Bayesi modell-összehasonlítás és alkalmazási irányok

Nyelvi adatok feldolgozása

Borbély Gábor

2020.05.18.

Gépi tanulás általában

- Legyen $\left(x_j, y_j\right)_{j=1}^{n}, x_j \in X, y_j \in Y$ valamilyen adat.

Gépi tanulás általában

- Legyen $(x_j, y_j)_{j=1}^{n}, x_j \in X, y_j \in Y$ valamilyen adat.
- Tegyük fel, hogy egy $\mathbb{P}(y_j \mid x_j, \mathbf{w})$ modellt akarok ($\mathbf{w}$ modell paraméter)

Gépi tanulás általában

- Legyen $(x_j, y_j)_{j=1}^{n}, x_j \in X, y_j \in Y$ valamilyen adat.
- Tegyük fel, hogy egy $\mathbb{P}(y_j \mid x_j, \mathbf{w})$ modellt akarok (**w** modell paraméter)
  - Nem csak döntést hoz, hanem valószínűségeket is ad.

Gépi tanulás általában

- Legyen $(x_j, y_j)_{j=1}^{n}, x_j \in X, y_j \in Y$ valamilyen adat.
- Tegyük fel, hogy egy $\mathbb{P}(y_j \mid x_j, \mathbf{w})$ modellt akarok (**w** modell paraméter)
  - Nem csak döntést hoz, hanem valószínűségeket is ad.
- Vagy a feladat legyen címkézetlen, felügyelet nélküli tanulás (representation learning)

Gépi tanulás általában

- Legyen $(x_j, y_j)_{j=1}^{n}, x_j \in X, y_j \in Y$ valamilyen adat.
- Tegyük fel, hogy egy $\mathbb{P}(y_j \mid x_j, \mathbf{w})$ modellt akarok ($\mathbf{w}$ modell paraméter)
  - Nem csak döntést hoz, hanem valószínűségeket is ad.
- Vagy a feladat legyen címkézetlen, felügyelet nélküli tanulás (representation learning)
  - Ekkor a feladatunk megtanulni az adatnak az X tér feletti eloszlását:

$$\mathbb{P}(x_j \mid \mathbf{w})$$

Gépi tanulás általában

- Legyen $(x_j, y_j)_{j=1}^{n}, x_j \in X, y_j \in Y$ valamilyen adat.
- Tegyük fel, hogy egy $\mathbb{P}(y_j \mid x_j, \mathbf{w})$ modellt akarok ($\mathbf{w}$ modell paraméter)
  - Nem csak döntést hoz, hanem valószínűségeket is ad.
- Vagy a feladat legyen címkézetlen, felügyelet nélküli tanulás (representation learning)
  - Ekkor a feladatunk megtanulni az adatnak az X tér feletti eloszlását:

$$\mathbb{P}(x_j \mid \mathbf{w})$$

  - a nyelvmodellezés is ilyen (pl. word2vec):

$$\max_{\text{szóvektorok}} \mathbb{P}(\text{szöveg} \mid \text{szóvektorok})$$

Gépi tanulás általában

- Legyen $(x_j, y_j)_{j=1}^{n}, x_j \in X, y_j \in Y$ valamilyen adat.
- Tegyük fel, hogy egy $\mathbb{P}(y_j \mid x_j, \mathbf{w})$ modellt akarok ($\mathbf{w}$ modell paraméter)
  - Nem csak döntést hoz, hanem valószínűségeket is ad.
- Vagy a feladat legyen címkézetlen, felügyelet nélküli tanulás (representation learning)
  - Ekkor a feladatunk megtanulni az adatnak az X tér feletti eloszlását:

$$\mathbb{P}(x_j \mid \mathbf{w})$$

  - a nyelvmodellezés is ilyen (pl. word2vec):

$$\max_{\text{szóvektorok}} \mathbb{P}(\text{szöveg} \mid \text{szóvektorok})$$

- A most következő módszer használható címkézett (felügyelt) tanításhoz is.

Representation Learning

- GAN, WGAN

Representation Learning

- GAN, WGAN
- LM, a szöveg perplexitásának minimalizálása (a modell paraméterekben)

Representation Learning

- GAN, WGAN
- LM, a szöveg perplexitásának minimalizálása (a modell paraméterekben)
- Mi van, ha több modell is van?

Representation Learning

- GAN, WGAN
- LM, a szöveg perplexitásának minimalizálása (a modell paraméterekben)
- Mi van, ha több modell is van?
- Legyen $\mathbf{w}_i \in \mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ több tér, melyek mindegyikében még $\mathbf{w}_i$-t is ki kell választanom.

Representation Learning

- GAN, WGAN
- LM, a szöveg perplexitásának minimalizálása (a modell paraméterekben)
- Mi van, ha több modell is van?
- Legyen $\mathbf{w}_i \in \mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ több tér, melyek mindegyikében még $\mathbf{w}_i$-t is ki kell választanom.
- Melyik a *jobb*?

Representation Learning

- GAN, WGAN
- LM, a szöveg perplexitásának minimalizálása (a modell paraméterekben)
- Mi van, ha több modell is van?
- Legyen $\mathbf{w}_i \in \mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ több tér, melyek mindegyikében még $\mathbf{w}_i$-t is ki kell választanom.
- Melyik a *jobb*?
  - Gondolhatnánk, hogy ha jobban illeszkedik az adatra, akkor jobb.

$$\operatorname{dist}\left(\mathbb{P}(x_j \mid \mathbf{w}_i), \mathbb{P}(x_j)\right) \to \min$$

Representation Learning

- GAN, WGAN
- LM, a szöveg perplexitásának minimalizálása (a modell paraméterekben)
- Mi van, ha több modell is van?
- Legyen $\mathbf{w}_i \in \mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ több tér, melyek mindegyikében még $\mathbf{w}_i$-t is ki kell választanom.
- Melyik a *jobb*?
  - Gondolhatnánk, hogy ha jobban illeszkedik az adatra, akkor jobb.

$$\mathrm{dist}\left(\mathbb{P}(x_j \mid \mathbf{w}_i), \mathbb{P}(x_j)\right) \to \min$$

  - De ekkor mindig a memorizálás a legjobb.

Representation Learning

- GAN, WGAN
- LM, a szöveg perplexitásának minimalizálása (a modell paraméterekben)
- Mi van, ha több modell is van?
- Legyen $\mathbf{w}_i \in \mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ több tér, melyek mindegyikében még $\mathbf{w}_i$-t is ki kell választanom.
- Melyik a *jobb*?
  - Gondolhatnánk, hogy ha jobban illeszkedik az adatra, akkor jobb.

$$\text{dist}\left(\mathbb{P}(x_j \mid \mathbf{w}_i), \mathbb{P}(x_j)\right) \to \min$$

  - De ekkor mindig a memorizálás a legjobb.
    - Érdekesség: egy neurális háló (néha) nagyobb, mint az az adat, amin feltanították.

Representation Learning

- GAN, WGAN
- LM, a szöveg perplexitásának minimalizálása (a modell paraméterekben)
- Mi van, ha több modell is van?
- Legyen $\mathbf{w}_i \in \mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ több tér, melyek mindegyikében még $\mathbf{w}_i$-t is ki kell választanom.
- Melyik a *jobb*?
  - Gondolhatnánk, hogy ha jobban illeszkedik az adatra, akkor jobb.

$$\text{dist}\left(\mathbb{P}(x_j \mid \mathbf{w}_i), \mathbb{P}(x_j)\right) \to \min$$

  - De ekkor mindig a memorizálás a legjobb.
    - Érdekesség: egy neurális háló (néha) nagyobb, mint az az adat, amin feltanították.
  - Például nyelvmodellezésben ha túl nagy *n*-gramokat veszek

Representation Learning

- GAN, WGAN
- LM, a szöveg perplexitásának minimalizálása (a modell paraméterekben)
- Mi van, ha több modell is van?
- Legyen $\mathbf{w}_i \in \mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ több tér, melyek mindegyikében még $\mathbf{w}_i$-t is ki kell választanom.
- Melyik a *jobb*?
  - Gondolhatnánk, hogy ha jobban illeszkedik az adatra, akkor jobb.

$$\text{dist}\left(\mathbb{P}(x_j \mid \mathbf{w}_i), \mathbb{P}(x_j)\right) \to \min$$

  - De ekkor mindig a memorizálás a legjobb.
    - Érdekesség: egy neurális háló (néha) nagyobb, mint az az adat, amin feltanították.
  - Például nyelvmodellezésben ha túl nagy *n*-gramokat veszek
- Regularizáció

Representation Learning

- GAN, WGAN
- LM, a szöveg perplexitásának minimalizálása (a modell paraméterekben)
- Mi van, ha több modell is van?
- Legyen $\mathbf{w}_i \in \mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ több tér, melyek mindegyikében még $\mathbf{w}_i$-t is ki kell választanom.
- Melyik a *jobb*?
  - Gondolhatnánk, hogy ha jobban illeszkedik az adatra, akkor jobb.

$$\mathrm{dist}\left(\mathbb{P}(x_j \mid \mathbf{w}_i), \mathbb{P}(x_j)\right) \to \min$$

  - De ekkor mindig a memorizálás a legjobb.
    - Érdekesség: egy neurális háló (néha) nagyobb, mint az az adat, amin feltanították.
  - Például nyelvmodellezésben ha túl nagy n-gramokat veszek
- Regularizáció
- Occam's razor: lehet rosszabb az illeszkedés, ha kisebb a modell.

Bayesi modell-összehasonlítás – I.

- Legyen $D \subseteq X$ adat, egy adatpont többször is szerepelhet.

Bayesi modell-összehasonlítás – I.

- Legyen $D \subseteq X$ adat, egy adatpont többször is szerepelhet.
  - multiset, multiplicitás: (x_j, n_j) párok, ahol $n_j \in \mathbb{N}$.

Bayesi modell-összehasonlítás – I.

- Legyen $D \subseteq X$ adat, egy adatpont többször is szerepelhet.
  - multiset, multiplicitás: (x_j, n_j) párok, ahol $n_j \in \mathbb{N}$.
- Adott $\mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ terek és mindegyikben egy valószínűségi modell:

$$\mathbb{P}(\bullet \mid \mathbf{w}_i) \quad \mathbf{w}_i \in \mathcal{H}_i$$

  Ezek közül kell választanunk.

Bayesi modell-összehasonlítás – I.

- Legyen $D \subseteq X$ adat, egy adatpont többször is szerepelhet.
  - multiset, multiplicitás: (x_j, n_j) párok, ahol $n_j \in \mathbb{N}$.
- Adott $\mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ terek és mindegyikben egy valószínűségi modell:

$$\mathbb{P}(\bullet \mid \mathbf{w}_i) \quad \mathbf{w}_i \in \mathcal{H}_i$$

  Ezek közül kell választanunk.
- Bayes tétel:

$$\text{“evidence”} = \mathbb{P}(\mathcal{H}_i \mid D) = \frac{\mathbb{P}(D \mid \mathcal{H}_i) \cdot \mathbb{P}(\mathcal{H}_i)}{\mathbb{P}(D)}$$

Bayesi modell-összehasonlítás – I.

- Legyen $D \subseteq X$ adat, egy adatpont többször is szerepelhet.
  - multiset, multiplicitás: (x_j, n_j) párok, ahol $n_j \in \mathbb{N}$.
- Adott $\mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ terek és mindegyikben egy valószínűségi modell:

$$\mathbb{P}(\bullet \mid \mathbf{w}_i) \quad \mathbf{w}_i \in \mathcal{H}_i$$

  Ezek közül kell választanunk.
- Bayes tétel:

$$\text{“evidence”} = \mathbb{P}(\mathcal{H}_i \mid D) = \frac{\mathbb{P}(D \mid \mathcal{H}_i) \cdot \mathbb{P}(\mathcal{H}_i)}{\mathbb{P}(D)}$$

- Ha $\mathbb{P}(\mathcal{H}_i)$ konstans (nincs a priori preferált modell)

$$\mathbb{P}(\mathcal{H}_i \mid D) \propto \mathbb{P}(D \mid \mathcal{H}_i) = \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i) \cdot \mathbb{P}(\mathbf{w}_i \mid \mathcal{H}_i)\, \mathrm{d}\mathbf{w}_i$$

Bayesi modell-összehasonlítás – I.

- Legyen $D \subseteq X$ adat, egy adatpont többször is szerepelhet.
  - multiset, multiplicitás: (x_j, n_j) párok, ahol $n_j \in \mathbb{N}$.
- Adott $\mathcal{H}_i \subseteq \mathbb{R}^{d_i}$ terek és mindegyikben egy valószínűségi modell:

  $$\mathbb{P}(\bullet \mid \mathbf{w}_i) \quad \mathbf{w}_i \in \mathcal{H}_i$$

  Ezek közül kell választanunk.
- Bayes tétel:

  $$\text{“evidence”} = \mathbb{P}(\mathcal{H}_i \mid D) = \frac{\mathbb{P}(D \mid \mathcal{H}_i) \cdot \mathbb{P}(\mathcal{H}_i)}{\mathbb{P}(D)}$$

- Ha $\mathbb{P}(\mathcal{H}_i)$ konstans (nincs a priori preferált modell)

  $$\mathbb{P}(\mathcal{H}_i \mid D) \propto \mathbb{P}(D \mid \mathcal{H}_i) = \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i) \cdot \mathbb{P}(\mathbf{w}_i \mid \mathcal{H}_i)\, \mathrm{d}\mathbf{w}_i$$

- Ha $\mathbb{P}(\mathbf{w}_i \mid \mathcal{H}_i)$ konstans (nincs a priori preferált paraméter érték)

  $$\max_i \frac{1}{\mathrm{Vol}(\mathcal{H}_i)} \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i)\, \mathrm{d}\mathbf{w}_i$$

Bayesi modell-összehasonlítás – II.

- $\frac{1}{\mathsf{Vol}(\mathcal{H}_i)} \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i)\, \mathrm{d}\mathbf{w}_i$

Bayesi modell-összehasonlítás – II.

- $\frac{1}{\mathsf{Vol}(\mathcal{H}_i)} \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i) \, \mathsf{d}\mathbf{w}_i$
- (MacKay, 2003)

Bayesi modell-összehasonlítás – II.

- $\frac{1}{\mathrm{Vol}(\mathcal{H}_i)} \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i)\, \mathrm{d}\mathbf{w}_i$
- (MacKay, 2003)
- statisztikus fizikában számolnak ilyeneket

Bayesi modell-összehasonlítás – II.

- $\frac{1}{\mathrm{Vol}(\mathcal{H}_i)} \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i) \, \mathrm{d}\mathbf{w}_i$
- (MacKay, 2003)
- statisztikus fizikában számolnak ilyeneket
- Lagrange integrál közelítő módszer

Bayesi modell-összehasonlítás – II.

- $\frac{1}{\mathrm{Vol}(\mathcal{H}_i)} \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i)\, \mathrm{d}\mathbf{w}_i$
- (MacKay, 2003)
- statisztikus fizikában számolnak ilyeneket
- Lagrange integrál közelítő módszer
- az alábbiak mind közrejátszanak a döntésben:

Bayesi modell-összehasonlítás – II.

- $\frac{1}{\mathrm{Vol}(\mathcal{H}_i)} \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i)\, \mathrm{d}\mathbf{w}_i$
- (MacKay, 2003)
- statisztikus fizikában számolnak ilyeneket
- Lagrange integrál közelítő módszer
- az alábbiak mind közrejátszanak a döntésben:
  - az adat mérete: n

Bayesi modell-összehasonlítás – II.

- $\frac{1}{\mathrm{Vol}(\mathcal{H}_i)} \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i)\, \mathrm{d}\mathbf{w}_i$
- (MacKay, 2003)
- statisztikus fizikában számolnak ilyeneket
- Lagrange integrál közelítő módszer
- az alábbiak mind közrejátszanak a döntésben:
  - az adat mérete: n
  - a modell mérete: d_i ($\mathrm{Vol}(\mathcal{H}_i)$)

Bayesi modell-összehasonlítás – II.

- $\frac{1}{\mathrm{Vol}(\mathcal{H}_i)} \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i)\,\mathrm{d}\mathbf{w}_i$
- (MacKay, 2003)
- statisztikus fizikában számolnak ilyeneket
- Lagrange integrál közelítő módszer
- az alábbiak mind közrejátszanak a döntésben:
  - az adat mérete: n
  - a modell mérete: d_i ($\mathrm{Vol}(\mathcal{H}_i)$)
  - az illesztés (relatív entrópia, KL divergencia)

Bayesi modell-összehasonlítás – II.

- $\frac{1}{\mathrm{Vol}(\mathcal{H}_i)} \int_{\mathcal{H}_i} \mathbb{P}(D \mid \mathbf{w}_i, \mathcal{H}_i)\, \mathrm{d}\mathbf{w}_i$
- (MacKay, 2003)
- statisztikus fizikában számolnak ilyeneket
- Lagrange integrál közelítő módszer
- az alábbiak mind közrejátszanak a döntésben:
    - az adat mérete: n
    - a modell mérete: d_i ($\mathrm{Vol}(\mathcal{H}_i)$)
    - az illesztés (relatív entrópia, KL divergencia)
- szükség van az illesztés (célfüggvény) Hesse mátrixára

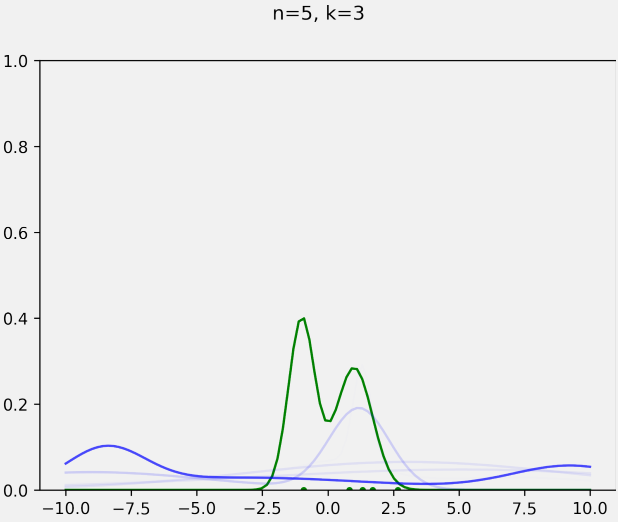
Bayesi modell-összehasonlítás – demo

Bayesi modell-összehasonlítás – demo

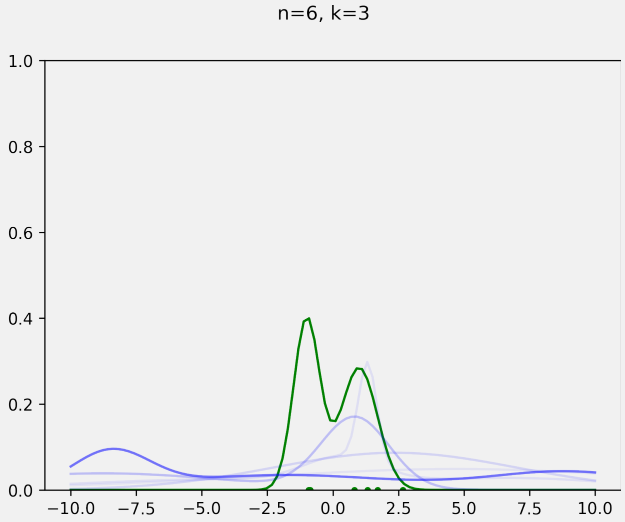
Bayesi modell-összehasonlítás – demo

Bayesi modell-összehasonlítás – demo

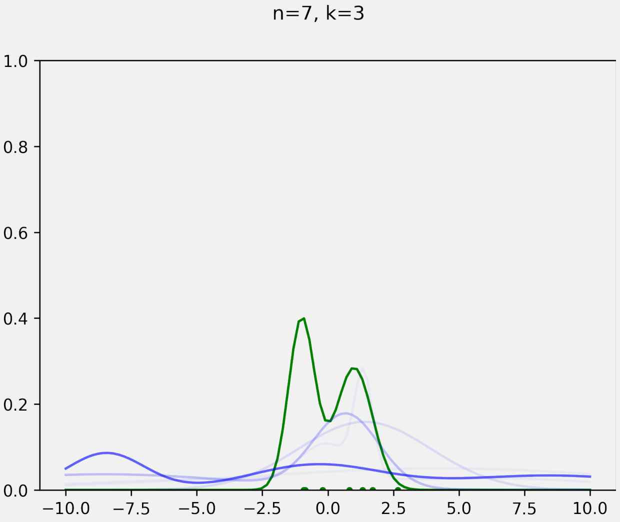
Bayesi modell-összehasonlítás – demo



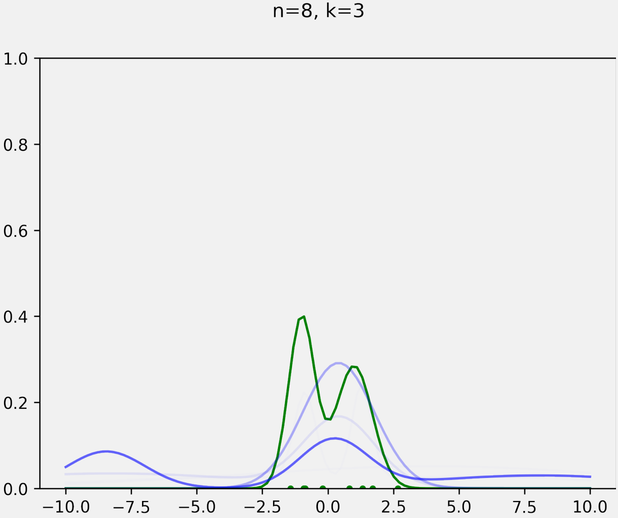
Bayesi modell-összehasonlítás – demo



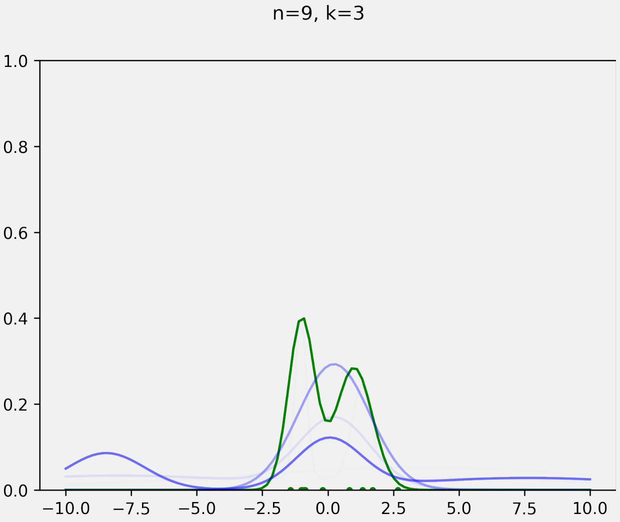
Bayesi modell-összehasonlítás – demo



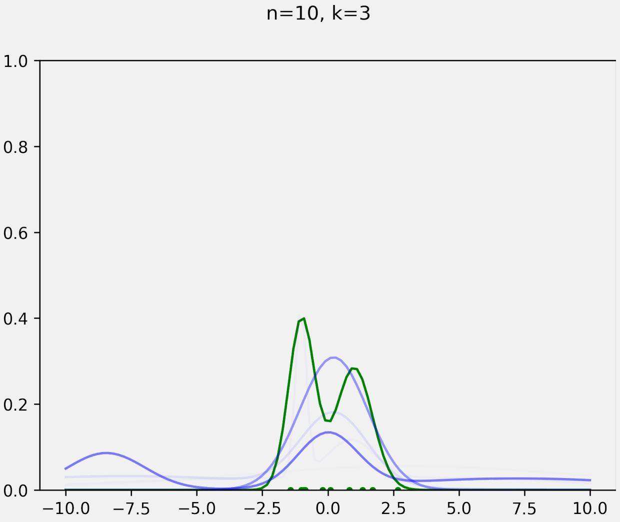
Bayesi modell-összehasonlítás – demo



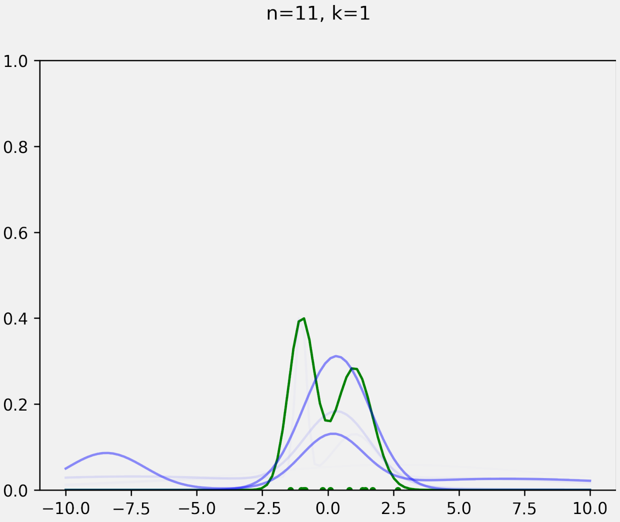
Bayesi modell-összehasonlítás – demo



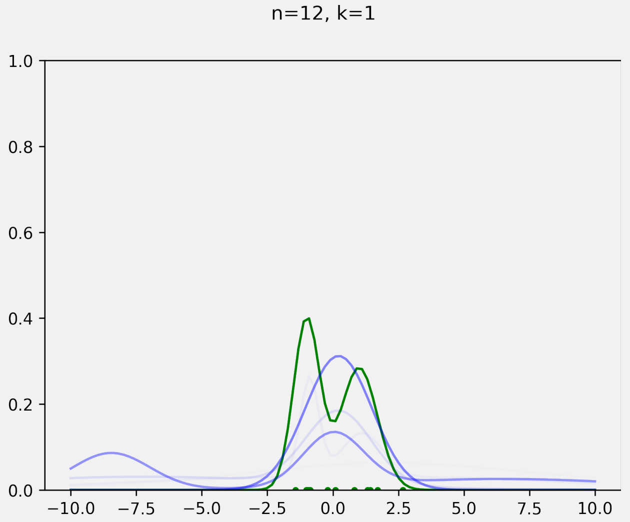
Bayesi modell-összehasonlítás – demo



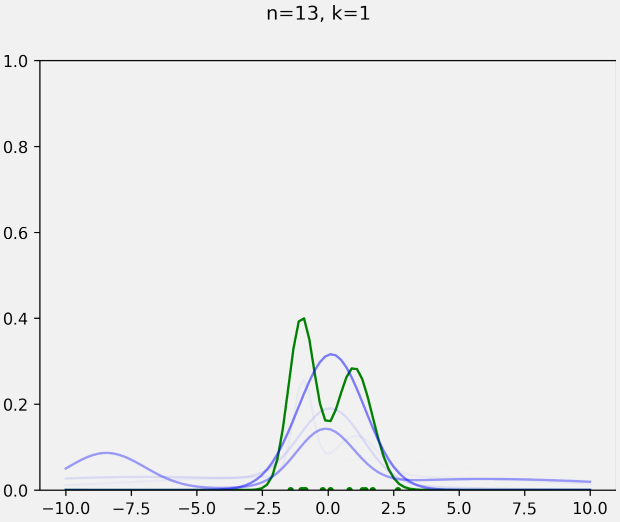
Bayesi modell-összehasonlítás – demo



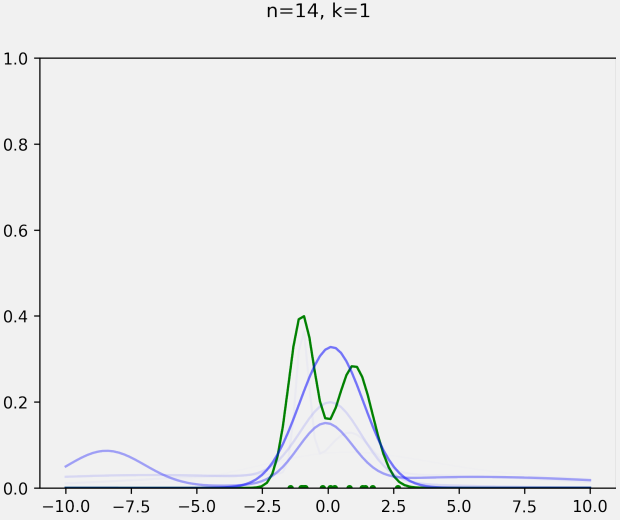
Bayesi modell-összehasonlítás – demo



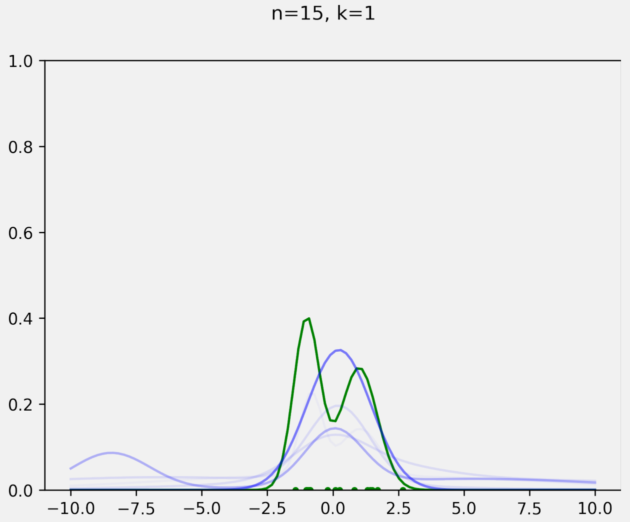
Bayesi modell-összehasonlítás – demo



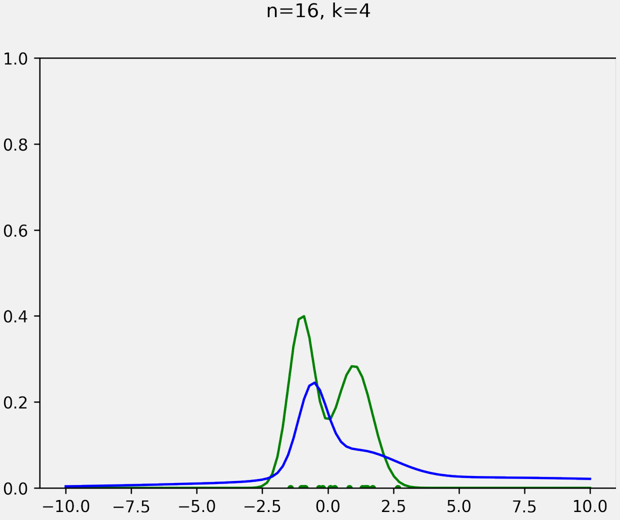
Bayesi modell-összehasonlítás – demo



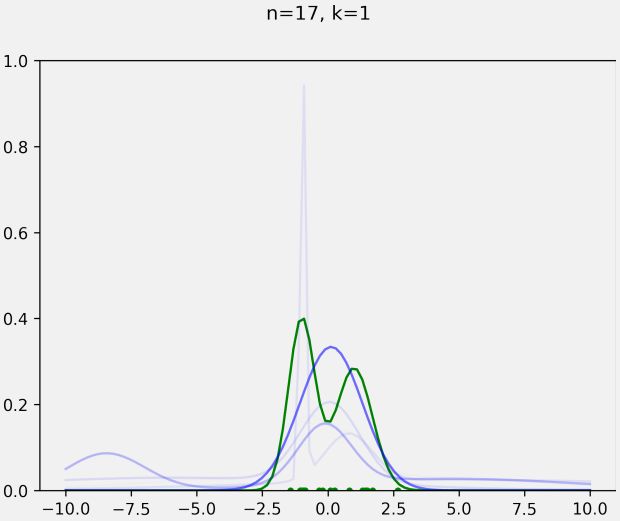
Bayesi modell-összehasonlítás – demo



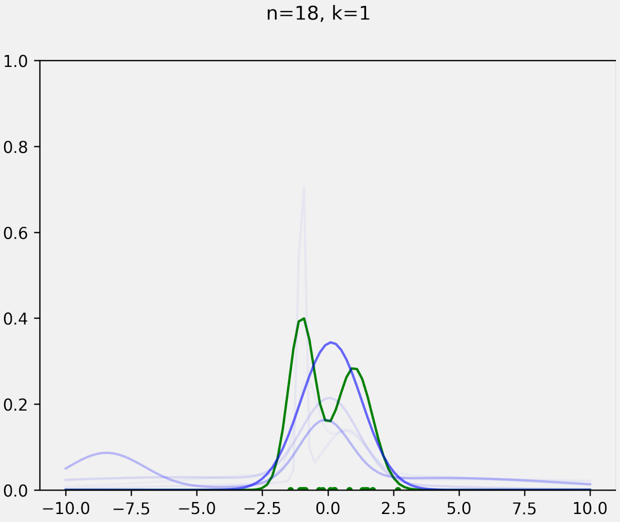
Bayesi modell-összehasonlítás – demo



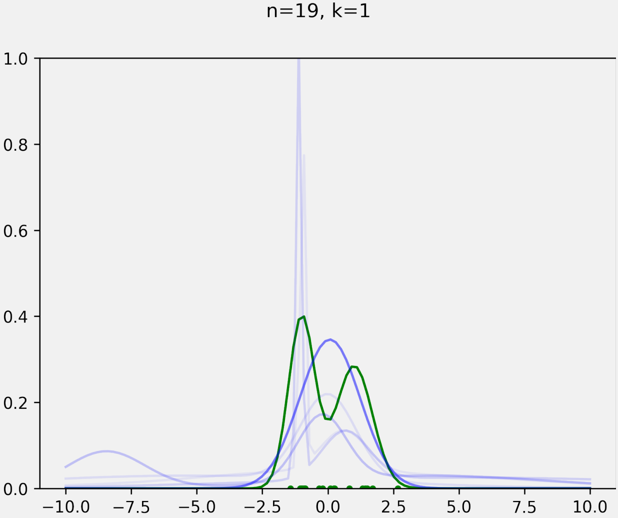
Bayesi modell-összehasonlítás – demo



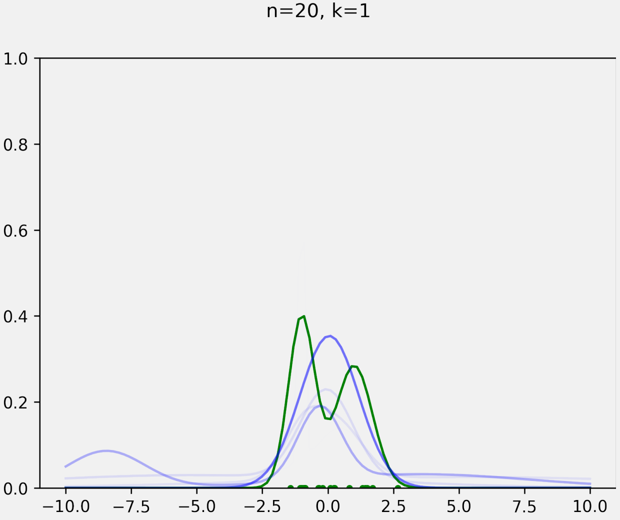
Bayesi modell-összehasonlítás – demo



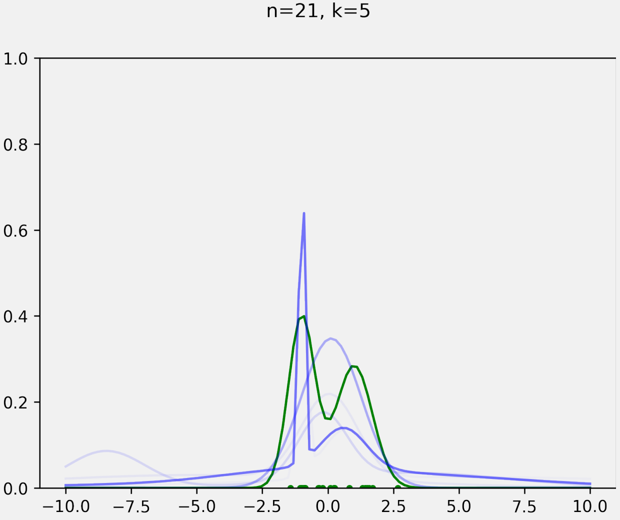
Bayesi modell-összehasonlítás – demo



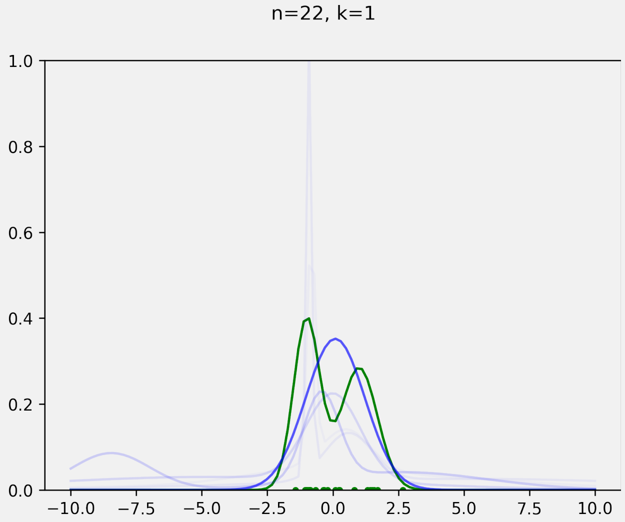
Bayesi modell-összehasonlítás – demo



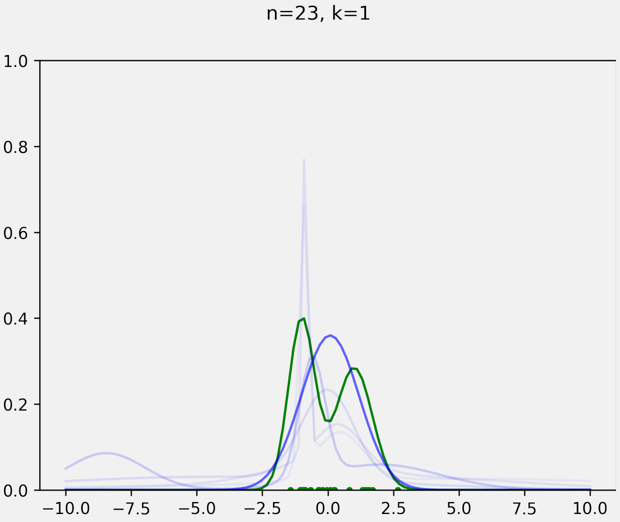
Bayesi modell-összehasonlítás – demo



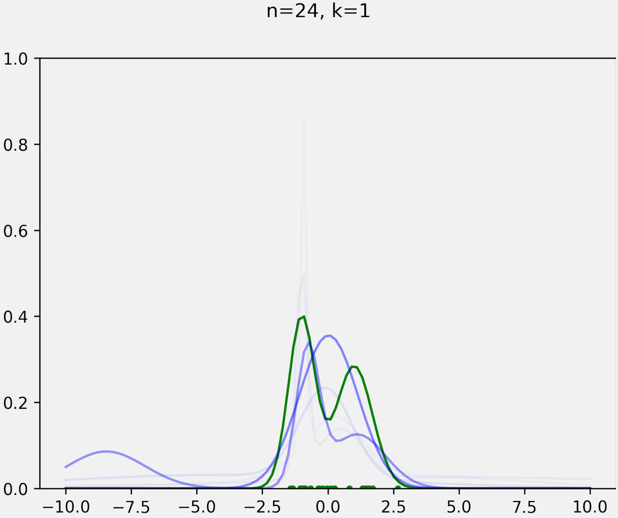
Bayesi modell-összehasonlítás – demo



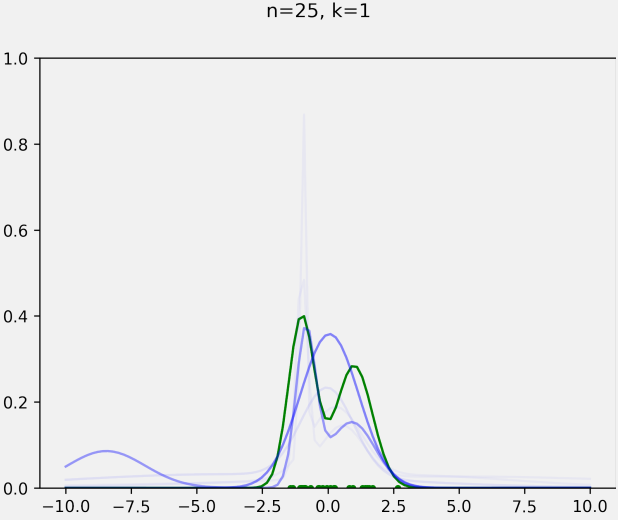
Bayesi modell-összehasonlítás – demo



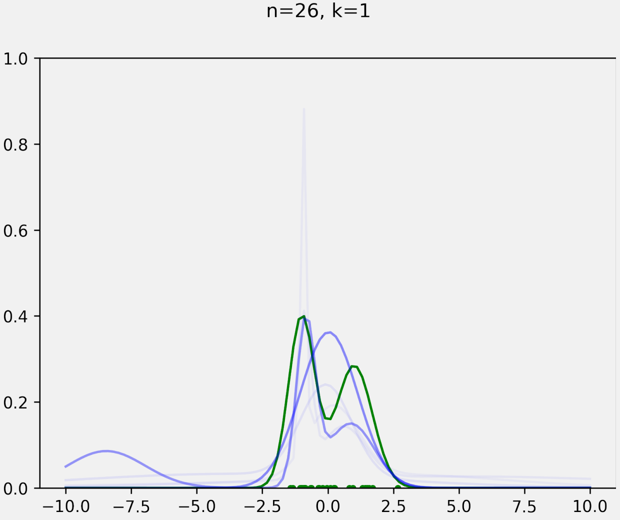
Bayesi modell-összehasonlítás – demo



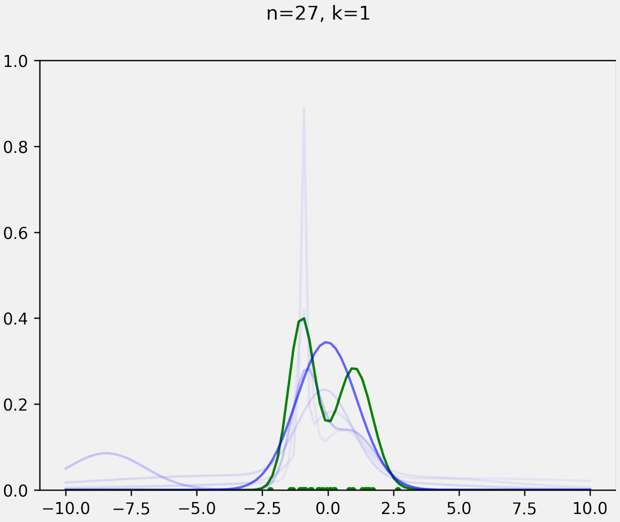
Bayesi modell-összehasonlítás – demo



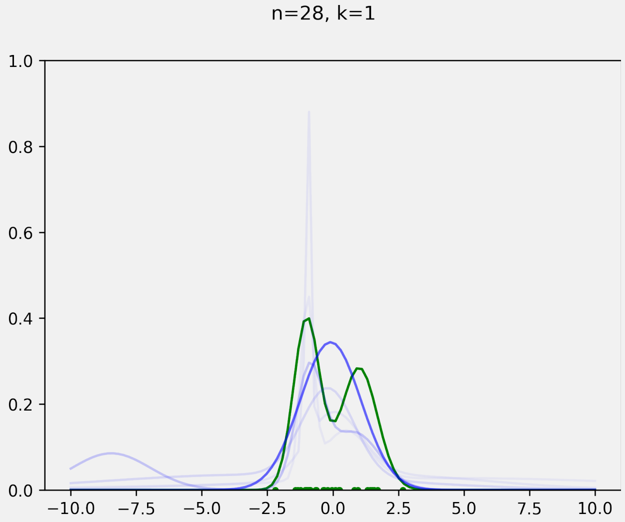
Bayesi modell-összehasonlítás – demo



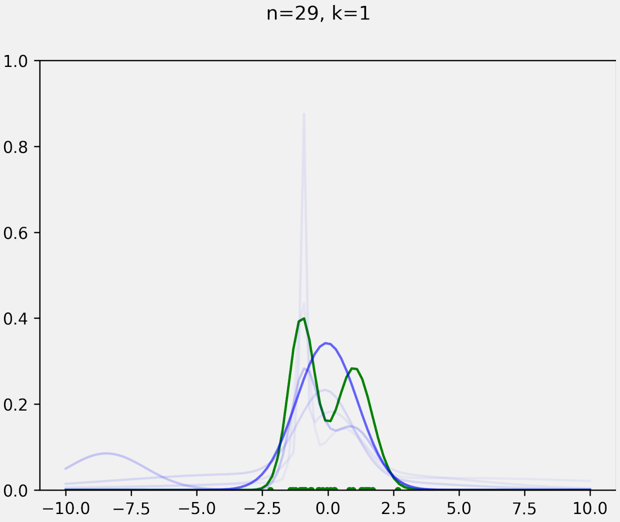
Bayesi modell-összehasonlítás – demo



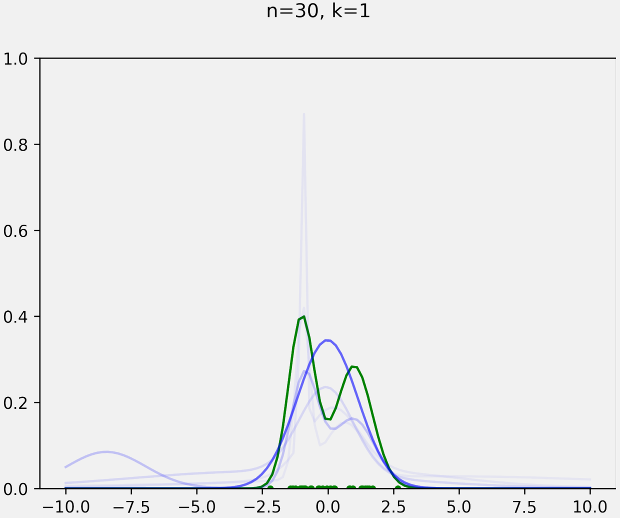
Bayesi modell-összehasonlítás – demo



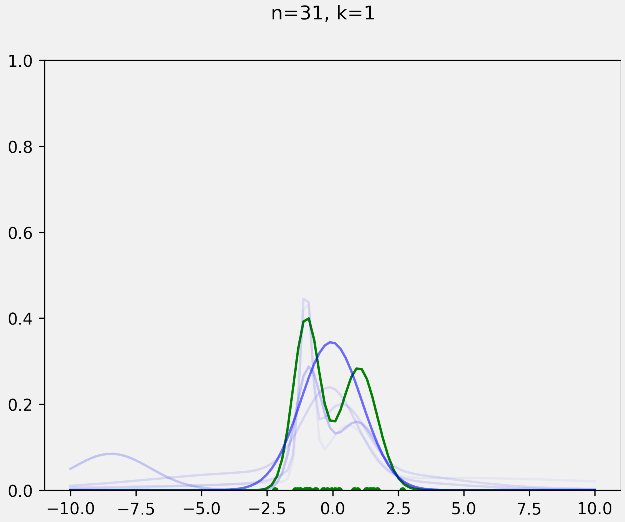
Bayesi modell-összehasonlítás – demo



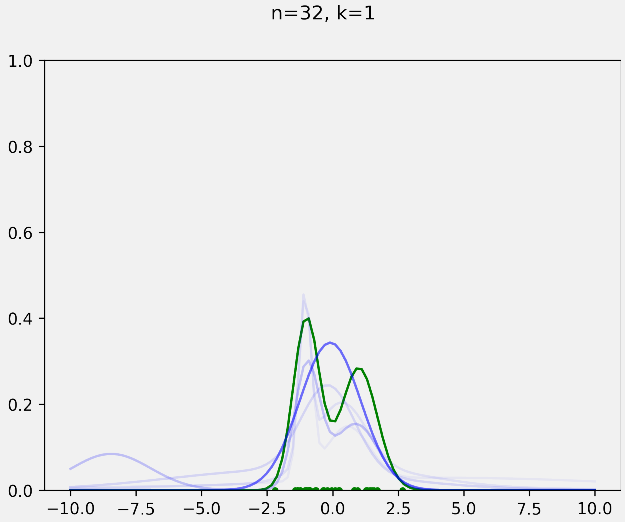
Bayesi modell-összehasonlítás – demo



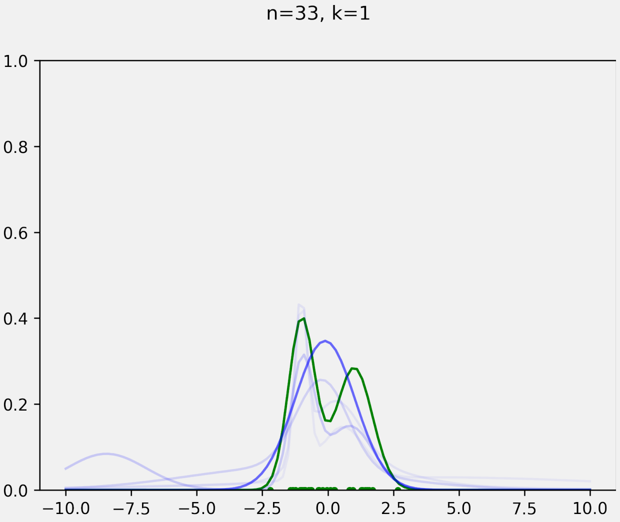
Bayesi modell-összehasonlítás – demo



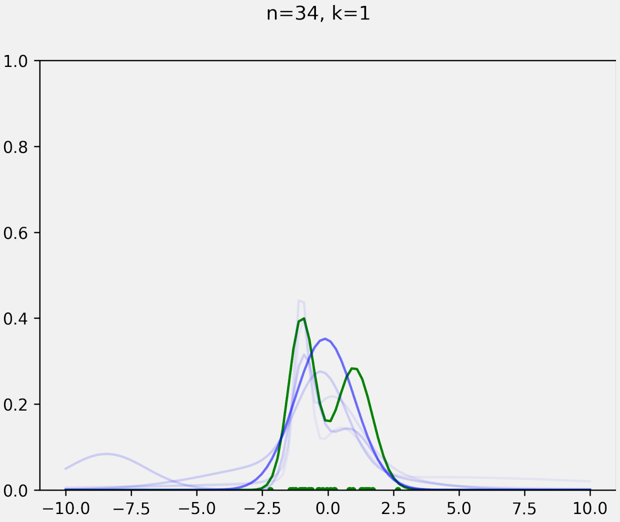
Bayesi modell-összehasonlítás – demo



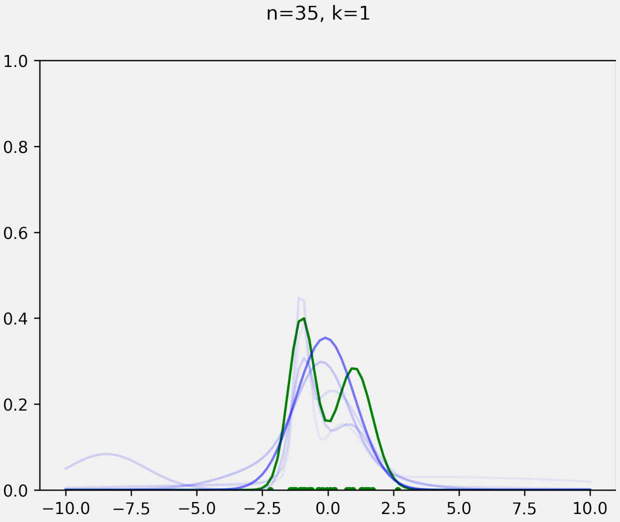
Bayesi modell-összehasonlítás – demo



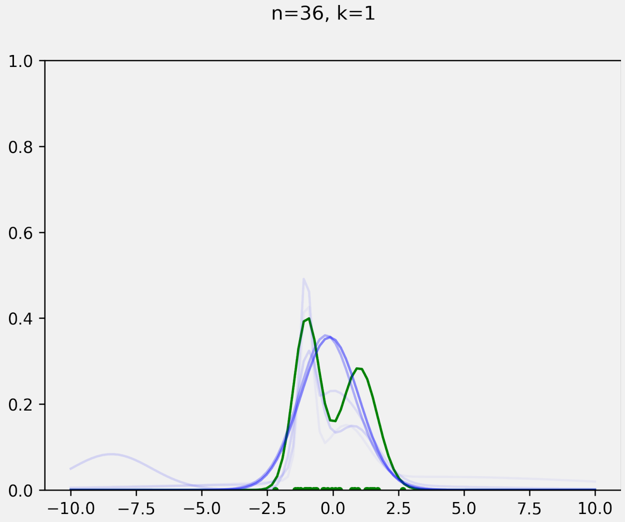
Bayesi modell-összehasonlítás – demo



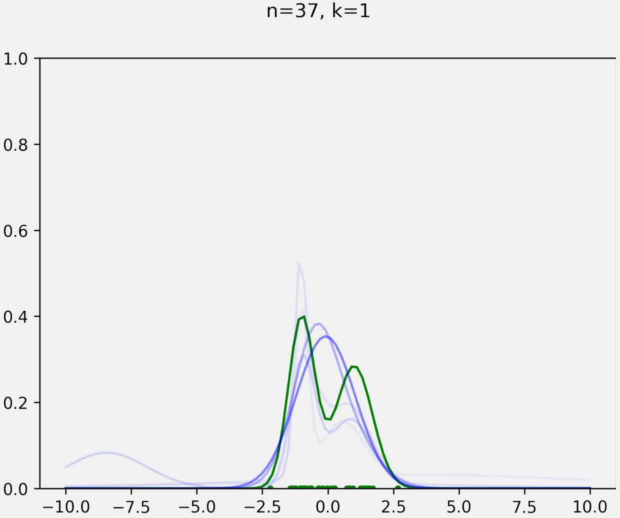
Bayesi modell-összehasonlítás – demo



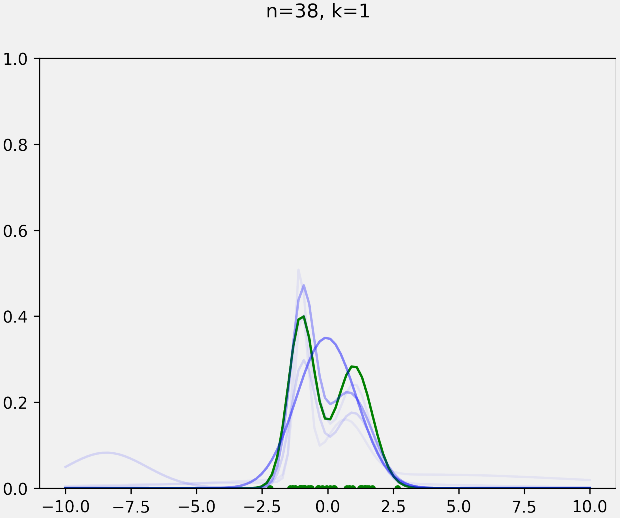
Bayesi modell-összehasonlítás – demo



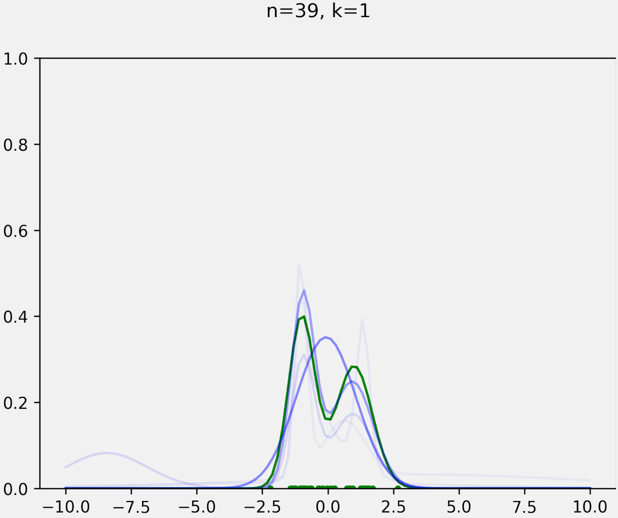
Bayesi modell-összehasonlítás – demo



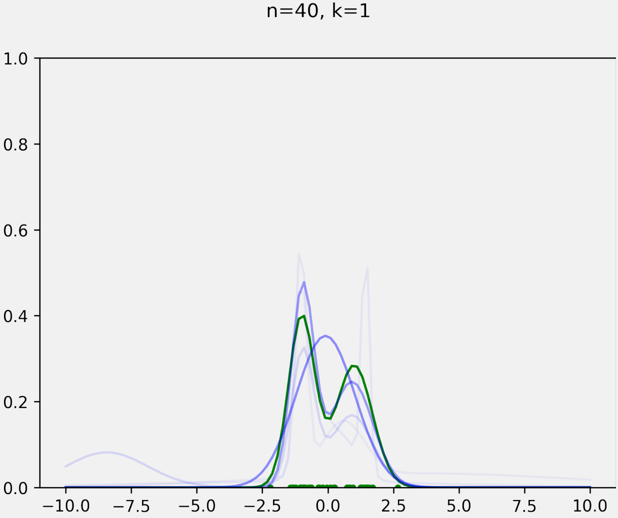
Bayesi modell-összehasonlítás – demo



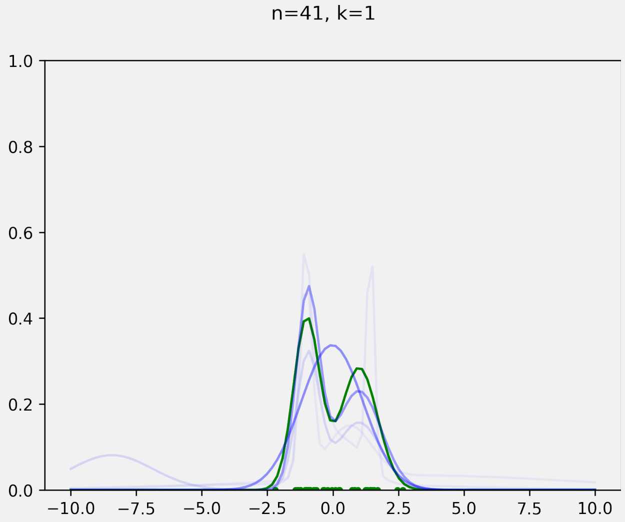
Bayesi modell-összehasonlítás – demo



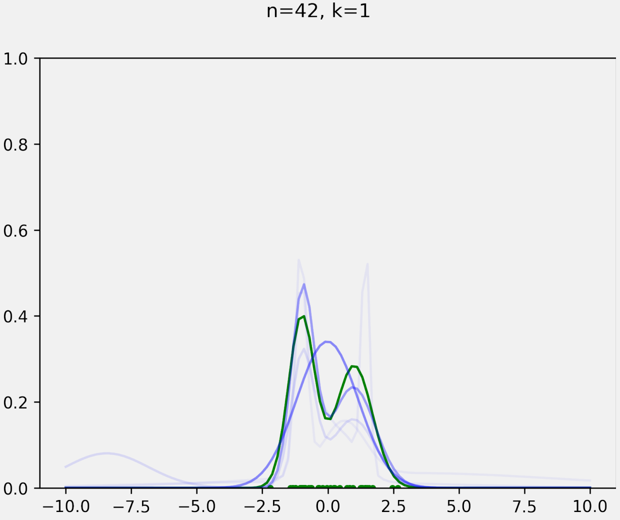
Bayesi modell-összehasonlítás – demo



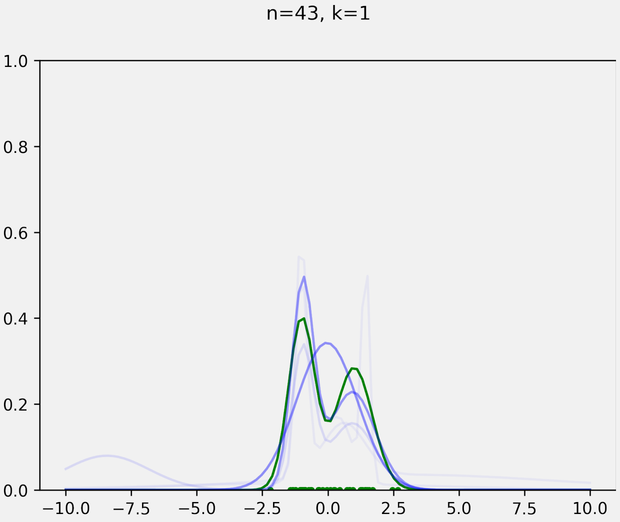
Bayesi modell-összehasonlítás – demo



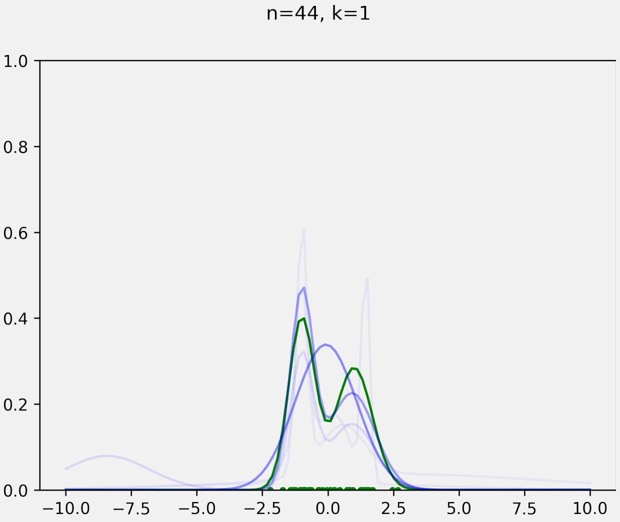
Bayesi modell-összehasonlítás – demo



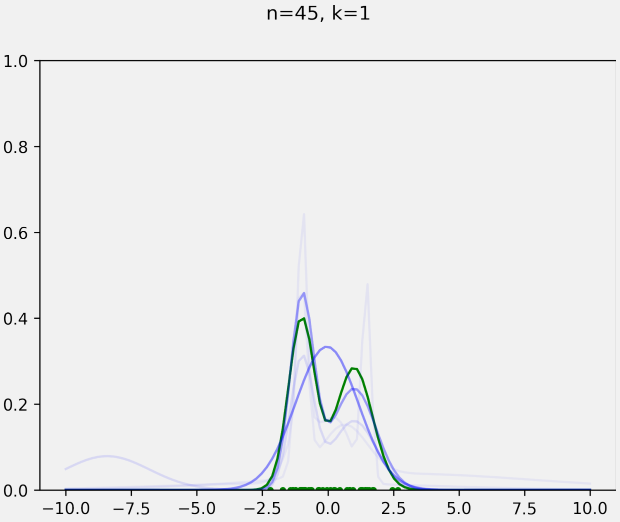
Bayesi modell-összehasonlítás – demo



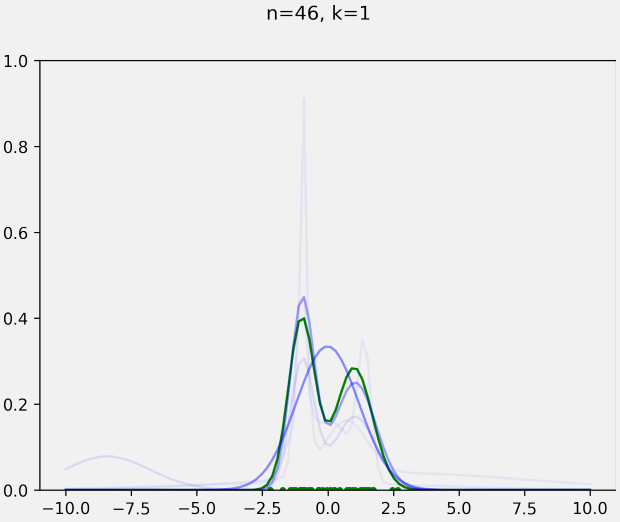
Bayesi modell-összehasonlítás – demo



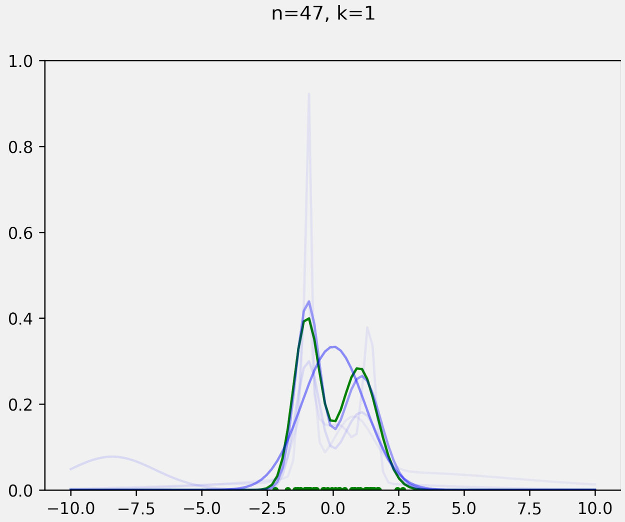
Bayesi modell-összehasonlítás – demo



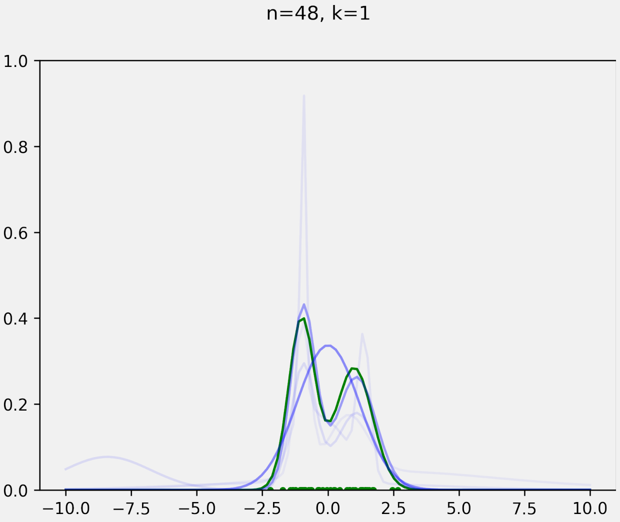
Bayesi modell-összehasonlítás – demo



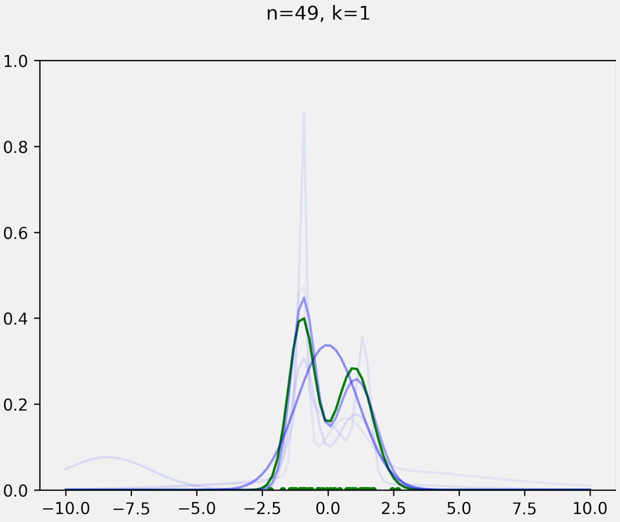
Bayesi modell-összehasonlítás – demo



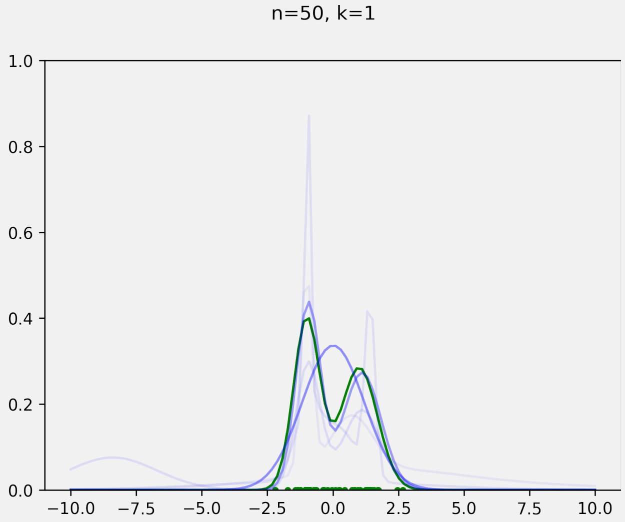
Bayesi modell-összehasonlítás – demo



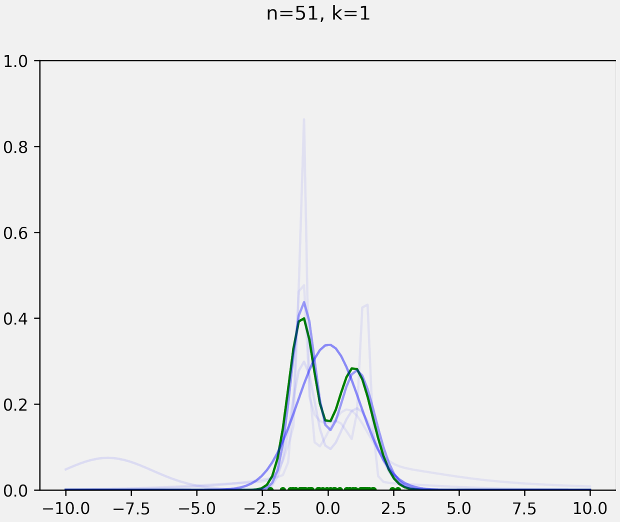
Bayesi modell-összehasonlítás – demo



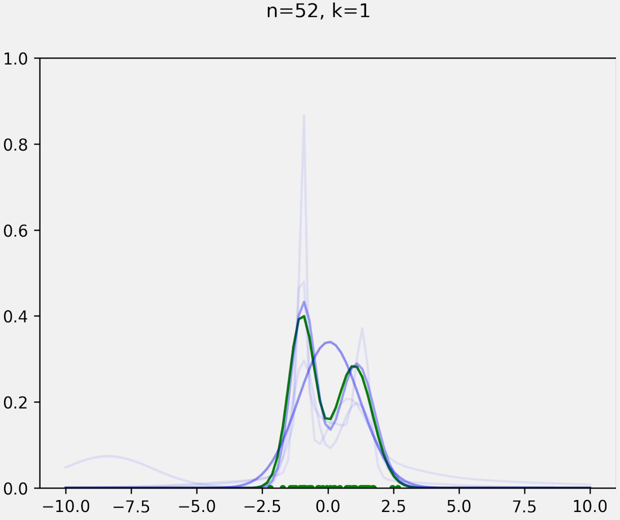
Bayesi modell-összehasonlítás – demo



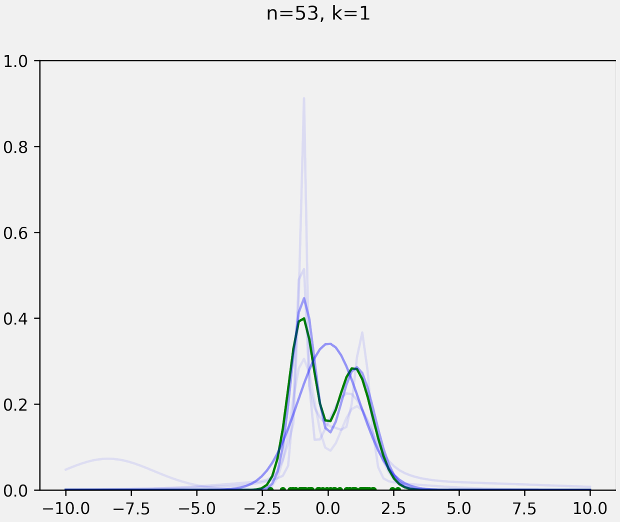
Bayesi modell-összehasonlítás – demo



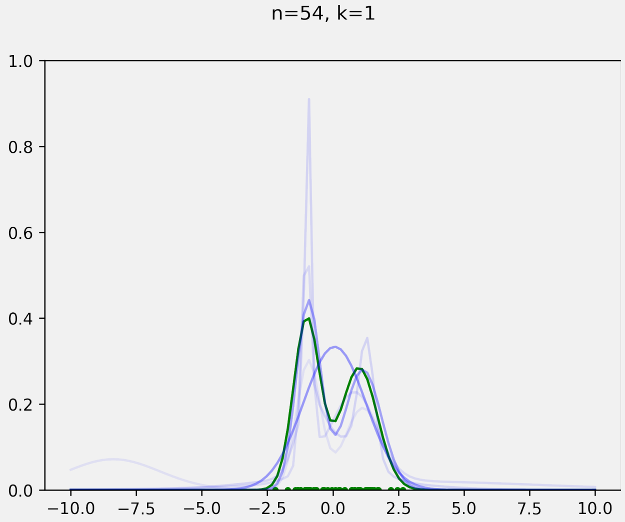
Bayesi modell-összehasonlítás – demo



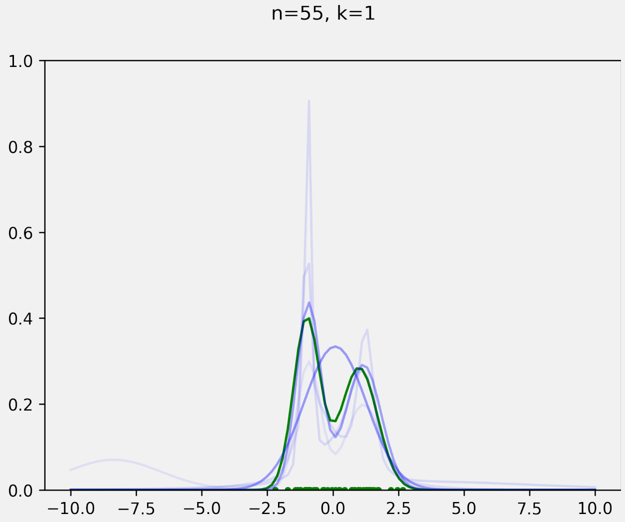
Bayesi modell-összehasonlítás – demo



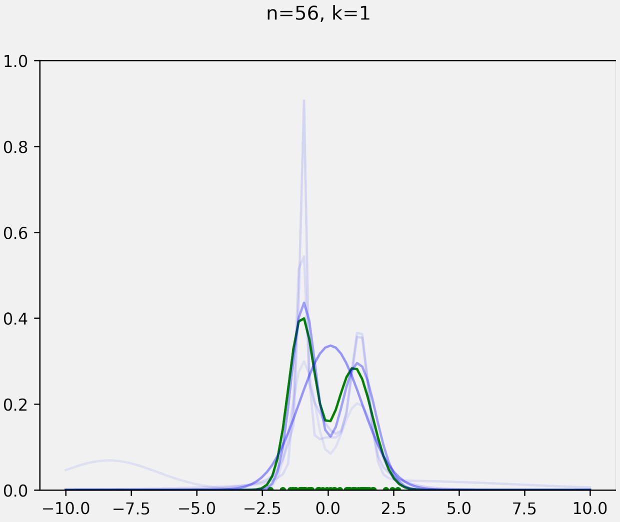
Bayesi modell-összehasonlítás – demo



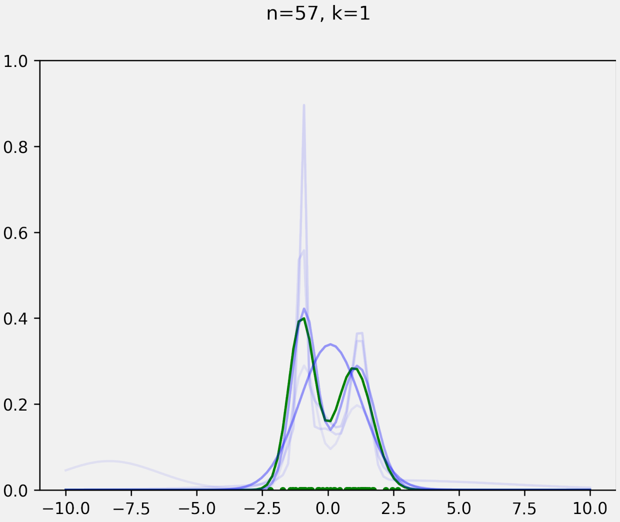
Bayesi modell-összehasonlítás – demo



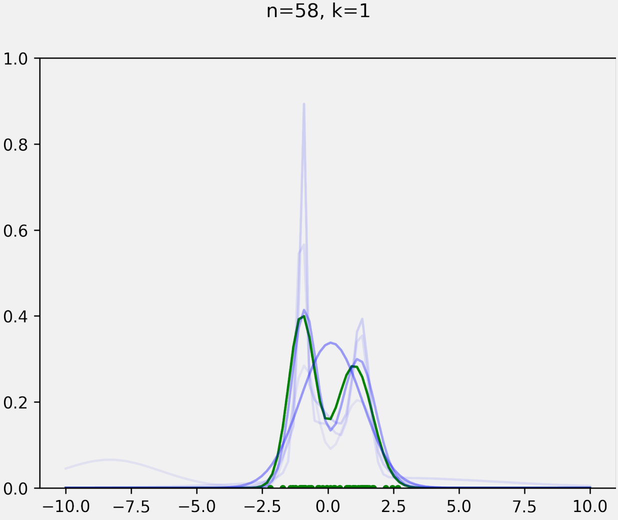
Bayesi modell-összehasonlítás – demo



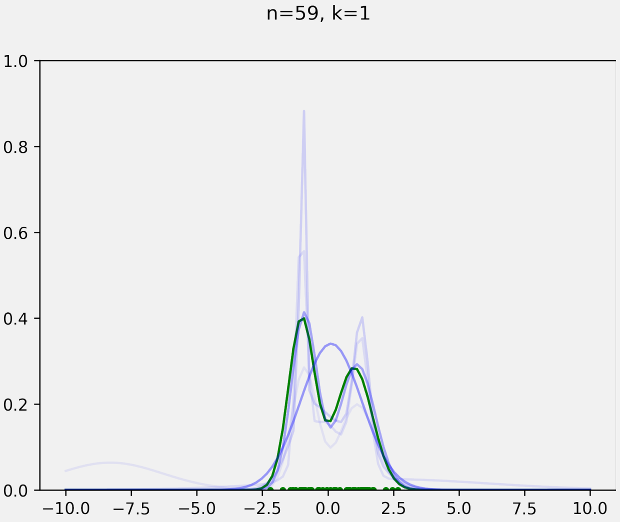
Bayesi modell-összehasonlítás – demo



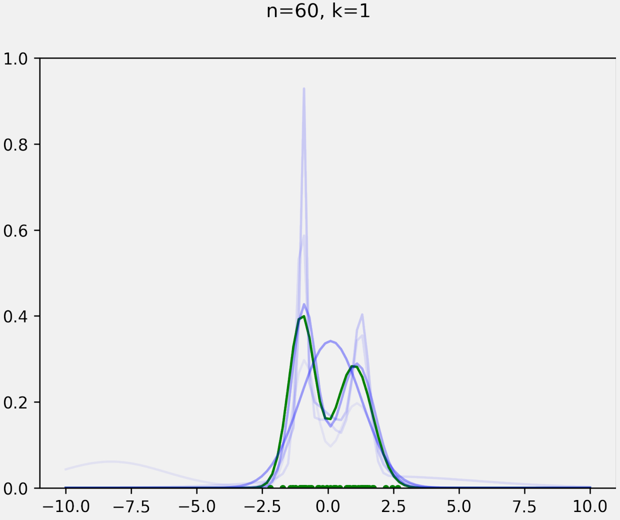
Bayesi modell-összehasonlítás – demo



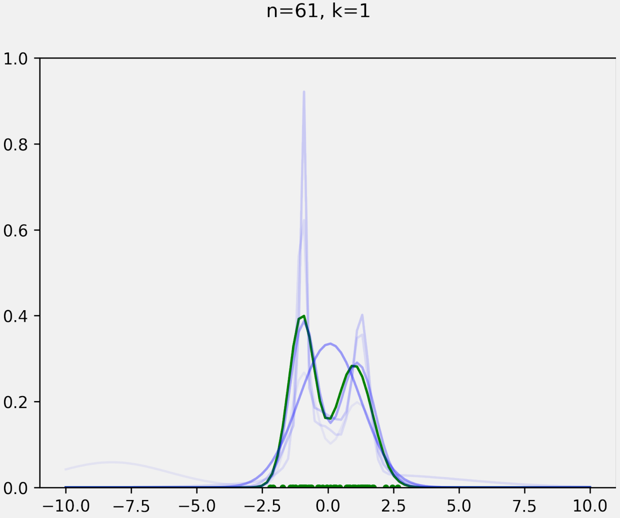
Bayesi modell-összehasonlítás – demo



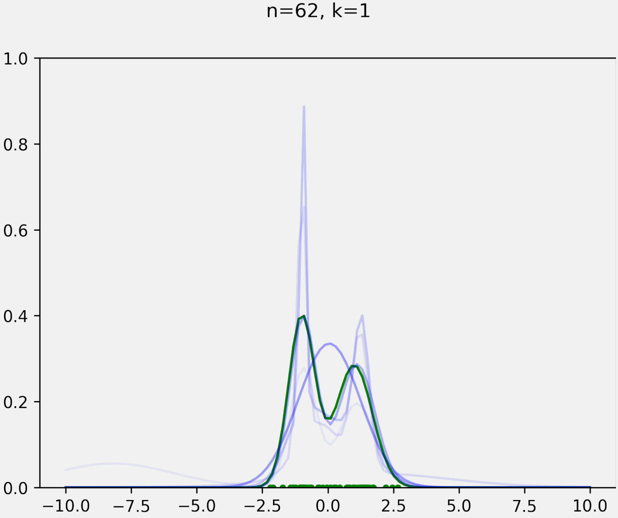
Bayesi modell-összehasonlítás – demo



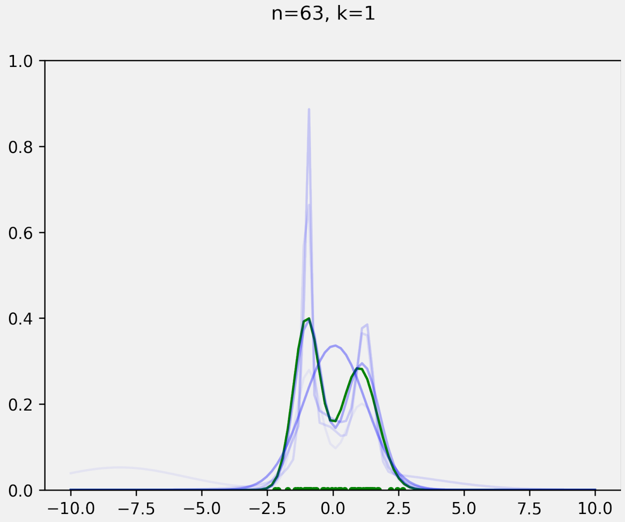
Bayesi modell-összehasonlítás – demo



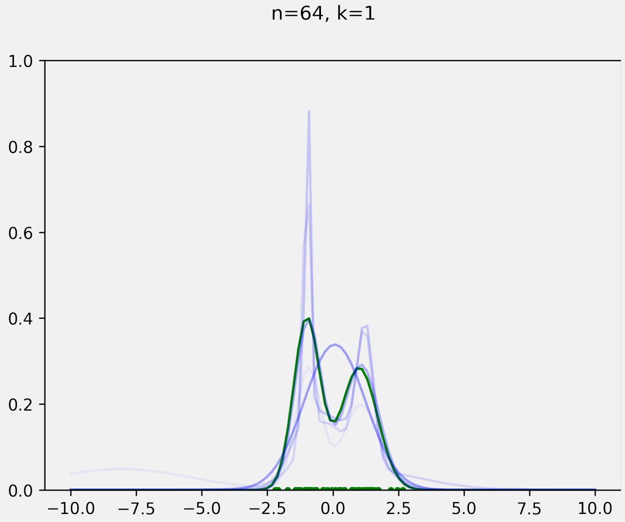
Bayesi modell-összehasonlítás – demo



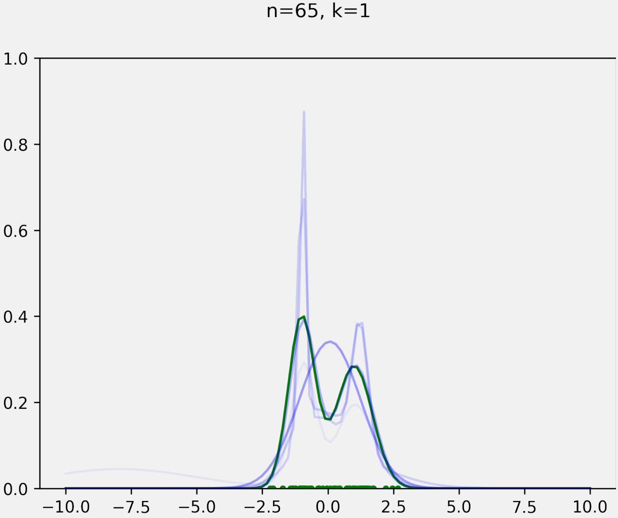
Bayesi modell-összehasonlítás – demo



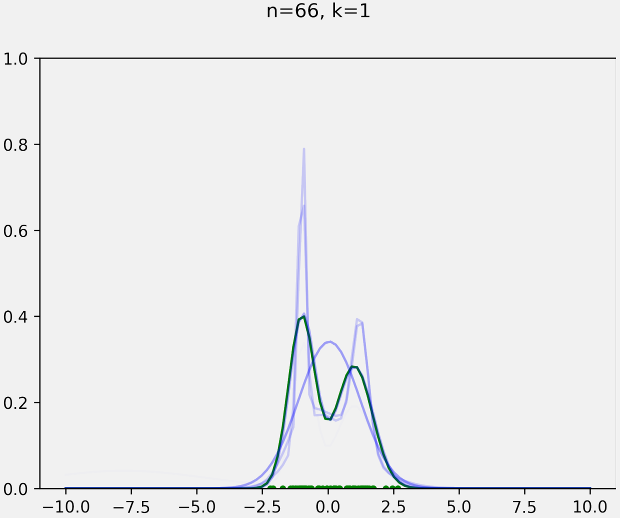
Bayesi modell-összehasonlítás – demo



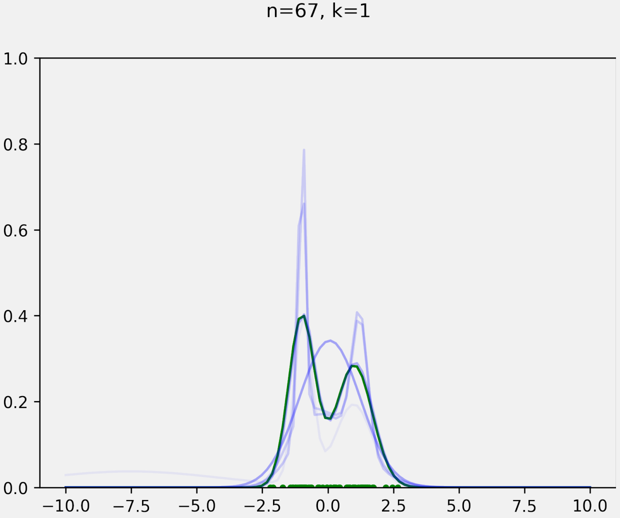
Bayesi modell-összehasonlítás – demo



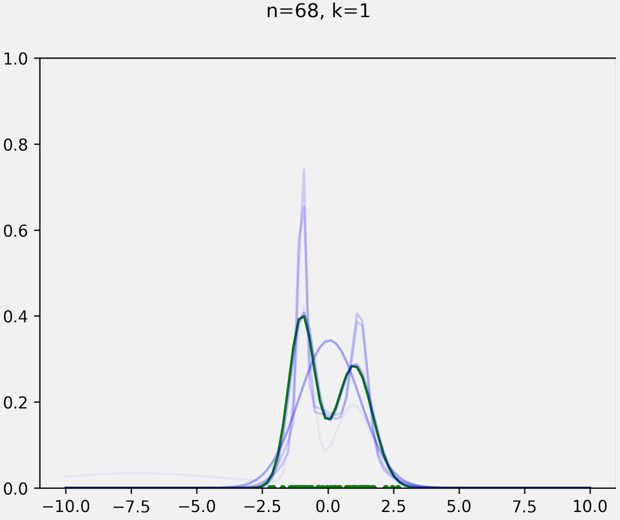
Bayesi modell-összehasonlítás – demo



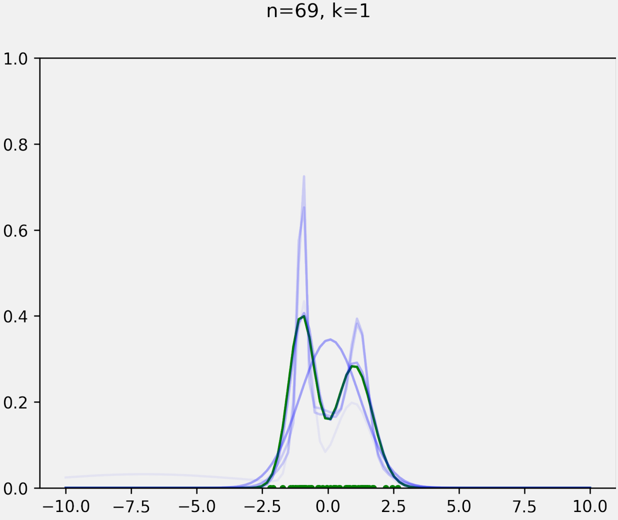
Bayesi modell-összehasonlítás – demo



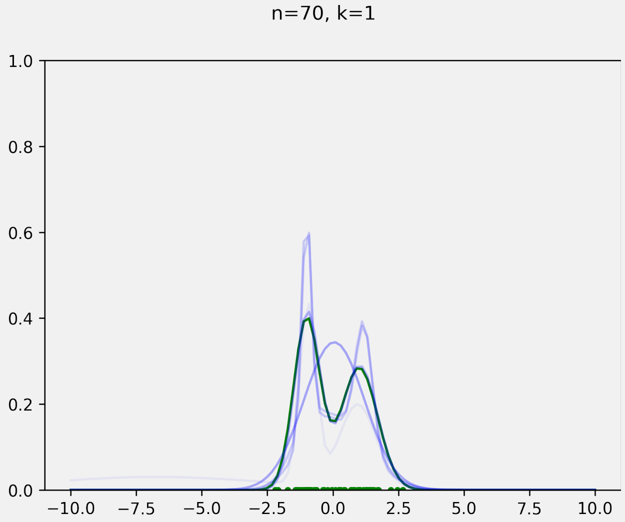
Bayesi modell-összehasonlítás – demo



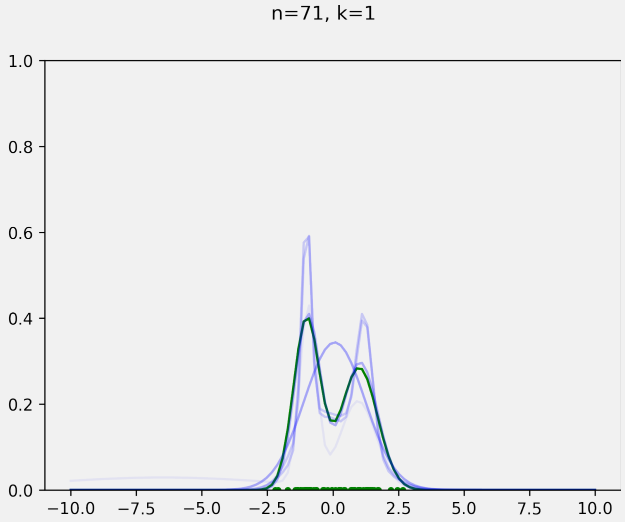
Bayesi modell-összehasonlítás – demo



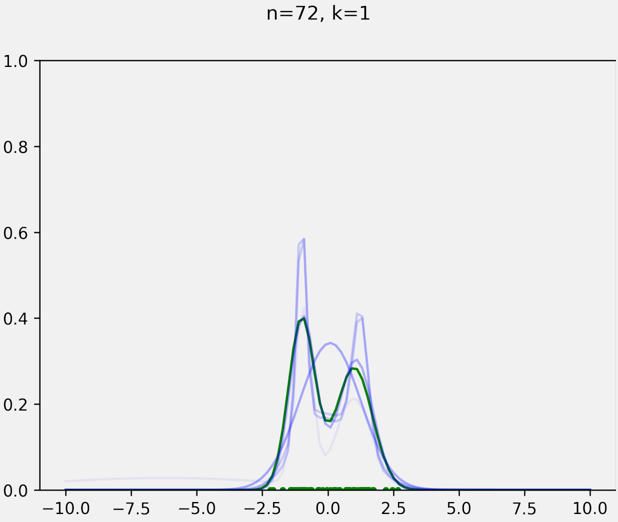
Bayesi modell-összehasonlítás – demo



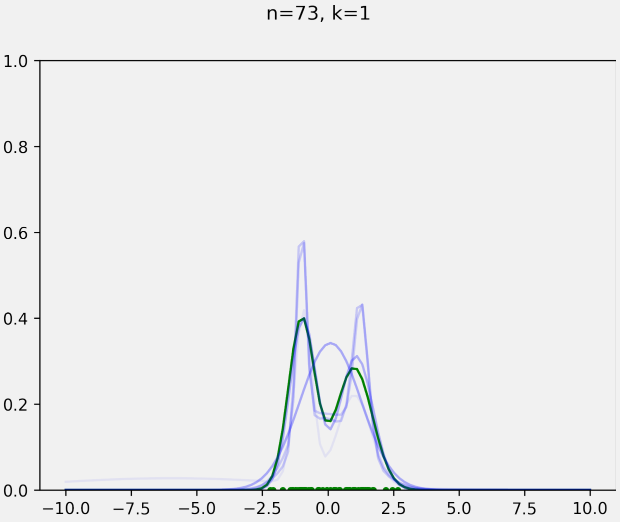
Bayesi modell-összehasonlítás – demo



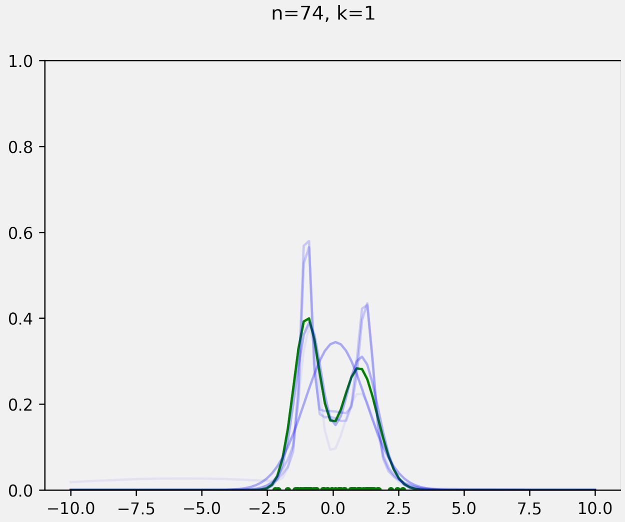
Bayesi modell-összehasonlítás – demo



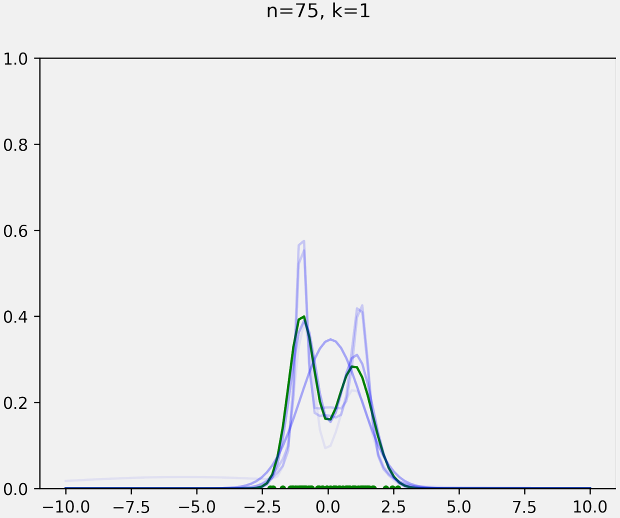
Bayesi modell-összehasonlítás – demo



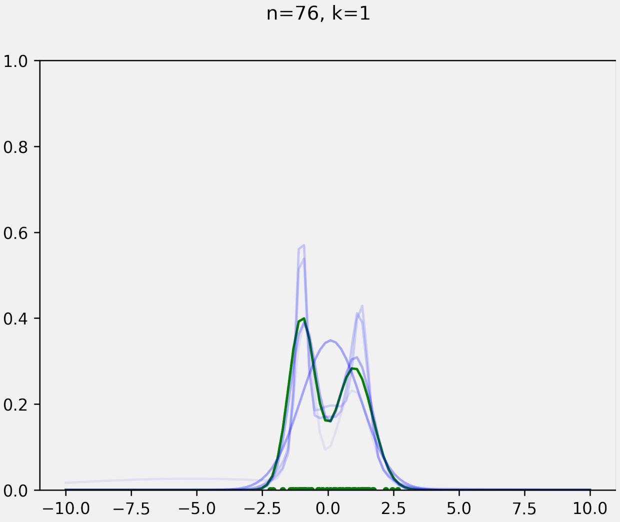
Bayesi modell-összehasonlítás – demo



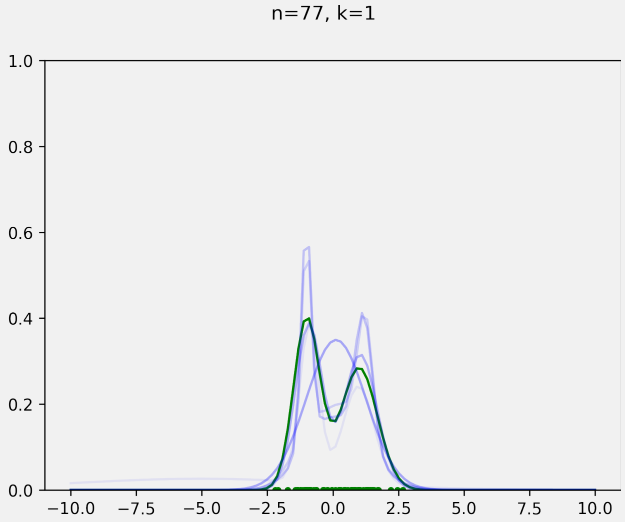
Bayesi modell-összehasonlítás – demo



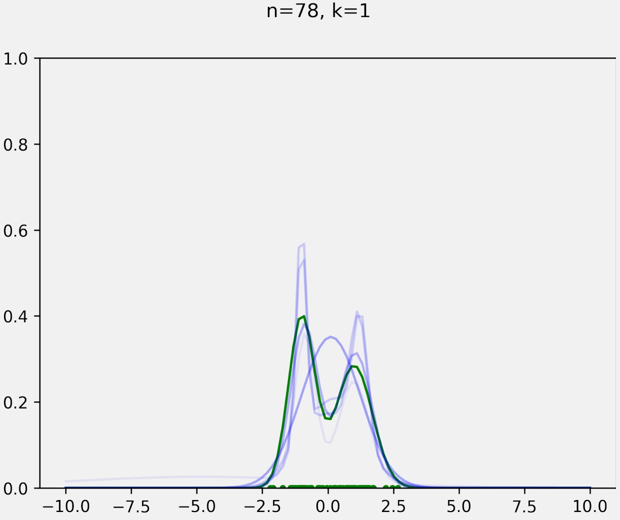
Bayesi modell-összehasonlítás – demo



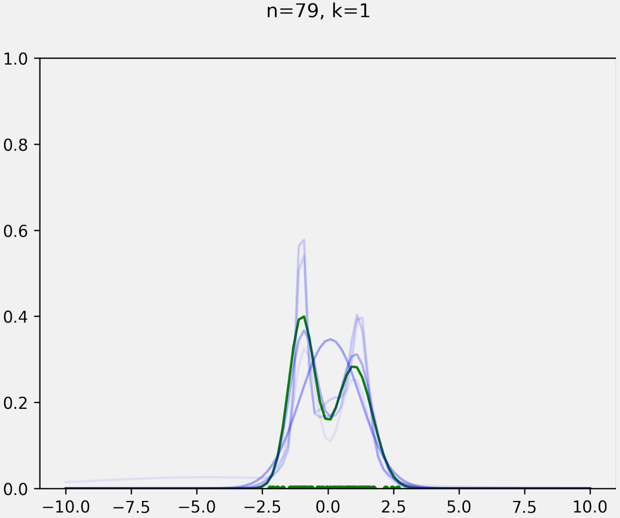
Bayesi modell-összehasonlítás – demo



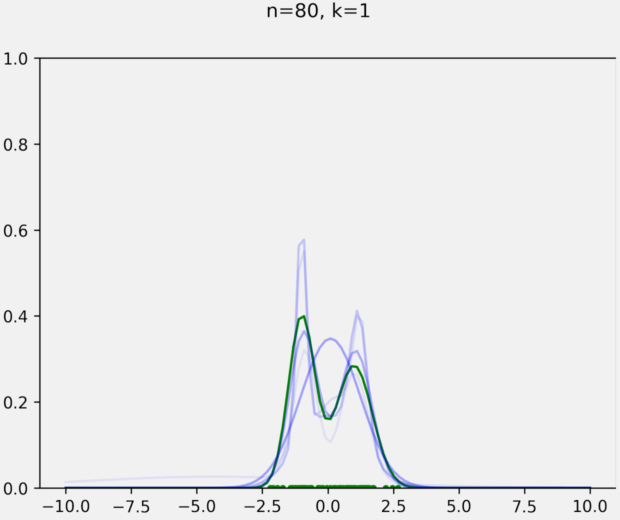
Bayesi modell-összehasonlítás – demo



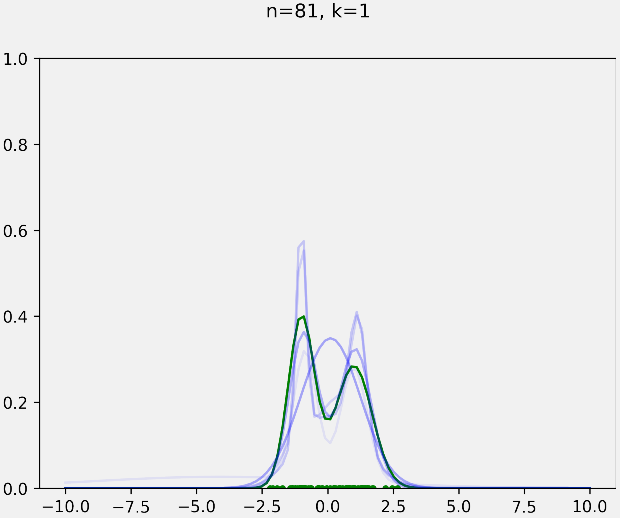
Bayesi modell-összehasonlítás – demo



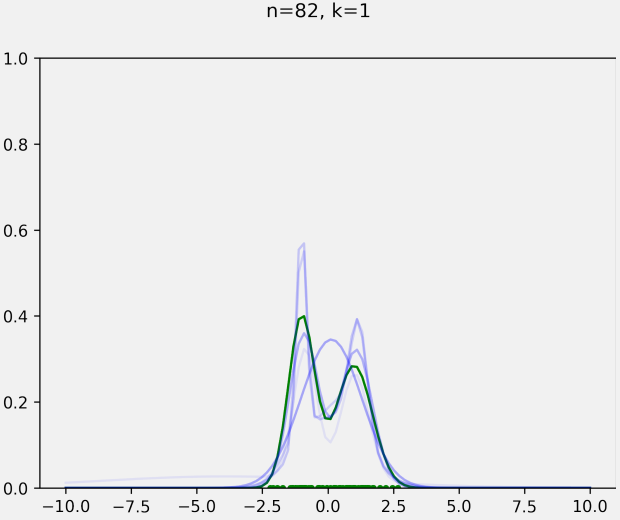
Bayesi modell-összehasonlítás – demo



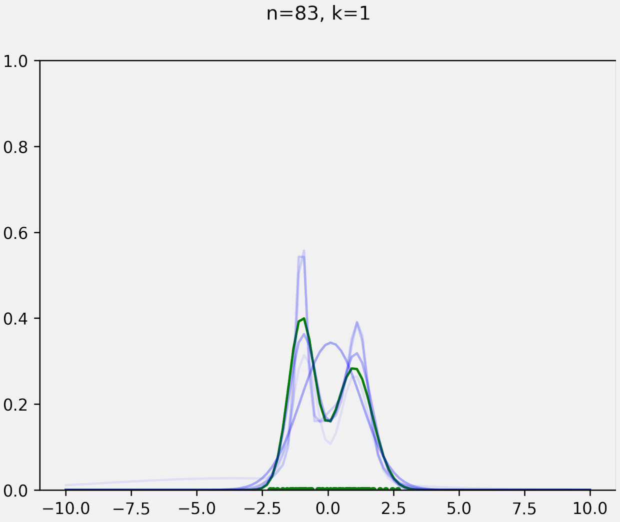
Bayesi modell-összehasonlítás – demo



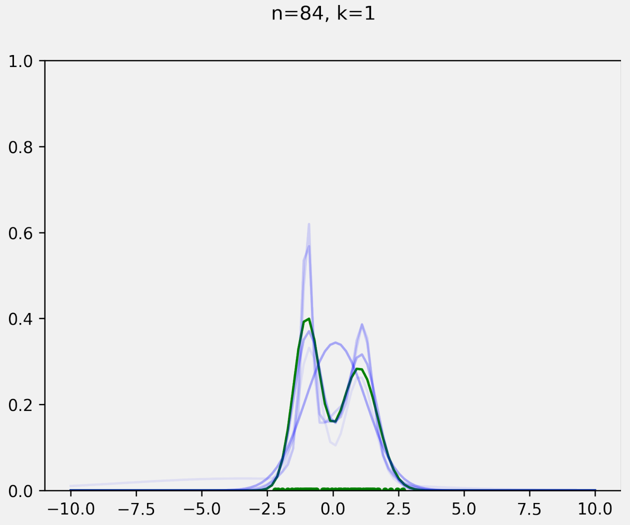
Bayesi modell-összehasonlítás – demo



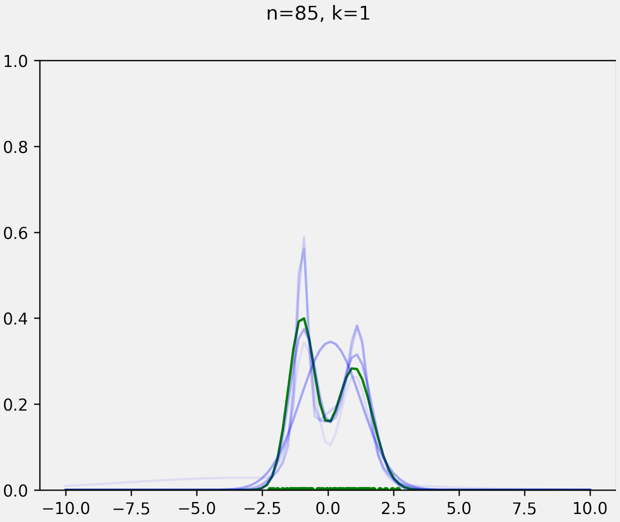
Bayesi modell-összehasonlítás – demo



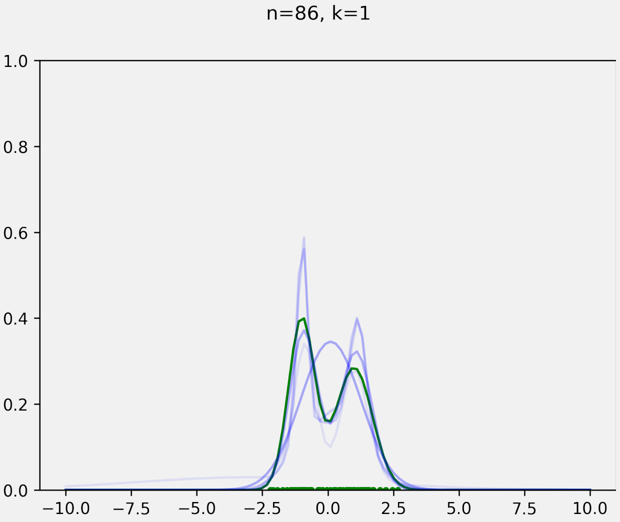
Bayesi modell-összehasonlítás – demo



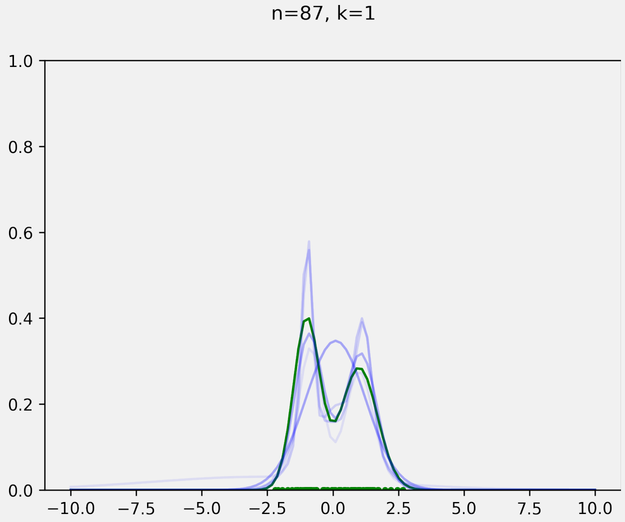
Bayesi modell-összehasonlítás – demo



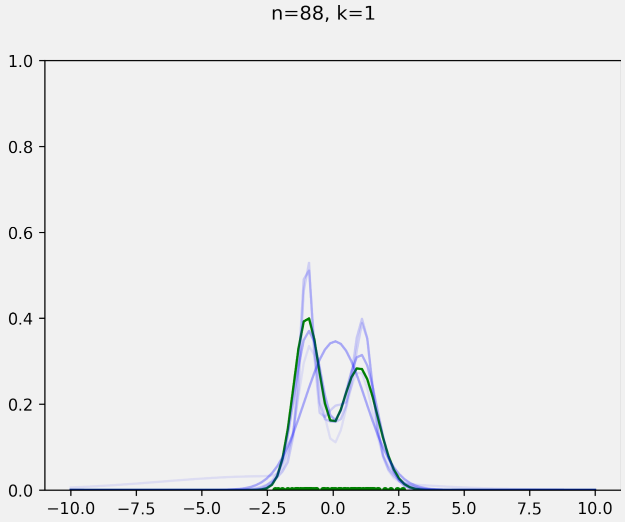
Bayesi modell-összehasonlítás – demo



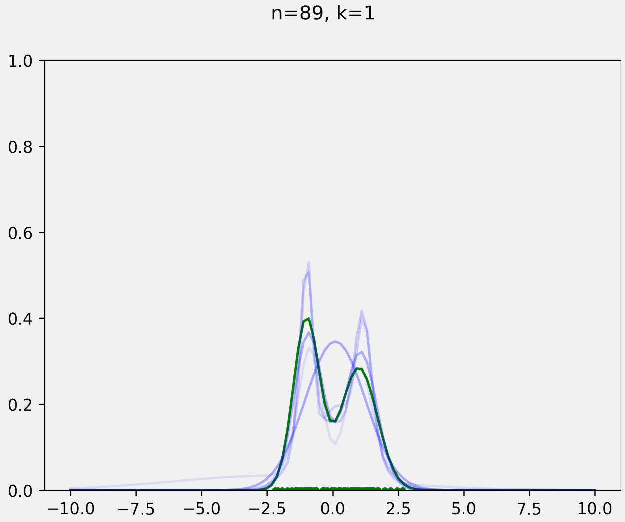
Bayesi modell-összehasonlítás – demo



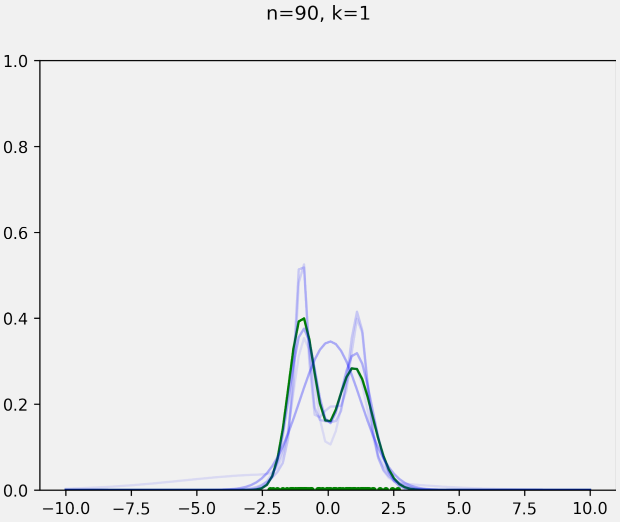
Bayesi modell-összehasonlítás – demo



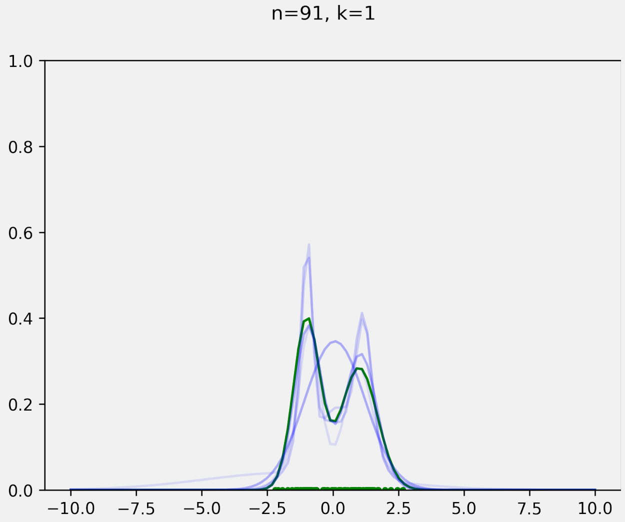
Bayesi modell-összehasonlítás – demo



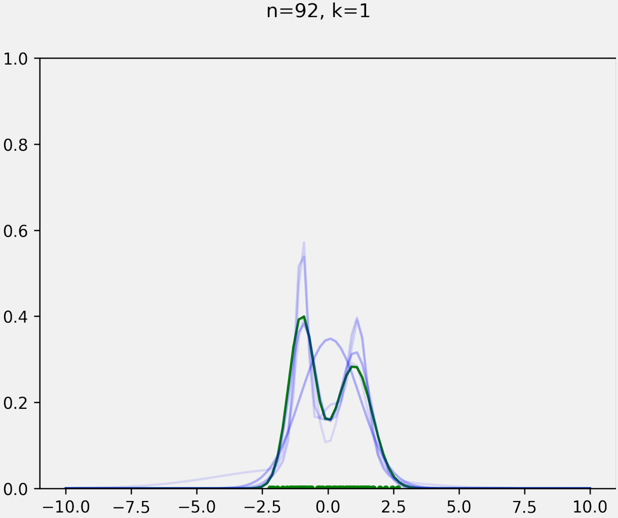
Bayesi modell-összehasonlítás – demo



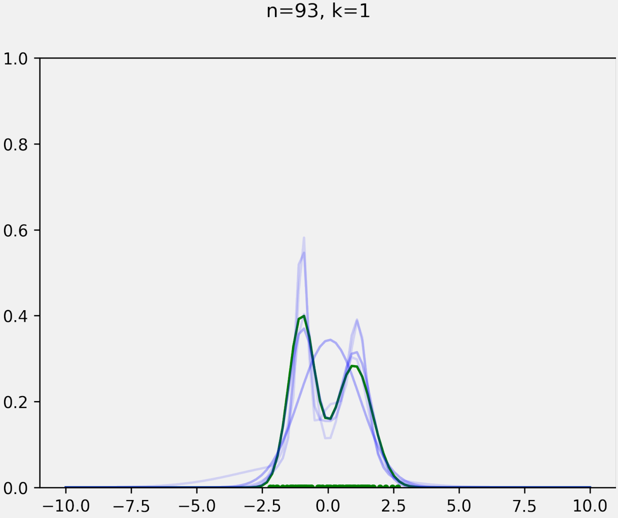
Bayesi modell-összehasonlítás – demo



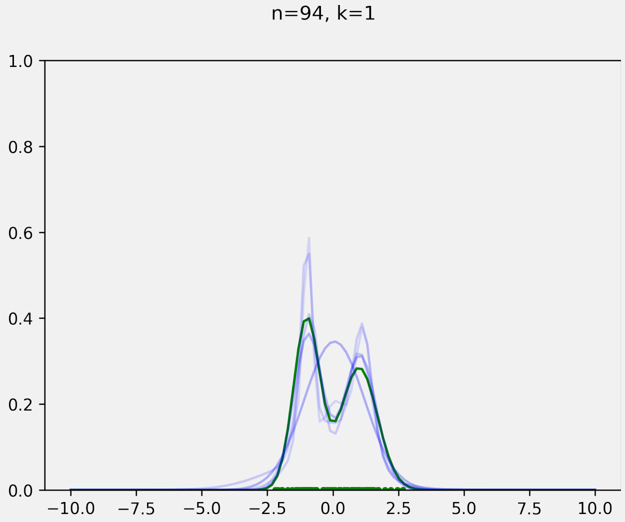
Bayesi modell-összehasonlítás – demo



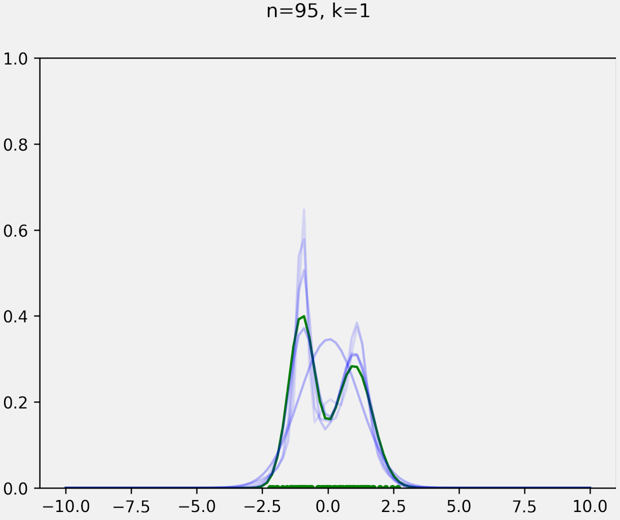
Bayesi modell-összehasonlítás – demo



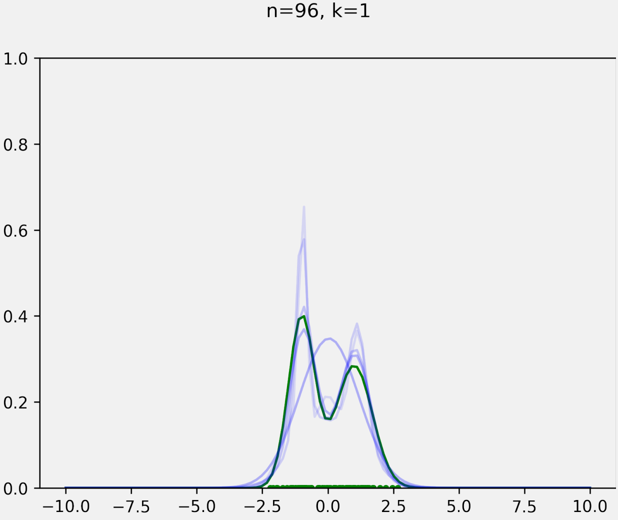
Bayesi modell-összehasonlítás – demo



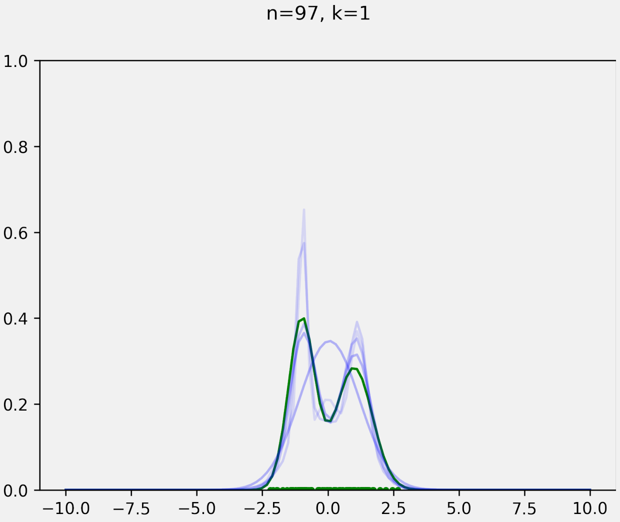
Bayesi modell-összehasonlítás – demo



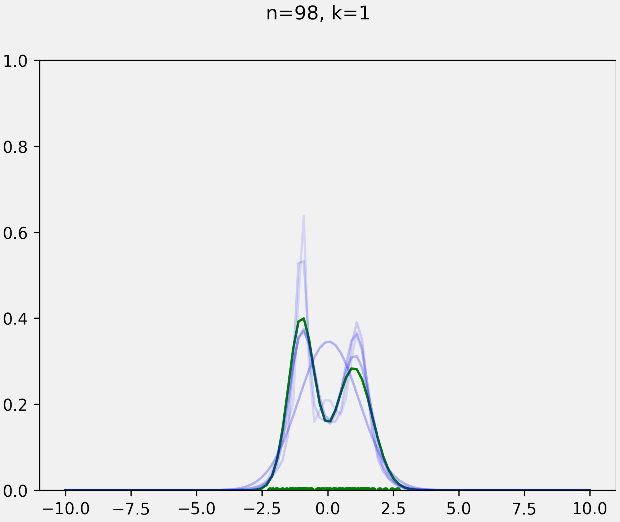
Bayesi modell-összehasonlítás – demo



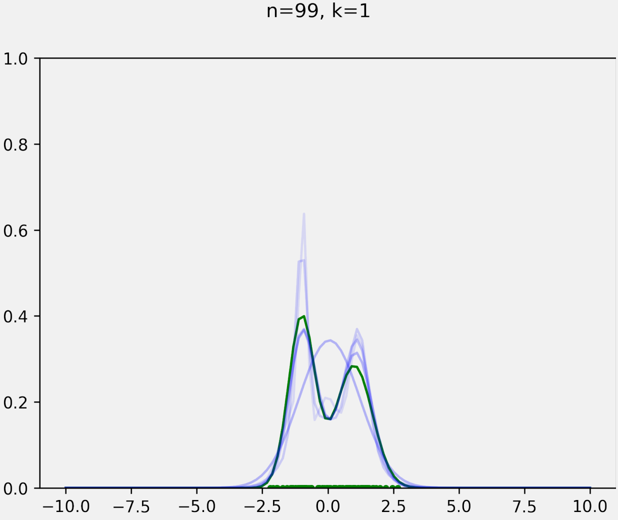
Bayesi modell-összehasonlítás – demo



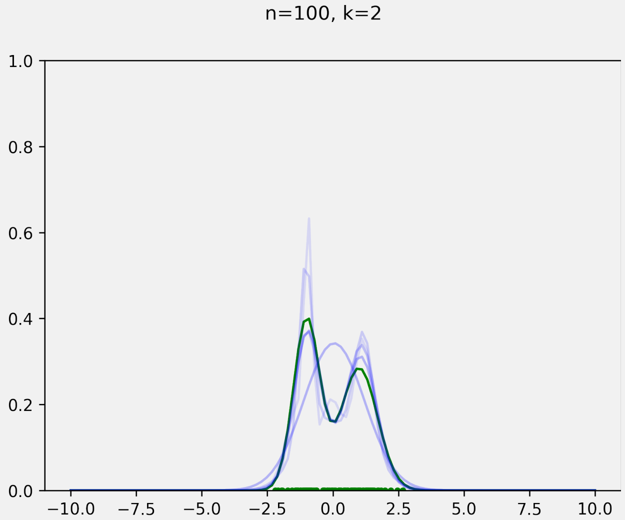
Bayesi modell-összehasonlítás – demo



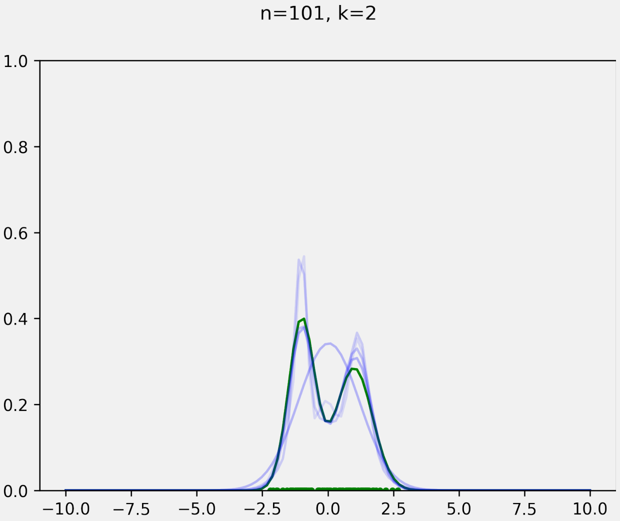
Bayesi modell-összehasonlítás – demo



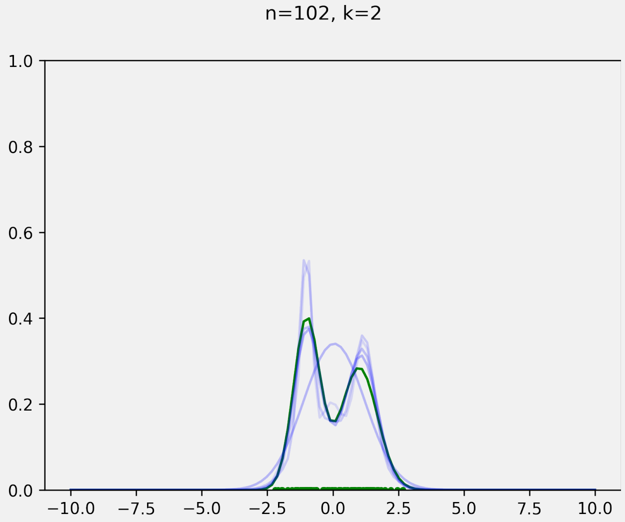
Bayesi modell-összehasonlítás – demo



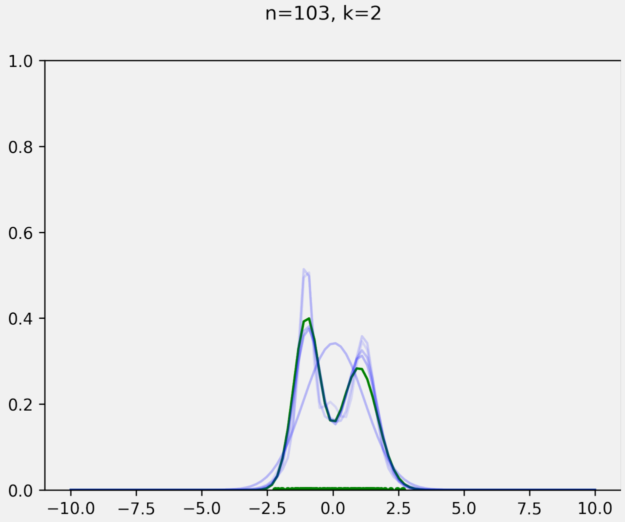
Bayesi modell-összehasonlítás – demo



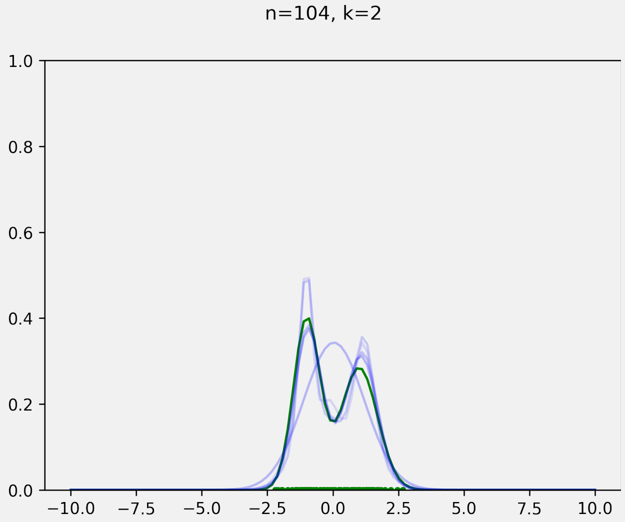
Bayesi modell-összehasonlítás – demo



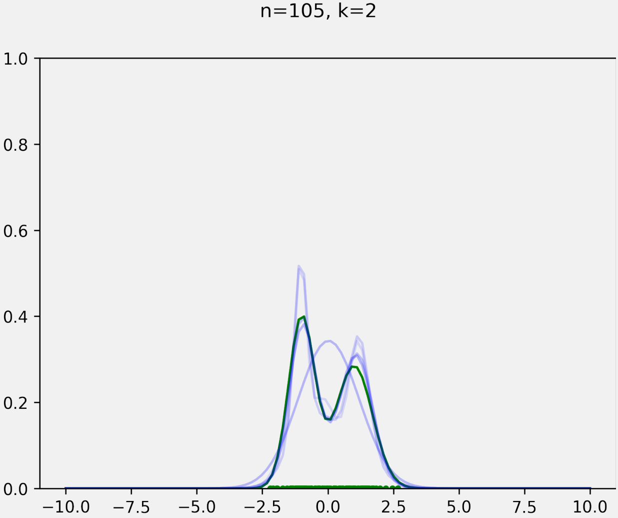
Bayesi modell-összehasonlítás – demo



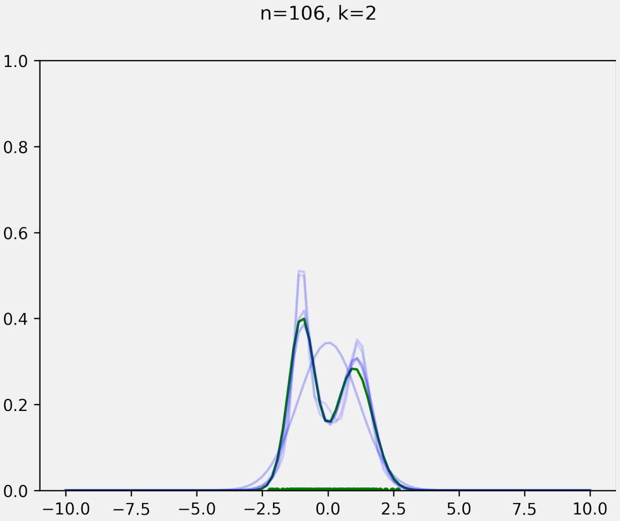
Bayesi modell-összehasonlítás – demo



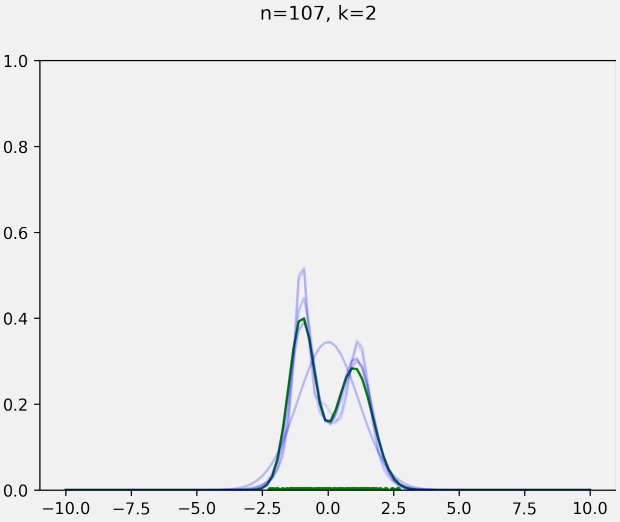
Bayesi modell-összehasonlítás – demo



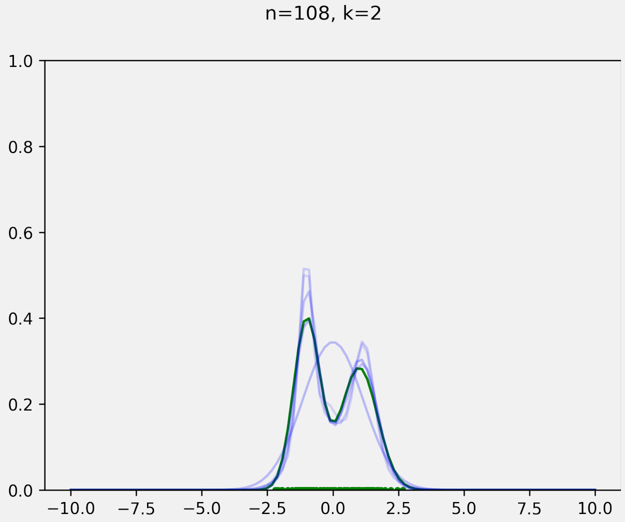
Bayesi modell-összehasonlítás – demo



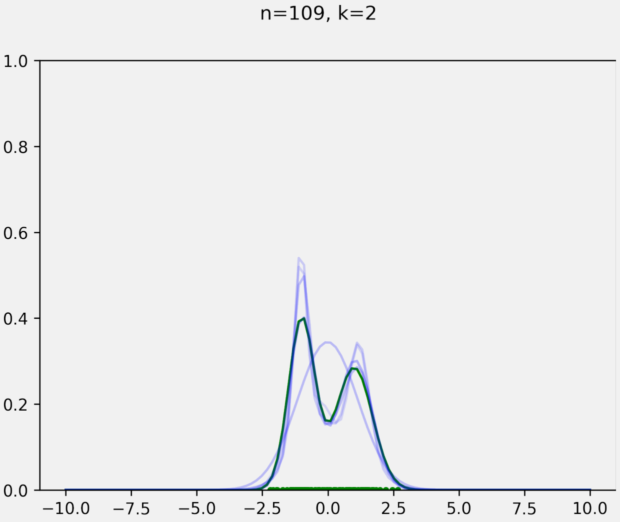
Bayesi modell-összehasonlítás – demo



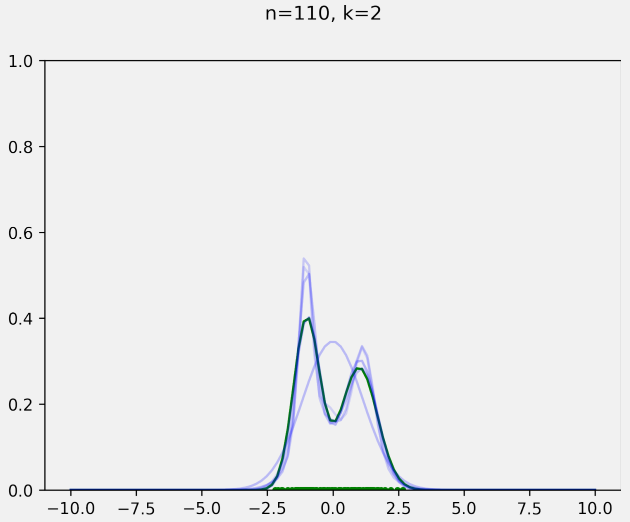
Bayesi modell-összehasonlítás – demo



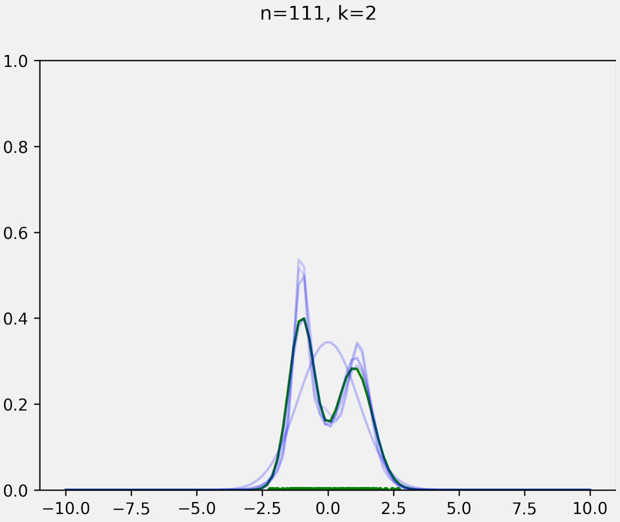
Bayesi modell-összehasonlítás – demo



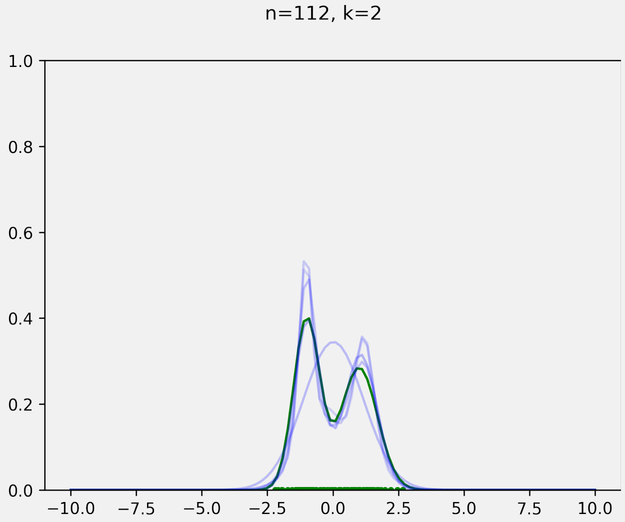
Bayesi modell-összehasonlítás – demo



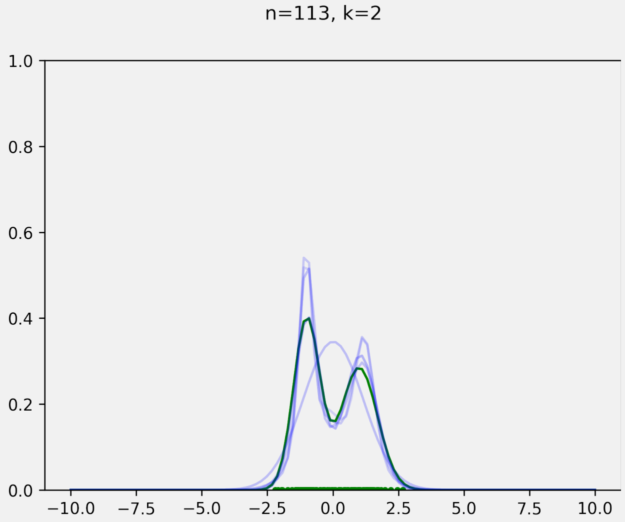
Bayesi modell-összehasonlítás – demo



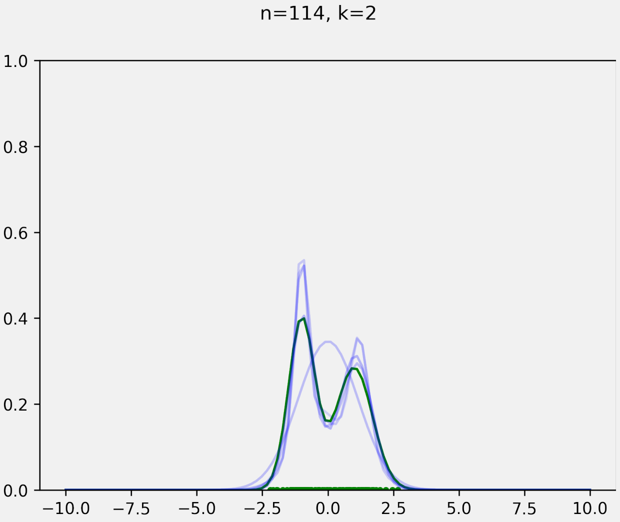
Bayesi modell-összehasonlítás – demo



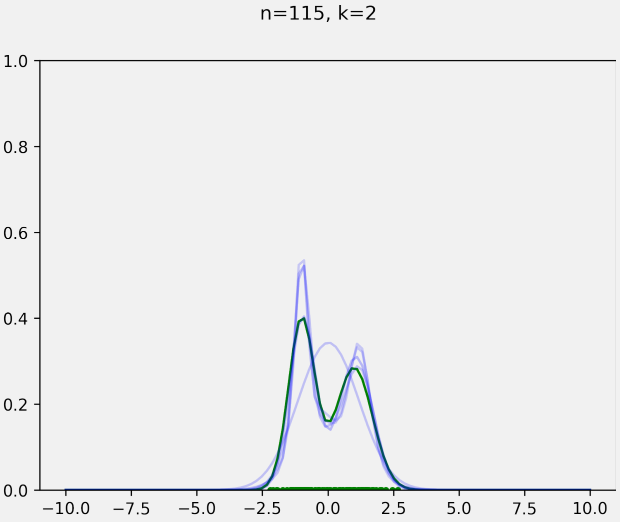
Bayesi modell-összehasonlítás – demo



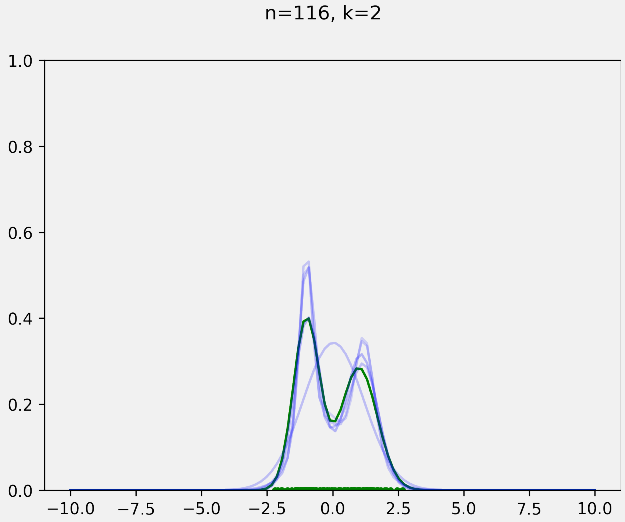
Bayesi modell-összehasonlítás – demo



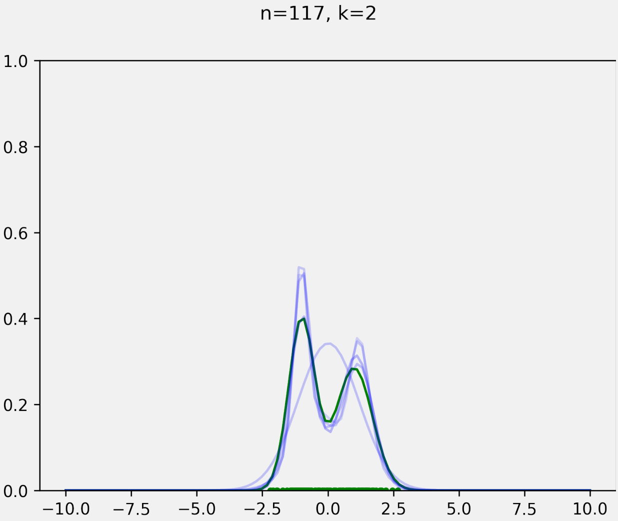
Bayesi modell-összehasonlítás – demo



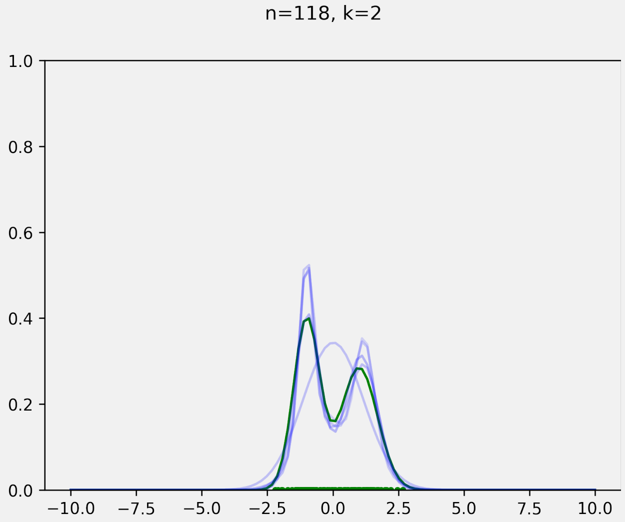
Bayesi modell-összehasonlítás – demo



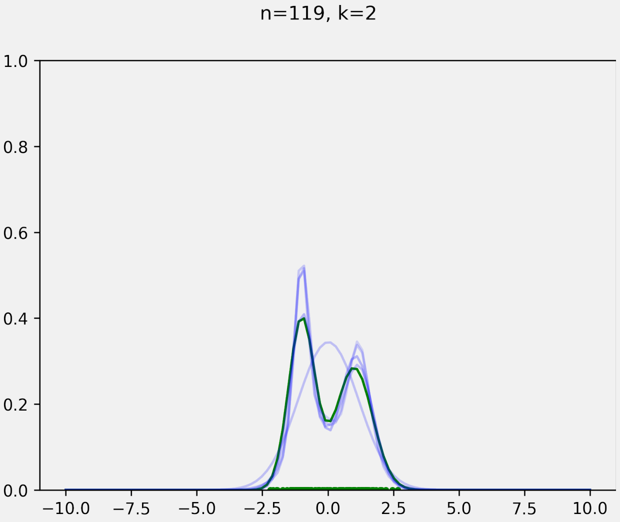
Bayesi modell-összehasonlítás – demo



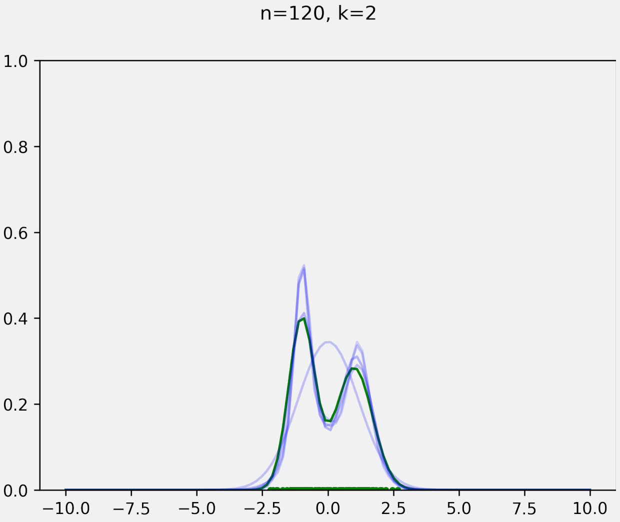
Bayesi modell-összehasonlítás – demo



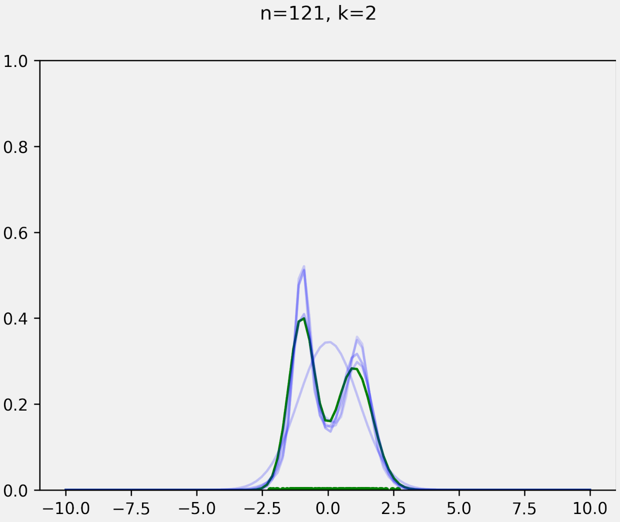
Bayesi modell-összehasonlítás – demo



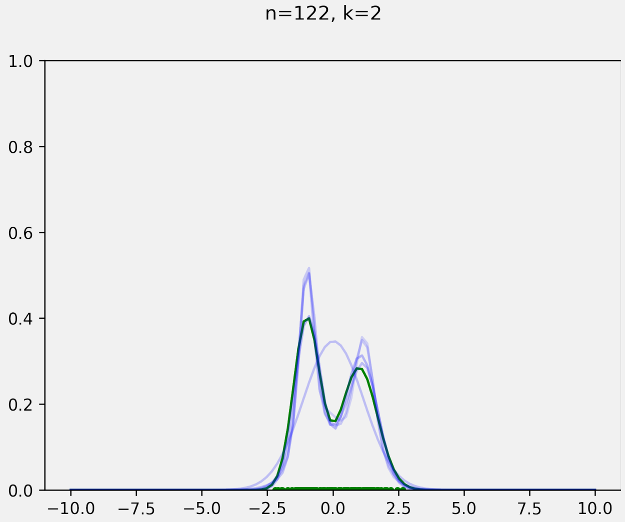
Bayesi modell-összehasonlítás – demo



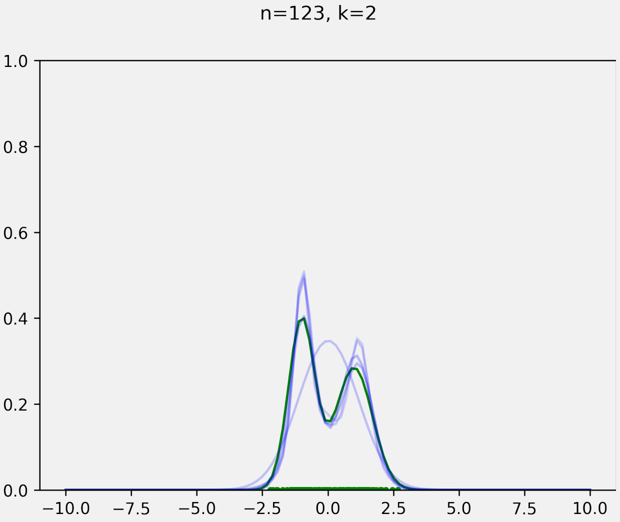
Bayesi modell-összehasonlítás – demo



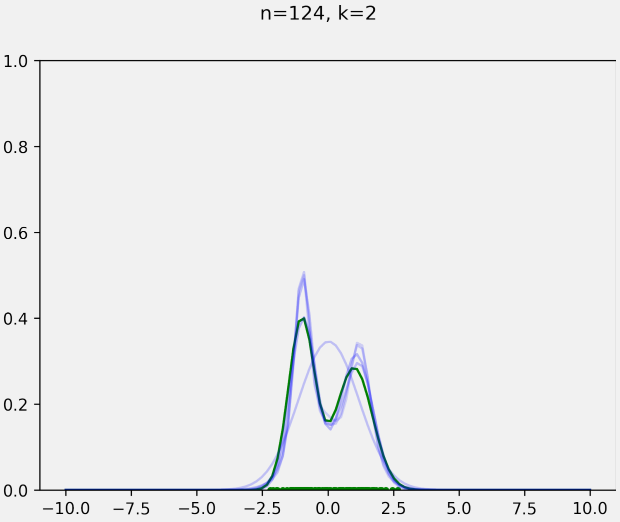
Bayesi modell-összehasonlítás – demo



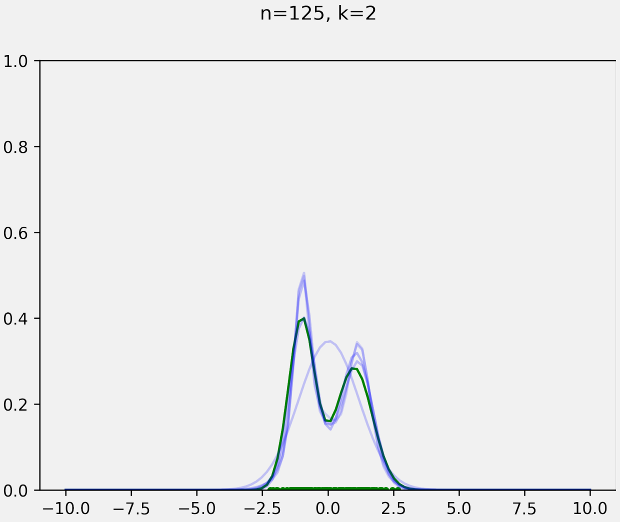
Bayesi modell-összehasonlítás – demo



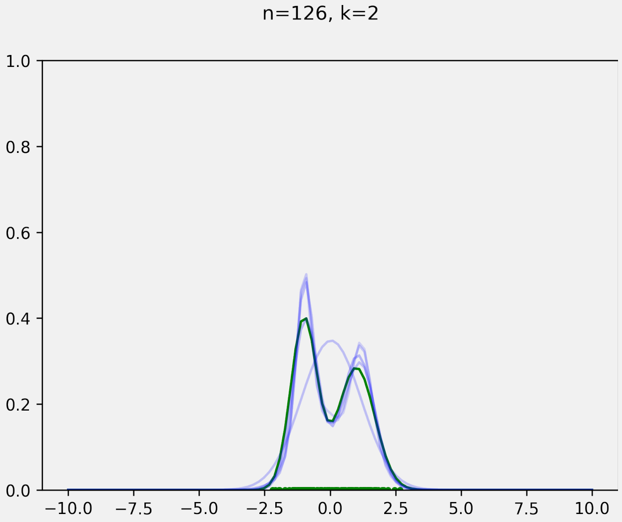
Bayesi modell-összehasonlítás – demo



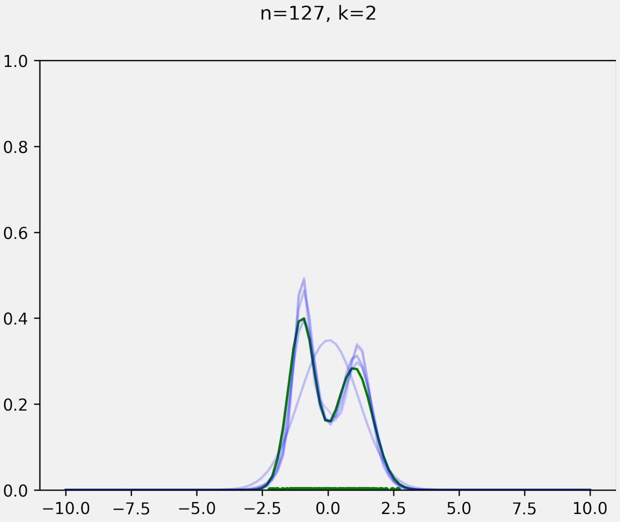
Bayesi modell-összehasonlítás – demo



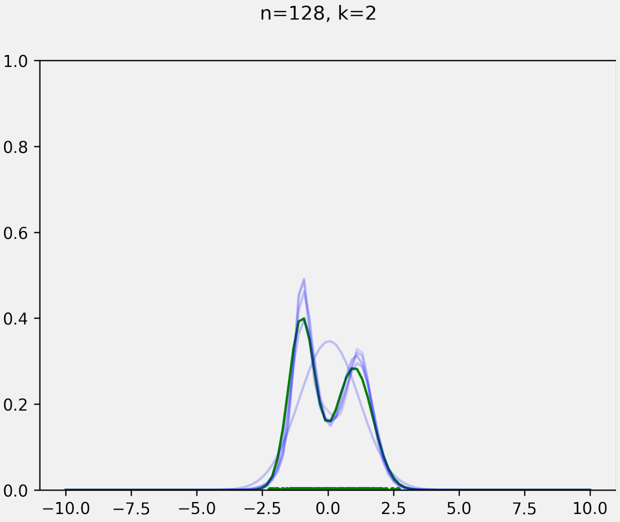
Bayesi modell-összehasonlítás – demo



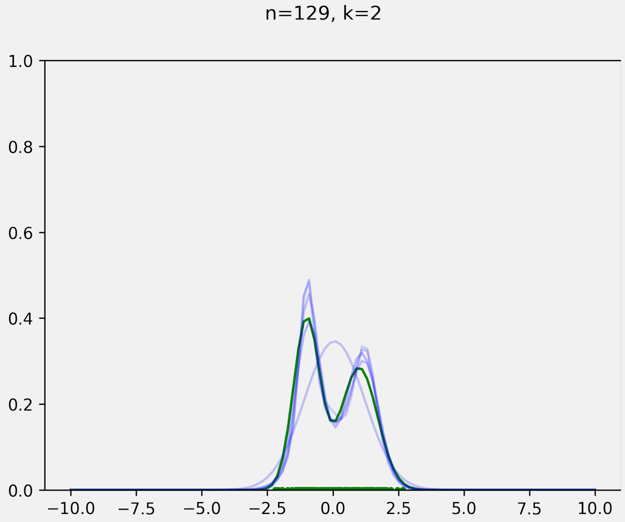
Bayesi modell-összehasonlítás – demo



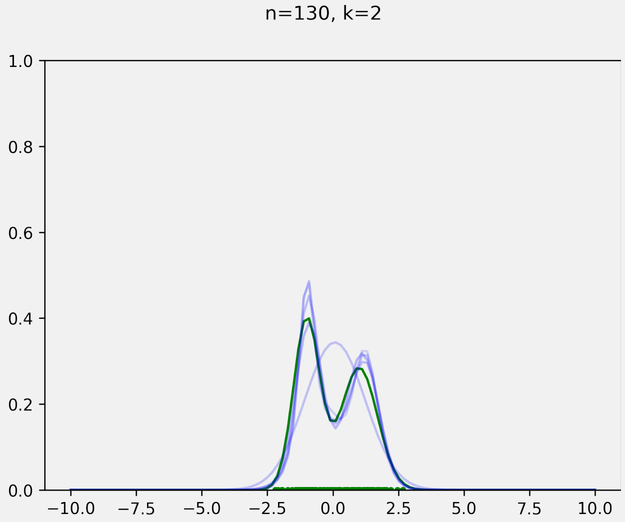
Bayesi modell-összehasonlítás – demo



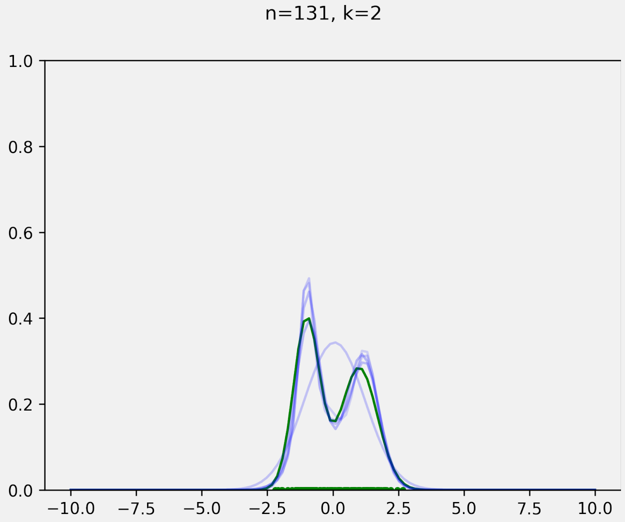
Bayesi modell-összehasonlítás – demo



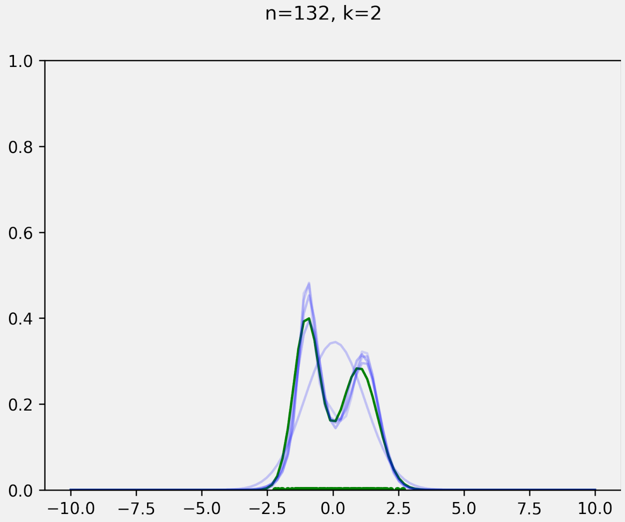
Bayesi modell-összehasonlítás – demo



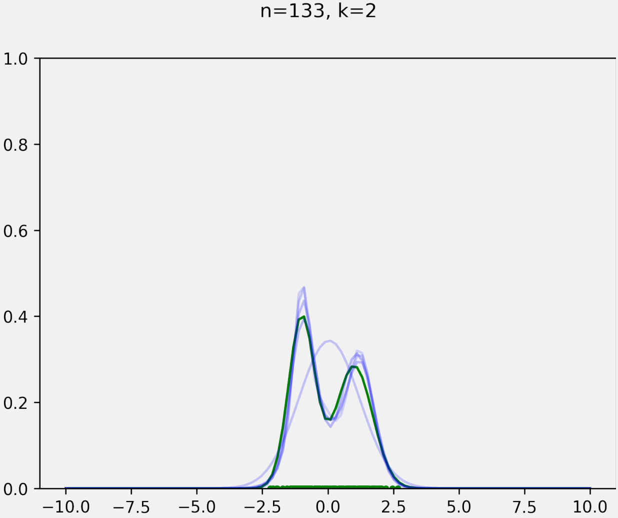
Bayesi modell-összehasonlítás – demo



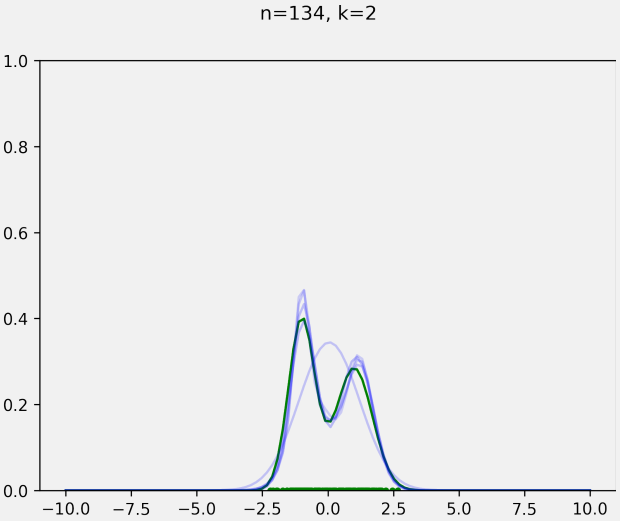
Bayesi modell-összehasonlítás – demo



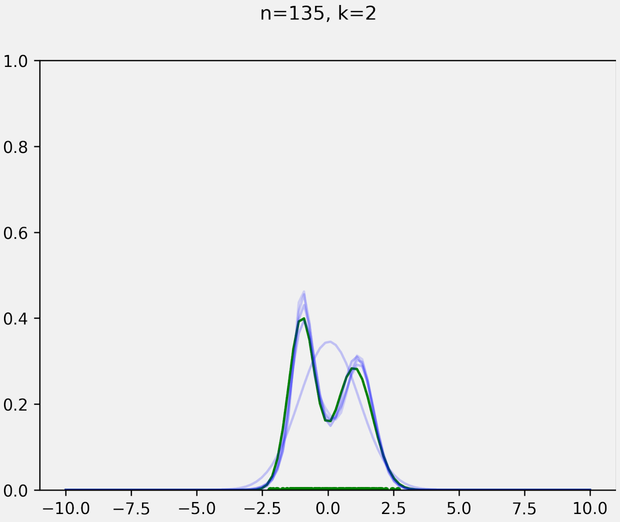
Bayesi modell-összehasonlítás – demo



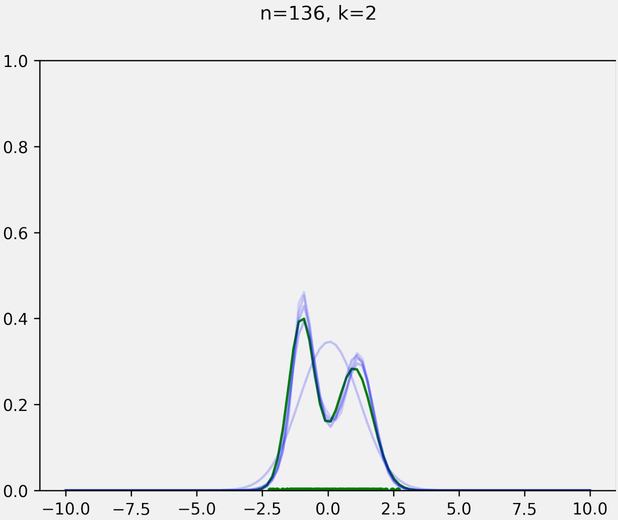
Bayesi modell-összehasonlítás – demo



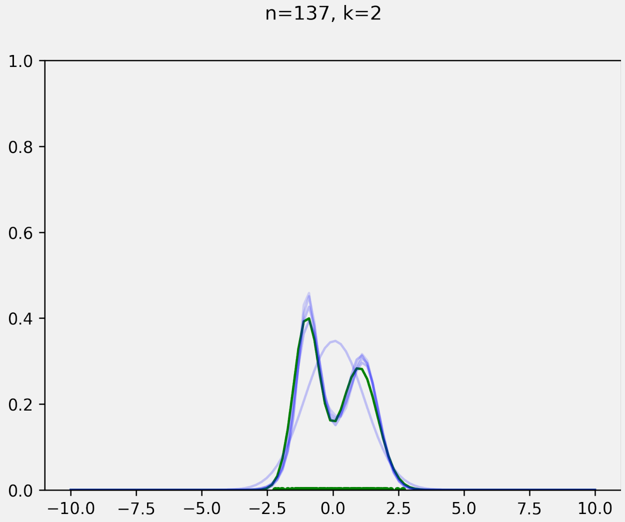
Bayesi modell-összehasonlítás – demo



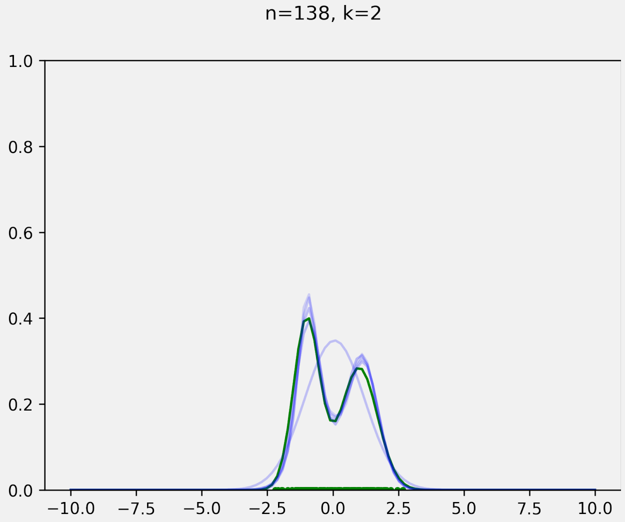
Bayesi modell-összehasonlítás – demo



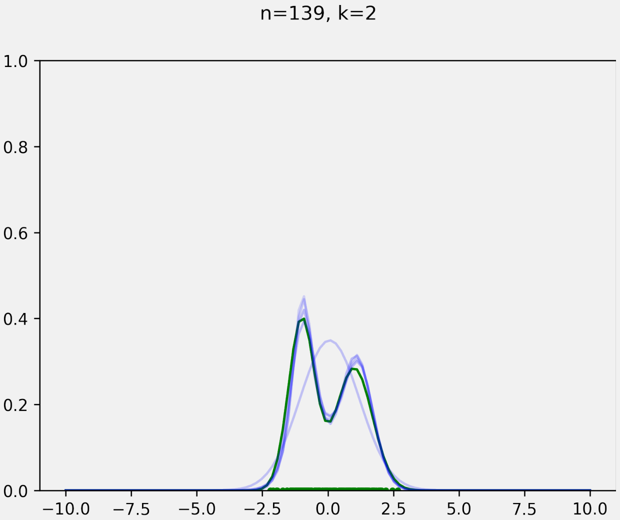
Bayesi modell-összehasonlítás – demo



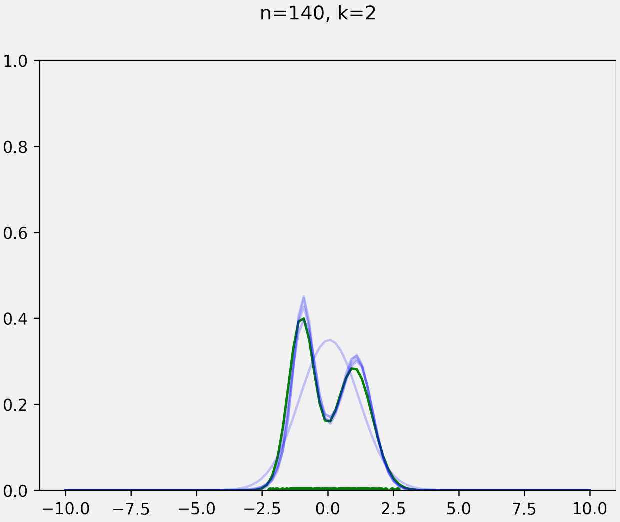
Bayesi modell-összehasonlítás – demo



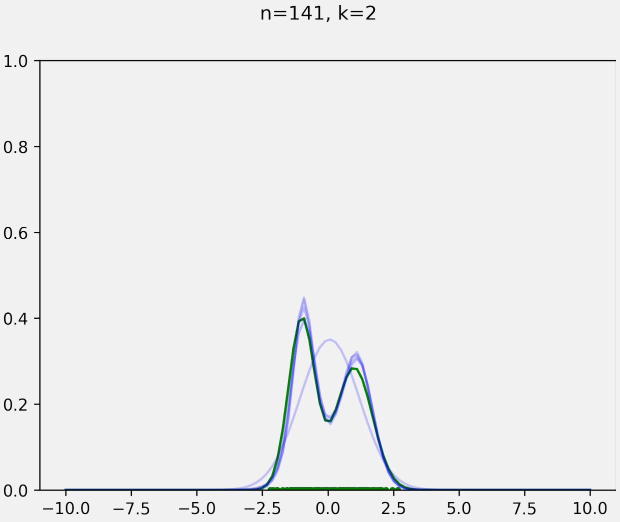
Bayesi modell-összehasonlítás – demo



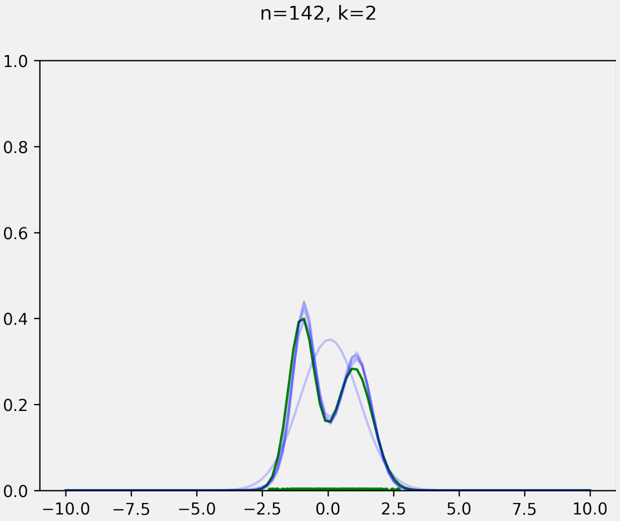
Bayesi modell-összehasonlítás – demo



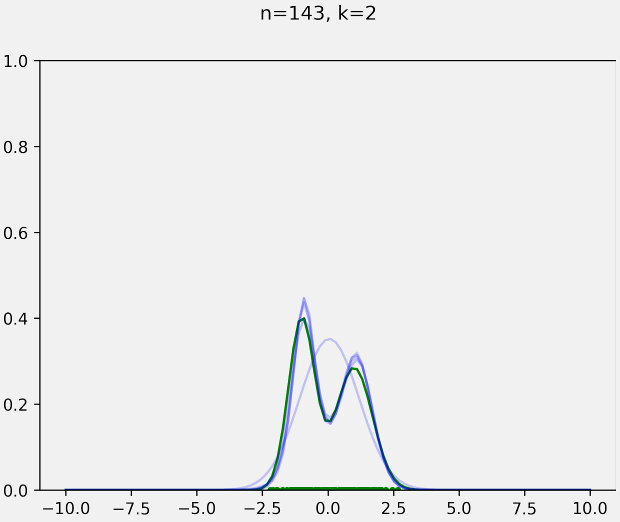
Bayesi modell-összehasonlítás – demo



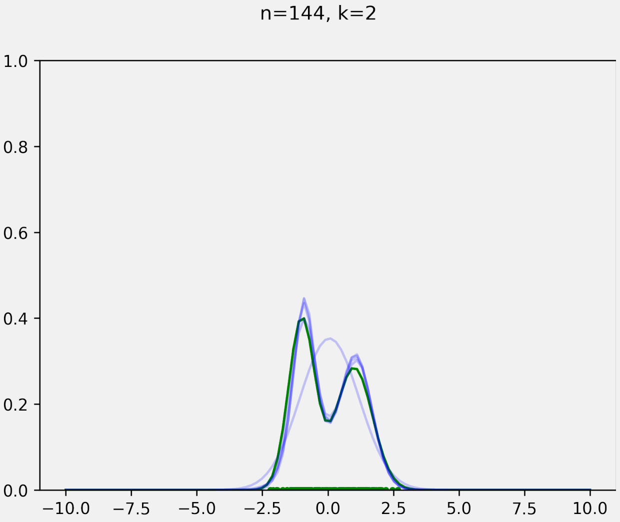
Bayesi modell-összehasonlítás – demo



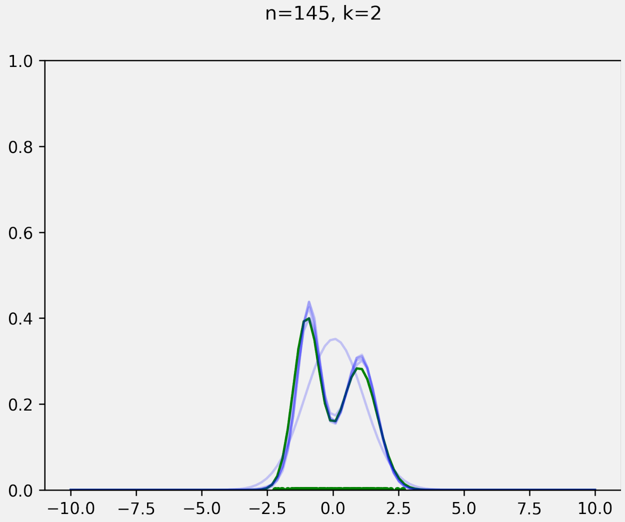
Bayesi modell-összehasonlítás – demo



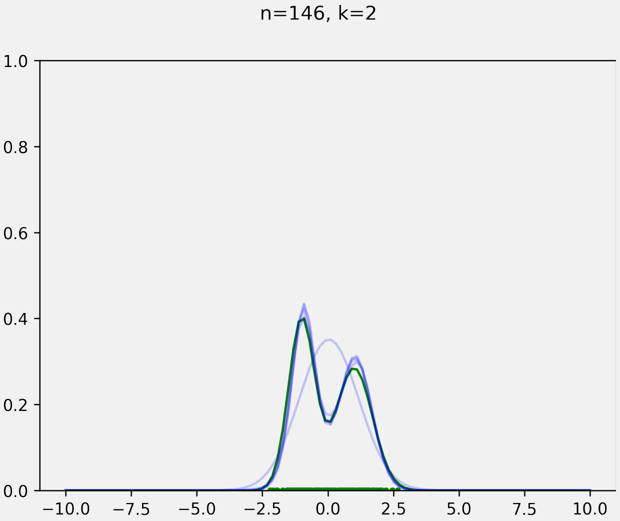
Bayesi modell-összehasonlítás – demo



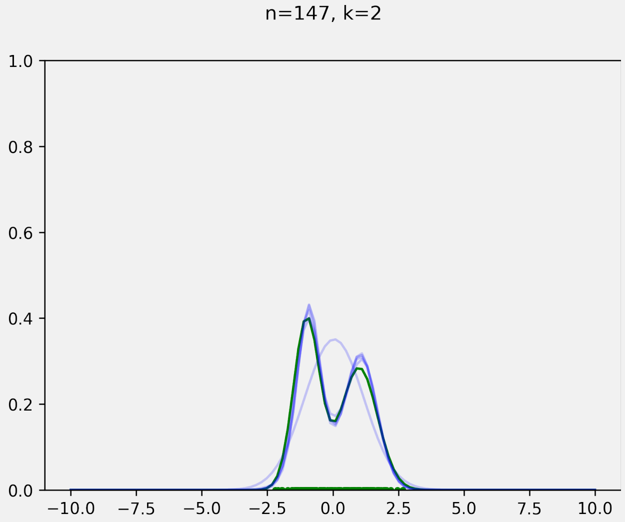
Bayesi modell-összehasonlítás – demo



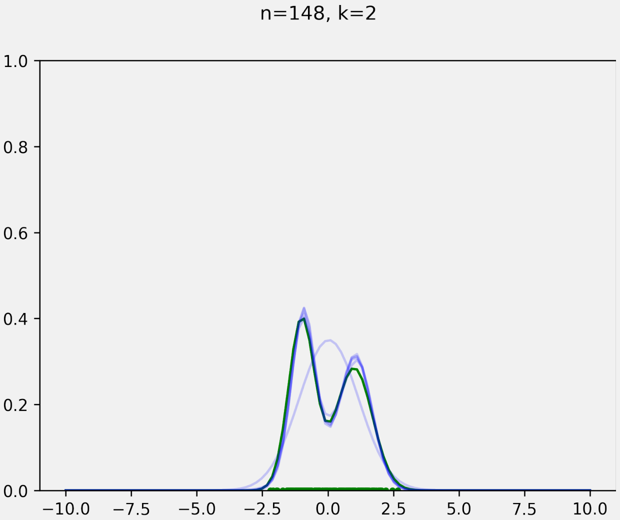
Bayesi modell-összehasonlítás – demo



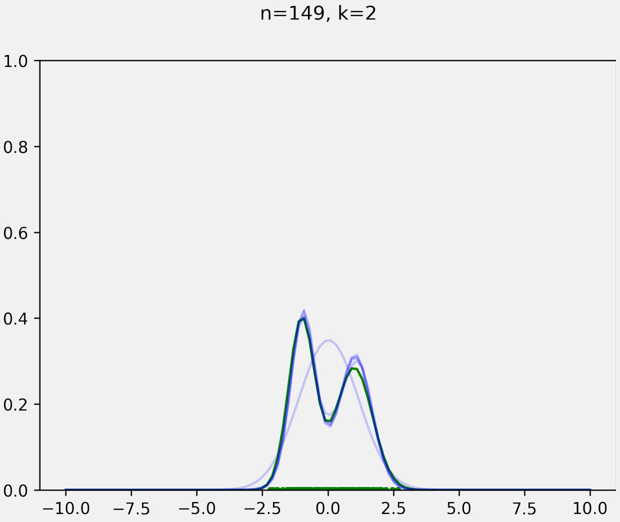
Bayesi modell-összehasonlítás – demo



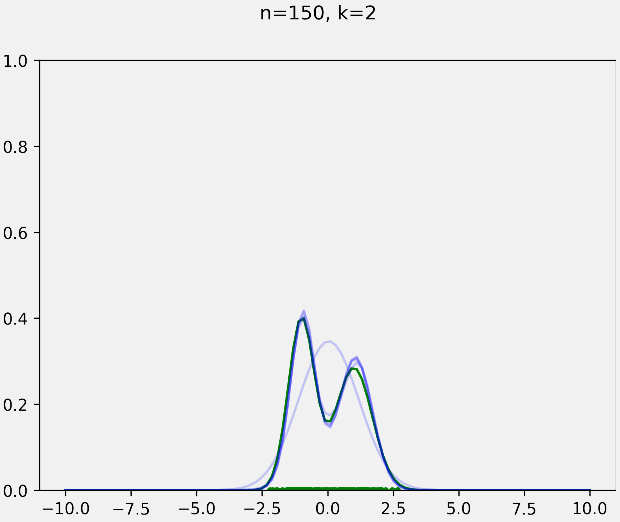
Bayesi modell-összehasonlítás – demo



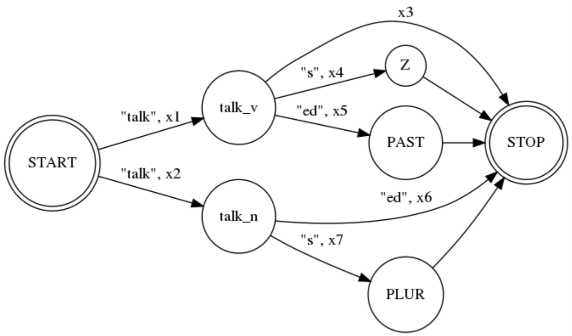
Bayesi modell-összehasonlítás – demo



Bayesi modell-összehasonlítás – demo

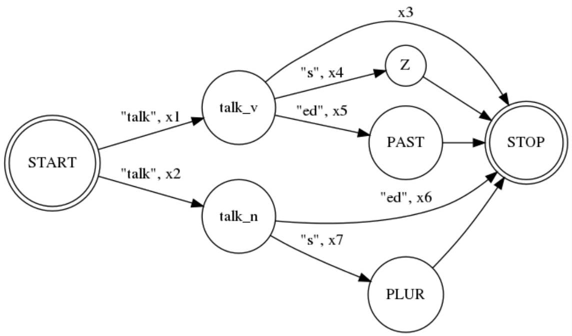


Véges állapotú morfológia



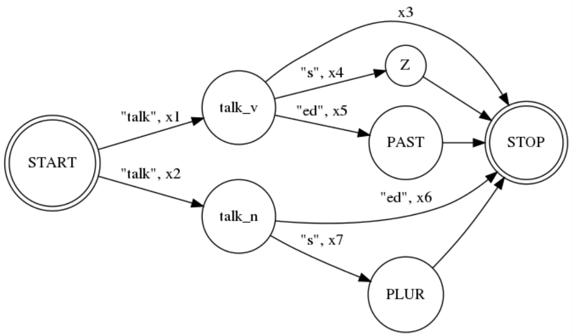
- Egy eloszlást definiál a $\{$*talk*, *talks*, *talked*$\}$ halmaz felett.

Véges állapotú morfológia



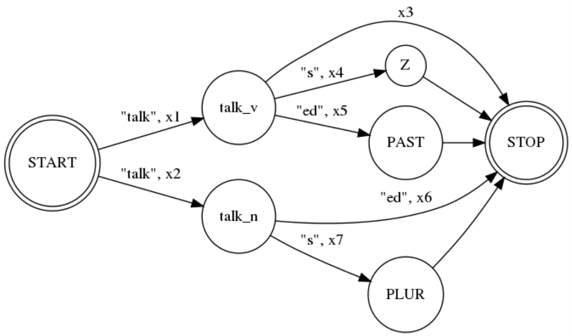
- Egy eloszlást definiál a $\{talk, talks, talked\}$ halmaz felett.
- az eloszlás differenciálható módon függ az $x_1, x_2 \ldots x_7$ paraméterektől.

Véges állapotú morfológia



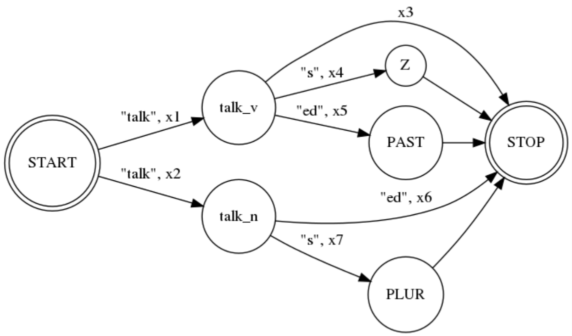
- Egy eloszlást definiál a $\{talk, talks, talked\}$ halmaz felett.
- az eloszlás differenciálható módon függ az $x_1, x_2 \ldots x_7$ paraméterektől.
- A paraméterekre megkötések vonatkoznak:

Véges állapotú morfológia



- Egy eloszlást definiál a $\{talk, talks, talked\}$ halmaz felett.
- az eloszlás differenciálható módon függ az $x_1, x_2 \ldots x_7$ paraméterektől.
- A paraméterekre megkötések vonatkoznak:
  - $x_i > 0$

Véges állapotú morfológia



- Egy eloszlást definiál a $\{talk, talks, talked\}$ halmaz felett.
- az eloszlás differenciálható módon függ az $x_1, x_2 \dots x_7$ paraméterektől.
- A paraméterekre megkötések vonatkoznak:
  - $x_i > 0$
  - $x_1 + x_2 = x_3 + x_4 + x_5 = x_6 + x_7 = 1$

Véges állapotú automata tanítása

- Mi legyen az architektúra?

Véges állapotú automata tanítása

- Mi legyen az architektúra?
  - diszkrét optimalizálási feladat

Véges állapotú automata tanítása

- Mi legyen az architektúra?
  - diszkrét optimalizálási feladat
- Adott architektúrára az élsúlyok megtalálása.

Véges állapotú automata tanítása

- Mi legyen az architektúra?
  - diszkrét optimalizálási feladat
- Adott architektúrára az élsúlyok megtalálása.
  - folytonos optimalizálási feladat

Véges állapotú automata tanítása

- Mi legyen az architektúra?
  - diszkrét optimalizálási feladat
- Adott architektúrára az élsúlyok megtalálása.
  - folytonos optimalizálási feladat
  - constrained optimization

Véges állapotú automata tanítása

- Mi legyen az architektúra?
  - diszkrét optimalizálási feladat
- Adott architektúrára az élsúlyok megtalálása.
  - folytonos optimalizálási feladat
  - constrained optimization
- Keressük az optimális automatát genetikus algoritmussal (Csikós, 2019)

Véges állapotú automata tanítása

- Mi legyen az architektúra?
  - diszkrét optimalizálási feladat
- Adott architektúrára az élsúlyok megtalálása.
  - folytonos optimalizálási feladat
  - constrained optimization
- Keressük az optimális automatát genetikus algoritmussal (Csikós, 2019)
  - melynek egy generációja egy bayesi modell-összehasonlítás.

Véges állapotú automata tanítása

- Mi legyen az architektúra?
  - diszkrét optimalizálási feladat
- Adott architektúrára az élsúlyok megtalálása.
  - folytonos optimalizálási feladat
  - constrained optimization
- Keressük az optimális automatát genetikus algoritmussal (Csikós, 2019)
  - melynek egy generációja egy bayesi modell-összehasonlítás.
- (Borbély és Kornai, 2019)

Hivatkozások

Borbély, G., Kornai, A.: Sentence length. In: Proceedings of the 16th Meeting on the Mathematics of Language. pp. 114–125. Association for Computational Linguistics, Toronto, Canada (18–19 Jul 2019), https://www.aclweb.org/anthology/W19-5710

Csikós, D.: Véges állapotú automaták tanulása genetikus algoritmussal. BSc Thesis, Budapest University of Technology and Economics (2019), https://math.bme.hu/~csikosd/Szakdolgozat.pdf

MacKay, D.J.: Information Theory, Inference, and Learning Algorithms. Cambridge University Press (2003)