Definicje

Kwantylem rzędu α $\left(\frac{1}{2n} \leqslant \alpha \leqslant 1 - \frac{1}{2n}\right)$ dla danych $x^T = [x_1, x_2, \ldots, x_n]$ jest liczba

$$q(\alpha) = q(\alpha, x) = (1-u)\, x_{(k)} + u x_{(k+1)},$$

$$k = [n\alpha + 0.5]\,,\ u = \{n\alpha + 0.5\}$$

$[a]$ jest częścią całkowitą a, $\{a\}$ - częścią ułamkową a

Rzutem kwantylowym k-wymiarowym wektora x o długości $l(x) = n$ $(k \leqslant n)$ nazywamy k-wymiarowy wektor $x^{[k]}$:

$$x_i^{[k]} = \begin{cases} x_i & k = n \\ q\left(\frac{2i-1}{2k}, x\right) & k < n \end{cases}$$

dla $i = 1, 2, \ldots, k$

Wektory x i y są ***kwantylowo równoważne*** $x \overset{q}{=} y$ gdy $x^{[k]} = y^{[k]}$ dla $k = \min(l(x), l(y))$

Środkiem i rozrzutem kwartylowym wektora x, nazywamy

$$M(x) = \frac{q(0.75, x) + q(0.25, x)}{2},$$

$$Q(x) = \frac{q(0.75, x) - q(0.25, x)}{2}$$

Wykresem kwantylowym (***QQ plot***) pary wektorów (x, y) jest zbiór

$$\left(x_i^{[k]}, y_i^{[k]}\right),\ k = \min(l(x), l(y)),\ i = 1, 2, \ldots, k$$

Dane

Dane 1. Policzono liczbę piór sterówek (piór ogonowych) u gołębi na Śląsku i na Mazowszu

Śląsk	14	17	25	30	31	35	38	42		
Mazowsze	24	28	32	34	36	36	38	39	40	43

Tabela 1. Dane 1. Liczba piór sterówek u gołębi

Zadania

1. Jaka jest interpretacja
$q(0.5)$,
$q(\alpha_k)\left(\alpha_k = \frac{2k-1}{2n},\ k = 1, 2, \ldots, n.\right)$
Ile jest elementów mniejszych od $q\left(\frac{k}{n}\right)$, $(1 \leqslant k < n)$ gdy $x_{(k)} < x_{(k+1)}$?

2. Niech $q(\alpha) = (1-u)\, x_{(k)} + u x_{(k+1)}$.
Oblicz $q(1-\alpha)$.
Pokaż, że $q(1-\alpha, x) = -q(\alpha, -x)$.
3. Pokaż, że wykres funkcji $q(\alpha, x)$ dla $\frac{1}{2n} \leqslant \alpha \leqslant 1 - \frac{1}{2n}$ jest niemalejącą łamaną o węzłach $\left(\alpha_k, x_{(k)}\right)$.
4. Gdy x i y mają tą samą liczbę elementów to $x \stackrel{q}{=} y$ wtedy i tylko wtedy, gdy $x_{(i)} = y_{(i)}$ dla każdego $i = 1, 2, \ldots, n$
5. Znajdź rzut kwantylowy:
(a) wektora$[1, 5, 5, 8]$ dla $k = 3$
(b) wektora Mazowsze dla $k = 8$
6. $\int_0^1 q(\alpha, x)\, d\alpha = \overline{x}$, gdzie $\overline{x}$ jest średnią arytmetyczną x.
7. $x \stackrel{q}{=} y$ wtedy i tylko wtedy, gdy ich wykres kwantylowy leży na prostej o równaniu $u = t$
8. $y \stackrel{q}{=} ax + b$, $a \geqslant 0$ wtedy i tylko wtedy, gdy wykres kwantylowy pary wektorów (x, y) leży na prostej o równaniu $u = at + b$.
9. Wykres kwantylowy pary wektorów (x, y) leży na prostej o równaniu $u = at + b$.
(a) Pokaż, że $a \geqslant 0$
(b) $y \stackrel{q}{=} ax + b$, $a \simeq \frac{Q\left(y^{[k]}\right)}{Q\left(x^{[k]}\right)}$, $M\left(y^{[k]}\right) \simeq aM\left(x^{[k]}\right) + b$ dla $k = \min\left(l(x), l(y)\right)$ (przybliżona prosta kwantylowa)
10. Narysuj wykresy kwantylowe wektora pary wektorów $(x, -x)$ gdy $x = [1, 2, 3]$ i gdy $x = [1, 2, 4]$. Czy oba wykresy leżą na jakiejś linii prostej?
11. Wykres kwantylowy pary wektorów $(x, -x)$ j leży na prostej o równaniu $u = at + b$. Wyznacz parametry a i b. Jakie są interpretacje a i b ?
12.

(a) Narysuj wykres kwantylowy dla pary (Śląsk, Mazowsze) (*Dane 1*) . Czy leży on w przybliżeniu na prostej?

(b) Jakie dane należy wykluczyć, aby wykres kwantylowy lepiej leżał na jakiejś prostej? Wyznacz parametry tej prostej. Co ta relacja oznacza?

(c) Korzystając z zad . 9(b) narysuj przybliżoną prostą kwantylową dla pary (Śląsk, Mazowsze) (*Dane 1*)