

Vysoké učení technické v Brně
Fakulta informačních technologií

Projekt
MSP – Statistika a pravděpodobnost

Číslo zadání: 1, 30
Autor: Filip Kočica

Skupina: Čtvrtek, 11 hod.

Za pomoci sázecího nástroje \LaTeX a tabulkového software LibreOffice calc.

1 Příklad číslo 1

Tabulka 1:

Číslo řádku	Statistický soubor	Uspořádaný soubor
1	0,72	-0,83
2	-0,55	-0,72
3	0,30	-0,71
4	0,12	-0,70
5	0,12	-0,59
6	0,80	-0,55
7	0,63	-0,52
8	-0,71	-0,42
9	-0,42	-0,42
10	-0,06	-0,39
11	0,46	-0,38
12	0,66	-0,30
13	0,63	-0,29
14	0,28	-0,27
15	-0,08	-0,24
16	0,08	-0,23
17	-0,02	-0,20
18	0,44	-0,16
19	0,40	-0,13
20	0,41	-0,13
21	-0,59	-0,08
22	-0,13	-0,06
23	0,33	-0,04
24	0,38	-0,02
25	-0,83	0,07

Tabulka 2:

Číslo řádku	Statistický soubor	Uspořádaný soubor
26	-0,13	0,08
27	0,55	0,12
28	-0,24	0,12
29	0,07	0,21
30	0,44	0,24
31	-0,20	0,28
32	-0,29	0,30
33	1,00	0,32
34	0,21	0,33
35	-0,52	0,38
36	-0,72	0,40
37	0,59	0,41
38	-0,70	0,44
39	-0,42	0,44
40	-0,30	0,44
41	-0,23	0,46
42	-0,04	0,49
43	0,32	0,55
44	-0,16	0,59
45	-0,27	0,63
46	0,49	0,63
47	0,44	0,66
48	0,24	0,72
49	-0,39	0,80
50	-0,38	1,00

1.1 Podsekce a)

$$x_{(1)} = \min_i x_i = -0.83$$

$$x_{(n)} = \max_i x_i = 1.00$$

$$\text{Variační obor: } \langle x_{(1)}, x_{(n)} \rangle = \langle -0.83, 1.00 \rangle$$

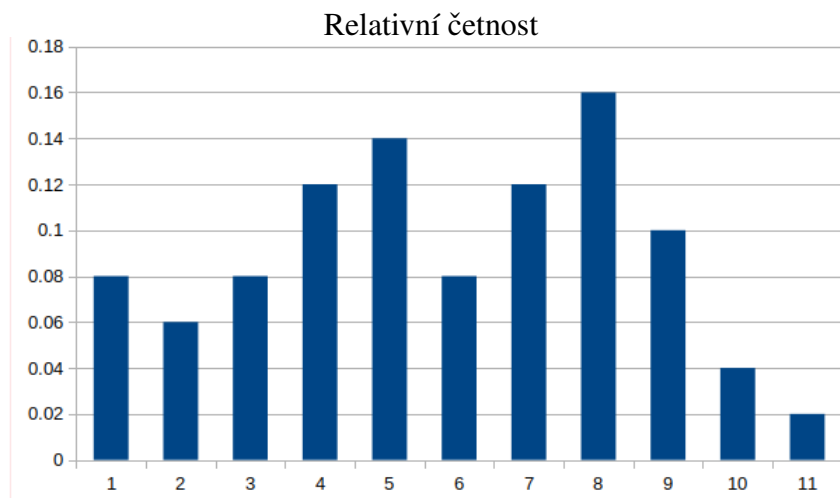
$$\text{Rozpětí: } x_{(n)} - x_{(1)} = 1,83$$

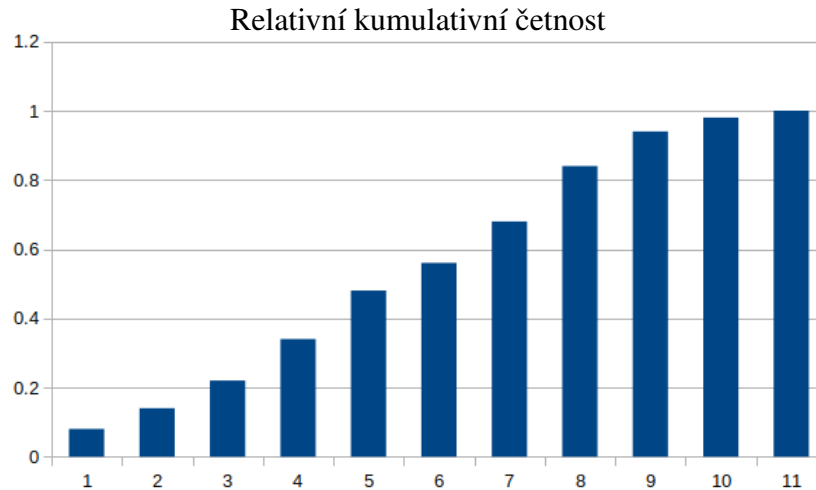
$$\text{Počet tříd } m = 11$$

$$\text{Délka třídy } \frac{x_{(n)} - x_{(1)}}{m} = 0,1663636364$$

Tabulka 3:

Třída	xi-	xi+	Střed třídy	Kum. čet.	Čet.	Relat. čet.	Relat. kum. čet.
1	-0.8300	-0.6636	-0.7468	4	4	0.08	0.08
2	-0.6636	-0.4973	-0.5805	7	3	0.06	0.14
3	-0.4973	-0.3309	-0.4141	11	4	0.08	0.22
4	-0.3309	-0.1645	-0.2477	17	6	0.12	0.34
5	-0.1645	0.0018	-0.0814	24	7	0.14	0.48
6	0.0018	0.1682	0.0850	28	4	0.08	0.56
7	0.1682	0.3345	0.2514	34	6	0.12	0.68
8	0.3345	0.5009	0.4177	42	8	0.16	0.84
9	0.5009	0.6673	0.5841	47	5	0.1	0.94
10	0.6673	0.8336	0.7505	49	2	0.04	0.98
11	0.8336	1.0000	0.9168	50	1	0.02	1





1.2 Podsekce b)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 0.0546$$

$$\text{Medián: } \tilde{x} = 0.075$$

$$\text{Modus: } \hat{x} = 0.44$$

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = 0.20312884$$

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = 0.4506981695$$

1.3 Podsekce c)

$$\text{Bodový odhad střední hodnoty: } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 0.0546$$

$$\text{Bodový odhad rozptylu: } s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 0.2072743265$$

$$\text{Bodový odhad směrodatné odchylky: } s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} = 0.4552739028$$

1.4 Podsekce d)

$$\text{Testovací kritérium: } t = \sum_{j=1}^m \frac{(f_j - \hat{f}_j)^2}{\hat{f}_j} = 3.77033159211426$$

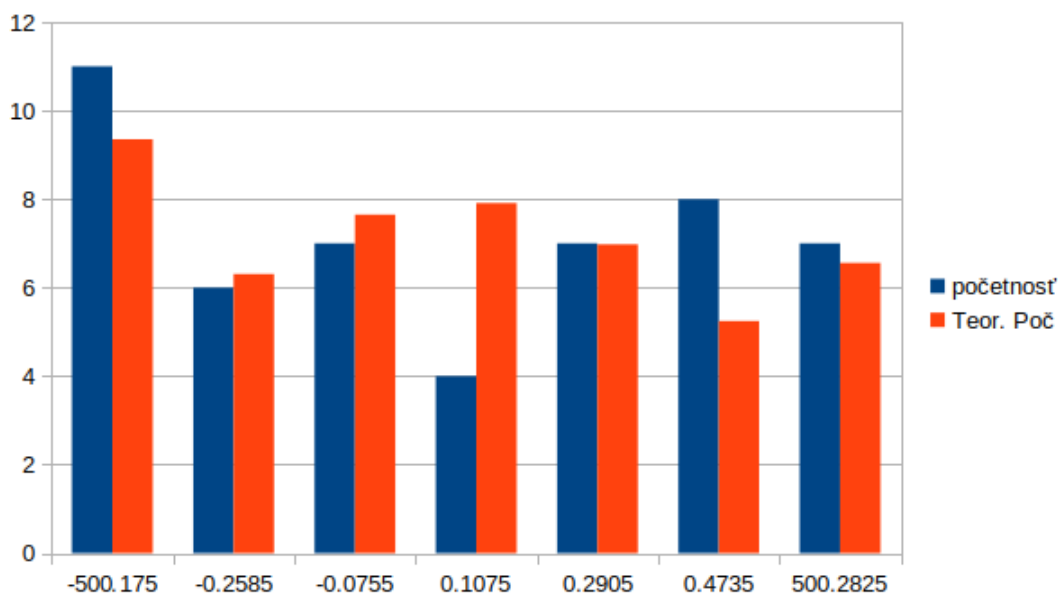
$$\text{Hodnota } \chi^2_{1-\alpha} \text{ pro } k = 7 - 2 - 1 \text{ stupňů volnosti: } 9.48772903678116$$

$$\text{Doplňek kritického oboru: } \overline{W}_\alpha = \langle 0, \chi^2_{1-\alpha} \rangle = \langle 0, 9.48772903678116 \rangle$$

Protože $t \in \overline{W}_\alpha$ se hypotéza: $X \sim N(0.0546, 0.2072743265)$ **nezamítá**.

Tabulka 4:

Třída	xi-	xi+	Střed třídy	Četnost	Teoretická četnost	Roz ² /Teor. četnost
1	-1000	-0.35	-500.175	11	9.35416487730508	0.28957937845079
2	-0.35	-0.167	-0.2585	6	6.30690075047425	0.014934129197225
3	-0.167	0.016	-0.0755	7	7.64975971221816	0.055189666016242
4	0.016	0.199	0.1075	4	7.91134169844347	1.93375466072874
5	0.199	0.382	0.2905	7	6.97627440188048	8.06883407534923e ⁻⁰⁵
6	0.382	0.565	0.4735	8	5.24524911457788	1.44676683127274
7	0.565	1000	500.2825	7	6.55630944510068	0.030026238107778



Modré - Četnost
Červené - Teoretická četnost

1.5 Podsekce e)

Předpoklad: $X \sim N(\mu, \sigma^2)$, σ^2 - neznámé

Bodový odhad střední hodnoty: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 0,0546$

Bodový odhad rozptylu: $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 0,2072743265$

Bodový odhad směrodatné odchylky: $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} = 0,4552739028$

Intervalový odhad parametru μ :

0,975 kvantil Studentova rozdělení $t_{1-\alpha/2}$ s $k = n - 1 = 50 - 1 = 49$ stupeň volnosti = 2.00957523712924

0,995 kvantil Studentova rozdělení $t_{1-\alpha/2}$ s $k = n - 1 = 50 - 1 = 49$ stupeň volnosti =

2.67995197363155

$$\alpha = 0.05 : \left\langle \bar{x} - t_{1-\alpha/2} \frac{s}{\sqrt{n}}; \bar{x} + t_{1-\alpha/2} \frac{s}{\sqrt{n}} \right\rangle = \langle -0.07478741155; 0.1839874116 \rangle$$

$$\alpha = 0.01 : \left\langle \bar{x} - t_{1-\alpha/2} \frac{s}{\sqrt{n}}; \bar{x} + t_{1-\alpha/2} \frac{s}{\sqrt{n}} \right\rangle = \langle -0.1179499213; 0.2271499213 \rangle$$

Intervalový odhad parametru σ^2 :

0,975 kvantil Pearsova rozdělení $\chi^2_{\alpha/2}$ s $k = n - 1 = 50 - 1 = 49$ stupeň volnosti = 31.55491646

0,975 kvantil Pearsova rozdělení $\chi^2_{1-\alpha/2}$ s $k = n - 1 = 50 - 1 = 49$ stupeň volnosti = 70.22241357

0,995 kvantil Pearsova rozdělení $\chi^2_{\alpha/2}$ s $k = n - 1 = 50 - 1 = 49$ stupeň volnosti = 27.24934907

0,995 kvantil Pearsova rozdělení $\chi^2_{1-\alpha/2}$ s $k = n - 1 = 50 - 1 = 49$ stupeň volnosti = 78.23070809

$$\alpha = 0.05 : \left\langle \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}; \frac{(n-1)s^2}{\chi^2_{\alpha/2}} \right\rangle = \langle 0.1446324825; 0.3218655962 \rangle$$

$$\alpha = 0.01 : \left\langle \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}; \frac{(n-1)s^2}{\chi^2_{\alpha/2}} \right\rangle = \langle 0.1298267937; 0.3727223712 \rangle$$

Intervalový odhad parametru σ :

$$\alpha = 0.05 : \left\langle \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}}; \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} \right\rangle = \langle 0.380305775; 0.5673319982 \rangle$$

$$\alpha = 0.01 : \left\langle \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}}; \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}}} \right\rangle = \langle 0.3603148536; 0.6105099272 \rangle$$

1.6 Podsekce f)

Testovací kritérium: $t = \frac{\bar{x} - \mu_0}{s} \sqrt{n} = 0.848017644303311$

Doplňek kritického oboru: $\overline{W}_\alpha = \langle -t_{1-\alpha/2}, t_{1-\alpha/2} \rangle$ pro alternativní hypotézu: $H_A : \mu \neq \mu_0$,

0,975 kvantil studentova rozdělení $t_{1-\alpha/2}$ s $k = n - 1 = 50 - 1 = 49$ stupeň volnosti = 2.00957523712924

$$\overline{W}_\alpha = \langle -t_{1-\alpha/2}, t_{1-\alpha/2} \rangle = \langle -2.00957523712924, 2.00957523712924 \rangle$$

Protože $t \in \overline{W}_\alpha$, tak se hypotéza $H_0 : \mu = 0$ **nezamítá**.

1.7 Podsekce g)

Tabulka 5:

Číslo	x1:20 - X	Číslo	x21:50 - Y
1	0.72	21	-0.59
2	-0.55	22	-0.13
3	0.30	23	0.33
4	0.12	24	0.38
5	0.12	25	-0.83
6	0.80	26	-0.13
7	0.63	27	0.55
8	-0.71	28	-0.24
9	-0.42	29	0.07
10	-0.06	30	0.44
11	0.46	31	-0.20
12	0.66	32	-0.29
13	0.63	33	1.00
14	0.28	34	0.21
15	-0.08	35	-0.52
16	0.08	36	-0.72
17	-0.02	37	0.59
18	0.44	38	-0.70
19	0.40	39	-0.42
20	0.41	40	-0.30
-	-	41	-0.23
-	-	42	-0.04
-	-	43	0.32
-	-	44	-0.16
-	-	45	-0.27
-	-	46	0.49
-	-	47	0.44
-	-	48	0.24
-	-	49	-0.39
-	-	50	-0.38

Tabulka 6:

	X	Y
n =	20	30
Průměr =	0.2105	-0.0493333333333333
Rozptyl s^2 =	0.17159475	0.197146222222222
Směrodatná odchylka =	0.414239966685978	0.444011511362287

Test rovnosti rozptylů – F-test

Testujeme hypotézu $H_0 : \sigma_X^2 = \sigma_Y^2$

Testovací kritérium: $t = \frac{s^2(X)}{s^2(Y)} = 0.870393295219115$

Doplňek kritického oboru: $\overline{W}_\alpha = F_{\frac{\alpha}{2}}(n-1, m-1), F_{1-\frac{\alpha}{2}}(n-1, m-1)$ pro $H_A : \sigma_X^2 \neq \sigma_Y^2$,

$F_{\frac{\alpha}{2}}(k_1, k_2), F_{1-\frac{\alpha}{2}}(k_1, k_2)$ jsou kvantily Fischerova-Snedecorova rozdělení s $k_1 = n-1$ a $k_2 = m-1$ stupni volnosti.

$$F_{\frac{\alpha}{2}}(19, 29) = 0.416329667$$

$$F_{1-\frac{\alpha}{2}}(19, 29) = 2.231274$$

$$\langle F_{\frac{\alpha}{2}}(n-1, m-1), F_{1-\frac{\alpha}{2}}(n-1, m-1) \rangle = \langle 0.416329667, 2.231274 \rangle$$

Protože $t \in \overline{W}_\alpha$ tedy hypotéza: $H_0 : \sigma_X^2 = \sigma_Y^2$ se **nezamítá**.

Studentův dvou-výběrový test:

Testujeme hypotézu $H_0 : \mu_X - \mu_Y = 0$ za podmínky $\sigma_X^2 = \sigma_Y^2$

Testovací kritérium: $t = \frac{\bar{x} - \bar{y} - \mu_0}{\sqrt{(n-1)s^2(X) + (m-1)s^2(Y)}} \frac{n*m(n+n-2)}{n+m} = 2.08126516493011$

Doplňek kritického oboru: $\overline{W}_\alpha = \langle -t_{1-\alpha/2}, t_{1-\alpha/2} \rangle$ pro $H_A : \mu_X - \mu_Y \neq 0$,

$t_{1-\alpha/2}$ - kvantil studentova rozdělení s $k = n + m - 2 = 20 + 30 - 2 = 48$ stupni volnosti,

$$t_{1-\alpha/2} = 2.010634758$$

$$\overline{W}_\alpha = \langle -t_{1-\alpha/2}, t_{1-\alpha/2} \rangle = \langle -2.010634758, 2.010634758 \rangle$$

Protože $t \notin \overline{W}_\alpha$ tedy hypotéza: $H_0 : \mu_X - \mu_Y = 0$ se **zamítá** a alternativní hypotéza $H_A : \mu_X - \mu_Y \neq 0$ se **nezamítá**.

2 Příklad číslo 2

Tabulka 7:

Výška	Váha
151	40.8576400048682
173	75.1338545520097
191	75.2529782352983
154	32.4327128568723
199	82.5215982536171
153	44.2118439774441
176	63.4449950025515
167	51.7859808918974
164	34.1283159765198
180	77.2391810363254
181	76.9020606464738
192	91.8069790869439
199	90.5576694567858
183	65.8915250607786
172	73.022553621128
151	33.5223067138509
177	65.671310222896
154	37.8481331642597
153	43.5502393034276
175	50.8950956694866

2.1 Podsekce a)

$$r = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sqrt{(\sum_{i=1}^n x_i^2 - n \bar{x}^2)(\sum_{i=1}^n y_i^2 - n \bar{y}^2)}} = 0.908209289297506$$

2.2 Podsekce b)

Testujeme hypotézu $H_0 : \rho = 0$:

$$t = \frac{|r| \sqrt{n-2}}{\sqrt{1-r^2}} = 9.20682361724213$$

Doplňek kritického oboru $\bar{W}_\alpha = \langle 0, t_{1-\alpha/2} \rangle$ pro alternativní hypotézu $H_A : \rho \neq 0$,

$$t_{1-\alpha/2}(n-2) = 2.100922037$$

Protože $t \notin \bar{W}_\alpha$, tedy hypotéza $H_0 : \rho = 0$ se **zamítá**.

2.3 Podsekce c)

Regresní analýza.

Tabulka 8:

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
151	41	22801	1669.346746	6169.503640
173	75	29929	5645.096099	12998.15683
191	75	36481	5663.010733	14373.31884
154	32	23716	1051.880863	4994.637779
199	83	39601	6809.814178	16421.79805
153	44	23409	1954.687147	6764.412128
176	63	30976	4025.267390	11166.31912
167	52	27889	2681.787816	8648.258808
164	34	26896	1164.741951	5597.043820
180	77	32400	5965.891087	13903.05258
181	77	32761	5913.926931	13919.27297
192	92	36864	8428.521409	17626.93998
199	91	39601	8200.691497	18020.97622
183	66	33489	4341.693074	12058.14908
172	73	29584	5332.293337	12559.87922
151	34	22801	1123.745047	5061.868313
177	66	31329	4312.720986	11623.82190
154	38	23716	1432.481184	5828.612507
153	44	23409	1896.623343	6663.186613
175	51	30625	2590.310763	8906.641742
$\sum = 3445$	$\sum = 1206.67697$	$\sum = 598277$	$\sum = 80204.53159$	$\sum = 213305.85019$

2.3.1 Podpodsekce 1)

$$\det(H) = n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 = 97515$$

$$b_2 = \frac{1}{\det(H)} \left(n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i \right) = 1.11895430884232$$

$$b_1 = \bar{y} - b_2 \bar{x} = -132.406031011417$$

$$y = b_1 + b_2 x = -132.406031011417 + 1.11895430884232x$$

$$S_{min}^* = \sum_{i=1}^n y_i^2 - b_1 \sum_{i=1}^n y_i - b_2 \sum_{i=1}^n x_i y_i = 1296.34021631919$$

$$s^2 = \frac{S_{min}^*}{n-2} = \frac{S_{min}^*}{20-2} = 72.0189009066218$$

2.3.2 Podpodsekce 2)

$$H : \beta_0 = -100, H_A : \beta_0 \neq -100$$

$$h^{11} = \frac{\sum_{i=1}^n x_i^2}{\det(H)} = 6.13523047736246$$

$$t = \frac{b_1 - (-100)}{s\sqrt{h_{11}}} = -1.54165486431127$$

$$t_{1-\alpha/2}(n-2) = t_{0.975}(20-2) = 2.100922037$$

$t \in \overline{W} = \langle -2.100922037; 2.100922037 \rangle$, a tedy $H : \beta_0 = -100$ se **nezamítá**.

$$H : \beta_1 = 1, H_A : \beta_1 \neq 1$$

$$h^{22} = \frac{n}{\det(H)} = 0.000205096651797$$

$$t = \frac{b_2 - 1}{s\sqrt{h_{22}}} = 0.978763235788905$$

$$t_{1-\alpha/2}(n-2) = t_{0.975}(20-2) = 2.100922037$$

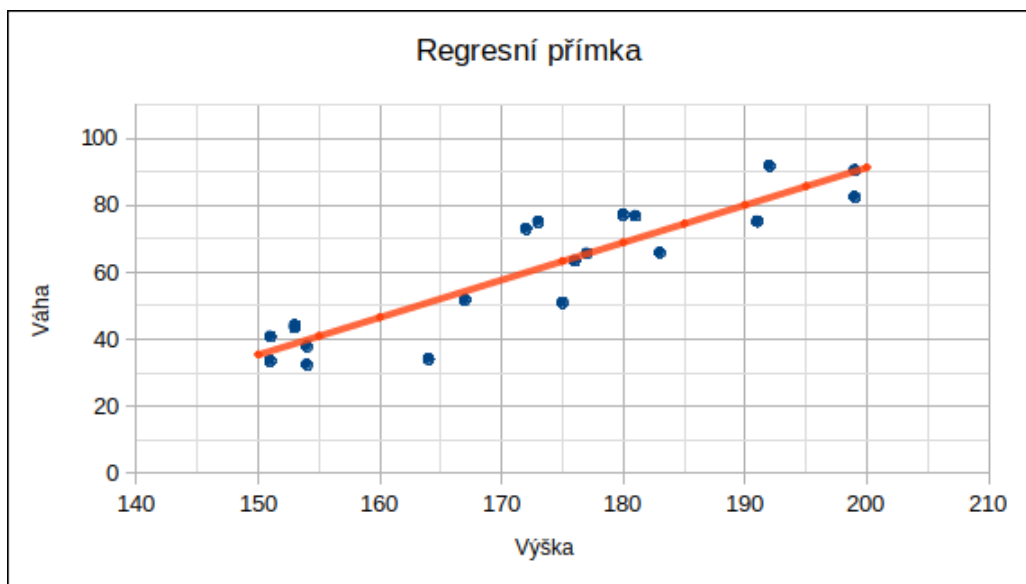
$t \in \overline{W} = \langle -2.100922037; 2.100922037 \rangle$, a tedy $H : \beta_1 = 1$ se **nezamítá**.

2.3.3 Podpodsekce 3)

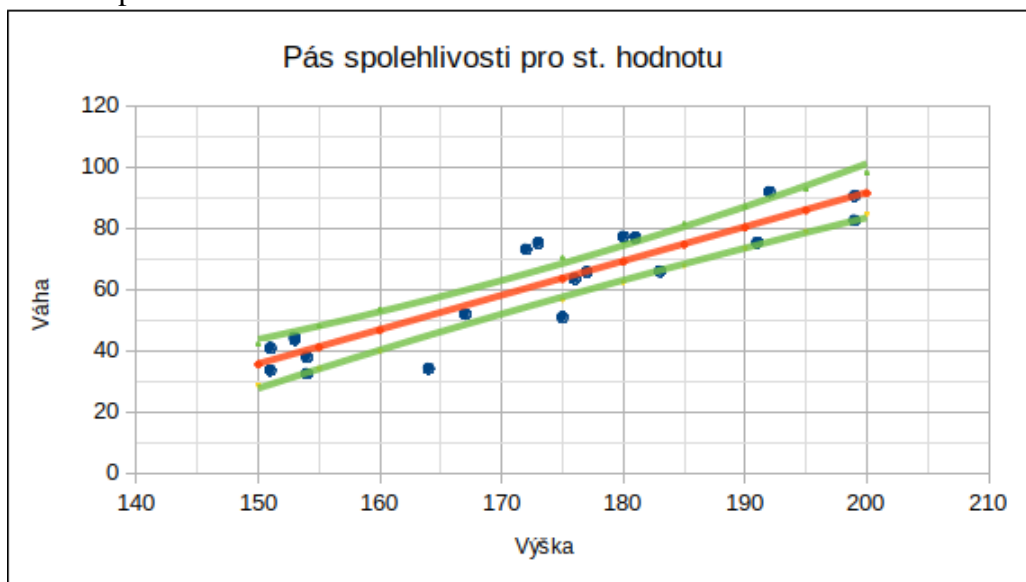
Tabulka 9:

x_i	y_i	h^*	y_i – dolní	y_i – horní	y_i – dolní	y_i – horní
			Střední	Střední	Individuální	Individuální
150	35.4371153	0.1515357	28.4966178	42.3776129	16.3046085	54.5696221
155	41.0318869	0.1110291	34.0913893	47.9723844	21.8993800	60.1643937
160	46.6266584	0.0807773	39.6861609	53.5671560	27.4941516	65.7591652
175	63.4109730	0.0515510	56.4704755	70.3514706	44.2784662	82.5434799
180	69.0057446	0.0623186	62.0652470	75.9462421	49.8732377	88.1382514
185	74.6005161	0.0833410	67.6600186	81.5410137	55.4680093	93.7330230
190	80.1952877	0.1146183	73.2547901	87.1357852	61.0627808	99.3277945
195	85.7900592	0.1561503	78.8495617	92.7305568	66.6575524	104.9225660
200	91.3848308	0.2079372	84.4443332	98.3253283	72.2523239	110.5173376

Regresní přímka



Pás spolehlivosti - střední hodnota



Pás spolehlivosti - individuální hodnota

