

**VŠB – Technická univerzita Ostrava**  
**Fakulta elektrotechniky a informatiky**

**PRAVDĚPODOBNOST A STATISTIKA**

**Zadání 8**

JMÉNO STUDENTKY/STUDENTA:

OSOBNÍ ČÍSLO:

JMÉNO CVIČÍCÍ/CVIČÍCÍHO:

	DATUM ODEVZDÁNÍ	HODNOCENÍ
DOMÁCÍ ÚKOL 1:		
DOMÁCÍ ÚKOL 2:		
DOMÁCÍ ÚKOL 3:		
DOMÁCÍ ÚKOL 4:		
CELKEM:	-----	

**Ostrava, AR 2016/2017**

**Popis datového souboru**

Běžné zářivky trpí efektem pomalého nabíhání, tedy plného výkonu dosáhnou až po jisté době provozu. Toto chování je ovlivněno okolní teplotou, což v praxi znamená, že v chladném prostředí může zářivkám trvat výrazně déle než dosáhnou maximálního výkonu.

Pro test náběhu zářivek na plný světelný výkon bylo vybráno celkem 350 zářivek od čtyř různých výrobců (Amber, Bright, Clear, Dim). Všechny zářivky měly deklarovaný maximální světelný tok 1000 lm. U každé zářivky byla změřena svítivost po 30 sekundách od zapnutí, nejprve při teplotě 5 °C a poté při teplotě 22 °C.

V souboru [ukol\\_8.xlsx](#) jsou pro každou z testovaných zářivek uvedeny následující údaje:

- pořadové číslo zářivky,
- výrobce – Amber (A), Bright (B), Clear (C), Dim (D),
- naměřená svítivost v lumenech při okolní teplotě 5 °C,
- naměřená svítivost v lumenech při okolní teplotě 22 °C.

**Obecné pokyny:**

- Úkoly zpracujte dle obecně známých typografických pravidel.
- Všechny tabulky i obrázky musí být opatřeny titulkem.
- Do úkolů nekládejte tabulky a obrázky, na něž se v doprovodném textu nebudete odkazovat.
- Bude-li to potřeba, citujte zdroje dle mezinárodně platné citační normy ČSN ISO 690.

**Úkol 1**

- Pomocí nástrojů explorační analýzy analyzujte svítivost zářivek výrobce Amber po 30 sekundách od zapnutí při teplotách 5 °C a 22 °C. Data vhodně graficky prezentujte (krabicový graf, histogram, q-q graf) a doplňte následující tabulky a text.

Svítivost zářivek, výrobce Amber			Po odstranění odlehlých pozorování	
	při teplotě 5 °C	při teplotě 22 °C	při teplotě 5 °C	při teplotě 22 °C
rozsah souboru				
<b>Míry polohy</b>				
minimum				
dolní kvartil				
medián				
průměr				
horní kvartil				
maximum				
<b>Míry variability</b>				
směrodatná odchylka				
variační koeficient (%)				
<b>Míry šikmosti a špičatosti</b>				
šikmost				
špičatost				
<b>Identifikace odlehlých pozorování - vnitřní hradby</b>				
dolní mez				
horní mez				

Jméno:

Číslo domácího úkolu: 8

**Grafická prezentace (krabicový graf, histogram, q-q graf):**

**Analýza svítivosti zářivek výrobce Amber po 30 sekundách od zapnutí při teplotě 5 °C**

Během testu byla měřena svítivost ..... kusů zářivek výrobce Amber. Naměřená svítivost při teplotě 5 °C se pohybovala v rozmezí ..... až ..... lm. Svítivosti žárovek č. .... byly na základě metody vnitřních hradeb identifikovány jako odlehlá pozorování a nebudou zahrnuty do dalšího zpracování. Možné příčiny vzniku odlehlých pozorování jsou: ..... / Žádné z měření nebylo identifikováno jako odlehlé pozorování. Dále uvedené výsledky tedy pocházejí z analýzy svítivosti ..... kusů zářivek. Jejich průměrná svítivost byla ..... lm, směrodatná odchylka pak ..... lm. U poloviny testovaných zářivek svítivost nepřekročila ..... lm. V polovině měření se svítivost pohybovala v rozmezí ..... až ..... lm. Vzhledem k hodnotě variačního koeficientu (.....%) lze / nelze analyzovaný soubor považovat za homogenní.

**Analýza svítivosti zářivek výrobce Amber po 30 sekundách od zapnutí při teplotě 22 °C**

Během testu byla měřena svítivost ..... kusů zářivek výrobce Amber. Naměřená svítivost při teplotě 22 °C se pohybovala v rozmezí ..... až ..... lm. Svítivosti žárovek č. .... byly na základě metody vnitřních hradeb identifikovány jako odlehlá pozorování a nebudou zahrnuty do dalšího zpracování. Možné příčiny vzniku odlehlých pozorování jsou: ..... / Žádné z měření nebylo identifikováno jako odlehlé pozorování. Dále uvedené výsledky tedy pocházejí z analýzy svítivosti ..... kusů zářivek. Jejich průměrná svítivost byla ..... lm, směrodatná odchylka pak ..... lm. U poloviny testovaných zářivek svítivost nepřekročila ..... lm. V polovině měření se svítivost pohybovala v rozmezí ..... až ..... lm. Vzhledem k hodnotě variačního koeficientu (.....%) lze / nelze analyzovaný soubor považovat za homogenní.

**Ověření normality svítivosti zářivek výrobce Amber po 30 sekundách od zapnutí při teplotě 5 °C na základě explorační analýzy**

Na základě grafického zobrazení (viz ..... ) a výběrové šikmosti a špičatosti (výběrová šikmost i špičatost leží / neleží v intervalu  $(-2; 2)$  ) lze / nelze předpokládat, že svítivost zářivek výrobce Amber při teplotě 5 °C má normální rozdělení. Dle pravidla  $3\sigma$  / Čebyševovy nerovnosti lze tedy očekávat, že přibližně 95 % / více než 75 % zářivek bude mít svítivost v rozmezí ..... až ..... lm.

**Ověření normality svítivosti zářivek výrobce Amber po 30 sekundách od zapnutí při teplotě 22 °C na základě explorační analýzy**

Na základě grafického zobrazení (viz ..... ) a výběrové šikmosti a špičatosti (výběrová šikmost i špičatost leží / neleží v intervalu  $(-2; 2)$  ) lze / nelze předpokládat, že svítivost zářivek výrobce Amber při teplotě 22 °C má normální rozdělení. Dle pravidla  $3\sigma$  / Čebyševovy nerovnosti lze tedy očekávat, že přibližně 95 % / více než 75 % zářivek bude mít svítivost v rozmezí ..... až ..... lm.

**Úkol 2**

Porovnejte pokles svítivosti po 30 sekundách od zapnutí při snížení okolní teploty z 22 °C na 5 °C u zářivek od výrobců Amber a Bright. Nezapomeňte, že použité metody mohou vyžadovat splnění určitých předpokladů. Pokud tomu tak bude, okomentujte splnění/nesplnění těchto předpokladů jak na základě explorační analýzy (např. s odkazem na histogram apod.), tak exaktně pomocí metod statistické indukce.

a) Graficky prezentujte srovnání poklesů svítivosti zářivek výrobců Amber a Bright při snížení okolní teploty (vícenásobný krabicový graf, histogramy, q-q grafy).

b) Určete bodové a 95% intervalové odhady pro střední hodnoty (resp. mediány) poklesů svítivosti zářivek výrobců Amber i Bright při snížení okolní teploty.

c) Na hladině významnosti 5 % rozhodněte, zda jsou střední poklesy (resp. mediány poklesů) svítivosti zářivek výrobců Amber a Bright statisticky významné.

d) Určete bodový a 95% intervalový odhad rozdílu středních hodnot (resp. mediánů) poklesů svítivosti zářivek výrobců Amber a Bright (při snížení okolní teploty).

e) Na hladině významnosti 5% rozhodněte, zda je rozdíl středních hodnot (mediánů) poklesů svítivosti zářivek výrobců Amber a Bright (při snížení okolní teploty) statisticky významný.

**Úkol 3**

Na hladině významnosti 5 % rozhodněte, zda se svítivost zářivek při teplotě 5 °C liší v závislosti na tom, od kterého výrobce pocházejí. Posouzení proveďte nejprve na základě explorační analýzy a následně pomocí vhodného statistického testu, včetně ověření potřebných předpokladů. V případě, že se svítivost zářivek jednotlivých výrobců statisticky významně liší, určete pořadí výrobců dle svítivosti zářivek při 5°C.

a) Daný problém vhodným způsobem graficky prezentujte (vícenásobný krabicový graf, histogramy, q-q grafy).

b) Ověřte normalitu svítivosti zářivek při teplotě 5 °C u všech čtyř výrobců (empiricky i exaktně).

c) Ověřte homoskedasticitu (shodu rozptylů) svítivosti zářivek při teplotě 5 °C jednotlivých výrobců (empiricky i exaktně).



- d) Určete bodové a 95% intervalové odhady střední hodnoty (resp. mediánu) svítivosti zářivek při teplotě 5 °C pro všechny srovnávané výrobce. (Nezapomeňte na ověření předpokladů pro použití příslušných intervalových odhadů.)
- e) Čistým testem významnosti ověřte, zda je pozorovaný rozdíl středních hodnot (resp. mediánů) svítivosti zářivek při teplotě 5 °C u všech výrobců statisticky významný na hladině významnosti 5%. Pokud ano, zjistěte, zda lze některé skupiny výrobců označit (z hlediska svítivosti zářivek po 30 sekundách od zapnutí při teplotě 5 °C) za homogenní, tj. určete pořadí výrobců dle svítivosti zářivek při 5°C. (Nezapomeňte na ověření předpokladů pro použití zvoleného testu.)

**Úkol 4**

Všichni čtyři výrobci udávají, že jejich zářivky dosáhnou po 30 sekundách od zapnutí alespoň osmdesáti procent deklarovaneho maximálního světelného toku (tj. 80% z 1000 lm). (Definujte si novou dichotomickou proměnnou Splnění požadavku na deklarovany světelný tok po 30 s, která bude nabývat hodnot {ANO, NE}.)

a) Srovnajte skladbu zářivek dle toho, zda splňují deklarovany světelný tok po 30 s od zapnutí při teplotě 5 °C pro jednotlivé výrobce (Amber, Bright, Clear, Dim). Výsledky prezentujte pomocí kontingenční tabulky, vhodného grafu a vhodné míry kontingence.

b) V případě výrobce Bright určete bodový i 95% intervalový odhad pravděpodobnosti, že při teplotě 5 °C zářivka nedosáhne po 30 sekundách požadované svítivosti (80% deklarovaneho maximálního světelného toku).

c) Určete bodový i 95% intervalový odhad relativního rizika, že při teplotě 5 °C zářivka nedosáhne po 30 sekundách požadované svítivosti (80% deklarovaneho maximálního světelného toku), pro „nejhoršího“ výrobce (vzhledem k „nejlepšímu“ výrobcu). Výsledky slovně interpretujte.

- d) Určete bodový i 95% intervalový odhad poměru šancí, že při teplotě 5 °C zářivka nedosáhne po 30 sekundách požadované svítivosti (80% deklarovaného maximálního světelného toku), pro „nejhoršího“ výrobce (vzhledem k „nejlepšímu“ výrobcí). Výsledky slovně interpretujte.
- e) Pomocí chí-kvadrát testu nezávislosti rozhodněte, jestli pravděpodobnost, že při teplotě 5 °C zářivka nedosáhne po 30 sekundách požadované svítivosti (80% deklarovaného maximálního světelného toku), závisí statisticky významně na tom, od kterého výrobce zářivka pochází.

## Jak identifikovat, zda jsou v datech odlehlá pozorování?

### Empirické posouzení:

- použití vnitřních (vnějších) hradeb, resp.  $z$  – souřadnice, resp. mediá nová suř adnice,
- vizuální posouzení krabicového grafu.

### Exaktní posouzení:

- Grubbsův test (parametrický test - vyžaduje normalitu dat)
- Deanův - Dixonův test (neparametrický test)

Jak naložit s odlehlými hodnotami by měl definovat hlavně zadavatel analýzy (expert na danou problematiku).

## Jak ověřit normalitu dat?

### Empirické posouzení:

- vizuální posouzení histogramu,
- vizuální posouzení grafu odhadu hustoty pravděpodobnosti,
- Q-Q graf,
- P-P graf,
- posouzení výběrové šikmosti a výběrové špičatosti.

### Exaktní posouzení:

- testy normality (např. Shapirův - Wilkův test, Andersonův-Darlingův test, Lillieforsův test, ...)

## Jak ověřit homoskedasticitu (shodu rozptylů)?

### Empirické posouzení:

- poměr největší a nejmenší směrodatné odchylky,
- vizuální posouzení krabicového grafu.

### Exaktní posouzení:

- $F$  - test (parametrický dvouvýběrový test),
- Bartlettův test (parametrický vícevýběrový test),
- Leveneův test (neparametrický test).