
ChD	0	10	0,0000	100,0000

ad b) Statistika – Základní statistiky a tabulky – Tabulky četností – OK – Proměnné X, Y – OK – vybereme Select Cases - zaškrtneme Zapnout filtr – do okénka některé, vybrané pomocí výrazu napíšeme $Z = 1$, OK, Výpočet.

Variační řada známek z matematiky pro muže:

Kategorie	Tabulka četností: X: známka z matematiky (znamky) Zhrnout podmínku: Z=1			
	Četnost	Kumulativní četnost	Rel.četnost	Kumulativní rel.četnost
výborně	2	2	20,00000	20,0000
velmi dobře	1	3	10,00000	30,0000
dobře	1	4	10,00000	40,0000
neprospěl	6	10	60,00000	100,0000
ChD	0	10	0,00000	100,0000

Variační řada známek z angličtiny pro muže:

Kategorie	Tabulka četností: Y: známka z angličtiny (znamky) Zhrnout podmínku: Z=1			
	Četnost	Kumulativní četnost	Rel.četnost	Kumulativní rel.četnost
velmi dobře	2	2	20,00000	20,0000
dobře	6	8	60,00000	80,0000
neprospěl	2	10	20,00000	100,0000
ChD	0	10	0,00000	100,0000

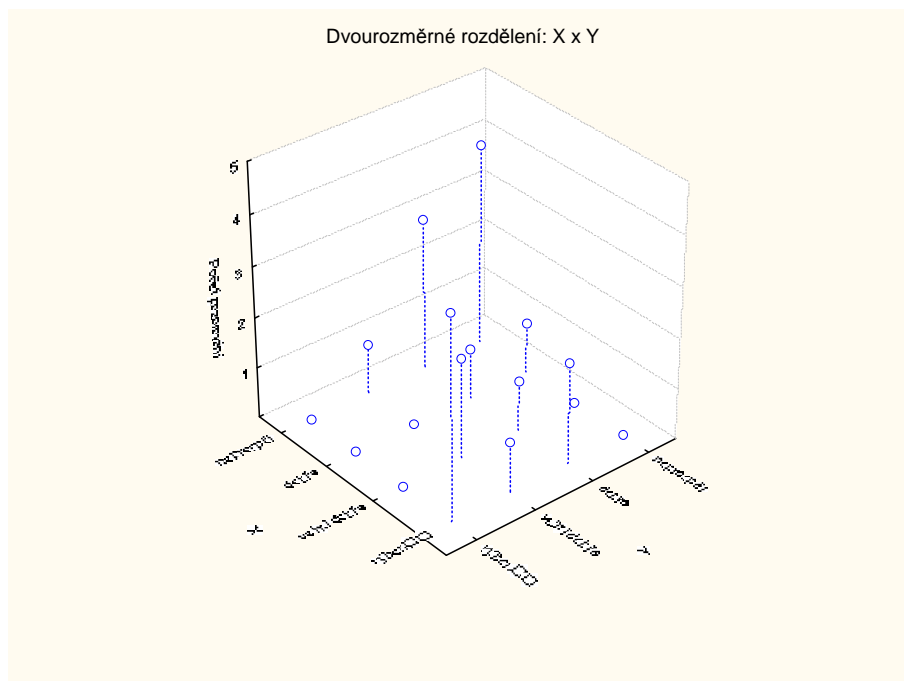
Úkol 6.: Nadále pracujte s celým datovým souborem. Vytvořte kontingenční tabulku absolutních četností znaků X a Y a graf simultánní četnostní funkce.

Návod: Statistiky – Základní statistiky/tabulky – odškrtneme Zapnout filtr – OK - Kontingenční tabulky – OK – Specif. tabulky - List 1 X, List 2 Y, OK, Výpočet.

Kontingenční tabulka (znamky) Četnost označených buněk > 10 (Marginální součty nejsou označeny)					
X	Y	Y	Y	Y	Řádk. součty
	v ý b o r n ě	v e l m i d o b ř e	d o b ř e	n e p r o s p ě l	
v ý b o r n ě	4	1	2	0	7
v e l m i d o b ř e	0	2	1	0	3
d o b ř e	0	0	1	1	2
n e p r o s p ě l	0	1	3	4	8
V š . s k u p .	4	4	7	5	20

Vidíme, že ve výběrovém souboru byly 4 studenti, kteří měli z obou předmětů „výborně“, jeden student, který měl z matematiky „výborně“ a z angličtiny „velmi dobře“ atd. až 4 studenti, kteří z obou předmětů neprospěli.

Vytvoření grafu simultánní četnostní funkce: Na liště aktivujeme Výsledky: kontingenční tabulky – Detaily - 3D histogramy. Vzniklý graf je třeba upravit: 2x klikneme myší na pozadí grafu – Graf: Vzhled – Typ – Špičky – OK.



Graf lze natáčet pomocí Zorného bodu.

Úkol 7.: Vytvořte kontingenční tabulku sloupcově a řádkově podmíněných relativních četností znaků X a Y.

Návod: Aktivujeme na liště Výsledky: kontingenční tabulky – Možnosti - zaškrtneme ve sloupci Výpočet tabulek volbu Procenta z počtu ve sloupci (resp. Procenta z počtu v řádku) – Výpočet.

Kontingenční tabulka sloupcově podmíněných relativních četností :

Kontingenční tabulka (znamky)						
Četnost označených buněk > 10						
(Marginální součty nejsou označeny)						
	X	Y	Y	Y	Y	Řádk. součty
		výborně	velmi dobře	dobře	neprospěl	
Četnost	výborně	4	1	2	0	7
Sloupc. četn.		100,00%	25,00%	28,57%	0,00%	
Četnost	velmi dobře	0	2	1	0	3
Sloupc. četn.		0,00%	50,00%	14,29%	0,00%	
Četnost	dobře	0	0	1	1	2
Sloupc. četn.		0,00%	0,00%	14,29%	20,00%	
Četnost	neprospěl	0	1	3	4	8
Sloupc. četn.		0,00%	25,00%	42,86%	80,00%	
Četnost	Vš.skup.	4	4	7	5	20

Interpretace např. 4. řádku ve 2. sloupci: V souboru byli 4 studenti, kteří měli velmi dobře z angličtiny. Mezi nimi byl jeden, který neprospěl z matematiky, což představuje $1/4 = 25\%$.

Kontingenční tabulka řádkově podmíněných relativních četností:

Kontingenční tabulka (znamky) Četnost označených buněk > 10 (Marginální součty nejsou označeny)						
	X	Y	Y	Y	Y	Řádk. součty
		v ýborn ě	velmi dobře	dobře	neprospěl	
Četnost	v ýborn ě	4	1	2	0	7
Řádk. četn.		57,14%	14,29%	28,57%	0,00%	
Četnost	velmi dobře	0	2	1	0	3
Řádk. četn.		0,00%	66,67%	33,33%	0,00%	
Četnost	dobře	0	0	1	1	2
Řádk. četn.		0,00%	0,00%	50,00%	50,00%	
Četnost	neprospěl	0	1	3	4	8
Řádk. četn.		0,00%	12,50%	37,50%	50,00%	
Četnost	Vš.skup.	4	4	7	5	20

Interpretace např. 2. sloupce ve 4. řádku: V souboru bylo 8 studentů, kteří neprospěli z matematiky. Mezi nimi byl jeden, který měl velmi dobře z angličtiny, což představuje $1/8 = 12,5 \%$.