

Experimental Design and Statistics - AGA46E

M. Maciak (Czech University of Life Sciences, Prague)

Lab Session 4 - Summer Term 2015

1 Prehled nekterych prikazu

- generatory nahodnych cisel s diskretnym rozdelenim: `rbinom()`, `rpois()`, `geom()`, a dalsi...
- funkce pro vykresleni histogramu a empiricke distribucni funkce: `hist()`, `ecdf()`;
- nektere funkce dulezite pro praci s faktorovymi promennymi: `factor()`, `levels()`; a nektere dodatecne parametre: `labels()`;
- pro overeni typu promenne je prikaz `class()`;

2 Spojite nahodne veliciny a generatory v Rku

- pomoci google vyhladajte potrebne prikazy pro generovani hodnot z rovnomernyho, exponencialneho a normalneho rozdeleni;
- jake jsou dodatecne parametre pozadovane pro korektni specifikaci? (vyuzite `help` v Rku)
- vyzkousejte ruzne nastaveni (hodnoty) potrebnych parametru a vygenerujte nekolik sekvencii nahodnych realizaci; pomoci grafickyh nástroju je nasledne porovnejte; (napr. funkce `hist`, `ecdf()`, `stepfun()`, a dalsi.)
- jaky zakladny efekt je mozne sledovat v obrazcich, jestlize velikost vyberu se porad zvetsuje?

3 Spojite promenne v Rku

- jak rozlysit spojite a diskretne nahodne veliciny pri praci s realnymi datami?
- spojite promenne by meli byt dostatecne identifikovatelne z povahy samotneho experimentu; samotne pozorovani je totiz porad mozne chapat pouze jako diskretni hodnoty, kedze nahodny vyber je pomerne maly (hlavne v porovnani s nekonecne velkou populaci); (ruzne graficke nástroje se pouzivaji pro vykresleni spojitych, diskretnych a faktorovych promennych)
- pozite data `ChickWeight`, ktere jsou k dispozici v Rku a vykreslite spojite promenne;
- ktere promenne v `ChickWeight` lze chapat jako zavislou promennou a ktere promenne spise jako nezávisle (vysvetlujici) promenne?
- jaky je zakladny rozdíl v interpretaci pozorovani jako faktoru a jako spojitej promennej? (podivejte se napr. na rozdíl v nasledujim)

```
> plot(weight ~ as.numeric(Chick)) ### Chickens are identified with numbers 1 to 50
> plot(weight ~ as.factor(Chick))  ### Chickens are "labeled" with numbers 1 to 50
```

- neboli stejne rozdíl mezi

```
> plot(as.factor(Chick))
> plot(as.numeric(Chick))
```

- vyzkousejte nasledující část kódu v Rku a udelejte graf $weight \sim Time$ o něco hezčeji

```
> attach(ChickWeight)
>
> plot(weight ~ Time, pch="*", col = "gray", xlab = "Time in months", ylab = "Chick's weight")
> for (i in levels(Chick)){
>   lines(weight[Chick == i] ~ Time[Chick == i], col = i)
> }
```

- jak lze vysvětlit jednotlivé řádky v části kódu, který jsme využili?
- vyzkousejte analogický kód v Rku pro data DNase nebo Orange a pokuste se udelet stejný graf (oba datové soubory jsou k dispozici v Rku - příkaz `data()`);

4 Real Data with Continuous Variables

- z webové stránky stáhnete data `passengerData.csv` a načtete je to Rka:

```
> data <- read.table("passengerData.csv", header = T, sep = " ", dec = ".")
> attach(data)
```

- jednoduchý souhrn ohledem datového souboru se v Rku získá pomocí příkazu `summary()`; (zakazdym je vhodné používat tenhle příkaz k prvnímu nahlédnutí na data)
- využijte grafické nástroje v Rku a nakreslete několik obrázků k datovému souboru `passengerData.csv`; (napr. příkazy `plot()`, `boxplot()`, `barplot()`, a pod.)