Shinybrms - rozszerzenie shiny do uprawiania statystyki Bayesowskiej

Marta Szuwarska

Wprowadzenie do statystyki Bayesowskiej

czyli co to za twór?

Przykład 1


legenda
burze
silny deszcz
grad
porywy wiatru
trąby powietrzne
słabe burze
burze
silne burze
bardzo silne burze
prawdopodobieństwo
powyżej 60%
poniżej 60%
IMGW
od 08:00 dnia 09.07.2021
do 20:00 dnia 09.07.2021
Biuro Prognoz Warszawa

Przykład 1

Przykład 2


Kwiecień 2019
1
FRIDAY, OCTOBER 1, 2021 AT 12 PM UTC+02 – 3 PM UTC+02
W 2021 zjem na stołówce MiNI
Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska
Styczeń 2021
1
FRIDAY, OCTOBER 1, 2021 AT 12 PM UTC+02 – 3 PM UTC+02
W 2021 zjem na stołówce MiNI
Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska

Przykład 2

Przykład 3

Twierdzenie Bayesa

$Niech\ A, B \in \mathcal{F}\ będą\ zdarzeniami\ oraz\ \mathbb{P}(\mathrm{B}) > 0.$

Wtedy:

$$\mathbb{P}\,(A \mid B) = \frac{\mathbb{P}(B\,|A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

Twierdzenie Bayesa

$Niech\ \theta\ będzie\ wektorem\ parametrów.$

Wtedy:

$$\mathbb{P}\left(\theta \mid dane\right) = \frac{\mathbb{P}(dane\,|\theta)\mathbb{P}(\theta)}{\mathbb{P}(dane)}$$

Twierdzenie Bayesa

Niech θ będzie wektorem parametrów.

Wtedy:

Wiarygodność

Prawdopodobieństwo a priori

$$\mathbb{P}(\theta \mid dane) = \frac{\mathbb{P}(dane \mid \theta)\mathbb{P}(\theta)}{\mathbb{P}(dane)}$$

Prawdopodobieństwo a posteriori

Stała normalizująca

Twierdzenie Bayesa

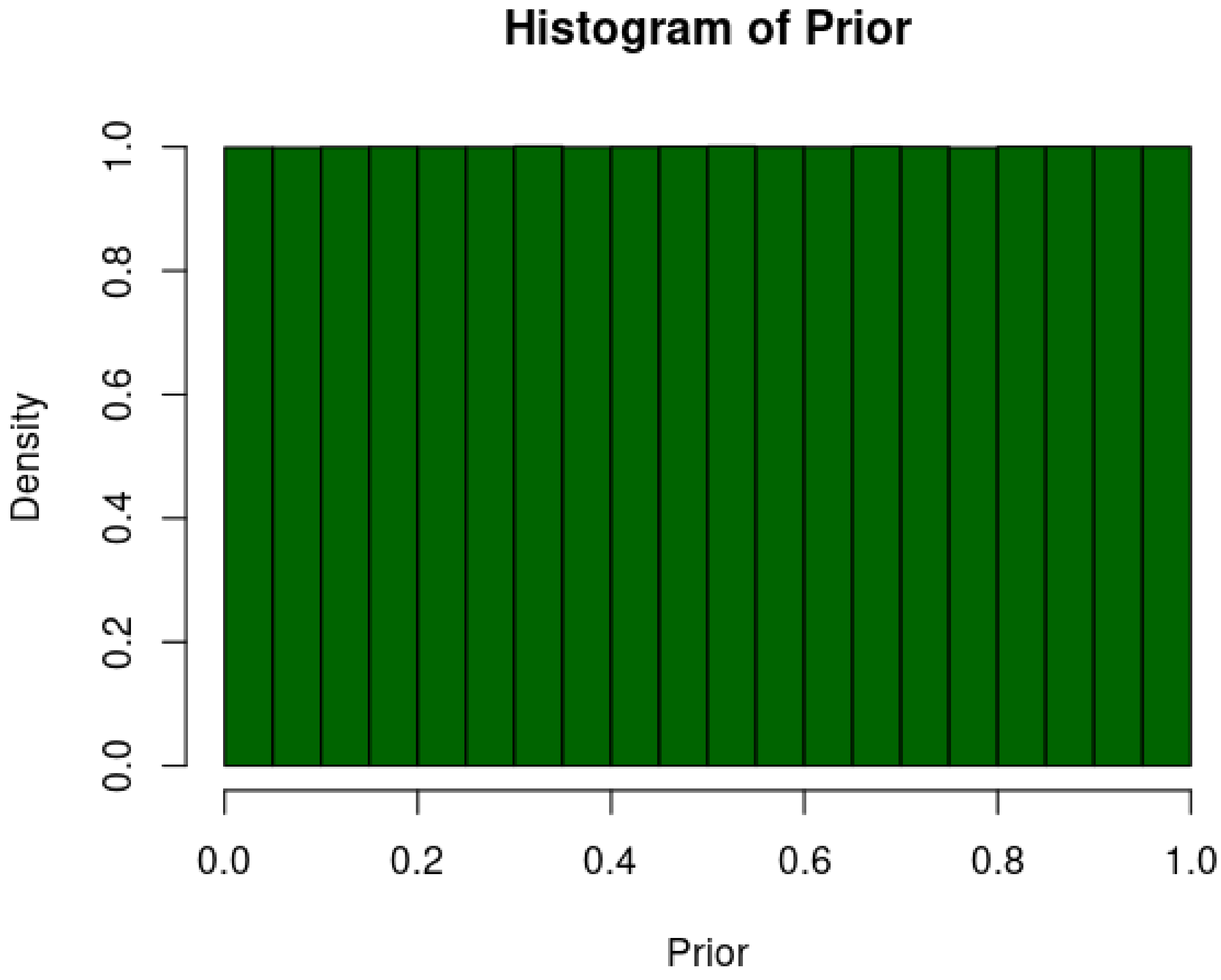
$Niech\ \theta\ będzie\ wektorem\ parametrów.$

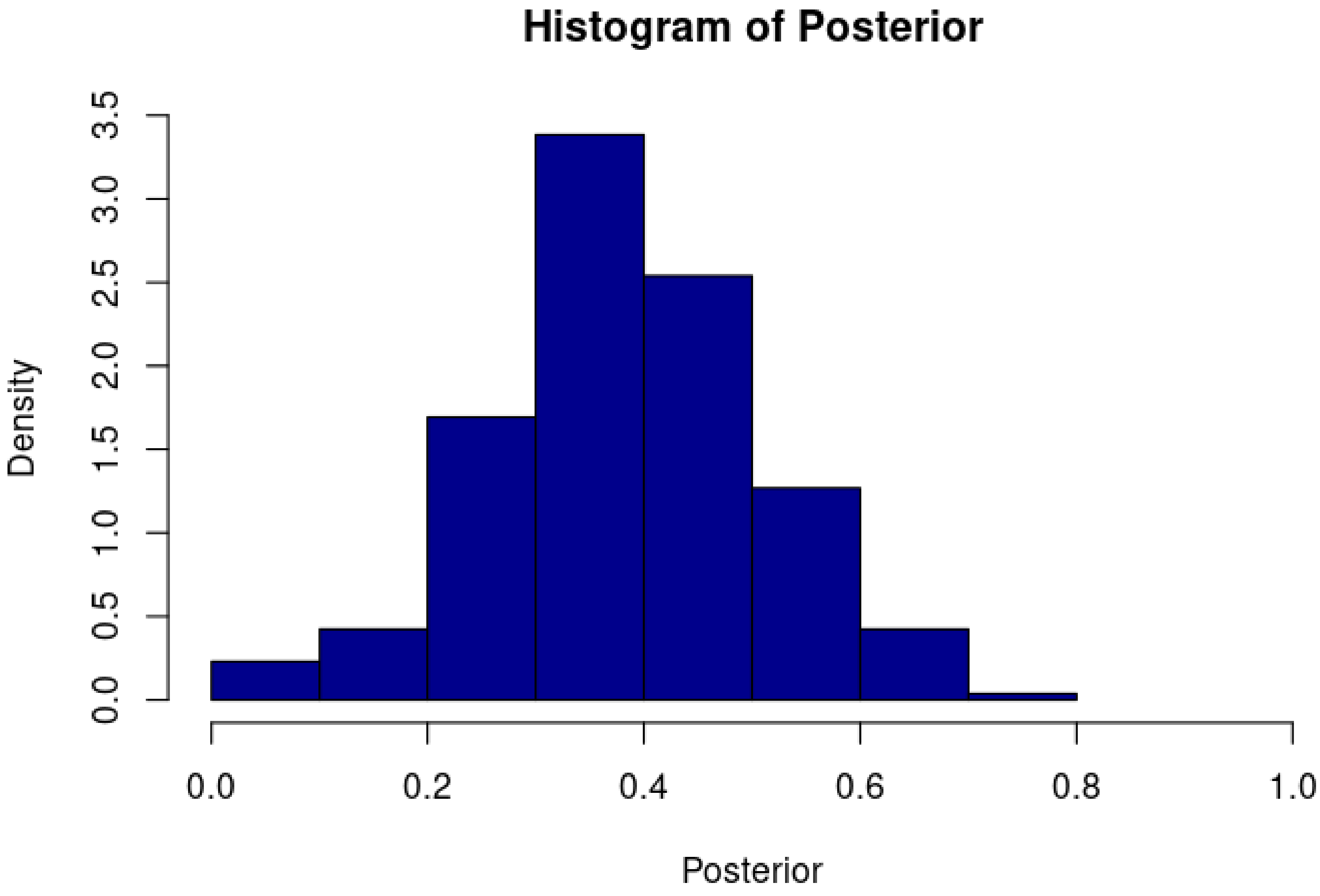
Wtedy:

$$\mathbb{P}\,(\theta \mid dane) = \frac{\mathbb{P}(dane\,|\theta)\mathbb{P}(\theta)}{\mathbb{P}(dane)}$$

$$\mathbb{P}\,(\theta \mid dane) = \frac{\mathbb{P}(dane\,|\theta)\mathbb{P}(\theta)}{\int_{\Theta}\ \mathbb{P}\,(dane|\theta)\mathbb{P}(\theta)d\theta}$$

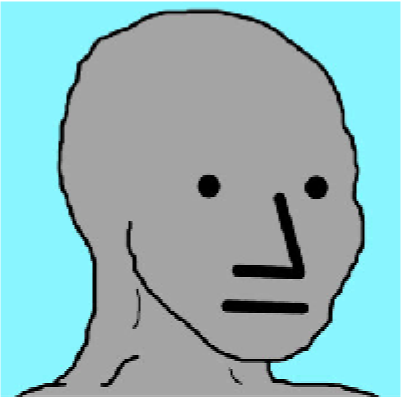
A priori i a posteriori na przykładzie


Histogram of Prior
Density
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Prior


Histogram of Posterior
Density
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.0
0.2
0.4
0.6
0.8
1.0
Posterior

Podejście Bayesowskie a klasyczne

Podejście klasyczne	Podejście Bayesowskie
Parametry ustalone	Parametry zmienne
Dane zmienne	Dane ustalone



Zalety podejścia Bayesa

czyli dlaczego warto zostać bayesistą?

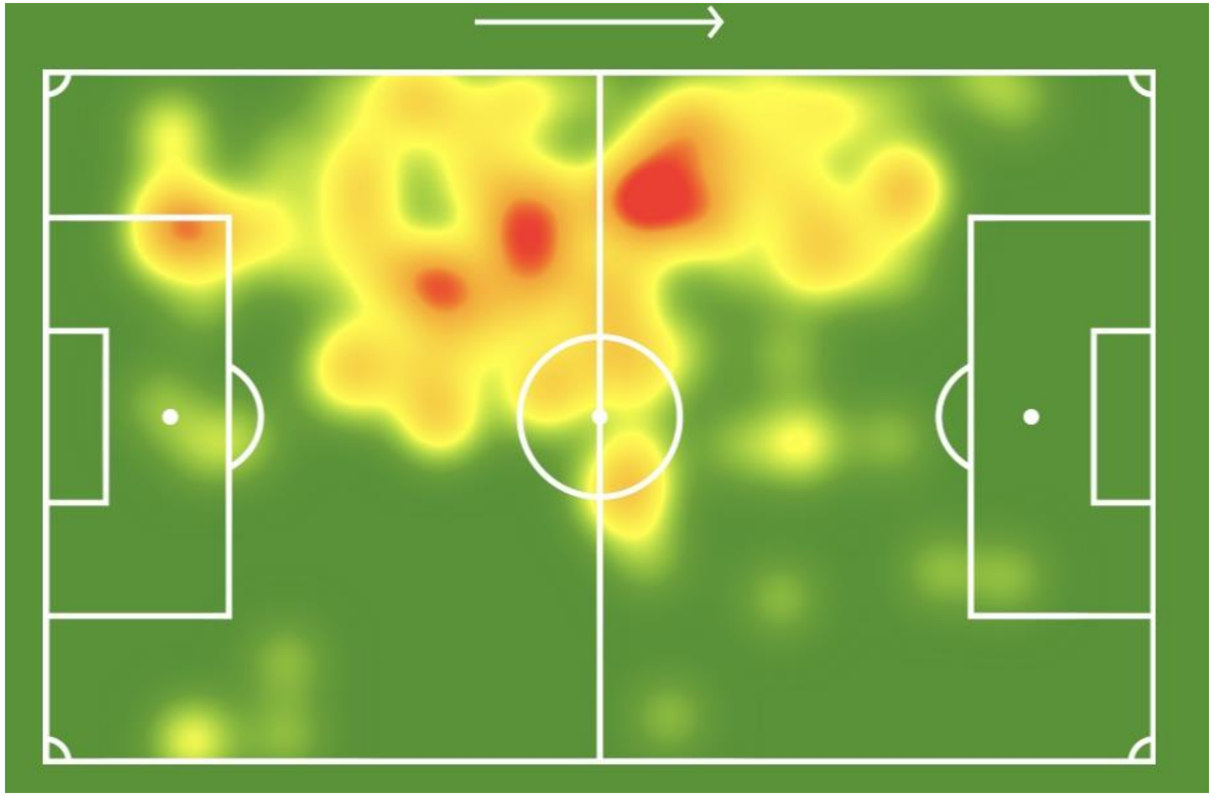
- Możliwość uwzględnienia wiedzy eksperckiej
- Możliwość łatwego narzucenia ograniczeń parametrów
- Możliwość wyliczenia pełnego rozkładu nawet przy mało licznym zbiorze
- Bardziej intuicyjna interpretacja wyników
- Łatwe wykluczenie parametrów zakłócających (*ang. nuisance parameters*)
- Rzadziej dochodzi do testowania istotności hipotezy zerowej
- Podejście bayesowskie nigdy nie jest gorsze niż klasyczne
- Można łatwo przeprowadzić diagnostykę modelu

Stosowane algorytmy

czyli gdzie tu optymalizacja?

Algorytmy MCMC

- Metropolis
- Metropolis-Hastings
- Próbnik Gibbsa
- Hamiltonian Monte Carlo
- no-U-turn sampler



NUTS

- eksploruje rozkład w każdą stronę,
- nie musi wracać,
- jest bardziej efektywny niż tradycyjne algorytmy,
- automatycznie określa długość i kierunek symulowanych trajektorii.

Powody powstania pakietu shinybrms

czyli czy naprawdę potrzebujemy kolejnego pakietu do R?

Starsze pakiety

GUI name	Commercial	Algorithm (for inferring the posterior)							Algorithm choice
		Non-MCMC				MCMC			
		Analytic	Numerical	MC	BB	Non-HMC	Static HMC	NUTS	
WinBUGS (Lunn et al., 2000)	no	no	no	no	no	yes	no	no	no
OpenBUGS (Spiegelhalter et al., 2014)	no	no	no	no	no	yes	yes	no	no
BugsXLA (Woodward, 2011)	no	no	no	no	no	yes	yes	no	no
IBM SPSS Amos (Arbuckle, 2020)	yes	no	no	no	no	yes	yes	no	yes
TEET (Qian, 2011)	no[(i)]	yes	yes	yes	no	yes	no	no	no
JASP (JASP Team, 2022)	no	yes	yes	yes	no	yes	no	yes[(ii)]	no
BRNPM (Karabatsos, 2015, 2017)	no	no	no	no	no	yes	no	no	no
Stata (StataCorp, 2019b)	yes	no	no	no	no	yes	no	no	yes
BayES (Emvalomatis, 2020)	no	no	no	no	no	yes	no	no	no
IBM SPSS (IBM Corp., 2020)	yes	yes	yes	yes	no	no	no	no	no
BEsmarter (BEsmarter Team, 2020a,b; Ramírez-Hassan and Graciano-Londoño, 2021)	no	no	no	no	yes	yes	no	no	yes[(iii)]

Testujemy apkę na przykładzie

czyli czy opatrunek z plazmy działa?

Podsumowanie

czyli czy warto zainstalować pakiet shinybrms?

Dziękuję za uwagę