

STATISZTIKAI TANULÁS AZ
IDEGRENDSZERBEN

KÖZELÍTŐ INFERENCIA II.
MONTE CARLO MÓDSZEREK

ISMÉTLÉS

- ▶ Egy valószínűségi modellben a következtetéseinket a látensek vagy a paraméterek fölötti poszterior írja le.
- ▶ Ezek kiszámítása gyakran nehéz.
- ▶ Ekkor használunk közelítő módszereket:
 - ▶ Variációs módszerek (egyszerűbb eloszlással közelítés)
 - ▶ pl. Pontbecslések (pl. EM, k-means)
 - ▶ Monte Carlo módszerek

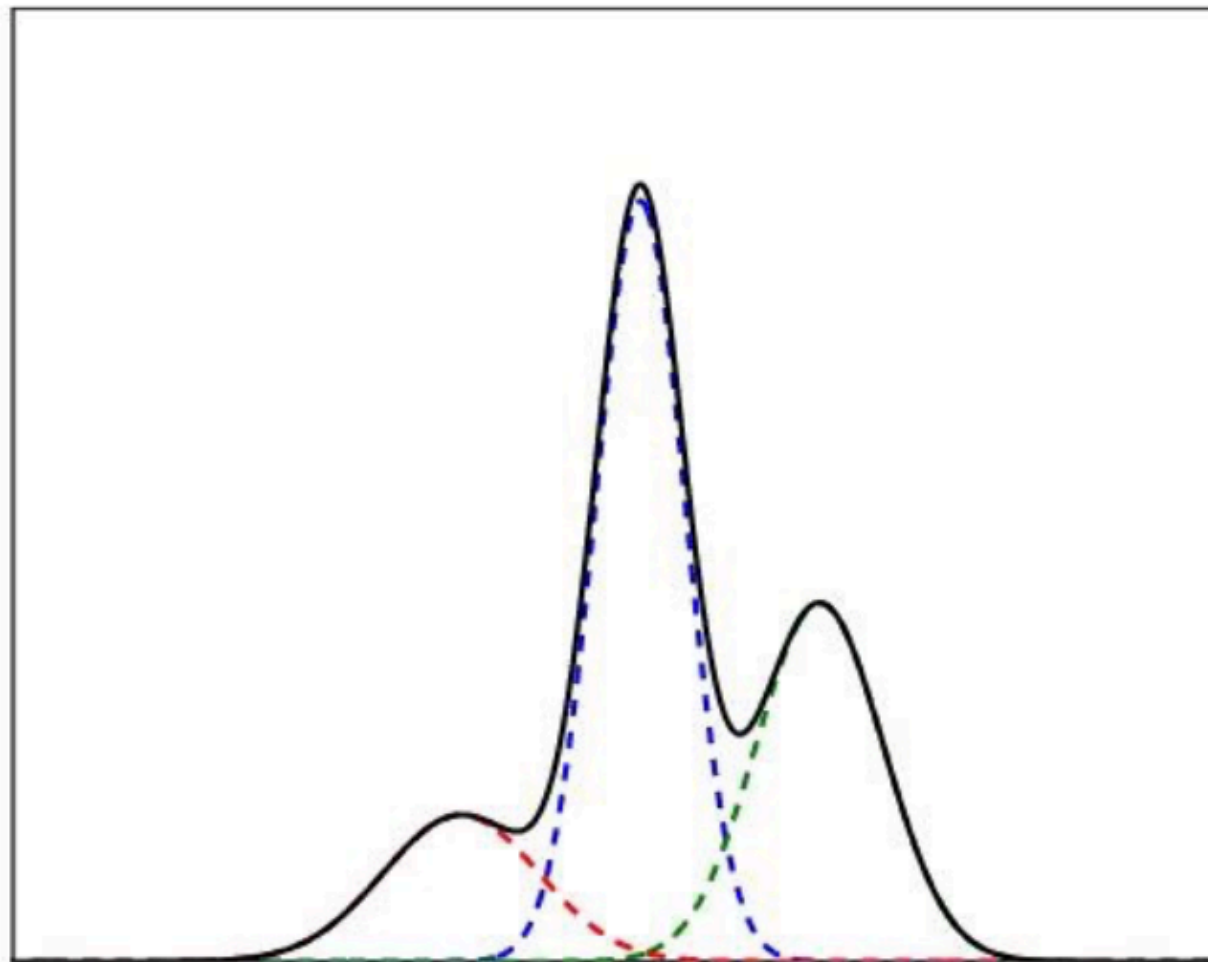
INFERENCIA II.

- ▶ Közelítés szükségessége
- ▶ Mintavételezés
- ▶ Egyszerűbb Monte Carlo módszerek
- ▶ Markov Chain Monte Carlo módszerek

INFERENCIA II.

- ▶ Közelítés szükségessége
- ▶ Mintavételezés
- ▶ Egyszerűbb Monte Carlo módszerek
- ▶ Markov Chain Monte Carlo módszerek

MIXTURE OF GAUSSIANS

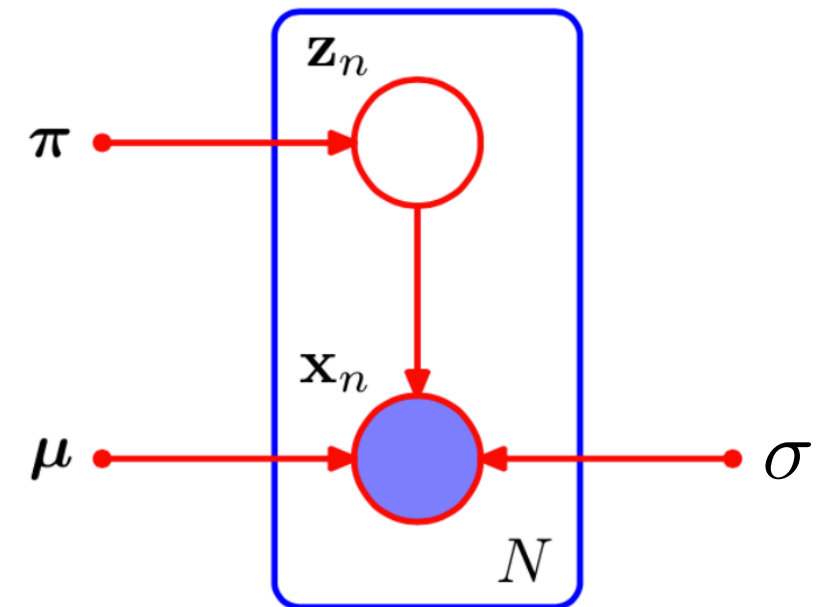
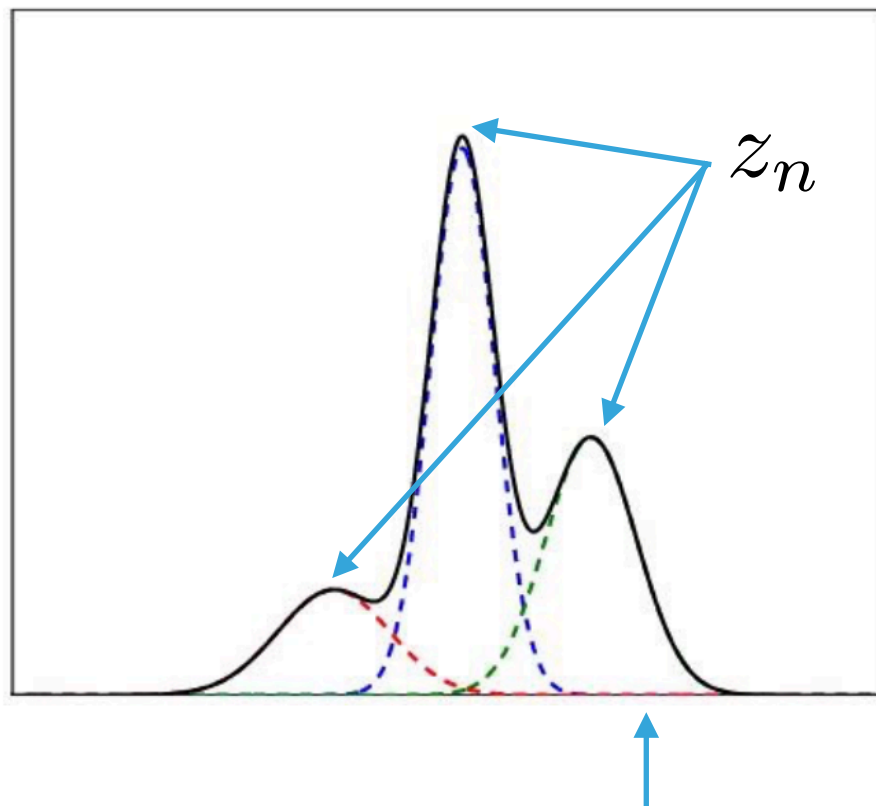


MIXTURE OF GAUSSIANS

- Valószínűségi modell:

$$p(z_n = i | \pi) = \pi_i$$

$$p(x_n | \mu, \sigma, z_n) = \mathcal{N}(x; \mu_{z_n}, \sigma_{z_n})$$

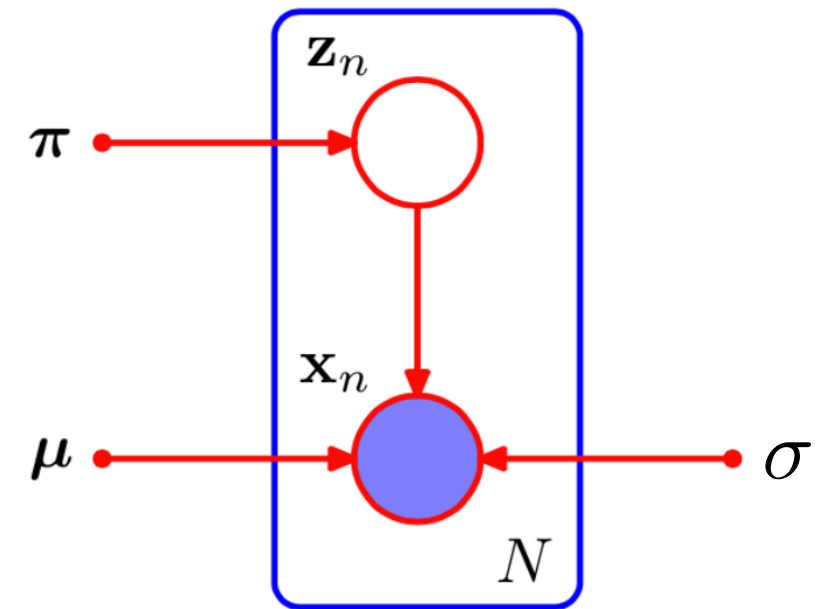


MIXTURE OF GAUSSIANS

- Valószínűségi modell:

$$p(z_n = i | \pi) = \pi_i$$

$$p(x_n | \mu, \sigma, z_n) = \mathcal{N}(x; \mu_{z_n}, \sigma_{z_n})$$



- Poszterior:

$$p(z, \pi, \mu, \sigma | x) = \frac{p(x | z, \mu, \sigma) p(z | \pi) p(\mu, \sigma, \pi)}{p(x)}$$

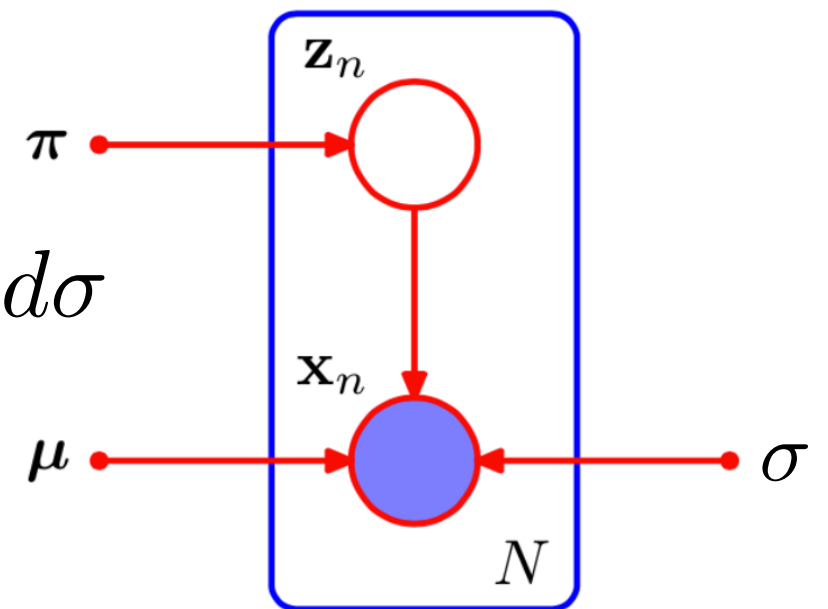
$$p(x | z, \mu, \sigma) = \prod_{n=1}^N p(x_n | z_n, \mu, \sigma)$$

$$= \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma_{z_n}^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma_{z_n}^2}\right)$$

MIXTURE OF GAUSSIANS

$$p(\pi, \mu, \sigma | x) = ? \sum p(z, \pi, \mu, \sigma | x)$$

$$p(x) = ? \int \int \int \sum_z p(x, \pi, \mu, \sigma, z) d\pi d\mu d\sigma$$



► Posztterior:

$$p(z, \pi, \mu, \sigma | x) = \frac{p(x | z, \mu, \sigma) p(z | \pi) p(\mu, \sigma, \pi)}{p(x)}$$

$$p(x | z, \mu, \sigma) = \prod_{n=1}^N p(x_n | z_n, \mu, \sigma)$$

$$= \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma_{z_n}^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma_{z_n}^2}\right)$$

KÖZELÍTÉS

- ▶ Pontbecslés
 - ▶ optimalizációs probléma
 - ▶ k-means, EM
- ▶ Variációs módszerek
 - ▶ Egyszerűbb modell, melyet könnyebb illeszteni
- ▶ Monte Carlo módszerek
 - ▶ Az eloszlások mintákkal való reprezentálása

INFERENCIA II.

- ▶ Közelítés szükségessége
- ▶ Mintavételezés
- ▶ Egyszerűbb Monte Carlo módszerek
- ▶ Markov Chain Monte Carlo módszerek

MINTAVÉTELEZÉS

► Problémák:

- **1. Független mintákat szeretnénk venni egy eloszlásból (pl. a poszteriorból)**

$$x^{(r)} \sim P(x)$$

- **2. Egy tetszőleges függvény várhatóértékének kiszámítása**

$$\mathbb{E}(f(x)) = \langle f \rangle_{P(x)} = \int f(x) P(x) dx$$

MINTAVÉTELEZÉS

- ▶ Tegyük fel, hogy valaki egy (bonyolult) kockajátékra invitál.
- ▶ Megadja, hogy milyen esetben mennyi nyereményt fizet.
- ▶ Azt is megmondja, hogy mennyibe kerül egy játék.
- ▶ Azt szeretnénk eldönteni, hogy akarunk-e játszani.
- ▶ A játék várhatóértékét szeretnénk kiszámolni.
- ▶ **Ötlet:** játsszuk le a játékot (magunkban) sokszor és vegyük az eredmények átlagát.

MINTAVÉTELEZÉS

- ▶ Játsszuk le a játékot (magunkban) sokszor és vegyük az eredmények átlagát.

Nyeremény várhatóértéke

$$F = \mathbb{E}(f(x))$$

Nyeremény egy x kimenetelnél

$$x_n \sim P(x)$$

Minták a játékból

$$\hat{F} = \sum_{n=1}^N \frac{1}{N} f(x_n)$$

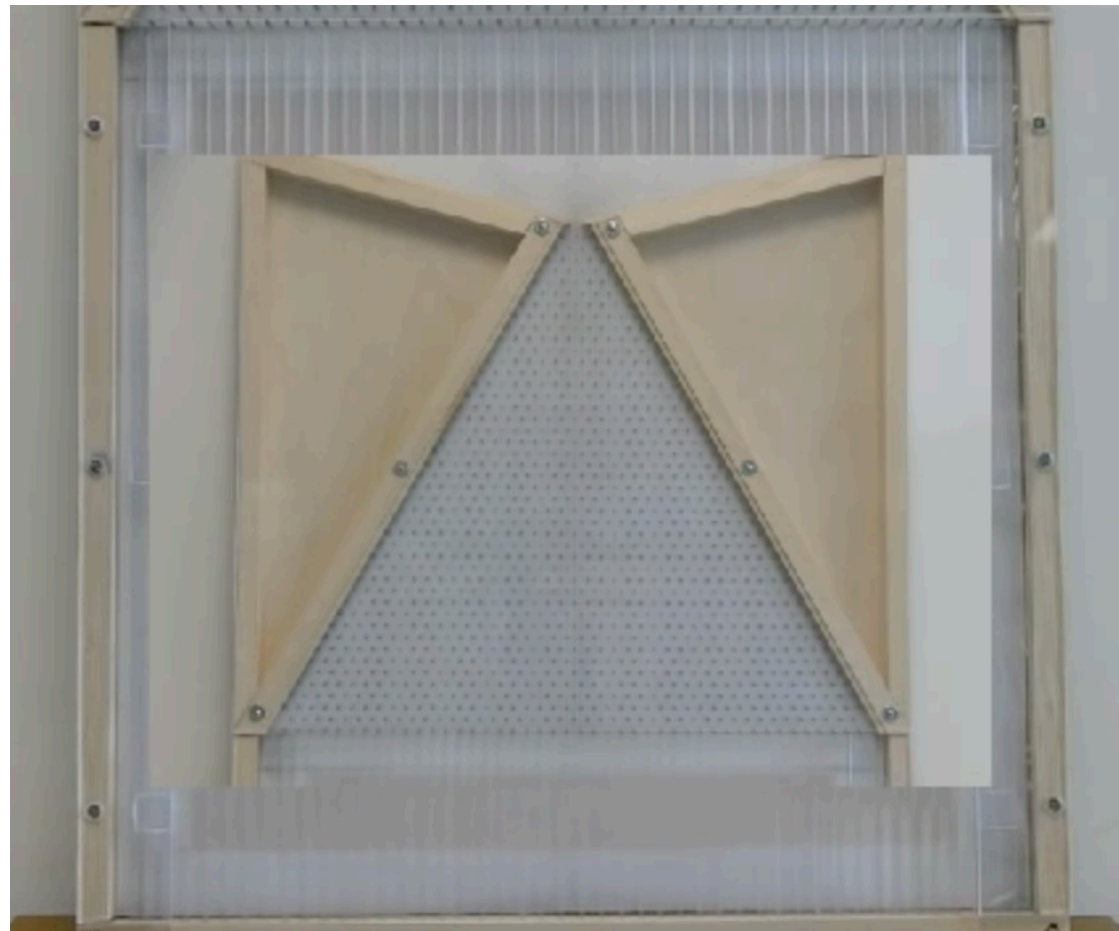
Becslés a várhatóértékre

MINTAVÉTELEZÉS

- ▶ Problémák:

- ▶ **1. Független mintákat szeretnénk venni egy eloszlásból (pl. a poszteriorból)**
- ▶ 2. Egy tetszőleges függvény várhatóértékének kiszámítása

GALTON-DESZKA



<https://www.youtube.com/watch?v=EMK5WPRjGgA>

MILYEN ELOSZLÁSOKBÓL TUDUNK KÖNNYEN MINTÁKAT VENNI?

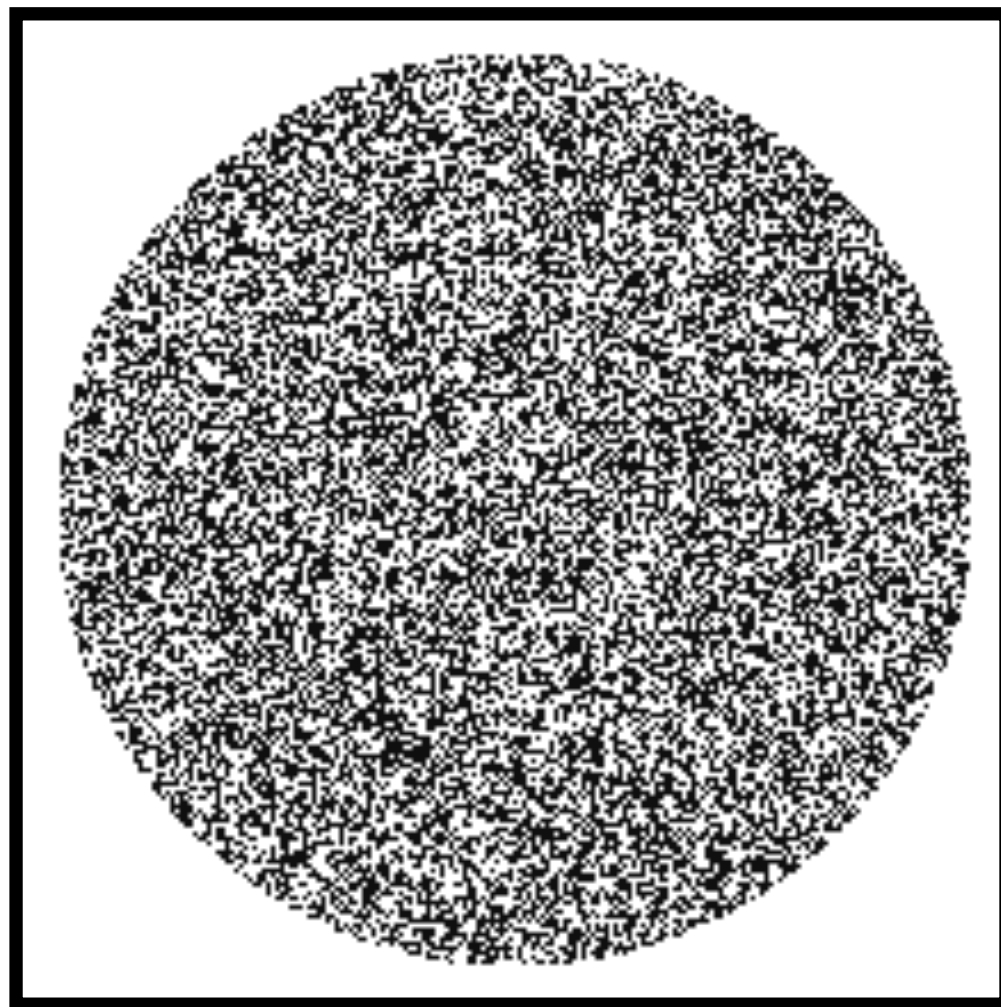
- ▶ Dobókocka (egyenletes eloszlás egy hat elemű halmazon)
- ▶ Egyenletes eloszlás a $[0,1]$ intervallumon (pseudo-random generátor)
- ▶ Pénzérme feldobás
- ▶ Tudjuk-e ezeket használni, hogy más eloszlásokból vegyünk mintákat?

INFERENCIA II.

- ▶ Közelítés szükségessége
- ▶ Mintavételezés
- ▶ Egyszerűbb Monte Carlo módszerek
- ▶ Markov Chain Monte Carlo módszerek

REJECTION SAMPLING

- ▶ Hogyan mintavételezzünk pontokat egyenletesen egy körlapon?



MINTAVÉTELEZÉS

► Problémák:

- 1. Független mintákat szeretnénk venni egy eloszlásból (pl. a poszteriorból)

$$x^{(r)} \sim P(x)$$

- Tegyük fel, hogy csak P valamely többszöröséhez férünk hozzá:

$$P^*(x) = Z \cdot P(x)$$

$$P(x|Data) \propto \boxed{P(Data|x)P(x)} = P^*(x)$$

$$P(Data) = \int \int \int \sum_z P(\mu, \sigma, \pi, z, Data) d\mu d\sigma d\pi$$

MINTAVÉTELEZÉS

- ▶ Problémák:

- ▶ 1. Független mintákat szeretnénk venni egy eloszlásból (pl. a poszteriorból)

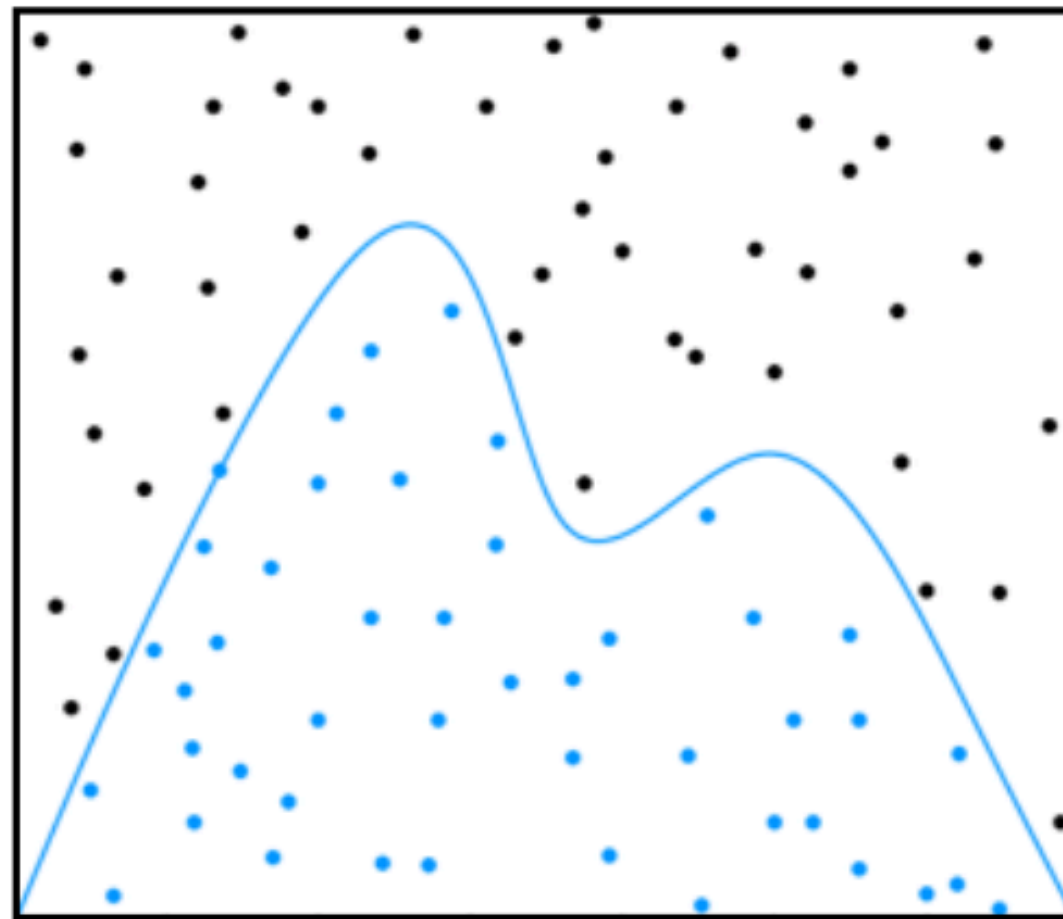
$$x^{(r)} \sim P(x)$$

- ▶ Tegyük fel, hogy csak P valamely többszöröséhez férünk hozzá:

$$P^*(x) = Z \cdot P(x)$$

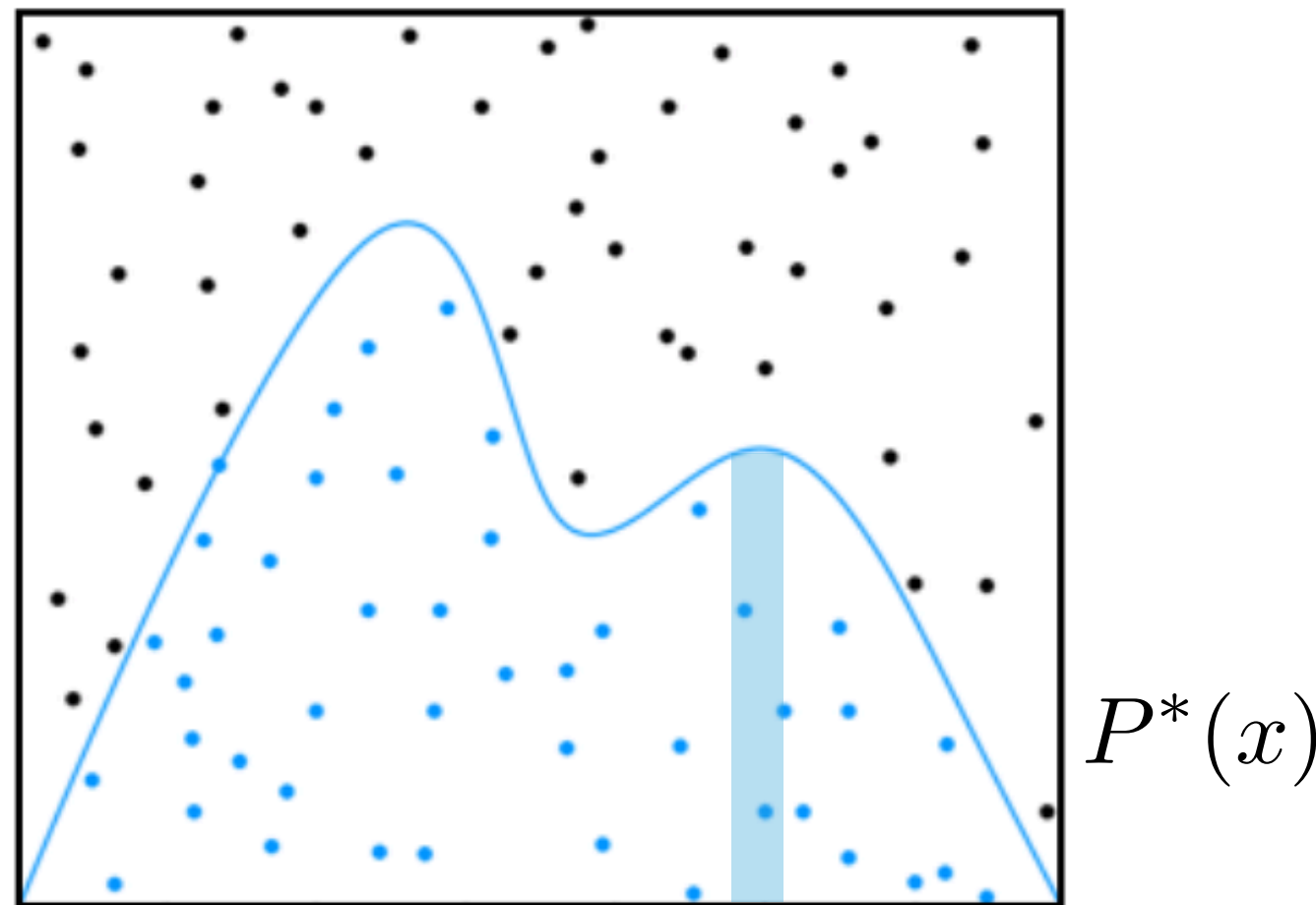
- ▶ Lényeg: ki tudjuk értékelni $P^*(x)$ -et tetszőleges x -ben.

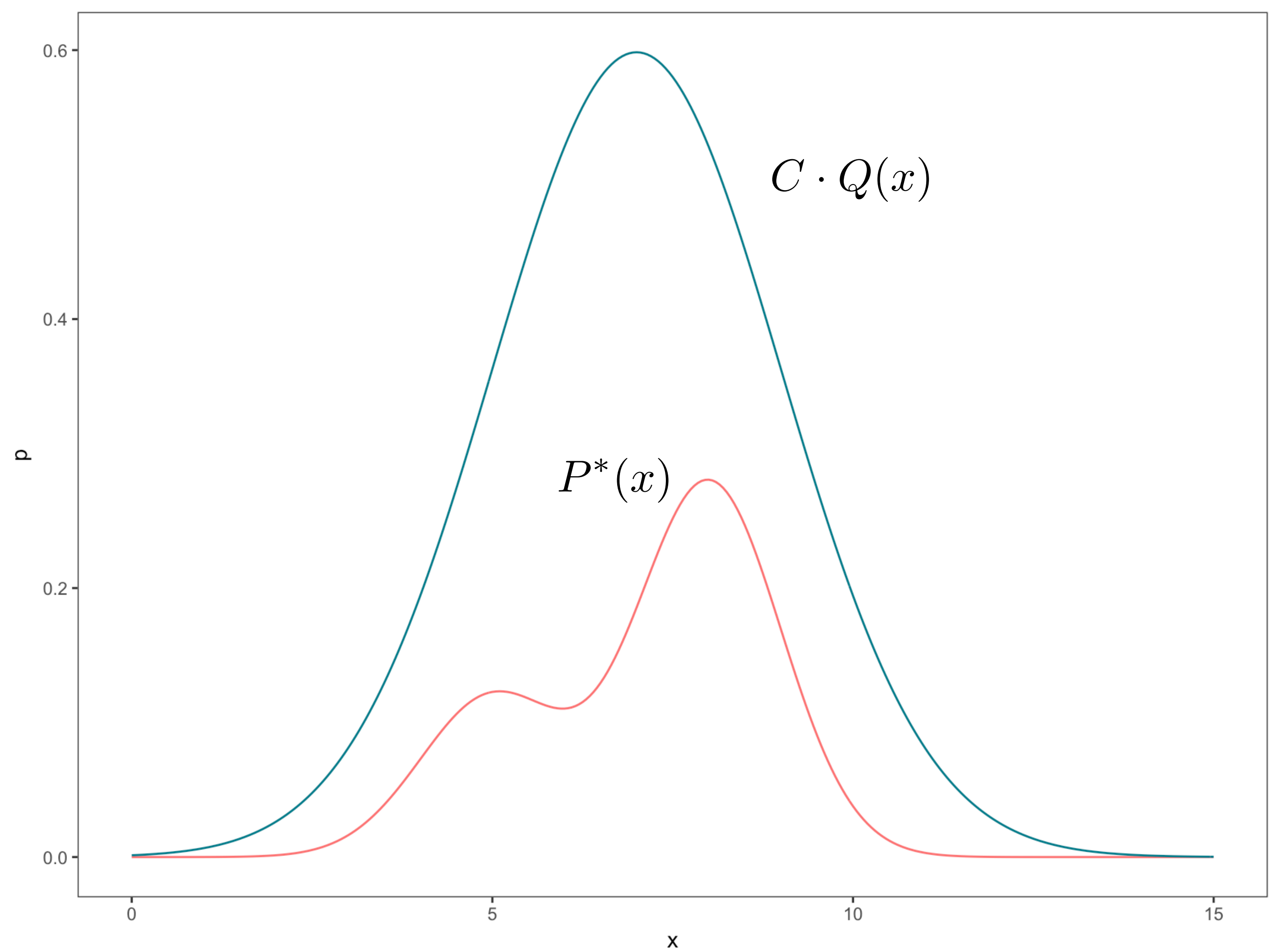
REJECTION SAMPLING

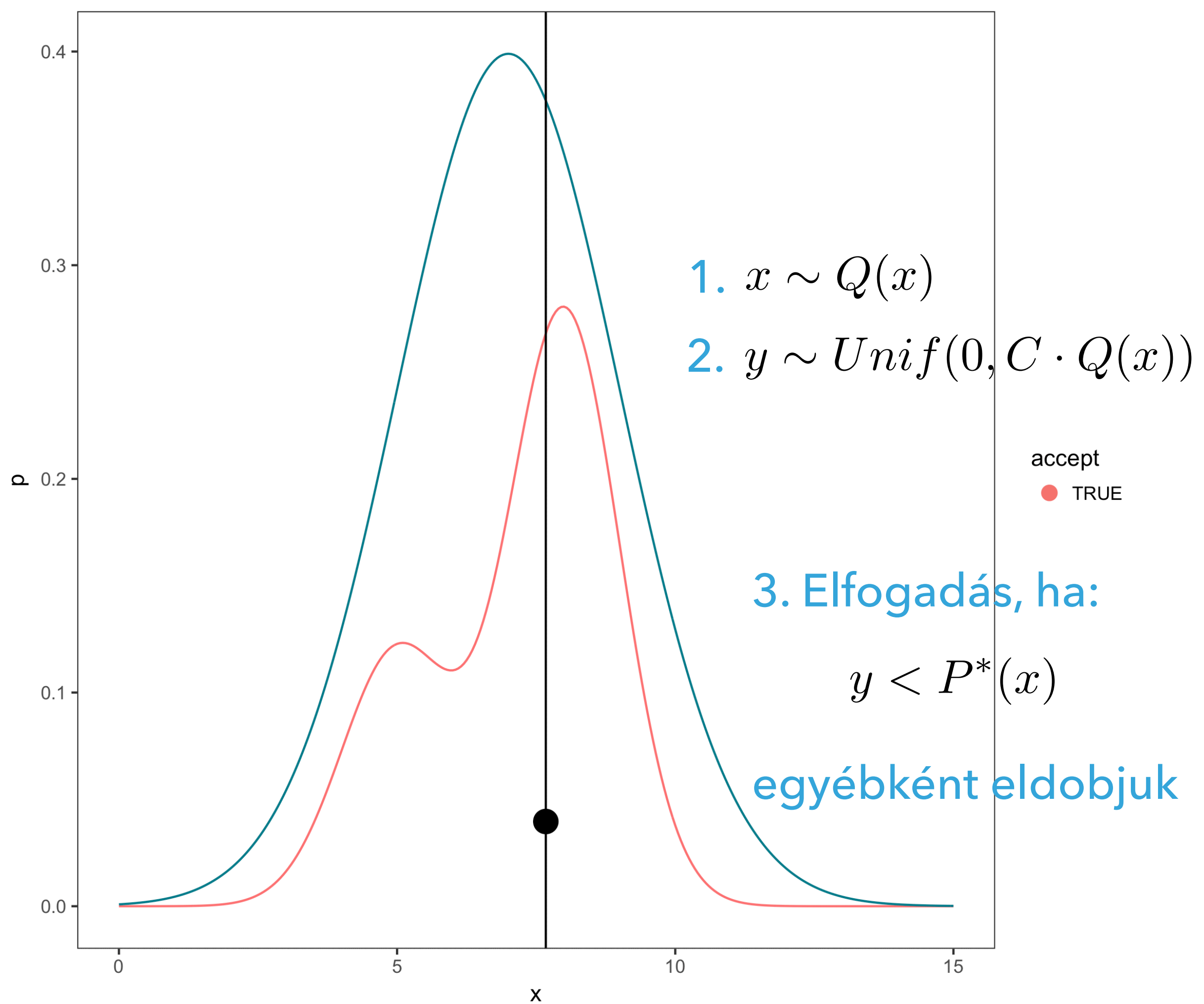


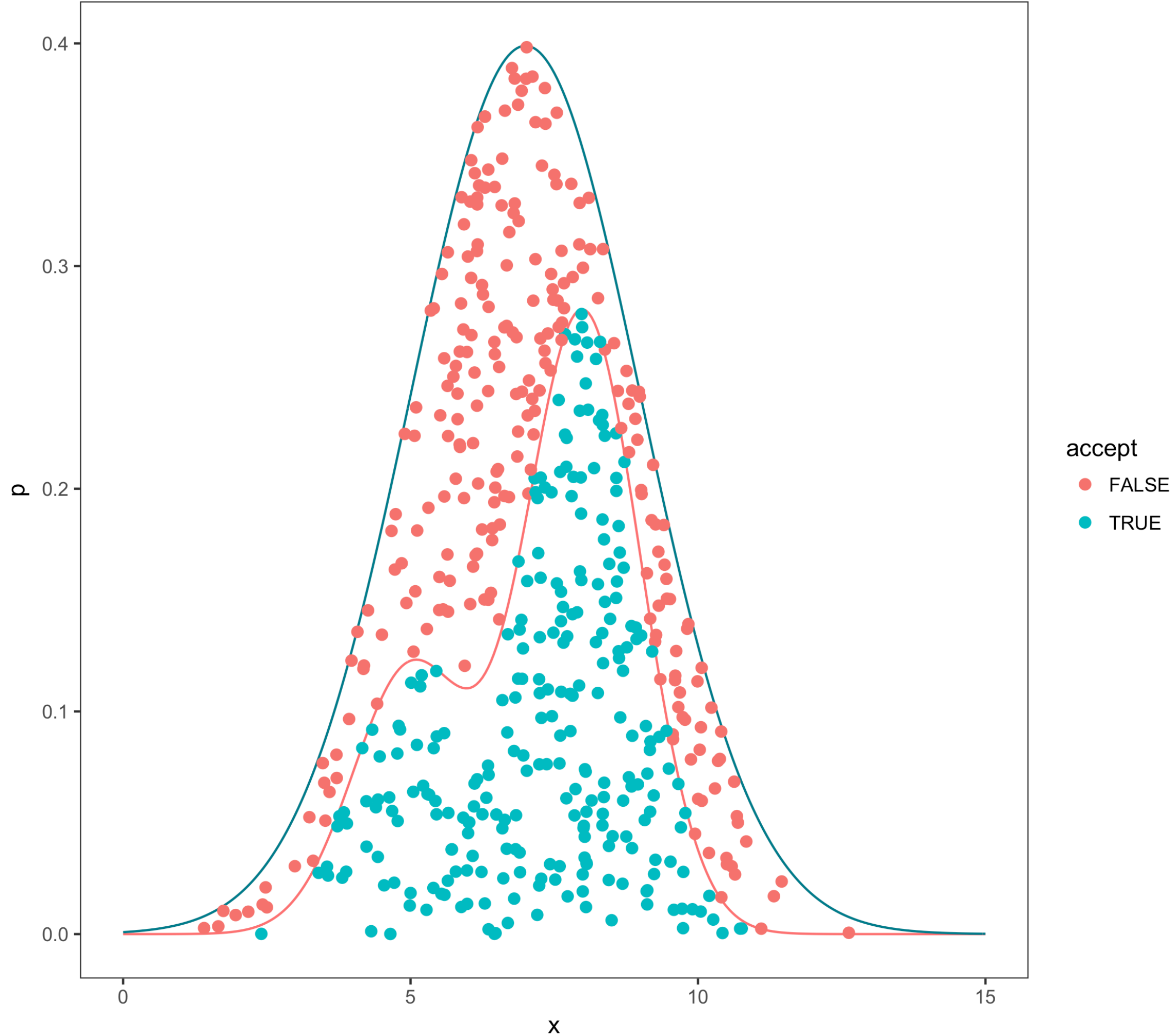
$$P^*(x)$$

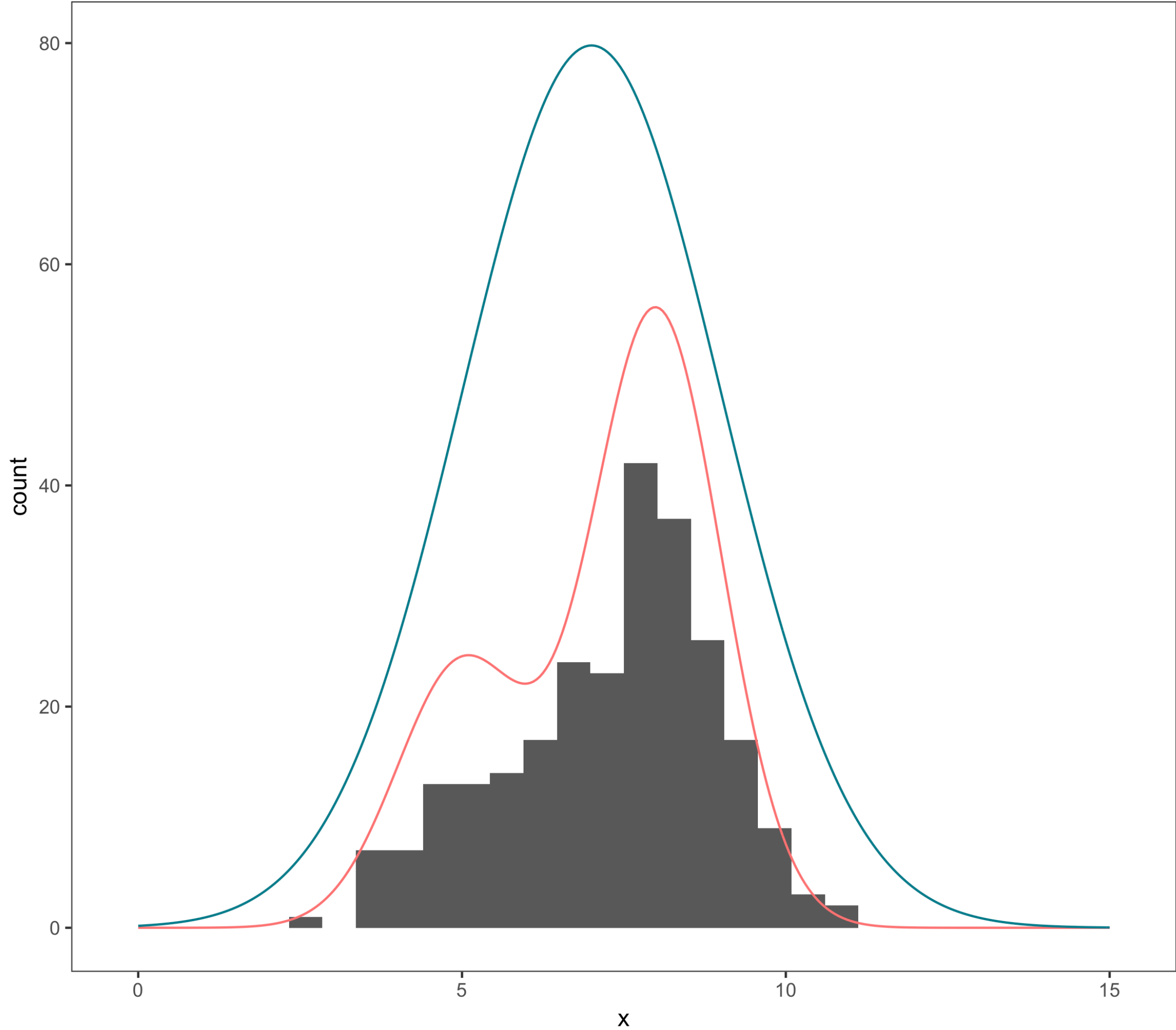
REJECTION SAMPLING









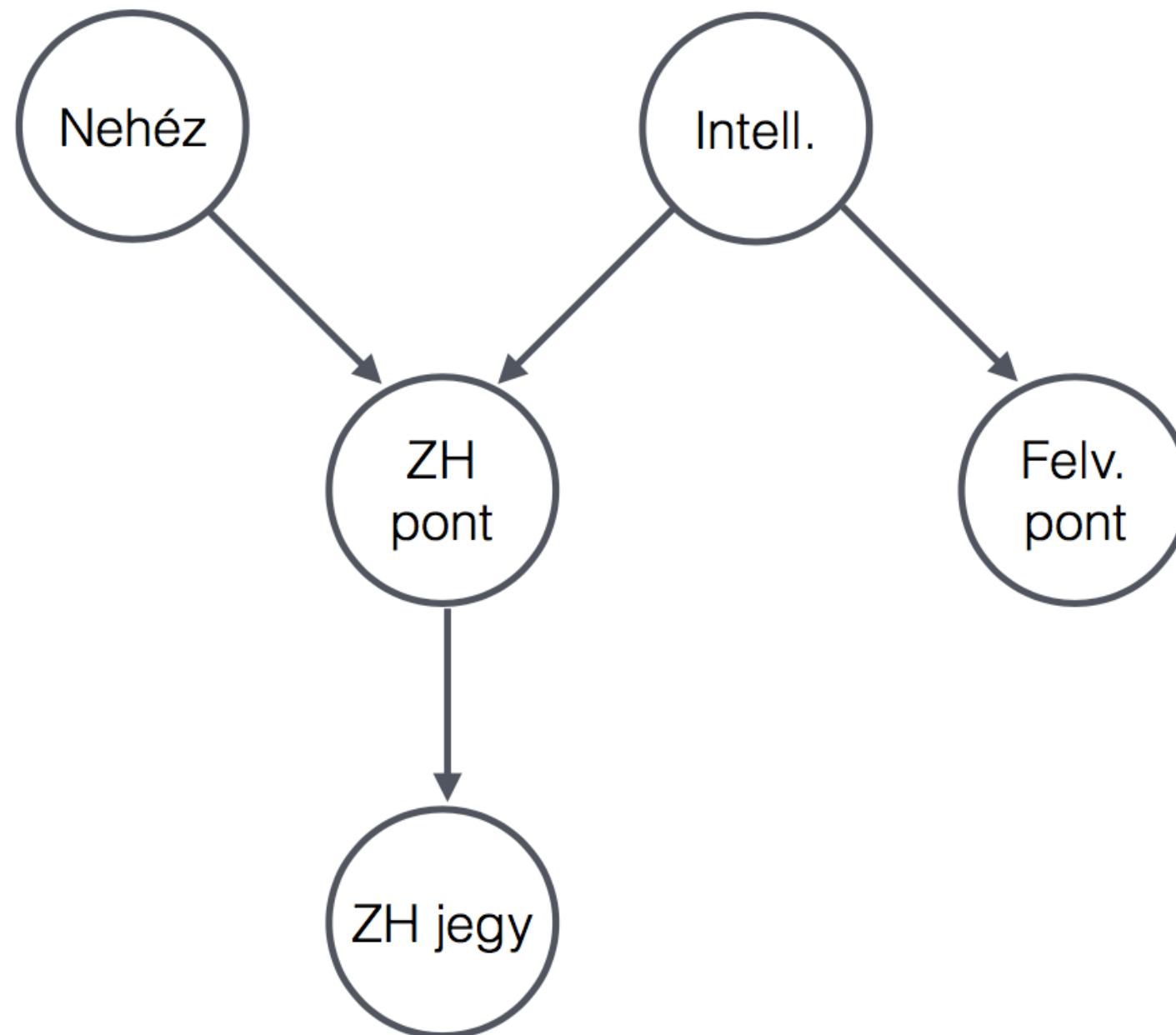


REJECTION SAMPLING

- ▶ Problémák:

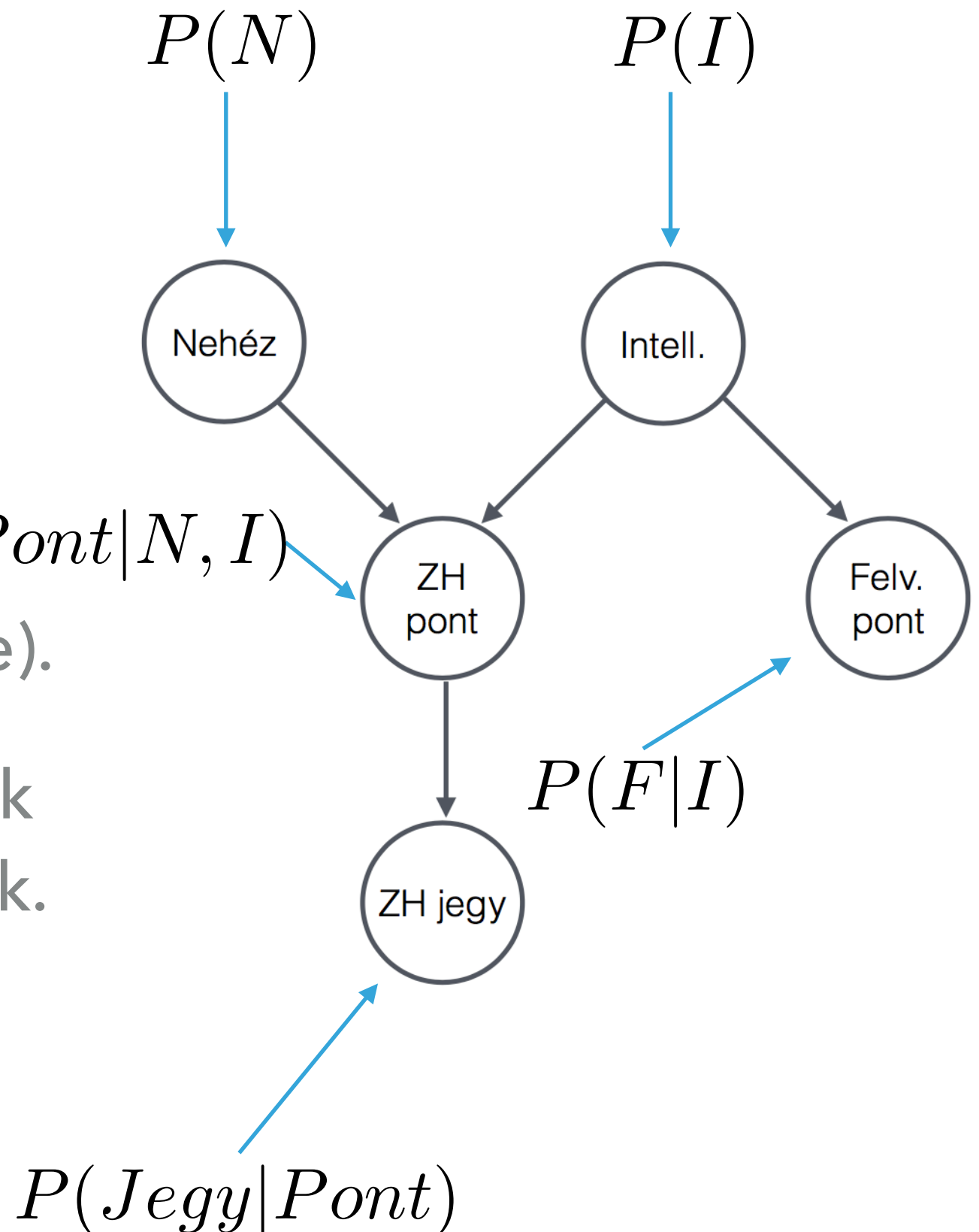
- ▶ Ha $CQ(x)$ nem mindenhol nagyobb, mint $P^*(x)$, akkor nem a megfelelő eloszlásból mintavételezünk.

ANCESTOR SAMPLING



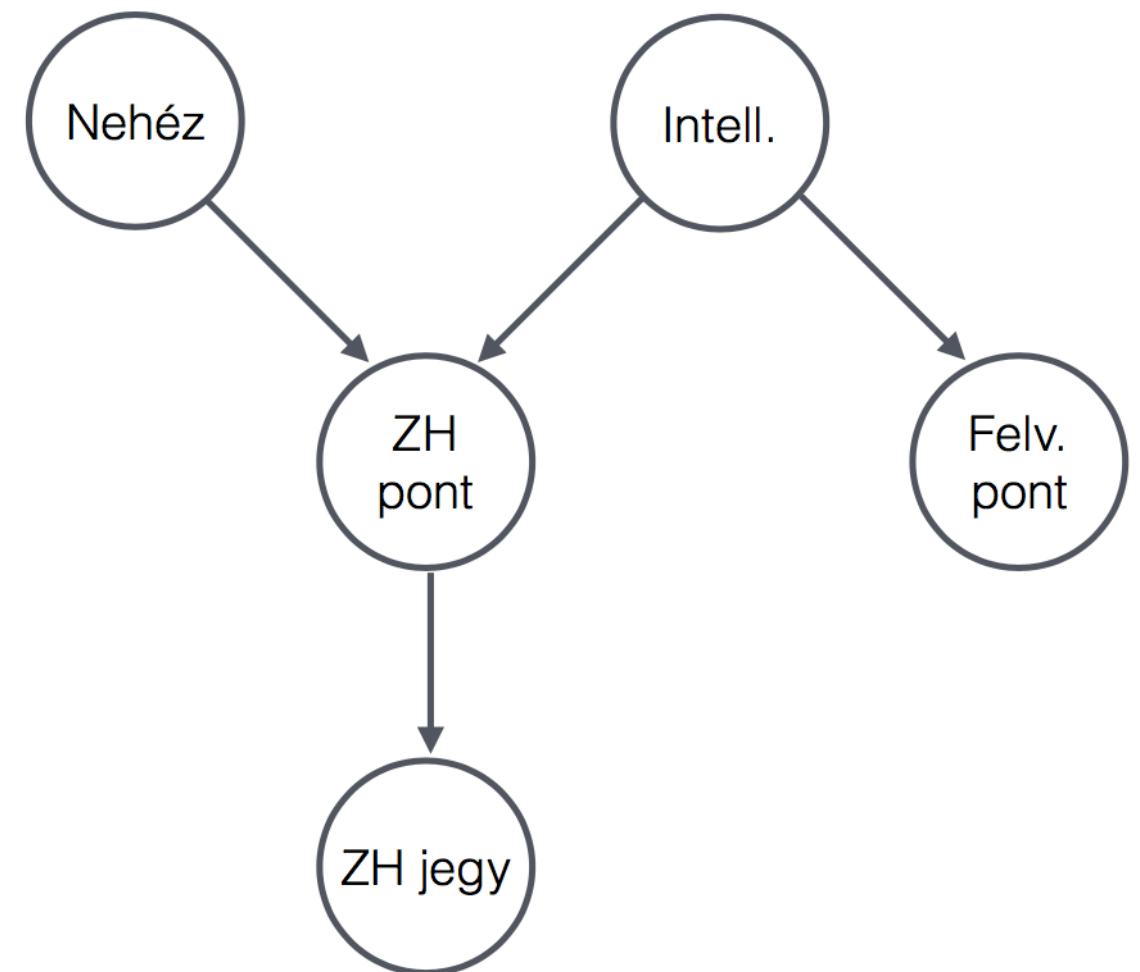
ANCESTOR SAMPLING

- ▶ Szeretnénk egy mintát az együttes eloszlásból.
- ▶ 1. Először mintát veszünk a "legfelső változókból" $P(Pont|N, I)$ (amelyeknek nincsen szülője).
- ▶ 2. Majd azokból, amelyeknek a szüleit már mintavételeztük.
- ▶ 2. lépést ismételjük



ANCESTOR SAMPLING

- ▶ Ha tudunk mintákat venni $P(N, I, F, Pont, Jegy)$ eloszlásból, hogy tudunk inferenciát végezni, vagyis, kondicionálisokat számolni?
- ▶ pl: $P(I | Jegy = 5) = ?$
- ▶ Rejection sampling!
Azokat a mintákat nézzük, amelyekben a Jegy a megadott értékű



REJECTION SAMPLING PROBLÉMA

- ▶ Dimenzió növekedésével csökken az elfogadott minták aránya
- ▶ Nehéz az eloszlást felülről becsülni.

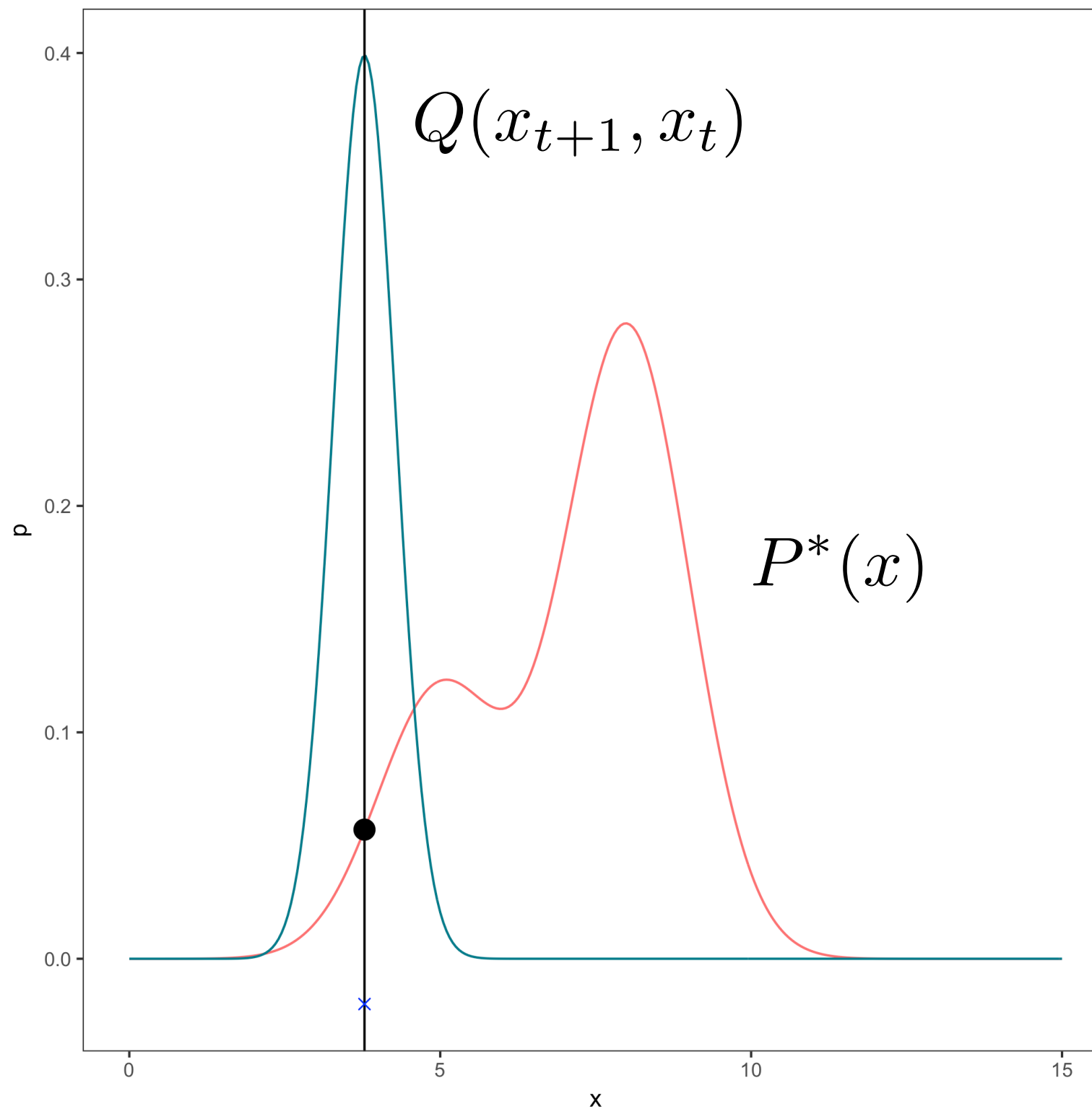
INFERENCIA II.

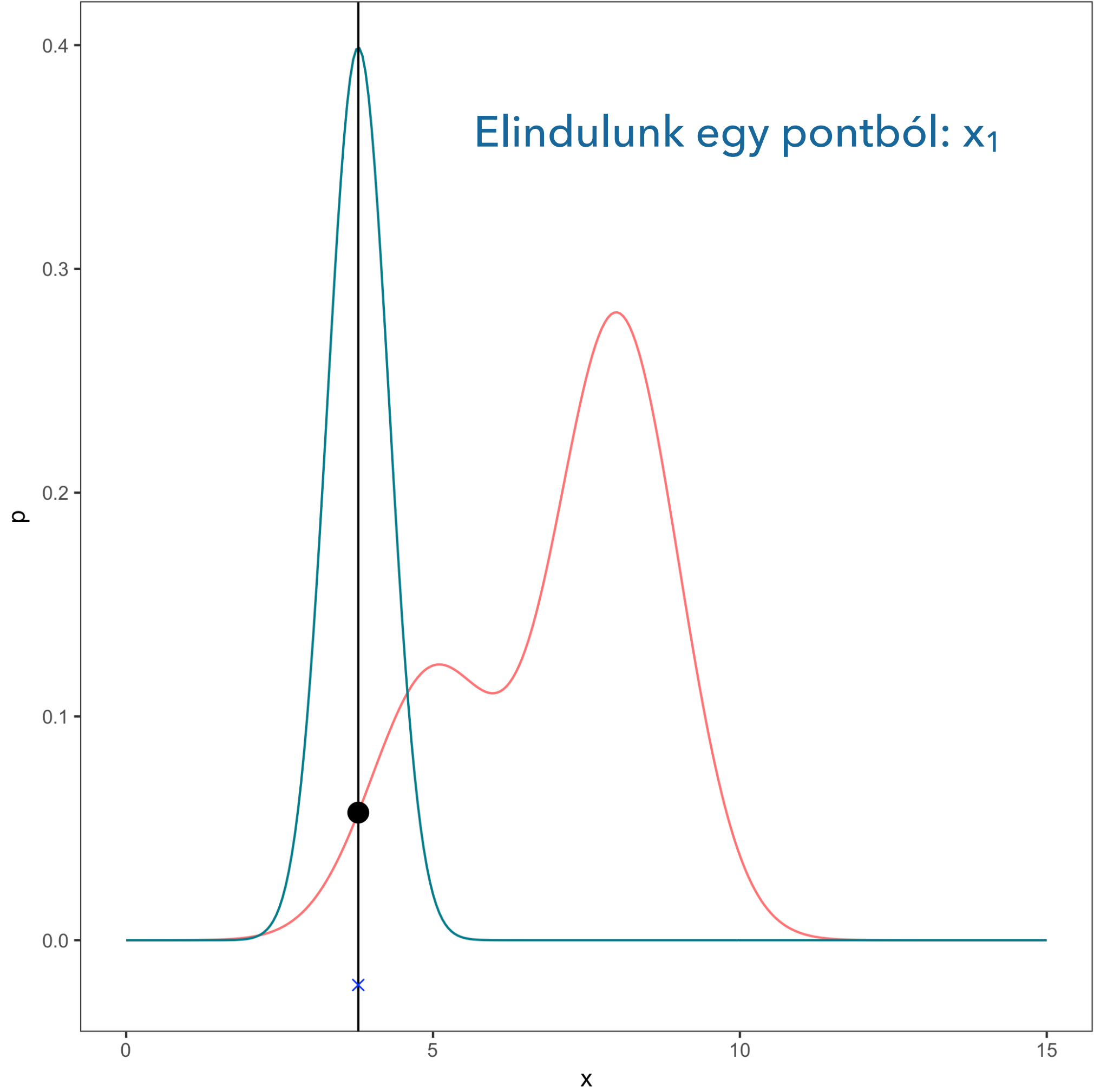
- ▶ Közelítés szükségessége
- ▶ Mintavételezés
- ▶ Egyszerűbb Monte Carlo módszerek
- ▶ Markov Chain Monte Carlo módszerek

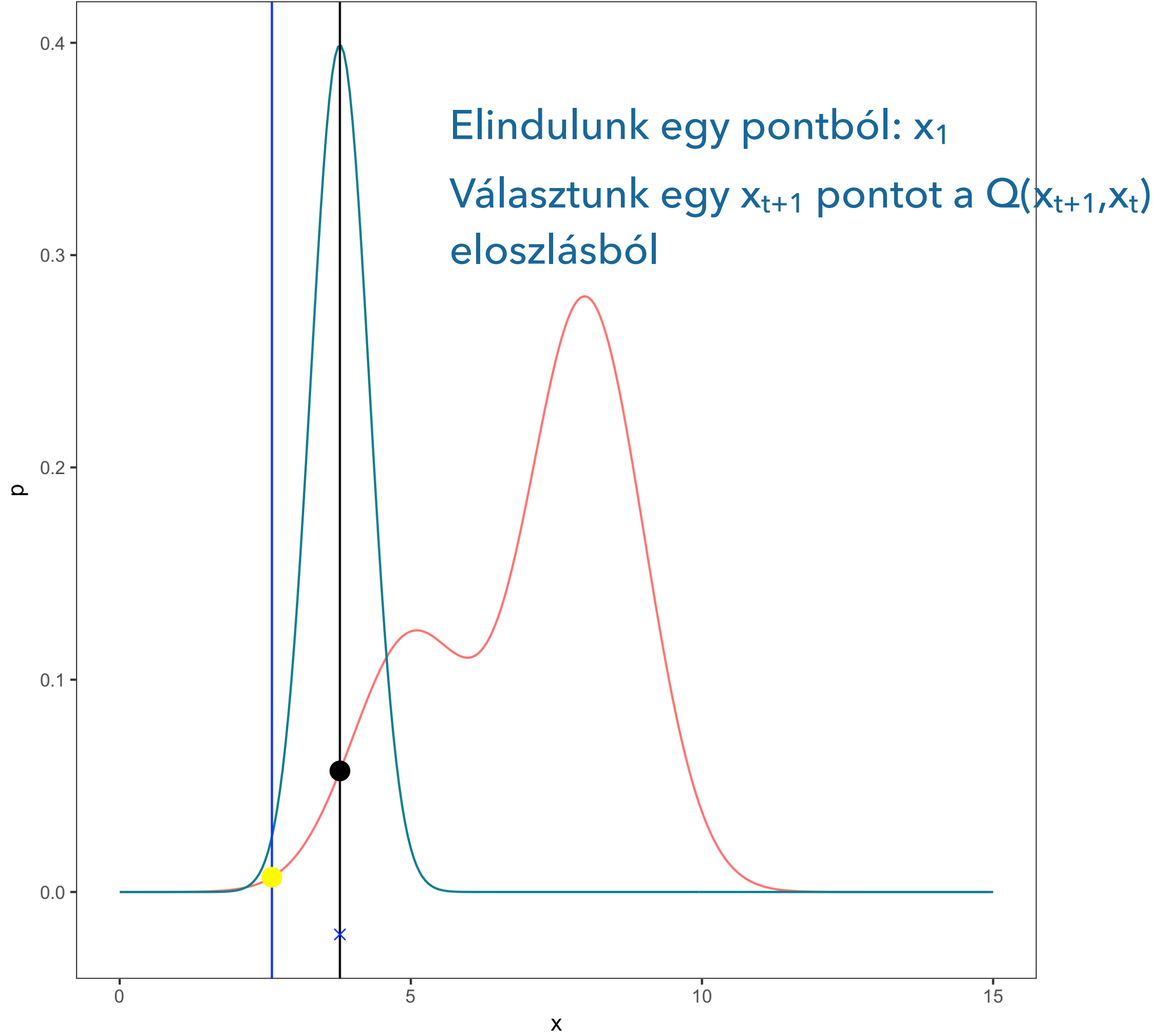
MARKOV CHAIN MONTE CARLO

- ▶ Ahhoz, hogy a dimenzionalitás problémáját kikerüljük, megpróbálhatjuk felhasználni az eddigi mintáinkban rejlő információt.
- ▶ Viszont elveszítjük a függetlenséget!

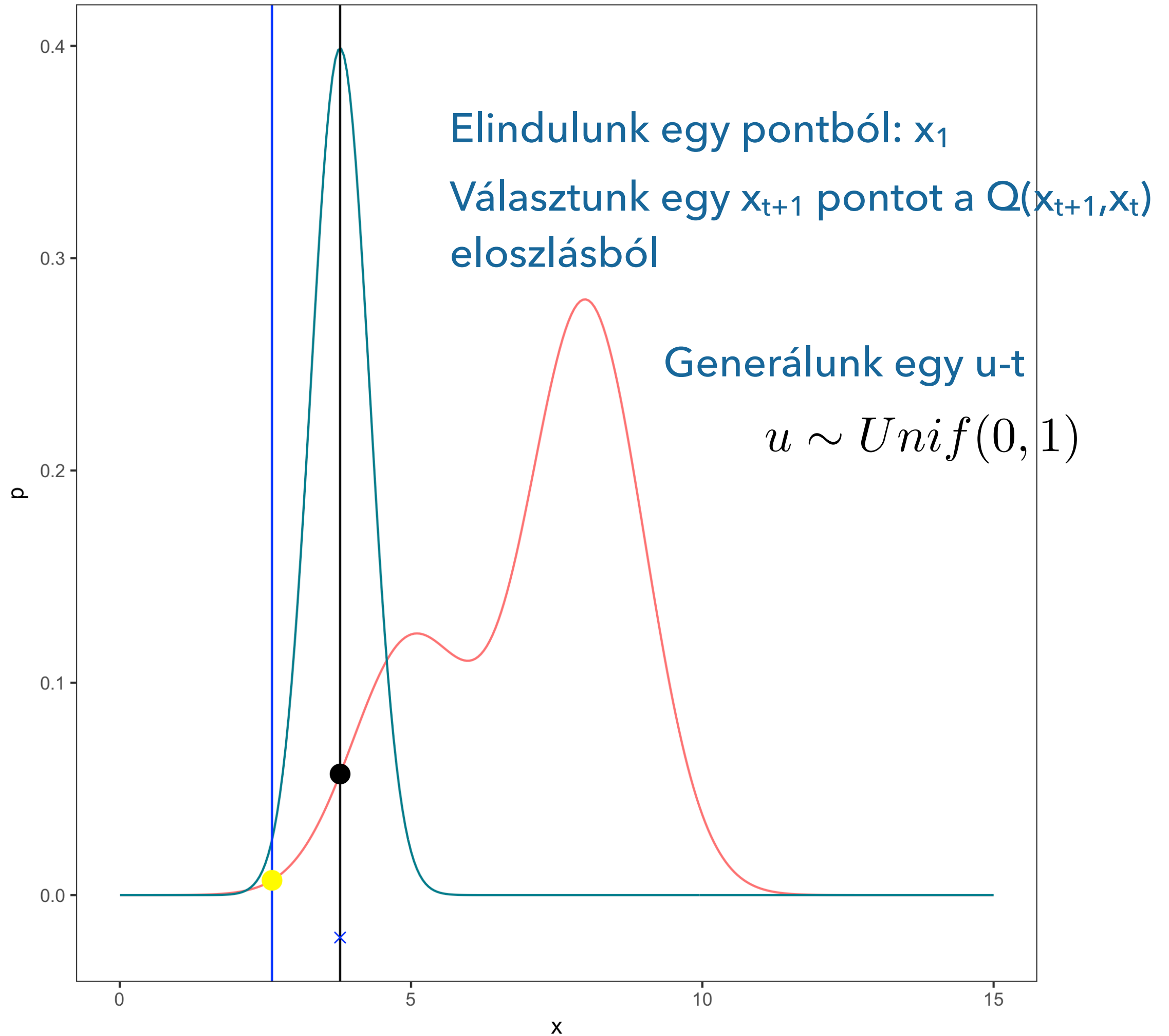
METROPOLIS-HASTINGS ALGORITMUS

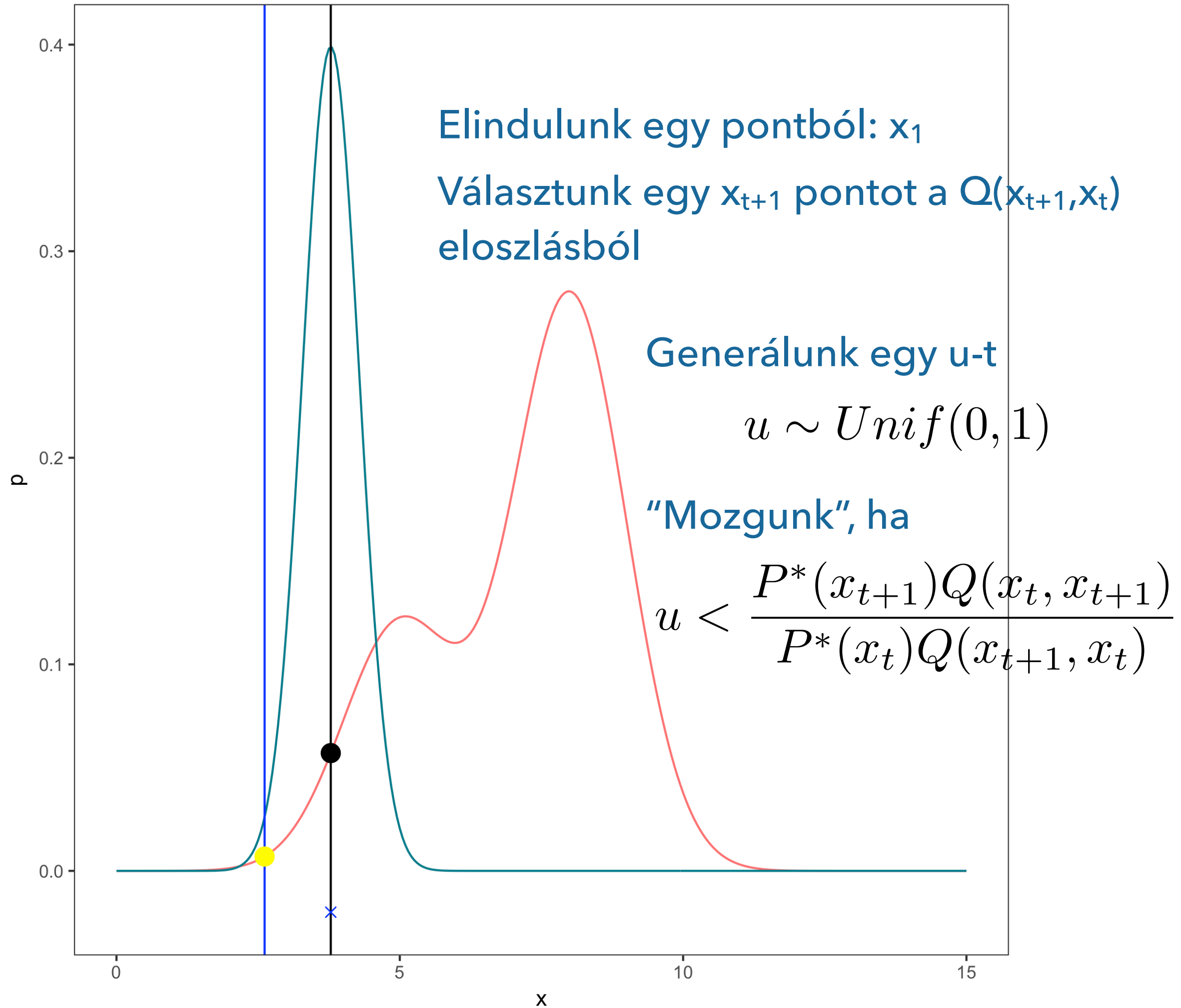


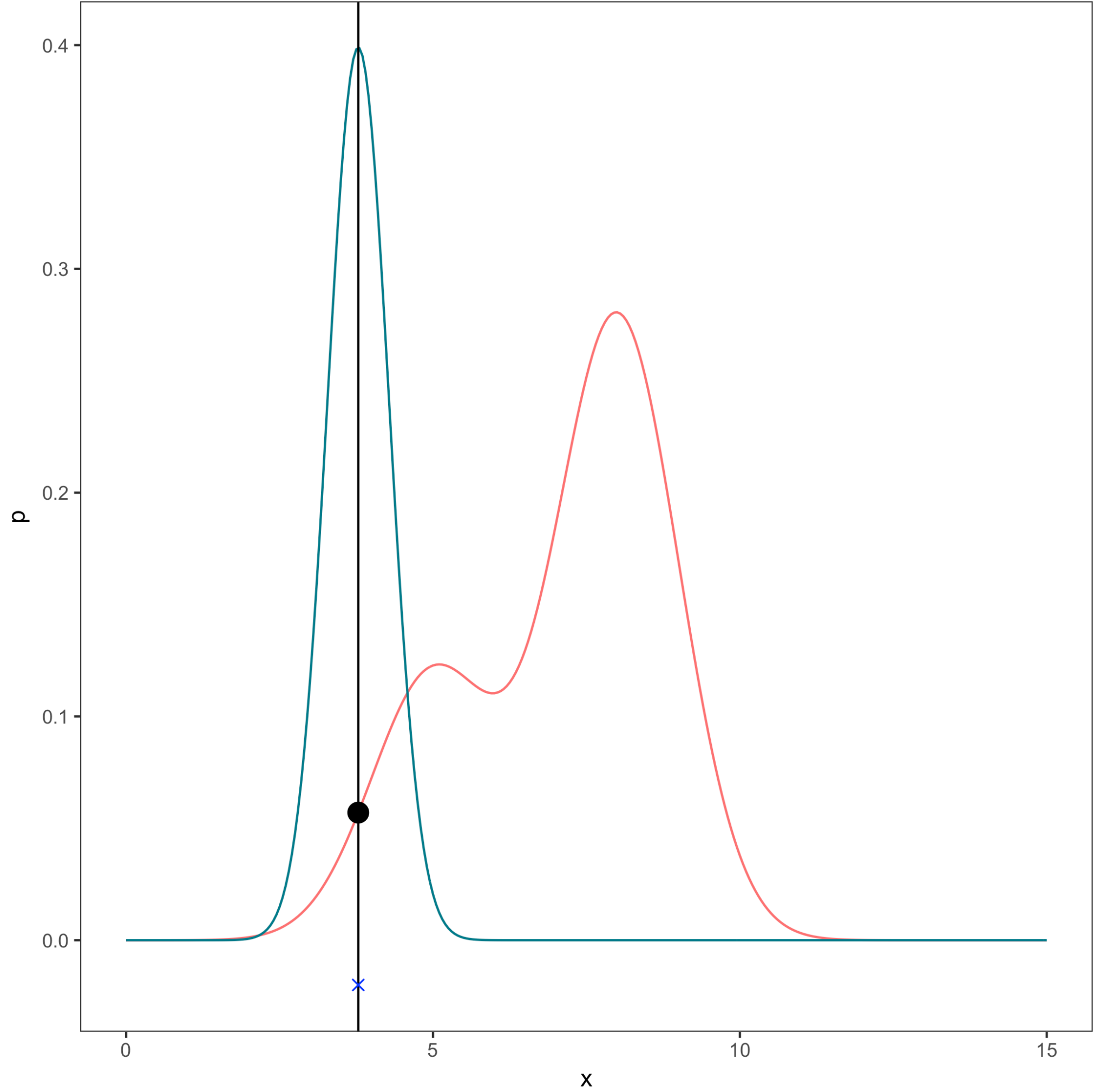


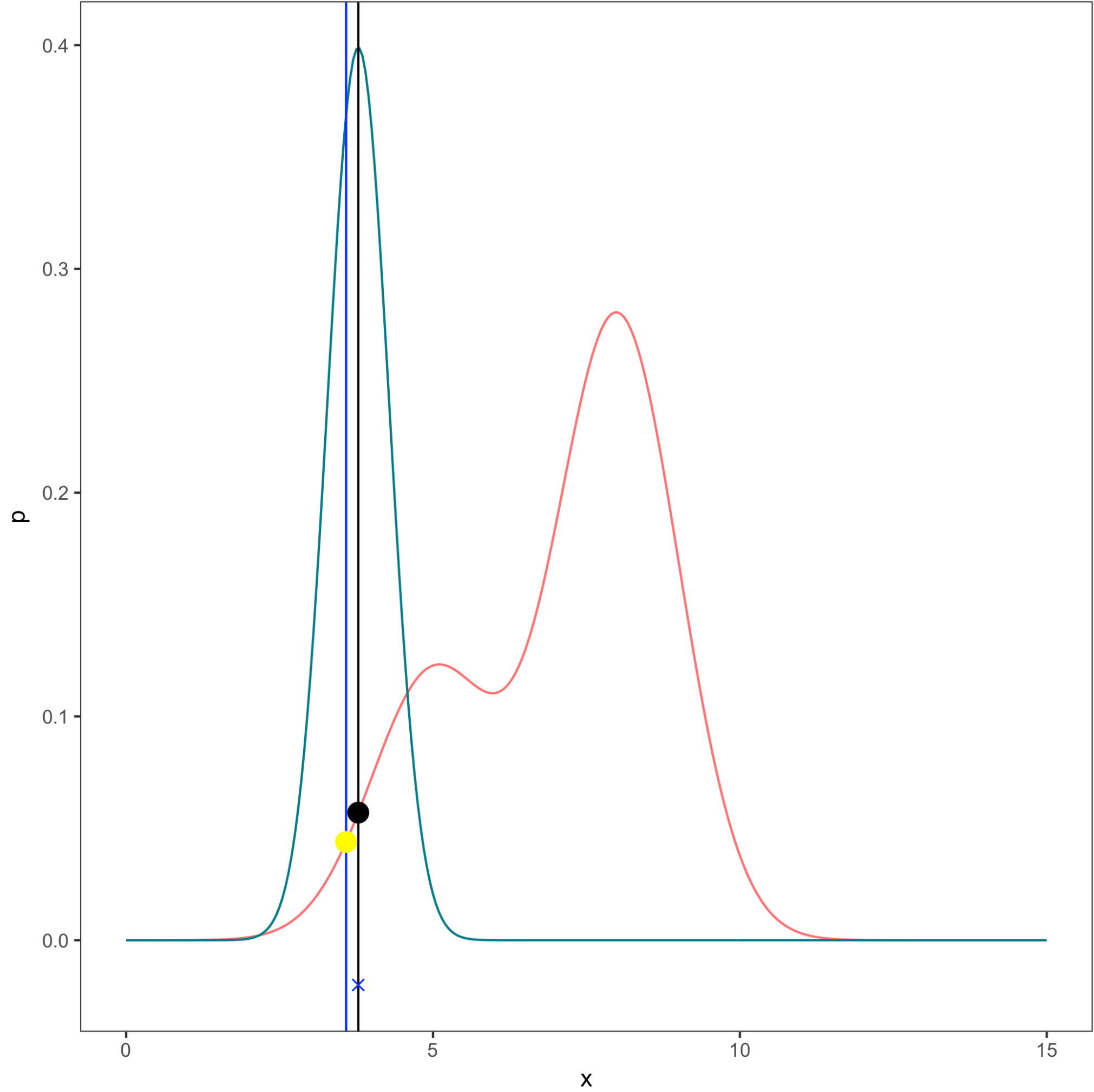


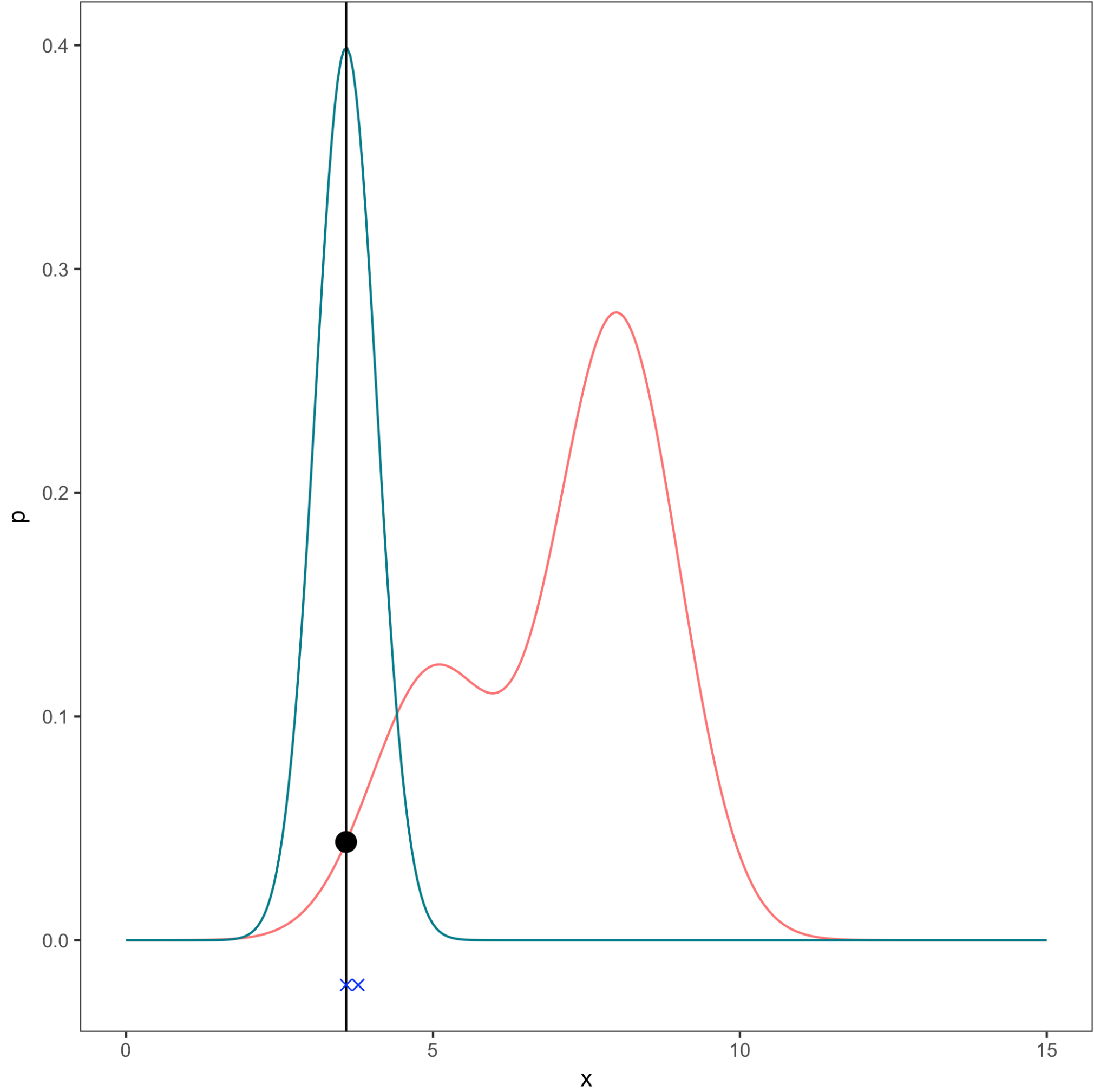
Elindulunk egy pontból: x_1
Választunk egy x_{t+1} pontot a $Q(x_{t+1}, x_t)$
eloszlásból

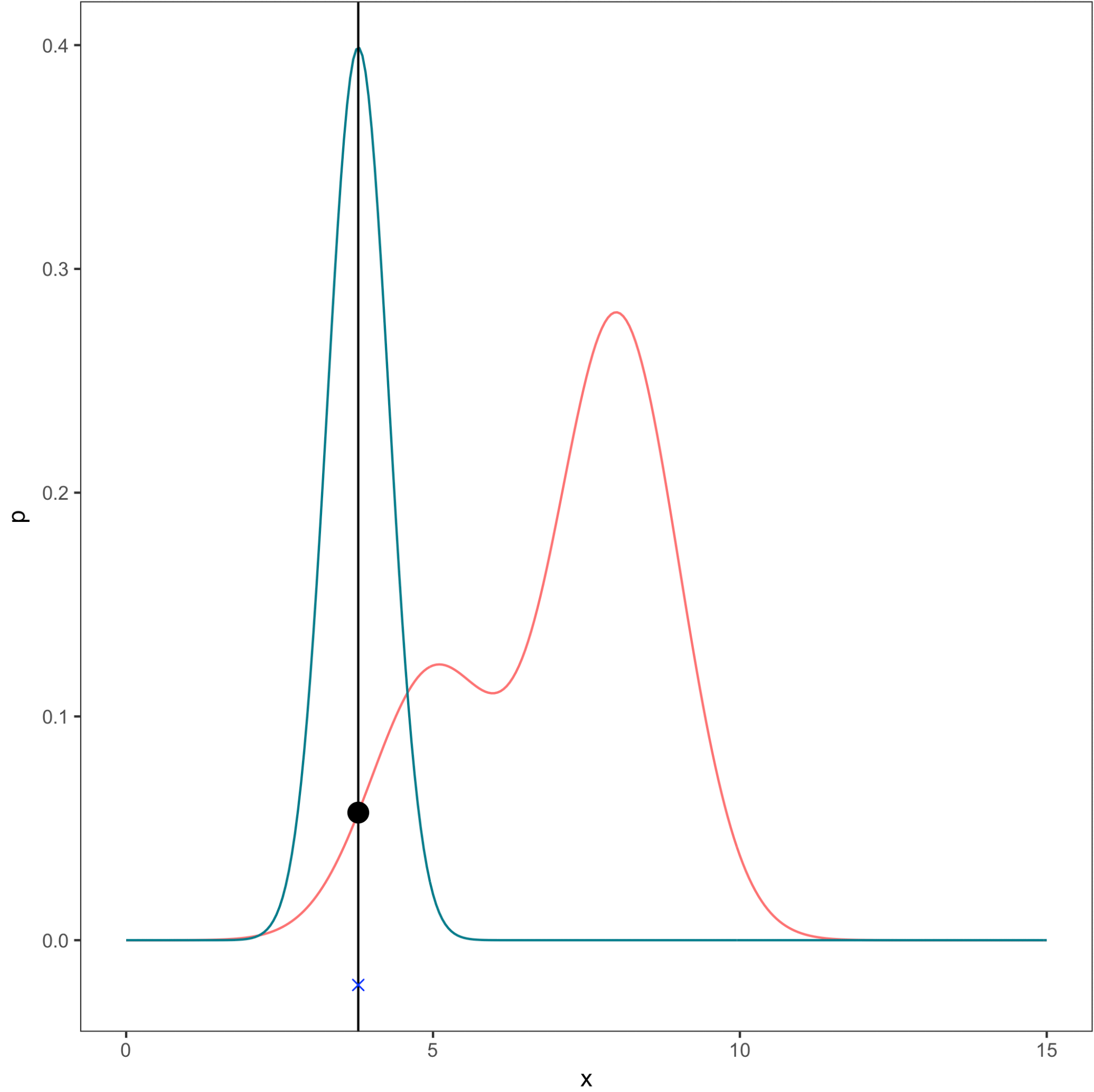


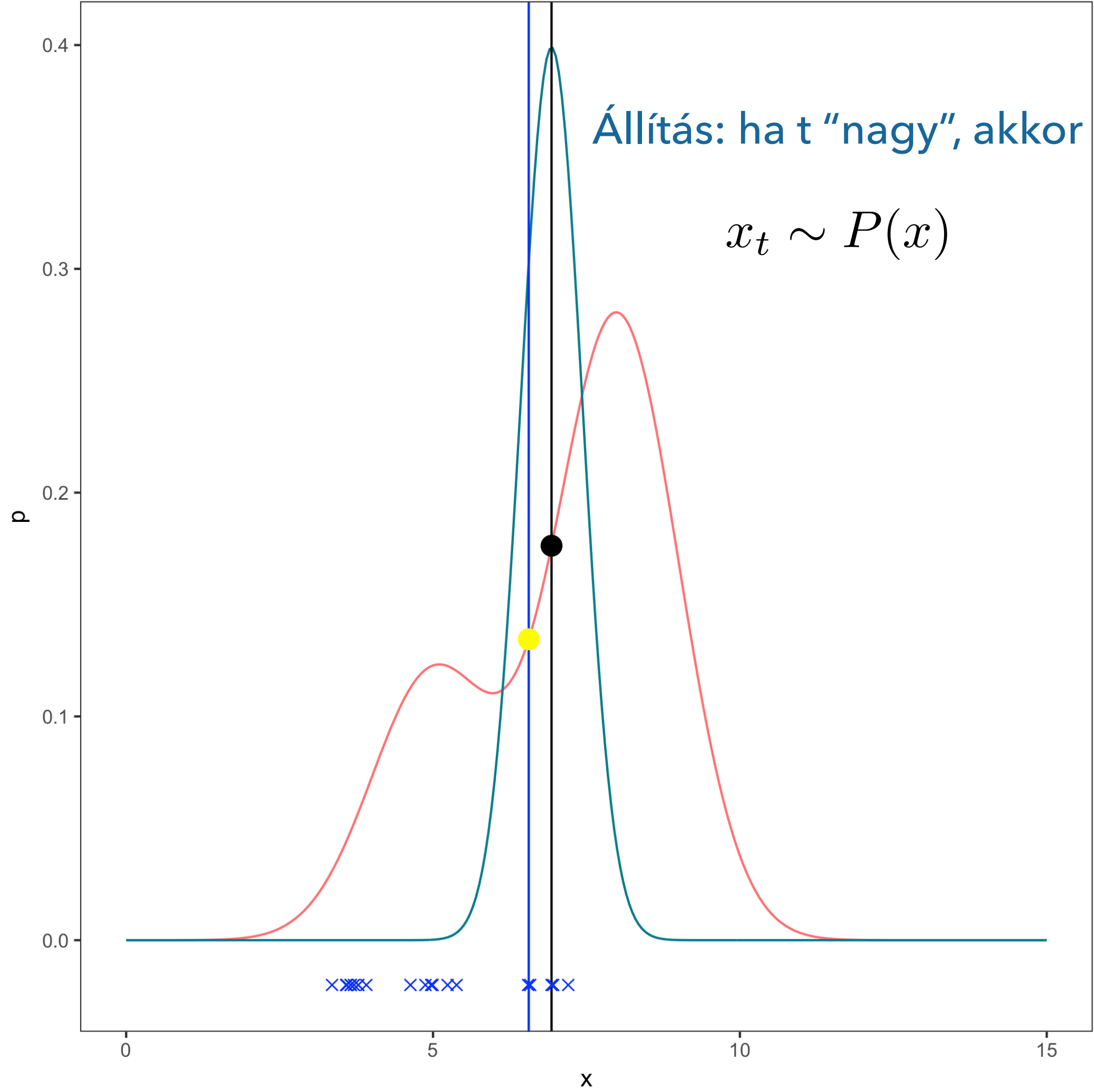


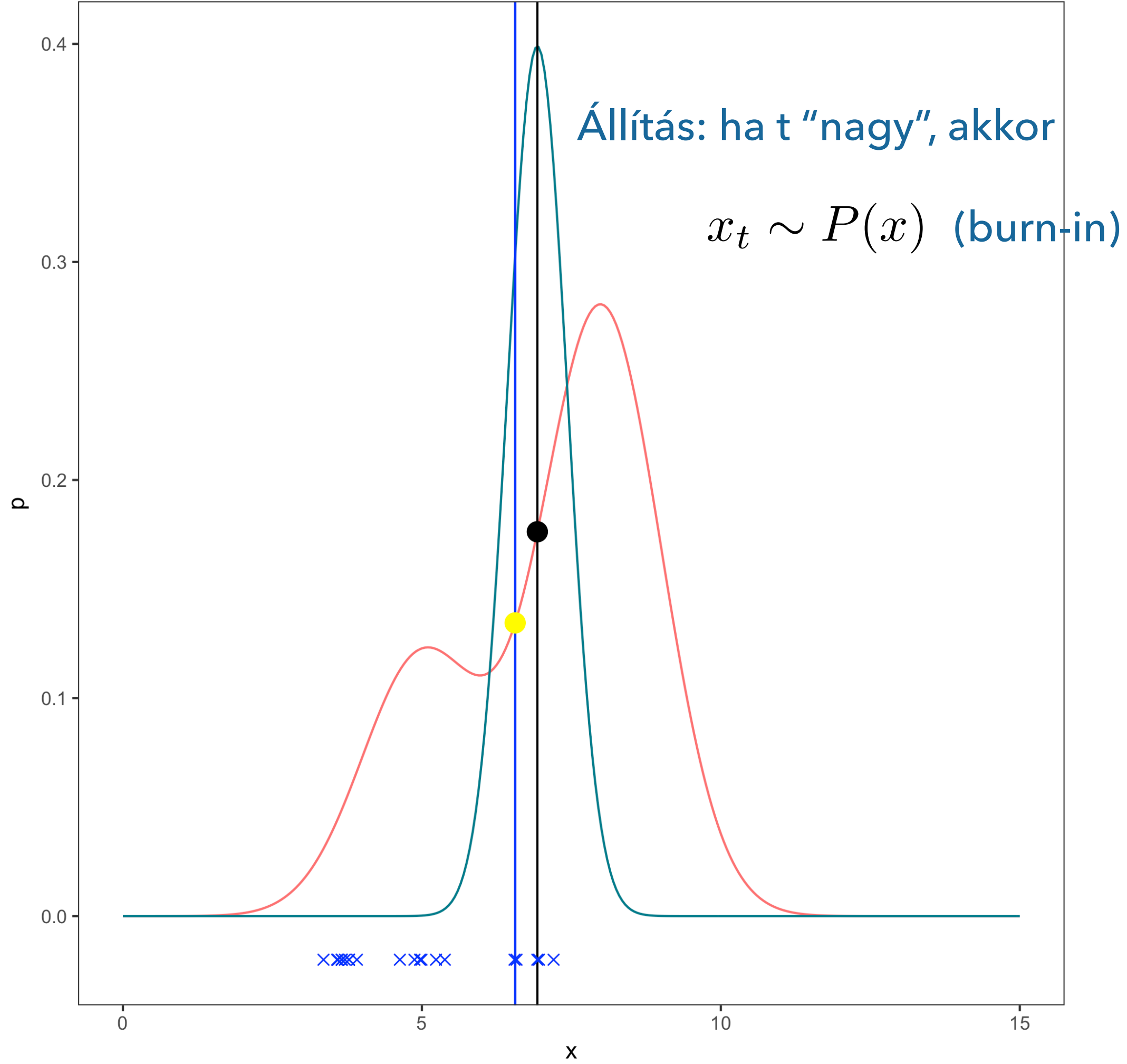


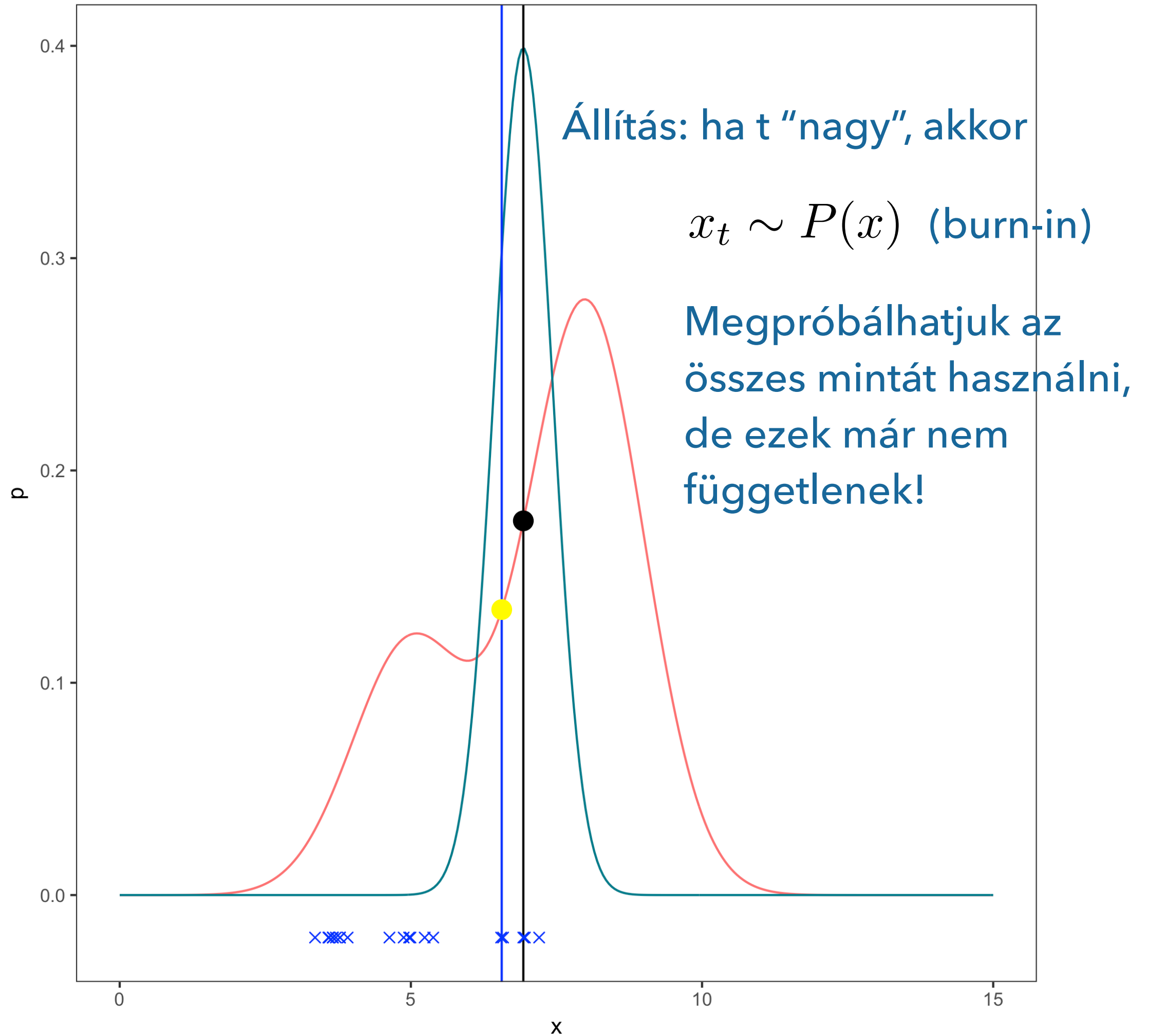


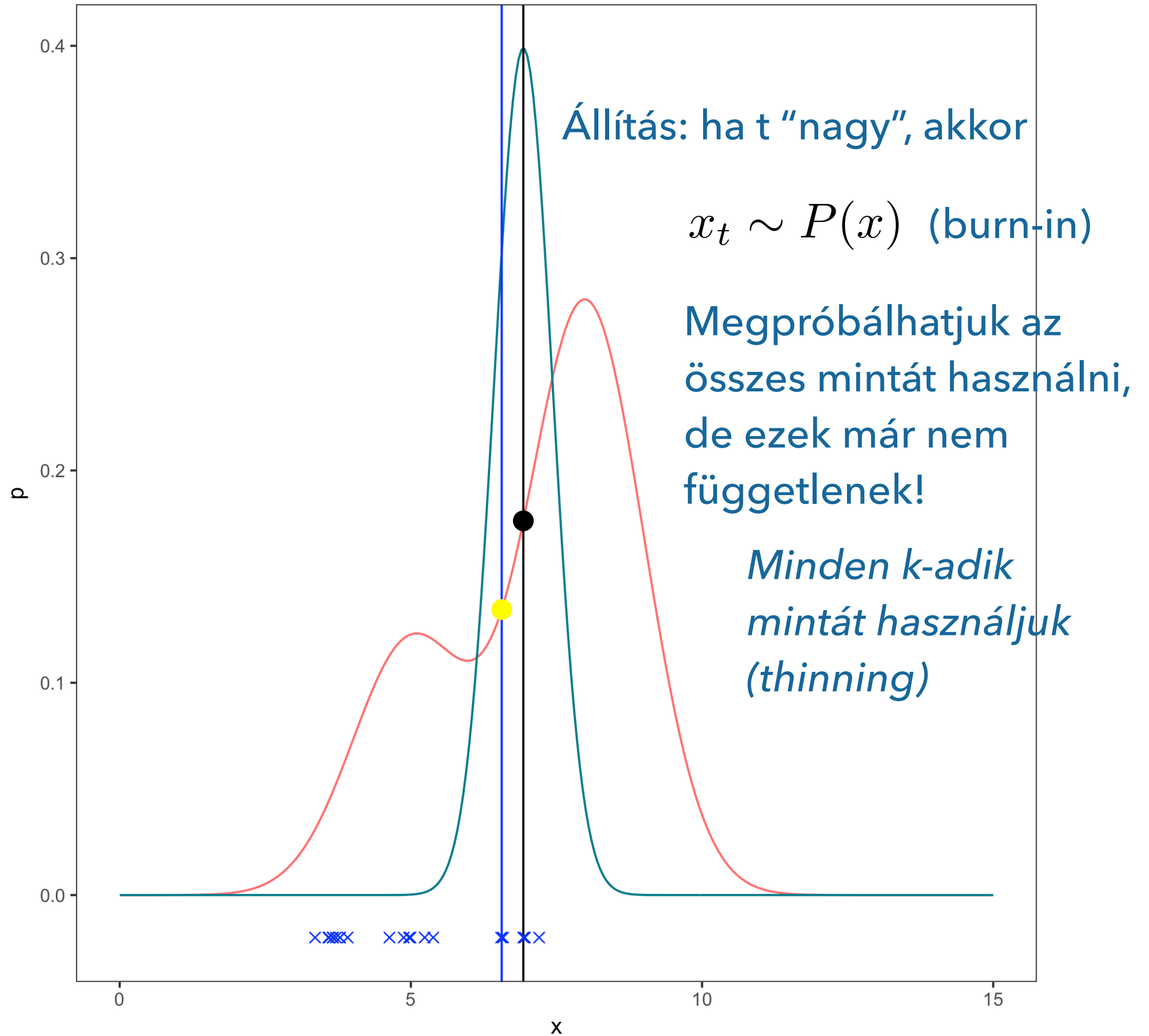












METROPOLIS HASTINGS VS. REJECTION SAMPLING

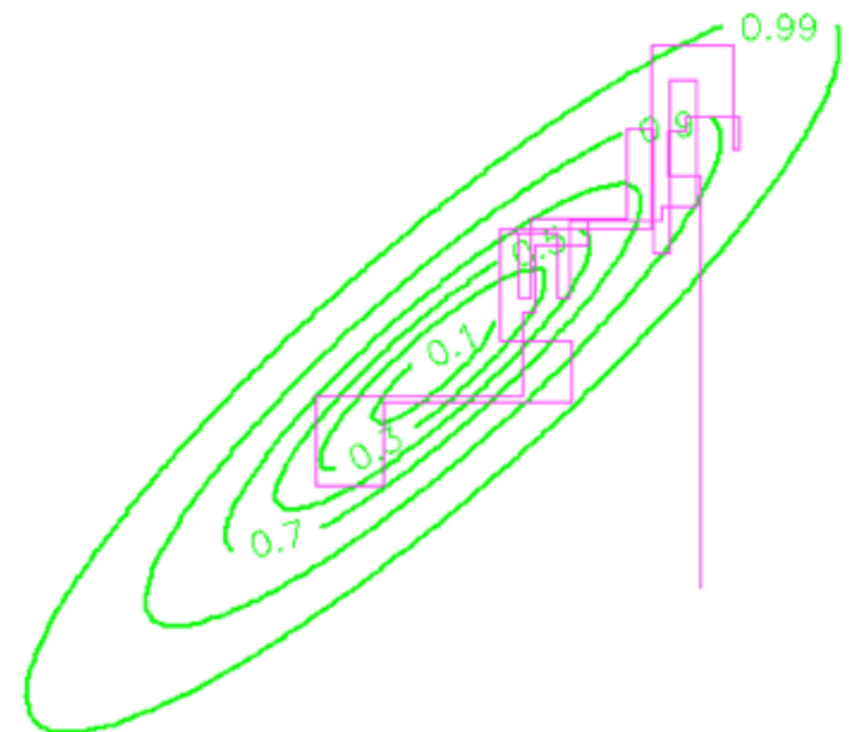
- ▶ MH: az algoritmikus lépések nem függenek a céleloszlás $P(x)$ alakjától. Emiatt általános inferencia-algoritmus fejleszthető belőle. (Általános, mert bármilyen modellre alkalmazható, lásd: STAN)
 - ▶ a rejection sampling esetében megfelelő $Q(x)$ mintavételező eloszlás kell ($CQ(x) > P^*(x)$)
- ▶ Kevésbé sújtja a “dimenzionalitás átka”

GIBBS SAMPLING

- ▶ Az egyes változókat sorban mintavételezzük, a többi változóra kondicionálva:

$$x_{t+1}^{(1)} \sim p(x^{(1)} | x_t^{(2)})$$

$$x_{t+1}^{(2)} \sim p(x^{(2)} | x_{t+1}^{(1)})$$



ÖSSZEFOGLALÁS

- ▶ Valószínűségi modellben végzünk inferenciát
 - ▶ Egzakt poszterior-számítás nem lehetséges
 - ▶ Közelítő inferenciát használunk (pontbecslés, variációs módszer, mintavételezés)
- ▶ Mintavételezés

ÖSSZEFOGLALÁS

- ▶ Mintavételezés:
 - ▶ Egyszerű eloszlásokból minták (kockadobás, uniform)
 - ▶ Rejection sampling (speciálisan pl. kondicionálisok ancestor samplinggal)
 - ▶ Metropolis-Hastings algoritmus
 - ▶ Gibbs-sampling

SEGÉDANYAGOK

- ▶ David Mackay video lectures 12. rész
https://www.youtube.com/watch?v=sN_0iGWcyLI
- ▶ Bishop: Pattern recognition and machine learning,
11. fejezet