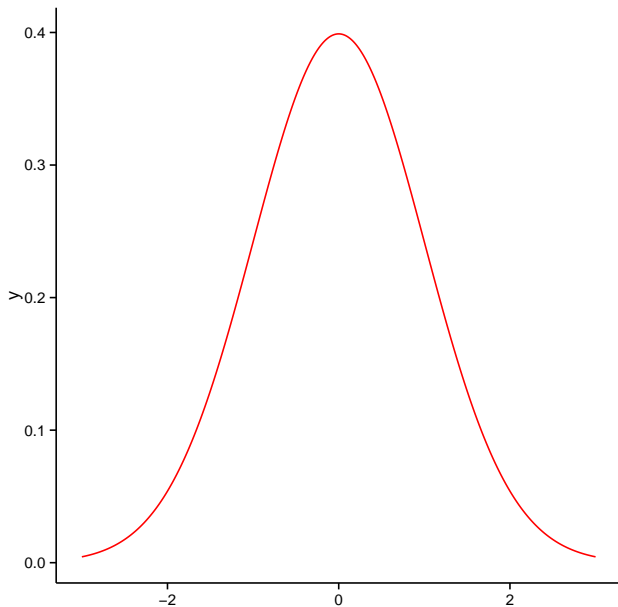


Statystyka w analizie i planowaniu eksperymentu

Paweł Błażej

24 marca 2017

Jak sprawdzić czy dane pochodzą z rozkładu normalnego?



Jak sprawdzić czy nasze dane pochodzą z rozkładu normalnego?

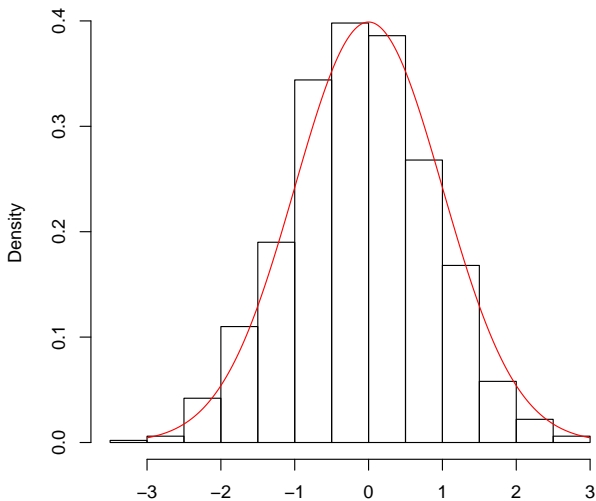
Jest to oczywiste pytanie ponieważ:

wszystkie metody, które do tej pory poznaliśmy opierają się na tym założeniu.

(a) histogram;

- (a) histogram;
- (b) wykres Q-Q plot.

histogram



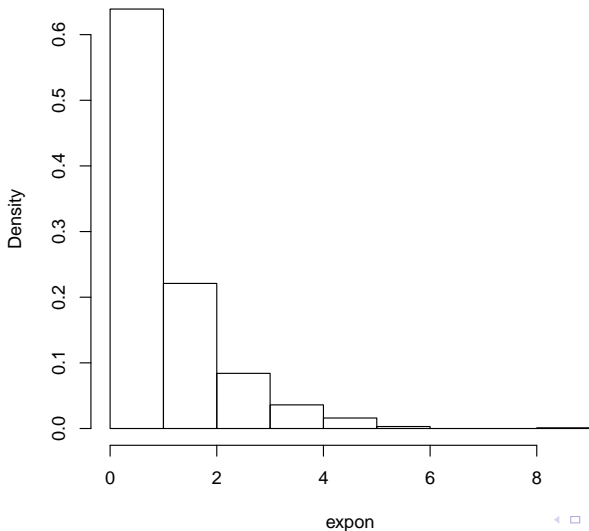
normalny

Paweł Błażej

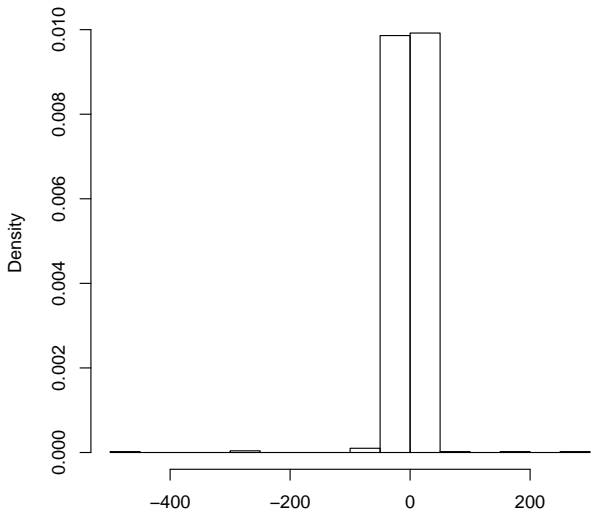
Statystyka w analizie i planowaniu eksperymentu



histogram

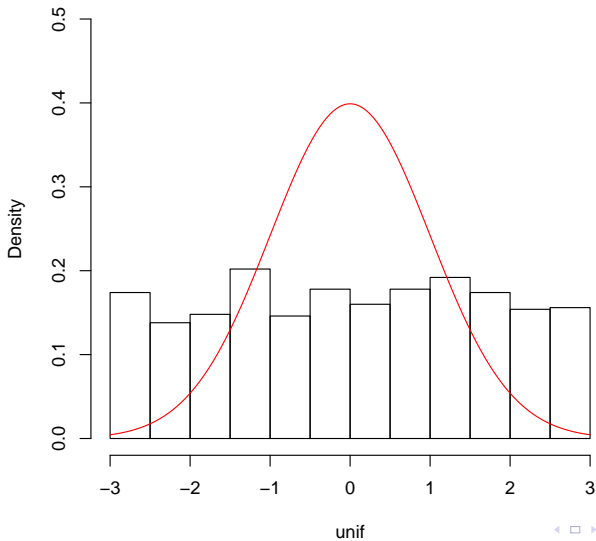


histogram

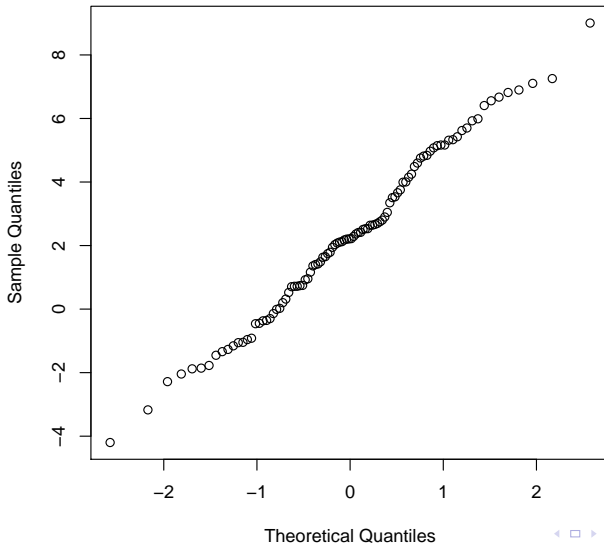


cauchy

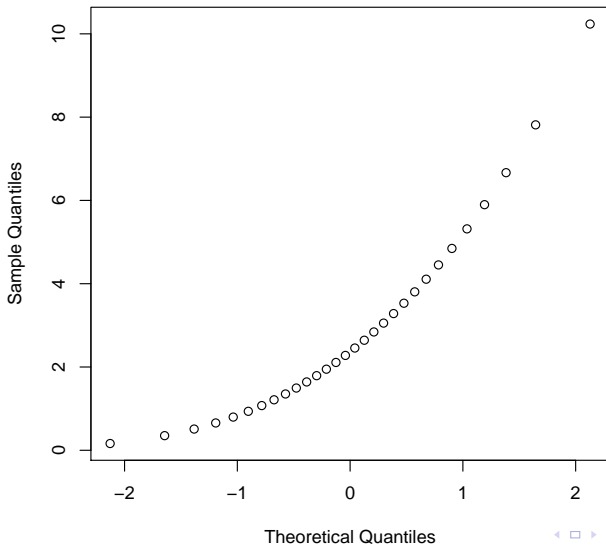
histogram



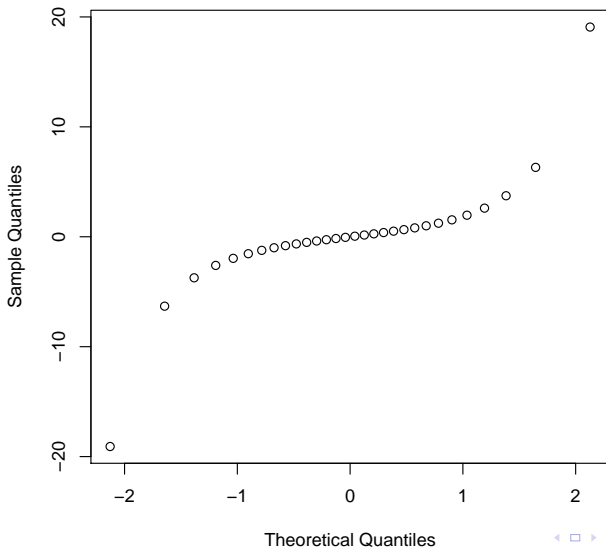
Normal Q-Q Plot



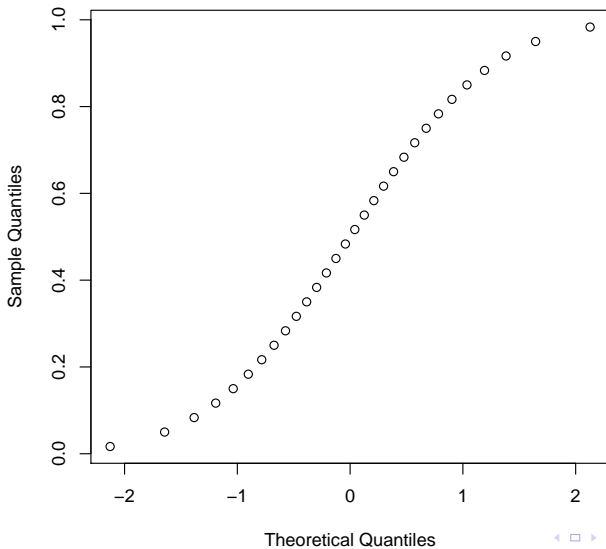
Normal Q-Q Plot



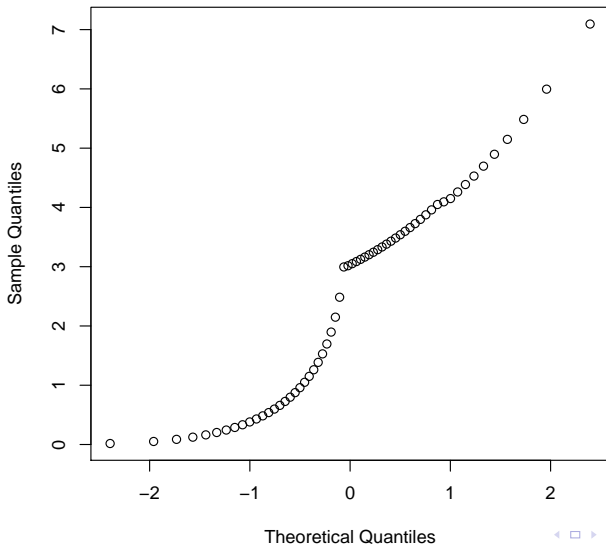
Normal Q-Q Plot



Normal Q-Q Plot



Normal Q-Q Plot



(a) zalety;

- (a) zalety;
- (b) wady.

Kurtoza - miara spłaszczenia rozkładu

$$K = \frac{E(X - E[X])^4}{\sigma^4} - 3$$

Uwaga

Dla rozkładu normalnego kurtoza wynosi 0.

Rozkłady prawdopodobieństwa można podzielić ze względu na wartość kurtozy na rozkłady:

- 1 **mezokurtyczne** - wartość kurtozy wynosi 0, spłaszczenie rozkładu jest podobne do spłaszczenia rozkładu normalnego (dla którego kurtoza wynosi dokładnie 0);

Rozkłady prawdopodobieństwa można podzielić ze względu na wartość kurtozy na rozkłady:

- 1 **mezokurtyczne** - wartość kurtozy wynosi 0, spłaszczenie rozkładu jest podobne do spłaszczenia rozkładu normalnego (dla którego kurtoza wynosi dokładnie 0);
- 2 **leptokurtyczne** - kurtoza jest dodatnia, wartości cechy bardziej skoncentrowane niż przy rozkładzie normalnym;

Rozkłady prawdopodobieństwa można podzielić ze względu na wartość kurtozy na rozkłady:

- 1 **mezokurtyczne** - wartość kurtozy wynosi 0, spłaszczenie rozkładu jest podobne do spłaszczenia rozkładu normalnego (dla którego kurtoza wynosi dokładnie 0);
- 2 **leptokurtyczne** - kurtoza jest dodatnia, wartości cechy bardziej skoncentrowane niż przy rozkładzie normalnym;
- 3 **platokurtyczne** - kurtoza jest ujemna, wartości cechy mniej skoncentrowane niż przy rozkładzie normalnym.

$$\gamma_3 = \frac{E(X - E[X])^3}{\sigma^3}$$

(a) $\gamma_3 = 0, Mo = Me = \mu$

(a) $\gamma_3 = 0, Mo = Me = \mu$

(b) $\gamma_3 < 0, \mu < Me < Mo$

(a) $\gamma_3 = 0, Mo = Me = \mu$

(b) $\gamma_3 < 0, \mu < Me < Mo$

(c) $\gamma_3 > 0, \mu > Me > Mo$

Jak to zrobić w R lub w Excelu?

- (1) H_0 : dane pochodzą z rozkładu normalnego;
- (2) H_1 : dane nie pochodzą z rozkładu normalnego;

(a) test Kołmogorowa-Smirnowa;

- (a) test Kołmogorowa-Smirnowa;
- (b) test Shapiro-Wilka;

Jak to zrobić w R lub w Excelu?