

Statystyka Matematyczna

Anna Janicka

wykład XIII, 30.05.2016

STATYSTYKA BAYESOWSKA

Plan na dzisiaj

1. Statystyka Bayesowska

- rozkłady a priori i a posteriori
- estymacja Bayesowska:
 - Bayesowski Estymator Największej Wiarogodności (BENW)
 - Estymator Bayesowski przy zadanej funkcji straty



Statystyka Bayesowska vs. tradycyjna

Tradycyjna: nieznane parametry są ustalone,
dane (obserwacje) są zmiennymi losowymi

Bayesowska: dane (obserwacje) są ustalone,
nieznane parametry są losowe



Statystyka Bayesowska

Różnica w stosunku do „tradycyjnej” statystyki i metod wnioskowania: naszą wiedzę o (nieznanych) parametrach opisujemy przy pomocy **rozkładów prawdopodobieństwa**, przy czym dodatkowa wiedza może wpływać na nasz opis. Wiedza:

- ☐ ogólna
- ☐ dotycząca konkretnego przypadku

Przykład: Rzut monetą



Model Bayesowski

- zmienne $X_1, \dots, X_n$, pochodzące z rozkładu P_θ , o gęstości $f_\theta(\mathbf{x})$ – gęstość warunkowa pod warunkiem konkretnej wartości θ (funkcja wiarygodności).
- P – rodzina rozkładów prawdopodobieństw P_θ , indeksowana parametrem $\theta \in \Theta$
- Wiedza ogólna: rozkład p-stwa Π na przestrzeni parametrów Θ , zadany przez $\pi(\theta)$ – tzw. rozkład **a priori** parametru θ ,

Model Bayesowski – cd.

Wiedza dodatkowa, szczególna, kontekstualna: wynika z obserwacji. Mamy rozkład łączny obserwacji i parametru θ :

$$f(x_1, x_2, \dots, x_n, \theta) = f(x_1, x_2, \dots, x_n \mid \theta) \pi(\theta)$$

z którego możemy wyznaczyć warunkowy rozkład dla θ (po uwzględnieniu obserwacji)

$$\pi(\theta \mid x_1, \dots, x_n) = \frac{f(x_1, \dots, x_n \mid \theta) \pi(\theta)}{m(x_1, \dots, x_n)},$$

gdzie

$$m(x_1, \dots, x_n) = \int_{\Theta} f(x_1, \dots, x_n \mid \theta) \pi(\theta) d\theta$$



Model Bayesowski – rozkład a posteriori

Rozkład $\pi(\theta \mid x_1, \dots, x_n)$ nazywany jest rozkładem **a posteriori**, ozn. Π_x

Rozkład a posteriori obrazuje całą wiedzę: ogólną (wstępną) oraz szczególną (wynikającą z konkretnych danych).

Jest podstawą wnioskowania Bayesowskiego



Rozkłady a priori i a posteriori: przykłady

1. Niech: $X_1, \dots, X_n$ IID z rozkładu 0-1 z p-stwem sukcesu θ , niech

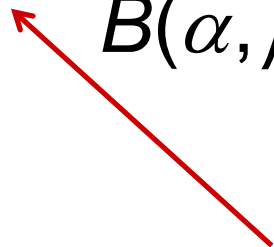
$$\pi(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha, \beta)}$$

dla $\theta \in (0, 1)$

gdzie $B(\alpha, \beta) = \int_0^1 u^{\alpha-1}(1-u)^{\beta-1} du = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$

oraz $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} \exp(-u) du = (\alpha-1)\Gamma(\alpha-1)$

rozkład
Beta(α, β)

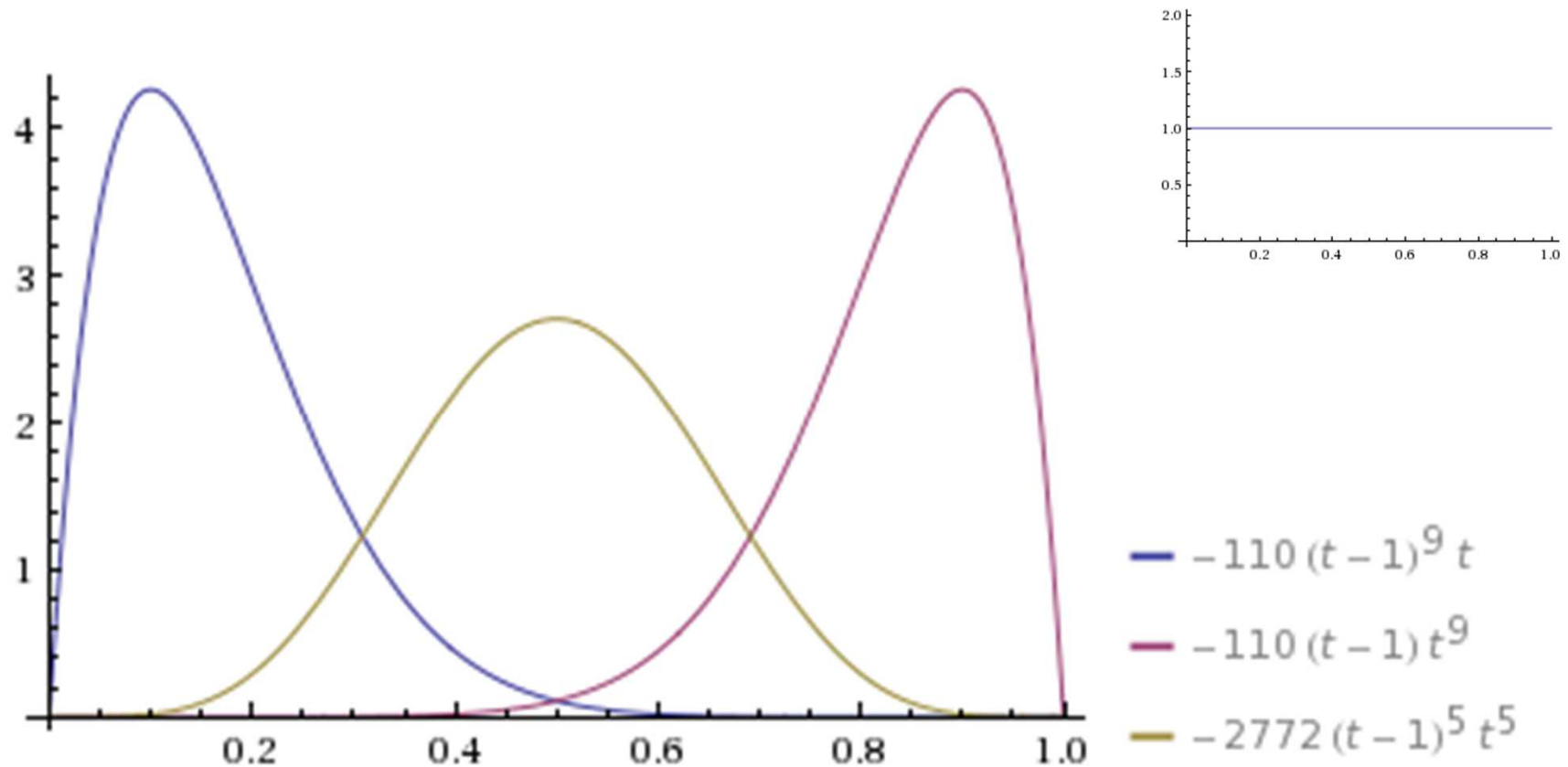


wówczas rozkład a posteriori:

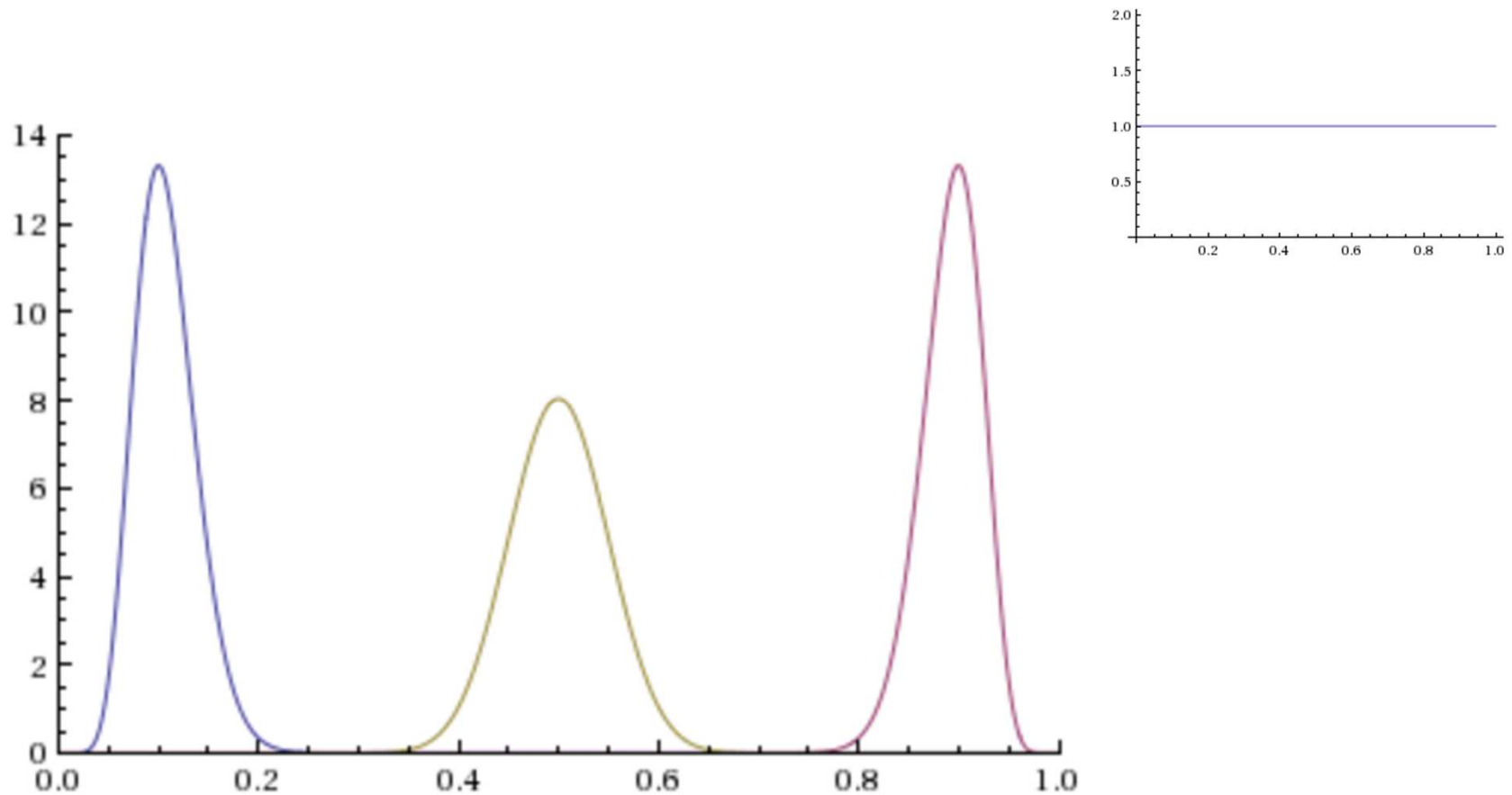
$$\text{Beta}\left(\sum_{i=1}^n x_i + \alpha, n - \sum_{i=1}^n x_i + \beta\right)$$



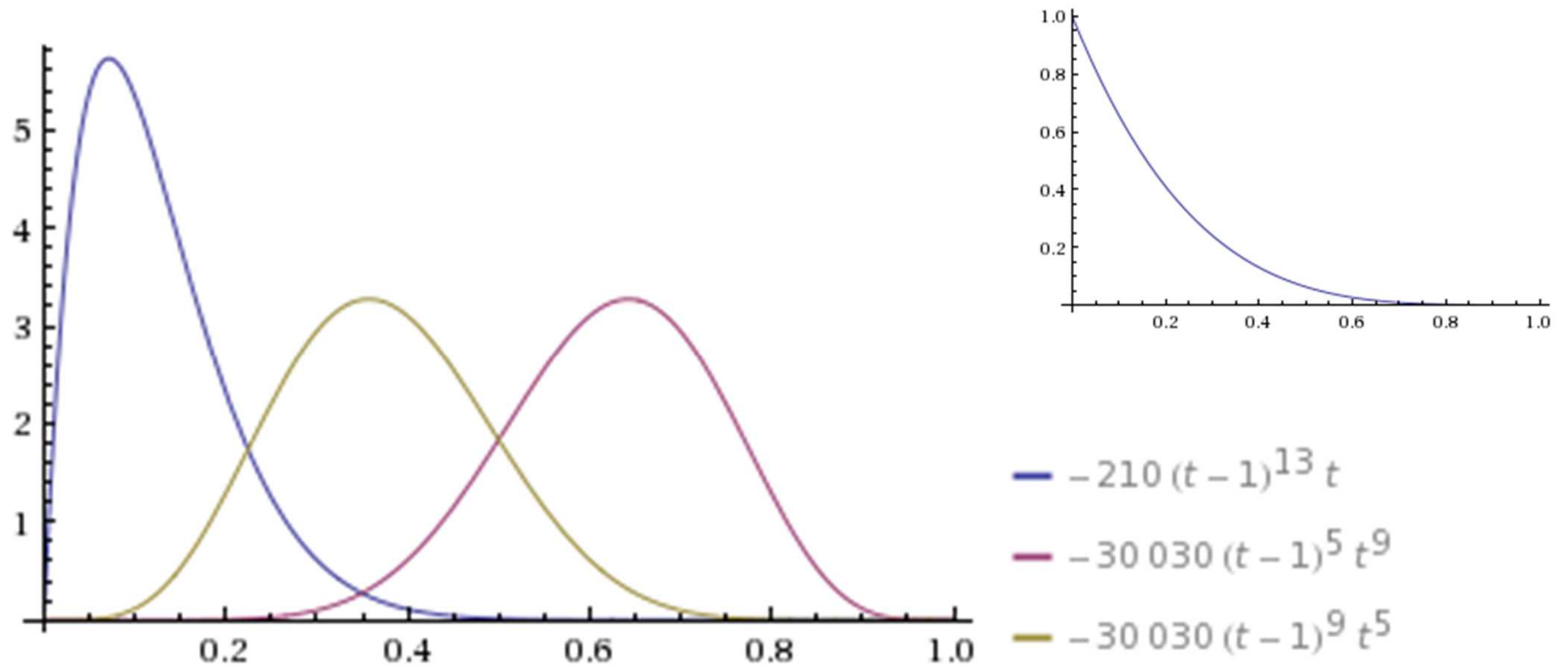
Dla Beta (1,1) a priori i danych: $n=10$ i 1, 5, 9 sukcesów



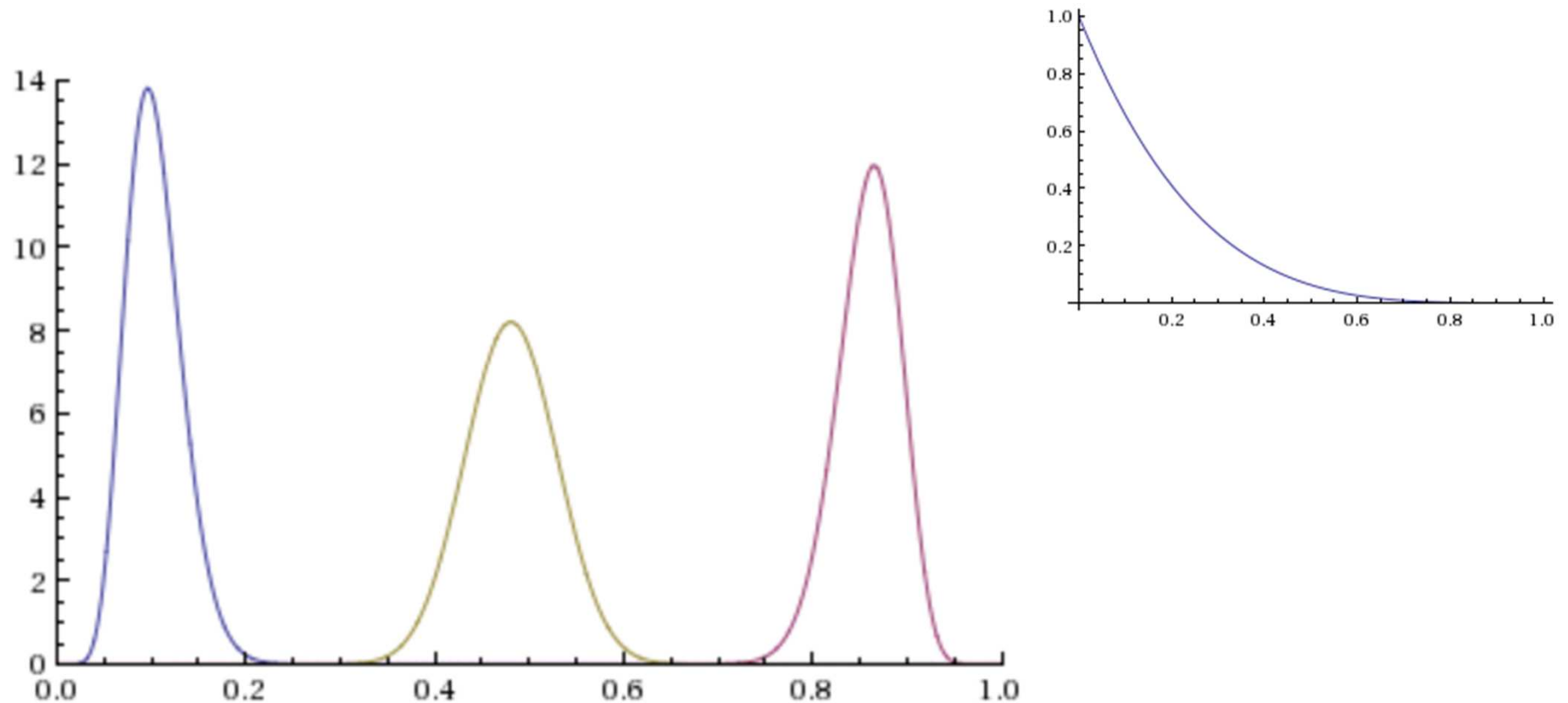
Dla Beta (1,1) a priori i danych: $n=100$ i 10, 50, 90 sukcesów



Dla Beta (1,5) a priori i danych: $n=10$ i 1, 5, 9 sukcesów



Dla Beta (1,5) a priori i danych: $n=100$ i 10, 50, 90 sukcesów



Rozkłady a priori i a posteriori: przykłady (2)

2. Niech: $X_1, \dots, X_n$ IID z rozkładu $N(\theta, \sigma^2)$, przy czym σ^2 znane; $\theta \sim N(m, \tau^2)$ dla pewnych m, τ znanych.

wówczas rozkład a posteriori dla θ :

$$N\left(\frac{n \frac{1}{\sigma^2} \bar{X} + \frac{1}{\tau^2} m}{n \frac{1}{\sigma^2} + \frac{1}{\tau^2}}, \frac{1}{n \frac{1}{\sigma^2} + \frac{1}{\tau^2}}\right)$$



Statystyka Bayesowska

Korzystając z modeli Bayesowskich, możemy

- ☐ wyznaczać estymatory
- ☐ wyznaczać odpowiedniki przedziału ufności
- ☐ weryfikować hipotezy

- ☐ dokonywać predykcji



Bayesowski Estymator Największej Wiarygodności (BENW)

Wyznaczamy tak, by maksymalizować prawdopodobieństwo a posteriori (moda rozkładu):

$$\pi(\hat{\theta}_{BENW} \mid x_1, \dots, x_n) = \max_{\theta} \pi(\theta \mid x_1, \dots, x_n)$$

czyli

$$BENW(\theta) = \hat{\theta}_{BENW} = \operatorname{argmax}_{\theta} \pi(\theta \mid x_1, \dots, x_n)$$



BENW: przykłady

1. Niech: $X_1, \dots, X_n$ IID z rozkładu 0-1 z p-stwem sukcesu θ ; niech dla $\theta \in (0,1)$ $\pi(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha, \beta)}$ znamy rozkład a posteriori:

Beta($\sum_{i=1}^n x_i + \alpha, n - \sum_{i=1}^n x_i + \beta$)
max dla

$$BENW(\theta) = \frac{\sum_{i=1}^n x_i + \alpha - 1}{n + \beta + \alpha - 2}$$

rozkład Beta(α, β);
jego moda
= $(\alpha-1)/(\alpha + \beta - 2)$
dla $\alpha > 1, \beta > 1$

np. dla 5 sukcesów zaobserwowanych w 10 próbach i dla rozkładu a priori $U(0,1)$ (czyli Beta(1,1)), mamy $BENW(\theta) = 5/10 = 1/2$

a dla 9 sukcesów zaobserwowanych w 10 próbach i tego samego rozkładu a priori, mamy $BENW(\theta) = 9/10$



BENW: przykłady (2)

2. Niech: $X_1, \dots, X_n$ IID z rozkładu $N(\theta, \sigma^2)$, przy czym σ^2 znane; $\theta \sim N(m, \tau^2)$ dla pewnych m, τ znanych.

Wówczas rozkład a posteriori dla θ :

a zatem

$$BENW(\theta) = \frac{n \frac{1}{\sigma^2} \bar{X} + \frac{1}{\tau^2} m}{n \frac{1}{\sigma^2} + \frac{1}{\tau^2}}$$

$$N\left(\frac{n \frac{1}{\sigma^2} \bar{X} + \frac{1}{\tau^2} m}{n \frac{1}{\sigma^2} + \frac{1}{\tau^2}}, \frac{1}{n \frac{1}{\sigma^2} + \frac{1}{\tau^2}}\right)$$

np. jeśli próba 5 obserwacji 1.2; 1.7 ; 1.9 ; 2.1; 3.1 z rozkładu $N(\theta, 4)$ a rozkład a priori $\theta \sim N(1, 1)$, to

$$BENW(\theta) = (5/4 * 2 + 1)/(5/4 + 1) = 14/9 \approx 1.56$$

a gdyby rozkład a priori $\theta \sim N(3, 1)$, to

$$BENW(\theta) = (5/4 * 2 + 1*3)/(5/4 + 1) = 22/9 \approx 2.44$$



Estymator Bayesowski przy zadanej funkcji straty

$L(\theta, a)$ – **funkcja straty** zależna od prawdziwej wartości parametru θ oraz decyzji a .

np. jeśli estymujemy wielkość $g(\theta)$:

$L(\theta, a) = (g(\theta) - a)^2$ – kwadratowa funkcja straty

$L(\theta, a) = |g(\theta) - a|$ – modułowa funkcja straty

Definiujemy też **ryzyko a posteriori**:

$$R(\Pi, \hat{g}(x)) = E(L(\theta, \hat{g}(x)) | X = x) = \int_{\Theta} L(\theta, \hat{g}(x)) \pi(\theta | x) d\theta$$

(średnia strata estymatora przy ustalonym rozkładzie a priori i danych, tj. przy wyliczonym rozkładzie a posteriori)



Estymator Bayesowski przy zadanej funkcji straty – cd.

Estymator Bayesowski przy danej funkcji straty $L(\theta, a)$ to $\hat{g}_B$ t. że

$$\forall x \quad R(\Pi, \hat{g}_B(x)) = \min_a R(\Pi, a)$$

Przy kwadratowej funkcji straty $(\theta - a)^2$:

$$\hat{\theta}_B = E(\theta \mid X = x) = E(\Pi_x)$$

Przy modułowej funkcji straty $|\theta - a|$:

$$\hat{\theta}_B = \text{Med}(\Pi_x)$$

ogólniej: $E(g(\theta)|x)$



Estymator Bayesowski: przykłady

1. Niech: $X_1, \dots, X_n$ IID z rozkładu 0-1 z p-stwem sukcesu θ ; niech $\pi(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha, \beta)}$ dla $\theta \in (0, 1)$ znamy rozkład a posteriori:

$$\text{Beta}\left(\sum_{i=1}^n x_i + \alpha, n - \sum_{i=1}^n x_i + \beta\right)$$

rozkład $\text{Beta}(\alpha, \beta)$;
jego średnia
 $= \alpha / (\alpha + \beta)$

a zatem estymator Bayesowski przy kwadratowej funkcji straty to

$$\hat{\theta}_B = \frac{\sum_{i=1}^n x_i + \alpha}{n + \beta + \alpha}$$

np. dla 10 sukcesów zaobserwowanych w 20 próbach i dla rozkładu a priori $U(0, 1)$ (czyli $\text{Beta}(1, 1)$), mamy $\theta_B = 11/22 = 1/2$

a dla 15 sukcesów przy tym samym rozkładzie a priori:

$$\theta_B = 16/22 = 8/11$$



BENW: przykłady

1. Niech: $X_1, \dots, X_n$ IID z rozkładu 0-1 z p-stwem sukcesu θ ; niech dla $\theta \in (0,1)$ znamy rozkład a posteriori:
- $$\pi(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha, \beta)}$$

Beta($\sum_{i=1}^n x_i + \alpha, n - \sum_{i=1}^n x_i + \beta$)
max dla

$$BENW(\theta) = \frac{\sum_{i=1}^n x_i + \alpha - 1}{n + \beta + \alpha - 2}$$

rozkład Beta(α, β);
jego moda
= $(\alpha-1)/(\alpha + \beta - 2)$
dla $\alpha > 1, \beta > 1$

np. dla 10 sukcesów zaobserwowanych w 20 próbach i dla rozkładu a priori U(0,1) (czyli Beta(1,1)), mamy $BENW(\theta) = 10/20 = 1/2$
a dla 15 sukcesów zaobserwowanych w 20 próbach i tego samego rozkładu a priori, mamy $BENW(\theta) = 15/20 = 3/4$



Estymator Bayesowski: przykłady (2)

2. Niech: $X_1, \dots, X_n$ IID z rozkładu $N(\theta, \sigma^2)$, przy czym σ^2 znane; $\theta \sim N(m, \tau^2)$ dla pewnych m, τ znanych.

Wówczas rozkład a posteriori dla θ : $N\left(\frac{n\frac{1}{\sigma^2}\bar{X} + \frac{1}{\tau^2}m}{n\frac{1}{\sigma^2} + \frac{1}{\tau^2}}, \frac{1}{n\frac{1}{\sigma^2} + \frac{1}{\tau^2}}\right)$

a zatem Bayesowski estymator przy kwadratowej i modułowej funkcji straty to

$$\hat{\theta}_B = \frac{n\frac{1}{\sigma^2}\bar{X} + \frac{1}{\tau^2}m}{n\frac{1}{\sigma^2} + \frac{1}{\tau^2}}$$

np. jeśli próba 5 obserwacji 1.2; 1.7 ; 1.9 ; 2.1; 3.1 z rozkładu $N(\theta, 1)$ a rozkład a priori $\theta \sim N(1, 1)$, to

$$\theta_B = (5/1 * 2 + 1)/(5 + 1) = 11/6 \approx 1.83$$

a gdyby rozkład a priori $\theta \sim N(3, 1)$, to

$$\theta_B = (5/1 * 2 + 1*3)/(5 + 1) = 13/6 \approx 2.17$$



BENW: przykłady (2)

2. Niech: $X_1, \dots, X_n$ IID z rozkładu $N(\theta, \sigma^2)$, przy czym σ^2 znane; $\theta \sim N(m, \tau^2)$ dla pewnych m, τ znanych.

Wówczas rozkład a posteriori dla θ :

a zatem

$$BENW(\theta) = \frac{n \frac{1}{\sigma^2} \bar{X} + \frac{1}{\tau^2} m}{n \frac{1}{\sigma^2} + \frac{1}{\tau^2}}$$

$$N\left(\frac{n \frac{1}{\sigma^2} \bar{X} + \frac{1}{\tau^2} m}{n \frac{1}{\sigma^2} + \frac{1}{\tau^2}}, \frac{1}{n \frac{1}{\sigma^2} + \frac{1}{\tau^2}}\right)$$

np. jeśli próba 5 obserwacji 1.2; 1.7 ; 1.9 ; 2.1; 3.1 z rozkładu $N(\theta, 1)$ a rozkład a priori $\theta \sim N(1, 1)$, to

$$BENW(\theta) = (5 / 1 * 2 + 1) / (5 / 1 + 1) = 11 / 6 \approx 1.83$$

a gdyby rozkład a priori $\theta \sim N(3, 1)$, to

$$BENW(\theta) = (5 / 1 * 2 + 1 * 3) / (5 + 1) = 13 / 6 \approx 2.17$$





WARSAW UNIVERSITY
Faculty of Economic Sciences