Asymptotyka wartości ekstremalnych i rekordowych

Wiesław Dziubdziela

Instytut Matematyczny Uniwersytetu Jana Kochanowskiego w Kielcach

22 sierpnia 2016

1. Teoria wartości ekstremalnych. Ciąg niezależnych zmiennych losowych o jednakowym rozkładzie.
2. Teoria wartości rekordowych. Ciąg niezależnych zmiennych losowych o jednakowym rozkładzie.
3. Badania statystyczne w oparciu o wartości rekordowe
4. Zakończenie

Niech $X_1, X_2, \ldots$ będzie ciągiem niezależnych (i.i.d.) zmiennych losowych o dystrybuancie F. Niech $M_n = \max\{X_1, \ldots, X_n\}$ oznacza maksimum pierwszych n zmiennych losowych i niech $\omega(F) = \sup\{x : F(x) < 1\}$. Ponieważ

$$P(M_n \leq x) = P(X_1 \leq x, \ldots, X_n \leq x) = F^n(x),$$

więc M_n jest zbieżne prawie na pewno do $\omega(F)$ bez względu na to czy ω jest skończona czy też nieskończona.

Rozpoczniemy od badania zbieżności $P\{M_n \leq u_n\}$, gdy $n \longrightarrow \infty$, gdzie u_n jest pewnym ciągiem liczb rzeczywistych.

Rozpoczniemy od badania zbieżności $P\{M_n \le u_n\}$, gdy $n \longrightarrow \infty$, gdzie u_n jest pewnym ciągiem liczb rzeczywistych.

- W monografii: M.R. Leadbetter, Georg Lindgren, Holger Rootzén, *Extremes and related properties of random sequences and processes*, Springer, New York 1983, można znaleźć następujące twierdzenie:

Rozpoczniemy od badania zbieżności $P\{M_n \le u_n\}$, gdy $n \longrightarrow \infty$, gdzie u_n jest pewnym ciągiem liczb rzeczywistych.

- W monografii: M.R. Leadbetter, Georg Lindgren, Holger Rootzén, *Extremes and related properties of random sequences and processes*, Springer, New York 1983,
  można znaleźć następujące twierdzenie:
- **Twierdzenie**. Niech X_n będzie i.i.d. ciągiem. Niech $0 \le \tau \le \infty$ i zakładamy, że u_n jest ciągiem liczb rzeczywistych takim, że

$$n(1 - F(u_n)) \to \tau, \text{ gdy } n \to \infty \tag{1}$$

wtedy

$$P\{M_n \le u_n\} \to e^{-\tau}, \text{ gdy } n \to \infty. \tag{2}$$

Odwrotnie, jeżeli (2) zachodzi dla pewnego τ, $0 \le \tau \le \infty$, to zachodzi także (1).

Podstawowe pytanie klasycznej teorii wartości ekstremalnych

Czy istnieją ciągi stałych normujących $a_n > 0$ i $b_n \in \mathbf{R}$ takie, że $\frac{M_n - b_n}{a_n}$ jest zbieżne według rozkładu do niezdegenerowanej dystrybuanty granicznej? Jakiej postaci są a_n i b_n i jakie są możliwe rozkłady graniczne?

Pierwsze odpowiedzi:

Podstawowe pytanie klasycznej teorii wartości ekstremalnych

Czy istnieją ciągi stałych normujących $a_n > 0$ i $b_n \in \mathbf{R}$ takie, że $\frac{M_n - b_n}{a_n}$ jest zbieżne według rozkładu do niezdegenerowanej dystrybuanty granicznej? Jakiej postaci są a_n i b_n i jakie są możliwe rozkłady graniczne?

Pierwsze odpowiedzi:

- R.A. Fisher, L.H.C. Tippet, Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Phil. Soc.* **24**(1928), 180 -190

Podstawowe pytanie klasycznej teorii wartości ekstremalnych

Czy istnieją ciągi stałych normujących $a_n > 0$ i $b_n \in \mathbf{R}$ takie, że $\frac{M_n - b_n}{a_n}$ jest zbieżne według rozkładu do niezdegenerowanej dystrybuanty granicznej? Jakiej postaci są a_n i b_n i jakie są możliwe rozkłady graniczne?

Pierwsze odpowiedzi:

- R.A. Fisher, L.H.C. Tippet, Limiting forms of the frequency distribution of the largest or smallest member of a sample,*Proc. Cambridge Phil. Soc.* **24**(1928), 180 -190
- B.V. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.* **44** (1943), 423 -453.,

Zanim odpowiemy na zadane wcześniej pytania, wprowadzimy definicję typu rozkładu i przedstawimy twierdzenie Khintchina

Zanim odpowiemy na zadane wcześniej pytania, wprowadzimy definicję typu rozkładu i przedstawimy twierdzenie Khintchina

- A. Khinchin *Limit laws for sums of independent random variables* (po rosyjsku), Moskwa 1938.

Twierdzenie o zbieżności typów rozkładów

Zanim odpowiemy na zadane wcześniej pytania, wprowadzimy definicję typu rozkładu i przedstawimy twierdzenie Khintchina

- A. Khinchin *Limit laws for sums of independent random variables* (po rosyjsku), Moskwa 1938.
- **Definicja typu rozkładu.** Niech X i Y będą dwiema zmiennymi losowymi z rozkładami μ i ν odpowiednio. Mówimy, że μ i ν są tego samego typu, jeżeli istnieją $a > 0$ i $b \in \mathbf{R}$ takie, że $aX + b$ ma taki sam rozkład jak Y.

Podamy teraz twierdzenie o zbieżności typów rozkładów nazywane twierdzeniem Khintchina.

Podamy teraz twierdzenie o zbieżności typów rozkładów nazywane twierdzeniem Khintchina.

- **Twierdzenie.** Niech W_n będzie ciągiem zmiennych losowych zbieżnym słabo do W, oraz dla pewnych $a_n > 0$ i $b_n \in \mathbf{R}$,

$$a_n W_n + b_n \Rightarrow W',$$

gdzie W i W' są niezdegenerowane. Wtedy $a_n \to a$ i $b_n \to b$ dla pewnych $a > 0$ i $b \in \mathbf{R}$.
Równoważnie, jeżeli F_n, F i F' są dystrybuantami, przy czym F i F' są niezdegenerowane, i istnieją $a_n, a'_n > 0$ i $b_n, b'_n \in \mathbf{R}$ takie, że $F_n(a_n x + b_n) \xrightarrow{w} F(x)$ i $F_n(a'_n x + b'_n) \xrightarrow{w} F'(x)$ wtedy $\frac{a_n}{a'_n} \to a > 0$, $\frac{(b_n - b'_n)}{a'_n} \to b \in \mathbf{R}$, przy czym

$$F'(ax + b) = F(x)$$

dla wszystkich $x \in \mathbf{R}$.

Zdefiniujemy teraz rozkłady maksymaknie stabilne i podamy ich pewną charakteryzację.

Zdefiniujemy teraz rozkłady maksymaknie stabilne i podamy ich pewną charakteryzację.

- Mówimy, że dystrybuanta F jest maksymalnie stabilna, jeżeli istnieją stałe $a_n > 0,\ b_n \in \mathbf{R}$ takie, że dla każdego $n \geq 1$

$$F^n(x) = F(a_n x + b_n).$$

Zdefiniujemy teraz rozkłady maksymaknie stabilne i podamy ich pewną charakteryzację.

- Mówimy, że dystrybuanta F jest maksymalnie stabilna, jeżeli istnieją stałe $a_n > 0, \; b_n \in \mathbf{R}$ takie, że dla każdego $n \geq 1$

$$F^n(x) = F(a_n x + b_n).$$

- **Twierdzenie (Rozkłady maksymalnie stabilne jako słabe granice liniowo unormowanych maksimów).** Niezdegenerowana dystrybuanta F jest maksymalnie stabilna wtedy i tylko wtedy, gdy istnieje ciąg $X_1, X_2, \ldots$ niezależnych o jednakowym rozkładzie G zmiennych losowych oraz istnieją stałe normujące $a_n > 0, b_n$ taka, że

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) = G^n(a_n x + b_n) \xrightarrow{w} F(x)$$

(symbol $\xrightarrow{w}$ oznacza słabą zbieżność dystrybuant, czyli w punktach x ciągłości dystrybuanty granicznej F), gdy $n \to \infty$.

Twierdzenie o ekstremalnych typach. Każda maksymalnie stabilna dystrybuanta F jest ekstremalnego typu, tzn. jest takiego samego typu jak jedna z następujących dystrybuant:

Twierdzenie o ekstremalnych typach. Każda maksymalnie stabilna dystrybuanta F jest ekstremalnego typu, tzn. jest takiego samego typu jak jedna z następujących dystrybuant:

- **Typ I, Rozkład Gumbela**

$$\Lambda(x) = \exp\{-\exp(-x)\}, \quad -\infty < x < \infty;$$

Twierdzenie o ekstremalnych typach. Każda maksymalnie stabilna dystrybuanta F jest ekstremalnego typu, tzn. jest takiego samego typu jak jedna z następujących dystrybuant:

- **Typ I, Rozkład Gumbela**

$$\Lambda(x) = \exp\{-\exp(-x)\}, \quad -\infty < x < \infty;$$

- **Typ II, Rozkład Frécheta**

$$\Phi_{\alpha}(x) = \exp\left\{-x^{-\alpha}\right\}, \quad x \geq 0,$$

dla pewnego $\alpha > 0$;

Twierdzenie o ekstremalnych typach. Każda maksymalnie stabilna dystrybuanta F jest ekstremalnego typu, tzn. jest takiego samego typu jak jedna z następujących dystrybuant:

- **Typ I, Rozkład Gumbela**

$$\Lambda(x) = \exp\left\{-\exp(-x)\right\}, \quad -\infty < x < \infty;$$

- **Typ II, Rozkład Frécheta**

$$\Phi_\alpha(x) = \exp\left\{-x^{-\alpha}\right\}, \quad x \geq 0,$$

  dla pewnego $\alpha > 0$;

- **Typ III, Rozkład Weibulla**

$$\Psi_\alpha(x) = \exp\left\{-(-x)^\alpha\right\}, \quad x < 0,$$

  dla pewnego $\alpha > 0$.

Szeroko akceptowaną standardową reprezentacją rozkładów ekstremalnych jest reprezentacja Jenkinsona-von Misesa postaci

$$
\begin{aligned}
G_\gamma(x) = EV_\gamma(x) &= \exp\left\{-(1+\gamma x)^{-1/\gamma}\right\} \ \texttt{if}\ 1+\gamma x > 0, \gamma \neq 0, \\
&= \exp\left\{-e^{-x}\right\} \ \texttt{if}\ x \in \mathbf{R}, \gamma = 0.
\end{aligned}
$$

Dystrybuanta EV_γ jednolicie zapisuje trzy rodzaje ekstremalnych rozkładów.

Szeroko akceptowaną standardową reprezentacją rozkładów ekstremalnych jest reprezentacja Jenkinsona-von Misesa postaci

$$\begin{aligned} G_\gamma(x) = EV_\gamma(x) &= \exp\left\{-\left(1+\gamma x\right)^{-1/\gamma}\right\} \ \texttt{if } 1+\gamma x > 0, \gamma \neq 0, \\ &= \exp\left\{-e^{-x}\right\} \ \texttt{if } x \in \mathbf{R}, \gamma = 0. \end{aligned}$$

Dystrybuanta EV_γ jednolicie zapisuje trzy rodzaje ekstremalnych rozkładów.

- **Rozkład Gumbela,**

$$\Lambda(x) = EV_0(x) = \exp\{-\exp(-x)\}, \quad -\infty < x < \infty, \gamma = 0;$$

Uogólniony rozkład ekstremalny

Szeroko akceptowaną standardową reprezentacją rozkładów ekstremalnych jest reprezentacja Jenkinsona-von Misesa postaci

$$G_\gamma(x) = EV_\gamma(x) = \exp\left\{-(1+\gamma x)^{-1/\gamma}\right\} \texttt{ if } 1+\gamma x > 0, \gamma \neq 0,$$
$$= \exp\left\{-e^{-x}\right\} \texttt{ if } x \in \mathbf{R}, \gamma = 0.$$

Dystrybuanta EV_γ jednolicie zapisuje trzy rodzaje ekstremalnych rozkładów.

- **Rozkład Gumbela,**

  $$\Lambda(x) = EV_0(x) = \exp\{-\exp(-x)\}, \quad -\infty < x < \infty, \gamma = 0;$$

- **Rozkład Frécheta,**

  $$\Phi_\alpha(x) = EV_{1/\alpha}(\alpha(x-1)) = \exp\left\{-x^{-\alpha}\right\}, \; x > 0, \gamma = 1/\alpha > 0;$$

Szeroko akceptowaną standardową reprezentacją rozkładów ekstremalnych jest reprezentacja Jenkinsona-von Misesa postaci

$$G_\gamma(x) = EV_\gamma(x) = \exp\left\{-(1+\gamma x)^{-1/\gamma}\right\} \ \texttt{if}\ 1+\gamma x > 0, \gamma \neq 0,$$
$$= \exp\left\{-e^{-x}\right\} \ \texttt{if}\ x \in \mathbf{R}, \gamma = 0.$$

Dystrybuanta EV_γ jednolicie zapisuje trzy rodzaje ekstremalnych rozkładów.

- **Rozkład Gumbela,**

  $$\Lambda(x) = EV_0(x) = \exp\{-\exp(-x)\}, \quad -\infty < x < \infty, \gamma = 0;$$

- **Rozkład Frécheta,**

  $$\Phi_\alpha(x) = EV_{1/\alpha}(\alpha(x-1)) = \exp\left\{-x^{-\alpha}\right\}, \ x > 0, \gamma = 1/\alpha > 0;$$

- **Rozkład Weibulla,**

  $$\Psi_\alpha(x) = EV_{-1/\alpha}(\alpha(x+1)) = \exp\{-(-x)^\alpha\},$$
  $$x < 0, \gamma = -1/\alpha < 0.$$

Niech X będzie zmienną losową o dystrybuancie F oraz niech $p > 0$ będzie "bardzo małe".

Niech X będzie zmienną losową o dystrybuancie F oraz niech $p > 0$ będzie "bardzo małe".

- **ogon dystrybuanty** [tail probability]

$$p = P\left(X > x_p\right) = 1 - F\left(x_p\right);$$

Niech X będzie zmienną losową o dystrybuancie F oraz niech $p > 0$ będzie "bardzo małe".

- **ogon dystrybuanty** [tail probability]

$$p = P\left(X > x_p\right) = 1 - F\left(x_p\right);$$

- **wysokie kwantyle** [high quantiles]

$$(1 - p) - \texttt{kwantyl}\ x_p;$$

Niech X będzie zmienną losową o dystrybuancie F oraz niech $p > 0$ będzie "bardzo małe".

- **ogon dystrybuanty** [tail probability]

$$p = P(X > x_p) = 1 - F(x_p);$$

- **wysokie kwantyle** [high quantiles]

$$(1-p) - \texttt{kwantyl } x_p;$$

- **prawe ograniczenie suportu** [right endpoint]

$$x^F = \sup\{x : F(x) < 1\}.$$

Definicja. Mówimy, ż dystrybuanta F należy do obszaru przysiągania (dla maksimów) rozkładu ekstremalnego $G_\gamma(x) = EV_\gamma(x)$ i piszemy $F \in D(EV_\gamma)$, gdy istnieją stałe $a_n > 0,\ b_n \in \mathbf{R}$ takie, że dla $n \to \infty$

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) = F^n(a_n x + b_n) \xrightarrow{w} EV_\gamma(x).$$

Konsekwencją istnienia powyższej granicy jest aproksymacja wykorzystywana w zastosowaniach

$$P(M_n \leq x) = F^n(x) \approx EV_\gamma\left(\frac{x - b_n}{a_n}\right).$$

Aproksymacja uogólnionym rozkładem ekstremalnym

Zapiszmy aproksymację dystrybuanty maksimum $P(M_n \leq x)$ w postaci

$$P(M_n \leq x) \approx \exp\left\{-\left(1+\gamma\frac{x-\mu}{\sigma}\right)^{-1/\gamma}\right\},$$

gdzie

$$1+\gamma\frac{x-\mu}{\sigma} > 0;\;\; \mu \in \mathbf{R}, \sigma > 0 \text{ i } \gamma \in \mathbf{R}.$$

Przypadek $\gamma = 0$ interpretujemy jako granicę $\gamma \to 0$. Otrzymujemy wtedy aproksymację rozkładem Gumbela

$$P(M_n \leq x) \approx \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}, \text{ gdzie } -\infty < x < \infty.$$

Wykorzystując aproksymację uogólnionym rozkładem ekstremalnym, otrzymujemy

$$G_\gamma(x_p) = 1 - p \Rightarrow x_p = \mu - \frac{\sigma}{\gamma}\left[1 - \{-\ln(1-p)\}^{-\gamma}\right].$$

Parametry μ i σ interpretujemy jako parametry położenia i skali, odpowiednio.
Natomiast wartości parametru γ wyznaczają zachowanie asymptotyczne ogonów dystrybuanty. Parametr ten jest nazywany parametrem kształtu. Mierzy on grubość prawego ogona $\overline{F}(x) = 1 - F(x)$.

Interpretacje parametru γ:

Interpretacje parametru γ:

- **Gumbel.** Jeżeli $\gamma = 0$, to ogon jest wykładniczy, przy czy $x^F = \sup \{x : F(x) < 1\}$ jest skończony albo nieskończony.

Interpretacje parametru γ:

- **Gumbel.** Jeżeli $\gamma = 0$, to ogon jest wykładniczy, przy czy $x^F = \sup\{x : F(x) < 1\}$ jest skończony albo nieskończony.
- **Fréchet.** Jeżeli $\gamma > 0$, to ogon jest wielomianowy, gruby, przy czym x^F jest nieskończone.

Interpretacje parametru γ:

- **Gumbel.** Jeżeli $\gamma = 0$, to ogon jest wykładniczy, przy czy $x^F = \sup\{x : F(x) < 1\}$ jest skończony albo nieskończony.
- **Fréchet.** Jeżeli $\gamma > 0$, to ogon jest wielomianowy, gruby, przy czym x^F jest nieskończone.
- **Weibull.** Jeżeli $\gamma < 0$, to ogon jest krótki, przy czym x^F jest skończone.

Wyznaczenie parametru γ teoretyczne lub statystyczne jest jednym z najważniejszych problemów współczesnej teorii wartości ekstremalnych. Poświęcono mu wiele uwagi.

Problem: Wyznaczanie granicznej dystrybuanty (o ile istnieje) dla liniowo unormowanego maksimum na podstawie znajomoći dystrybuanty F?
Innaczej mówiąc, pytamy się do jakiego obszaru przyciągania $D(EV_\gamma)$ należy dystrybuanta F?
Pierwszą odpowiedzią są tzw. **dostateczne warunki von Mises'a**.

Problem: Wyznaczanie granicznej dystrybuanty (o ile istnieje) dla liniowo unormowanego maksimum na podstawie znajomoći dystrybuanty F?
Innaczej mówiąc, pytamy się do jakiego obszaru przyciągania $D(EV_\gamma)$ należy dystrybuanta F?
Pierwszą odpowiedzią są tzw. **dostateczne warunki von Mises'a**.

- R. von Mises (1936), La distribution de la plus grande de n valeurs, Reprinted in *Selected Papers Volumen II, American Mathematical Society*, Providence, R.I., 1954, str. 271-294.

Problem: Wyznaczanie granicznej dystrybuanty (o ile istnieje) dla liniowo unormowanego maksimum na podstawie znajomoći dystrybuanty F?
Innaczej mówiąc, pytamy się do jakiego obszaru przyciągania $D(EV_\gamma)$ należy dystrybuanta F?
Pierwszą odpowiedzią są tzw. **dostateczne warunki von Mises'a**.

- R. von Mises (1936), La distribution de la plus grande de n valeurs, Reprinted in *Selected Papers Volumen II, American Mathematical Society*, Providence, R.I., 1954, str. 271-294.
- W tym fragmencie prezentacji opieram się między innymi na rozdziale 4 pt. *Extreme Value Theory: An Introductory Overview* , napisanym przez Maria Isabel Fraga Alves i Cáudia Neves monografii *Extreme Events in Finance: A Handbook of Extreme Value Theory and its Applications*, edited by Francois Longin, Wiley, October 2016.

Załóżmy, że istnieją gęstość $f(x) = F'(x)$ oraz $F''(x)$. Definiujemy

$$h(x) = \frac{f(x)}{1 - F(x)} \text{ and } r(x) = \frac{1 - F(x)}{f(x)}.$$

Jeżeli

$$\lim_{x \to x^F} r'(x) = \gamma,$$

to

$$F \in D(EV_\gamma).$$

ze stałymi normującymi

$$b_n = F^{\leftarrow}\left(1 - \frac{1}{n}\right) = U(n) \text{ i } a_n = r(b_n) = \frac{1}{nf(b_n)} = nU'(n),$$

gdzie

$$F^{\leftarrow}(u) = \inf\{x : F(x) \geq u\} \ i \ U(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right), \ t \in [1, \infty].$$

Przedstawię także pierwsze ogólne warunki podane przez Gnedenko (1943)

Przedstawię także pierwsze ogólne warunki podane przez Gnedenko (1943)

- B.V. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.* **44** (1943), 423 -453.

Przedstawię także pierwsze ogólne warunki podane przez Gnedenko (1943)

- B.V. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.* **44** (1943), 423 -453.
- $F \in D(EV_\gamma)$, $\gamma \in \mathbf{R}$,
  wtedy i tylko wtedy, gdy

Przedstawię także pierwsze ogólne warunki podane przez Gnedenko (1943)

- B.V. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.* **44** (1943), 423 -453.
- $F \in D(EV_\gamma)$, $\gamma \in \mathbf{R}$, wtedy i tylko wtedy, gdy
- dla $\gamma > 0 : x^F = \infty$ i $\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-1/\gamma}$, dla wszystkich $x > 0$; tzn. $F \in RV_{-1/\gamma}$;

Przedstawię także pierwsze ogólne warunki podane przez Gnedenko (1943)

- B.V. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.* **44** (1943), 423 -453.
- $F \in D(EV_\gamma)$, $\gamma \in \mathbf{R}$,
  wtedy i tylko wtedy, gdy
- dla $\gamma > 0 : x^F = \infty$ i $\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-1/\gamma}$, dla wszystkich $x > 0$; tzn. $F \in RV_{-1/\gamma}$;
- Przypomnijmy, jeżeli dodatnia funkcja h jest taka, że $\lim_{t\to\infty} \frac{h(tx)}{h(t)} = x^\beta$, dla wszystkich $x > 0$, mówimy, że h jest regularnie zmieniająca z indeksem β (w nieskończoności) i piszemy $h \in RV_\beta$;
  jeżeli $\lim_{t\to\infty} \frac{a(tx)}{a(t)} = 1, \mathtt{dla}\ x > 0$, wtedy a jest wolno zmieniająca się, piszemy $a \in RV_0$.

Następne warunki są postaci:

Następne warunki są postaci:

- $F \in D(EV_\gamma)$, $\gamma \in \mathbf{R}$, wtedy i tylko wtedy, gdy

Następne warunki są postaci:

- $F \in D(EV_\gamma)$, $\gamma \in \mathbf{R}$, wtedy i tylko wtedy, gdy
- dla $\gamma < 0 : x^F < \infty$ i $\lim_{t \downarrow 0} \frac{1-F(x^F-tx)}{1-F(x^F-t)} = x^{-1/\gamma}$ dla wszystkich $x > 0$;

Następne warunki są postaci:

- $F \in D(EV_\gamma)$, $\gamma \in \mathbf{R}$, wtedy i tylko wtedy, gdy
- dla $\gamma < 0$: $x^F < \infty$ i $\lim_{t\downarrow 0} \frac{1-F(x^F - tx)}{1-F(x^F - t)} = x^{-1/\gamma}$ dla wszystkich $x > 0$;
- dla $\gamma = 0$: $x^F < \infty$ albo $x^F = \infty$ i

$$\lim_{t\uparrow x^F} \frac{1 - F(t + xg(t))}{1 - F(t)} = e^{-x}$$

dla wszystkich $x > 0$, przy czym $\int_t^{x^F} (1 - F(s))\, ds < \infty$, natomiast g może być wybrane jako

$$g(t) = \frac{\int_t^{x^F} (1 - F(s))\, ds}{1 - F(t)} = E[X - t | X > t] \text{ dla } t < x^F.$$

Niech $U(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right),\ t \in [1, \infty]$.

Niech $U(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right), \ t \in [1, \infty]$.

- $F \in D(EV_\gamma), \ \gamma \in \mathbf{R} \ \Leftrightarrow \ \lim_{t\to\infty} \frac{U(tx)-U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}$, dla pewnej dodatniej funkcji $a(\cdot)$ i $x > 0$.

Niech $U(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right), \ t \in [1, \infty]$.

- $F \in D(EV_\gamma), \ \gamma \in \mathbf{R} \ \Leftrightarrow \ \lim_{t\to\infty} \frac{U(tx)-U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}$, dla pewnej dodatniej funkcji $a(\cdot)$ i $x > 0$.
- Dla $\gamma = 0$ przyjmujemy $\frac{x^\gamma - 1}{\gamma} = \ln x$.

Niech $U(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right)$, $t \in [1, \infty]$.

- $F \in D(EV_{\gamma})$, $\gamma \in \mathbf{R} \Leftrightarrow \lim_{t \to \infty} \frac{U(tx) - U(t)}{a(t)} = \frac{x^{\gamma} - 1}{\gamma}$, dla pewnej dodatniej funkcji $a(\cdot)$ i $x > 0$.
- Dla $\gamma = 0$ przyjmujemy $\frac{x^{\gamma}-1}{\gamma} = \ln x$.
- L. de Haan, A. Ferreira, *Extreme Value Theory: An Introduction*, Springer 2006.

Niech $U(t) = F^{\leftarrow}\left(1 - \frac{1}{t}\right), \; t \in [1, \infty]$.

- $F \in D(EV_\gamma), \; \gamma \in \mathbf{R} \; \Leftrightarrow \; \lim_{t\to\infty} \frac{U(tx)-U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}$, dla pewnej dodatniej funkcji $a(\cdot)$ i $x > 0$.
- Dla $\gamma = 0$ przyjmujemy $\frac{x^\gamma - 1}{\gamma} = \ln x$.
- L. de Haan, A. Ferreira, *Extreme Value Theory: An Introduction*, Springer 2006.
- L. de Haan, *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*, Mathematical Centre Tract 32, Mathematics Centre, Amsterdam 1970.

Grube ogony opisują zjawiska dla których prawdopodobieństwa wystąpienia bardzo dużych wartości są relatywnie duże.

Grube ogony opisują zjawiska dla których prawdopodobieństwa wystąpienia bardzo dużych wartości są relatywnie duże.

- W podręczniku
  Sidney I. Resnick, *Heavy-Tail Phenomena. Probabilistic and Statistical Modeling*, Springer 2007.
  mamy następującą definicję:

Grube ogony opisują zjawiska dla których prawdopodobieństwa wystąpienia bardzo dużych wartości są relatywnie duże.

- W podręczniku
  Sidney I. Resnick, *Heavy-Tail Phenomena. Probabilistic and Statistical Modeling*, Springer 2007.
  mamy następującą definicję:
- Zmienna losowa X ma gruby (prawy) ogon, jeżeli istnieje dodatni parametr $\alpha > 0$ taki, że $P(X > x) \propto x^{-\alpha}, \ x \to \infty$.

Grube ogony opisują zjawiska dla których prawdopodobieństwa wystąpienia bardzo dużych wartości są relatywnie duże.

- W podręczniku
  Sidney I. Resnick, *Heavy-Tail Phenomena. Probabilistic and Statistical Modeling*, Springer 2007.
  mamy następującą definicję:
- Zmienna losowa X ma gruby (prawy) ogon, jeżeli istnieje dodatni parametr $\alpha > 0$ taki, że $P(X > x) \propto x^{-\alpha}, \ x \to \infty$.
- Wykorzystujemy oznaczenie

$$f(x) \propto g(x), \ x \to \infty,$$

jako skrót dla

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 1.$$

Ogólniejszą definicję można znaleźć w monografii:

Ogólniejszą definicję można znaleźć w monografii:

- S.Foss, D. Korshunov, S. Zachary, *An Introduction to Heavy-Tailed and Subexponential Distributions*, Springer, 2013.

Ogólniejszą definicję można znaleźć w monografii:

- S.Foss, D. Korshunov, S. Zachary, *An Introduction to Heavy-Tailed and Subexponential Distributions*, Springer, 2013.
- Zmienna losowa X ma gruby (prawy) ogon, jeżeli wszystkie dodatnie wykładnicze momenty są nieskończone,

$$Ee^{cX} \equiv \int e^{cx} dP(x) = \infty, \ \texttt{dla wszystkich} \ c > 0,$$

(w przeciwnym razie X ma cieńki ogon).

Ogólniejszą definicję można znaleźć w monografii:

- S.Foss, D. Korshunov, S. Zachary, *An Introduction to Heavy-Tailed and Subexponential Distributions*, Springer, 2013.
- Zmienna losowa X ma gruby (prawy) ogon, jeżeli wszystkie dodatnie wykładnicze momenty są nieskończone,

$$Ee^{cX} \equiv \int e^{cx} dP(x) = \infty, \ \texttt{dla wszystkich}\ c > 0,$$

  (w przeciwnym razie X ma cieńki ogon).
- Oczywiście jeżeli X ma gruby ogon w sensie Resnika, to i ma gruby ogon w sensie Fossa. Ta druga klasa jest szersza.

Dystrybuanta F należy do obszaru przyciągania rozkładu Fréchet'a, gdy $1 - F(x) = x^{-\alpha} L_F(x)$, gdzie $\alpha = \frac{1}{\gamma}$, $\gamma > 0$, a $L_F(x)$ jest funkcją wolno zmieniającą się w nieskończoności. Zatem F ma **gruby ogon**.

Niektórzy autorzy przyjmują, że dystrybuanta F ma gruby ogon, gdy $0 < \alpha < 2$.

Dystrybuanta F należy do obszaru przyciągania rozkładu Fréchet'a, gdy $1 - F(x) = x^{-\alpha} L_F(x)$, gdzie $\alpha = \frac{1}{\gamma}$, $\gamma > 0$, a $L_F(x)$ jest funkcją wolno zmieniającą się w nieskończoności. Zatem F ma **gruby ogon**.

Niektórzy autorzy przyjmują, że dystrybuanta F ma gruby ogon, gdy $0 < \alpha < 2$.

- Jeżeli $1 \leq \alpha = \frac{1}{\gamma} < 2$, to nie istnieje wartość oczekiwana $E(X)$, a wariancja $\mathtt{Var}(X) = \infty$.

Dystrybuanta F należy do obszaru przyciągania rozkładu Fréchet'a, gdy $1 - F(x) = x^{-\alpha} L_F(x)$, gdzie $\alpha = \frac{1}{\gamma}$, $\gamma > 0$, a $L_F(x)$ jest funkcją wolno zmieniającą się w nieskończoności. Zatem F ma **gruby ogon**.

Niektórzy autorzy przyjmują, że dystrybuanta F ma gruby ogon, gdy $0 < \alpha < 2$.

- Jeżeli $1 \leq \alpha = \frac{1}{\gamma} < 2$, to nie istnieje wartość oczekiwana $E(X)$, a wariancja $\mathtt{Var}(X) = \infty$.
- Jeżeli $0 < \alpha < 1$, to nie istnieje wartość oczekiwana $E(X)$.

Dystrybuanta F należy do obszaru przyciągania rozkładu Fréchet'a, gdy $1 - F(x) = x^{-\alpha} L_F(x)$, gdzie $\alpha = \frac{1}{\gamma}$, $\gamma > 0$, a $L_F(x)$ jest funkcją wolno zmieniającą się w nieskończoności. Zatem F ma **gruby ogon**.

Niektórzy autorzy przyjmują, że dystrybuanta F ma gruby ogon, gdy $0 < \alpha < 2$.

- Jeżeli $1 \leq \alpha = \frac{1}{\gamma} < 2$, to nie istnieje wartość oczekiwana $E(X)$, a wariancja $\mathtt{Var}(X) = \infty$.
- Jeżeli $0 < \alpha < 1$, to nie istnieje wartość oczekiwana $E(X)$.
- Momenty $E\left(X_+^{\alpha+\delta}\right)$, $\delta > 0$, są nieskończone.

Obszar przyciągania rozkładu Fréchet'a cd.

Przykłady rozkładów należących do obszaru przyciągania rozkładu Fréchet'a:

Przykłady rozkładów należących do obszaru przyciągania rozkładu Fréchet'a:

- **Pareto** $\mathtt{Pa}\,(\alpha)$: $F\,(x) = 1 - x^{-\alpha}, \ x > 1, \alpha > 0$; EVI: $\gamma = \frac{1}{\alpha}$;

Przykłady rozkładów należących do obszaru przyciągania rozkładu Fréchet'a:

- **Pareto** $\mathtt{Pa}\,(\alpha)$: $F\,(x) = 1 - x^{-\alpha},\ x > 1, \alpha > 0$; EVI: $\gamma = \frac{1}{\alpha}$;
- **Uogólniony Pareto** $\mathtt{GP}\,(\sigma, \gamma)$: $F\,(x) = 1 - \left(1 + \gamma\frac{x}{\sigma}\right)^{-\frac{1}{\gamma}},\ x > 0, \sigma, \gamma > 0$; EVI: γ;

Przykłady rozkładów należących do obszaru przyciągania rozkładu Fréchet'a:

- **Pareto** $\texttt{Pa}(\alpha)$: $F(x) = 1 - x^{-\alpha}, \; x > 1, \alpha > 0$; EVI: $\gamma = \frac{1}{\alpha}$;
- **Uogólniony Pareto** $\texttt{GP}(\sigma, \gamma)$: $F(x) = 1 - \left(1 + \gamma \frac{x}{\sigma}\right)^{-\frac{1}{\gamma}}, \; x > 0, \sigma, \gamma > 0$; EVI: γ;
- **Burr** $\texttt{Burr}(\eta, \tau, \lambda)$: $F(x) = 1 - \left(\frac{\eta}{\eta + x^{\tau}}\right)^{\lambda}, \; x > 0, \eta, \tau, \lambda > 0$; EVI: $\gamma = \frac{1}{\lambda\tau}$;

Przykłady rozkładów należących do obszaru przyciągania rozkładu Fréchet'a:

- **Pareto** $\mathtt{Pa}(\alpha)$: $F(x) = 1 - x^{-\alpha}$, $x > 1, \alpha > 0$; EVI: $\gamma = \frac{1}{\alpha}$;
- **Uogólniony Pareto** $\mathtt{GP}(\sigma, \gamma)$:
  $F(x) = 1 - \left(1 + \gamma \frac{x}{\sigma}\right)^{-\frac{1}{\gamma}}$, $x > 0, \sigma, \gamma > 0$; EVI: γ;
- **Burr** $\mathtt{Burr}(\eta, \tau, \lambda)$: $F(x) = 1 - \left(\frac{\eta}{\eta + x^{\tau}}\right)^{\lambda}$, $x > 0, \eta, \tau, \lambda > 0$;
  EVI: $\gamma = \frac{1}{\lambda\tau}$;
- **Fréchet** $\mathtt{Fréchet}(\alpha)$: $F(x) = \exp\left(-x^{-\alpha}\right)$, $x > 0, \alpha > 0$;
  EVI: $\gamma = \frac{1}{\alpha}$;

Przykłady rozkładów należących do obszaru przyciągania rozkładu Fréchet'a:

- **Pareto** $\mathtt{Pa}(\alpha)$: $F(x) = 1 - x^{-\alpha}, \ x > 1, \alpha > 0$; EVI: $\gamma = \frac{1}{\alpha}$;
- **Uogólniony Pareto** $\mathtt{GP}(\sigma, \gamma)$:
  $F(x) = 1 - \left(1 + \gamma\frac{x}{\sigma}\right)^{-\frac{1}{\gamma}}, \ x > 0, \sigma, \gamma > 0$; EVI: γ;
- **Burr** $\mathtt{Burr}(\eta, \tau, \lambda)$: $F(x) = 1 - \left(\frac{\eta}{\eta + x^{\tau}}\right)^{\lambda}, \ x > 0, \eta, \tau, \lambda > 0$;
  EVI: $\gamma = \frac{1}{\lambda\tau}$;
- **Fréchet** $\mathtt{Fréchet}(\alpha)$: $F(x) = \exp\left(-x^{-\alpha}\right), \ x > 0, \alpha > 0$;
  EVI: $\gamma = \frac{1}{\alpha}$;
- **Student-t** z ν stopniami swobody: EVI: $\gamma = \frac{1}{\nu}$;

Dalsze przykłady:

Dalsze przykłady:

- **Stabilne z indeksem** $\alpha < 2$: EVI: $\gamma = \frac{1}{\alpha}$;

Dalsze przykłady:

- **Stabilne z indeksem** $\alpha < 2$: EVI: $\gamma = \frac{1}{\alpha}$;
- **Cauchy**: $F(x) = \frac{1}{2} + \frac{1}{\pi} \operatorname{arc\,tg} x, \ x \in \mathbf{R}$: EVI: $\gamma = 1$;

Obszar przyciągania rozkładu Fréchet'a cd.

Dalsze przykłady:

- **Stabilne z indeksem** $\alpha < 2$: EVI: $\gamma = \frac{1}{\alpha}$;
- **Cauchy**: $F(x) = \frac{1}{2} + \frac{1}{\pi} \operatorname{arc\,tg} x,\ x \in \mathbf{R}$: EVI: $\gamma = 1$;
- **Log-Gamma** (α, λ): $F(x) = \int_1^x \frac{\lambda^\alpha}{\Gamma(\alpha)} (\ln t)^{\alpha-1}\, t^{-\lambda-1} dt$: EVI: $\gamma = \frac{1}{\lambda}$.

Następujące rozkłady należą do obszary przyciągania rozkładu Weibulla $EV_\gamma(x)$ z $\gamma < 0$:

Następujące rozkłady należą do obszary przyciągania rozkładu Weibulla $EV_{\gamma}(x)$ z $\gamma < 0$:

- **Jednostajny** $U(0, 1)$: $F(x) = x, \; 0 < x < 1$: EVI: $\gamma = -1$;

Następujące rozkłady należą do obszary przyciągania rozkładu Weibulla $EV_\gamma(x)$ z $\gamma < 0$:

- **Jednostajny** $U(0,1)$: $F(x) = x, \; 0 < x < 1$: EVI: $\gamma = -1$;
- **Beta** `Beta` (a,b): $F(x) = \int_0^x \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{a-1}(1-u)^{b-1}\,du$, $x < 1, \; a, b > 0$; EVI: $\gamma = \frac{-1}{b}$;

Następujące rozkłady należą do obszary przyciągania rozkładu Weibulla $EV_\gamma(x)$ z $\gamma < 0$:

- **Jednostajny** $U(0,1)$: $F(x) = x,\ 0 < x < 1$: EVI: $\gamma = -1$;
- **Beta** `Beta` (a, b): $F(x) = \int_0^x \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{a-1} (1-u)^{b-1}\, du$, $x < 1,\ a, b > 0$; EVI: $\gamma = \frac{-1}{b}$;
- **Odwrócony Burr** `Reversed Burr` : (β, τ, λ): $F(x) = 1 - \left(\frac{\beta}{\beta + (-x)^{-\tau}}\right)^{\lambda}$, $x < 0,\ \beta, \tau, \lambda > 0$: EVI: $\gamma = \frac{-1}{\lambda\tau}$;

Następujące rozkłady należą do obszary przyciągania rozkładu Weibulla $EV_\gamma(x)$ z $\gamma < 0$:

- **Jednostajny** $U(0,1)$: $F(x) = x, \ 0 < x < 1$: EVI: $\gamma = -1$;
- **Beta** `Beta` (a,b): $F(x) = \int_0^x \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{a-1}(1-u)^{b-1}\, du$, $x < 1, \ a, b > 0$; EVI: $\gamma = \frac{-1}{b}$;
- **Odwrócony Burr** `Reversed Burr` : (β, τ, λ): $F(x) = 1 - \left(\frac{\beta}{\beta + (-x)^{-\tau}}\right)^{\lambda}$, $x < 0, \ \beta, \tau, \lambda > 0$: EVI: $\gamma = \frac{-1}{\lambda\tau}$;
- **Weibull dla maksimów**: $F(x) = 1 - \exp\left(-x^{-\alpha}\right), \ X > 0$, $\alpha > 0$: EVI: $\gamma = \frac{-1}{\alpha}$.

Następujące rozkłady należą do obszary przyciągania rozkładu Gumbela $EV_0(x)$:

Następujące rozkłady należą do obszary przyciągania rozkładu Gumbela $EV_0(x)$:

- **Wykładniczy** `EXP` (1): $F(x) = 1 - \exp(-x),\ x > 0$;

Następujące rozkłady należą do obszary przyciągania rozkładu Gumbela $EV_0(x)$:

- **Wykładniczy** $\mathtt{EXP}(1)$: $F(x) = 1 - \exp(-x)$, $x > 0$;
- **Weibull dla minimów**: $F(x) = 1 - \exp(-\lambda x^{\tau})$, $x > 0$, $\lambda, \tau > 0$;

Następujące rozkłady należą do obszary przyciągania rozkładu Gumbela $EV_0(x)$:

- **Wykładniczy** $\mathtt{EXP}(1)$: $F(x) = 1 - \exp(-x)$, $x > 0$;
- **Weibull dla minimów**: $F(x) = 1 - \exp(-\lambda x^{\tau})$, $x > 0$, $\lambda, \tau > 0$;
- **Logistyczny**: $F(x) = 1 - \frac{1}{1+\exp(x)}$, $x \in \mathbf{R}$;

Obszar przyciągania rozkładu Gumbela

Następujące rozkłady należą do obszary przyciągania rozkładu Gumbela $EV_0\left(x\right)$:

- **Wykładniczy** $\mathtt{EXP}\left(1\right)$: $F\left(x\right) = 1 - \exp\left(-x\right),\ x > 0;$
- **Weibull dla minimów**: $F\left(x\right) = 1 - \exp\left(-\lambda x^{\tau}\right),$ $x > 0,\ \ \lambda, \tau > 0;$
- **Logistyczny**: $F\left(x\right) = 1 - \frac{1}{1+\exp(x)},\ x \in \mathbf{R};$
- **Gumbel**: $\Lambda\left(x\right) = \exp\left(-\exp\left(-x\right)\right),\ x \in \mathbf{R};$

Następujące rozkłady należą do obszary przyciągania rozkładu Gumbela $EV_0(x)$:

- **Wykładniczy** $\mathtt{EXP}(1)$: $F(x) = 1 - \exp(-x)$, $x > 0$;
- **Weibull dla minimów**: $F(x) = 1 - \exp(-\lambda x^{\tau})$, $x > 0$, $\lambda, \tau > 0$;
- **Logistyczny**: $F(x) = 1 - \frac{1}{1+\exp(x)}$, $x \in \mathbf{R}$;
- **Gumbel**: $\Lambda(x) = \exp(-\exp(-x))$, $x \in \mathbf{R}$;
- **Normalny**: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-t^2\right) dt$, $x \in \mathbf{R}$;

Następujące rozkłady należą do obszary przyciągania rozkładu Gumbela $EV_0(x)$:

- **Wykładniczy** $\mathtt{EXP}(1)$: $F(x) = 1 - \exp(-x)$, $x > 0$;
- **Weibull dla minimów**: $F(x) = 1 - \exp(-\lambda x^{\tau})$, $x > 0$, $\lambda, \tau > 0$;
- **Logistyczny**: $F(x) = 1 - \frac{1}{1+\exp(x)}$, $x \in \mathbf{R}$;
- **Gumbel**: $\Lambda(x) = \exp(-\exp(-x))$, $x \in \mathbf{R}$;
- **Normalny**: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-t^2\right) dt$, $x \in \mathbf{R}$;
- **Log-Normalny**: zmienna losowa X ma rozkład log-normalny, jeżeli $\ln X$ jest zmienną losową o rozkładzie normalnym;

Dalsze przykłady:

Dalsze przykłady:

- **Gamma** $\mathtt{Gamma}\,(\alpha, \beta)$: $F(x) = \frac{\alpha^{\beta}}{\Gamma(\alpha)} \int_0^x (\ln t)^{\beta-1}\, t^{-\alpha-1} dt$, $x > 0$;

Dalsze przykłady:

- **Gamma** $\mathtt{Gamma}\,(\alpha,\beta)$: $F\,(x) = \frac{\alpha^\beta}{\Gamma(\alpha)} \int_0^x (\ln t)^{\beta-1}\, t^{-\alpha-1} dt$, $x > 0$;
- **Fréchet dla minimów**: $\Phi^*_\alpha\,(x) = 1 - \Phi_\alpha\,(-x) = 1 - \exp\left(-\,(-x)^{-\alpha}\right)$, $x < 0$, $\alpha > 0$.

W zastosowaniach teorii wartości ekstremalnych szeroko wykorzystywana jest tzw. metoda POT [Peaks-Over-Threshold]. Niech X będzie zmienną losową o dystrybuancie F.

W zastosowaniach teorii wartości ekstremalnych szeroko wykorzystywana jest tzw. metoda POT [Peaks-Over-Threshold]. Niech X będzie zmienną losową o dystrybuancie F.

- Warunkowy rozkład wielkości przewyższenia poziomu u:

$$F_u(y) = P(X - u \leq y \,|X > u) = \frac{F(u+y) - F(u)}{1 - F(u)}.$$

W zastosowaniach teorii wartości ekstremalnych szeroko wykorzystywana jest tzw. metoda POT [Peaks-Over-Threshold]. Niech X będzie zmienną losową o dystrybuancie F.

- Warunkowy rozkład wielkości przewyższenia poziomu u:

$$F_u(y) = P(X - u \leq y \,|X > u) = \frac{F(u+y) - F(u)}{1 - F(u)}.$$

- **Uogólniony rozkład Pareto**:

$$\begin{aligned} H_\gamma(x) &= 1 - (1 + \gamma x)^{-1/\gamma},\ 1 + \gamma x,\ x > 0,\ \text{gdy } \gamma \neq 0, \\ &= 1 - \exp(-x),\ x > 0,\ \text{gdy } \gamma = 0. \end{aligned}$$

Uogólniony rozkład Pareto można zdefiniować ogólniej wprowadzając parametr położenia $u \in \mathbf{R}$ i skali $\sigma > 0$. Dla wartości $x > u$, definiujemy

$$H_\gamma\left(x; u, \sigma\right) = H_\gamma\left(\frac{x-u}{\sigma}\right).$$

i

$$H_\gamma\left(x; \sigma\right) = H_\gamma\left(\frac{x}{\sigma}\right).$$

Uogólniony rozkład Pareto można zdefiniować ogólniej wprowadzając parametr położenia $u \in \mathbf{R}$ i skali $\sigma > 0$. Dla wartości $x > u$, definiujemy

$$H_\gamma(x; u, \sigma) = H_\gamma\left(\frac{x-u}{\sigma}\right).$$

i

$$H_\gamma(x; \sigma) = H_\gamma\left(\frac{x}{\sigma}\right).$$

- Pickands (1975) znalazł związek pomiędzy uogólnionymi rozkładami ekstremalnymi a uogólnionymi rozkładami Pareto.

Uogólniony rozkład Pareto można zdefiniować ogólniej wprowadzając parametr położenia $u \in \mathbf{R}$ i skali $\sigma > 0$.
Dla wartości $x > u$, definiujemy

$$H_\gamma\left(x; u, \sigma\right) = H_\gamma\left(\frac{x-u}{\sigma}\right).$$

i

$$H_\gamma\left(x; \sigma\right) = H_\gamma\left(\frac{x}{\sigma}\right).$$

- Pickands (1975) znalazł związek pomiędzy uogólnionymi rozkładami ekstremalnymi a uogólnionymi rozkładami Pareto.
- J. Pickands, Statistical inference using extreme order statistics, *Annals of Statistics* **3**, (1975), 119-131.

Podamy teraz wynik Pickandsa. Niektórzy autorzy wymieniają także nazwiska Balkemy i de Haana (1974).

Podamy teraz wynik Pickandsa. Niektórzy autorzy wymieniają także nazwiska Balkemy i de Haana (1974).

- $$F \in D\left(EV_{\gamma}\right), \gamma \in \mathbf{R} \;\Leftrightarrow\; \lim_{u \to x^F} \sup_{0<y<x^F-u} \left|F_u\left(y\right) - H_{\gamma}\left(y; \sigma_u\right)\right|,$$

  gdzie $\sigma = \sigma_u$ jest parametrem skali zależnym od poziomu u.

Podamy teraz wynik Pickandsa. Niektórzy autorzy wymieniają także nazwiska Balkemy i de Haana (1974).

- $$F \in D\left(EV_\gamma\right), \gamma \in \mathbf{R} \;\Leftrightarrow\; \lim_{u \to x^F} \sup_{0<y<x^F-u} \left|F_u\left(y\right) - H_\gamma\left(y; \sigma_u\right)\right|,$$

  gdzie $\sigma = \sigma_u$ jest parametrem skali zależnym od poziomu u.
- Z powyższego otrzymujemy przybliżenie:

  $$F_u\left(y\right) = P\left(X - u \leq y \,|X > u\right) \approx H_\gamma\left(y; \sigma_u\right),$$

  gdzie $y \in \left[0, x^F - u\right]$, gdy $\gamma \geq 0$, oraz $y \in \left[0, -\frac{\sigma_u}{\gamma}\right]$, gdy $\gamma < 0$.

Celem jest statystyczne badanie nieznanej dystrybuanty F zmiennej losowej X.

Celem jest statystyczne badanie nieznanej dystrybuanty F zmiennej losowej X.

- Mamy

$$F_u(y) = P(X - u \leq y | X > u) = \frac{F(u+y) - F(u)}{1 - F(u)},$$

Celem jest statystyczne badanie nieznanej dystrybuanty F zmiennej losowej X.

- Mamy
$$F_u(y) = P(X - u \leq y \,|X > u) = \frac{F(u+y) - F(u)}{1 - F(u)},$$
- zatem, przyjmując $x = u + y$,
$$1 - F(x) = (1 - F(u))(1 - F_u(x-u)),$$

Uogólniony rozkład Pareto; zastosowania w statystyce

Celem jest statystyczne badanie nieznanej dystrybuanty F zmiennej losowej X.

- Mamy
$$F_u(y) = P(X - u \leq y \,|X > u) = \frac{F(u+y) - F(u)}{1 - F(u)},$$
- zatem, przyjmując $x = u + y$,
$$1 - F(x) = (1 - F(u))(1 - F_u(x-u)),$$
- następnie wykorzystując aproksymację
$$F_u(y) \approx H_\gamma(y; \sigma_u),$$

Celem jest statystyczne badanie nieznanej dystrybuanty F zmiennej losowej X.

- Mamy

$$F_u(y) = P(X - u \leq y \,|X > u) = \frac{F(u+y) - F(u)}{1 - F(u)},$$

- zatem, przyjmując $x = u + y$,

$$1 - F(x) = (1 - F(u))(1 - F_u(x-u)),$$

- następnie wykorzystując aproksymację

$$F_u(y) \approx H_\gamma(y; \sigma_u),$$

- otrzymujemy

$$\overline{F}(x) \approx \overline{F}(u)\left(1 - H_\gamma\left(\frac{x-u}{\sigma_u}\right)\right).$$

Podamy przykłady zastosowań

Podamy przykłady zastosowań

- **Estymacja prawdopodobieństwa przewyższenia wysokiego poziomu** u

$$\overline{F}(x) \approx \overline{F}(u)\left(1+\gamma\frac{x-u}{\sigma_u}\right)^{-\frac{1}{\gamma}}.$$

Podamy przykłady zastosowań

- **Estymacja prawdopodobieństwa przewyższenia wysokiego poziomu** u

$$\overline{F}(x) \approx \overline{F}(u)\left(1+\gamma\frac{x-u}{\sigma_u}\right)^{-\frac{1}{\gamma}}.$$

- Estymujemy $\overline{F}(u)$ przez częstość przekroczeń $\frac{N_u}{n}$, otrzymujemy

$$\widehat{\overline{F}}(x) = \frac{N_u}{n}\left(1+\widehat{\gamma}\frac{x-u}{\widehat{\sigma}_u}\right)^{-\frac{1}{\widehat{\gamma}}}.$$

Dalsze przykłady zastosowań:

Dalsze przykłady zastosowań:

- **Estymacja wysokiego kwantyla dystrybnuanty** F; $U\left(\frac{1}{p}\right) = F^{\leftarrow}(1-p)$

$$\widehat{U}\left(\frac{1}{p}\right) = u + \frac{\widehat{\sigma}_u}{\widehat{\gamma}}\left(\left(\frac{np}{N_u}\right)^{-\widehat{\gamma}} - 1\right).$$

Dalsze przykłady zastosowań:

- **Estymacja wysokiego kwantyla dystrybnuanty** F; $U\left(\frac{1}{p}\right) = F^{\leftarrow}(1-p)$

$$\widehat{U}\left(\frac{1}{p}\right) = u + \frac{\widehat{\sigma}_u}{\widehat{\gamma}}\left(\left(\frac{np}{N_u}\right)^{-\widehat{\gamma}} - 1\right).$$

- **Estymacja prawego kresu suportu dystrybuanty** F dla $\gamma < 0$

$$\widehat{x}^F = \widehat{U}(\infty) = u - \frac{\widehat{\sigma}_u}{\widehat{\gamma}}.$$

Uogólniony rozkład Pareto; estymacja paramerów

W ostatnich dekadach zaobserwowano bezprecedensowy wzrost rozmiarów zbiorów danych dostępnych w rozlicznych zastosowaniach, takich, jak finanse, ubezpieczenia, nauki komputerowe, komunikacja itp. Między innymi, estymacja parametrów uogólnionego rozkładu Pareto przyciągała dużą uwagę. Istnieje wiele metod takiej estymacji i wciąż powstają nowe. Dla przykładu zwróćmy uwagę na pracę:

W ostatnich dekadach zaobserwowano bezprecedensowy wzrost rozmiarów zbiorów danych dostępnych w rozlicznych zastosowaniach, takich, jak finanse, ubezpieczenia, nauki komputerowe, komunikacja itp. Między innymi, estymacja parametrów uogólnionego rozkładu Pareto przyciągała dużą uwagę. Istnieje wiele metod takiej estymacji i wciąż powstają nowe. Dla przykładu zwróćmy uwagę na pracę:

- Myung Hyun Park, Joseph H.T. Kim, Estimating extreme tail risk measures with generalized Pareto distribution, *Computational Statistics and Data Analysis*, **98**, (2016), 91-104.

Uogólniony rozkład Pareto; estymacja paramerów

W ostatnich dekadach zaobserwowano bezprecedensowy wzrost rozmiarów zbiorów danych dostępnych w rozlicznych zastosowaniach, takich, jak finanse, ubezpieczenia, nauki komputerowe, komunikacja itp. Między innymi, estymacja parametrów uogólnionego rozkładu Pareto przyciągała dużą uwagę. Istnieje wiele metod takiej estymacji i wciąż powstają nowe. Dla przykładu zwróćmy uwagę na pracę:

- Myung Hyun Park, Joseph H.T. Kim, Estimating extreme tail risk measures with generalized Pareto distribution, *Computational Statistics and Data Analysis*, **98**, (2016), 91-104.
- Autorzy znajdują estymator $(\widehat{\gamma}, \widehat{\sigma})$ oparty o nieliniową ważoną metodę najmniejszych kwadratów, minimalizującą sumę kwadratów różnic pomiędzy dystrybuantą empiryczną a teoretyczną dystrybuantą uogólnionego rozkładu Pareto.

Największe zainteresowanie budzi estymacja indeksu ekstremalnego γ. Poświęcono temu zagadnieniu dziesiątki prac. Na przykład, jedną z ostatnich jest praca:

Największe zainteresowanie budzi estymacja indeksu ekstremalnego γ. Poświęcono temu zagadnieniu dziesiątki prac. Na przykład, jedną z ostatnich jest praca:

- M. Ivette Gomes, Lígia Henriques-Rodrigues, Competitive estimation of the extreme value index, *Statistics and Probability Letters*, (2016), Article in Press.

Największe zainteresowanie budzi estymacja indeksu ekstremalnego γ. Poświęcono temu zagadnieniu dziesiątki prac. Na przykład, jedną z ostatnich jest praca:

- M. Ivette Gomes, Lígia Henriques-Rodrigues, Competitive estimation of the extreme value index, *Statistics and Probability Letters*, (2016), Article in Press.
- Estymatory opierają się o statystyki pozycyjne.

Największe zainteresowanie budzi estymacja indeksu ekstremalnego γ. Poświęcono temu zagadnieniu dziesiątki prac. Na przykład, jedną z ostatnich jest praca:

- M. Ivette Gomes, Lígia Henriques-Rodrigues, Competitive estimation of the extreme value index, *Statistics and Probability Letters*, (2016), Article in Press.
- Estymatory opierają się o statystyki pozycyjne.
- Niech $X_1, X_2, \ldots, X_n$ będzie próbą losową o dystrybuancie F. Oznaczny przez

$$X_{1:n} = \min\{X_1, X_2, \ldots, X_n\} \leq X_{2:n} \leq \cdots \leq X_{n-1:n} \leq X_{n:n} = \max\{X_1, X_2, \ldots, X_n\}$$

statystyki pozycyjne w ciągu $X_1, X_2, \ldots, X_n$, gdzie $n \geq 1$.

W przypadku $\gamma > 0$, najbardziej popularnym estymatorem indeksu ekstremalnego γ jest estymator Hilla.

Estymacja indeksu ekstremalnego γ cd.

W przypadku $\gamma > 0$, najbardziej popularnym estymatorem indeksu ekstremalnego γ jest estymator Hilla.

- B.M. Hill, A simple general approach to inference about the tail of a distribution, *Ann. Statist.* **13** (1975), 1163-1174.

Estymacja indeksu ekstremalnego γ cd.

W przypadku $\gamma > 0$, najbardziej popularnym estymatorem indeksu ekstremalnego γ jest estymator Hilla.

- B.M. Hill, A simple general approach to inference about the tail of a distribution, *Ann. Statist.* **13** (1975), 1163-1174.
- Estymator Hilla oparty jest o

$$V_{ik} = \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \leq i \leq k < n,$$

Estymacja indeksu ekstremalnego γ cd.

W przypadku $\gamma > 0$, najbardziej popularnym estymatorem indeksu ekstremalnego γ jest estymator Hilla.

- B.M. Hill, A simple general approach to inference about the tail of a distribution, *Ann. Statist.* **13** (1975), 1163-1174.
- Estymator Hilla oparty jest o

$$V_{ik} = \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \leq i \leq k < n,$$

- i jest geometryczną średnią (średnią rzędu 0) zmiennych losowych

$$U_{ik} = \frac{X_{n-i+1:n}}{X_{n-k:n}}, \quad 1 \leq i \leq k < n,$$

Estymacja indeksu ekstremalnego γ cd.

W przypadku $\gamma > 0$, najbardziej popularnym estymatorem indeksu ekstremalnego γ jest estymator Hilla.

- B.M. Hill, A simple general approach to inference about the tail of a distribution, *Ann. Statist.* **13** (1975), 1163-1174.
- Estymator Hilla oparty jest o

$$V_{ik} = \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \leq i \leq k < n,$$

- i jest geometryczną średnią (średnią rzędu 0) zmiennych losowych

$$U_{ik} = \frac{X_{n-i+1:n}}{X_{n-k:n}}, \quad 1 \leq i \leq k < n,$$

- to znaczy

$$\widehat{\gamma}_k^H = \frac{1}{k}\sum_{i=1}^{k} V_{ik} = \ln\left(\prod_{i=1}^{k} \frac{X_{n-i+1:n}}{X_{n-k:n}}\right)^{\frac{1}{k}}.$$

We wspomnianej wcześniej pracy Gomes i Henriques-Rodrigues badane są estymatory indeksu ekstremalnego γ oparte o średnie rzędu p:

We wspomnianej wcześniej pracy Gomes i Henriques-Rodrigues badane są estymatory indeksu ekstremalnego γ oparte o średnie rzędu p:

- Dla $p \in \mathbf{R}$

$$M_p(k) = \left(\frac{1}{k} \sum_{i=1}^{k} U_{ik}^{p} \right)^{\frac{1}{k}}, \quad p \neq 0,$$

i

$$M_p(k) = \left(\prod_{i=1}^{k} U_{ik} \right)^{\frac{1}{k}}, \quad p = 0.$$

We wspomnianej wcześniej pracy Gomes i Henriques-Rodrigues badane są estymatory indeksu ekstremalnego γ oparte o średnie rzędu p:

- Dla $p \in \mathbf{R}$

$$M_p(k) = \left(\frac{1}{k} \sum_{i=1}^{k} U_{ik}^{p} \right)^{\frac{1}{k}}, \quad p \neq 0,$$

i

$$M_p(k) = \left(\prod_{i=1}^{k} U_{ik} \right)^{\frac{1}{k}}, \quad p = 0.$$

- Estymatory są postaci

$$\widehat{\gamma}_k^{H_p} = \left(1 - M_p^{-p}(k)\right)/p, \quad p < 1/\gamma,$$

i

$$\widehat{\gamma}_k^{H_p} = \ln M_0(k) = \widehat{\gamma}_k^{H}, \quad p = 0.$$

Dany jest ciąg zmiennych losowych $X_1, X_2, \ldots, X_n$.

Dany jest ciąg zmiennych losowych $X_1, X_2, \ldots, X_n$.

- **Definicja.** Statystyką pozycyjną $X_{k:n}$, $k = 1, 2, \ldots, n$, nazywamy zmienną losową będącą funkcją wektora losowego $(X_1, X_2, \ldots, X_n)$ określoną w następujący sposób: Dla każdego zdarzenia elementarnego ω ciąg realizacji $X_1(\omega) = x_1, X_2(\omega) = x_2, \ldots, X_n(\omega) = x_n$ porządkujemy wzrastająco i otrzymujemy $z_1 \leq z_2 \leq \cdots \leq z_n$. W tym ciągu z_k jest realizacją zmiennej losowej $X_{k:n}$, tzn. $X_{k:n}(\omega) = z_k$.

Dany jest ciąg zmiennych losowych $X_1, X_2, \ldots, X_n$.

- **Definicja.** Statystyką pozycyjną $X_{k:n}$, $k = 1, 2, \ldots, n$, nazywamy zmienną losową będącą funkcją wektora losowego $(X_1, X_2, \ldots, X_n)$ określoną w następujący sposób: Dla każdego zdarzenia elementarnego ω ciąg realizacji $X_1(\omega) = x_1, X_2(\omega) = x_2, \ldots, X_n(\omega) = x_n$ porządkujemy wzrastająco i otrzymujemy $z_1 \leq z_2 \leq \cdots \leq z_n$. W tym ciągu z_k jest realizacją zmiennej losowej $X_{k:n}$, tzn. $X_{k:n}(\omega) = z_k$.
- **Definicja.** Zmienną losową $X_{n-k+1:n}$, gdzie k jest ustaloną liczbą naturalną, $k = 1, 2, \ldots, n$, nazywamy k-tą estremalną statystyką pozycyjną.

Dany jest ciąg zmiennych losowych $X_1, X_2, \ldots, X_n$.

- **Definicja.** Statystyką pozycyjną $X_{k:n}$, $k = 1, 2, \ldots, n$, nazywamy zmienną losową będącą funkcją wektora losowego $(X_1, X_2, \ldots, X_n)$ określoną w następujący sposób: Dla każdego zdarzenia elementarnego ω ciąg realizacji $X_1(\omega) = x_1, X_2(\omega) = x_2, \ldots, X_n(\omega) = x_n$ porządkujemy wzrastająco i otrzymujemy $z_1 \leq z_2 \leq \cdots \leq z_n$. W tym ciągu z_k jest realizacją zmiennej losowej $X_{k:n}$, tzn. $X_{k:n}(\omega) = z_k$.
- **Definicja.** Zmienną losową $X_{n-k+1:n}$, gdzie k jest ustaloną liczbą naturalną, $k = 1, 2, \ldots, n$, nazywamy k-tą estremalną statystyką pozycyjną.
- Oczywiście, $X_{n:n} = \max(X_1, X_2, \ldots, X_n)$, $X_{1:n} = \min(X_1, X_2, \ldots, X_n)$.

Niech teraz ciąg zmiennych losowych $X_1, X_2, \ldots, X_n$ będzie ciągiem niezależnych zmiennych losowych o jednakowej dystrybuancie F.

Niech teraz ciąg zmiennych losowych $X_1, X_2, \ldots, X_n$ będzie ciągiem niezależnych zmiennych losowych o jednakowej dystrybuancie F.

- Dla dowolnego $x \in \mathbf{R}$, zdefiniujmy zdarzenia

$$E_i(x) = \{X_i > x\}, \; i = 1, 2, \ldots, n,$$

i oznaczny indykatory tych zdarzeń przez

$$I_{E_i(x)}, \; i = 1, 2, \ldots, n.$$

Niech teraz ciąg zmiennych losowych $X_1, X_2, \ldots, X_n$ będzie ciągiem niezależnych zmiennych losowych o jednakowej dystrybuancie F.

- Dla dowolnego $x \in \mathbf{R}$, zdefiniujmy zdarzenia

$$E_i(x) = \{X_i > x\}, \ \ i = 1, 2, \ldots, n,$$

i oznaczny indykatory tych zdarzeń przez

$$I_{E_i(x)}, \ \ i = 1, 2, \ldots, n.$$

- Oznaczmy także

$$I_n(x) = \sum_{i=1}^{n} I_{E_i(x)}.$$

Niech teraz ciąg zmiennych losowych $X_1, X_2, \ldots, X_n$ będzie ciągiem niezależnych zmiennych losowych o jednakowej dystrybuancie F.

- Dla dowolnego $x \in \mathbf{R}$, zdefiniujmy zdarzenia

$$E_i(x) = \{X_i > x\}, \; i = 1, 2, \ldots, n,$$

  i oznaczny indykatory tych zdarzeń przez

$$I_{E_i(x)}, \; i = 1, 2, \ldots, n.$$

- Oznaczmy także

$$I_n(x) = \sum_{i=1}^{n} I_{E_i(x)}.$$

- Oczywiście zmienna losowa $I_n(x)$ ma rozkład dwumianowy z parametrami

$$n, \; p(x) = 1 - F(x).$$

Zauważmy, że

Zauważmy, że

- $$\{X_{n-k+1:n} \leq x\} = \{I_n(x) < k\}.$$

Rozkład k-tej ekstremalnej statystyki pozycyjnej

Zauważmy, że

- $$\{X_{n-k+1:n} \leq x\} = \{I_n(x) < k\}.$$

- Zatem dystrybuanta k-tej ekstremalnej statystyki pozycyjnej $X_{n-k+1:n}$ jest postaci

$$F_{n-k+1:n}(x) = P(X_{n-k+1:n} \leq x) = P(I_n(x) < k),$$

Zauważmy, że

- $$\{X_{n-k+1:n} \leq x\} = \{I_n(x) < k\}.$$

- Zatem dystrybuanta k-tej ekstremalnej statystyki pozycyjnej $X_{n-k+1:n}$ jest postaci
$$F_{n-k+1:n}(x) = P(X_{n-k+1:n} \leq x) = P(I_n(x) < k),$$

- stąd
$$F_{n-k+1:n}(x) = \sum_{r=0}^{k-1} \binom{n}{r} (1 - F(x))^r F^{n-r}(x),$$

Rozkład k-tej ekstremalnej statystyki pozycyjnej

Zauważmy, że

- $$\{X_{n-k+1:n} \leq x\} = \{I_n(x) < k\}.$$

- Zatem dystrybuanta k-tej ekstremalnej statystyki pozycyjnej $X_{n-k+1:n}$ jest postaci
$$F_{n-k+1:n}(x) = P(X_{n-k+1:n} \leq x) = P(I_n(x) < k),$$

- stąd
$$F_{n-k+1:n}(x) = \sum_{r=0}^{k-1} \binom{n}{r} (1 - F(x))^r F^{n-r}(x),$$

- a także
$$F_{n-k+1:n}(x) = (n-k+1)\binom{n}{k-1} \int_0^{F(x)} u^{n-k}(1-u)^{k-1}\, du.$$

Zbadamy asymptotykę $P\{X_{n-k+1:n} \le u_n\}$, gdy $n \longrightarrow \infty$, gdzie u_n jest pewnym ciągiem liczb rzeczywistych.

Zbadamy asymptotykę $P\{X_{n-k+1:n} \le u_n\}$, gdy $n \longrightarrow \infty$, gdzie u_n jest pewnym ciągiem liczb rzeczywistych.

- W monografii: M.R. Leadbetter, Georg Lindgren, Holger Rootzén, *Extremes and related properties of random sequences and processes*, Springer, New York 1983,
  można znaleźć następujące twierdzenia:

Twierdzenie. Niech X_n będzie i.i.d. ciągiem. Niech $0 \leq \tau \leq \infty$ i zakładamy, że u_n jest ciągiem liczb rzeczywistych takim, że

$$n\left(1 - F\left(u_n\right)\right) \to \tau, \text{ gdy } n \to \infty \tag{3}$$

wtedy, dla $k = 0, 1, 2, \dots,$

$$P\left\{S_n \leq k\right\} \to e^{-\tau} \sum_{s=0}^{k} \frac{\tau^s}{s!}, \text{ gdy } n \to \infty. \tag{4}$$

(prawa strona w powyższej granicy równa się zero, gdy $\tau = \infty$), gdzie

$$S_n = I_n\left(u_n\right).$$

Odwrotnie, jeżeli (4) zachodzi dla pewnego ustalonego k, to zachodzi także (3) (i (4) stąd zachodzi dla wszystkich k).

Asymptotyczny rozkład k-tej ekstremalnej statystyki pozycyjnej

Twierdzenie. Niech X_n będzie i.i.d. ciągiem. Jeżeli u_n jest ciągiem liczb rzeczywistych takim, że dla pewnego τ, $0 \leq \tau \leq \infty$,

$$n(1 - F(u_n)) \to \tau, \text{ gdy } n \to \infty \tag{5}$$

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

$$P\{X_{n-k+1} \leq u_n\} \to e^{-\tau} \sum_{s=0}^{k} \frac{\tau^s}{s!}, \text{ gdy } n \to \infty. \tag{6}$$

Odwrotnie, jeżeli (6) zachodzi dla pewnego ustalonego k, to zachodzi także (5) oraz (6) dla wszystkich k.

Asymptotyczny rozkład k-tej ekstremalnej statystyki pozycyjnej cd.

Twierdzenie. Załóżmy, że

$$P\left\{a_n\left(M_n - b_n\right) \leq x\right\} \xrightarrow{w} G\left(x\right)$$

dla pewnej niezdegenerowanej (a zatem ekstremalnej) dystrybuanty G. Wtedy, dla każdego $k = 1, 2, \ldots,$

$$P\left\{a_n\left(X_{n-k+1:n} - b_n\right) \leq x\right\} \xrightarrow{w} G\left(x\right)\sum_{s=0}^{k}\frac{\left(-\log G\left(x\right)\right)^s}{s!},$$

gdzie $G\left(x\right) > 0$ (i zero, gdzie $G\left(x\right) = 0$).

Asymptotyczny rozkład k-tej ekstremalnej statystyki pozycyjnej cd.

Odwrotnie, jeżeli dla pewnego ustalonego k,

$$P\left\{a_n\left(X_{n-k+1:n}-b_n\right)\leq x\right\}\xrightarrow{w}H\left(x\right)$$

dla pewnej niezdegenerowanej dystrybuanty H, to $H\left(x\right)$ musi być postaci

$$G\left(x\right)\sum_{s=0}^{k}\frac{\left(-\log G\left(x\right)\right)^s}{s!},$$

gdzie

$$P\left\{a_n\left(M_n-b_n\right)\leq x\right\}\xrightarrow{w}G\left(x\right)$$

zachodzi z tymi samymi G, a_n, b_n. Zatem

$$P\left\{a_n\left(X_{n-k+1:n}-b_n\right)\leq x\right\}\xrightarrow{w}G\left(x\right)\sum_{s=0}^{k}\frac{\left(-\log G\left(x\right)\right)^s}{s!},$$

dla wszystkich k.

1. Teoria wartości ekstremalnych. Ciąg niezależnych zmiennych losowych o jednakowym rozkładzie.
2. Teoria wartości rekordowych. Ciąg niezależnych zmiennych losowych o jednakowym rozkładzie.
3. Badania statystyczne w oparciu o wartości rekordowe
4. Zakończenie

Niech $X_1, X_2, \ldots$ będzie ciągiem zmiennych losowych. Oznaczmy przez

$$\begin{aligned} X_{1:n} = & \min\{X_1, X_2, \ldots, X_n\} \leq X_{2:n} \leq \cdots \leq X_{n-1:n} \\ & \leq X_{n:n} = \max\{X_1, X_2, \ldots, X_n\} \end{aligned}$$

statystyki pozycyjne w ciągu $X_1, X_2, \ldots, X_n$, gdzie $n \geq 1$. Niech $k \geq 1$ będzie ustaloną liczbą naturalną.
Podamy definicję k-tej statystyki (wartości) rekordowej.

Definicja k-tych statystyk rekordowych

Niech $X_1, X_2, \ldots$ będzie ciągiem zmiennych losowych. Oznaczmy przez

$$\begin{aligned} X_{1:n} = & \min\{X_1, X_2, \ldots, X_n\} \leq X_{2:n} \leq \cdots \leq X_{n-1:n} \\ & \leq X_{n:n} = \max\{X_1, X_2, \ldots, X_n\} \end{aligned}$$

statystyki pozycyjne w ciągu $X_1, X_2, \ldots, X_n$, gdzie $n \geq 1$. Niech $k \geq 1$ będzie ustaloną liczbą naturalną.
Podamy definicję k-tej statystyki (wartości) rekordowej.

- W. Dziubdziela, B. Kopociński, Limiting properties of the k-th record values, *Zastosowania Matematyki* **15** (1976) , 187-190.

Niech $X_1, X_2, \ldots$ będzie ciągiem zmiennych losowych. Oznaczmy przez

$$X_{1:n} = \min\{X_1, X_2, \ldots, X_n\} \leq X_{2:n} \leq \cdots \leq X_{n-1:n}$$
$$\leq X_{n:n} = \max\{X_1, X_2, \ldots, X_n\}$$

statystyki pozycyjne w ciągu $X_1, X_2, \ldots, X_n$, gdzie $n \geq 1$. Niech $k \geq 1$ będzie ustaloną liczbą naturalną.
Podamy definicję k-tej statystyki (wartości) rekordowej.

- W. Dziubdziela, B. Kopociński, Limiting properties of the k-th record values, *Zastosowania Matematyki* **15** (1976) , 187-190.
- Rozważmy ciąg wektorów losowych

$$\mathbf{X}_{n-k+1:n} = (X_{n-k+1:n}, X_{n-k+2:n}, \ldots, X_{n:n}), \ \ n \geq k.$$

In 1976, Dziubdziela and Kopociński wprowadzili k-te czasy rekordowe $L(1,k)$, $n \geq k$, jako kolejne momenty, w których ciąg $\{\mathbf{X}_{n-k+1:n}\}$ zmienia wartości. Niech $L(1,k) = k$. Oczywiście, mamy

In 1976, Dziubdziela and Kopociński wprowadzili k-te czasy rekordowe $L(1,k)$, $n \geq k$, jako kolejne momenty, w których ciąg $\{\mathbf{X}_{n-k+1:n}\}$ zmienia wartości. Niech $L(1,k) = k$. Oczywiście, mamy

- $$L(n+1,k) = \min\{j > L(n,k) : X_j > X_{j-k:j-1}\},$$

  gdzie $n \geq 1$.

In 1976, Dziubdziela and Kopociński wprowadzili k-te czasy rekordowe $L(1,k)$, $n \geq k$, jako kolejne momenty, w których ciąg $\{\mathbf{X}_{n-k+1:n}\}$ zmienia wartości. Niech $L(1,k) = k$. Oczywiście, mamy

- $$L(n+1,k) = \min\{j > L(n,k) : X_j > X_{j-k:j-1}\},$$
  gdzie $n \geq 1$.
- k-te wartości rekordowe (k-te statystyki rekordowe, k-rekordy) są zdefiniowane przez
  $$X(n,k) = X_{L(n,k)-k+1:L(n,k)}, \ \ n \geq 1,$$
  gdzie $k \geq 1$ jest stałą liczbą naturalną.

Definicja k-tych statystyk rekordowych cd.

In 1976, Dziubdziela and Kopociński wprowadzili k-te czasy rekordowe $L(1,k)$, $n \geq k$, jako kolejne momenty, w których ciąg $\{\mathbf{X}_{n-k+1:n}\}$ zmienia wartości. Niech $L(1,k) = k$. Oczywiście, mamy

- $$L(n+1,k) = \min\{j > L(n,k) : X_j > X_{j-k:j-1}\},$$

  gdzie $n \geq 1$.
- k-te wartości rekordowe (k-te statystyki rekordowe, k-rekordy) są zdefiniowane przez

  $$X(n,k) = X_{L(n,k)-k+1:L(n,k)}, \;\; n \geq 1,$$

  gdzie $k \geq 1$ jest stałą liczbą naturalną.
- W przypadku $k = 1$, otrzymujemy dobrze znane czasy rekordowe i rekordy.

Przykładowa literatura:

Przykładowa literatura:

- M. Ahsanullah, V. B. Nevzorov, *Records via Probability Theory*, Atlantis Press, Amsterdam 2015.

Przykładowa literatura:

- M. Ahsanullah, V. B. Nevzorov, *Records via Probability Theory*, Atlantis Press, Amsterdam 2015.
- B. C. Arnold, N. Balakrishnan, H. N. Nagaraja, *Records*, Wiley, New York 1998.

Przykładowa literatura:

- M. Ahsanullah, V. B. Nevzorov, *Records via Probability Theory*, Atlantis Press, Amsterdam 2015.
- B. C. Arnold, N. Balakrishnan, H. N. Nagaraja, *Records*, Wiley, New York 1998.
- V. B. Nevzorov, *Records: Mathematical Theory*, Transl. Math. Monographs 194, Amer. Math. Soc., Providence RI, 2001.

Możemy spojrzeć na definicję k-tych wartości rekordowych w inny sposób.

Definicja k-tych statystyk rekordowych, inne spojrzenie

Możemy spojrzeć na definicję k-tych wartości rekordowych w inny sposób.

- Rozważmy niemalejący ciąg k-tych statystyk pozycyjnych

  $$-\infty < X_{1:k} \leq X_{2:k+1} \leq \ldots \leq X_{n-k:n-1} \leq X_{n-k+1:n} \leq \ldots.$$

  Usuwając elementy powtarzające się w tym ciągu, otrzymujemy ściśle rosnący ciąg kolejnych k-tych wartości rekordowych

Definicja k-tych statystyk rekordowych, inne spojrzenie

Możemy spojrzeć na definicję k-tych wartości rekordowych w inny sposób.

- Rozważmy niemalejący ciąg k-tych statystyk pozycyjnych

  $$-\infty < X_{1:k} \leq X_{2:k+1} \leq \ldots \leq X_{n-k:n-1} \leq X_{n-k+1:n} \leq \ldots .$$

  Usuwając elementy powtarzające się w tym ciągu, otrzymujemy ściśle rosnący ciąg kolejnych k-tych wartości rekordowych

- $$-\infty < X_{1:k} < X_{L(2,k)-k+1:L(2,k)} < X_{L(3,k)-k+1:L(3,k)} < \cdots .$$

Definicja k-tych statystyk rekordowych, równoważna definicja

W pracy Dziubdziela and Kopociński (1976) rozważano ciąg $X_1, X_2, \ldots$ niezależnych zmiennych losowych z jednakową ciągłą dystrybuantą. Z założenia ciągłości wynika, że definicja k-tych czasów rekordowych i k-tych rekordowych wartości może być zapisana w postaci

Definicja k-tych statystyk rekordowych, równoważna definicja

W pracy Dziubdziela and Kopociński (1976) rozważano ciąg $X_1, X_2, \ldots$ niezależnych zmiennych losowych z jednakową ciągłą dystrybuantą. Z założenia ciągłości wynika, że definicja k-tych czasów rekordowych i k-tych rekordowych wartości może być zapisana w postaci

- $$\widehat{L}(1,k) = 1,$$
  $$\widehat{L}(n+1,k) = \min\left\{j > \widehat{L}(n,k) : X_{\widehat{L}(n,k):\widehat{L}(n,k)+k-1} < X_{j:j+k-1}\right\},$$
  $$n \geq 1, \quad \text{oraz}$$
  $$\widehat{X}(n,k) = X_{\widehat{L}(n,k):\widehat{L}(n,k)+k-1}, \;\; n \geq 1.$$

Definicja k-tych statystyk rekordowych, rozkład dyskretny

W roku 2005 Dembicka and López-Blázquez zauważyli, że w przypadku dyskretnych rozkładów powyższe definicje k-tych czasów rekordowych i k-tych rekordowych wartości nie są równoważne.

Definicja k-tych statystyk rekordowych, rozkład dyskretny

W roku 2005 Dembicka and López-Blázquez zauważyli, że w przypadku dyskretnych rozkładów powyższe definicje k-tych czasów rekordowych i k-tych rekordowych wartości nie są równoważne.

- A. Dembińska, F. López-Blázquez, k-th records from discrete distributions, *Statistics and Probability Letters* **71** (2005), 203-214.
  Podali oni następujący przykład:

W roku 2005 Dembicka and López-Blázquez zauważyli, że w przypadku dyskretnych rozkładów powyższe definicje k-tych czasów rekordowych i k-tych rekordowych wartości nie są równoważne.

- A. Dembińska, F. López-Blázquez, k-th records from discrete distributions, *Statistics and Probability Letters* **71** (2005), 203-214.
  Podali oni następujący przykład:
- **Przykład.** Niech ciąg $X_1, X_2, \dots$ ma realizację $0, 1, 1, 1, 0, 4, 0, 3, \dots$. Wtedy

$$\{L(n,2), n \geq 1\} = \{2, 3, 6, 8, \dots\}$$

i

$$\left\{\widehat{L}(n,2), n \geq 1\right\} = \{1, 2, 7, \dots\}.$$

Definicja k-tych statystyk rekordowych, rozkład dyskretny cd.

Zatem

Definicja k-tych statystyk rekordowych, rozkład dyskretny cd.

Zatem

- $$\left\{X_{L(n,2)-2+1:L(n,2)}, n \geq 1\right\} = \{0, 1, 1, 3, \ldots\}$$

  i

  $$\left\{X_{\widehat{L}(n,2):\widehat{L}(n,2)+2-1}, n \geq 1\right\} = \{0, 1, 3, \ldots\}$$

Definicja k-tych statystyk rekordowych, rozkład dyskretny cd.

Zatem

- $$\left\{X_{L(n,2)-2+1:L(n,2)}, n \geq 1\right\} = \{0, 1, 1, 3, \dots\}$$

  i

  $$\left\{X_{\widehat{L}(n,2):\widehat{L}(n,2)+2-1}, n \geq 1\right\} = \{0, 1, 3, \dots\}$$

- Wynika stąd, że zbiory o 2-gich rekordów nie są równe.

Kończąc fragment o definicjach, przedstawimy definicję opierającą się o pojęcie rangi.

Definicja k-tych statystyk rekordowych, rangi

Kończąc fragment o definicjach, przedstawimy definicję opierającą się o pojęcie rangi.

- Sekwencyjna ranga I_n to ranga X_n wśród $X_1, X_2, \ldots, X_n$, tzn.

$$X_n = X_{I_n:n}, \ n \geq 1$$

albo

Kończąc fragment o definicjach, przedstawimy definicję opierającą się o pojęcie rangi.

- Sekwencyjna ranga I_n to ranga X_n wśród $X_1, X_2, \ldots, X_n$, tzn.

$$X_n = X_{I_n:n}, \ n \geq 1$$

  albo

- $$I_n = \sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq X_n\}}, \ n \geq 1,$$

  gdzie $\mathbf{1}_A$ oznacza indykator zdarzenia losowego A.

Definicja k-tych statystyk rekordowych, rangi

Kończąc fragment o definicjach, przedstawimy definicję opierającą się o pojęcie rangi.

- Sekwencyjna ranga I_n to ranga X_n wśród $X_1, X_2, \ldots, X_n$, tzn.

$$X_n = X_{I_n:n}, \; n \geq 1$$

  albo

- $$I_n = \sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq X_n\}}, \; n \geq 1,$$

  gdzie $\mathbf{1}_A$ oznacza indykator zdarzenia losowego A.

- Teraz, definiujemy k-te czasy rekordowe przez

$$L(1,k) = k,$$

$$L(n+1,k) = \min\{j > L(n,k) : n-k+1 \leq I_j \leq n\}, \; n \geq 1.$$

Reprezentacja k-tych statystyk rekordowych

Poprzednio oznaczaliśmy przez $X(n, k)$ k-te statystyki rekordowe w ciągu $X_1, X_2, \ldots$ niezależnych zmiennych losowych o jednakowej ciągłej dystrybuancie F.

Reprezentacja k-tych statystyk rekordowych

Poprzednio oznaczaliśmy przez $X(n,k)$ k-te statystyki rekordowe w ciągu $X_1, X_2, \ldots$ niezależnych zmiennych losowych o jednakowej ciągłej dystrybuancie F.

- W szczególnym przypadku, przez $Z(n,k)$ oznaczyny k-te statystyki rekordowe w ciągu $Z_1, Z_2, \ldots$ niezależnych zmiennych losowych o jednakowej standardowej wykładniczej dystrybuancie

Reprezentacja k-tych statystyk rekordowych

Poprzednio oznaczaliśmy przez $X(n,k)$ k-te statystyki rekordowe w ciągu $X_1, X_2, \ldots$ niezależnych zmiennych losowych o jednakowej ciągłej dystrybuancie F.

- W szczególnym przypadku, przez $Z(n,k)$ oznaczyny k-te statystyki rekordowe w ciągu $Z_1, Z_2, \ldots$ niezależnych zmiennych losowych o jednakowej standardowej wykładniczej dystrybuancie
-
$$F(x) = 0, \; x < 0,$$
$$F(x) = 1 - e^{-x}, \; x \geq 0.$$

Reprezentacja k-tych statystyk rekordowych

Poprzednio oznaczaliśmy przez $X(n,k)$ k-te statystyki rekordowe w ciągu $X_1, X_2, \ldots$ niezależnych zmiennych losowych o jednakowej ciągłej dystrybuancie F.

- W szczególnym przypadku, przez $Z(n,k)$ oznaczyny k-te statystyki rekordowe w ciągu $Z_1, Z_2, \ldots$ niezależnych zmiennych losowych o jednakowej standardowej wykładniczej dystrybuancie
-
$$F(x) = 0, \; x < 0,$$
$$F(x) = 1 - e^{-x}, \; x \geq 0.$$

- Dla dowolnej dystrybuanty F przez $F^{\leftarrow}$ oznaczamy funkcję odwrotną

$$F^{\leftarrow}(s) = \inf\{x : F(x) \geq s\}, \; 0 < s < 1.$$

Prawdziwa jest następująca reprezentacja:

Reprezentacja k-tych statystyk rekordowych cd.

Prawdziwa jest następująca reprezentacja:

- **Reprezentacja 1.** Dla dowolnego $n = 1, 2, \dots$ i dowolnego $k = 1, 2, \dots$

$$\left(X\left(1, k\right), \dots, X\left(n, k\right)\right) \stackrel{d}{=} \left(H\left(Z\left(1, k\right)\right), \dots, H\left(Z\left(n, k\right)\right)\right),$$

gdzie $H\left(x\right) = F^{\leftarrow}\left(1 - \exp\left(-x\right)\right)$.

Prawdziwa jest następująca reprezentacja:

- **Reprezentacja 1.** Dla dowolnego $n = 1, 2, \ldots$ i dowolnego $k = 1, 2, \ldots$

$$(X(1,k), \ldots, X(n,k)) \stackrel{d}{=} (H(Z(1,k)), \ldots, H(Z(n,k))),$$

gdzie $H(x) = F^{\leftarrow}(1 - \exp(-x))$.

- W szczególnym przypadku dla pierwszych rekordów, mamy **Reprezentacja 1a.** Dla dowolnego $n = 1, 2, \ldots$ i dowolnego $k = 1, 2, \ldots$

$$(X(1,1), \ldots, X(n,1)) \stackrel{d}{=} (H(Z(1,1)), \ldots, H(Z(n,1))).$$

Następnie

Następnie

- Z wyników Tata (1969)
  M.N. Tata, On outstanding values in a sequence of random variables, *Z. Wahrsch. verw. Gebiete*, Ser. B **12** (1969),9-20.

Następnie

- Z wyników Tata (1969)
  M.N. Tata, On outstanding values in a sequence of random variables, *Z. Wahrsch. verw. Gebiete*, Ser. B **12** (1969),9-20.
- Otrzymujemy, że zmienne losowe

$$Z(1,1), Z(2,1) - Z(1,1), Z(3,1) - Z(2,1), \ldots$$

są niezależne i mają standardowy rozkład wykładniczy.

Następnie

- Z wyników Tata (1969)
  M.N. Tata, On outstanding values in a sequence of random variables, *Z. Wahrsch. verw. Gebiete*, Ser. B **12** (1969),9-20.
- Otrzymujemy, że zmienne losowe

$$Z(1,1), Z(2,1) - Z(1,1), Z(3,1) - Z(2,1), \ldots$$

  są niezależne i mają standardowy rozkład wykładniczy.
- **Twierdzenie.** Mamy

$$\{X(n,1)\}_{n=1}^{\infty} \stackrel{d}{=} \{H(\omega_1 + \cdots + \omega_n)\}_{n=1}^{\infty},$$

  gdzie $H(x) = F^{\leftarrow}(1 - \exp(-x))$ i $\omega_1, \omega_2, \ldots$ jest ciągiem niezależnych zmiennych losowych o jednakowym standardowym rozkładzie wykładniczym.

Dziubdziela i Kopociński (1976) otrzymali następujące uogólnienie reprezentacji Tata:

Dziubdziela i Kopociński (1976) otrzymali następujące uogólnienie reprezentacji Tata:

- **Twierdzenie.** Niech $Z_1, Z_2, \ldots$ będą niezależnymi zmiennymi o jednakowym standardowym rozkładzie wykładniczym i niech $Z\left(n,k\right), n = 1, 2, \ldots$, będą odpowiadającymi mu k-tymi wartościami rekordowymi. Wtedy dla $k = 1, 2, \ldots$

Dziubdziela i Kopociński (1976) otrzymali następujące uogólnienie reprezentacji Tata:

- **Twierdzenie.** Niech $Z_1, Z_2, \dots$ będą niezależnymi zmiennymi o jednakowym standardowym rozkładzie wykładniczym i niech $Z(n,k)$, $n = 1, 2, \dots$, będą odpowiadającymi mu k-tymi wartościami rekordowymi. Wtedy dla $k = 1, 2, \dots$
- $$\{Z(n,k)\}_{n=1}^{\infty} \stackrel{d}{=} \left\{ \frac{\omega_1 + \cdots + \omega_n}{k} \right\}_{n=1}^{\infty},$$

  gdzie $\omega_1, \omega_2, \dots$ jest ciągiem niezależnych zmiennych losowych o jednakowym standardowym rozkładzie wykładniczym.

Z omówionego uogólnienia otrzymujemy twierdzenia:

Z omówionego uogólnienia otrzymujemy twierdzenia:

- **Twierdzenie.** Dla $k = 1, 2, \ldots$

$$\{X(n,k)\}_{n=1}^{\infty} \stackrel{d}{=} \left\{ H\left(\frac{\omega_1 + \cdots + \omega_n}{k} \right) \right\}_{n=1}^{\infty},$$

Z omówionego uogólnienia otrzymujemy twierdzenia:

- **Twierdzenie.** Dla $k = 1, 2, \ldots$

$$\{X(n,k)\}_{n=1}^{\infty} \stackrel{d}{=} \left\{ H\left(\frac{\omega_1 + \cdots + \omega_n}{k}\right)\right\}_{n=1}^{\infty},$$

- **Twierdzenie.** Ciąg $X(1,k), X(2,k), \ldots$ jest łańcuchem Markowa i dla $k \geq 1, n \geq 1$ oraz $x > u$

$$P\{X(n+1,k) > x \,|X(n,k) = u\} = \left(\frac{1 - F(x)}{1 - F(u)}\right)^k.$$

Rożważmy dwa ciągi niezależnych o jednakowym rozkładzie zmiennych losowych:

Rożważmy dwa ciągi niezależnych o jednakowym rozkładzie zmiennych losowych:

- $X_1, X_2, \ldots$ z ciągłą dystrybuantą F
  i

Rożważmy dwa ciągi niezależnych o jednakowym rozkładzie zmiennych losowych:

- $X_1, X_2, \ldots$ z ciągłą dystrybuantą F
  i
-

$$Y_1 = \min\{X_1, \ldots, X_k\}, Y_2 = \min\{X_{k+1}, \ldots, X_{2k}\}, \ldots .$$

Rożważmy dwa ciągi niezależnych o jednakowym rozkładzie zmiennych losowych:

- $X_1, X_2, \ldots$ z ciągłą dystrybuantą F
  i
- $$Y_1 = \min\{X_1, \ldots, X_k\}, Y_2 = \min\{X_{k+1}, \ldots, X_{2k}\}, \ldots.$$
- z dystrybuantą

$$G(x) = 1 - (1 - F(x))^k.$$

Rożważmy dwa ciągi niezależnych o jednakowym rozkładzie zmiennych losowych:

- $X_1, X_2, \dots$ z ciągłą dystrybuantą F
  i
- $$Y_1 = \min\{X_1, \dots, X_k\}, Y_2 = \min\{X_{k+1}, \dots, X_{2k}\}, \dots.$$
- z dystrybuantą
  $$G(x) = 1 - (1 - F(x))^k.$$
- Niech $Y(n, 1)$ oznaczają wartości rekordowe w ciągu $Y_1, Y_2, \dots$.

Prawdziwe jest twierdzenie

Prawdziwe jest twierdzenie

- **Twierdzenie.** Dla dowolnego $k = 1, 2, \ldots$,

$$\{X(n,k)\}_{n=1}^{\infty} \stackrel{d}{=} \{Y(n,1)\}_{n=1}^{\infty}$$

Prawdziwe jest twierdzenie

- **Twierdzenie.** Dla dowolnego $k = 1, 2, \ldots$,

$$\{X(n, k)\}_{n=1}^{\infty} \stackrel{d}{=} \{Y(n, 1)\}_{n=1}^{\infty}$$

- Wiadomo, że wartość rekordowa $X(n, 1)$ ma dystrybuantę

$$P\{X(n, 1) \leq x\} = \frac{1}{(n-1)!} \int_0^{-\log(1-F(x))} u^{n-1} e^{-u} du.$$

Prawdziwe jest twierdzenie

- **Twierdzenie.** Dla dowolnego $k = 1, 2, \ldots$,

$$\{X(n,k)\}_{n=1}^{\infty} \stackrel{d}{=} \{Y(n,1)\}_{n=1}^{\infty}$$

- Wiadomo, że wartość rekordowa $X(n,1)$ ma dystrybuantę

$$P\{X(n,1) \leq x\} = \frac{1}{(n-1)!} \int_0^{-\log(1-F(x))} u^{n-1} e^{-u} du.$$

- Podstawiając za F dystrybuantę $1 - (1 - F(x))^k$ otrzymujemy dystrybuantę k-tej wartości rekordowej

$$P\{X(n,k) \leq x\} = \frac{1}{(n-1)!} \int_0^{-k\log(1-F(x))} u^{n-1} e^{-u} du.$$

Rozkłady asymptotyczne k-tych rekordów zostały w przypadku $k = 1$ otrzymane prrzez S.I. Resnicka w 1973 r, a w 1976 r. W. Dziubdziela i B. Kopociński rozszerzyli ten wynik na dowone $k = 1, 2, \ldots$.

Rozkłady asymptotyczne k-tych rekordów zostały w przypadku $k = 1$ otrzymane prrzez S.I. Resnicka w 1973 r, a w 1976 r. W. Dziubdziela i B. Kopociński rozszerzyli ten wynik na dowone $k = 1, 2, \ldots$.

- Sidney I Resnick, Limit Laws for Record Values, *Stochastic Processes and their Applications*, **1** (1973), 67-82.

Rozkłady asymptotyczne k-tych rekordów zostały w przypadku $k = 1$ otrzymane prrzez S.I. Resnicka w 1973 r, a w 1976 r. W. Dziubdziela i B. Kopociński rozszerzyli ten wynik na dowone $k = 1, 2, \ldots$.

- Sidney I Resnick, Limit Laws for Record Values, *Stochastic Processes and their Applications*, **1** (1973), 67-82.
- Wiesław Dziubdziela, Bolesław Kopociński, Limiting properties of the k-th record values, *Zastosowania Matematyki* **15** (1976) , 187-190.

Asymptotyczne rozkłady k-tych rekordów cd.

Niech $Z_1, Z_2, \ldots$ będą niezależnymi zmiennymi losowymi o jednakowym standardowym rozkładzie wykładniczym. Dla dowolnego ustalonego $k = 1, 2, \ldots$, rozważmy ciąg k-tych wartości rekordowych związanych z tym ciągiem

$$Z(1,k), Z(2,k), \ldots .$$

Asymptotyczne rozkłady k-tych rekordów cd.

Niech $Z_1, Z_2, \ldots$ będą niezależnymi zmiennymi losowymi o jednakowym standardowym rozkładzie wykładniczym. Dla dowolnego ustalonego $k = 1, 2, \ldots$, rozważmy ciąg k-tych wartości rekordowych związanych z tym ciągiem

$$Z(1,k), Z(2,k), \ldots .$$

- Pamiętamy, że

$$\{Z(n,k)\}_{n=1}^{\infty} \stackrel{d}{=} \left\{ \frac{\omega_1 + \cdots + \omega_n}{k} \right\}_{n=1}^{\infty} .$$

Asymptotyczne rozkłady k-tych rekordów cd.

Niech $Z_1, Z_2, \ldots$ będą niezależnymi zmiennymi losowymi o jednakowym standardowym rozkładzie wykładniczym. Dla dowolnego ustalonego $k = 1, 2, \ldots$, rozważmy ciąg k-tych wartości rekordowych związanych z tym ciągiem

$$Z(1, k), Z(2, k), \ldots .$$

- Pamiętamy, że
$$\{Z(n, k)\}_{n=1}^{\infty} \stackrel{d}{=} \left\{\frac{\omega_1 + \cdots + \omega_n}{k}\right\}_{n=1}^{\infty} .$$
- Zatem, dla każdego $n = 1, 2, \ldots$
$$Z(n, k) \stackrel{d}{=} \xi_1 + \cdots + \xi_n,$$
gdzie $\xi_1, \xi_2, \ldots$ są niezależnymi zmiennymi o jednakowym rozkładzie wykładniczym.

Ponieważ

Ponieważ

- $$\mathtt{E}\xi_n = \frac{1}{k},\ \mathtt{Var}\xi_n = \frac{1}{k^2},\ n = 1, 2, \ldots,$$

Asymptotyczne rozkłady k-tych rekordów cd.

Ponieważ

- $$\mathtt{E}\xi_n = \frac{1}{k},\ \mathtt{Var}\xi_n = \frac{1}{k^2},\ n = 1, 2, \ldots,$$

- więc, z Centralnego Twierdzenia Granicznego zastosowanego do sumy $\xi_1 + \cdots + \xi_n$, otrzymujemy

$$P\left\{\frac{Z(n,k) - \frac{n}{k}}{\frac{n^{1/2}}{k}} \leq x\right\} \xrightarrow{w} \Phi(x),\ \mathtt{gdy}\ n \to \infty,$$

gdzie Φ jest dystrybuantą standardowego rozkładu normalnego.

W szczególnym przypadku $k = 1$, otrzymujemy

W szczególnym przypadku $k = 1$, otrzymujemy

- $$P\left\{\frac{Z(n,1)-n}{n^{1/2}} \leq x\right\} \xrightarrow{w} \Phi(x)\,,\ \text{gdy } n \to \infty,$$

W szczególnym przypadku $k = 1$, otrzymujemy

- $$P\left\{\frac{Z(n,1)-n}{n^{1/2}} \leq x\right\} \xrightarrow{w} \Phi(x), \text{ gdy } n \to \infty,$$

- zatem

$$P\left\{Z(n,1) \leq \sqrt{n}x + n\right\} = \frac{1}{(n-1)!}\int_0^{\sqrt{n}x+n} u^{n-1}e^{-u}du \xrightarrow{w} \Phi(x).$$

Asymptotyczne rozkłady k-tych rekordów cd.

W szczególnym przypadku $k = 1$, otrzymujemy

- $$P\left\{\frac{Z(n,1)-n}{n^{1/2}} \le x\right\} \xrightarrow{w} \Phi(x)\,, \text{ gdy } n \to \infty,$$

- zatem

$$P\left\{Z(n,1) \le \sqrt{n}x + n\right\} = \frac{1}{(n-1)!}\int_0^{\sqrt{n}x+n} u^{n-1}e^{-u}du \xrightarrow{w} \Phi(x)\,.$$

- bo

$$Z(n,1) = \omega_1 + \cdots + \omega_n$$

 ma rozkład gamma z parametrem n.

Wprowadźmy przekształcenie

Wprowadźmy przekształcenie

- $$R^{(k)}(x) = -k \ln(1 - F(x)).$$

Wprowadźmy przekształcenie

- $$R^{(k)}(x) = -k \ln (1 - F(x)).$$

- Wtedy

$$P\left\{R^{(k)}(X(n,k)) \leq x\right\} = P\{Z(n,1) \leq x\}$$

Wprowadźmy przekształcenie

- $$R^{(k)}(x) = -k \ln(1 - F(x)).$$

- Wtedy
$$P\left\{R^{(k)}(X(n,k)) \leq x\right\} = P\{Z(n,1) \leq x\}$$

- i
$$P\left\{R^{(k)}(X(n,k)) \leq x\right\} = \frac{1}{(n-1)!}\int_0^x u^{n-1}e^{-u}du.$$

Zatem

Zatem

- $$P\left\{R^{(k)}\left(X\left(n,k\right)\right) \leq \sqrt{n}x + n\right\} \xrightarrow{w} \Phi\left(x\right).$$

Zatem

- $$P\left\{R^{(k)}\left(X\left(n,k\right)\right)\leq\sqrt{n}x+n\right\}\xrightarrow{w}\Phi\left(x\right).$$

- Sformułujemy to co powiedzieliśmy jako twierdzenie:
Twierdzenie. Aby istniał ciągi normujące $a_n>0$ i $b_n\in\mathbf{R}$ oraz niezdegenerowana dystrybuanta Ψ taka, że

$$P\left\{Z\left(n,k\right)\leq a_nx+b_n\right\}\xrightarrow{w}\Psi\left(x\right)$$

potrzeba i wystarcza, by istniała niemalejąca funkcja $g\left(x\right)$ (być może przyjmująca wartości nieskończone) z więcej niż jednym punktem wzrostu taka, że

$$\lim_{n\to\infty}\frac{R^{(k)}\left(a_nx+b_n\right)-n}{\sqrt{n}}=g\left(x\right)$$

w punktach ciągłości $g\left(x\right)$.

Ponadto

Ponadto

- $$\Psi(x) = \Phi(g(x)),$$

  gdzie $\Phi(x)$ jest dystrybuantą standardowego rozkładu normalnego.

Ponadto

- $$\Psi(x) = \Phi(g(x)),$$

  gdzie $\Phi(x)$ jest dystrybuantą standardowego rozkładu normalnego.
- Pokazuje się, że funkcja $g(x)$ ma postać

  $$g(x) = -2\ln(-\ln G(x)),$$

  gdzie $G(x)$ jest dystrybuantą rozkładu ekstremalnego.

Asymptotyczne rozkłady k-tych rekordów cd.

Ponadto

- $$\Psi(x) = \Phi(g(x)),$$

  gdzie $\Phi(x)$ jest dystrybuantą standardowego rozkładu normalnego.
- Pokazuje się, że funkcja $g(x)$ ma postać

  $$g(x) = -2\ln(-\ln G(x)),$$

  gdzie $G(x)$ jest dystrybuantą rozkładu ekstremalnego.
- Wykorzystujemy pracę:
  Sidney I Resnick, Limit Laws for Record Values, *Stochastic Processes and their Applications*, **1** (1973), 67-82.

Oznaczyliśmy poprzednio (dla $k = 1, 2, \ldots$)

Oznaczyliśmy poprzednio (dla $k = 1, 2, \ldots$)

- $$R^{(k)}(x) = -k \ln(1 - F(x)).$$

Obszary przyciągania dla k-tych rekordów

Oznaczyliśmy poprzednio (dla $k = 1, 2, \ldots$)

- $$R^{(k)}(x) = -k \ln(1 - F(x)).$$

- Dystrybuantę F można zapisać jako
$$F(x) = 1 - \frac{1}{k} \exp\left(-R^{(k)}(x)\right).$$

Oznaczyliśmy poprzednio (dla $k = 1, 2, \ldots$)

- $$R^{(k)}(x) = -k \ln(1 - F(x)).$$

- Dystrybuantę F można zapisać jako
$$F(x) = 1 - \frac{1}{k} \exp\left(-R^{(k)}(x)\right).$$

- Wiążemy z nią tzw. *dystrybuantę stowarzyszoną*
$$H(x) = 1 - \frac{1}{k} \exp\left(-\sqrt{R^{(k)}(x)}\right).$$

Obszary przyciągania dla k-tych rekordów cd.

We następnym wzorze parametr dyskretny n zastępujemy parametrem ciągłym s

We następnym wzorze parametr dyskretny n zastępujemy parametrem ciągłym s

- $$\lim_{n\to\infty} \frac{R^{(k)}\left(a_n x + b_n\right) - n}{\sqrt{n}} = g\left(x\right).$$

We następnym wzorze parametr dyskretny n zastępujemy parametrem ciągłym s

- $$\lim_{n\to\infty} \frac{R^{(k)}\left(a_n x + b_n\right) - n}{\sqrt{n}} = g\left(x\right).$$
- Otrzymujemy
  $$\lim_{s\to\infty} \frac{R^{(k)}\left(a_s x + b_s\right) - s}{\sqrt{s}} = g\left(x\right)$$
  w punktach ciągłości g.

Obszary przyciągania dla k-tych rekordów cd.

We następnym wzorze parametr dyskretny n zastępujemy parametrem ciągłym s

- $$\lim_{n\to\infty} \frac{R^{(k)}(a_n x + b_n) - n}{\sqrt{n}} = g(x).$$
- Otrzymujemy
$$\lim_{s\to\infty} \frac{R^{(k)}(a_s x + b_s) - s}{\sqrt{s}} = g(x)$$
w punktach ciągłości g.
- Stąd, dla x takich, że $|g(x)| < \infty$,
$$\lim_{s\to\infty} \frac{R^{(k)}(a_s x + b_s) - s}{s} = 0.$$

Stąd

Stąd

- $$\lim_{s\to\infty} \frac{R^{(k)}\left(a_s x + b_s\right)}{s} = 1.$$

Obszary przyciągania dla k-tych rekordów cd.

Stąd

- $$\lim_{s\to\infty} \frac{R^{(k)}\left(a_s x + b_s\right)}{s} = 1.$$

- Zauważmy, że
$$\frac{R^{(k)}\left(a_s x + b_s\right) - s}{\sqrt{s}} =$$
$$\left(\sqrt{R^{(k)}\left(a_s x + b_s\right)} - \sqrt{s}\right) \frac{\sqrt{R^{(k)}\left(a_s x + b_s\right)} + \sqrt{s}}{\sqrt{s}},$$
$$\lim_{s\to\infty} \sqrt{\frac{R^{(k)}\left(a_s x + b_s\right)}{s}} = 1.$$

Stąd

- $$\lim_{s\to\infty} \frac{R^{(k)}\left(a_s x + b_s\right)}{s} = 1.$$

- Zauważmy, że
$$\frac{R^{(k)}\left(a_s x + b_s\right) - s}{\sqrt{s}} =$$
$$\left(\sqrt{R^{(k)}\left(a_s x + b_s\right)} - \sqrt{s}\right) \frac{\sqrt{R^{(k)}\left(a_s x + b_s\right)} + \sqrt{s}}{\sqrt{s}},$$
$$\lim_{s\to\infty} \sqrt{\frac{R^{(k)}\left(a_s x + b_s\right)}{s}} = 1.$$

- Ostatecznie
$$\lim_{s\to\infty} \sqrt{R^{(k)}\left(a_s x + b_s\right)} - \sqrt{s} = \frac{1}{2} g\left(x\right).$$

Z poprzedniego otrzymujemy

Z poprzedniego otrzymujemy

- $$\lim_{s\to\infty} \exp\left[\sqrt{s}\right] \exp\left[-\sqrt{R^{(k)}\left(a_s x + b_s\right)}\right] = \exp\left[-\frac{1}{2} g\left(x\right)\right].$$

Z poprzedniego otrzymujemy

- $$\lim_{s\to\infty} \exp\left[\sqrt{s}\right] \exp\left[-\sqrt{R^{(k)}\left(a_s x + b_s\right)}\right] = \exp\left[-\frac{1}{2}g\left(x\right)\right].$$

- Podstawiając $y = \exp\left[\sqrt{s}\right]$, mamy
$$\lim_{y\to\infty} y\left(1 - H\left(a\left(\ln^2 y\right)x + b\left(\ln^2 y\right)\right)\right) = \exp\left[-\frac{1}{2}g\left(x\right)\right],$$

Obszary przyciągania dla k-tych rekordów cd.

Z poprzedniego otrzymujemy

- $$\lim_{s\to\infty} \exp\left[\sqrt{s}\right] \exp\left[-\sqrt{R^{(k)}\left(a_s x + b_s\right)}\right] = \exp\left[-\frac{1}{2} g\left(x\right)\right].$$

- Podstawiając $y = \exp\left[\sqrt{s}\right]$, mamy
$$\lim_{y\to\infty} y\left(1 - H\left(a\left(\ln^2 y\right)x + b\left(\ln^2 y\right)\right)\right) = \exp\left[-\frac{1}{2} g\left(x\right)\right],$$

- Zatem $H\left(x\right)$ należy do obszaru przyciągania rozkładu ekstremalnego i
$$\exp\left[-\sqrt{g\left(x\right)}\right] = -\ln G\left(x\right),$$
gdzie $G\left(x\right)$ jest ekstremalną dystrybuantą.

Graniczne rozkłady k-tych rekordowych statystyk należą do jednego z trzech typów rozkładów:

Graniczne rozkłady k-tych rekordowych statystyk należą do jednego z trzech typów rozkładów:

- $$\Psi_1(x) = \Phi(x), \; -\infty < x < \infty,$$

Graniczne rozkłady k-tych rekordowych statystyk należą do jednego z trzech typów rozkładów:

- $$\Psi_1(x) = \Phi(x), \; -\infty < x < \infty,$$

- $$\Psi_{2,\alpha}(x) = 0, \; x \leq 0,$$
  $$\Psi_{2,\alpha}(x) = \Phi(\ln x^{\alpha}), \; x > 0, \alpha > 0,$$

Graniczne rozkłady k-tych rekordowych statystyk należą do jednego z trzech typów rozkładów:

- $$\Psi_1(x) = \Phi(x), \; -\infty < x < \infty,$$

- $$\Psi_{2,\alpha}(x) = 0, \; x \leq 0,$$
  $$\Psi_{2,\alpha}(x) = \Phi(\ln x^{\alpha}), \; x > 0, \alpha > 0,$$

- $$\Psi_{3,\alpha}(x) = \Phi\left(\ln(-x)^{-\alpha}\right), \; x \leq 0, \alpha > 0,$$
  $$\Psi_{3,\alpha}(x) = 1, \; x > 0.$$

Rozkład wykładniczy

Rozkład wykładniczy

- $$F(x) = 0, \; x \leq 0,$$
  $$F(x) = 1 - e^{-\lambda x}, \; x > 0, \lambda > 0$$
  należy do obszaru przyciągania rozkładu Gumbela

Rozkład wykładniczy

- $$F(x) = 0, \; x \leq 0,$$
  $$F(x) = 1 - e^{-\lambda x}, \; x > 0, \lambda > 0$$
  należy do obszaru przyciągania rozkładu Gumbela
- $$\Lambda(x) = \exp\{-\exp(-x)\}, \; -\infty < x < \infty.$$

Rozkład wykładniczy

- $$F(x) = 0,\ x \leq 0,$$
  $$F(x) = 1 - e^{-\lambda x},\ x > 0, \lambda > 0$$
  należy do obszaru przyciągania rozkładu Gumbela

- $$\Lambda(x) = \exp\{-\exp(-x)\},\ -\infty < x < \infty.$$

- Mamy bowiem
  $$n\left(1 - F\left(\frac{1}{\lambda}x + \frac{1}{\lambda}\ln n\right)\right) = e^{-x},\ x > -\ln n,$$

Rozkład wykładniczy

- $$F(x) = 0,\ x \leq 0,$$
  $$F(x) = 1 - e^{-\lambda x},\ x > 0, \lambda > 0$$
  należy do obszaru przyciągania rozkładu Gumbela
- $$\Lambda(x) = \exp\{-\exp(-x)\},\ -\infty < x < \infty.$$
- Mamy bowiem
  $$n\left(1 - F\left(\frac{1}{\lambda}x + \frac{1}{\lambda}\ln n\right)\right) = e^{-x},\ x > -\ln n,$$
- a stąd
  $$\lim_{n\to\infty} n\left(1 - F\left(\frac{1}{\lambda}x + \frac{1}{\lambda}\ln n\right)\right) = e^{-x},\ -\infty < x < \infty.$$

W tym przypadku, dla $k = 1, 2, \ldots$,

Obszary przyciągania dla k-tych rekordów. Przykład

W tym przypadku, dla $k = 1, 2, \dots$,

- $$R^{(k)}(x) = -k \ln\left(e^{-\lambda x}\right) = k\lambda x$$

W tym przypadku, dla $k = 1, 2, \ldots$,

- $$R^{(k)}(x) = -k \ln\left(e^{-\lambda x}\right) = k\lambda x$$

- i

$$g(x) = \lim_{n\to\infty} \frac{R^{(k)}\left(\frac{\sqrt{n}}{k\lambda}x + \frac{n}{k\lambda}\right)}{\sqrt{n}} = x.$$

Obszary przyciągania dla k-tych rekordów. Przykład

W tym przypadku, dla $k = 1, 2, \dots$,

- $$R^{(k)}(x) = -k \ln\left(e^{-\lambda x}\right) = k\lambda x$$

- i
$$g(x) = \lim_{n\to\infty} \frac{R^{(k)}\left(\frac{\sqrt{n}}{k\lambda}x + \frac{n}{k\lambda}\right)}{\sqrt{n}} = x.$$

- Ostatecznie
$$\lim_{n\to\infty} P\left\{X(n,k) \le \frac{\sqrt{n}}{k\lambda}x + \frac{n}{k\lambda}\right\} = \Phi(x).$$

Twierdzenie (Duality Theorem). Niech $F(x) = 1 - \exp[-R(x)]$ będzie ciągłą dystrybuantą i $H^{(k)}(x) = 1 - \exp\left[-\sqrt{R^{(k)}(x)}\right]$, gdzie $R^{(k)}(x) = kR(x)$.

Twierdzenie (Duality Theorem). Niech $F(x) = 1 - \exp[-R(x)]$ będzie ciągłą dystrybuantą i $H^{(k)}(x) = 1 - \exp\left[-\sqrt{R^{(k)}(x)}\right]$, gdzie $R^{(k)}(x) = kR(x)$.

- $R^{(k)}(x)$ należy od obszaru przyciągania rozkładu Ψ_1 wtedy i tylko wtedy, gdy $H^{(k)}(x)$ należy do obszaru przyciągania rozkładu Gumbela Λ,

Obszary przyciągania dla k-tych rekordów. Twierdzenie

Twierdzenie (Duality Theorem). Niech $F(x) = 1 - \exp[-R(x)]$ będzie ciągłą dystrybuantą i $H^{(k)}(x) = 1 - \exp\left[-\sqrt{R^{(k)}(x)}\right]$, gdzie $R^{(k)}(x) = kR(x)$.

- $R^{(k)}(x)$ należy od obszaru przyciągania rozkładu Ψ_1 wtedy i tylko wtedy, gdy $H^{(k)}(x)$ należy do obszaru przyciągania rozkładu Gumbela Λ,
- przy czym stałe normujące są postaci

$$a_n = \left[R^{(k)}\right]^{\leftarrow}(n + \sqrt{n}) - \left[R^{(k)}\right]^{\leftarrow}(n), \; b_n = \left[R^{(k)}\right]^{\leftarrow}(n).$$

Obszary przyciągania dla k-tych rekordów. Twierdzenie

Twierdzenie (Duality Theorem). Niech $F(x) = 1 - \exp\left[-R(x)\right]$ będzie ciągłą dystrybuantą i $H^{(k)}(x) = 1 - \exp\left[-\sqrt{R^{(k)}(x)}\right]$, gdzie $R^{(k)}(x) = kR(x)$.

- $R^{(k)}(x)$ należy od obszaru przyciągania rozkładu Ψ_1 wtedy i tylko wtedy, gdy $H^{(k)}(x)$ należy do obszaru przyciągania rozkładu Gumbela Λ,
- przy czym stałe normujące są postaci

$$a_n = \left[R^{(k)}\right]^{\leftarrow}\left(n + \sqrt{n}\right) - \left[R^{(k)}\right]^{\leftarrow}(n), \; b_n = \left[R^{(k)}\right]^{\leftarrow}(n).$$

- W rozważanym wcześniej przykładzie:

$$R^{(k)}(x) = k\lambda x, a_n = \frac{n+\sqrt{n}}{k\lambda} - \frac{n}{k\lambda} = \frac{\sqrt{n}}{k\lambda}, \; b_n = \frac{n}{k\lambda}.$$

Twierdzenie (Duality Theorem).

Obszary przyciągania dla k-tych rekordów. Twierdzenie cd.

Twierdzenie (Duality Theorem).

- $R^{(k)}(x)$ należy od obszaru przyciągania rozkładu $\Psi_{2,\alpha}$ wtedy i tylko wtedy, gdy $H^{(k)}(x)$ należy do obszaru przyciągania rozkładu Frécheta $\Phi_{\frac{\alpha}{2}}$,

Obszary przyciągania dla k-tych rekordów. Twierdzenie cd.

Twierdzenie (Duality Theorem).

- $R^{(k)}(x)$ należy od obszaru przyciągania rozkładu $\Psi_{2,\alpha}$ wtedy i tylko wtedy, gdy $H^{(k)}(x)$ należy do obszaru przyciągania rozkładu Frécheta $\Phi_{\frac{\alpha}{2}}$,
- przy czym stałe normujące są postaci

$$a_n = \left[R^{(k)}\right]^{\leftarrow}(n), \; b_n = 0.$$

Twierdzenie (Duality Theorem).

- $R^{(k)}(x)$ należy od obszaru przyciągania rozkładu $\Psi_{2,\alpha}$ wtedy i tylko wtedy, gdy $H^{(k)}(x)$ należy do obszaru przyciągania rozkładu Frécheta $\Phi_{\frac{\alpha}{2}}$,
- przy czym stałe normujące są postaci

$$a_n = \left[R^{(k)}\right]^{\leftarrow}(n)\,,\ b_n = 0.$$

- $R^{(k)}(x)$ należy od obszaru przyciągania rozkładu $\Psi_{2,\alpha}$ wtedy i tylko wtedy, gdy $H^{(k)}(x)$ należy do obszaru przyciągania rozkładu Weibulla $\Psi_{\frac{\alpha}{2}}$,

Obszary przyciągania dla k-tych rekordów. Twierdzenie cd.

Twierdzenie (Duality Theorem).

- $R^{(k)}(x)$ należy od obszaru przyciągania rozkładu $\Psi_{2,\alpha}$ wtedy i tylko wtedy, gdy $H^{(k)}(x)$ należy do obszaru przyciągania rozkładu Frécheta $\Phi_{\frac{\alpha}{2}}$,
- przy czym stałe normujące są postaci

$$a_n = \left[R^{(k)}\right]^{\leftarrow}(n)\,,\ b_n = 0.$$

- $R^{(k)}(x)$ należy od obszaru przyciągania rozkładu $\Psi_{2,\alpha}$ wtedy i tylko wtedy, gdy $H^{(k)}(x)$ należy do obszaru przyciągania rozkładu Weibulla $\Psi_{\frac{\alpha}{2}}$,
- przy czym stałe normujące są postaci

$$a_n = x^F - \left[R^{(k)}\right]^{\leftarrow}(n)\,,\ b_n = x^F,$$

gdzie x^F jest skończone.

1. Teoria wartości ekstremalnych. Ciąg niezależnych zmiennych losowych o jednakowym rozkładzie.
2. Teoria wartości rekordowych. Ciąg niezależnych zmiennych losowych o jednakowym rozkładzie.
3. **Badania statystyczne w oparciu o wartości rekordowe**
4. Zakończenie

Rozkład łączny k-tych wartości rekordowch

Niech $X_1, X_2, \ldots$ będzie ciągiem niezależnych zmiennych losowych o jednakowej ciągłej dystrybuancie F i gęstości f. Dla $k = 1, 2, \ldots$, niech

$$X(1, k), X(2, k), \cdots, X(n, k)$$

będą n pierwszymi k-tymi wartościami rekordowymi w ciągu $X_1, X_2, \ldots$.

Rozkład łączny k-tych wartości rekordowch

Niech $X_1, X_2, \ldots$ będzie ciągiem niezależnych zmiennych losowych o jednakowej ciągłej dystrybuancie F i gęstości f. Dla $k = 1, 2, \ldots$, niech

$$X(1, k), X(2, k), \cdots, X(n, k)$$

będą n pierwszymi k-tymi wartościami rekordowymi w ciągu $X_1, X_2, \ldots$.

- Interesują nas metody statystyczne pozwalające, na podstawie znajomości tych wartości rekordowych, przeprowadzać wnioskowania statystyczne dotyczące dytstrybuanty F. W tym celu potrzebne będą rozkłady łączne wartości rekordowych.

Rozkład łączny k-tych wartości rekordowch

Niech $X_1, X_2, \ldots$ będzie ciągiem niezależnych zmiennych losowych o jednakowej ciągłej dystrybuancie F i gęstości f. Dla $k = 1, 2, \ldots$, niech

$$X(1,k), X(2,k), \cdots, X(n,k)$$

będą n pierwszymi k-tymi wartościami rekordowymi w ciągu $X_1, X_2, \ldots$.

- Interesują nas metody statystyczne pozwalające, na podstawie znajomości tych wartości rekordowych, przeprowadzać wnioskowania statystyczne dotyczące dytstrybuanty F. W tym celu potrzebne będą rozkłady łączne wartości rekordowych.
- W pracy
  W. Dziubdziela, B. Kopociński, Limiting properties of the k-th record values, *Zastosowania Matematyki* **15** (1976) , 187-190.
  znaleziono rozkład k-tej wartości rekordowej:

Gęstość rozkładu k-tej wartości rekordowej $X(n,k)$:

Rozkład łączny k-tych wartości rekordowch cd.

Gęstość rozkładu k-tej wartości rekordowej $X(n,k)$:

- $$f_{X(n,k)}(x) = \frac{k^n}{(n-1)!}\left[H(x)\right]^{n-1}\left[(1-F(x))\right]^{k-1} f(x),$$

  gdie

  $$H(x) = -\ln(1-F(x)).$$

Rozkład łączny k-tych wartości rekordowch cd.

Gęstość rozkładu k-tej wartości rekordowej $X(n,k)$:

- $$f_{X(n,k)}(x) = \frac{k^n}{(n-1)!}\left[H(x)\right]^{n-1}\left[(1-F(x))\right]^{k-1} f(x),$$

  gdie

  $$H(x) = -\ln(1-F(x)).$$

- W pracy
  Z. Grudzień, Charakterystyka rozkładów w terminach statystyk rekordowych oraz rozkłady i momenty statystyk porządkowych i rekordowych z prób o losowej liczebności, *Praca doktorska, UMCS Lublin* (1982).
  znaleziono łączny rozkład dwóch k-tych wartości rekordowych:

Gęstość łączna k-tej wartości rekordowej $X(m,k)$ i k-tej wartości rekordowej $X(n,k)$, $1 \leq m < n$, $n = 1, 2, \ldots$:

Rozkład łączny k-tych wartości rekordowch cd.

Gęstość łączna k-tej wartości rekordowej $X(m,k)$ i k-tej wartości rekordowej $X(n,k)$, $1 \leq m < n$, $n = 1, 2, \ldots$:

- $$f_{X(m,k),X(n,k)}(x,y) =$$
$$\frac{k^n}{(m-1)!\,(n-m-1)!}\left[H(y) - H(x)\right]^{n-m-1}\left[H(x)\right]^{m-1} \times$$
$$\times h(x)\left[(1-F(y))\right]^{k-1} f(y)\,, \; x < y,$$

  gdzie
$$h(x) = H'(x)\,.$$

Rozkład łączny k-tych wartości rekordowch cd.

Gęstość łączna k-tej wartości rekordowej $X(m,k)$ i k-tej wartości rekordowej $X(n,k)$, $1 \leq m < n$, $n = 1, 2, \ldots$:

- $$f_{X(m,k),X(n,k)}(x,y) =$$
  $$\frac{k^n}{(m-1)!\,(n-m-1)!}\left[H(y)-H(x)\right]^{n-m-1}\left[H(x)\right]^{m-1} \times$$
  $$\times h(x)\left[(1-F(y))\right]^{k-1} f(y)\,, \;\; x < y,$$
  gdzie
  $$h(x) = H'(x)\,.$$

- W pracy
  U. Kamps, A concept of generalized order statistics, *Journal of Statistical Planning and Inference* **48** (1995) , 1-23.
  znaleziono łączny rozkład n k-tych wartości rekordowych:

Gęstość łączna pierwszych n k-tych wartości rekordowych:

Gęstość łączna pierwszych n k-tych wartości rekordowych:

- $$f_{X(1,k),X(2,k),\dots,X(n,k)}(y_1, y_2, \dots, y_n) = k^n \prod_{i=1}^{n-1} \frac{f(y_i)}{1 - F(y_i)} (1 - F(y_n))^{k-1} f(y_n),$$

$$y_1 < y_2 < \cdots < y_n.$$

Rozkład łączny k-tych wartości rekordowch cd.

Gęstość łączna pierwszych n k-tych wartości rekordowych:

- $$f_{X(1,k),X(2,k),\ldots,X(n,k)}\left(y_1,y_2,\ldots,y_n\right)=$$
  $$k^n\prod_{i=1}^{n-1}\frac{f\left(y_i\right)}{1-F\left(y_i\right)}\left(1-F\left(y_n\right)\right)^{k-1}f\left(y_n\right),$$
  $$y_1<y_2<\cdots<y_n.$$

- Powyższą gęstość wykorzystuje się przy wnioskowaniach opartych o funkcję wiarogodności.

Rozkład łączny k-tych wartości rekordowch cd.

Gęstość łączna pierwszych n k-tych wartości rekordowych:

- $$f_{X(1,k),X(2,k),\ldots,X(n,k)}\left(y_1,y_2,\ldots,y_n\right)=$$
$$k^n\prod_{i=1}^{n-1}\frac{f\left(y_i\right)}{1-F\left(y_i\right)}\left(1-F\left(y_n\right)\right)^{k-1}f\left(y_n\right),$$
$$y_1<y_2<\cdots<y_n.$$

- Powyższą gęstość wykorzystuje się przy wnioskowaniach opartych o funkcję wiarogodności.
- Podamy teraz przykłady.

W pracach:

W pracach:

- J. Saran, A.Pandey, Estimation of parameters of a two-parameter exponential distribution and its characterization by *k-th* record values, *International Journal of Statistical Sciences*, **4** (2005), 1-12.
  J. Ahmadi, M. Doostparast, Statistical inference base on *k*-records, *Mashhad R. J. Math. Sci.*, **1** (2008), 67-82.

W pracach:

- J. Saran, A.Pandey, Estimation of parameters of a two-parameter exponential distribution and its characterization by *k-th* record values, *International Journal of Statistical Sciences*, **4** (2005), 1-12.
  J. Ahmadi, M. Doostparast, Statistical inference base on *k*-records, *Mashhad R. J. Math. Sci.*, **1** (2008), 67-82.
- rozważany jest rozkład wykładniczy.
  Definicja. Mówimy, że zmienna losowa X ma rozkład wykładniczy $Exp\,(\mu, \sigma)$ o dwóch parametrach, jeżeli jej gęstość jest postaci

$$f\,(x) = \frac{1}{\sigma} \exp\left\{ -\frac{X-\mu}{\sigma} \right\}, \ \mu \leq x < \infty, \ \sigma > 0.$$

Stąd obliczamy dystrybuantę

Stąd obliczamy dystrybuantę

- $$F(x) = 1 - \exp\left\{-\frac{X-\mu}{\sigma}\right\}, \; \mu \leq x < \infty, \; \sigma > 0.$$

Przykłady. Rozkład wykładniczy cd.

Stąd obliczamy dystrybuantę

- $$F(x) = 1 - \exp\left\{-\frac{X-\mu}{\sigma}\right\},\ \mu \leq x < \infty,\ \sigma > 0.$$

- Logarytm funkcji wiarogodności dla k-tych wartości rekordowych ma postać

$$\ln L = n \ln k + \sum_{i=1}^{n-1} \ln \frac{f(y_i)}{1 - F(y_i)} +$$

$$(k-1)(1 - F(y_n)) + \ln f(y_n),$$

$$y_1 < y_2 < \ldots < y_n.$$

W przypadku rozkładu wykładniczego mamy

W przypadku rozkładu wykładniczego mamy

- $$f(x) = \frac{1}{\sigma}(1 - F(x)), \ \mu \leq x < \infty, \ \sigma > 0,$$

W przypadku rozkładu wykładniczego mamy

- $$f(x) = \frac{1}{\sigma}(1 - F(x)), \; \mu \leq x < \infty, \; \sigma > 0,$$

- Zatem logarytm funkcji wiarogodności przyjmuje postać

$$\ln L = n \ln k + \sum_{i=1}^{n-1} \ln \frac{1}{\sigma} + (k-1)(1 - F(y_n)) + \ln\left(\frac{1}{\sigma}(1 - F(y_n))\right),$$

$$\mu \leq y_1 < y_2 < \ldots < y_n < \infty.$$

W przypadku rozkładu wykładniczego mamy

- $$f(x) = \frac{1}{\sigma}(1 - F(x)), \ \mu \leq x < \infty, \ \sigma > 0,$$

- Zatem logarytm funkcji wiarogodności przyjmuje postać

$$\ln L = n \ln k + \sum_{i=1}^{n-1} \ln \frac{1}{\sigma} + (k-1)(1 - F(y_n)) + \ln\left(\frac{1}{\sigma}(1 - F(y_n))\right),$$

$$\mu \leq y_1 < y_2 < \ldots < y_n < \infty.$$

- Ostatecznie

$$\ln L = -n \ln \sigma + n \ln k - k \frac{y_n - \mu}{\sigma}, \ \mu \leq y_1 < y_2 < \ldots < y_n < \infty.$$

Przyjmijmy, że parametry μ i σ są nieznane. Wtedy ich estymatory największej wiarogodności są postaci:

Przykłady. Rozkład wykładniczy cd.

Przyjmijmy, że parametry μ i σ są nieznane. Wtedy ich estymatory największej wiarogodności są postaci:

- $$\hat{\mu} = X(1, k)$$

Przyjmijmy, że parametry μ i σ są nieznane. Wtedy ich estymatory największej wiarogodności są postaci:

- $$\hat{\mu} = X(1, k)$$

- i
$$\hat{\sigma} = \frac{k}{n}\left(X(n, k) - X(1, k)\right).$$

Przyjmijmy, że parametry μ i σ są nieznane. Wtedy ich estymatory największej wiarogodności są postaci:

- $$\hat{\mu} = X(1,k)$$
- i
  $$\hat{\sigma} = \frac{k}{n}\left(X(n,k) - X(1,k)\right).$$
- Łatwo sprawdzić, że
  $$X(1,k) \texttt{ ma rozkład } Exp\left(\mu, \frac{\sigma}{k}\right),$$

Przyjmijmy, że parametry μ i σ są nieznane. Wtedy ich estymatory największej wiarogodności są postaci:

- $$\hat{\mu} = X(1,k)$$
- i $$\hat{\sigma} = \frac{k}{n}\left(X(n,k) - X(1,k)\right).$$
- Łatwo sprawdzić, że $$X(1,k) \texttt{ ma rozkład } Exp\left(\mu, \frac{\sigma}{k}\right),$$
- $$X(n,k) - X(1,k) \texttt{ i } X(1,k)$$ są niezależnymi zmiennymi losowymi.

A także

A także

- $$X(n,k) - X(1,k)$$

  ma rozkład gamma z parametrami $n-1$ i $\frac{k}{\sigma}$.

A także

- $$X(n,k) - X(1,k)$$
  ma rozkład gamma z parametrami $n-1$ i $\frac{k}{\sigma}$.
- Z własności z poprzedniego slajdu
  $$E(\hat{\mu}) = \mu + \frac{\sigma}{k},$$
  $$MSE(\hat{\mu}) = 2\frac{\sigma^2}{k^2},$$
  gdzie MSE (mean square error) jest średnim błędem kwadratowym
  $$MSE(\hat{\mu}) = E\left[(\hat{\mu} - \mu)^2\right].$$

Natomiast dla estymatora $\hat{\sigma}$ mamy

Natomiast dla estymatora $\hat{\sigma}$ mamy

- $$E\left(\hat{\sigma}\right) = \frac{n-1}{n}\sigma,$$

Natomiast dla estymatora $\hat{\sigma}$ mamy

- $$E(\hat{\sigma}) = \frac{n-1}{n}\sigma,$$

- $$MSE(\hat{\sigma}) = \frac{\sigma^2}{n},$$

  przy czym błąd nie zależy od k,

Natomiast dla estymatora $\hat{\sigma}$ mamy

- $$E(\hat{\sigma}) = \frac{n-1}{n}\sigma,$$

- $$MSE(\hat{\sigma}) = \frac{\sigma^2}{n},$$

  przy czym błąd nie zależy od k,

- oraz
  $$Cov(\hat{\mu}, \hat{\sigma}) = 0.$$

Zauważmy, że $\hat{\mu}$ jest obciążonym estymatorem parametru μ.

Zauważmy, że $\hat{\mu}$ jest obciążonym estymatorem parametru μ.

- Nieobciążony estymator jest postaci

$$\tilde{\mu} = \frac{n+k-1}{n-1} X(1,k) - \frac{k}{n-1} X(n,k).$$

1. Teoria wartości ekstremalnych. Ciąg niezależnych zmiennych losowych o jednakowym rozkładzie.
2. Teoria wartości rekordowych. Ciąg niezależnych zmiennych losowych o jednakowym rozkładzie.
3. Badania statystyczne w oparciu o wartości rekordowe
4. Zakończenie

Omówimy kilka prac opublikowanych w latach 2014-2016, w których rozwijana jest teoria k-tych wartości rekordowych.

Omówimy kilka prac opublikowanych w latach 2014-2016, w których rozwijana jest teoria k-tych wartości rekordowych.

- J. Paul, P. Y. Thomas, On generalized upper(k)record values from Weibull distribution, *Statistica* **75** (2015), 313-330.

Omówimy kilka prac opublikowanych w latach 2014-2016, w których rozwijana jest teoria k-tych wartości rekordowych.

- J. Paul, P. Y. Thomas, On generalized upper(k)record values from Weibull distribution, *Statistica* **75** (2015), 313-330.
- Rozkład Weibulla ma dystrybuantę

$$F(x) = 1 - \exp\left\{-\left(\frac{x-\theta}{\sigma}\right)^c\right\}, \quad x > \theta,$$

$$\sigma > 0, -\infty < \theta < \infty, c > 0.$$

Omówimy kilka prac opublikowanych w latach 2014-2016, w których rozwijana jest teoria k-tych wartości rekordowych.

- J. Paul, P. Y. Thomas, On generalized upper(k)record values from Weibull distribution, *Statistica* **75** (2015), 313-330.
- Rozkład Weibulla ma dystrybuantę

$$F(x) = 1 - \exp\left\{-\left(\frac{x-\theta}{\sigma}\right)^c\right\}, \; x > \theta,$$

$$\sigma > 0, -\infty < \theta < \infty, c > 0.$$

- W oparciu o k-te wartości rekordowe scharakteryzowano rozkład Weibulla, estymowano jego parametry (estymatory BLUE [*Best Linear Unbiased Estimation*]) i przewidywano wartości przyszłych k-tych rekordów.

Następnie.

Następnie.

- P. Y. Thomas, J. Paul, On generalized lower(k)record values from the Fréchet distribution, *J. Japan Statist. Soc.* **44** (2014), 157-178.

Następnie.

- P. Y. Thomas, J. Paul, On generalized lower(k)record values from the Fréchet distribution, *J. Japan Statist. Soc.* **44** (2014), 157-178.
- Rozkład Frécheta ma dystrybuantę

$$F(x) = \exp\left\{-\left(\frac{x-\theta}{\sigma}\right)^c\right\}, \; x > \theta,$$

$$\sigma > 0, -\infty < \theta < \infty, c > 0.$$

Następnie.

- P. Y. Thomas, J. Paul, On generalized lower(k)record values from the Fréchet distribution, *J. Japan Statist. Soc.* **44** (2014), 157-178.
- Rozkład Frécheta ma dystrybuantę

$$F(x) = \exp\left\{-\left(\frac{x-\theta}{\sigma}\right)^c\right\}, \; x > \theta,$$

$$\sigma > 0, -\infty < \theta < \infty, c > 0.$$

- W oparciu o dolne k-te wartości rekordowe estymowano parametry rozkładu Frécheta (estymatory BLUE) i przewidywano wartości przyszłych k-tych rekordów.

Dalej.

Dalej.

- H. P. Singh, V. Mehta, A class of shrinkage estimators of scale parameter of uniform distribution based on k-record values, *Natl. Acad. Sci Lett.* (2016), on line.

Dalej.

- H. P. Singh, V. Mehta, A class of shrinkage estimators of scale parameter of uniform distribution based on k-record values, *Natl. Acad. Sci Lett.* (2016), on line.
- Rozkład jednostajny ma gęstość

$$f(x) = \frac{1}{\sigma}, \ \mu < x < \mu + \sigma, \ \sigma > 0,$$

gdzie σ jest parametrem skali.

Dalej.

- H. P. Singh, V. Mehta, A class of shrinkage estimators of scale parameter of uniform distribution based on k-record values, *Natl. Acad. Sci Lett.* (2016), on line.
- Rozkład jednostajny ma gęstość

$$f(x) = \frac{1}{\sigma}, \ \mu < x < \mu + \sigma, \ \sigma > 0,$$

  gdzie σ jest parametrem skali.
- W oparciu o k-te wartości rekordowe estymowano parametr skali rozkładu jednostajnego, w przypadku posiadania informacji *a'priori*.

W pracy

W pracy

- V. Kumar, Some results on Tsallis entropy measure and k-record values, *Physica A* (2016), on line.

W pracy

- V. Kumar, Some results on Tsallis entropy measure and k-record values, *Physica A* (2016), on line.
- entropię Tsalisa definiuje się przez

$$H_{\alpha}(X) = \frac{1}{1-\alpha}\left\{\int_0^{\infty} f^{\alpha}(x)\,dx - 1\right\},\ \alpha \neq 1, \alpha > 0,$$

gdzie f jest gęstością zmiennej losowej X.

W pracy

- V. Kumar, Some results on Tsallis entropy measure and k-record values, *Physica A* (2016), on line.
- entropię Tsalisa definiuje się przez

$$H_\alpha(X) = \frac{1}{1-\alpha}\left\{\int_0^\infty f^\alpha(x)\,dx - 1\right\},\ \alpha \neq 1, \alpha > 0,$$

  gdzie f jest gęstością zmiennej losowej X.
- Wyznaczono tam entropię Tsallisa dla k-ch wartości rekordowych z prób z wybranych rozkładów prawdopodobieństwa.

W pracy

W pracy

- D.-T. Kang, L. Yan, On the dynamic cumulative residual quantile entropy ordering, *Statistical Methodology* **32** (2016), 14-35.

W pracy

- D.-T. Kang, L. Yan, On the dynamic cumulative residual quantile entropy ordering, *Statistical Methodology* **32** (2016), 14-35.
- zdefiniowano nowy stochastyczny porządek w oparciu o entropię DCRQE [*dynamic cumulative residual quantile entropy*]

$$\varphi(p) = -\int_{F^{\leftarrow}(p)}^{\infty} \frac{1-F(x)}{1-p} \ln\left[\frac{1-F(x)}{1-p}\right] dx, \;\; p \in (0,1),$$

gdzie F jest dystrybuantą nieujemnej zmiennej losowej X.

W pracy

- D.-T. Kang, L. Yan, On the dynamic cumulative residual quantile entropy ordering, *Statistical Methodology* **32** (2016), 14-35.
- zdefiniowano nowy stochastyczny porządek w oparciu o entropię DCRQE [*dynamic cumulative residual quantile entropy*]

$$\varphi(p) = -\int_{F^{\leftarrow}(p)}^{\infty} \frac{1-F(x)}{1-p} \ln\left[\frac{1-F(x)}{1-p}\right] dx, \;\; p \in (0,1),$$

  gdzie F jest dystrybuantą nieujemnej zmiennej losowej X.
- Rozważano tam własności zachowania tego porządku w modelach k-tych wartości rekordowych.