

A bayesi gondolkodás alapjai

**Kehl Dániel,
a STAB tagja, a PTE Közgazdaságtudományi Kar adjunktusa**

Budapest, 2022. június 16.

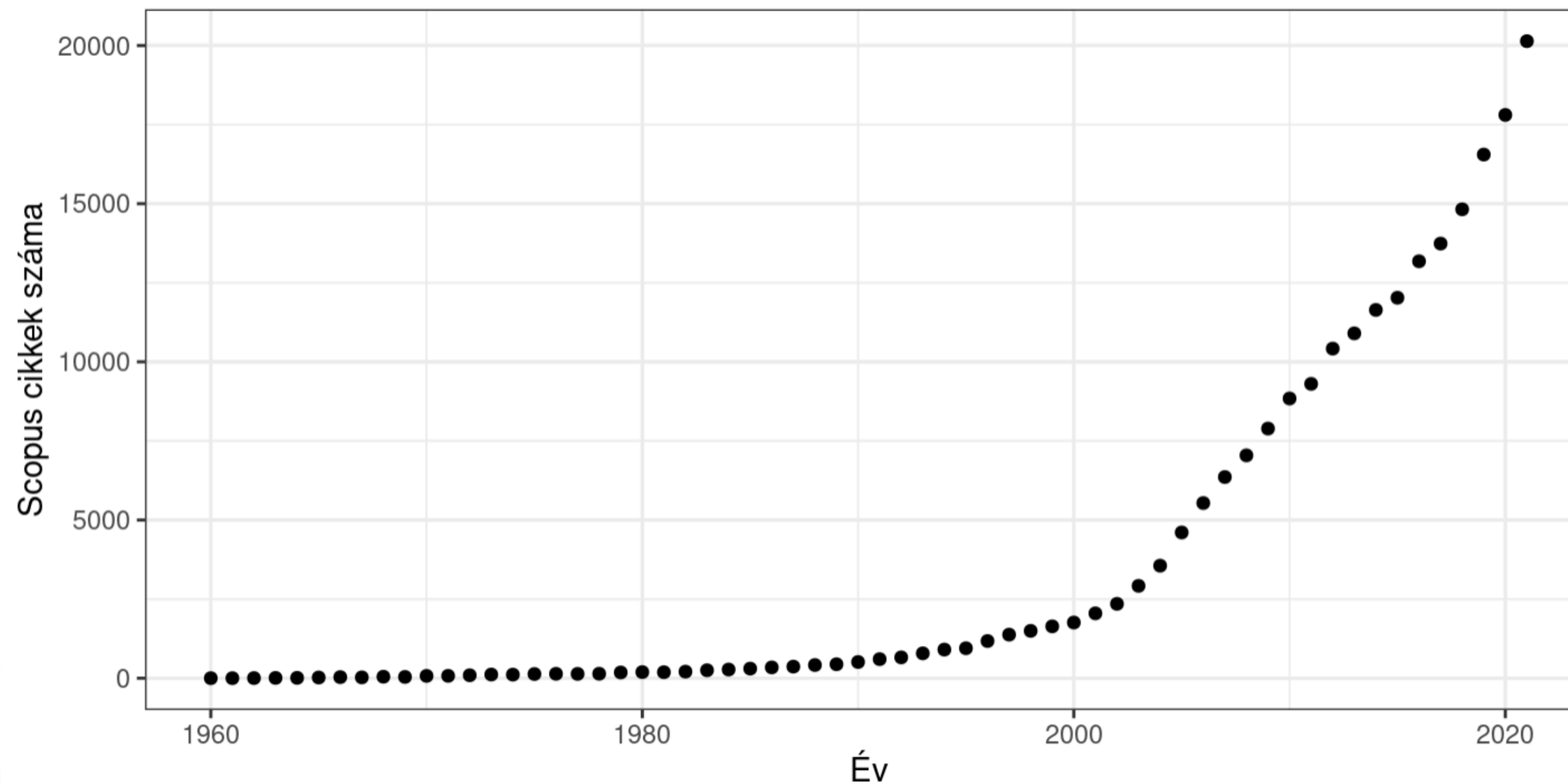
Bayesi statisztika

- a névadó Bayes (1763), az elnevezés R. A. Fishertől származik
- egy világ, ahol nincs CLT, megbízhatósági szint, p-érték, ...

Tulajdonság Hunyadi (2011)	Klasszikus/frekventista statisztika	Bayesi statisztika
Paraméter	Rögzített	Valószínűségi változó
Valószínűség	Objektív	Szubjektív is lehet
Külső információ	Nincs explicit megjelenítve	Lehet és pontosan dokumentált
Minta	Feltételezzük az ismételt mintavétel lehetőségét	Csak egyetlen minta van

Bayesi statisztikai alkalmazások

- a számítástechnikai kapacitás fejlődésével hatalmas lendület
- Magyarországon szigetszerű alkalmazásokkal



Bayes-tétel

- a középiskolában/egyetemen megismert formájában

$$\mathbf{P}(A \mid B) = \frac{\mathbf{P}(B \mid A)\mathbf{P}(A)}{\mathbf{P}(B)}$$

- a bayesi statisztikában alkalmazott formájában

$$f(\theta \mid y) = \frac{f(y \mid \theta)f(\theta)}{f(y)} \propto f(y \mid \theta)f(\theta)$$

- ahol y jelöli az adatokat (observables), θ a paramétereket (unobservables), $f(\theta)$ a prior, $f(y \mid \theta)$ a likelihood, $f(\theta \mid y)$ pedig a poszterior

összetevők

- likelihood
 - sűrűségfgv értékek szorzata, a modell specifikációja és az adatok megfigyelése után a paraméterek függvénye
 - nem a bayesi statisztika sajátja, nem közvetlenül a maximumát keressük
- prior
 - a paraméterre vonatkozó előzetes információinkat sűríti magában
 - nem célja az öncélú szubjektivitás
 - lehet konjugált, informatív/nem informatív, proper/improper, stb.
- poszterior
 - jellemzően bonyolult, sokdimenziós valószínűségi változó
 - előállítani könnyű, kiértékelni nehéz lehet

klasszikus példa – pénzfeldobás, fejek aránya (θ)

- likelihood (legyen n a feldobások száma, y a kapott fejek száma)

$$f(y \mid \theta) = \binom{n}{y} \theta^y (1 - \theta)^{n-y} \propto \theta^y (1 - \theta)^{n-y}$$

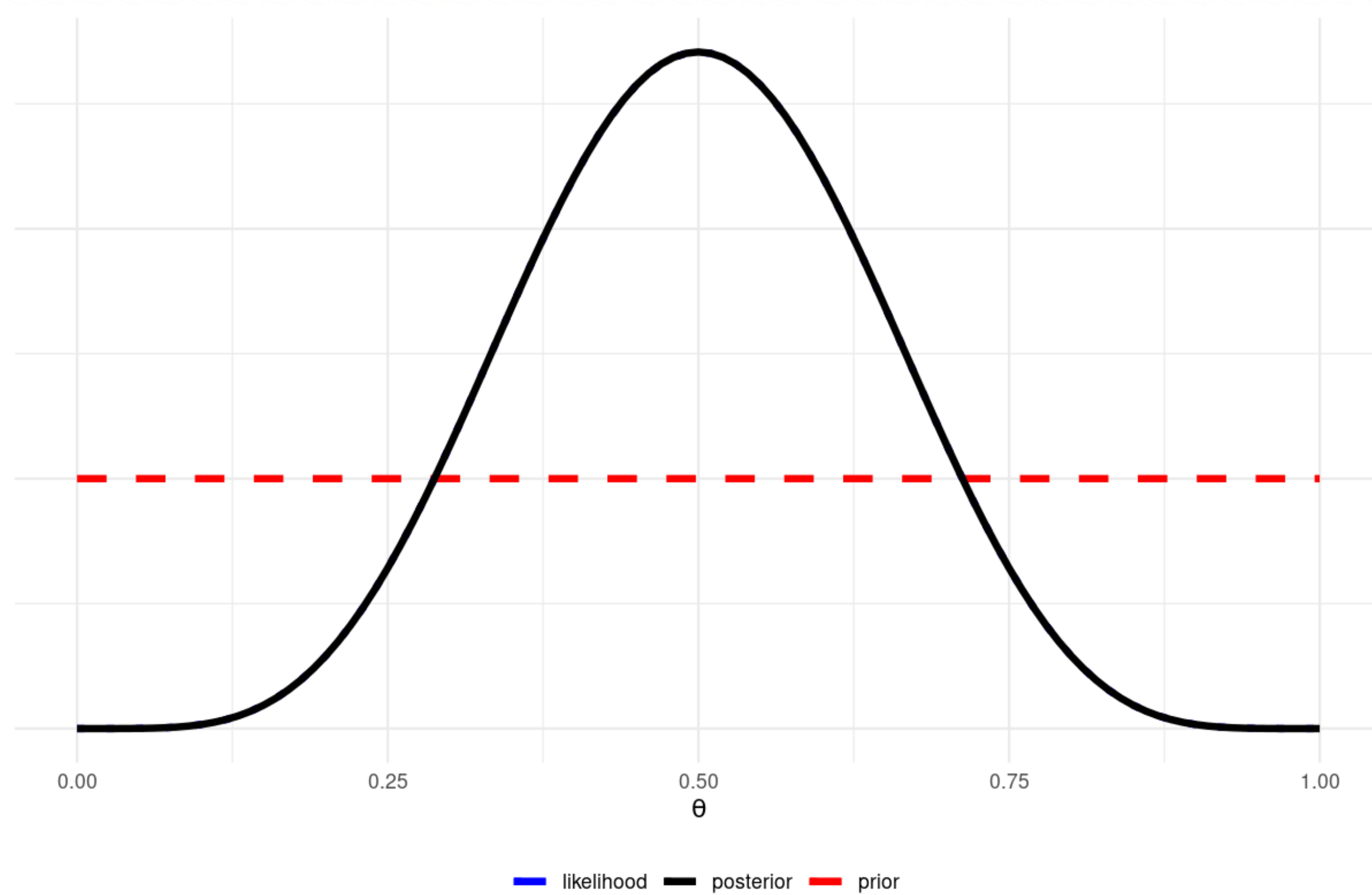
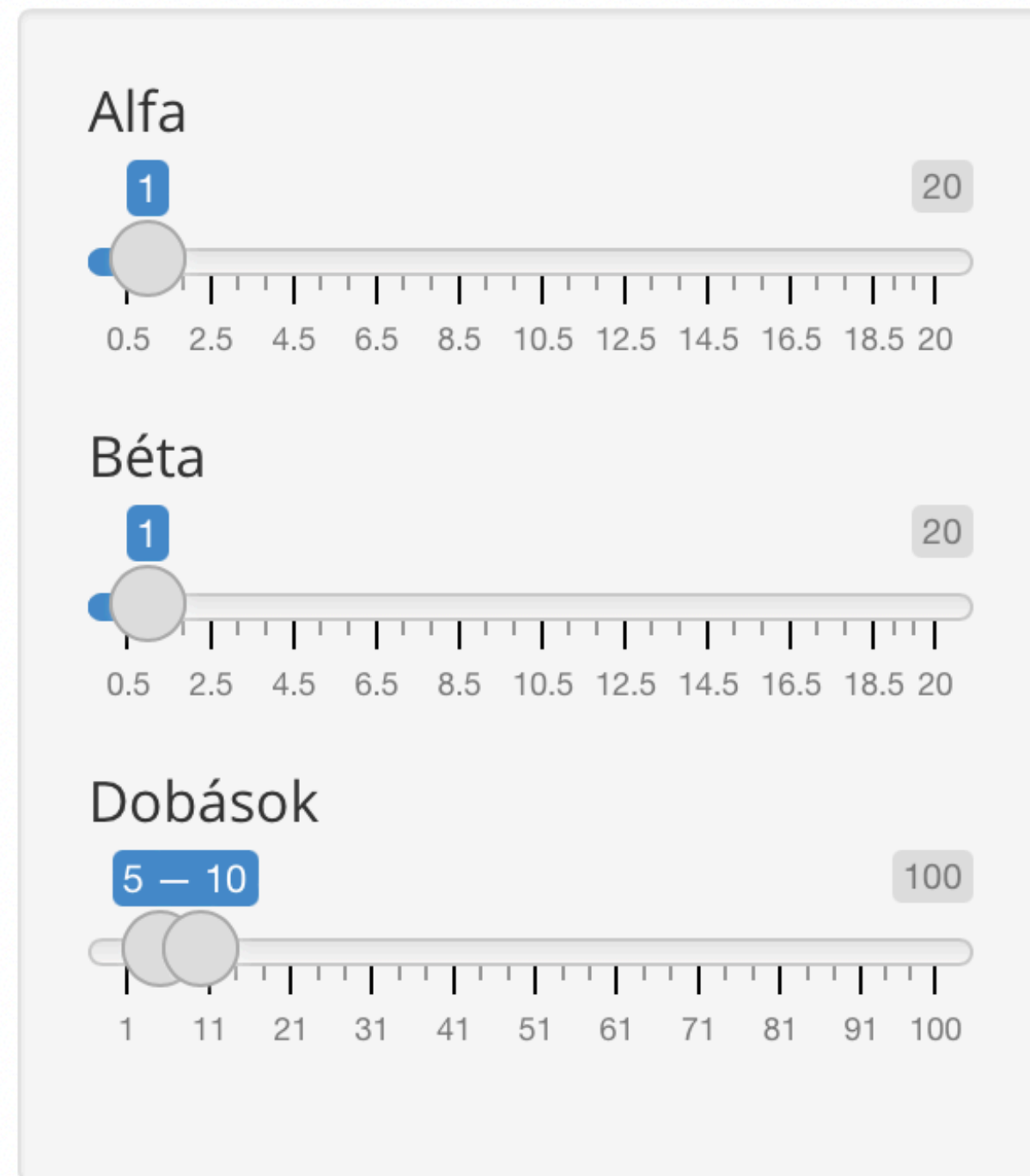
- prior (praktikus választás a béta eloszlás)

$$f(\theta) \propto \theta^{\alpha-1} (1 - \theta)^{\beta-1} \quad \theta \in (0, 1), \alpha > 0, \beta > 0$$

- poszterior (szintén béta eloszlás, $\alpha + y$ és $\beta + n - y$ paraméterekkel)

$$f(\theta \mid y) \propto \theta^{y+\alpha-1} (1 - \theta)^{n-y+\beta-1}$$

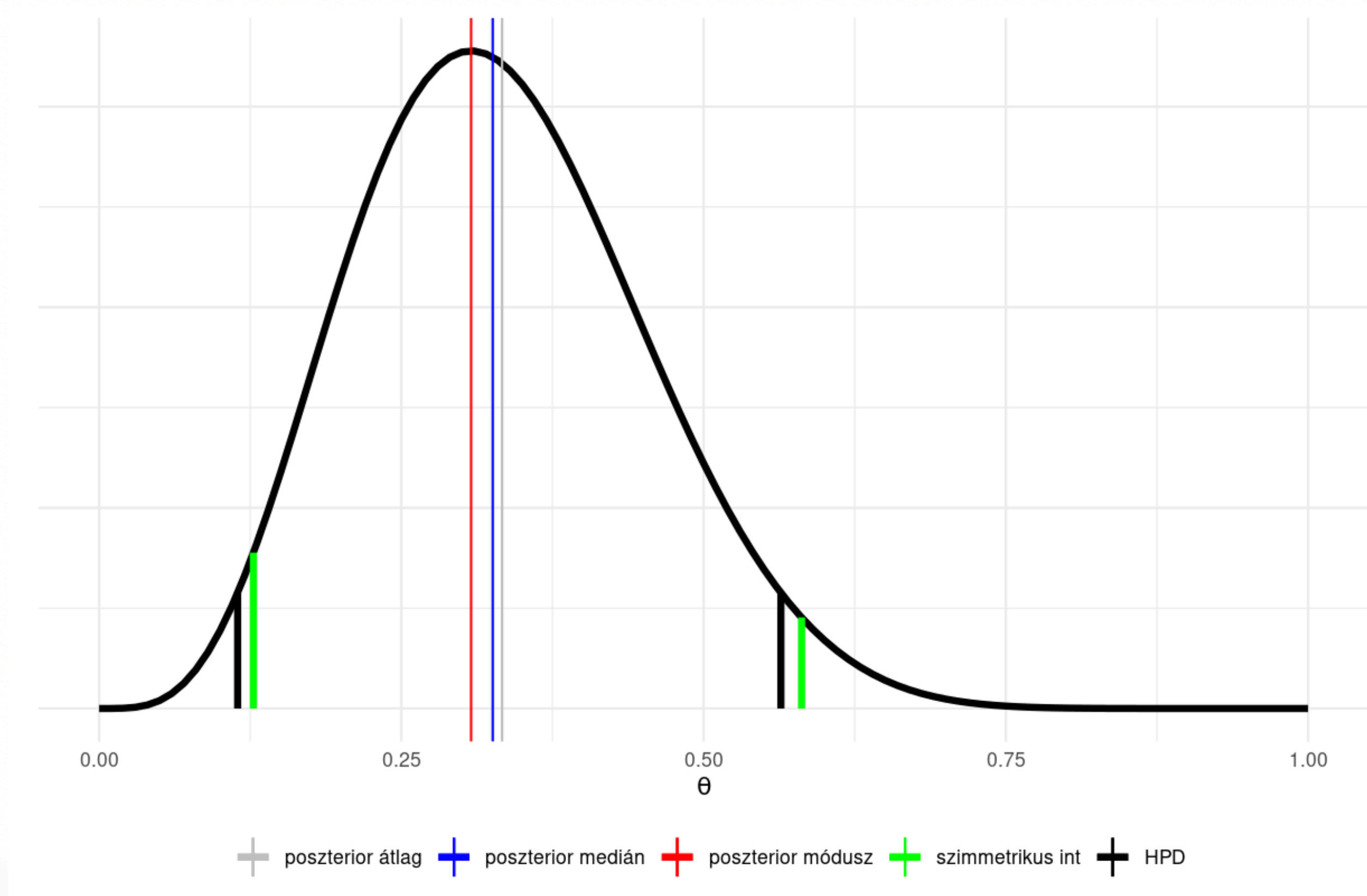
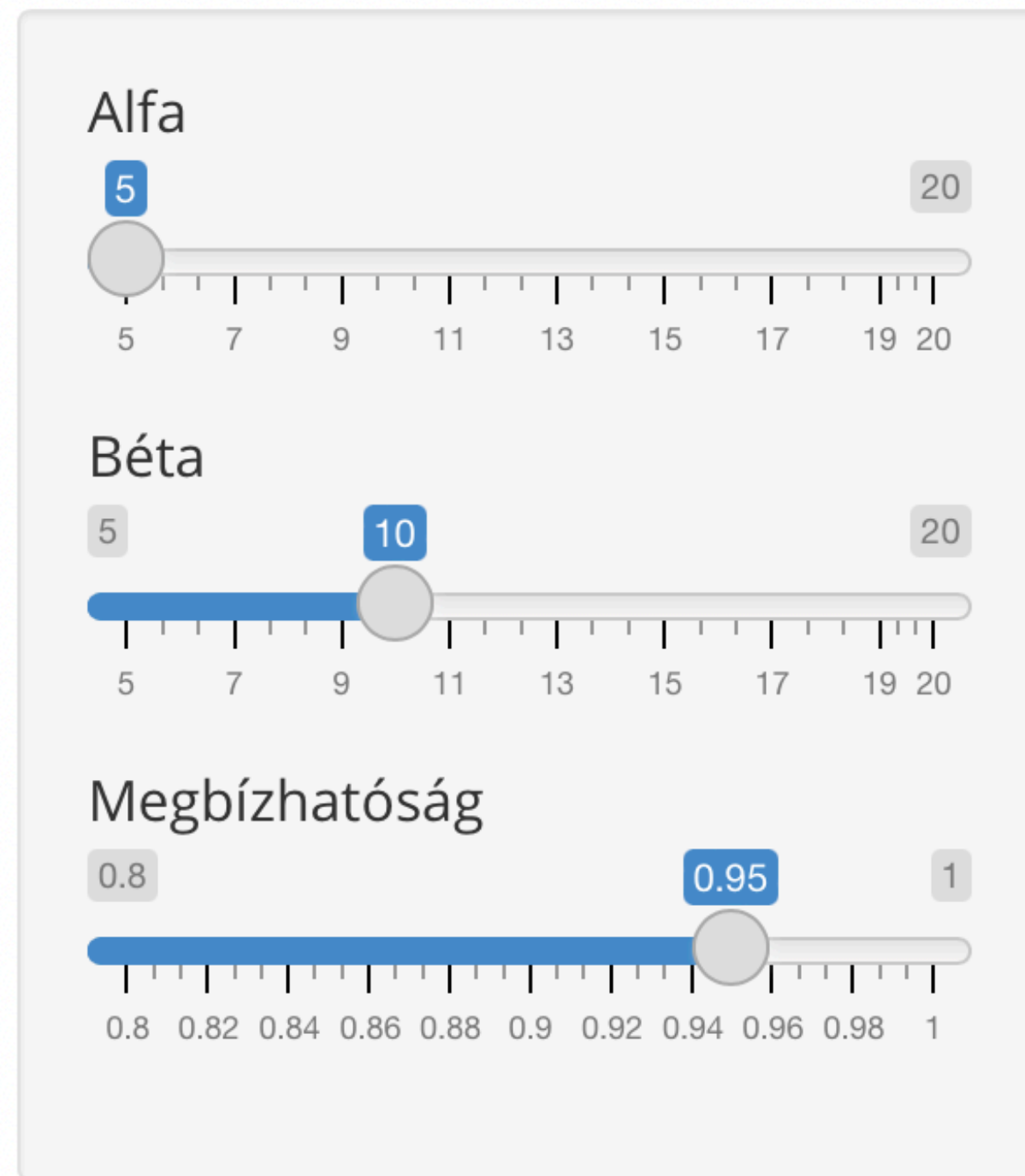
pénzfeldobás animáció



poszterior kiértékelése

- tartalmazza tehát a paraméterekre vonatkozó előzetes ismereteket (a prioron keresztül), az adatokban rejlő információt (a likelihoodon keresztül)
- egy általában sokdimenziós kvázi sűrűségfüggvény
 - “intervallum becslés”
 - credible set, ha ez összefüggő, akkor credible interval
 - highest posterior density (HPD) meghatározása nem triviális
 - “pontbecsléshez” veszteségfüggvényre van szükség
 - négyzetes $\rightarrow$ várható érték
 - abszolút $\rightarrow$ medián
 - mindent-vagy-semmit $\rightarrow$ módusz
- napjainkban ezt már a gép/szoftver automatikusan elvégzi!

poszterior kiértékelése a pénzérme dobás esetén



egyéb poszterior kiértékelési módszerek

- egydimenziós esetben
 - akár analitikus megoldás
 - rejection sampling
- sok becsülendő paraméter esetén MCMC algoritmusok
 - Markov lánc(ok) létrehozása, melyek “végigsétálnak” a kívánt eloszláson
 - hátránya, hogy a mintaelemek nem függetlenek
 - fontos feladat a konvergencia megállapítása
 - Metropolis-Hastings algoritmus (Metropolis et al. (1953), Hastings (1970))
 - Gibbs mintavétel (Geman and Geman (1984))
- <https://chi-feng.github.io/mcmc-demo/app.html>

összefoglalás

- a modell és a prior meghatározása után a poszterior előállítása egyszerű
- a poszteriorban rejlő információ kinyerése sokszor számításigényes
- egyre fejlettebb algoritmusok és támogató szoftverek
- képzésbe integrálva egyre több bayesi gondolat/cikk születése várható

hivatkozások

Bayes, T. 1763. "An Essay Towards Solving a Problem in the Doctrine of Chances." *Phil. Trans. Of the Royal Soc. Of London* 53: 370–418.

Geman, Stuart, and Donald Geman. 1984. "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images." *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6 (6): 721–41.
<https://doi.org/10.1109/TPAMI.1984.4767596>.

Hastings, Wilfred Keith. 1970. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." *Biometrika* 57 (1): 97–109.

Hunyadi, L. 2011. "Bayesi Gondolkodás a Statisztikában." *Statisztikai Szemle* 89 (10-11): 1150–71.

Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. 1953. "Equation of State Calculations by Fast Computing Machines." *The Journal of Chemical Physics* 21 (6): 1087–92.
<https://doi.org/10.1063/1.1699114>.

