

Téma 1.: Bodové a intervalové rozložení četností

Nejprve budeme pracovat s datovým souborem znamky.sta, který obsahuje údaje o známkách z matematiky, angličtiny a pohlaví 20 studentů 1. ročníku.

Úkol 1.: Načtěte soubor znamky.sta. Proměnným X, Y, Z vytvořte návěští (X - známka z matematiky, Y - známka z angličtiny, Z - pohlaví studenta). 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).

Návod: Soubor – Otevřít – vybereme příslušný adresář se souborem znamky.sta – Otevřít. 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 2.: Vytvořte

- variační řadu známek z matematiky a známek z angličtiny,
- sloupkový diagram absolutních četností znaků X a Y,
- polygon absolutních četností znaků X a Y

Návod:

ad a) Statistika – Základní statistiky a tabulky – Tabulky četností – OK – Proměnné X, Y – OK - Výpočet.

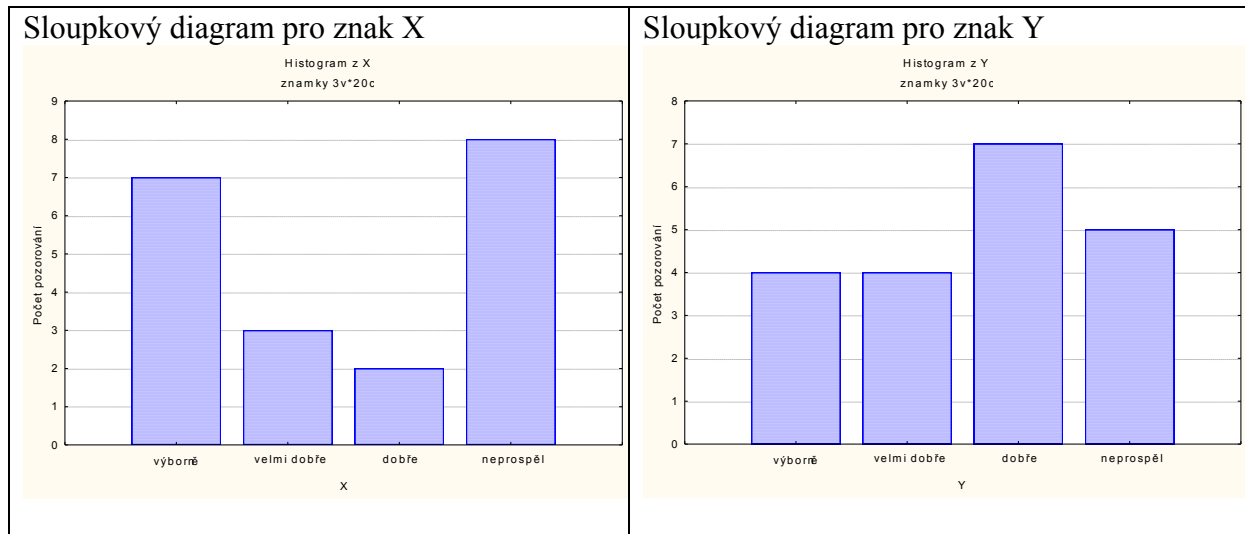
Variační řada známek z matematiky

Kategorie	Tabulka četností: X: známka z M (znamky .sta)			
	Četnost	Kumulativní četnost	Rel. četnost	Kumulativní rel. četnost
vyborne	7	7	35,00000	35,0000
velmi dobre	3	10	15,00000	50,0000
dobre	2	12	10,00000	60,0000
neprospel	8	20	40,00000	100,0000
ChD	0	20	0,00000	100,0000

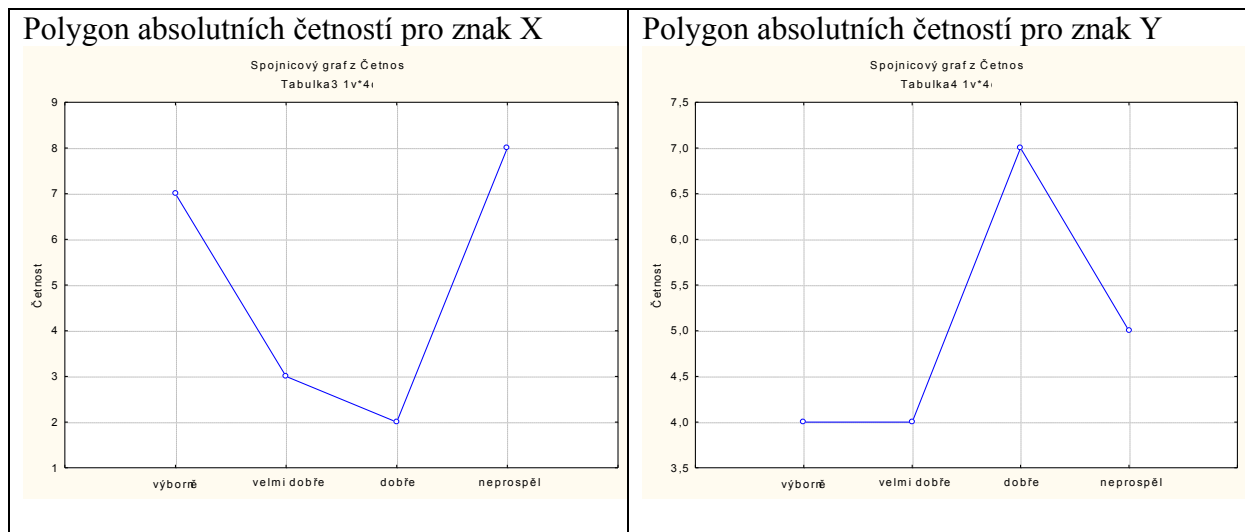
Variační řada známek z angličtiny

Kategorie	Tabulka četností: Y: známka z A (znamky .sta)			
	Četnost	Kumulativní četnost	Rel. četnost	Kumulativní rel. četnost
vyborne	4	4	20,00000	20,0000
velmi dobre	4	8	20,00000	40,0000
dobre	7	15	35,00000	75,0000
neprospel	5	20	25,00000	100,0000
ChD	0	20	0,00000	100,0000

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



ad c) 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.



Úkol 2.: 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 zapíšeme Z = 0, OK, Výpočet.

Variační řada známek z matematiky pro ženy:

Tabulka četností: X: známka z matematiky (znamky) Zhrnout podmínku: Z=0				
Kategorie	Četnost	Kumulativní četnost	Rel. četnost	Kumulativní rel. četnost
v ýbom ě	5	5	50,00000	50,0000
velmi dobře	2	7	20,00000	70,0000
dobře	1	8	10,00000	80,0000
neprospěl	2	10	20,00000	100,0000
ChD	0	10	0,00000	100,0000

Variační řada známek z angličtiny pro ženy:

Tabulka četností: Y: známka z angličtiny (znamky) Zhrnout podmínku: Z=0				
Kategorie	Četnost	Kumulativní četnost	Rel. četnost	Kumulativní rel. četnost
v ýbom ě	4	4	40,00000	40,0000
velmi dobře	2	6	20,00000	60,0000
dobře	1	7	10,00000	70,0000
neprospěl	3	10	30,00000	100,0000
ChD	0	10	0,00000	100,0000

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

Variační řada známek z matematiky pro muže:

Tabulka četností: X: známka z matematiky (znamky) Zhrnout podmínku: Z=1				
Kategorie	Četnost	Kumulativní četnost	Rel. četnost	Kumulativní rel. četnost
v ýbom ě	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:

Tabulka četností: Y: známka z angličtiny (znamky) Zhrnout podmínku: Z=1				
Kategorie	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 3.: Nadále budeme pracovat s celým datovým souborem. Vytvoříme kontingenční tabulku simultánních absolutních četností znaků X a Y a graf simultánní četnostní funkce.

Návod: Statistiky – Základní statistiky/tabulky – odškrtneme Zapnout filtr – OK - Kontingenční tabulky – OK – Select cases — 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.
	v ýbom ě	velmi dobře	dobře	neprospěl	součty
v ýbom ě	4	1	2	0	7
velmi dobře	0	2	1	0	3
dobře	0	0	1	1	2
neprospěl	0	1	3	4	8
Vš. skup.	4	4	7	5	20

Vidíme, že ve výběrovém souboru byli 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.

Úkol 4.: 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.
	v ýbom ě	velmi dobře	dobře	neprospěl		součty
Četnost	v ýbom ě	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 ýbom ě	ř	velmi dobře	dobře	neprospěl	
Četnost	v ýbom 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 $1/8 = 12,5\%$.

Nyní se budeme věnovat datovému souboru ocel.sta. Obsahuje údaje o mezi plasticity a mezi pevnosti 60 vzorků oceli.

Úkol 5.: Načtete soubor ocel.sta. Proměnným X a Y vytvořte návěští „mez plasticity“ a „mez pevnosti“. Podle Sturgersova pravidla najdete optimální počet třídících intervalů pro znaky X a Y a vhodně stanovíte meze třídících intervalů.

Návod: Soubor – Otevřít – vybereme příslušný adresář se souborem ocel.sta – Otevřít. Kurzor nastavíme na X – 2x klikneme myší – Dlouhé jméno mez plasticity – OK, kurzor nastavíme na Y – 2x klikneme myší – Dlouhé jméno mez pevnosti – OK.

Protože případů je 60, podle Sturgersova pravidla je optimální počet třídících intervalů 7.

Musíme zjistit minimum a maximum, abychom vhodně stanovili třídící intervaly: Statistiky - Základní statistiky/tabulky – Popisné statistiky - OK - Proměnné X, Y – OK – Detailní výsledky – ponecháme zaškrtnuté Minimum&maximum – Výpočet.

Pro X je minimum 33 a maximum 160, tedy dolní mez prvního třídícího intervalu volíme 30, horní mez posledního třídícího intervalu 170. Celkem tedy třídící intervaly pro znak X budou: (30,50>, (50,70>, (70,90>, (90,110>, (110,130>, (130,150>, (150,170>

Pro Y je minimum 52 a maximum 189, tedy dolní mez prvního třídícího intervalu volíme 50, horní mez posledního třídícího intervalu 190. Celkem tedy třídící intervaly pro znak Y budou: (50,70>, (70,90>, (90,110>, (110,130>, (130,150>, (150,170>, (170,190>.

Úkol 6.: Proved'te zakódování hodnot proměnných X a Y do příslušných třídících intervalů. Všem hodnotám proměnné X, které leží v intervalu (30,50>, přiřaďte hodnotu 1 atd. až všem hodnotám proměnné X, které leží v intervalu (170,190>, přiřaďte hodnotu 7. Analogicky pro proměnnou Y.

Návod: Vytvoříme dvě nové proměnné: Vložit – Přidat proměnné – 2 – Za Y – OK – přejmenujeme je na RX a RY. Nastavíme se kurzorem na RX – Data – Překódovat - vyplníme podmínky pro všech 7 kategorií. (Pozor – podmínky píšeme ve tvaru $X > 30$ and $X \leq 50$ atd.). Pak klepneme na OK.

Překódovat hodnoty proměnné 3: RX

Kategorie 1 Zahmout pokud: $X > 30 \text{ and } X \leq 50$	Nová hodnota 1 <input checked="" type="radio"/> hodnota: 1 <input type="radio"/> Kód ChD
Kategorie 2 Zahmout pokud: $X > 50 \text{ and } X \leq 70$	Nová hodnota 2 <input checked="" type="radio"/> hodnota: 2 <input type="radio"/> Kód ChD
Kategorie 3 Zahmout pokud: $X > 70 \text{ and } X \leq 90$	Nová hodnota 3 <input checked="" type="radio"/> hodnota: 3 <input type="radio"/> Kód ChD
Kategorie 4 Zahmout pokud: $X > 90 \text{ and } X \leq 110$	Nová hodnota 4 <input checked="" type="radio"/> hodnota: 4 <input type="radio"/> Kód ChD

Jiné
Pokud nejsou splněny žádné podmínky, nastavit hodnoty na:
 Kód ChD
 hodnota:
 nezměněna

OK
Storno
Zrušit vše
Otevřít...
Uložit jako...
Proměnná...

Analogicky překódujeme hodnoty proměnné Y do proměnné RY.

Překódovat hodnoty proměnné 4: RY

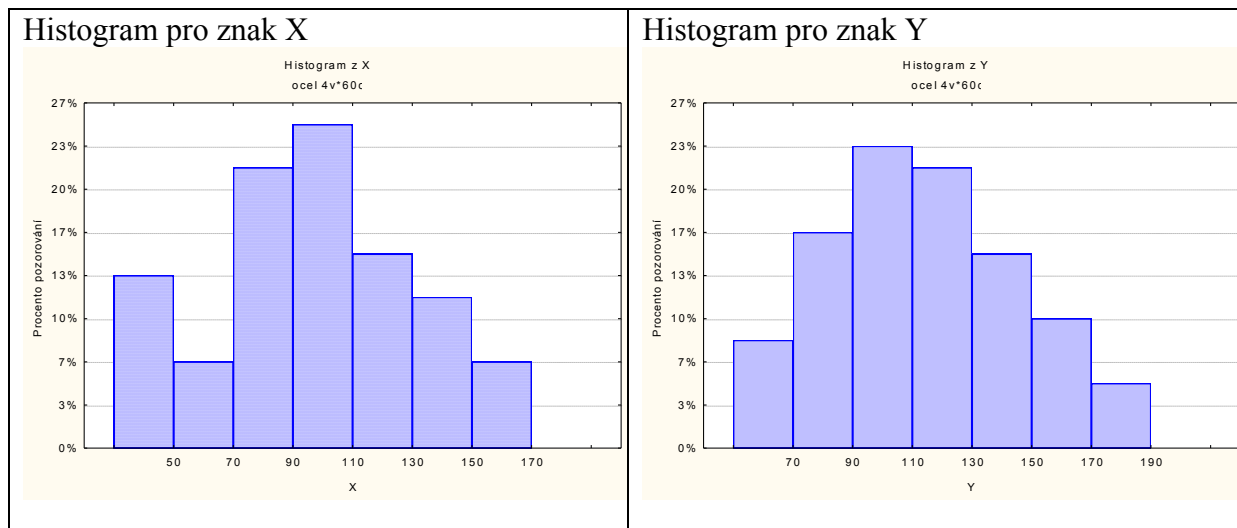
Kategorie 1 Zahmout pokud: $Y > 50 \text{ and } Y \leq 70$	Nová hodnota 1 <input checked="" type="radio"/> hodnota: 1 <input type="radio"/> Kód ChD
Kategorie 2 Zahmout pokud: $Y > 70 \text{ and } Y \leq 90$	Nová hodnota 2 <input checked="" type="radio"/> hodnota: 2 <input type="radio"/> Kód ChD
Kategorie 3 Zahmout pokud: $Y > 90 \text{ and } Y \leq 110$	Nová hodnota 3 <input checked="" type="radio"/> hodnota: 3 <input type="radio"/> Kód ChD
Kategorie 4 Zahmout pokud: $Y > 110 \text{ and } Y \leq 130$	Nová hodnota 4 <input checked="" type="radio"/> hodnota: 4 <input type="radio"/> Kód ChD

Jiné
Pokud nejsou splněny žádné podmínky, nastavit hodnoty na:
 Kód ChD
 hodnota:
 nezměněna

OK
Storno
Zrušit vše
Otevřít...
Uložit jako...
Proměnná...

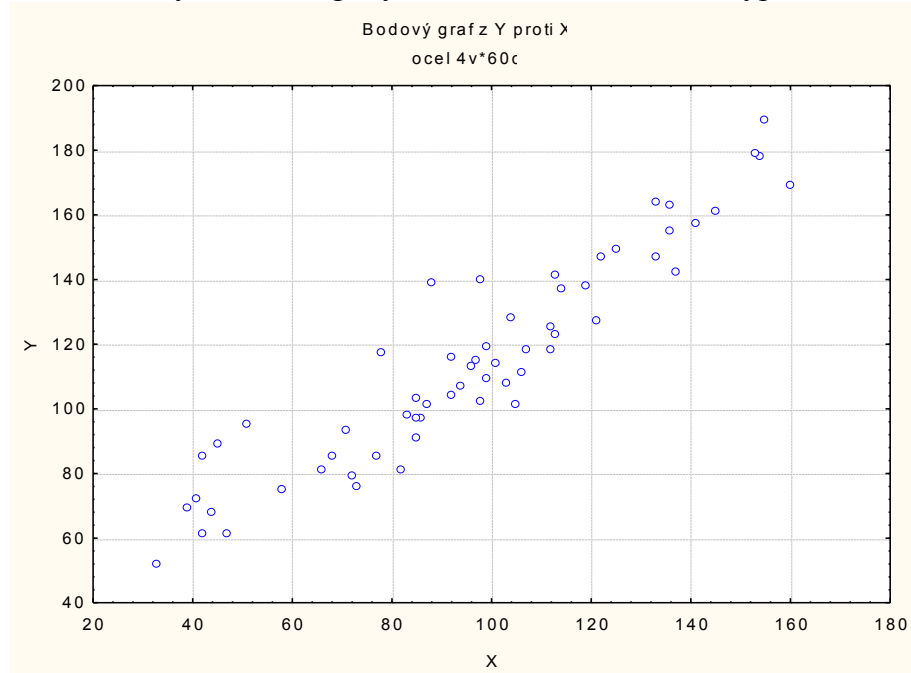
Úkol 7.: Vytvořte histogram pro X a pro Y.

Návod: Grafy – Histogramy – Proměnné X – vypneme Normální proložení – Details – zaškrtneme Hranice – Určit hranice – 50 70 90 110 130 150 170 OK – Osa Y %. Po vykreslení histogramu lze 2 x klepnout na pozadí grafu a ve volbě Všechny možnosti měnit různé vlastnosti grafu.



Úkol 8.: Nakreslete dvourozměrný tečkový diagram pro (X,Y).

Návod: Grafy – Bodové grafy – Proměnné X,Y – OK - vypneme Lineární proložení – OK.



Vidíme, že mezi oběma proměnnými existuje určitý stupeň přímé lineární závislosti – s růstem hodnot meze plasticity vesměs rostou hodnoty meze pevnosti a naopak.