

Úloha 1

[10b]

Naměřené koncentrace vápníku ve vzorcích sušené citronové kůry (v gramech Ca na 100g kůry) jsou:

0.863	0.921	0.952	0.976	0.983	0.997	0.992	0.995	1.032	1.068
0.969	0.972	0.981	0.952	1.001	1.016	0.937	0.948	1.041	1.138
0.876	0.978	1.003	0.989	0.962	1.071	0.982	1.056	0.933	0.949

Zpracujte tento datový soubor:

- Určete rozsah, variační rozpětí, variační obor, medián, kvartily.
- Sestrojte krabicový graf.
- Určete průměr, rozptyl a směrodatnou odchylku.
- Soubor roztrďte vhodným počtem tříd: vytvořte tabulku četností a sestrojte histogram. Určete modus.
- Jaké procento dat padne do intervalu průměr ± 1 směrodatná odchylka? Jakým kvantilem odpovídají pozice průměr ± 1 směrodatná odchylka?

Úloha 2

[10b]

Určete korelaci mezi proměnnými x a y z následujícího souboru dat. Určete rovnici regresní přímky a nakreslete přímku do grafu. Vypočítejte koeficient determinace R^2 . Vykreslete graf residuí.

x	y
5.479	10.892
4.919	10.167
3.561	8.852
3.466	8.937
3.864	8.292
6.104	10.413
4.525	9.392
4.034	9.208
4.289	10.466
4.424	9.410
5.620	11.273
4.381	8.786
5.075	10.205
6.653	11.500
4.097	8.212
4.010	10.046
5.722	10.473
4.650	9.932
5.032	9.204
5.707	10.763

Dejme tomu, že transformujeme hodnoty x i y na z-skóre, a pak určíme rovnici regresní přímky těchto transformovaných bodů. Jaká bude směrnice a a úsek b regresní přímky?

Pozn.: Je možné použít libovolný software: R Studio, MS Excel, Python (numpy, scipy, pandas, etc.), Matlab, ... I při použití software je ale potřeba chápat, co čísla a výsledky znamenají, a eventuálně mít představu, jestli jsou správné – jestli dávají smysl.