

Pokyny

Utvořte skupiny po třech. Úlohu odevzdejte ve formě písemné zprávy, buďto vytištěné na posledním cvičení před Vánoci, nebo dle dohody s vaším cvičícím. Zpráva nechť obsahuje řešení zadaných úkolů spolu s vysvětlením vašeho postupu, grafy, diskusí výsledků a klíčovými částmi kódu.

Poznámka: Pokud chcete, můžete se rozhodnout některé z úkolů nevypracovat a získat za ně nula bodů. Tuto skutečnost indikujte.

Data a parametry

Ze své trojice vyberte reprezentanta. Úlohu odevzdejte cvičícímu, ke kterému reprezentant přísluší. Parametry úlohy spočtete následovně:

- K = den narození reprezentanta skupiny (1-31),
- L = počet písmen v příjmení reprezentanta,
- $M = (K * L * 47) \bmod(12)$.

Načtete datový soubor z R knihovny Sleuth2 podle následující tabulky:

M	Datový soubor	Popis
0	case0101	skóre v kreativitě dle typu motivace
1	case0102	mzda dle pohlaví
2	case0201	délka humeru dle přežití vrabců
3	case0202	objem hipokampu dle schizofrenie
4	case0301	srážky dle ošetření mraků
5	case0302	koncentrace dioxinu dle vojenského působiště
6	case0401	incidenty raketoplánů dle teploty
7	case0402	doba řešení úkolu dle typu výuky geometrie
8	ex0112	krevní tlak dle diety
9	ex0211	přežití morčat dle typu experimentu
10	ex0221	váha dle přežití vrabců
11	ex0222	hladina cholesterolu dle prostředí.

Každý z datových souborů reprezentuje pozorovaná data z reálného experimentu, kdy byla určitá spojitá veličina měřena zvlášť pro dvě skupiny.

Knihovnu lze stáhnout zde:

<https://cran.r-project.org/package=Sleuth2>

Dokumentace k datovým souborům je k dispozici zde:

<https://cran.r-project.org/web/packages/Sleuth2/Sleuth2.pdf>

Úkoly

- 1) (1b) Načtěte datový soubor a rozdělte sledovanou proměnnou na příslušné dvě pozorované skupiny. Data stručně popište a pro každou skupinu zvlášť odhadněte střední hodnotu, rozptyl a medián příslušného rozdělení.
- 2) (1b) Pro každou skupinu zvlášť odhadněte hustotu a distribuční funkci pomocí histogramu a empirické distribuční funkce.
- 3) (3b) Pro každou skupinu zvlášť najděte nejbližší rozdělení. Odhadněte parametry normálního, exponenciálního a rovnoměrného rozdělení a zanechte příslušné hustoty s odhadnutými parametry do grafů histogramu. Diskutujte, které z rozdělení odpovídá pozorovaným datům nejlépe.
- 4) (1b) Pro každou skupinu zvlášť vygenerujte náhodný výběr o 100 hodnotách z rozdělení zvoleného jako nejbližší v předchozím bodě, s parametry odhadnutými shora. Porovnejte histogram simulovaných hodnot s pozorovanými daty.
- 5) (1b) Pro každou skupinu zvlášť spočítejte oboustranný 95% konfidenční interval pro střední hodnotu.
- 6) (1b) Pro každou skupinu zvlášť otestujte na hladině významnosti 5% hypotézu, zda je střední hodnota rovná hodnotě K , proti oboustranné alternativě. Můžete použít jak výsledek z předešlého bodu, tak výstup z příslušné vestavěné funkce vašeho softwaru a výsledky porovnat.
- 7) (2b) Na hladině spolehlivosti 5% otestujte, jestli mají pozorované skupiny stejnou střední hodnotu. Typ testu a alternativy stanovte tak, aby vaše volba nejlépe korespondovala s povahou zkoumaného problému.

Software

Úlohu můžete řešit buďto ve freeware statistickém softwaru R, nebo v libovolném jiném softwaru, který má potřebné funkcionality (Mathematica, Matlab, Python, Excel, ...). Výpočty můžete provádět buď ručně z příslušných vzorců, nebo pomocí statistických balíčků a funkcí. Vždy je potřeba postup vysvětlit a výsledek řádně interpretovat.

R reference

Instalace	http://www.r-project.org/
R reference card	http://cran.r-project.org/doc/contrib/Short-refcard.pdf
R tutorial	http://www.cyclismo.org/tutorial/R/index.html
R guide	http://cran.r-project.org/doc/contrib/Owen-TheRGuide.pdf
Simple R	http://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf
Introduction to R	http://cran.r-project.org/doc/contrib/Lam-IntroductionToR_LHL.pdf

Důležité příkazy v R

Instalace knihovny	<code>install.packages("Sleuth2")</code>
Načtení knihovny	<code>library(Sleuth2)</code>
Export dat	<code>write.table(ex0331,"D:/my_data.csv",row.names=F,sep=";",dec=",")</code>
Řádky, sloupce a prvky z tabulky	<code>ex0331[k,]; ex0331[,l]; ex0331[k,l];</code>
Další důležité příkazy	<code>subset();mean();ecdf();hist();t.test();</code>
Dokumentace k příslušnému příkazu	<code>?subset</code>