

Téma 2: Intervalové zpracování četností

Vzorový příklad: U 60 vzorků oceli byly zjišťovány hodnoty meze plasticity a meze pevnosti v kpcm^2 (viz skripta Popisná statistika, př. 2.5). Datový soubor se jmenuje **ocel.sta**. Proveďte intervalové zpracování četností.

Postup ve STATISTICE:

1. Načtete soubor **ocel.sta**. Proměnným X a Y vytvořte návěští „mez plasticity“ a „mez pevnosti“.
2. Pro X a Y použijeme intervalové zpracování četností. Podle Sturgesova pravidla je optimální počet třídících intervalů 7. Musíme zjistit minimum a maximum, abychom vhodně stanovili třídící intervaly.

Návod: Statistics – Basic Statistics/Tables – Descriptive statistics - Variables X,Y – zaškrtneme Minimum&maximum – Summary. (Pro X je minimum 33 a maximum 160, tedy vhodná volba třídících intervalů je $(30,50>$, $(50,70>$, ..., $(150,170>$, pro Y je minimum 52 a maximum 189, tedy třídící intervaly zvolíme $(50,70>$, $(70,90>$, ... , $(170,190>$)

Řešení:

Variable	Descriptive Statistics (c)	
	Minimum	Maximum
X	33,00000	160,0000
Y	52,00000	189,0000

U znaku X volíme dolní mez prvního třídícího intervalu 30, horní mez posledního třídícího intervalu 170. U znaku Y volíme dolní mez prvního třídícího intervalu 50, horní mez posledního třídícího intervalu 190.

Celkem tedy třídící intervaly znak X budou:

$(30,50>$, $(50,70>$, $(70,90>$, $(90,110>$, $(110,130>$, $(130,150>$, $(150,170>$

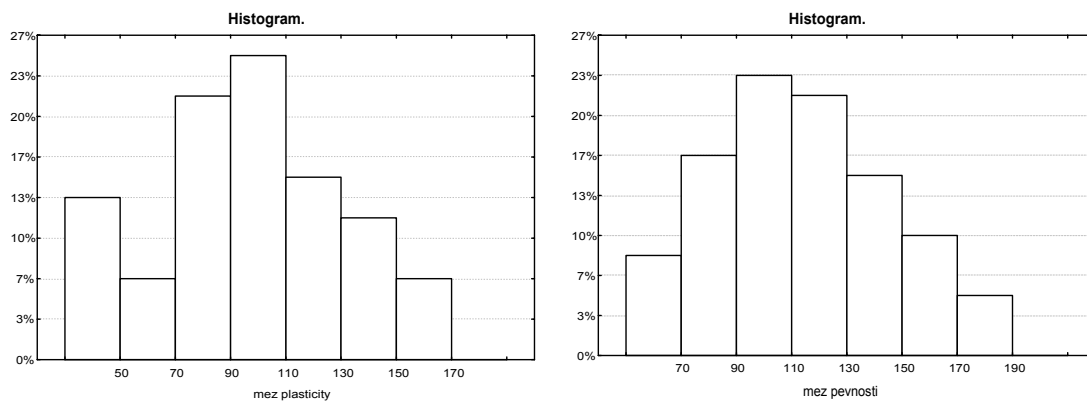
a pro znak Y:

$(50,70>$, $(70,90>$, $(90,110>$, $(110,130>$, $(130,150>$, $(150,170>$, $(170,190>$.

3. Vytvořte histogram pro X a pro Y.

Návod: Graphs – Histograms – Variables X – vypneme Normal fit – Advanced – zaškrtneme Boundaries – Specify Boundaries – 50 70 90 110 130 150 170 OK – Y Axis %.
Po vykreslení histogramu lze 2 x klepnout na pozadí grafu a ve volbě All Options měnit různé vlastnosti grafu.

Řešení:



Komentář: Rozložení četností jak pro mez plasticity tak pro mez pevnosti je lehce nesymetrické. Navíc v histogramu pro mez plasticity je vidět, že interval od 50 do 70 má velmi malé četnostní zastoupení. Vysvětlení této skutečnosti je ovšem mimomatematická záležitost.

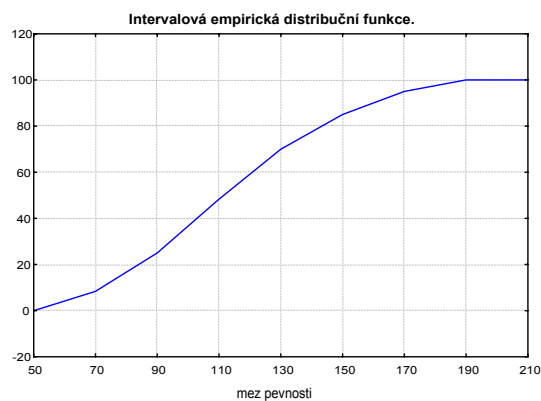
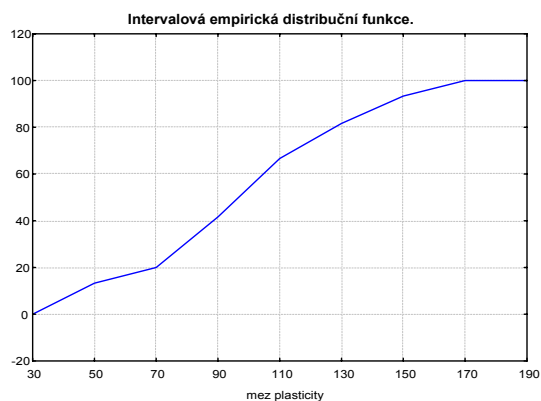
4. Proveďte zakódování hodnot proměnných X a Y do příslušných třídicích intervalů.
Návod: Insert – Add Variables – 2 – After Y – OK – přejmenujeme je na RX a RY.
Nastavíme se kurzorem na RX – Data – Recode - vyplníme podmínky pro všech 7 kategorií.
(Pozor – podmínky se musí psát ve tvaru $X > 30$ and $X \leq 50$ atd.). Pak klepneme na OK.
Analogicky pro Y.
5. Vytvořte graf intervalové empirické distribuční funkce pro X.
Návod: Vytvoříme Frequency table pro RX. Před 1. případ vložíme řádek, kde do Category napíšeme 0 a do Cumulative Percent také 0. Nastavíme se kurzorem na Cumulative Percent – Graphs – Graphs of Block Data – Custom Graph from Block by Column – Line Plots (Variables) – OK.

Řešení:

Kategorie	Tabulka četností:RX (ocel)			
	Četnost	Kumulativní četnost	Rel.četnost	Kumulativní rel.četnost
0				0,0000
1	8	8	13,33333	13,3333
2	4	12	6,66667	20,0000
3	13	25	21,66667	41,6667
4	15	40	25,00000	66,6667
5	9	49	15,00000	81,6667
6	7	56	11,66667	93,3333
7	4	60	6,66667	100,0000
ChD	0	60	0,00000	100,0000

Kategorie	Tabulka četností:RY (ocel)			
	Četnost	Kumulativní četnost	Rel.četnost	Kumulativní rel.četnost
0				0,0000
1	5	5	8,33333	8,3333
2	10	15	16,66667	25,0000
3	14	29	23,33333	48,3333
4	13	42	21,66667	70,0000
5	9	51	15,00000	85,0000
6	6	57	10,00000	95,0000
7	3	60	5,00000	100,0000
ChD	0	60	0,00000	100,0000

Komentář: U meze plasticity je nejvíce zastoupena kategorie 4, tj. interval od 90 do 110, u meze pevnosti kategorie 3, tj. interval od 90 do 110.



6. Sestavte kontingenční tabulky absolutních četností (relativních četností, sloupcově a řádkově podmíněných relativních četností) dvourozměrných třídících intervalů pro (X,Y).
Návod: Viz úkol č. 7 v tématu 1, kde budeme pracovat s proměnnými RX a RY.

Řešení:

Kontingenční tabulky absolutních a relativních četností.

		Summary Frequency Table (ocel)							
		Table: RX(7) x RY(7)							
	RX	RY	RY	RY	RY	RY	RY	RY	Row Totals
		(50,70>	(70,90>	(90,110>	110,130	130,150	150,170	170,190	
Count	(30,50>	5	3	0	0	0	0	0	8
Total Percent		8,33%	5,00%	0,00%	0,00%	0,00%	0,00%	0,00%	13,33%
Count	(50,70>	0	3	1	0	0	0	0	4
Total Percent		0,00%	5,00%	1,67%	0,00%	0,00%	0,00%	0,00%	6,67%
Count	(70,90>	0	4	7	1	1	0	0	13
Total Percent		0,00%	6,67%	11,67%	1,67%	1,67%	0,00%	0,00%	21,67%
Count	(90,110>	0	0	6	8	1	0	0	15
Total Percent		0,00%	0,00%	10,00%	13,33%	1,67%	0,00%	0,00%	25,00%
Count	110,130>	0	0	0	4	5	0	0	9
Total Percent		0,00%	0,00%	0,00%	6,67%	8,33%	0,00%	0,00%	15,00%
Count	(130,150	0	0	0	0	2	5	0	7
Total Percent		0,00%	0,00%	0,00%	0,00%	3,33%	8,33%	0,00%	11,67%
Count	(150,170	0	0	0	0	0	1	3	4
Total Percent		0,00%	0,00%	0,00%	0,00%	0,00%	1,67%	5,00%	6,67%
Count	All Grps	5	10	14	13	9	6	3	60
Total Percent		8,33%	16,67%	23,33%	21,67%	15,00%	10,00%	5,00%	

Kontingenční tabulky řádkově a sloupcově podmíněných relativních četností.

		Summary Frequency Table (ocel)							
		Table: RX(7) x RY(7)							
	RX	RY	RY	RY	RY	RY	RY	RY	Row Totals
		1	2	3	4	5	6	7	
Count	1	5	3	0	0	0	0	0	8
Row Percent		62,50%	37,50%	0,00%	0,00%	0,00%	0,00%	0,00%	
Count	2	0	3	1	0	0	0	0	4
Row Percent		0,00%	75,00%	25,00%	0,00%	0,00%	0,00%	0,00%	
Count	3	0	4	7	1	1	0	0	13
Row Percent		0,00%	30,77%	53,85%	7,69%	7,69%	0,00%	0,00%	
Count	4	0	0	6	8	1	0	0	15
Row Percent		0,00%	0,00%	40,00%	53,33%	6,67%	0,00%	0,00%	
Count	5	0	0	0	4	5	0	0	9
Row Percent		0,00%	0,00%	0,00%	44,44%	55,56%	0,00%	0,00%	
Count	6	0	0	0	0	2	5	0	7
Row Percent		0,00%	0,00%	0,00%	0,00%	28,57%	71,43%	0,00%	
Count	7	0	0	0	0	0	1	3	4
Row Percent		0,00%	0,00%	0,00%	0,00%	0,00%	25,00%	75,00%	
Count	All Grps	5	10	14	13	9	6	3	60

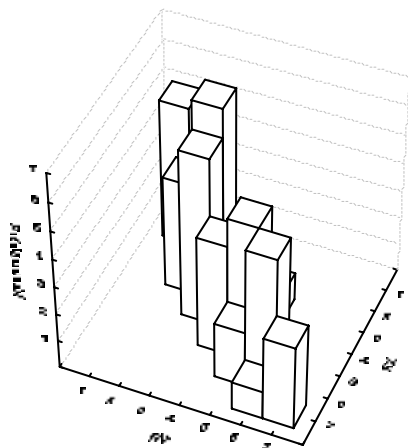
Summary Frequency Table (ocel)									
Table: RX(7) x RY(7)									
	RX	RY 1	RY 2	RY 3	RY 4	RY 5	RY 6	RY 7	Row Totals
Count	1	5	3	0	0	0	0	0	8
Column Percent		100,00%	30,00%	0,00%	0,00%	0,00%	0,00%	0,00%	
Count	2	0	3	1	0	0	0	0	4
Column Percent		0,00%	30,00%	7,14%	0,00%	0,00%	0,00%	0,00%	
Count	3	0	4	7	1	1	0	0	13
Column Percent		0,00%	40,00%	50,00%	7,69%	11,11%	0,00%	0,00%	
Count	4	0	0	6	8	1	0	0	15
Column Percent		0,00%	0,00%	42,86%	61,54%	11,11%	0,00%	0,00%	
Count	5	0	0	0	4	5	0	0	9
Column Percent		0,00%	0,00%	0,00%	30,77%	55,56%	0,00%	0,00%	
Count	6	0	0	0	0	2	5	0	7
Column Percent		0,00%	0,00%	0,00%	0,00%	22,22%	83,33%	0,00%	
Count	7	0	0	0	0	0	1	3	4
Column Percent		0,00%	0,00%	0,00%	0,00%	0,00%	16,67%	100,00%	
Count	All Grps	5	10	14	13	9	6	3	60

7. Vytvořte stereogram pro (RX,RY).

Návod: V tabulce Crosstabulation Tables Result zaškrtneme 3D histograms. Ve volbě Axis Scaling (pro RX i pro RY) změním Mode na Manual – Minimum 0.

Řešení:

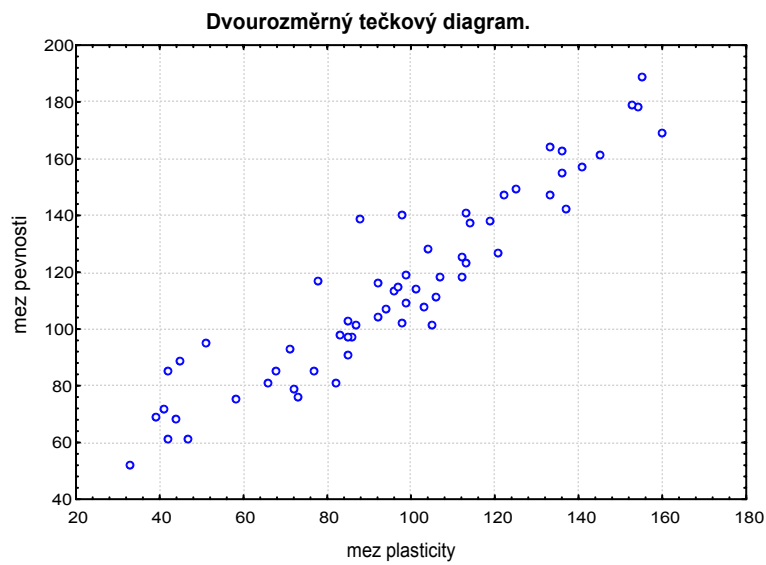
Stereogram: RY x RX



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

Návod: Graphs – Scatterplots – Variables X,Y – OK vypneme Linear fit – OK.

Řešení:



Komentář: Z dvourozměrného tečkového diagramu je patrné, že s růstem (poklesem) hodnot meze plasticity vesměs rostou (klesají) hodnoty meze pevnosti. Znamená to, že mezi oběma znaky existuje určitý stupeň přímé lineární závislosti. Intenzitu této lineární závislosti posuzujeme pomocí koeficientu korelace (viz Téma 4).