3. Normální rozdělení

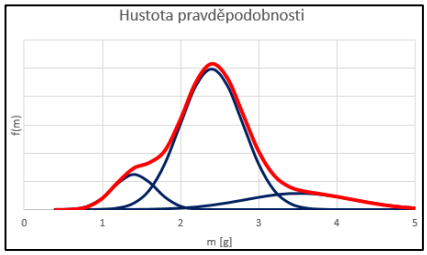
Motivace:

Nejčastějším typem rozdělení náhodných dat je Normální rozdělení. V této úloze budeme zpracovávat hmotnosti arašídů. Ale můžeme si představit, že jsou to hmotnosti rohlíků nebo elektrická vodivost polymerních granulí v balících dodávky. V poslední části, budete testovat, zda se „jumbo“ arašídy liší od těch obyčejných. Jednou v praxi budete řešit podobný problém, třeba jestli se liší jednotlivé šarže od jednoho dodavatele nebo produkty různých dodavatelů, případně zda dodávka odpovídá specifikaci dodavatele (jestli nám nedodává příliš lehké rohlíky nebo málo vodivé granule).

Poznámky k arašídům:

V úloze používáme dva druhy arašídů. Arašídy od IBK Trade prodávané v Bille a Jumbo Peanuts od ALESTO z Lidlu. Na balení není garantována velikost oříšků, ani se netvrdí, že „jumbo“ znamená větší. Přesto slovo „jumbo“ znamená v angličtině „obrovský“ a budeme tedy testovat, jestli jsou takto pojmenované arašídy větší.
Budeme vybírat arašídy obsahující právě dvě jádra, kterých je v balení přes 90 %. V grafu vpravo jsou modře hustoty pravděpodobnosti arašídů s jedním, dvěma a třemi jádry a tlustě červeně výsledná hustota pravděpodobnosti. Pokud bychom započítali i jádra s jedním nebo třemi a více jádry, nemají hmotnosti Normální rozdělení (viz červená čára v grafu vlevo). Zdá se, že normální rozdělení mají spíš hmotnosti jader. Pokud bychom vážili všechny arašídy, lišilo by se rozložení jejich hmotnosti výrazně od normálního a testy vycházející z předpokladu normality by nebyly použitelné.

Víte, že slovo „Jumbo“ se do angličtiny dostalo jako jméno velkého (4 m, 5900 kg) slona afrického vystupujícího 1882-1885 v Barnumově cirkusu v USA? Slon zemřel mlád, ve věku 24 let, po srážce s lokomotivou.



Požadované znalosti:

1. Střední hodnota, směrodatná odchylka souboru, směrodatná odchylka průměru.
2. Sestrojení histogramu četností a distribuční funkce.
3. Interpretace grafu hustoty pravděpodobnosti Normálního rozdělení.
4. Grubbsův test odlehlých hodnot.
5. Test rovnosti středních hodnot.

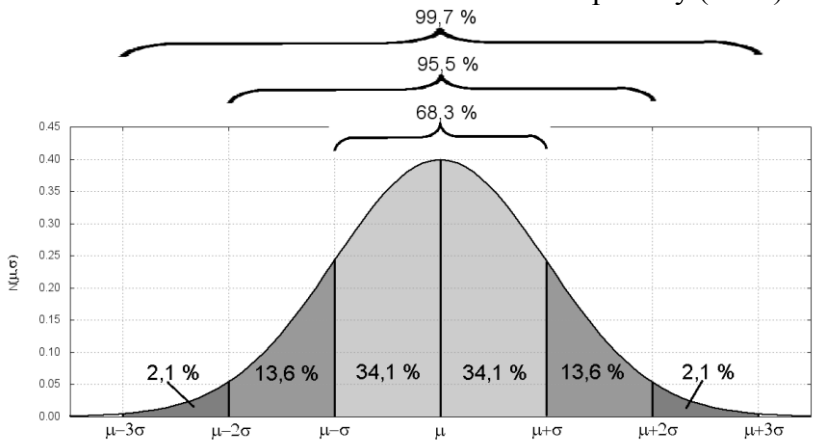
Úkoly měření:

1. Změřte hmotnosti 10 „jumbo“ arašídů (vyberte arašídy obsahující právě dvě jádra) a použijte hmotnosti 100 běžných arašídů z předchozí úlohy.
2. Odhadněte pro každou skupinku zvlášť její střední hodnotu a směrodatnou odchylku.
3. Otestujte větší skupinku na výskyt odlehlých hodnot a ty případně vyřaďte. Použijte Grubbsův test.
4. Odhadněte pro každou skupinku směrodatnou odchylku průměru a zvažte, zda má význam použít Studentovy koeficienty.
5. Vypočítejte intervaly spolehlivosti pro průměrnou hmotnost obou skupinek předmětů tak, aby průměr ležel uvnitř intervalu s pravděpodobností 0,683, 0,95 a 0,997.

6. Otestujte hypotézu, že obě skupiny mají stejnou střední hodnotu.
7. Vypočítejte, kolik hodnot z větší skupinky by mělo být v intervalu o šířce 1, 2 a 3 směrodatných odchylek a srovnejte s realitou.

Použité přístroje a pomůcky:

1. Větší množství arašídů, váhy.

Seznam doporučené literatury:

[1] http://fyzika.ft.utb.cz/ze/pred3.pdf
[2] Halliday D., Resnick R., Walker J.: Fyzika, VUT v Brně, Nakladatelství VUTIUM, (2000).
[3] Meloun M., Militký J,: Statistická analýza experimentálních dat, Academia, Nakladatelství Akademie věd České republiky (2004).