

Report: Štatistická analýza

Pohybová aktivita sumca

Autor: MM

Instituce: MFF UK

Datum: 01.11.2019

1 Úvod

V úvode by mala byť nejaká stručná motivácia, mali by sme predstaviť a odôvodniť vedeckú otázku, kôli ktorej vznikol experiment, data, aj samotná štatistická analýza.

2 Metody

V tejto sekcii by mal byť dostatočne podrobne popísaný experiment aj následná štatistická analýza. Popis by mal byť dostatočne podrobný na to, aby bolo možné experiment aj analýzu reprodukovat'.

V tomto príklade si namiesto toho ukážeme, ako používať knižnicu Sweave na prípravu PDF reportu a ako implementovať jednotlivé príkazy. V prvom kroku je potrebné načítať podkladové data, ale pre výsledného čitateľa je konkrétny spôsob načítania dat zbytočná/nepodstatná informácia.

Príkaz pre načítanie dat sa preto vo výslednom PDF reporte neobjaví.

Analogickými príkazmi je možné data upraviť do tvaru, ktorý je potrebný pre samotnú štatistickú analýzu. Sekcia Metody by mala ideálne končiť informáciou o tom, v akom programe bola analýza spracovaná. V tomto názornom príklade boli data spracované pomocou programu R (Team Development Core 2020).

3 Výsledky

V tejto časti by sa mali nachádzať jednak výsledky popisných charakteristík (prvá časť výsledkov) a jednak popis a interpretácia samotného pravdepodobnostného/statistického modelu (druhá časť výsledkov).

V tejto ukážke si predstavíme pouze prvú časť a uvedieme si niekoľko popisných charakteristík vo forme textu, tabuliek a obrázkov (všetko toto patrí do sekcii exploratívnej analýzy). Zameriame sa pritom na konkrétne spôsoby použitia knižnice *Sweave*, ne na formálnu korektnosť prezentovaných výsledkov.

Napríklad, popisné charakteristiky premenných zaznamenajúcich pohybovú aktivitu môžeme získať nasledujúcim spôsobom, hoci samotná estetická/vizuálna stránka tabuľky je pomerne slabá.

	A_biting	A_chasing	A_latdsipl	A_frontdispl
Min.	:0.0000	Min. :0.0000	Min. :0.00000	Min. :0.00000
1st Qu.	:0.0000	1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.00000
Median	:0.0000	Median :0.0000	Median :0.00000	Median :0.00000
Mean	:0.3553	Mean :0.1842	Mean :0.09211	Mean :0.06579
3rd Qu.	:0.0000	3rd Qu.:0.0000	3rd Qu.:0.00000	3rd Qu.:0.00000
Max.	:8.0000	Max. :6.0000	Max. :2.00000	Max. :2.00000

Alebo dokonca v krajšej L^AT_EX tabuľke pomocou R knižnice `xtable`:

	A_biting	A_chasing	A_latdsipl	A_frontdispl
X	Min. :0.0000	Min. :0.0000	Min. :0.00000	Min. :0.00000
X.1	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.00000
X.2	Median :0.0000	Median :0.0000	Median :0.00000	Median :0.00000
X.3	Mean :0.3553	Mean :0.1842	Mean :0.09211	Mean :0.06579
X.4	3rd Qu.:0.0000	3rd Qu.:0.0000	3rd Qu.:0.00000	3rd Qu.:0.00000
X.5	Max. :8.0000	Max. :6.0000	Max. :2.00000	Max. :2.00000

Správne by ale mala tabuľka obsahovať aj popisok a pomocou popisku by malo byť možné sa na príslušnú tabuľku v texte kedykoľvek odkázať (vid' nižšie).

	A_biting	A_chasing	A_latdsipl	A_frontdispl
X	Min. :0.0000	Min. :0.0000	Min. :0.00000	Min. :0.00000
X.1	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:0.00000	1st Qu.:0.00000
X.2	Median :0.0000	Median :0.0000	Median :0.00000	Median :0.00000
X.3	Mean :0.3553	Mean :0.1842	Mean :0.09211	Mean :0.06579
X.4	3rd Qu.:0.0000	3rd Qu.:0.0000	3rd Qu.:0.00000	3rd Qu.:0.00000
X.5	Max. :8.0000	Max. :6.0000	Max. :2.00000	Max. :2.00000

Table 1: Popisné charakteristiky pre štyri agresívne aktivity zaznamenané v priebehu dňa: `A_biting` - pohryzanie, `A_chasing` - naháňanie, `A_lat dsipl` - útok zo strany a `A_frontdisp` - útok z predu.

	00:00 - 06:00	06:00 - 12:00	12:00 - 18:00	18:00 - 24:00
Pohryzenie	0.50 (0.84)	0.47 (1.57)	0.00 (0.00)	0.35 (0.75)
Naháňanie	0.00 (0.00)	0.34 (1.18)	0.00 (0.00)	0.12 (0.33)
Bočný útok	0.00 (0.00)	0.09 (0.30)	0.00 (0.00)	0.15 (0.46)
Čelný útok	0.33 (0.52)	0.06 (0.18)	0.05 (0.00)	0.07 (0.39)

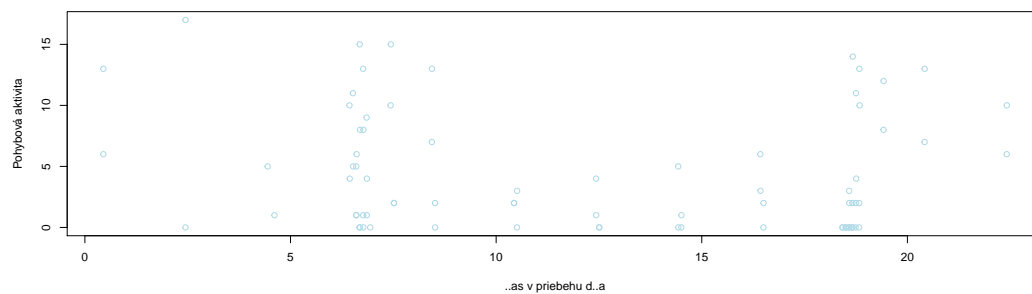
Table 2: Tabuľka pomocou príkazu `cat()`.

Iná a často najlepšia možnosť (keďže poskytuje najvyššiu mieru voľnosti a užívateľského prispôsobenia) je vytvoriť tabuľku v programe R príkazom `cat()` a výsledok v zdrojovom `tex` súbore načítať pomocou príkazu

```
\input{tabulka.tex}
```

Príslušný zdrojový kód, pomocou ktorého bola tabuľka v programe R vytvorená, je k dispozícii na webovej stránke tohto semináru (viď seminár č.4).

Pomocou popisku a čísla tabuľky sa môžeme na tabuľku pri interpretácii výsledkov náležite odkázať. Napríklad: Čelný útok sa najčastejšie vyskytuje v ranných hodinách - viď Tabuľka 2.



Obrázok...

Samozrejme aj obrázok by mal obsahovať dostatočne podrobný popisok a poradové číslo, aby bolo možné analogicky sa odkazovať aj na obrázky.

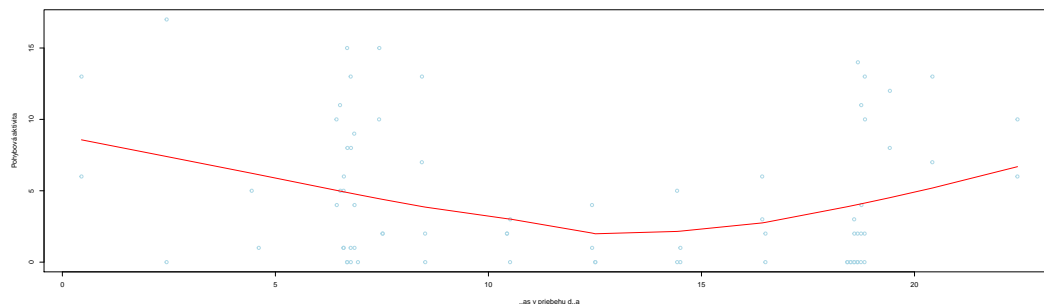


Figure 1: Celková pohybová aktivita meraná v závislosti na čase zaznamenania (od času 0:00 až do 24:00).

Z Obrázku 1 je zjavné, že pohybová aktivita sa v priebehu dňa mení. Priemerná pohybová aktivita je na úrovni 4.51315789473684, čo je ale asi lepšie prezentovať ako 4.51 (směrodatná chyba 4.86). Tento spôsob zápisu umožňuje automaticky update pri zmene datového súboru (má to samozrejme svoje výhody aj nevýhody).

4 Záver

Report zo štatistickej analýzy by mal obsahovať nejaké celkové zhrnutie a záver. V prípade použitia akejkoľvek odbornej literatúry, alebo iného zdroja, je nutné doplniť tieto zdroje do záverečného zoznamu literatúry.