KU LEUVEN

Bayesian D- and I-optimal designs for choice experiments with mixtures using a multinomial logit model

Mario Becerra
Peter Goos

Outline

1. Choice modeling and choice experiments
2. Mixture experiments
3. Combining choice models and mixture models
4. Optimality criteria for choice experiments
5. Results
6. Conclusions and future work

Discrete choice experiments

Discrete choice experiments

- Quantify consumer preferences

Discrete choice experiments

- Quantify consumer preferences
- Preference data is collected

Discrete choice experiments

- Quantify consumer preferences
- Preference data is collected
- Respondents are presented sets of alternatives (choice sets) and asked to choose

Discrete choice experiments

- Quantify consumer preferences
- Preference data is collected
- Respondents are presented sets of alternatives (choice sets) and asked to choose
- Example: a customer responding whether they prefer to buy product A, B or C

Discrete choice experiments

- Quantify consumer preferences
- Preference data is collected
- Respondents are presented sets of alternatives (choice sets) and asked to choose
- Example: a customer responding whether they prefer to buy product A, B or C
- Models assume a latent utility function used to derive the probability of each respondent making each decision

Mixtures

Mixtures

- Many products and services can be described as mixtures of ingredients

Mixtures

- Many products and services can be described as mixtures of ingredients
- Examples:
  - components of a mobility budget such as car and public transportation card

Mixtures

- Many products and services can be described as mixtures of ingredients
- Examples:
  - components of a mobility budget such as car and public transportation card
  - the wheat varieties used to make bread

Mixtures

- Many products and services can be described as mixtures of ingredients
- Examples:
  - components of a mobility budget such as car and public transportation card
  - the wheat varieties used to make bread
  - ingredients used to make a cocktail

Mixtures

- In mixture experiments, products are expressed as combinations of proportions of ingredients

Mixtures

- In mixture experiments, products are expressed as combinations of proportions of ingredients
- The researchers' interest is generally in one or more characteristics of the mixture

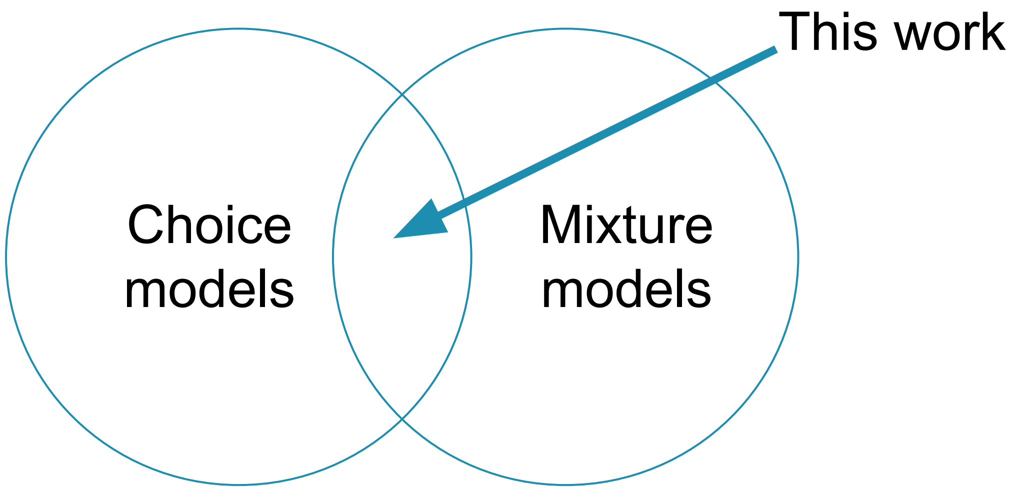
Mixtures

- In mixture experiments, products are expressed as combinations of proportions of ingredients
- The researchers' interest is generally in one or more characteristics of the mixture
- For us, the characteristic of interest is the **preference** of respondents

Mixtures

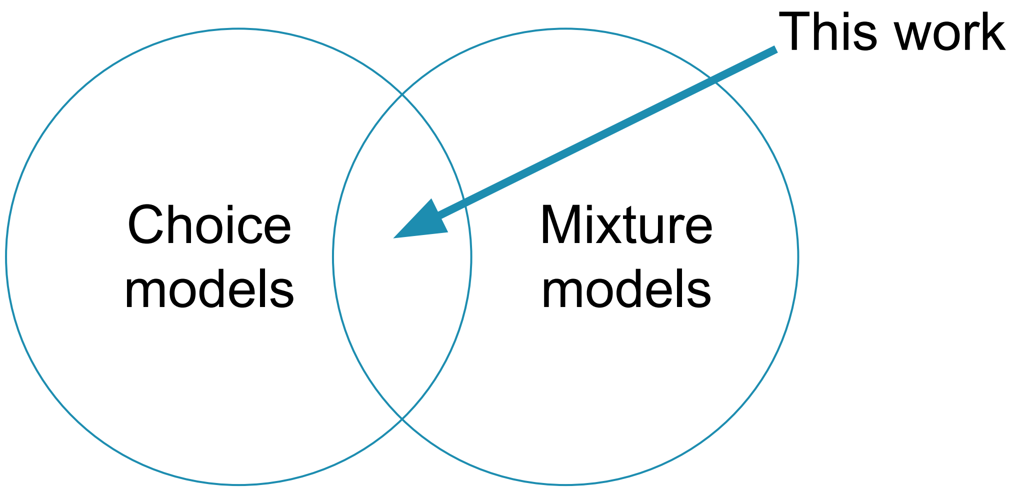
- In mixture experiments, products are expressed as combinations of proportions of ingredients
- The researchers' interest is generally in one or more characteristics of the mixture
- For us, the characteristic of interest is the **preference** of respondents
- Choice experiments are ideal to collect data for quantifying and modeling preferences for mixtures

Choice experiments with mixtures



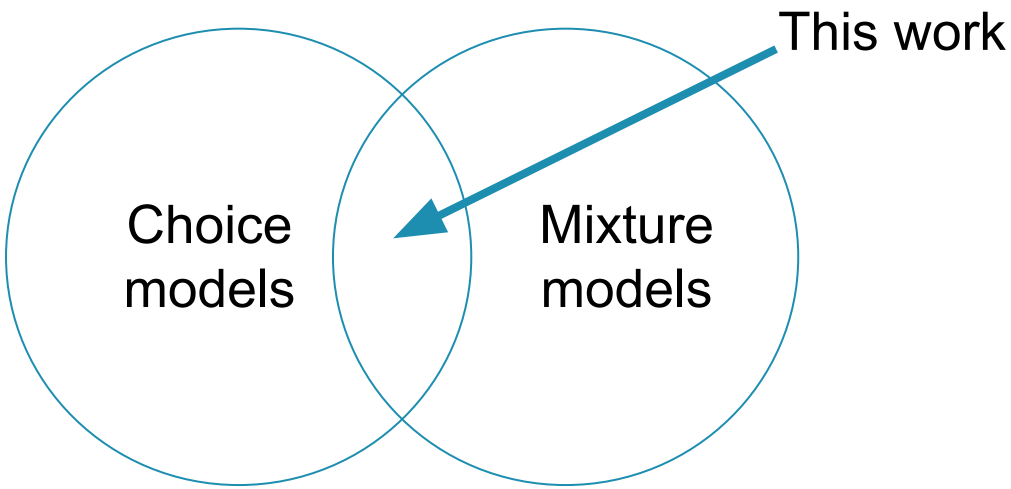
Choice experiments with mixtures

- First example by Courcoux and Séménou (1997)



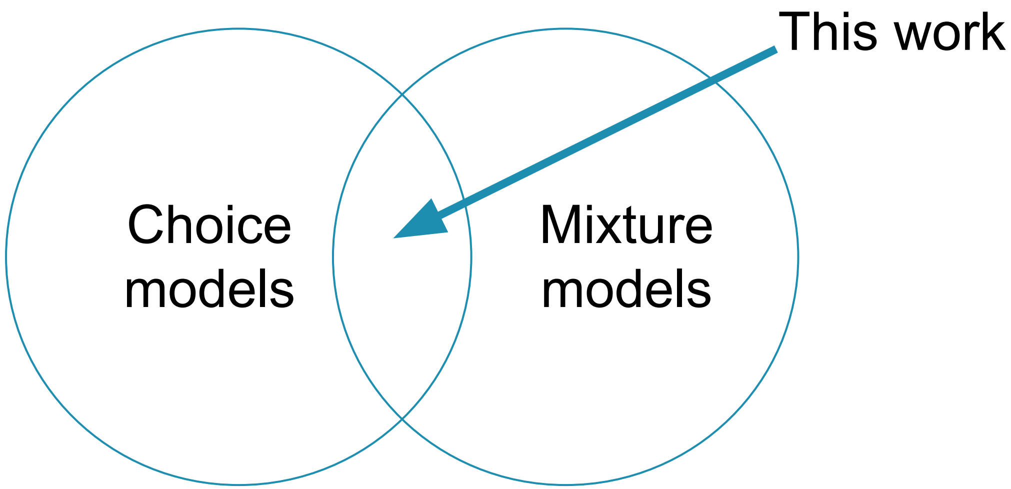
Choice experiments with mixtures

- First example by Courcoux and Séménou (1997)
- Preferences for cocktails involving different proportions of mango juice, lime juice, and blackcurrant syrup



Choice experiments with mixtures

- First example by Courcoux and Séménou (1997)
- Preferences for cocktails involving different proportions of mango juice, lime juice, and blackcurrant syrup
- Experimental data involved the responses of sixty people, each making eight pairwise comparisons of different cocktails



Designing choice experiments with mixtures

Designing choice experiments with mixtures

- Discrete choice experiments are expensive, cumbersome and time-consuming

Designing choice experiments with mixtures

- Discrete choice experiments are expensive, cumbersome and time-consuming
- We need efficient experimental designs so that the experiments provide reliable information

Designing choice experiments with mixtures

- Discrete choice experiments are expensive, cumbersome and time-consuming
- We need efficient experimental designs so that the experiments provide reliable information
- Optimal design of experiments is the branch of statistics that deals with the construction of efficient experimental designs

Optimal choice experiments with mixtures

Optimal choice experiments with mixtures

- Little research has been done concerning the optimal design of choice experiments with mixtures

Optimal choice experiments with mixtures

- Little research has been done concerning the optimal design of choice experiments with mixtures
- D-optimal experimental designs are good if we want low-variance estimators

Optimal choice experiments with mixtures

- Little research has been done concerning the optimal design of choice experiments with mixtures
- D-optimal experimental designs are good if we want low-variance estimators
- In experiments with mixtures, we want to optimize the composition of the mixture to maximize consumer preference

Optimal choice experiments with mixtures

- Little research has been done concerning the optimal design of choice experiments with mixtures
- D-optimal experimental designs are good if we want low-variance estimators
- In experiments with mixtures, we want to optimize the composition of the mixture to maximize consumer preference
- Precise predictions are crucial

Optimal choice experiments with mixtures

- Little research has been done concerning the optimal design of choice experiments with mixtures
- D-optimal experimental designs are good if we want low-variance estimators
- In experiments with mixtures, we want to optimize the composition of the mixture to maximize consumer preference
- Precise predictions are crucial
- I-optimal designs are more suitable

Models for data from mixture experiments

Models for data from mixture experiments

- Models assume two or more ingredients and a response variable that depends only on the **relative** proportions of the ingredients in the mixture

Models for data from mixture experiments

- Models assume two or more ingredients and a response variable that depends only on the **relative** proportions of the ingredients in the mixture
- Each mixture is described as a combination of q ingredient proportions, with the constraint that these proportions sum up to one

Models for data from mixture experiments

- Models assume two or more ingredients and a response variable that depends only on the **relative** proportions of the ingredients in the mixture
- Each mixture is described as a combination of q ingredient proportions, with the constraint that these proportions sum up to one
- Dedicated models are needed to avoid perfect collinearity

Models for data from mixture experiments

- Models assume two or more ingredients and a response variable that depends only on the **relative** proportions of the ingredients in the mixture
- Each mixture is described as a combination of *q* ingredient proportions, with the constraint that these proportions sum up to one
- Dedicated models are needed to avoid perfect collinearity
- Special-cubic Scheffé model:

$$Y = \sum_{i=1}^{q} \beta_i x_i + \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \beta_{ij} x_i x_j + \sum_{i=1}^{q-2} \sum_{j=i+1}^{q-1} \sum_{k=j+1}^{q} \beta_{ijk} x_i x_j x_k + \varepsilon$$

Multinomial logit model for choice data

Multinomial logit model for choice data

- A respondent faces S choice sets involving J alternatives each

Multinomial logit model for choice data

- A respondent faces S choice sets involving J alternatives each
- Respondent chooses the alternative that has the highest perceived utility

Multinomial logit model for choice data

- A respondent faces *S* choice sets involving *J* alternatives each
- Respondent chooses the alternative that has the highest perceived utility

	Car A	Car B
BRAND	BMW	Mercedes
MILEAGE	2 miles per gallon	10 miles per gallon
COLOR	British racing green	Mettalic Green
PRICE	$20,000	$100,000
which do you prefer		

Multinomial logit model for choice data

- A respondent faces *S* choice sets involving *J* alternatives each
- Respondent chooses the alternative that has the highest perceived utility

	Car A	Car B
BRAND	BMW	Mercedes
MILEAGE	2 miles per gallon	10 miles per gallon
COLOR	British racing green	Mettalic Green
PRICE	$20,000	$100,000
which do you prefer	○	○

- The probability that a respondent chooses alternative $j \in \{1, ..., J\}$ in choice set s is

$$p_{js} = \frac{\exp\left[\boldsymbol{f}^T(\boldsymbol{x}_{js})\boldsymbol{\beta}\right]}{\sum_{t=1}^{J} \exp\left[\boldsymbol{f}^T(\boldsymbol{x}_{ts})\boldsymbol{\beta}\right]}$$

Model for choice data concerning mixtures

Model for choice data concerning mixtures

- We assume vector $\boldsymbol{x}_{js}$ contains the *q* ingredient proportions and that $\boldsymbol{f}(\boldsymbol{x}_{js})$ represents the model expansion of these proportions

Model for choice data concerning mixtures

- We assume vector $\boldsymbol{x}_{js}$ contains the q ingredient proportions and that $\boldsymbol{f}(\boldsymbol{x}_{js})$ represents the model expansion of these proportions

- Perceived utility modeled as

$$U_{js} = \boldsymbol{f}^T(\boldsymbol{x}_{js})\boldsymbol{\beta} = \sum_{i=1}^{q-1} \beta_i^* x_{ijs} + \sum_{i=1}^{q-1} \sum_{k=i+1}^{q} \beta_{ik} x_{ijs} x_{kjs} + \sum_{i=1}^{q-2} \sum_{k=i+1}^{q-1} \sum_{l=k+1}^{q} \beta_{ikl} x_{ijs} x_{kjs} x_{ljs} + \varepsilon_{js}$$

D-optimal designs

D-optimal designs

- D-optimality criterion

$$\mathcal{D} = \log\left(\det\left(\left[\boldsymbol{I}^{-1}(\boldsymbol{X},\boldsymbol{\beta})\right)\right]^{\frac{1}{r}}\right)$$

D-optimal designs

- D-optimality criterion

$$\mathcal{D} = \log \left(\det \left(\left[\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta}) \right) \right]^{\frac{1}{r}} \right)$$

- Bayesian D-optimality criterion

$$\mathcal{D}_B = \log \left(\int_{\mathbb{R}^r} \left[\det \left(\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta}) \right) \right]^{\frac{1}{r}} \pi(\boldsymbol{\beta}) d\boldsymbol{\beta} \right)$$

I-optimal designs

I-optimal designs

- I-optimality criterion

$$\mathcal{I} = \mathrm{tr}\left[\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta})\boldsymbol{W}_u\right]$$

$$\boldsymbol{W}_u = \int_{\chi} \boldsymbol{f}(\boldsymbol{x}_{js})\boldsymbol{f}^T(\boldsymbol{x}_{js})d\boldsymbol{x}_{js}$$

I-optimal designs

- I-optimality criterion

$$\mathcal{I} = \mathrm{tr}\left[\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta})\boldsymbol{W}_u\right]$$

- Bayesian I-optimality criterion

$$\mathcal{I}_B = \int_{\mathbb{R}^r} \mathrm{tr}\left[\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta})\boldsymbol{W}_u\right] \pi(\boldsymbol{\beta}) d\boldsymbol{\beta}$$

$$\boldsymbol{W}_u = \int_{\chi} \boldsymbol{f}(\boldsymbol{x}_{js})\boldsymbol{f}^T(\boldsymbol{x}_{js})d\boldsymbol{x}_{js}$$

Cocktail preferences

Cocktail preferences

- Ruseckaite et al. revisited an experiment by Courcoux and Semenou in which respondents tasted fruit cocktails involving mango juice, blackcurrant syrup, and lemon juice

Cocktail preferences

- Ruseckaite et al. revisited an experiment by Courcoux and Semenou in which respondents tasted fruit cocktails involving mango juice, blackcurrant syrup, and lemon juice
- 60 consumers were asked to taste different pairs of seven fruit cocktails

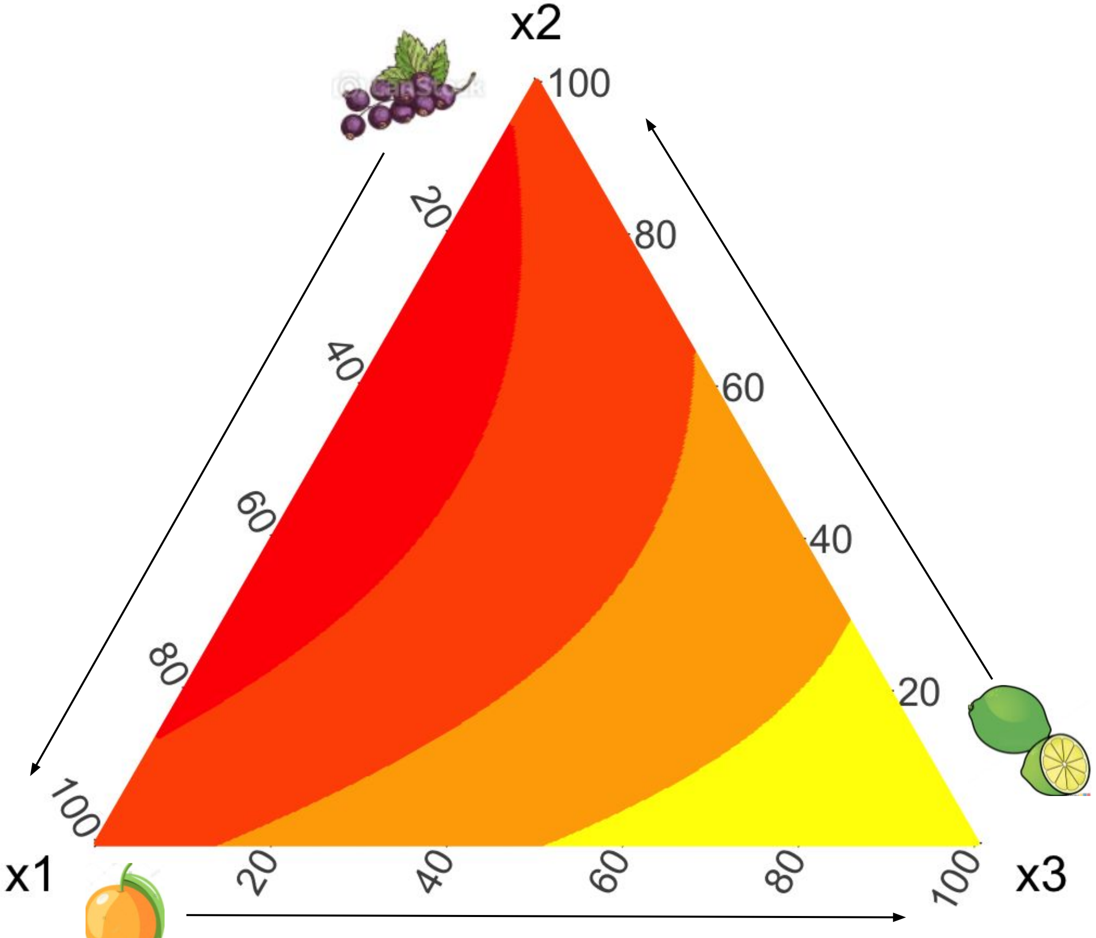
Cocktail preferences

- Ruseckaite et al. revisited an experiment by Courcoux and Semenou in which respondents tasted fruit cocktails involving mango juice, blackcurrant syrup, and lemon juice
- 60 consumers were asked to taste different pairs of seven fruit cocktails
- Ruseckaite et al. obtained a prior distribution for parameter vector $\boldsymbol{\beta}$ in a special-cubic Scheffé model

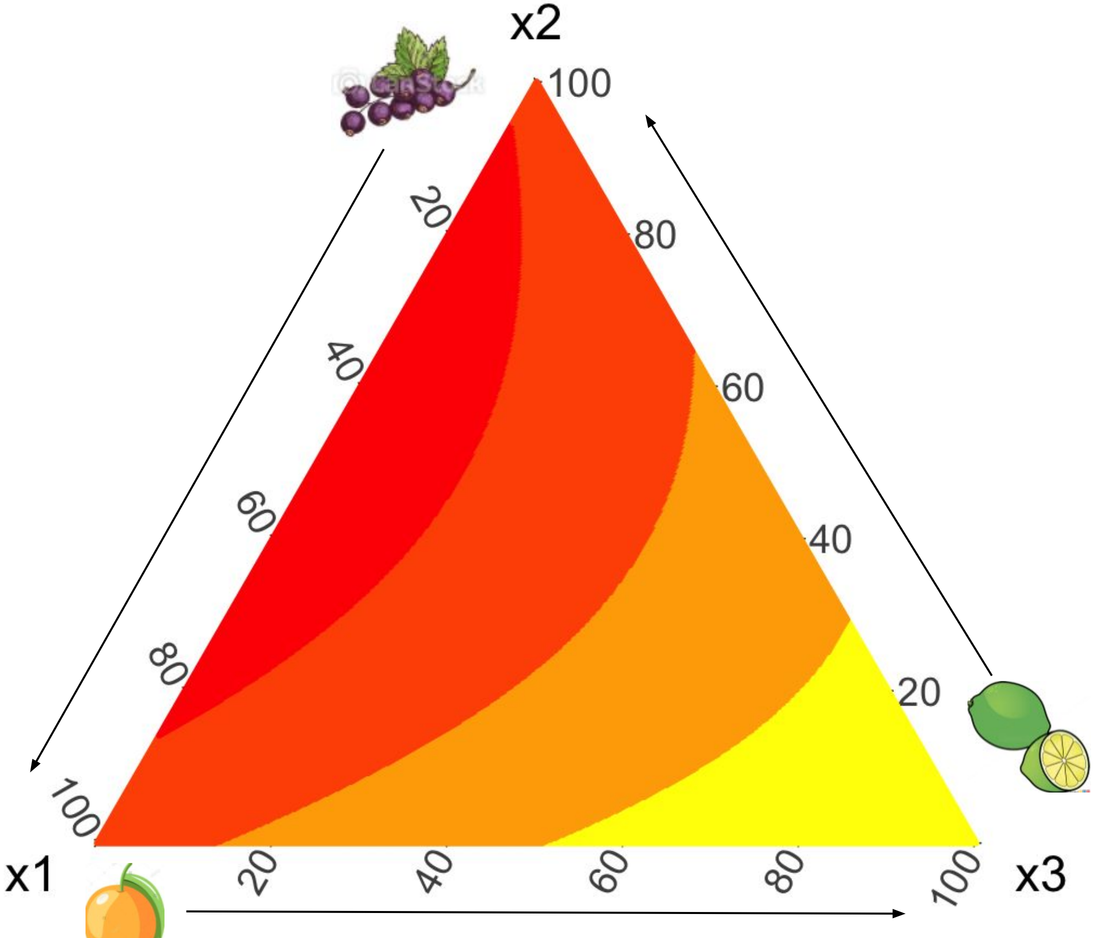
Cocktail preferences

- Ruseckaite et al. revisited an experiment by Courcoux and Semenou in which respondents tasted fruit cocktails involving mango juice, blackcurrant syrup, and lemon juice
- 60 consumers were asked to taste different pairs of seven fruit cocktails
- Ruseckaite et al. obtained a prior distribution for parameter vector $\boldsymbol{\beta}$ in a special-cubic Scheffé model
- We used the same prior distribution to compute Bayesian D- and I-optimal designs using a coordinate-exchange algorithm

Cocktail preferences



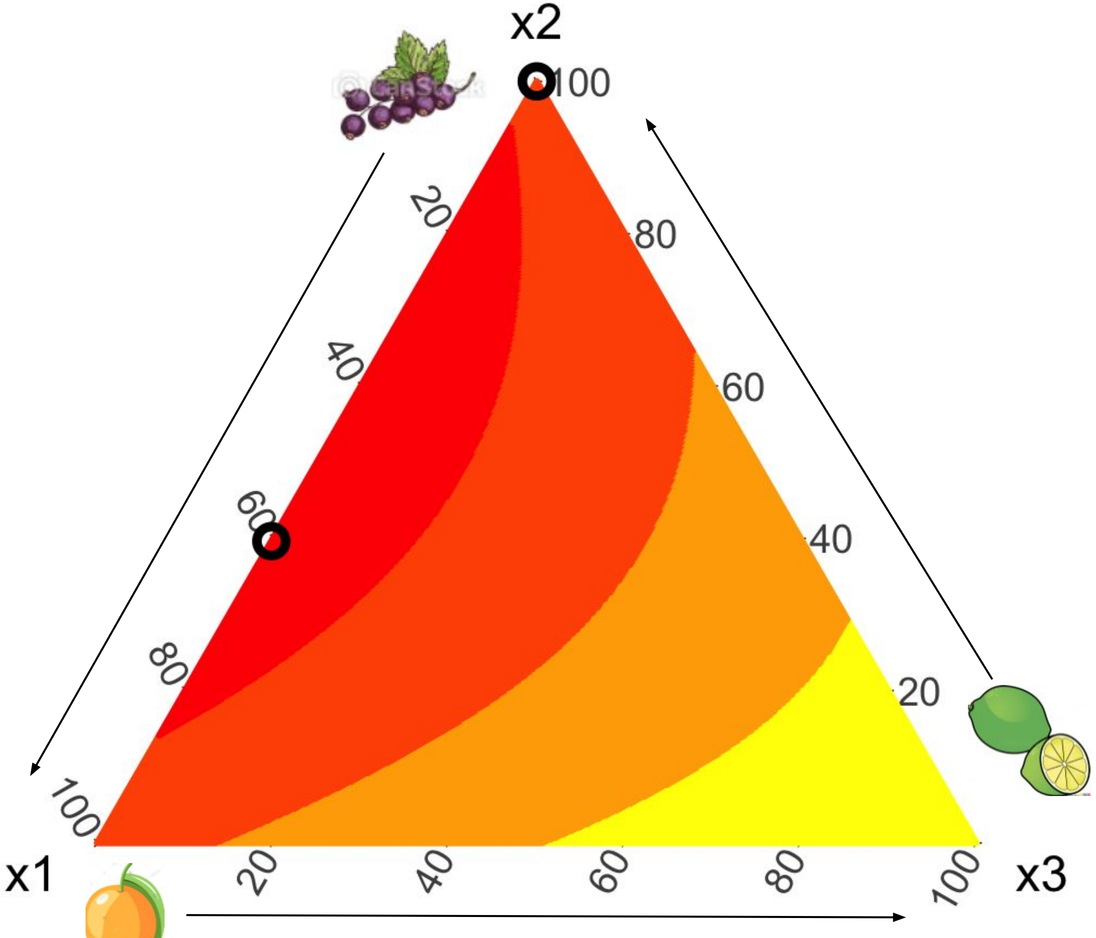
Bayesian D-optimal design



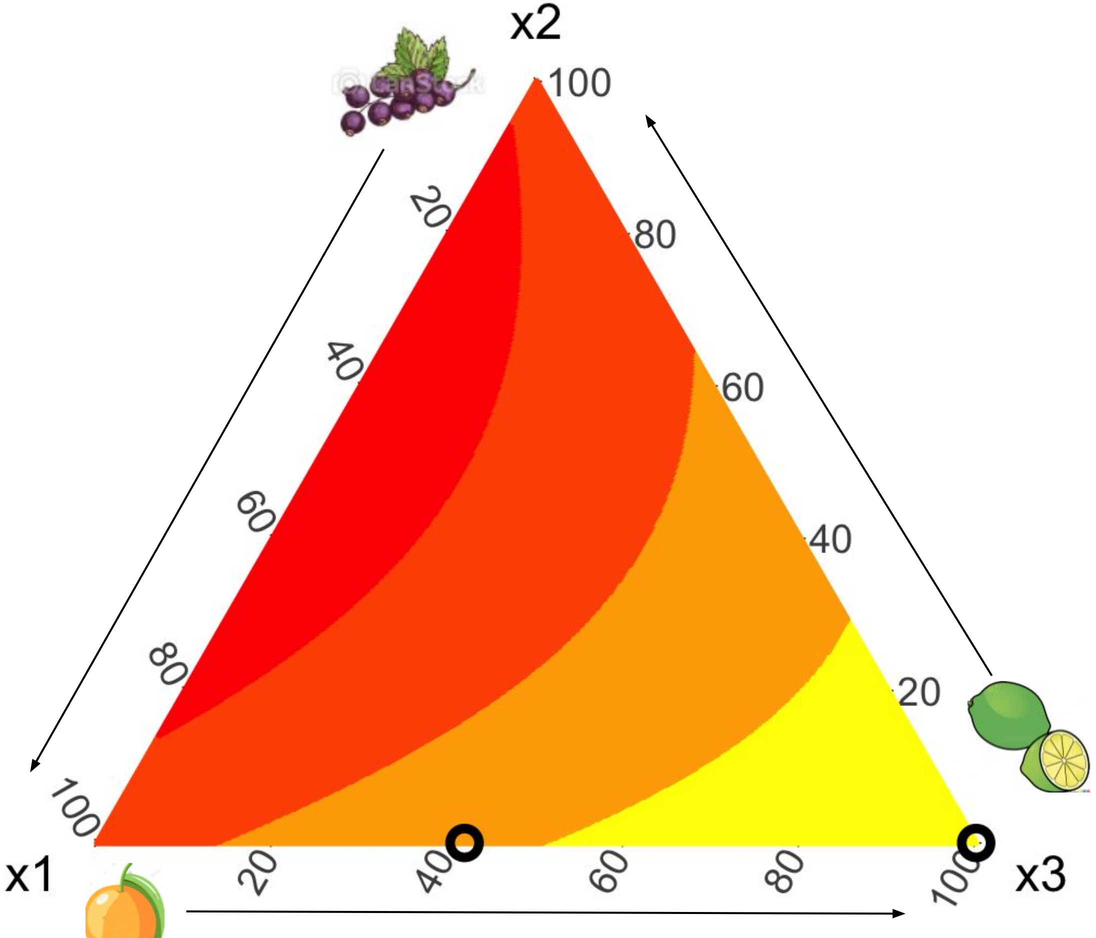
Bayesian I-optimal design

[0, 0.5625), [0.5625, 1.125), [1.125, 1.6875), [1.6875, 2.25)

Cocktail preferences



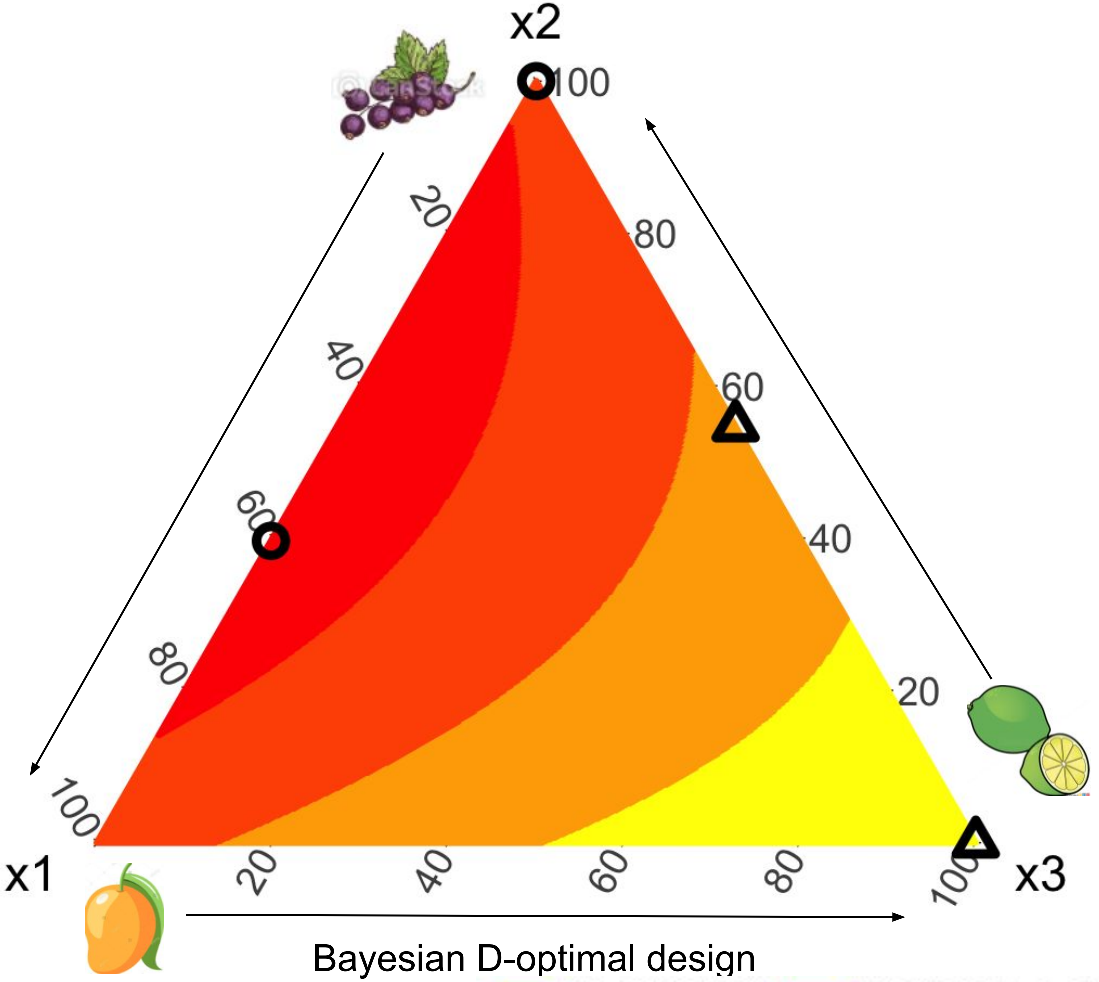
Bayesian D-optimal design



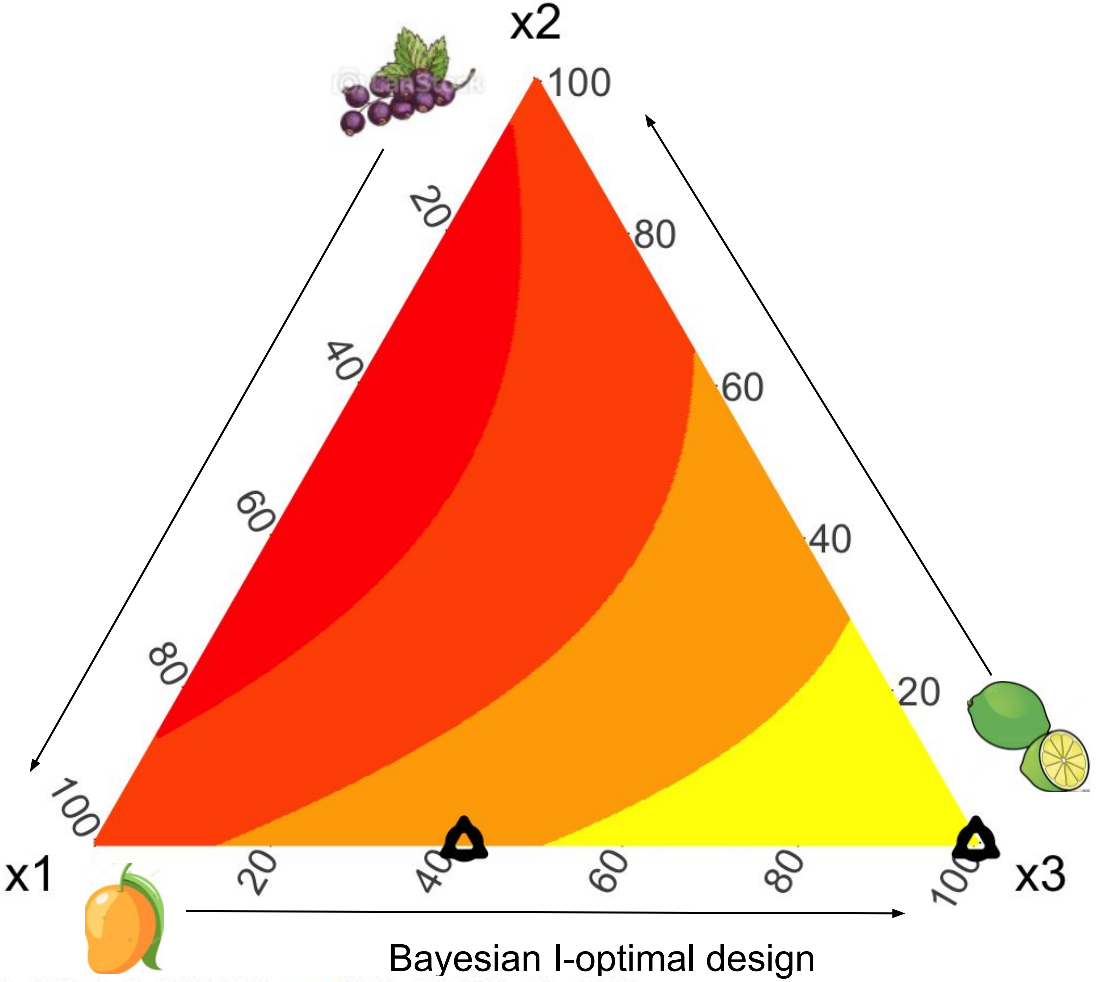
Bayesian I-optimal design

[0, 0.5625), [0.5625, 1.125), [1.125, 1.6875), [1.6875, 2.25)

Cocktail preferences



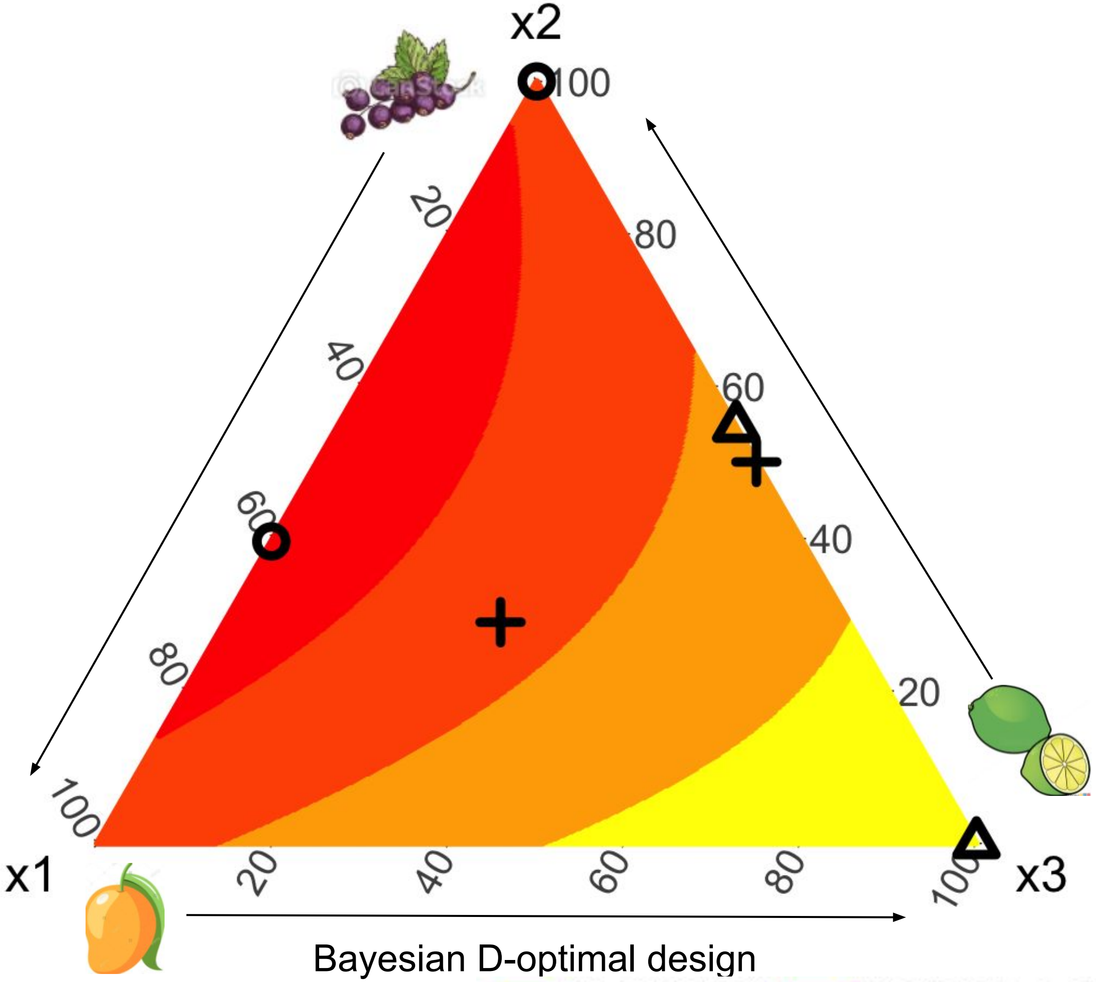
Bayesian D-optimal design

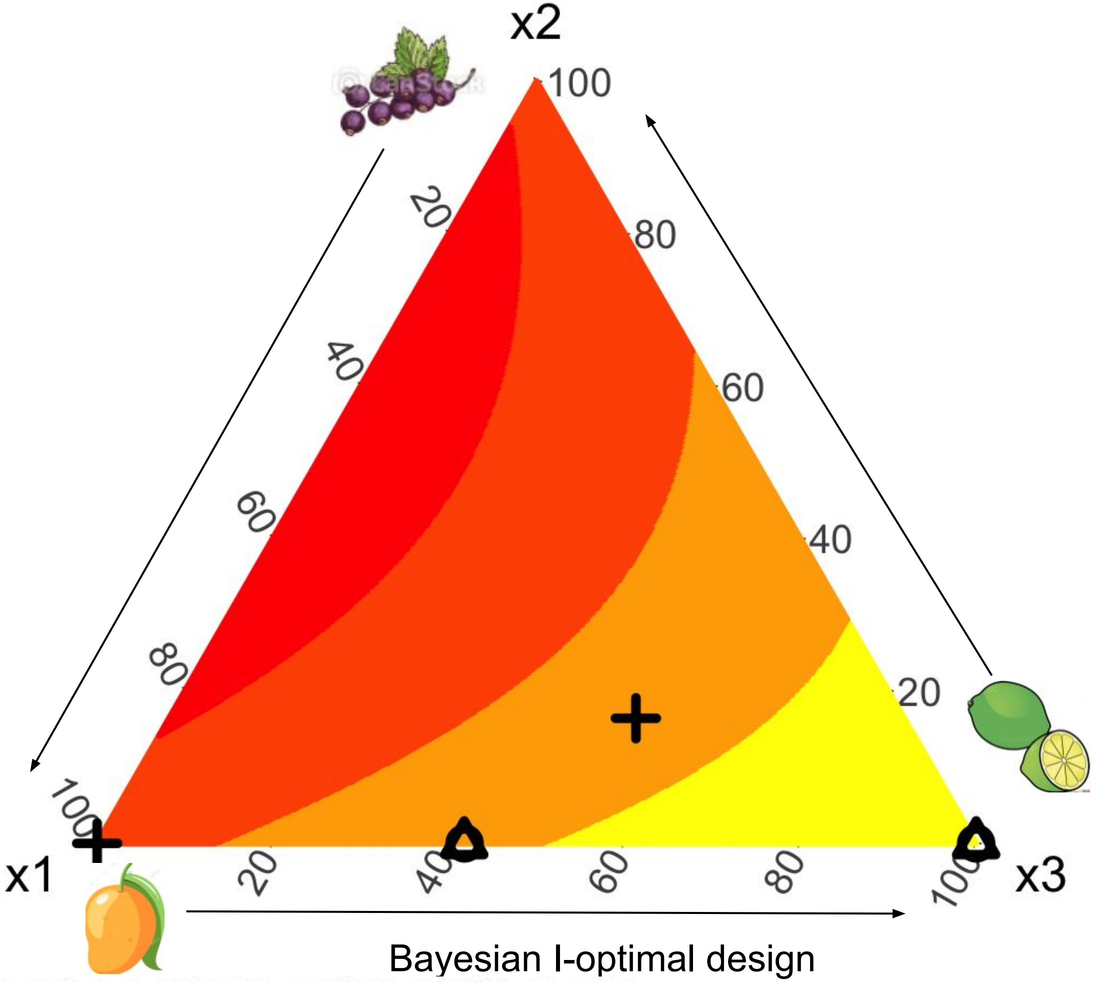


Bayesian I-optimal design

[0, 0.5625), [0.5625, 1.125), [1.125, 1.6875), [1.6875, 2.25)

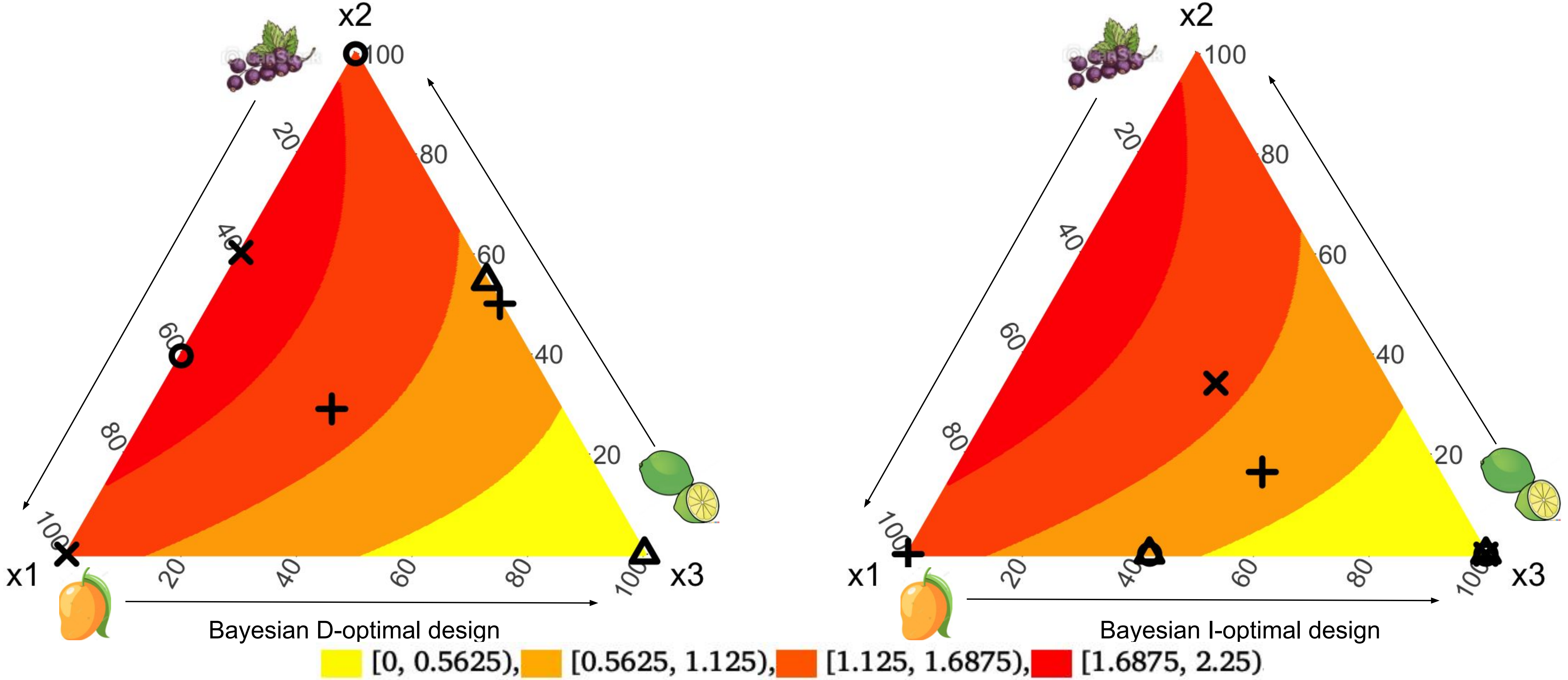
Cocktail preferences



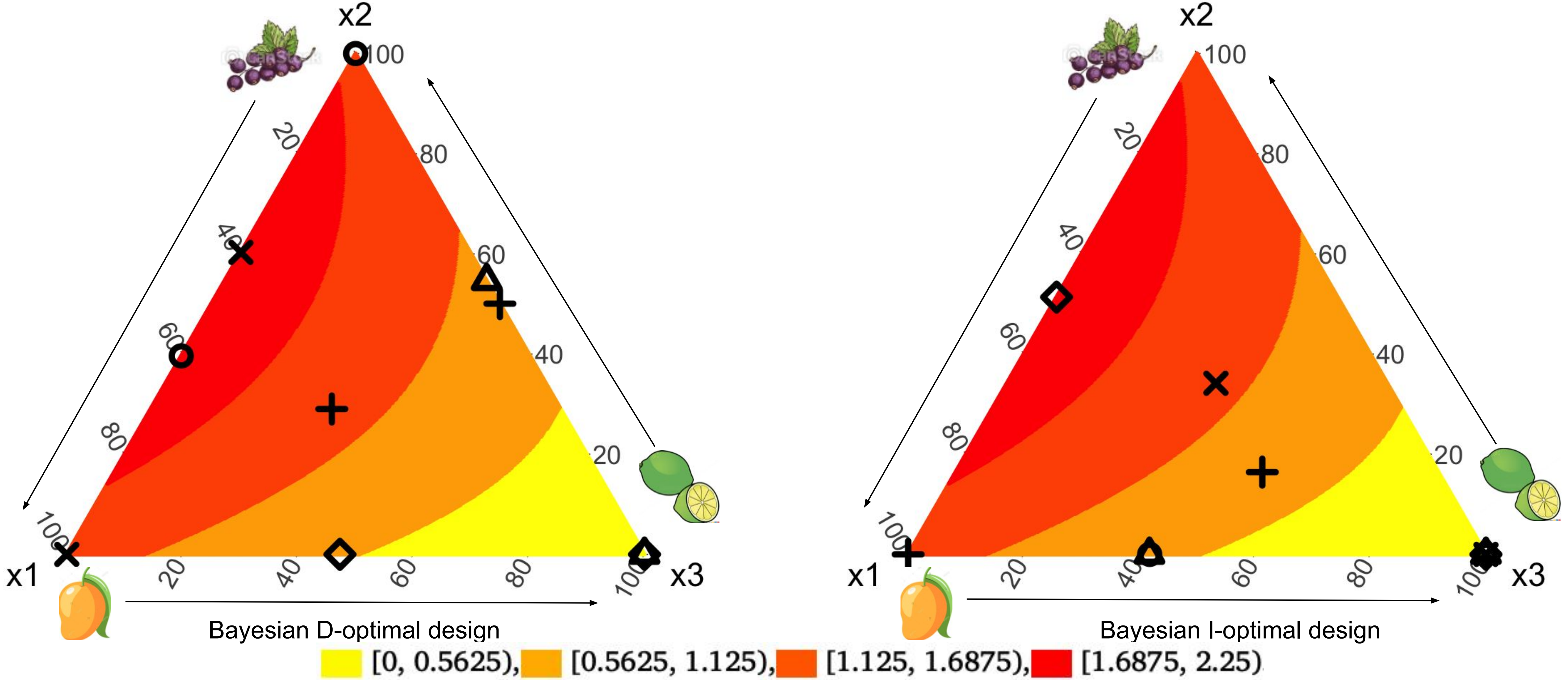


[0, 0.5625), [0.5625, 1.125), [1.125, 1.6875), [1.6875, 2.25)

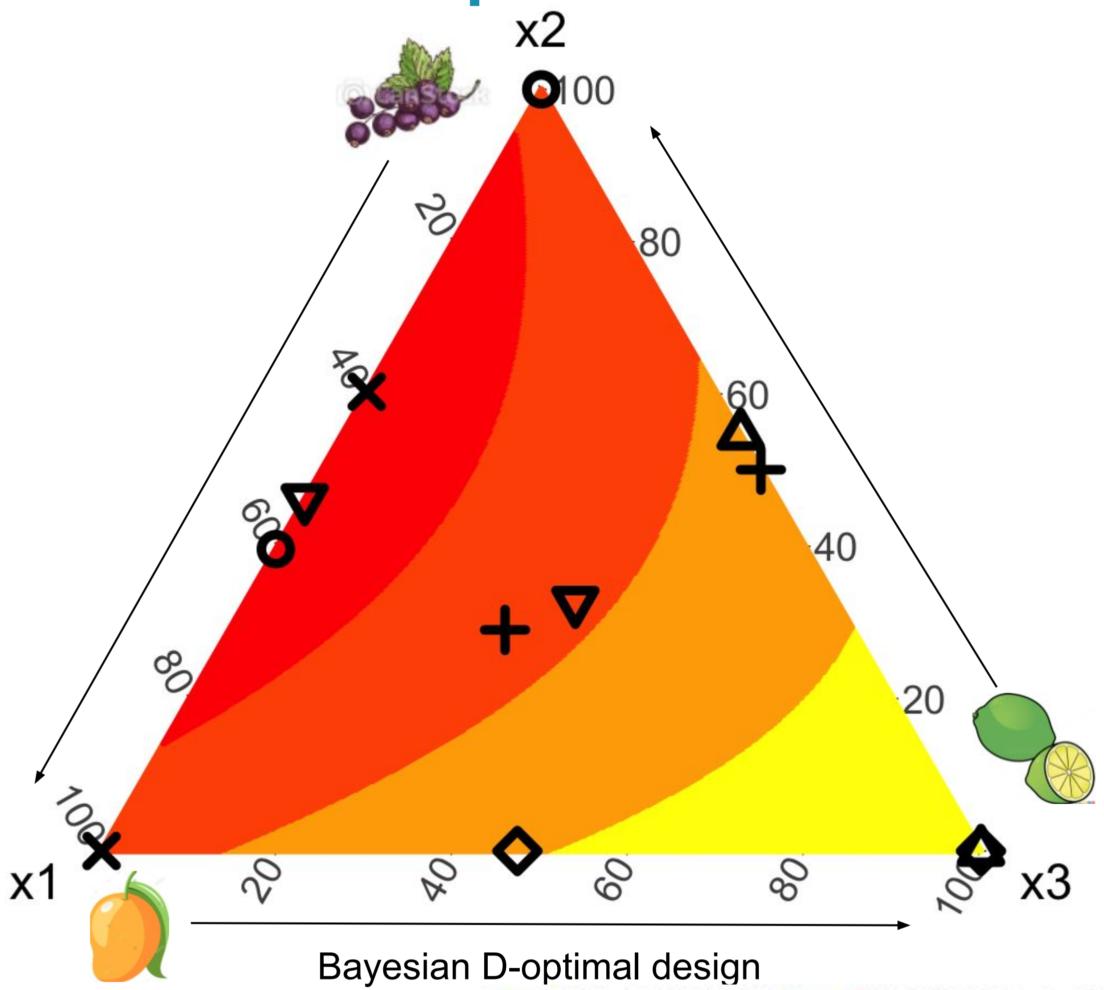
Cocktail preferences



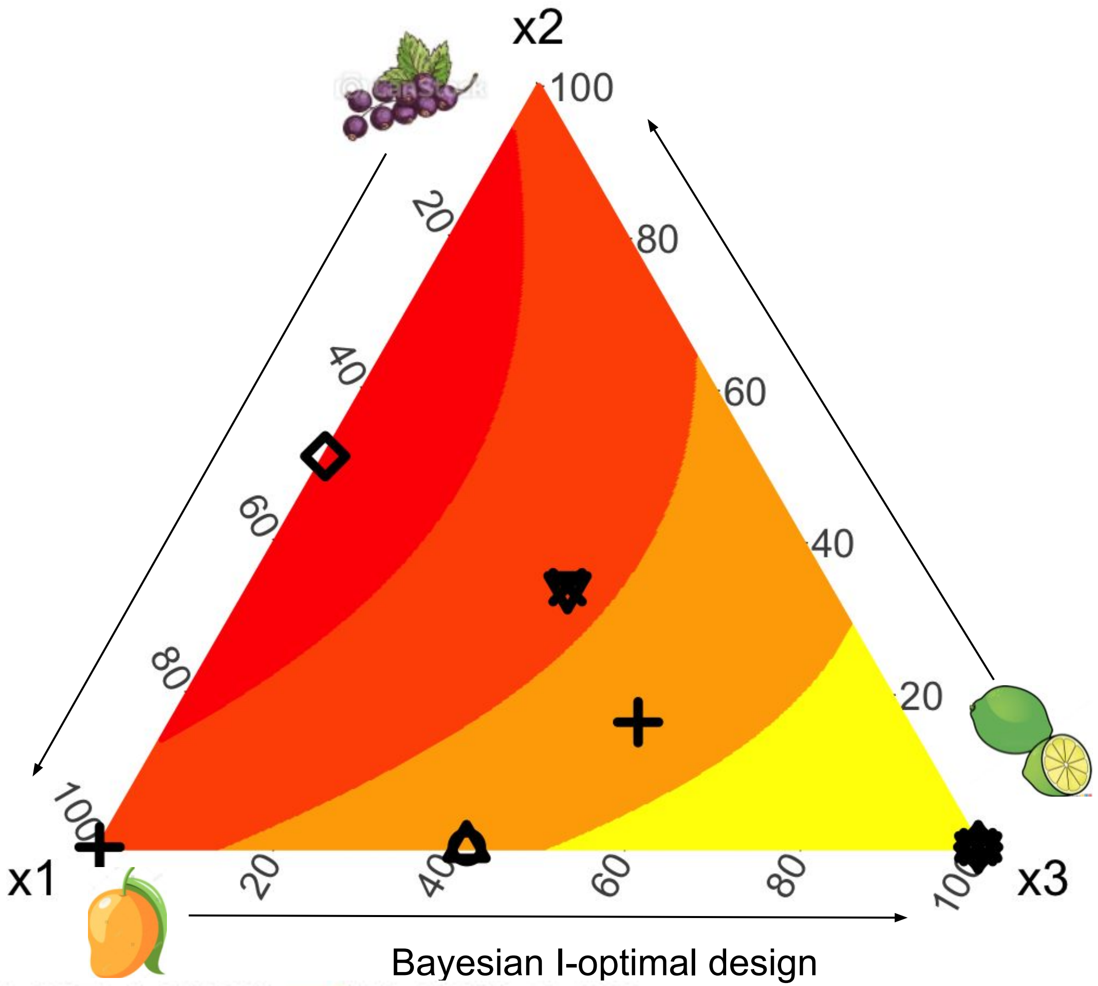
Cocktail preferences



Cocktail preferences



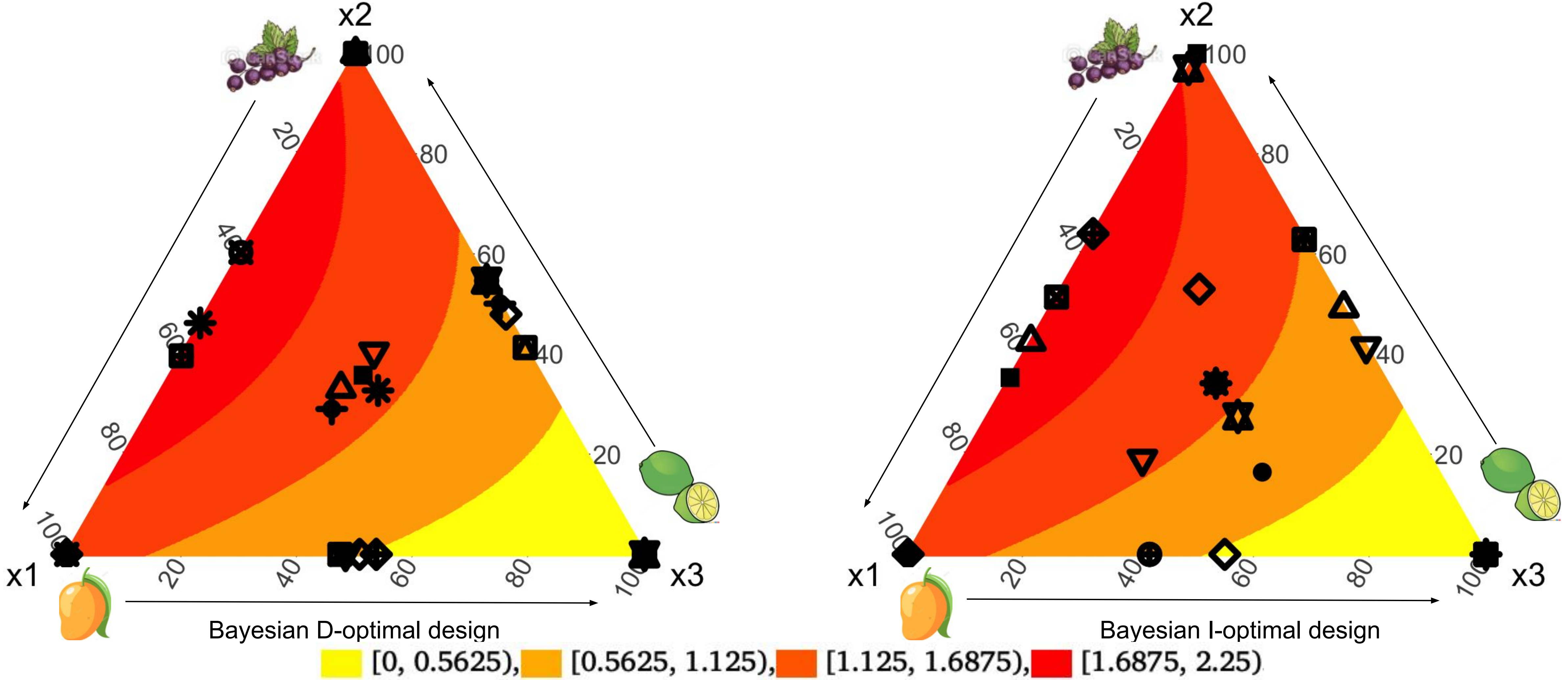
Bayesian D-optimal design



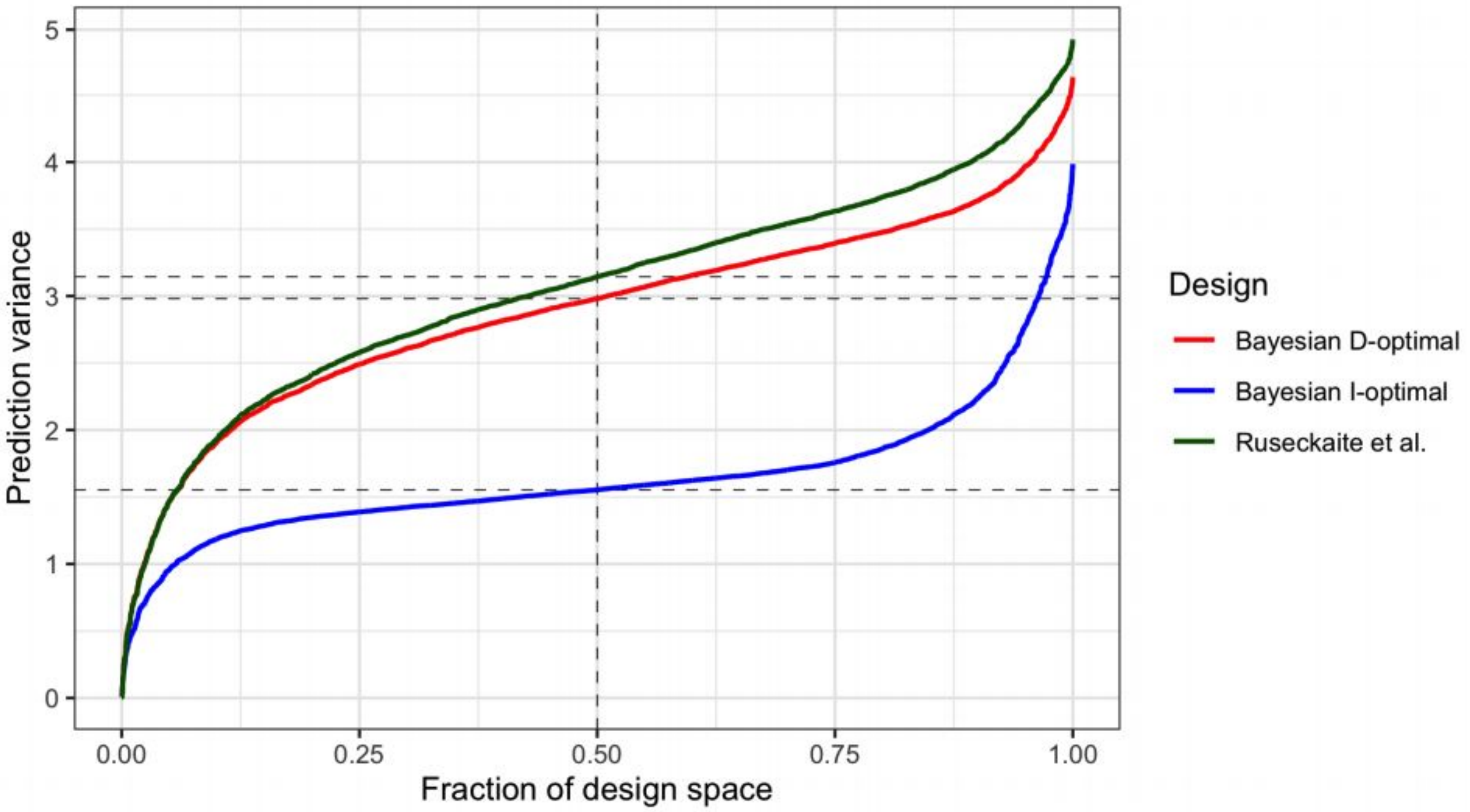
Bayesian I-optimal design

[0, 0.5625), [0.5625, 1.125), [1.125, 1.6875), [1.6875, 2.25)

Cocktail preferences



Cocktail preferences



Conclusions

Conclusions

Contributions

- Demonstrated that the I-optimal designs perform better than their D-optimal counterparts

Conclusions

Contributions

- Demonstrated that the I-optimal designs perform better than their D-optimal counterparts
- Created an accessible and open-source R package with our algorithms, including an efficient definition for I-optimal designs for choice experiments involving mixtures of ingredients

Conclusions

Contributions

- Demonstrated that the I-optimal designs perform better than their D-optimal counterparts
- Created an accessible and open-source R package with our algorithms, including an efficient definition for I-optimal designs for choice experiments involving mixtures of ingredients

Future work

- Add process variables

Conclusions

Contributions

- Demonstrated that the I-optimal designs perform better than their D-optimal counterparts
- Created an accessible and open-source R package with our algorithms, including an efficient definition for I-optimal designs for choice experiments involving mixtures of ingredients

Future work

- Add process variables
- Models that take into account possible presence of consumer heterogeneity

Conclusions

Contributions

- Demonstrated that the I-optimal designs perform better than their D-optimal counterparts
- Created an accessible and open-source R package with our algorithms, including an efficient definition for I-optimal designs for choice experiments involving mixtures of ingredients

Future work

- Add process variables
- Models that take into account possible presence of consumer heterogeneity
- Extend the work to other classes of models for data from mixture experiments

More information

- Becerra, Mario, and Peter Goos. *Bayesian I-optimal designs for choice experiments with mixtures.* Chemometrics and Intelligent Laboratory Systems 217 (2021): 104395. DOI: 10.1016/j.chemolab.2021.104395

- Mario Becerra's website (with links to paper, R package, and code to reproduce the paper): mariobecerra.github.io/

Thank you

Extra: Optimal design criteria

- D-optimal designs: low-variance estimators
- I-optimal designs: low-variance predictions
- Information matrix of multinomial logit model: $$\boldsymbol{I}(\boldsymbol{X},\boldsymbol{\beta}) = \sum_{s=1}^{S} \boldsymbol{X}_s^T(\boldsymbol{P}_s - \boldsymbol{p}_s\boldsymbol{p}_s^T)\boldsymbol{X}_s$$
- With

$$\boldsymbol{P}_s = \mathrm{diag}(\boldsymbol{p}_s)$$

$$\boldsymbol{p}_s = (p_{1s}, \dots, p_{Js})^T$$

$$\boldsymbol{X}_s^T = [\boldsymbol{f}(\boldsymbol{x}_{js})]_{j\in\{1,\dots,J\}}$$

$$\boldsymbol{X} = [\boldsymbol{X}_1, \dots, \boldsymbol{X}_S]$$

$$p_{js} = \frac{\exp\left[\boldsymbol{f}^T(\boldsymbol{x}_{js})\boldsymbol{\beta}\right]}{\sum_{t=1}^{J}\exp\left[\boldsymbol{f}^T(\boldsymbol{x}_{ts})\boldsymbol{\beta}\right]}$$

Extra: D-optimal designs

- D-optimality criterion

$$\mathcal{D} = \log\left(\det\left(\left[\boldsymbol{I}^{-1}(\boldsymbol{X},\boldsymbol{\beta})\right)\right]^{\frac{1}{r}}\right)$$

- Bayesian D-optimality criterion

$$\mathcal{D}_B = \log\left(\int_{\mathbb{R}^r}\left[\det\left(\boldsymbol{I}^{-1}(\boldsymbol{X},\boldsymbol{\beta})\right)\right]^{\frac{1}{r}}\pi(\boldsymbol{\beta})d\boldsymbol{\beta}\right)$$

- Numerical approximation to Bayesian D-optimality criterion

$$\mathcal{D}_B \approx \log\left(\frac{1}{R}\sum_{i=1}^{R}\left[\det\left(\boldsymbol{I}^{-1}(\boldsymbol{X},\boldsymbol{\beta}^{(i)})\right)\right]^{\frac{1}{r}}\right)$$

Extra: I-optimal designs

- I-optimality criterion

$$\mathcal{I} = \operatorname{tr}\left[\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta})\boldsymbol{W}_u\right]$$

- Bayesian I-optimality criterion

$$\mathcal{I}_B = \int_{\mathbb{R}^r} \operatorname{tr}\left[\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta})\boldsymbol{W}_u\right] \pi(\boldsymbol{\beta}) d\boldsymbol{\beta}$$

- Numerical approximation to Bayesian I-optimality criterion

$$\mathcal{I}_B \approx \frac{1}{R} \sum_{i=1}^{R} \operatorname{tr}\left[\boldsymbol{I}^{-1}(\boldsymbol{X}, \boldsymbol{\beta}^{(i)})\boldsymbol{W}_u\right]$$

$$\boldsymbol{W}_u = \int_{\chi} \boldsymbol{f}(\boldsymbol{x}_{js})\boldsymbol{f}^T(\boldsymbol{x}_{js})d\boldsymbol{x}_{js}$$

Extra: Model for choice data concerning mixtures

- The attributes of the alternatives in a choice experiment are the ingredients of a mixture
- Vector $\boldsymbol{x}_{js}$ contains the *q* ingredient proportions and that $\boldsymbol{f}(\boldsymbol{x}_{js})$ represents the model expansion of these proportions
- Most natural thing to do:

$$U_{js} = \sum_{i=1}^{q} \beta_i x_{ijs} + \sum_{i=1}^{q-1} \sum_{k=i+1}^{q} \beta_{ik} x_{ijs} x_{kjs} + \sum_{i=1}^{q-2} \sum_{k=i+1}^{q-1} \sum_{l=k+1}^{q} \beta_{ikl} x_{ijs} x_{kjs} x_{ljs} + \varepsilon_{js}$$

- Rewrite x_{qjs} as $1 - x_{1js} - \ldots - x_{q-1,js}$

$$U_{js} = \boldsymbol{f}^T(\boldsymbol{x}_{js})\boldsymbol{\beta} = \sum_{i=1}^{q-1} \beta_i^* x_{ijs} + \sum_{i=1}^{q-1} \sum_{k=i+1}^{q} \beta_{ik} x_{ijs} x_{kjs} + \sum_{i=1}^{q-2} \sum_{k=i+1}^{q-1} \sum_{l=k+1}^{q} \beta_{ikl} x_{ijs} x_{kjs} x_{ljs} + \varepsilon_{js}$$

- With

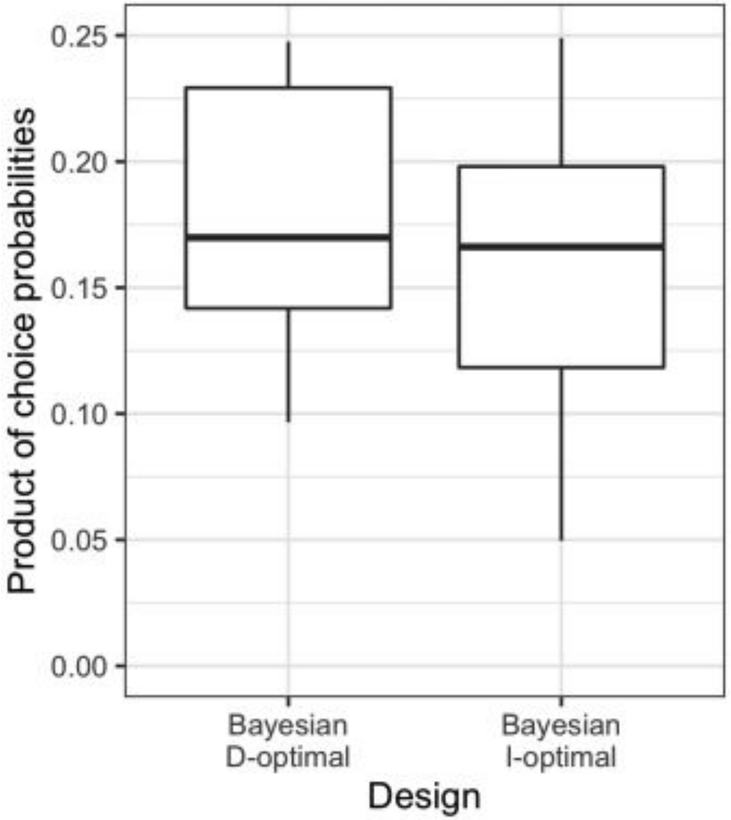
$$\boldsymbol{f}(\boldsymbol{x}_{js}) = (x_{1js}, x_{2js}, \ldots, x_{q-1,js}, x_{1js}x_{2js}, \ldots, x_{q-1,js}x_{qjs}, x_{1js}x_{2js}x_{3js}, \ldots, x_{q-2,js}x_{q-1,js}x_{qjs})^T$$

$$\beta_i^* = \beta_i - \beta_q \text{ for } i \in \{1, \ldots, q-1\}$$

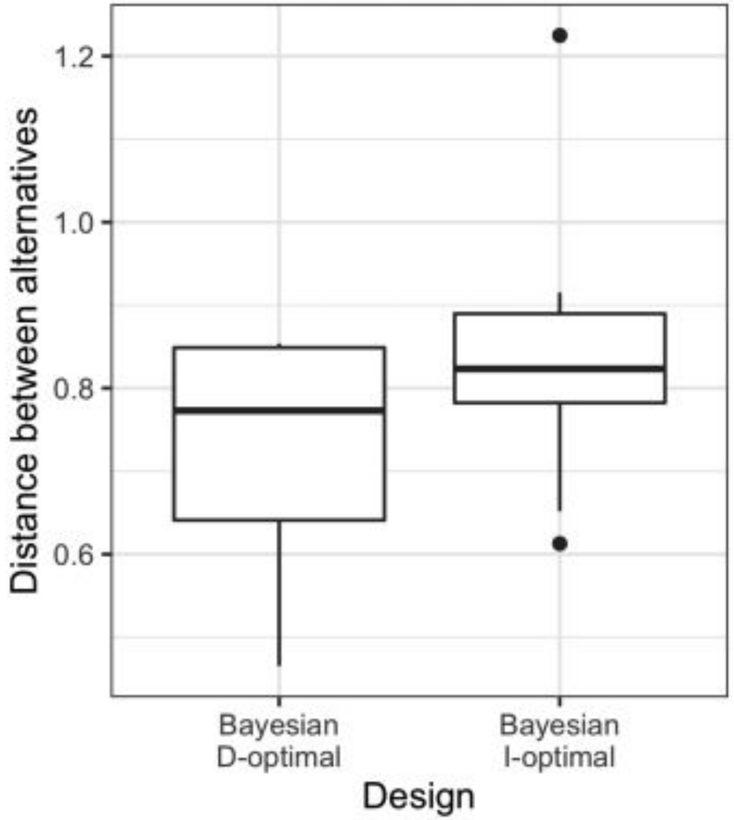
$$\boldsymbol{x}_{js} = (x_{1js}, x_{2js}, \ldots, x_{qjs})^T$$

$$\boldsymbol{\beta} = (\beta_1^*, \beta_2^*, \ldots, \beta_{q-1}^*, \beta_{1,2}, \ldots, \beta_{q-1,q}, \beta_{123}, \ldots, \beta_{q-2,q-1,q})^T$$

Extra results: cocktail preferences

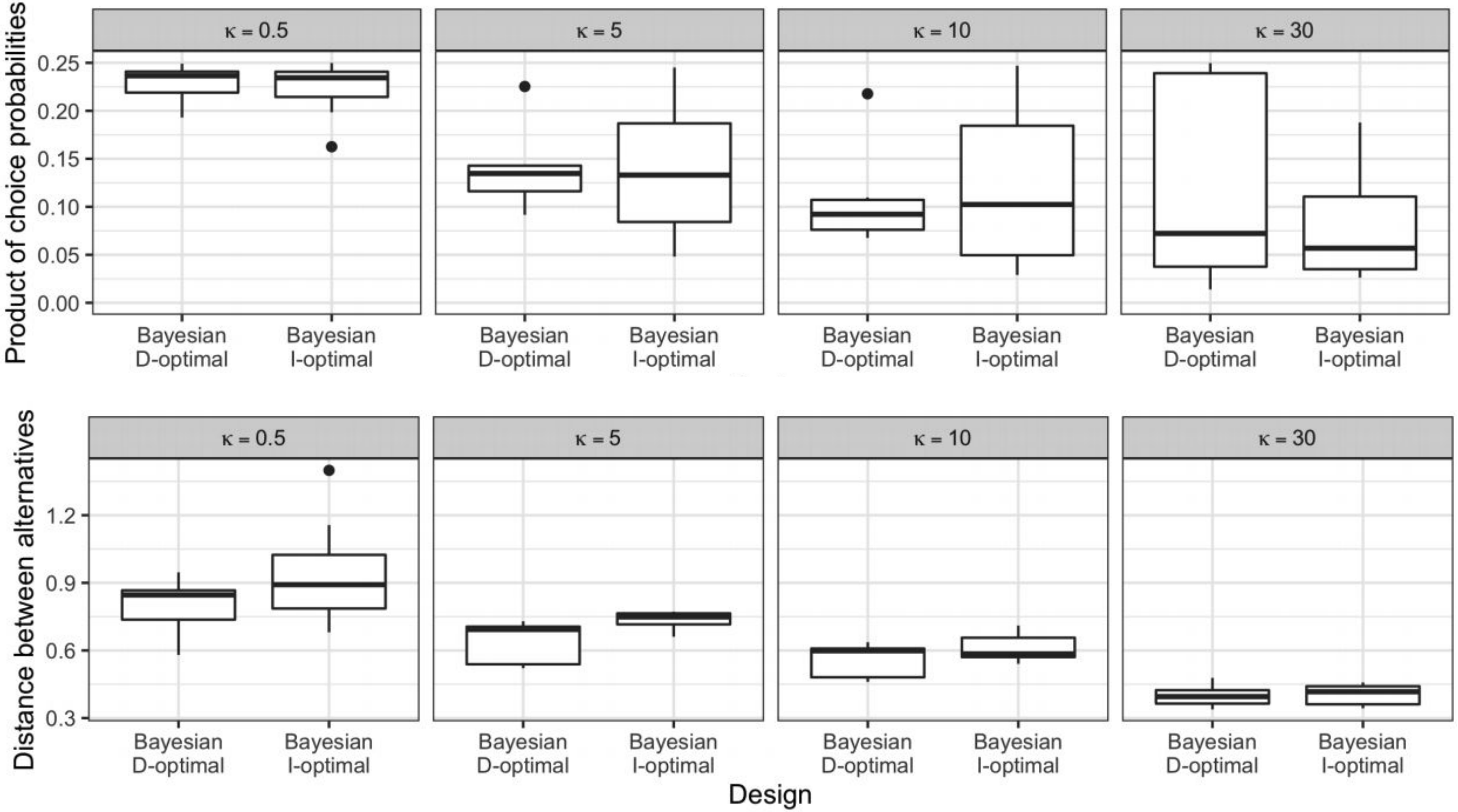


(a) Utility balance



(b) Euclidean distances

Extra results: artificial sweetener experiment

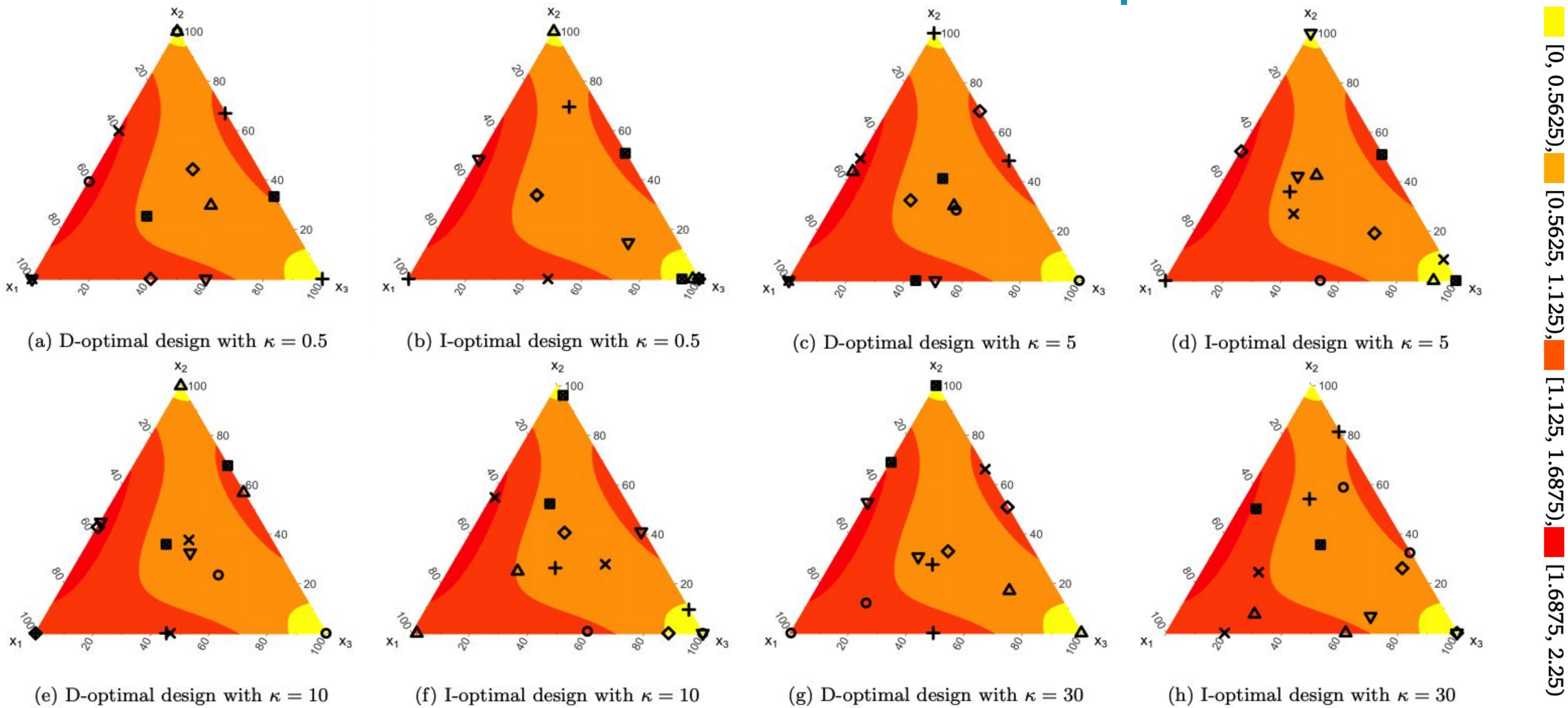


Extra results: artificial sweetener experiment

- Varying levels of uncertainty in parameter vector parameter. Values κ = 0.5, 5, 10 and 30.

$$\mathbf{\Sigma}_0' = \begin{pmatrix} 2\kappa & \kappa & 0 & 0 & 0 & 0 \\ \kappa & 2\kappa & 0 & 0 & 0 & 0 \\ 0 & 0 & \kappa & 0 & 0 & 0 \\ 0 & 0 & 0 & \kappa & 0 & 0 \\ 0 & 0 & 0 & 0 & \kappa & 0 \\ 0 & 0 & 0 & 0 & 0 & \kappa \end{pmatrix}$$

Extra results: artificial sweetener experiment



(a) D-optimal design with $\kappa = 0.5$

(b) I-optimal design with $\kappa = 0.5$

(c) D-optimal design with $\kappa = 5$

(d) I-optimal design with $\kappa = 5$

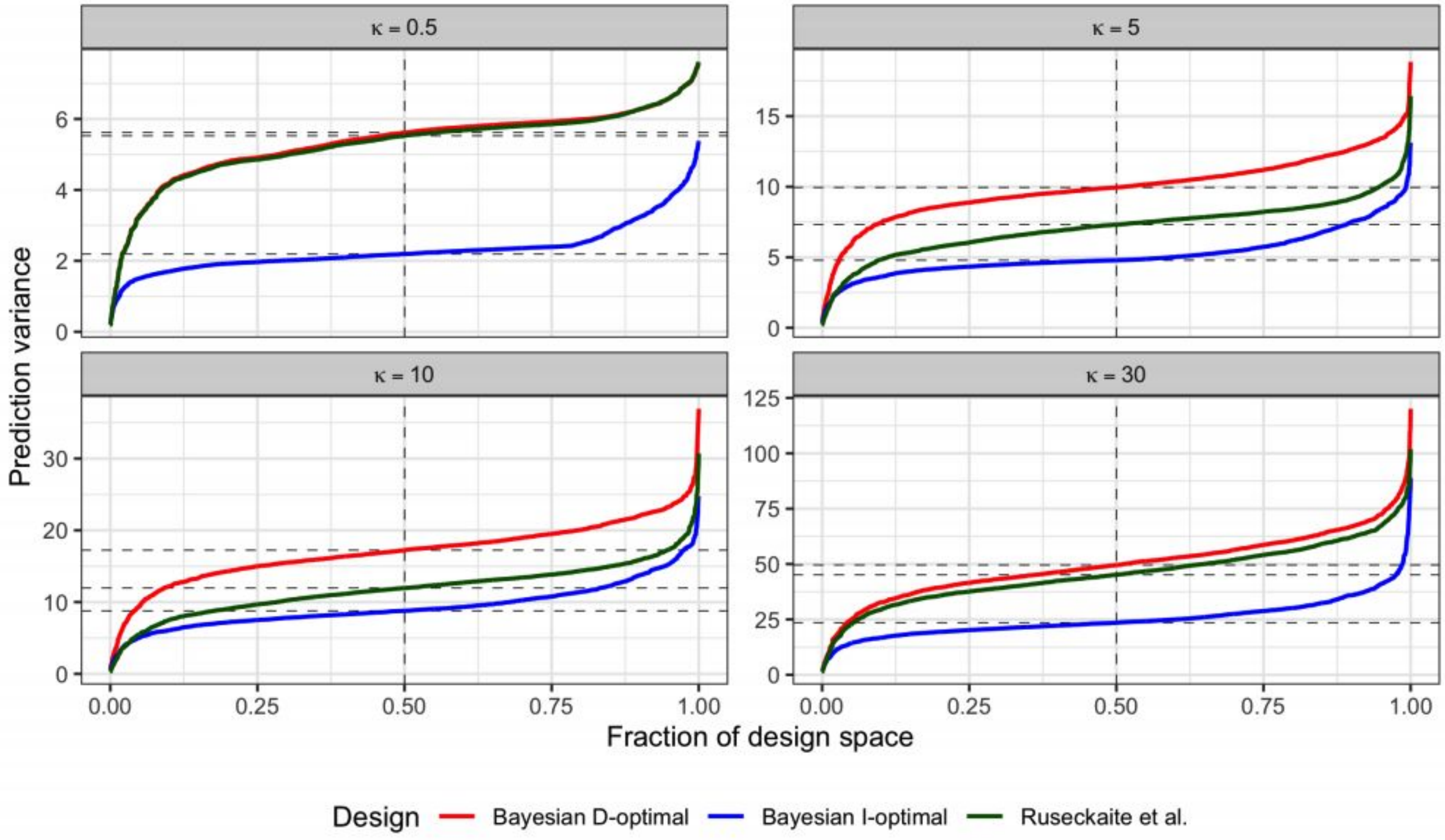
(e) D-optimal design with $\kappa = 10$

(f) I-optimal design with $\kappa = 10$

(g) D-optimal design with $\kappa = 30$

(h) I-optimal design with $\kappa = 30$

Extra results: artificial sweetener experiment



Artificial sweetener experiment

- We revisited a three-ingredient mixture experiment about artificial sweetener in a sports drink, originally by Cornell and reanalyzed by Ruseckaite et al.

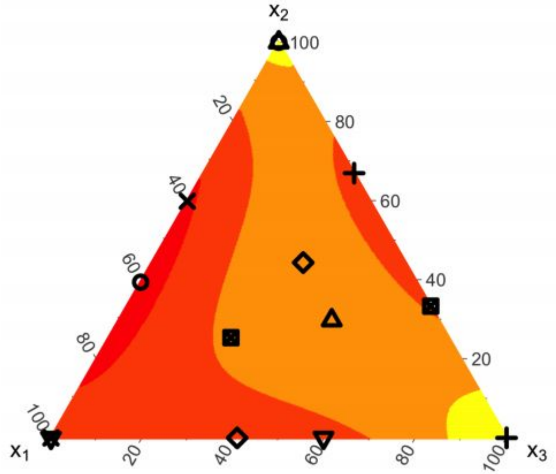
Artificial sweetener experiment

- We revisited a three-ingredient mixture experiment about artificial sweetener in a sports drink, originally by Cornell and reanalyzed by Ruseckaite et al.
- Original response: intensity of sweetness aftertaste

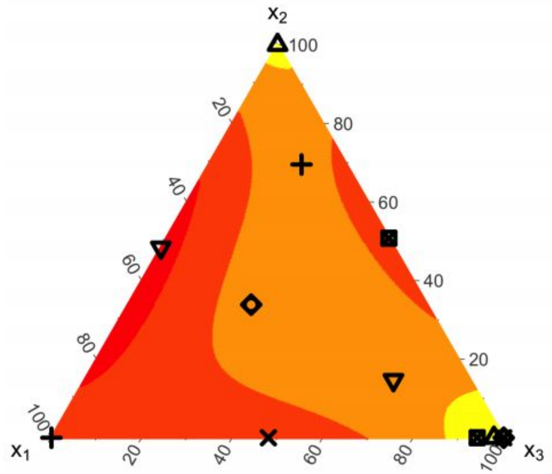
Artificial sweetener experiment

- We revisited a three-ingredient mixture experiment about artificial sweetener in a sports drink, originally by Cornell and reanalyzed by Ruseckaite et al.
- Original response: intensity of sweetness aftertaste
- Ruseckaite et al. obtained a prior distribution for parameter vector $\boldsymbol{\beta}$ in a special-cubic Scheffé model with varying levels of uncertainty

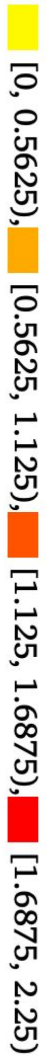
Artificial sweetener experiment



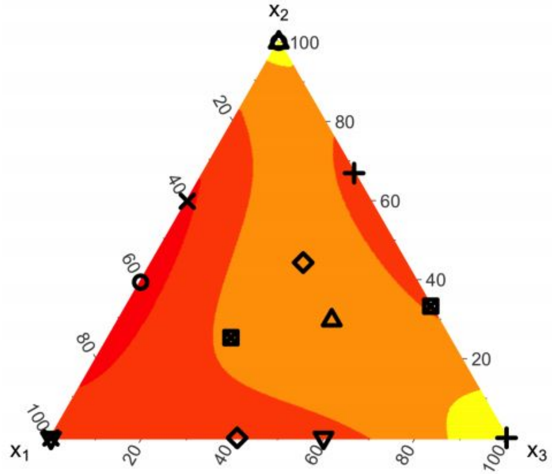
Bayesian D-optimal design with low uncertainty



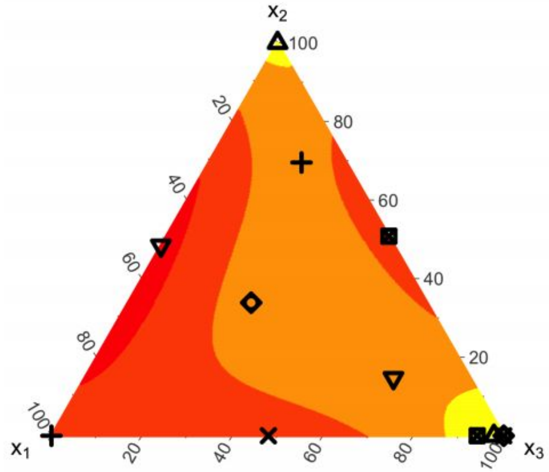
Bayesian D-optimal design with high uncertainty



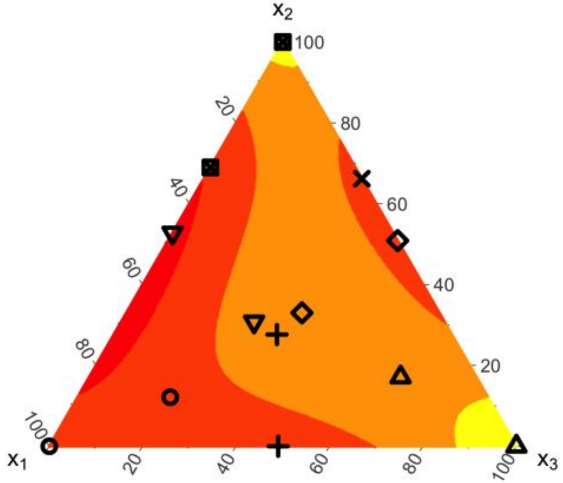
Artificial sweetener experiment



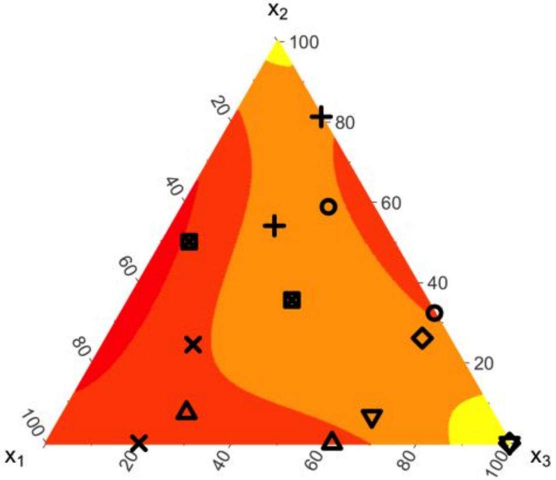
Bayesian D-optimal design with low uncertainty



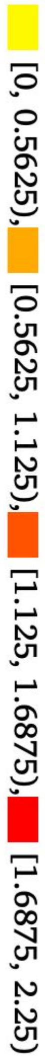
Bayesian D-optimal design with high uncertainty



Bayesian I-optimal design with low uncertainty



Bayesian I-optimal design with high uncertainty



Artificial sweetener experiment