Estymacja bayesowska parametrów pewnych rozkładów ze zniekształceniem i deformacją

Małgorzata Murat, Politechnika Lubelska

XXXV Konferencja Statystyka Matematyczna
Wisła 2009

PLAN REFERATU

- Wstęp
  - Rozważane rozkłady:
    ✓ zniekształcony, ✓ zdeformowany.
  - Rozkłady a priori
    - dla parametru zniekształcenia i deformacji: beta, dwustronny potęgowy
    - dla parametru rozkładu: beta, gamma, Pareto
  - Niesymetryczne funkcje straty:
    - liniowo-wykładnicza
    - uogólniona entropijna
- Rozkłady i momenty a posteriori parametrów rozkładu
  ✓ zniekształconego, ✓ zdeformowanego
- Przykłady
  - Uogólniony rozkład ujemnie dwumianowy.
  - Uogólniony rozkład Poissona.
- Podsumowanie.

- *funkcja prawdopodobieństwa*

$$P[X = x] = \frac{a(x)[g(\theta)]^x}{f(\theta)}$$

✓ $g(\theta)$-nieujemna, skończona, różniczkowalna funkcja

✓ $f(\theta) = \sum_x a(x)[g(\theta)]^x$,

✓ współczynniki $a(x)$ niezależne od parametru θ

◇ Gupta R.C., (1974), *Modified power series distribution and some of its applications*, Sankhyā, ser.B., vol. 36, 288-298.

- *Funkcja prawdopodobieństwa zmodyfikowanego rozkładu potęgowo szeregowego ze zniekształceniem w punkcie s*

$$P[X=x]=\begin{cases}\beta+\alpha\frac{a(s)[g(\theta)]^s}{f(\theta)}, & x=s,\\ \alpha\frac{a(x)[g(\theta)]^x}{f(\theta)}, & x\neq s,\end{cases}$$

- ✓ $g(\theta)$-nieujemna, skończona, różniczkowalna funkcja
- ✓ $f(\theta)=\sum_x a(x)[g(\theta)]^x$,
- ✓ współczynniki $a(x)$ niezależne od parametru θ
- ✓ $0\leqslant\alpha\leqslant 1$, $\beta=1-\alpha$

◇ Pandey, K.N., (1964-65), *Generalized inflated Poisson distribution*, J. Scienc. Res. Banares Hindu Univ., XV(2) , 157-162.

- *Funkcja prawdopodobieństwa zmodyfikowanego rozkładu potęgowo szeregowego z deformacją w punkcie s*

$$P[X=x]=\begin{cases}\left[1+\alpha g(\theta)\frac{a(s+1)}{a(s)}\right]\frac{a(s)[g(\theta)]^s}{f(\theta)}, & x=s,\\ \beta\,\frac{a(s+1)[g(\theta)]^{s+1}}{f(\theta)}, & x=s+1,\\ \frac{a(x)[g(\theta)]^x}{f(\theta)}, & x=0,1,...,s-1,s+2,...\end{cases}$$

- ✓ $g(\theta)$-nieujemna, skończona, różniczkowalna funkcja
- ✓ $f(\theta)=\sum_x a(x)[g(\theta)]^x$,
- ✓ współczynniki $a(x)$ niezależne od parametru θ
- ✓ $0\leqslant\alpha\leqslant 1$, $\beta=1-\alpha$

◇ Cohen, A. C., (1960), *Estimating the parameters of a modified Poisson distribution*, Amer. Statist. Assoc. Journal, 55, 139-143.

◇ Cohen, A. C., (1960), *Estimation in the Poisson distribution when sample values of $c+1$ are sometimes erreoneously reported as c*, Ann. Inst. Stat. Math., 9, 189-193.

Rozkłady a priori dla parametru α

(A) *rozkład Beta z parametrami* $u > 0$, $v > 0$

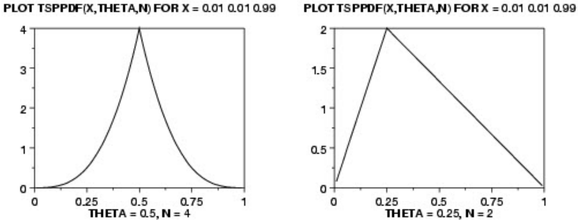
$$\varphi(\alpha|u,v) = \frac{\alpha^u(1-\alpha)^v}{B(u+1,v+1)}, \quad 0 < \alpha < 1,$$

(B) *dwustronny rozkład potęgowy z parametrami* $0 < \theta < 1$, $n > 0$

$$\varphi(\alpha|\theta,n) = \begin{cases} n\left(\frac{\alpha}{\theta}\right)^{n-1}, & 0 < \alpha \leqslant \theta, \\ n\left(\frac{1-\alpha}{1-\theta}\right)^{n-1}, & \theta < \alpha \leqslant 1, \end{cases}$$

- S. Kotz, J. R. van Dorp, "Beyond beta", World Scientific Publishing Co. Pte. Ltd. (2004)

www.vosesoftware.com



Rozkłady a priori dla parametru rozkładu θ

(I) *rozkład Beta z parametrami* $c > 0$, $d > 0$

(II) *rozkład Gamma z parametrami* $c > -1$, $d > 0$

$$\psi(\theta|c,d) = \frac{d^{c+1}\theta^c e^{-\theta d}}{\Gamma(c+1)}, \quad \theta > 0$$

(III) *uogólniony rozkład Pareto*

$$\psi(\theta \mid b) = \begin{cases} (1+b\theta)^{-1-\frac{1}{b}}, & \theta \geqslant 0, b > 0, \\ e^{-\theta}, & \theta \geqslant 0, b = 0 \\ (1+b\theta)^{-1-\frac{1}{b}}, & \theta \in (0, -\frac{1}{b}), b < 0. \end{cases}$$

(IV) *rozkład sprzężony (conjugate prior) z parametrami* c, d

$$\psi(\theta|c,d) = \frac{[g(\theta)]^c[f(\theta)]^{-d}}{D(c,d)}, \quad \text{gdzie } D(c,d) < \infty.$$

- *liniowo-wykładnicza funkcja straty*

$$L(\theta, \widehat{\theta}) = \gamma[\exp(\delta(\theta - \widehat{\theta})) - \delta(\theta - \widehat{\theta}) - 1], \text{ gdzie } \quad \delta \neq 0, \gamma > 0,$$

estymator bayesowski $\widehat{\theta}$ parametru θ

$$\widehat{\theta} = -\frac{1}{\delta} \ln \left[\sum_{r=0}^{\infty} \frac{(-\delta)^r}{r!} \mathbf{E}(\theta^r | \mathbf{x}) \right],$$

◇ Zellner A., (1986), *Bayesian estimation and prediction using asymmetric loss function*, J. Amer. Statist. Assoc. 81, 446-451.

Funkcje straty

- *uogólniona entropijna funkcja straty*

$$L(\theta, \theta^*) = \gamma \left[\left(\frac{\theta}{\theta^*} \right)^q - q \ln \left(\frac{\theta}{\theta^*} \right) - 1 \right], \text{ gdzie}, \gamma > 0,$$

estymator bayesowski θ^* parametru θ

$$\theta^* = \left[\mathbf{E}(\theta^{-q} | \mathbf{x}) \right]^{\frac{1}{q}}, \quad q > 0$$

◇ Soliman, A. A. (2006), *Estimators for the finite fixture of Rayleigh model based on progressively censored data*, Comm. Stat.-Theor. Meth. 35, 803-820.

Funkcja wiarogodności rozkładu zniekształconego

- ✓ dla $i = 1, 2, \ldots, m$ mamy n_i obserwacji równych x_i i różnych od s,
- ✓ mamy n_s obserwacji równych s,
- ✓ $N = \sum_{i=1}^{m} n_i + n_s$, $T = \sum_{i=1}^{N-n_s} x_i + sn_s$
- ✓ $\Psi(\psi(\theta); x, y) := \int\limits_{\Theta} \psi(\theta)[g(\theta)]^x [f(\theta)]^{-y} d\theta$
- ✓ $D(x, y) := \Psi(1; x, y) = \int\limits_{\Theta} [g(\theta)]^x [f(\theta)]^{-y} d\theta.$

- *Funkcja wiarogodności dla rozkładu zniekształconego w punkcie s*

$$L(\theta, \alpha | \mathbf{x}) = \prod_{i=1}^{N-n_s} a(x_i) \sum_{j=0}^{n_s} \binom{n_s}{j} \beta^j \alpha^{N-j} \frac{[a(s)]^{n_s-j}}{[f(\theta)]^{N-j}} [g(\theta)]^{T-sj}.$$

Rozkład a priori zniekształcenia ∼ beta (u, v)

- *funkcja prawdopodobieństwa rozkładu a posteriori*

$$h(\alpha,\theta|\mathbf{x}) = k_A(u,v;\psi(\theta)) \times \sum_{j=0}^{n_s} \binom{n_s}{j} [a(s)]^{-j} \alpha^{u+N-j} \beta^{v+j} \psi(\theta) [g(\theta)]^{T-sj} [f(\theta)]^{j-N}$$

$$[k_A(u,v;\psi(\theta))]^{-1} = \sum_{j=0}^{n_s} \binom{n_s}{j} [a(s)]^{-j} B(u+N-j+1, v+j+1) \Psi(\psi(\theta); T-sj, N-j)$$

- *momenty a posteriori*

$$\mathbf{E}(\alpha^l \theta^r|\mathbf{x}) = k_A(u,v;\psi(\theta)) \times \sum_{j=0}^{n_s} \binom{n_s}{j} [a(s)]^{-j} B(l+u+N-j+1, v+j+1) \Psi(\theta^r\psi(\theta); T-sj, N-j)$$

- *funkcja prawdopodobieństwa rozkładu a posteriori*

$$h(\alpha, \theta|\mathbf{x}) =$$
$$\begin{cases} k_B(u, v; \psi(\theta)) \sum_{j=0}^{n_s} \binom{n_s}{j} [a(s)]^{-j} \dfrac{\alpha^{v-1+N-j}\beta^j}{u^{1-v}} \dfrac{\psi(\theta)[g(\theta)]^{T-sj}}{[f(\theta)]^{N-j}}, & 0 < \alpha \leqslant u, \\ k_B(u, v; \psi(\theta)) \sum_{j=0}^{n_s} \binom{n_s}{j} [a(s)]^{-j} \dfrac{\alpha^{N-j}\beta^{v-1+j}}{(1-u)^{1-v}} \dfrac{\psi(\theta)[g(\theta)]^{T-sj}}{[f(\theta)]^{N-j}}, & u < \alpha \leqslant 1, \end{cases}$$

$$[k_B(u, v; \psi(\theta))]^{-1} = \sum_{j=0}^{n_s} \binom{n_s}{j} [a(s)]^{-j} \Psi(\psi(\theta); T - sj, N - j)$$
$$\times \left[u^{1-v} B_u(v + N - j, j + 1) + (1 - u)^{1-v} \widetilde{B_u}(N - j + 1, v + j)\right]$$

- *momenty a posteriori*

$$\mathbf{E}(\alpha^l \theta^r|\mathbf{x}) = k_B(u, v; \psi(\theta)) \sum_{j=0}^{n_s} \binom{n_s}{j} [a(s)]^{-j} \Psi(\theta^r \psi(\theta); T - sj, N - j)$$
$$\times \left[u^{1-v} B_u(l + v + N - j, j + 1) + (1 - u)^{1-v} \widetilde{B_u}(l + N - j + 1, v + j)\right].$$

Uogólniony rozkład ujemnie dwumianowy z parametrami θ, γ, m

- *funkcja prawdopodobieństwa*

$$P[X=x]=\frac{m\Gamma(m+\gamma x)[\theta(1-\theta)^{\gamma-1}]^{x}}{x!\Gamma(m+\gamma x-x+1)(1-\theta)^{-m}},$$

gdzie $0<\theta<1, \mid \theta\gamma \mid < 1$

- $a(x)= \frac{m\Gamma(m+\gamma x)}{x!\Gamma(m+\gamma x-x+1)}$
- $f(\theta)= (1-\theta)^{-m}$
- $g(\theta)= \theta(1-\theta)^{\gamma-1}$
- $\gamma=0$ - rozkład ujemnie dwumianowy
- $\gamma=1$ - rozkład dwumianowy

Model kolejkowy

- Consul P. C., Famoye F., "Lagrangian Probability Distributions", Birkhäuser, (2006)

- ✓ klienci przybywają do jednoosobowego punktu obsługi w grupach k
- ✓ k jest zmienną losową o rozkładzie dwumianowym z parametrami θ, γ
- ✓ podczas obsługi przybyłych k klientów przybywają kolejni klienci zgodnie z rozkładem dwumianowym z parametrami θ, m
- rozkład liczby klientów obsłużonych w czasie pierwszego nasilenia (firts busy period) ma uogólniony rozkład ujemnie dwumianowy z parametrami θ, γ, m

- ✓ rozkład a priori parametru $\alpha \sim$ dwustronny potęgowy (u, v),
- ✓ rozkład a priori parametru $\theta \sim$ beta (c, d),
- ✓ funkcja straty - liniowo-wykładnicza

$$\widehat{\theta} = -\frac{1}{\delta}\ln\left[\sum_{r=0}^{\infty}\frac{(-\delta)^r}{r!}\frac{\sum_{j=0}^{n_s}\binom{n_s}{j}A_j\mathcal{U}_j(0, c+r, d, c, d)}{\sum_{j=0}^{n_s}\binom{n_s}{j}A_j\mathcal{U}_j(0, 0, d, 0, 0)}\right],$$

$$\widehat{\alpha} = -\frac{1}{\delta}\ln\left[\sum_{l=0}^{\infty}\frac{(-\delta)^l}{l!}\frac{\sum_{j=0}^{n_s}\binom{n_s}{j}A_j\mathcal{U}_j(l, c, d, c, d)}{\sum_{j=0}^{n_s}\binom{n_s}{j}A_j\mathcal{U}_j(0, 0, d, 0, 0)}\right].$$

gdzie $A_j = \left[\frac{s!\Gamma(m+\gamma s-s+1)}{m\Gamma(m+\gamma s)}\right]^j$ oraz $\mathcal{U}_j(l, x, y, z, q) =$

$$\left[u^{1-v}B_u(l+v+N-j, j+1) + (1-u)^{1-v}\widetilde{B_u}(l+N-j+1, v+j)\right]$$
$$\times B(x+c+T-sj+1, y+(\gamma-1)(z+T-sj)+m(N-j+q)+1),$$

- *funkcja prawdopodobieństwa*

$$P[X = x] = e^{-\theta(1+\gamma x)} \frac{\theta^x (1+\gamma x)^{x-1}}{x!},$$

gdzie $\mid \theta\gamma \mid < 1$.

◇ Consul P. C., Famoye F., "Lagrangian Probability Distributions", Birkhäuser, (2006)

- $a(x) = \frac{(1+\gamma x)^{x-1}}{x!}$
- $f(\theta) = e^{\theta}$
- $g(\theta) = \theta e^{-\gamma\theta}$
- $\gamma = 0$ - rozkład Poissona

- rozkład graniczny uogólnionego rozkładu ujemnie dwumianowego z parametrami p, β, m, gdy m i β są tak duże, a p jest tak małe, że $mp = \theta$ i $\beta p = \gamma$

- ✓ rozkład a priori parametru $\alpha \sim$ dwustronny potęgowy (u, v),
- ✓ rozkład a priori parametru $\theta \sim$ beta (c, d),
- ✓ funkcja straty - uogólniona entropijna

$$\theta^* = \left[\frac{\sum_{j=0}^{n_s} \binom{n_s}{j} C_j \mathcal{W}_j(0, 2c + r, c, 2d)}{\sum_{j=0}^{n_s} \binom{n_s}{j} C_j \mathcal{W}_j(0, c, c, 2d)} \right]^{\frac{1}{q}},$$

$$\alpha^* = \left[\frac{\sum_{j=0}^{n_s} \binom{n_s}{j} C_j \mathcal{W}_j(l, 2c, c, 2d)}{\sum_{j=0}^{n_s} \binom{n_s}{j} C_j \mathcal{W}_j(0, c, c, 2d)} \right]^{\frac{1}{q}},$$

gdzie $C_j = \left[\frac{s!}{(1+\gamma s)^{s-1}} \right]^j$ oraz $\mathcal{W}_j(l, x, y, z) =$

$$\left[u^{1-v} B_u(l + v + N - j, j + 1) + (1-u)^{1-v} \widetilde{B_u}(l + N - j + 1, v + j) \right]$$

$$\times \left(\frac{x + T - sj}{\gamma(y + T - sj) + N - j + z} \right)^{x+T-sj+1} \Gamma(x + T - sj + 1)$$

- ✓ dla $i = 1, 2, \ldots, m$ mamy n_i obserwacji równych x_i i nierównych s oraz $s+1$
- ✓ mamy n_s obserwcji równych s
- ✓ mamy n_{s+1} obserwcji równych $s+1$
- ✓ $N = \sum_{i=1}^{m} n_i + n_s + n_{s+1}$, $T = \sum_{i=1}^{N-n_s-n_{s+1}} x_i + sn_s + (s+1)n_{s+1}$

- *Funkcja wiarogodności dla rozkładu zdeformowanego*

$$L(\theta, \alpha|\mathbf{x}) = \prod_{i=1}^{N} a(x_i) \sum_{j=0}^{n_s} \binom{n_s}{j} A_s(j) \alpha^{n_s-j} \beta^{n_s+1} \frac{[g(\theta)]^{T+n_s-j}}{[f(\theta)]^N},$$

gdzie $A_s(j) = \left[\frac{a(s+1)}{a(s)}\right]^{n_s-j}$.

Rozkład a priori deformacji $\alpha \sim$ beta (u, v)

- *funkcja prawdopodobieństwa rozkładu a posteriori*

$$h(\alpha,\theta|\mathbf{x}) = k_A(u,v;\psi(\theta)) \times \sum_{j=0}^{n_s} \binom{n_s}{j} A_s(j)\alpha^{u+n_s-j}\beta^{v+n_{s+1}}\psi(\theta)[g(\theta)]^{T+n_s-j}[f(\theta)]^{-N}$$

$$[k_A(u,v;\psi(\theta))]^{-1} = \sum_{j=0}^{n_s} \binom{n_s}{j} A_s(j) B(u+n_s-j+1, v+1+n_{s+1})\ \Psi(\psi(\theta);T+n_s-j,N)$$

- *momenty rozkładu a posteriori*

$$\mathbf{E}(\alpha^l\theta^r|\mathbf{x}) = k_A(u,v;\psi(\theta)) \sum_{j=0}^{n_s} \binom{n_s}{j} A_s(j) \times B(l+u+n_s-j+1, v+1+n_{s+1})\Psi(\theta^r\psi(\theta);T+n_s-j,N)$$

- *funkcja prawdopodobieństwa rozkładu a posteriori:* $h(\alpha,\theta|\mathbf{x}) =$

$$\begin{cases} k_B(u,v;\psi(\theta)) \sum_{j=0}^{n_s} \binom{n_s}{j} [a(s)]^{-j} \frac{\alpha^{v-1+N-j}\beta^j}{u^{1-v}} \frac{\psi(\theta)[g(\theta)]^{T-sj}}{[f(\theta)]^{N-j}}, & 0<\alpha\leqslant u, \\ k_B(u,v;\psi(\theta)) \sum_{j=0}^{n_s} \binom{n_s}{j} [a(s)]^{-j} \frac{\alpha^{N-j}\beta^{v-1+j}}{(1-u)^{1-v}} \frac{\psi(\theta)[g(\theta)]^{T-sj}}{[f(\theta)]^{N-j}}, & u<\alpha\leqslant 1, \end{cases}$$

$$[k_B(u,v;\psi(\theta))]^{-1} = \sum_{j=0}^{n_s} \binom{n_s}{j} A_s(j)\Psi(\psi(\theta); T+n_s-j, N)$$

$$\times \left[u^{1-v} B_u(v+n_s-j, 1+n_{s+1}) + (1-u)^{1-v}\widetilde{B_u}(n_s-j+1, v+n_{s+1})\right]$$

- *momenty rozkładu a posteriori:*

$$\mathbf{E}(\alpha^l\theta^r|\mathbf{x}) = k_B(u,v;\psi(\theta)) \sum_{j=0}^{n_s} \binom{n_s}{j} [a(s)]^{-j} \Psi(\theta^r\psi(\theta); T-sj, N-j)$$

$$\times \left[u^{1-v} B_u(l+v+N-j, j+1) + (1-u)^{1-v}\widetilde{B_u}(l+N-j+1, v+j)\right].$$

Estymatory bayesowskie parametrów uogólnionego rozkładu ujemnie dwumianowego zdeformowanego w punkcie s

- ✓ rozkład a priori parametru $\alpha \sim \mathrm{beta}(u, v)$,
- ✓ rozkład a priori parametru $\theta \sim$ sprzężony (c, d),
- ✓ uogólniona entropijna funkcja straty

$$\theta^* = \left[\sum_{r=0}^{\infty} \frac{(-\delta)^r}{r!} \frac{\sum_{j=0}^{n_s} \binom{n_s}{j} A_j \mathcal{M}_j(0, -q, c, dm+1)}{\sum_{j=0}^{n_s} \binom{n_s}{j} A_j \mathcal{M}_j(0, 0, c, dm+1)} \right]^{\frac{1}{q}},$$

$$\alpha^* = \left[\sum_{l=0}^{\infty} \frac{(-\delta)^l}{l!} \frac{\sum_{j=0}^{n_s} \binom{n_s}{j} A_j \mathcal{M}_j(-q, 0, c, dm+1)}{\sum_{j=0}^{n_s} \binom{n_s}{j} A_j \mathcal{M}_j(0, 0, c, dm+1)} \right]^{\frac{1}{q}}.$$

gdzie $A_j = \left[\frac{s!\Gamma(m+\gamma s-s+1)}{m\Gamma(m+\gamma s)}\right]^{n_s-j}$ oraz

$$\mathcal{M}_j(l, r, x, y) = B(r+c+T+n_s-j, (\gamma-1)(x+T+n_s-j)+mN+y)$$
$$\times \left[u^{1-v} B_u(l+v+n_s-j, 1+n_{s+1}) + (1-u)^{1-v} \widetilde{B_u}(l+n_s-j+1, v+n_{s+1})\right]$$

Wstęp
Rozważane rozkłady
Rozkłady a priori
Funkcje straty
Rozkłady i momenty a posteriori rozkładu zniekształconego
Przykłady
Rozkłady i momenty a posteriori rozkładu zdeformowanego
Przykłady
Podsumowanie

- ✓ rozkład a priori parametru $\alpha \sim$ beta (u, v)
- ✓ rozkład a priori parametru $\theta \sim$ sprzężony (c, d)
- ✓ funkcja straty - liniowo-wykładnicza

$$\widehat{\theta} = \left[\sum_{r=0}^{\infty} \frac{(-\delta)^r}{r!} \frac{\sum_{j=0}^{n_s} \binom{n_s}{j} C_j \mathcal{Q}_j(0, c-q, c, d)}{\sum_{j=0}^{n_s} \binom{n_s}{j} C_j \mathcal{Q}_j(0, c, c, d)} \right]^{\frac{1}{q}},$$

$$\widehat{\alpha} = \left[\sum_{l=0}^{\infty} \frac{(-\delta)^l}{l!} \frac{\sum_{j=0}^{n_s} \binom{n_s}{j} C_j \mathcal{Q}_j(-q, c, c, d)}{\sum_{j=0}^{n_s} \binom{n_s}{j} C_j \mathcal{Q}_j(0, c, c, d)} \right]^{\frac{1}{q}}.$$

gdzie $C_j = \left[\frac{s!}{(1+\gamma s)^{s-1}} \right]^{n_s - j}$

$$\mathcal{Q}_j(l, x, y, z) = B(l + u + n_s - j + 1, v + 1 + n_{s+1}) \times \left(\frac{x + T + n_s - j}{\gamma(y + T + n_s - j) + N + z} \right)^{x+T+n_s-j+1} \Gamma(x + T + n_s - j + 1)$$

PODSUMOWANIE

- Estymatory bayesowskie rozkładu:
  ✓ zniekształconego, ✓ zdeformowanego
- Rozkłady a priori:
  ✓ parametr zniekształcenia i deformacji: beta, dwustronny potęgowy,
  ✓ parametr rozkładu: beta, gamma, Pareto, sprzężony
- Funkcje straty: ✓ liniowo-wykładnicza, ✓ uogólniona entropijna
- Uogólnienie i rozszerzenie wyników:

✓ Bansal, A. K., Aggarwal, P. (2007), *Bayesian estimators for the zero inflated modified power series distribution*, Amer. J. Math. Man. Sci., vol. 27, Nos. 1 & 2, 5-23.

✓ Murat, M., Szynal, D. (1999), *Posterior distributions and moments of deformed modified power series distributions*, Disc. Math. Prob. and Stat. 19, 287-297.