

Statystyka
w
Geografii / Gospodarce Przestrzennej
r.a 2014/2015, semestr letni

Michał Marosz

Katedra Meteorologii i Klimatologii IG UG

21 lutego 2015



Spis treści

| | |
|--|----|
| 1 R is the answer... | 5 |
| 2 Podstawowe miary statystyczne - wprowadzenie | 7 |
| 3 Podstawowe miary statystyczne - szereg rozdzielczy | 11 |
| 4 Prawdopodobieństwo - wprowadzenie | 13 |
| 5 Rozkłady teoretyczne - podstawy | 15 |
| 6 Rozkład Gauss'a | 17 |
| 7 Miary korelacji/asocjacji | 19 |
| 8 Regresja | 21 |
| 9 Trend | 23 |
| 10 Kreślenie map w R - interpolacja | 25 |

Rozdział 1

R is the answer...

Poniżej zamieszczono przykładowe komendy Ra, które mogą zostać wykorzystane w analizach statystycznych i prezentacji wyników.

Rozdział 2

Podstawowe miary statystyczne - wprowadzenie

Pojęcia do praktycznego opanowania

- szereg szczegółowy,
- szereg rozdzielczy,
- wykres liniowy,
- histogram, dystrybuanta,
- rozkład zmiennej,
- rozkład teoretyczny, rozkład empiryczny

Miary tendencji centralnej

- średnia arytmetyczna, średnia arytmetyczna ważona,
- dominanta (moda),
- mediana, kwantyle (kwartyle, decyle),

Miary dyspersji

- wariancja,
- odchylenie standardowe,
- współczynnik zmienności,
- odchylenie przeciętne,
- rozstęp,
- rozstęp międzykwartyłowy,

Wybrane wzory

Średnia arytmetyczna

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Współczynnik zmienności

$$V = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

Wariancja (dla populacji generalnej)

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n}$$

Wariancja (dla próby)

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Odchylenie standardowe (dla populacji)

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}}$$

Odchylenie standardowe (dla próby)

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

Zadania

Zadanie 1

Wykorzystując zamieszczone poniżej dane (Tab.2) oraz informacje przekazane przez prowadzącego uzupełnij Tabelę 1.

Zadanie 2

Wykorzystując R oblicz powyższe miary dla poniższego szeregu szczegółowego. Wykonaj wykres liniowy.

12, 24, 9, 16, 12, 8, 6, 17, 11, 16, 21

Tablica 2.1:

| Miara statystyczna | Wartość | Wynik z R |
|--------------------|---------|-----------|
| \bar{x} | | |
| Me | | |
| Q1 | | |
| Q3 | | |
| V | | |
| s | | |
| IRQ | | |

Zadanie 3

Wygeneruj próbę (o liczebności 1000) o rozkładzie normalnym ($\bar{x}=7$ i $s=1,5$).

Wykonaj obliczenia podstawowych charakterystyk i uzupełnij Tablicę 1.

Przydatne polecenia w R:

```
dane=rnorm(1000, 7, 1.5)
plot(dane)
hist(dane)
mean(dane)
sd(dane)
var(dane)
summary(dane)
```

Proszę pamiętać, że polecenie dostępu do pomocy w R można uzyskać wpisując znak `?` przed nazwą funkcji np. `?hist`.

Wprowadzenie do R (<http://cran.r-project.org/doc/contrib/Komsta-Wprowadzenie.pdf>) autorstwa Łukasza Komsty pozwoli na pełniejsze zapoznanie się z możliwościami pakietu R, który jest darmowym środowiskiem programistycznym użytecznym w analizach statystycznych (<http://www.r-project.org/>) natomiast RStudio (<http://www.rstudio.com/>) sprawia, że praca w R staje się przyjemnością.

Tablica 2.2:

| x_i | $x_i - \bar{x}$ | $ x_i - \bar{x} $ | $(x_i - \bar{x})^2$ | Szereg uporządkowany | Przedziały | Częstość % | Częstość skumulowana w przedziałach % |
|----------|-----------------|-------------------|---------------------|----------------------|------------|------------|---------------------------------------|
| 0,1 | | | | | | | |
| 1,3 | | | | | | | |
| 1,6 | | | | | | | |
| 1,8 | | | | | | | |
| 0,9 | | | | | | | |
| 2,0 | | | | | | | |
| 0,8 | | | | | | | |
| 1,0 | | | | | | | |
| 1,8 | | | | | | | |
| 0,7 | | | | | | | |
| 1,3 | | | | | | | |
| 1,5 | | | | | | | |
| 1,8 | | | | | | | |
| 1,8 | | | | | | | |
| 1,3 | | | | | | | |
| 1,7 | | | | | | | |
| 3,0 | | | | | | | |
| 2,5 | | | | | | | |
| 1,2 | | | | | | | |
| 1,9 | | | | | | | |
| Σ | | | | | | | |

Rozdział 3

Podstawowe miary statystyczne - szereg rozdzielczy

Na podstawie zamieszczonych poniżej wzorów oraz tabeli zbiorczej narysuj histogram rozkładu oraz przeprowadź obliczenia i dokonaj interpretacji dla podanych poniżej miar statystycznych

Tablica 3.1:

| k | Wysokość w mnpm | L. przy- pad- ków n_i | L. sku- mu- lo- wane | Środek prze- działu | $x_i n_i$ | $(x_i - \bar{x}) n_i$ | $(x_i - \bar{x})^2 n_i$ | $(x_i - \bar{x})^3 n_i$ | $(x_i - \bar{x})^4 n_i$ |
|---|--------------------|-------------------------------------|----------------------------------|---------------------------|-----------|-----------------------|-------------------------|-------------------------|-------------------------|
| 1 | 30-49,9 | 10 | | 40 | | | | | |
| 2 | 50-69,9 | 16 | | 60 | | | | | |
| 3 | 70-89,9 | 27 | | 80 | | | | | |
| 4 | 90-109,9 | 19 | | 100 | | | | | |
| 5 | 110-129,9 | 6 | | 120 | | | | | |
| 6 | 130-149,9 | 2 | | 140 | | | | | |
| X | Σ | | | Σ | | | | | |

Wybrane wzory

Średnia

$$\bar{x} = \frac{1}{n} \sum_1^k x_i n_i$$

Dominanta

$$D = x_0 + \frac{n_0 - n_{-1}}{(n_0 - n_{n-1}) + (n_0 - n_{+1})} c_0$$

gdzie:

n_0 - liczebność przedziału mody, n_{-1} - liczebność przedziału poprzedzającego przedział mody, n_{+1} - liczebność przedziału następującego po przedziale mody, x_0 - dolna granica przedziału najliczniejszego, c_0 - rozpiętość przedziału najliczniejszego

Mediana

$$Me = x_0 + [0,5(n+1) - N_{x_{i-1}}] \frac{c_0}{n_0}$$

gdzie:

x_0 – dolna granica przedziału mediany, n_0 – liczebność przedziału mediany, c_0 – rozpiętość przedziału mediany, $N_{x_{i-1}}$ – liczebność szeregu skumulowanego do przedziału poprzedzającego ten, zawierający medianę

Kwartył 1 & 3

$$Q1 = x_0 + [0,25(n+1) - N_{x_{i-1}}] \frac{c_0}{n_0}, \quad Q3 = x_0 + [0,75(n+1) - N_{x_{i-1}}] \frac{c_0}{n_0}$$

Odchylenie standardowe

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2 n_i}{n}}$$

Odchylenie przeciętne

$$d = \frac{\sum_{i=1}^k |x_i - \bar{x}| n_i}{n}$$

Współczynnik zmienności

$$V(d) = \frac{d}{\bar{x}} 100$$

Asymetria

| | |
|---|---|
| Asymetria $\mu_3 = \frac{\sum_{i=1}^n (x_i - \mu)^3 n_i}{n}$ | Moment 3-ci względny $\alpha_3 = \frac{\mu_3}{\sigma^3} - 3$ |
|---|---|

Kurtoza

| | |
|---|---|
| Kurtoza $\mu_4 = \frac{\sum_{i=1}^n (x_i - \mu)^4 n_i}{n}$ | Moment 4-ty względny $\alpha_4 = \frac{\mu_4}{\sigma^4}$ |
|---|---|

Rozdział 4

Prawdopodobieństwo - wprowadzenie

Rozdział 5

Rozkłady teoretyczne - podstawy

Rozdział 6

Rozkład Gauss'a

Rozdział 7

Miary korelacji/asocjacji

Rozdział 8

Regresja

Rozdział 9

Trend

Rozdział 10

Kreślenie map w R - interpolacja