

Vlastní analýza datového souboru

Popis datového souboru

Pro dlouhodobý test tužkových NiMH akumulátorů od čtyř různých výrobců (A, B, C, D) bylo od každého výrobce pořízeno 100 akumulátorů. Všechny akumulátory měly deklarovanou kapacitu 2000 mAh. Kapacita každého akumulátoru byla změřena nejprve po 5 nabíjecích cyklech a následně po 100 nabíjecích cyklech.

V souboru <http://am-nas.vsb.cz/lit40/DATA/akumulatory.xlsx> jsou pro každý z akumulátorů uvedeny následující údaje: číslo akumulátoru, výrobce (A, B, C, D), změřená kapacita po 5 nabíjecích cyklech a změřená kapacita po 100 nabíjecích cyklech.

Úkoly

1. Tabulku se záznamy na listu Tabulka upravte do formátu vhodného pro import do statistického softwaru a uložte ji ve formátu csv.
2. Importujte data do software RkWard (resp. RStudio). Datový rámec nazvěte „data“.
3. Data transformujte do standardního datového formátu – „data.s“. (Nezapomeňte na odstranění chybějících hodnot (NA) vzniklých v důsledku různého počtu akumulátorů jednotlivých výrobců.)
4. Určete pro všechny akumulátory pokles kapacity po 100 nabíjecích cyklech oproti kapacitě po 5 nabíjecích cyklech a údaje uložte do nového sloupce s názvem „pokles“ v datovém rámci data.s.
5. Určete kolik akumulátorů jednotlivých výrobců je v našem datovém souboru a výsledek prezentujte sloupcovým grafem.
6. Vygenerujte vícenásobný krabicový graf, který prezentuje srovnání poklesů kapacit akumulátorů dle výrobců.
7. Určete vnitřní hradby pro identifikaci odlehlých pozorování (využijte funkci boxplot).
8. Na základě vícenásobného krabicového grafu identifikujte odlehlá pozorování v poklesech kapacit a do datového rámce data.s přidejte sloupec pokles.bez, do nějž uložte hodnoty poklesu kapacit bez odlehlých pozorování. (Odstranění odlehlých pozorování proveďte ručně s využitím interaktivního rozhraní RkWardu.)
9. Vygenerujte stejný vícenásobný krabicový graf po odstranění odlehlých pozorování.
10. Doplňte tabulku 1, nezapomeňte na správné zaokrouhlení číselných charakteristik. (Stačí, když vyplníte hodnoty pro výrobce A 😊)

Výběr. charakteristika	Výrobce A	Výrobce B	Výrobce C	Výrobce D
rozsah souboru				
Míry polohy				
minimum				
dolní kvartil				
medián				
průměr				
horní kvartil				
maximum				
Míry variability				
směrodatná odchylka				
variační koeficient (%)				
Míry šikmosti a špičatosti				
šikmost				
špičatost				

Tab. 1: Poklesy kapacit akumulátorů (mAh) po 100 nab. cyklech oproti kapacitám po 5 nab. cyklech

11. Na základě Tab. 1 a znalosti vnitřních hradeb a původních četnosti akumulátorů dle jednotlivých výrobců doplňte následující text.

Analýza poklesů kapacit akumulátorů výrobce A

Během testu byla měřena kapacita kusů akumulátorů výrobce A. Zjištěný pokles kapacit se pohyboval v rozmezí až mAh. Kapacity mimo interval (.....;.....) mAh byly identifikovány jako odlehlá pozorování a nebudou zahrnuty do dalšího zpracování. / Žádné z měření nebylo identifikováno jako odlehlé pozorování. Dále uvedené výsledky tedy pocházejí z analýzy kapacit kusů akumulátorů. Jejich průměrná kapacita byla mAh, směrodatná odchylka pak mAh. U poloviny testovaných akumulátorů kapacita nepřekročila mAh. V polovině měření se kapacita pohybovala v rozmezí až mAh. Vzhledem k hodnotě variačního koeficientu (.....%) lze / nelze analyzovaný soubor považovat za homogenní.

12. Vygenerujte čtveřici histogramů poklesů kapacit pro jednotlivé výrobce. (Nezapomeňte na stejné rozsahy os umožňující vzájemné srovnání histogramů.)
13. Vygenerujte čtveřici q-q grafů poklesů kapacit pro jednotlivé výrobce.
14. Posuďte normalitu poklesů kapacit akumulátorů jednotlivých výrobců na základě výběrových šikmostí a špičatostí a na základě příslušných histogramů a q-q grafů. Na základě posouzení normality odhadněte očekávané zastoupení dat v intervalu $\bar{x} \pm s$. (Doplňte následující text.)

Ověření normality poklesů kapacit akumulátorů výrobce A na základě explorační analýzy

Na základě grafického zobrazení (*tady by měl být odkaz na q-q grafy a histogramy*) a výběrové šikmosti a špičatosti (výběrová šikmost i špičatost leží / neleží v intervalu $(-2; 2)$) lze / nelze předpokládat, že poklesy kapacit akumulátorů výrobce A mají normální rozdělení. Dle pravidla 3σ / Čebyševovy nerovnosti lze tedy očekávat, že přibližně 95 % / více než 75 % akumulátorů bude vykazovat pokles kapacity v rozmezí až