Florence Bockting
Computational Statistics
TU Dortmund University

PREDICTIVE PRIOR ELICITATION

State of the Art & Current Challenges

Lund, 26.08.2025

BAYESIAN INFERENCE

Setting the starting point

data distribution

$$p(\theta \mid y) \propto p(y, \theta) = p(y \mid \theta)\, p(\theta \mid \lambda)$$

data model

prior distribution

BAYESIAN INFERENCE

Setting the starting point

prior specification

$$p(\theta \mid y) \propto p(y,\theta) = p(y \mid \theta)\,\boxed{p(\theta \mid \lambda)}$$

Non-informative, diffuse priors

- **maximize** the use of **information** derived from the **data** distribution

BAYESIAN INFERENCE

Setting the starting point

prior specification

$$p(\theta \mid y) \propto p(y, \theta) = p(y \mid \theta)\, \boxed{p(\theta \mid \lambda)}$$

Non-informative, diffuse priors

- **maximize** the use of **information** derived from the **data** distribution

Informative priors

- substantial **problem-specific knowledge**, ideally capturing all relevant information available before observing the data

BAYESIAN INFERENCE

Setting the starting point

prior specification

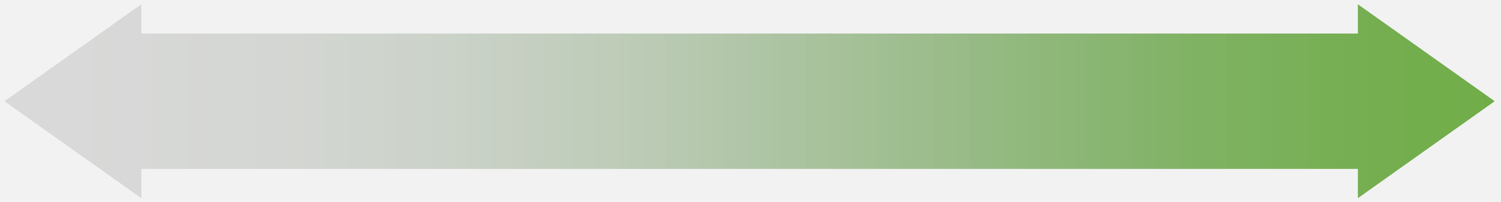
$$p(\theta \mid y) \propto p(y,\theta) = p(y \mid \theta)\,\boxed{p(\theta \mid \lambda)}$$

Non-informative, diffuse priors

- **maximize** the use of **information** derived from the **data** distribution

Informative priors

- substantial **problem-specific knowledge**, ideally capturing all relevant information available before observing the data



Weakly informative priors

- **general domain knowledge** applicable across a broad class of problems

BAYESIAN INFERENCE

Setting the starting point

prior specification

$$p(\theta \mid y) \propto p(y,\theta) = p(y \mid \theta)\,\boxed{p(\theta \mid \lambda)}$$

meta-
analyses

historical
data

domain
experts

previous
studies

Informative priors

- substantial **problem-specific knowledge**, ideally capturing all relevant information available before observing the data

BAYESIAN INFERENCE

Setting the starting point

Expert Prior Elicitation (EPE)

$$p(\theta \mid y) \propto p(y,\theta) = p(y \mid \theta)\boxed{p(\theta \mid \lambda)}$$

meta-analyses

historical data

domain experts

previous studies

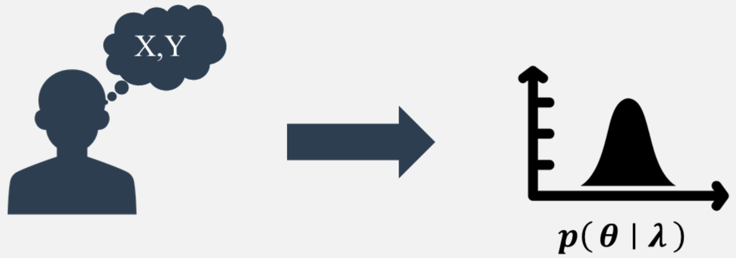
Informative priors

- substantial **problem-specific knowledge**, ideally capturing all relevant information available before observing the data

EXPERT PRIOR ELICITATION

***What it is** and why we need it*

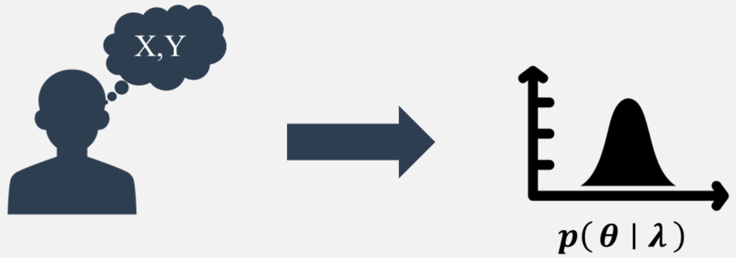
- structured process for **translating** an **individual's knowledge** and beliefs about one or more uncertain quantities **into a (joint) probability distribution**



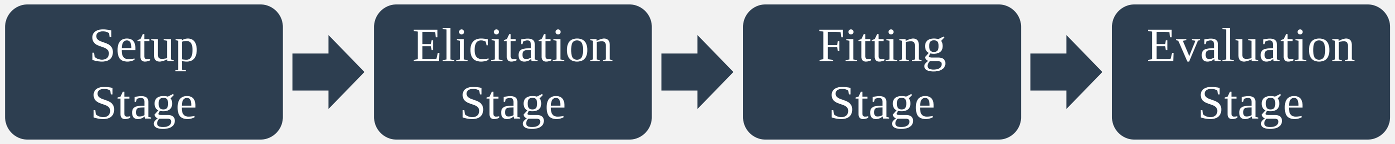
EXPERT PRIOR ELICITATION

What it is and why we need it

- structured process for **translating** an **individual's knowledge** and beliefs about one or more uncertain quantities **into a (joint) probability distribution**



- stages in an **expert prior elicitation process** according to Garthwaite et al. (2005)



EXPERT PRIOR ELICITATION

What it is and why we need it

- ✓ selection of experts
- ✓ data model (data distribution + prior)
- ✓ target quantities („What“)
- ✓ elicitation techniques („How“)

Setup Stage → Elicitation Stage → Fitting Stage → Evaluation Stage

EXPERT PRIOR ELICITATION

What it is and why we need it

- ✓ selection of experts
- ✓ data model (data distribution + prior)
- ✓ target quantities („What“)
- ✓ elicitation techniques („How“)

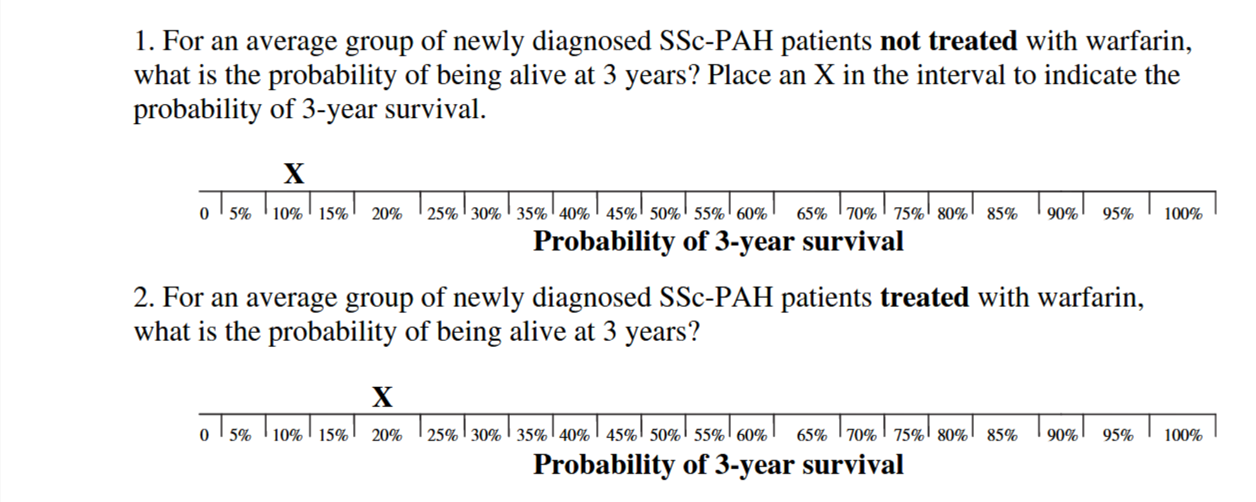
Setup Stage → Elicitation Stage → Fitting Stage → Evaluation Stage

- ✓ elicitation protocol
- ✓ training of experts
- ✓ questionnaire, interview
- ✓ on-site, online

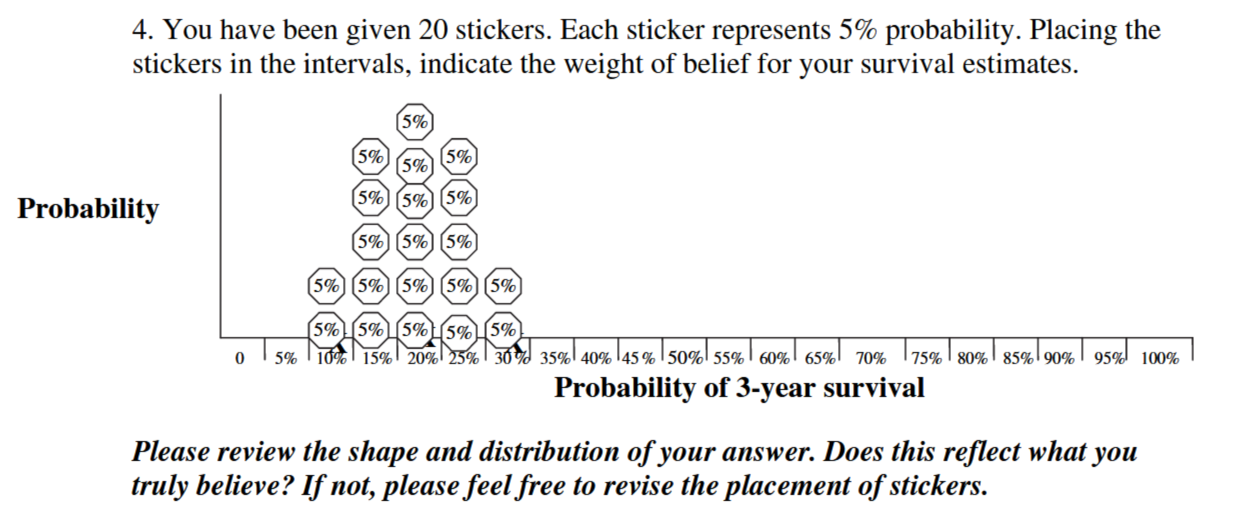
percentiles, histogram, moments, etc.

EXPERT PRIOR ELICITATION

What it is and why we need it


1. For an average group of newly diagnosed SSc-PAH patients **not treated** with warfarin, what is the probability of being alive at 3 years? Place an X in the interval to indicate the probability of 3-year survival.

2. For an average group of newly diagnosed SSc-PAH patients **treated** with warfarin, what is the probability of being alive at 3 years?


4. You have been given 20 stickers. Each sticker represents 5% probability. Placing the stickers in the intervals, indicate the weight of belief for your survival estimates.

Please review the shape and distribution of your answer. Does this reflect what you truly believe? If not, please feel free to revise the placement of stickers.

taken from Johnson et al. (2010)

EXPERT PRIOR ELICITATION

What it is and why we need it

- ✓ selection of experts
- ✓ data model (data distribution + prior)
- ✓ target quantities („What“)
- ✓ elicitation techniques („How“)

Setup Stage → Elicitation Stage → Fitting Stage → Evaluation Stage

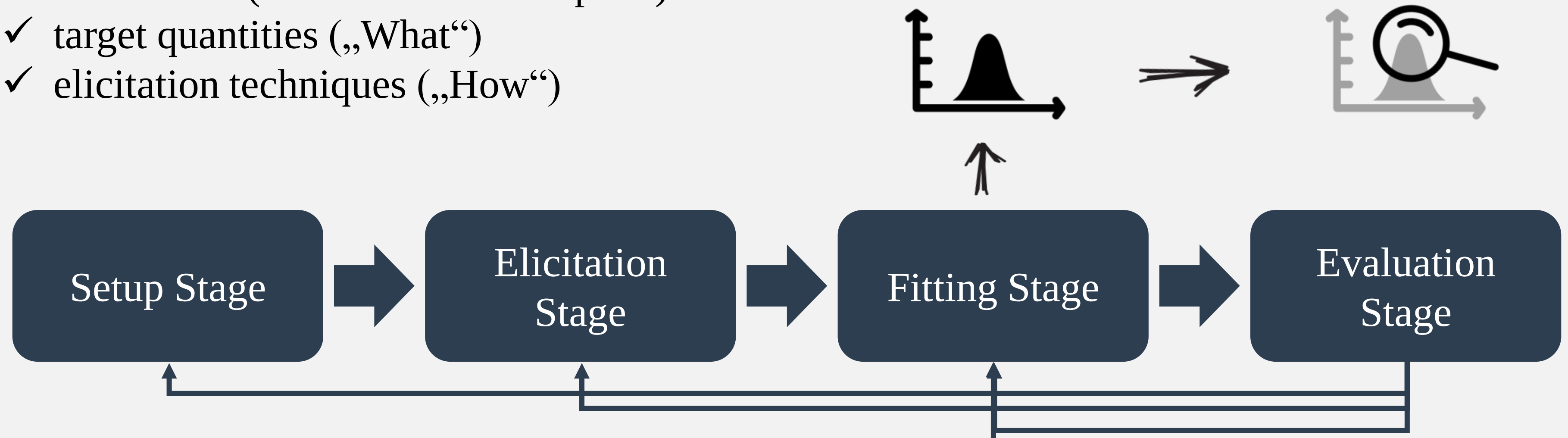
- ✓ elicitation protocol
- ✓ training of experts
- ✓ questionnaire, interview
- ✓ on-site, online

percentiles, histogram, moments, etc.

EXPERT PRIOR ELICITATION

What it is and why we need it

- ✓ selection of experts
- ✓ data model (data distribution + prior)
- ✓ target quantities („What“)
- ✓ elicitation techniques („How“)



- ✓ elicitation protocol
- ✓ training of experts
- ✓ questionnaire, interview
- ✓ on-site, online

percentiles, histogram, moments, etc.

EXPERT PRIOR ELICITATION

What it is and why we need it

- ✓ selection of experts
- ✓ data model (data distribution + prior)
- ✓ target quantities („What“)
- ✓ elicitation techniques („How“)

Setup Stage → Elicitation Stage → Fitting Stage → Evaluation Stage

- ✓ elicitation protocol
- ✓ training of experts
- ✓ questionnaire, interview
- ✓ on-site, online

percentiles, histogram, moments, etc.

EPE methods

EPE METHODS

Classification of EPE methods

EPE methods

Setup Stage → Elicitation Stage → Fitting Stage → Evaluation Stage

configurations in setup stage

characterization of an EPE method

✓ data model (data distribution + prior)

model-specific vs. model-agnostic
parametric vs. non-parametric
independent vs. joint

EPE METHODS

Classification of EPE methods

EPE methods

Setup Stage → Elicitation Stage → Fitting Stage → Evaluation Stage

configurations in setup stage

characterization of an EPE method

- ✓ data model (data distribution + prior)
- ✓ target quantities („What“)

model-specific vs. model-agnostic
parametric vs. non-parametric
independent vs. joint

structural vs. predictive vs. hybrid

EPE METHODS

Classification of EPE methods

EPE methods

Setup Stage → Elicitation Stage → Fitting Stage → Evaluation Stage

configurations in setup stage

characterization of an EPE method

- ✓ data model (data distribution + prior)
- ✓ target quantities („What“)
- ✓ elicitation techniques („How“)

model-specific vs. model-agnostic
parametric vs. non-parametric
independent vs. joint

structural vs. predictive vs. hybrid

fixed vs. flexible

EPE METHODS

Classification of EPE methods

EPE methods

Setup Stage → Elicitation Stage → Fitting Stage → Evaluation Stage

configurations in setup stage

characterization of an EPE method

- ✓ data model (data distribution + prior)
- ✓ target quantities („What“)
- ✓ elicitation techniques („How“)

model-specific vs. model-agnostic
parametric vs. non-parametric
independent vs. joint

structural vs. **predictive** vs. hybrid

fixed vs. flexible

PREDICTIVE PRIOR ELICITATION

The promise and the difficulty

✓ most predictive EPE methods are *model-specific*

Author	Model	Prior	Hypercube D3
Kadane et al. (1980)	NLR	NCP	O
Oman (1985)	NLR	NCP	O
Garthwaite and Dickey (1988)	NLR	NCP	O
Ibrahim and Laud (1994)	NLR	NCP	O
Bedrick et al. (1996)	GLM	CMP	O
Chen and Ibrahim (2003)	GLM	CP	O
Denham and Mengersen (2007)	GLM	NCP	H
Elfadaly and Garthwaite (2011)	NLR	NCP	O
Garthwaite et al. (2013)	PW-GLM	NCP	O
Elfadaly and Garthwaite (2015)	Gam-GLM	NCP/log-normal	O
Garthwaite and Dickey (1992)	NLR	mixture-NCP	H
Laud and Ibrahim (1995)	NLR	NCP	O
Chen et al. (1999)	LR	custom	O
Leamer (1992)	NLR	NCP	P
Hosack et al. (2017)	GLM	NCP	O
Carlin et al. (1992)	RE-LR	custom	P
Al-Hamzawi et al. (2011)	RE-BQR	power prior	P
Garthwaite and Al-Awadhi (2006)	PW-LR	NCP	O
Kadane et al. (1996)	AR	PW-CP	O
Garthwaite and Dickey (1991)	NLR	NCP	O
Chaloner et al. (1993)	PHR	adjusted-NCP	O
Ibrahim et al. (1999)	PHR	semi-parametric	O
Soare et al. (2016)	NLR	delta	P
Micallef et al. (2017)	NLR	half-normal	P

taken from
Mikkola et al. (2024)

PREDICTIVE PRIOR ELICITATION

The promise and the difficulty

✓ most predictive EPE methods are *model-specific*

✓ only recently several *model-agnostic* methods have been proposed, e.g.
- Hartmann and Agiashvili [2020]. Manderson and Goudie [2024], da Silva et al. [2023], Bockting et al. [2024]
- However, focus on simple, parametric prior distributions

✓ further work in predictive EPE focussing on more complex prior distributions include
- Gaussian processes [Oakley and O'Hagan, 2007]
- Quantile-parametrized distributions [Perepolkin et al., 2024]
- Normalizing Flows with preferential judgments [Mikkola et al., 2024]
- Extension of simulation-based EPE method [Bockting et al., 2025]

Why is the development of predictive EPE methods so challenging?

PREDICTIVE PRIOR ELICITATION

The promise and the difficulty

- ✓ selection of experts
- ✓ data model (data distribution + prior)
- ✓ ***target quantities („What“)***
- ✓ ***elicitation techniques („How“)***

✓ Interpretability
✓ Informativeness

Setup Stage → Elicitation Stage → Fitting Stage → Evaluation Stage

PREDICTIVE PRIOR ELICITATION

The promise and the difficulty

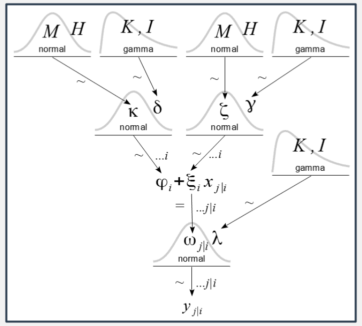
- ✓ selection of experts
- ✓ data model (data distribution + prior)
- ✓ ***target quantities („What“)***
- ✓ ***elicitation techniques („How“)***

✓ Interpretability
✓ Informativeness

Setup Stage → Elicitation Stage → Fitting Stage → Evaluation Stage

Interpretability

Informativeness



PREDICTIVE PRIOR ELICITATION

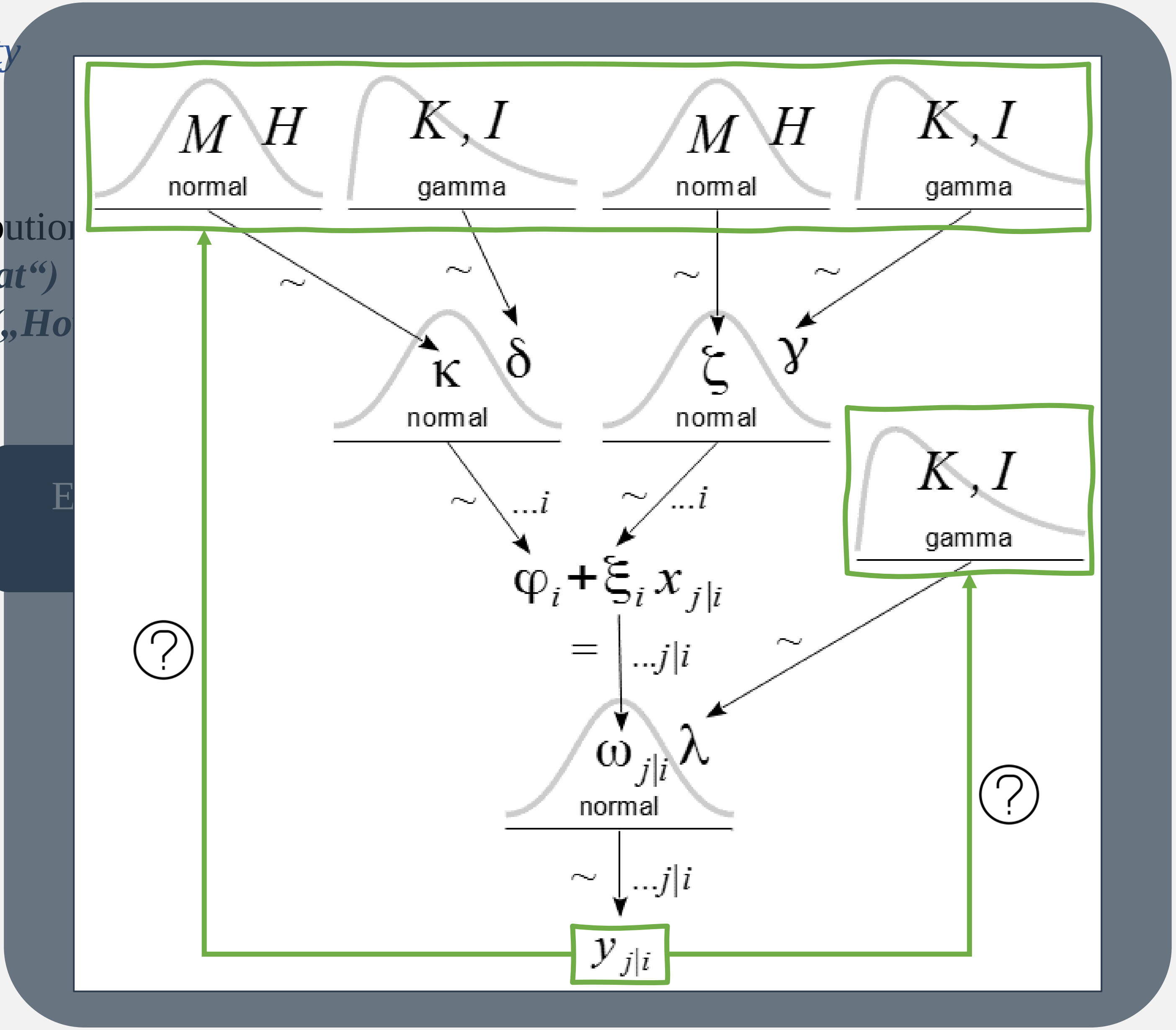
The promise and the difficulty

- ✓ selection of experts
- ✓ data model (data distributio
- ✓ ***target quantities („What“)***
- ✓ ***elicitation techniques („Ho***

Setup Stage

E

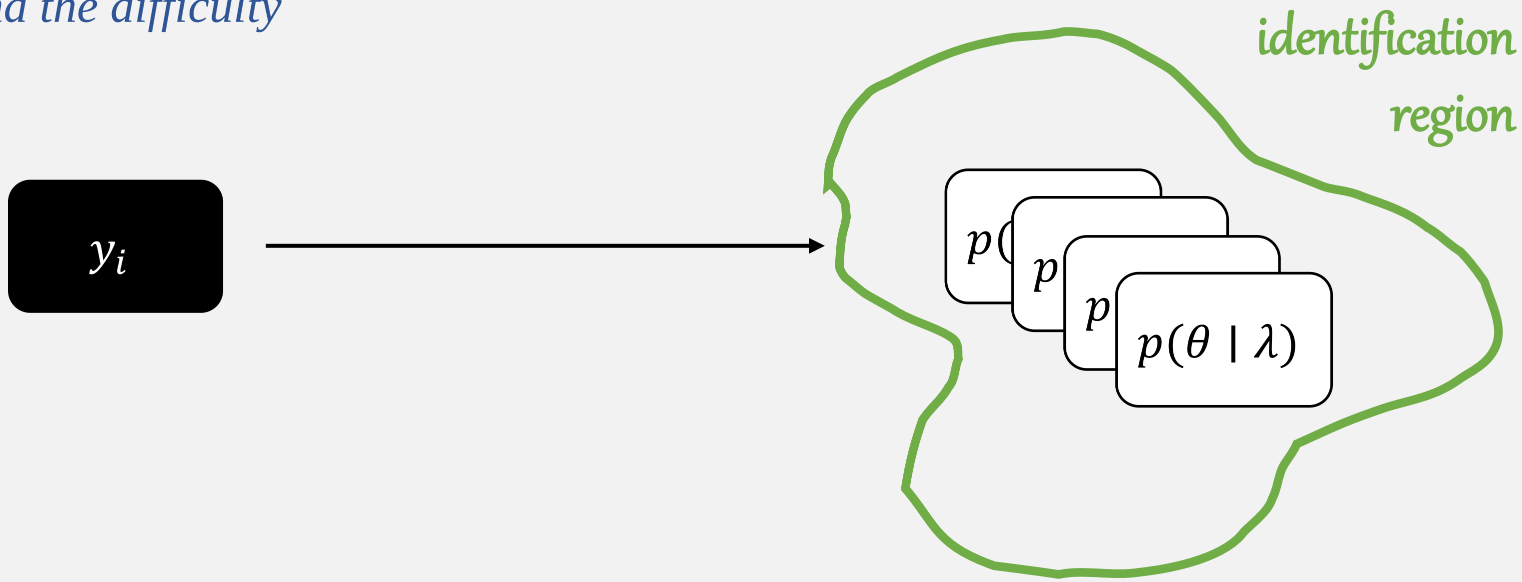
Interpretability



taken from Kruschke (2014)

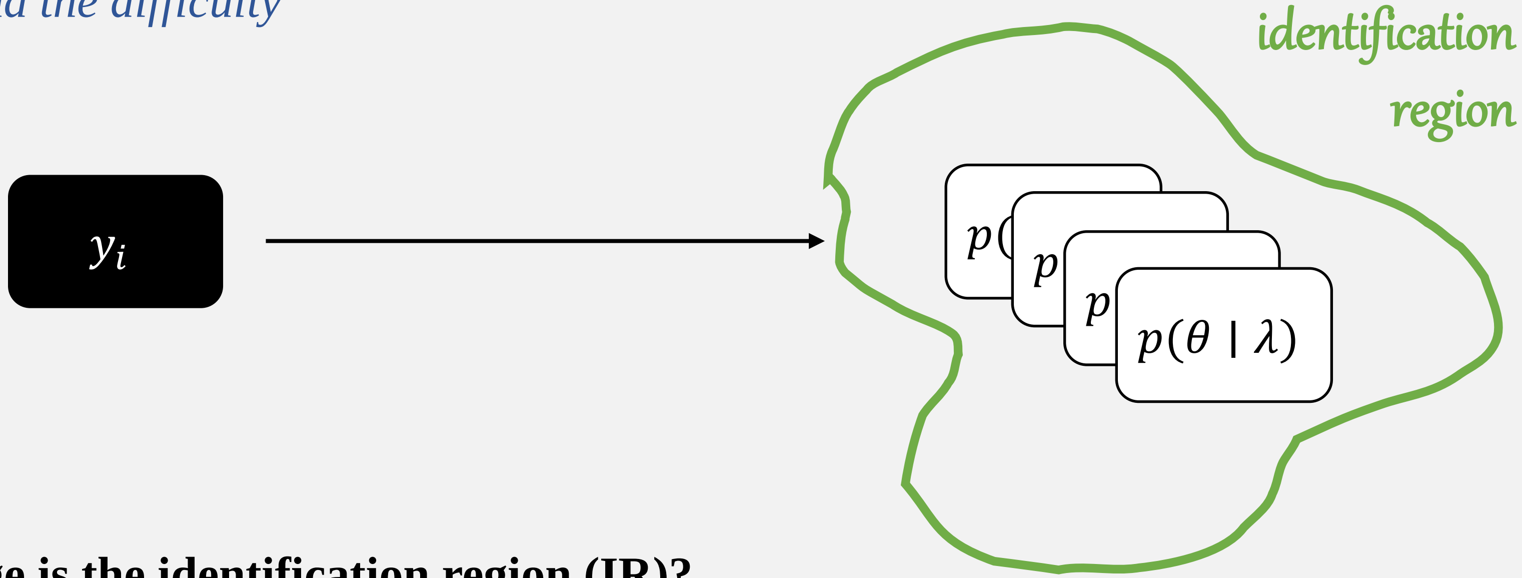
PREDICTIVE PRIOR ELICITATION

The promise and the difficulty



PREDICTIVE PRIOR ELICITATION

The promise and the difficulty



How large is the identification region (IR)?

- Naive approach: Sample from IR
- Principled approach: Prior on $p(\lambda)$ and compute posterior $p(\lambda \mid \theta)$

How can we reasonably constrain the identification region?

- Adjust prior $p(\lambda)$ (if exists)
- Adjust set of target quantities/elicitation techniques
- Add regularization term to the loss function

PREDICTIVE PRIOR ELICITATION

The promise and the difficulty

DESIDERATA – EPE METHODS

What we should aim for and where we are

- ✓ D-M1 accommodates a **flexible definition of target quantities**, supporting quantities defined in both the parameter space and the observable space.
- ✓ D-M2 accommodates a **flexible range of elicitation techniques**, such as moments, quantiles, and distributions.
- ✓ D-M3 is **agnostic** to the **model** formulation

DESIDERATA – EPE METHODS

What we should aim for and where we are

- ✓ D-M4 **propagates total uncertainty** from the elicitation process into the resulting prior distributions.
- ✓ D-M5 always returns a **learned prior** distribution, regardless of how **limited** the input **information** is.
- ✓ D-M6 **detects incoherent input information**, reconciling incoherence where possible or providing feedback on the incoherence.
- ✓ D-M7 returns the **same set of learned priors** when fitted to the **same set of expert-elicited summaries**, ensuring reproducibility

DESIDERATA – EPE WORKFLOW

The need to understand EPE methods in a broader context

- ✓ D-W1 **integrate** EPE methods within **EPE protocols**
- ✓ D-W2 **general evaluation framework** (standard set of diagnostics, evaluation metrics, …)
- ✓ D-W3 **benchmark data sets**; standardized **comparison between EPE methods**
- ✓ D-W4 **case studies** showcasing the use of EPE methods in **real-world situations** to challenge it with complexity of reality
- ✓ D-W5 **robustness analysis**, i.e., quantifying consequences of selecting specific prior for subsequent Bayesian inference task (change of posterior)

THE NEED FOR SOFTWARE

Great methods fail if no one can use them

- ✓ D-S1 **interfaces** compatible **with expert-friendly elicitation tools** accommodating different response formats.
- ✓ D-S2 **integrates into an elicitation protocol**, allowing immediate fitting of prior distributions to elicited summaries, delivery of informative visual feedback and diagnostics, …
- ✓ D-S3 **modular, open-source, and version-controlled**, facilitating community-driven development, easy modification, integration of extensions, and transparency.

THE NEED FOR SOFTWARE

Great methods fail if no one can use them

- ✓ D-S4 **integrated into the broader Bayesian workflow**, ensuring seamless exchange of information between the EPE and Bayesian workflows

- ✓ D-S5 **compatible** with **different probabilistic programming languages**.

- ✓ D-S6 **facilitator-friendly,** providing an **intuitive interface**, comprehensive **documentation**, tutorials, and case studies.

- ✓ D-S7 standard set of **evaluation metrics**, **diagnostics**, and **visualization** tools.

THANK YOU

Florence Bockting

https://github.com/florence-bockting

florence.bockting@tu-dortmund.de

https://florence-bockting.github.io/