Imprecise probabilistic models for inference in exponential families

Erik Quaeghebeur   Gert de Cooman

SYSTeMS reseach group
Ghent University

2 June 2006

Sketch of the context

- There is some stochastic process generating samples.
- Make inferences about 'things' depending on this process.
  - Parametric: depending on the value of its parameters.
  - Predictive: depending on the next sample(s).
- Inferences are typically expressed using
  - probabilities of events, or
  - previsions of gambles.
- The sample size is possibly small, so instead use
  - lower & upper probabilities of events, or
  - lower & upper previsions of gambles.
- A classical inference model structure is used:
  - Choose a prior model, and
  - update it with the sample data using (generalized) Bayes's rule: "prior and likelihood combine into a posterior".
  - Generate inferences with the posterior.

Sketch of the context

- There is some stochastic process generating samples.
- Make inferences about 'things' depending on this process.
  - Parametric: depending on the value of its parameters.
  - Predictive: depending on the next sample(s).
- Inferences are typically expressed using
  - probabilities of events, or
  - previsions of gambles.
- The sample size is possibly small, so instead use
  - lower & upper probabilities of events, or
  - lower & upper previsions of gambles.
- A classical inference model structure is used:
  - Choose a prior model, and
  - update it with the sample data using (generalized) Bayes's rule: "prior and likelihood combine into a posterior".
  - Generate inferences with the posterior.

Sketch of the context

- There is some stochastic process generating samples.
- Make inferences about 'things' depending on this process.
  - Parametric: depending on the value of its parameters.
  - Predictive: depending on the next sample(s).
- Inferences are typically expressed using
  - probabilities of events, or
  - previsions of gambles.
- The sample size is possibly small, so instead use
  - lower & upper probabilities of events, or
  - lower & upper previsions of gambles.
- A classical inference model structure is used:
  - Choose a prior model, and
  - update it with the sample data using (generalized) Bayes's rule: "prior and likelihood combine into a posterior".
  - Generate inferences with the posterior.

Sketch of the context

- There is some stochastic process generating samples.
- Make inferences about 'things' depending on this process.
  - Parametric: depending on the value of its parameters.
  - Predictive: depending on the next sample(s).
- Inferences are typically expressed using
  - probabilities of events, or
  - previsions of gambles.
- The sample size is possibly small, so instead use
  - lower & upper probabilities of events, or
  - lower & upper previsions of gambles.
- A classical inference model structure is used:
  - Choose a prior model, and
  - update it with the sample data using (generalized) Bayes's rule: "prior and likelihood combine into a posterior".
  - Generate inferences with the posterior.

Sketch of the context

- There is some stochastic process generating samples.
- Make inferences about 'things' depending on this process.
    - Parametric: depending on the value of its parameters.
    - Predictive: depending on the next sample(s).
- Inferences are typically expressed using
    - probabilities of events, or
    - previsions of gambles.
- The sample size is possibly small, so instead use
    - lower & upper probabilities of events, or
    - lower & upper previsions of gambles.
- A classical inference model structure is used:
    - Choose a prior model, and
    - update it with the sample data using (generalized) Bayes's rule: "prior and likelihood combine into a posterior".
    - Generate inferences with the posterior.

Sketch of the context

- There is some stochastic process generating samples.
- Make inferences about 'things' depending on this process.
  - Parametric: depending on the value of its parameters.
  - Predictive: depending on the next sample(s).
- Inferences are typically expressed using
  - probabilities of events, or
  - previsions of gambles.
- The sample size is possibly small, so instead use
  - lower & upper probabilities of events, or
  - lower & upper previsions of gambles.
- A classical inference model structure is used:
  - Choose a prior model, and
  - update it with the sample data using (generalized) Bayes's rule: "prior and likelihood combine into a posterior".
  - Generate inferences with the posterior.

Sketch of the context

- There is some stochastic process generating samples.
- Make inferences about 'things' depending on this process.
    - Parametric: depending on the value of its parameters.
    - Predictive: depending on the next sample(s).
- Inferences are typically expressed using
    - probabilities of events, or
    - previsions of gambles.
- The sample size is possibly small, so instead use
    - lower & upper probabilities of events, or
    - lower & upper previsions of gambles.
- A classical inference model structure is used:
    - Choose a prior model, and
    - update it with the sample data using (generalized) Bayes's rule: "prior and likelihood combine into a posterior".
    - Generate inferences with the posterior.

Exponential family sampling models

- Stochastic processes we look at:

  exponential family sampling models.

- Exponential families:

  Normal, Poisson, Exponential, Bernoulli,. . .

- Typical exponential family form:
  - For a sequence $\boldsymbol{x}$ of m samples,

$$\mathrm{Ef}_{\psi}(\boldsymbol{x}) = \mathrm{a}(\boldsymbol{x}) \exp^{m}\left(\langle \psi, \overline{\tau}(\boldsymbol{x}) \rangle - \mathrm{b}(\psi)\right).$$

- Other concepts:
  - $\mathrm{SEf}_{\boldsymbol{x}}(\psi) = \mathrm{Ef}_{\psi}(\boldsymbol{x})$,
  - Sufficient statistic $(m, \overline{\tau}(\boldsymbol{x}))$, and
  - The likelihood function

$$\mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})}(\psi) = \exp^{m}\left(\langle \psi, \overline{\tau}(\boldsymbol{x}) \rangle - \mathrm{b}(\psi)\right).$$

Exponential family sampling models

- Stochastic processes we look at:
  exponential family sampling models.
- Exponential families:
  Normal, Poisson, Exponential, Bernoulli,. . .
- Typical exponential family form:
  - For a sequence $\boldsymbol{x}$ of m samples,

$$\mathrm{Ef}_{\psi}(\boldsymbol{x}) = \mathrm{a}(\boldsymbol{x}) \exp^{m}\left(\langle\psi, \overline{\tau}(\boldsymbol{x})\rangle - \mathrm{b}(\psi)\right).$$

- Other concepts:
  - $\mathrm{SEf}_{\boldsymbol{x}}(\psi) = \mathrm{Ef}_{\psi}(\boldsymbol{x})$,
  - Sufficient statistic $(m, \overline{\tau}(\boldsymbol{x}))$, and
  - The likelihood function

$$\mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})}(\psi) = \exp^{m}\left(\langle\psi, \overline{\tau}(\boldsymbol{x})\rangle - \mathrm{b}(\psi)\right).$$

Exponential family sampling models

- Stochastic processes we look at:

  exponential family sampling models.
- Exponential families:

  Normal, Poisson, Exponential, Bernoulli,...
- Typical exponential family form:
  - For one sample x,

$$\mathrm{Ef}_{\psi}(x) = \mathrm{a}(x) \exp\left(\langle \psi, \tau(x) \rangle - \mathrm{b}(\psi)\right).$$

  - For a sequence $\boldsymbol{x}$ of m samples,

$$\mathrm{Ef}_{\psi}(\boldsymbol{x}) = \mathrm{a}(\boldsymbol{x}) \exp^{m}\left(\langle \psi, \overline{\tau}(\boldsymbol{x}) \rangle - \mathrm{b}(\psi)\right).$$

- Other concepts:
  - $\mathrm{SEf}_{\boldsymbol{x}}(\psi) = \mathrm{Ef}_{\psi}(\boldsymbol{x})$,
  - Sufficient statistic $(m, \overline{\tau}(\boldsymbol{x}))$, and
  - The likelihood function

$$\mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})}(\psi) = \exp^{m}\left(\langle \psi, \overline{\tau}(\boldsymbol{x}) \rangle - \mathrm{b}(\psi)\right).$$

Exponential family sampling models

- Stochastic processes we look at:

  exponential family sampling models.

- Exponential families:

  Normal, Poisson, Exponential, Bernoulli,. . .

- Typical exponential family form:
  - For a sequence $\boldsymbol{x}$ of m samples,

$$\mathrm{Ef}_{\psi}(\boldsymbol{x}) = \mathrm{a}(\boldsymbol{x}) \exp^{m} \left(\langle \psi, \overline{\tau}(\boldsymbol{x}) \rangle - \mathrm{b}(\psi) \right).$$

- Other concepts:
  - $\mathrm{SEf}_{\boldsymbol{x}}(\psi) = \mathrm{Ef}_{\psi}(\boldsymbol{x})$,
  - Sufficient statistic $(m, \overline{\tau}(\boldsymbol{x}))$, and
  - The likelihood function

$$\mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})}(\psi) = \exp^{m} \left(\langle \psi, \overline{\tau}(\boldsymbol{x}) \rangle - \mathrm{b}(\psi) \right).$$

Exponential family sampling models

- Stochastic processes we look at:
  exponential family sampling models.
- Exponential families:
  Normal, Poisson, Exponential, Bernoulli, . . .
- Typical exponential family form:
  - For a sequence $\boldsymbol{x}$ of m samples,

$$\mathrm{Ef}_{\psi}(\boldsymbol{x}) = \mathrm{a}(\boldsymbol{x}) \exp^{m}\left(\langle \psi, \overline{\tau}(\boldsymbol{x}) \rangle - \mathrm{b}(\psi)\right).$$

- Other concepts:
  - $\mathrm{SEf}_{\boldsymbol{x}}(\psi) = \mathrm{Ef}_{\psi}(\boldsymbol{x})$,
  - Sufficient statistic $(m, \overline{\tau}(\boldsymbol{x}))$, and
  - The likelihood function

$$\mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})}(\psi) = \exp^{m}\left(\langle \psi, \overline{\tau}(\boldsymbol{x}) \rangle - \mathrm{b}(\psi)\right).$$

Exponential family sampling models

- Stochastic processes we look at:

  exponential family sampling models.
- Exponential families:

  Normal, Poisson, Exponential, Bernoulli,...
- Typical exponential family form:
  - For a sequence $\boldsymbol{x}$ of m samples,

$$\mathrm{Ef}_{\psi}(\boldsymbol{x}) = \mathrm{a}(\boldsymbol{x}) \exp^{m}\left(\langle \psi, \overline{\tau}(\boldsymbol{x})\rangle - \mathrm{b}(\psi)\right).$$

- Other concepts:
  - $\mathrm{SEf}_{\boldsymbol{x}}(\psi) = \mathrm{Ef}_{\psi}(\boldsymbol{x})$,
  - Sufficient statistic $(m, \overline{\tau}(\boldsymbol{x}))$, and
  - The likelihood function

$$\mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})}(\psi) = \exp^{m}\left(\langle \psi, \overline{\tau}(\boldsymbol{x})\rangle - \mathrm{b}(\psi)\right).$$

The foundations: conjugate & predictive distributions

- Basis of the prior & posterior models:

  conjugate families of distributions.

- Typical conjugate family form:

$$\mathrm{CEf}_{n,y}(\psi) = \mathrm{c}(n,y) \exp^{n}\left(\langle \psi, y \rangle - \mathrm{b}(\psi)\right).$$

- Updating is done using Bayes's rule:

$$\mathrm{CEf}_{n_0,y_0} \mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})} \propto \mathrm{CEf}_{n_0+m, \frac{n_0 y_0 + m\overline{\tau}(\boldsymbol{x})}{n_0+m}}.$$

- $\mathrm{CEf}_{n,y}$ is the basis for the parametric inference models.
- The predictive family of distributions is derived from the conjugate family:

$$\mathrm{PEf}_{n,y}(\boldsymbol{x}) = \int_{\Psi} \mathrm{CEf}_{n,y} \mathrm{SEf}_{\boldsymbol{x}} = \frac{\mathrm{c}(n,y)\mathrm{a}(\boldsymbol{x})}{\mathrm{c}\left(n+m, \frac{ny+m\overline{\tau}(\boldsymbol{x})}{n+m}\right)}.$$

- $\mathrm{PEf}_{n,y}$ is the basis for the parametric inference models.

The foundations: conjugate & predictive distributions

- Basis of the prior & posterior models:

  conjugate families of distributions.

- Typical conjugate family form:

$$\mathrm{CEf}_{n,y}(\psi) = \mathrm{c}(n,y)\exp^{n}\left(\langle\psi,y\rangle - \mathrm{b}(\psi)\right).$$

- Updating is done using Bayes's rule:

$$\mathrm{CEf}_{n_0,y_0}\mathrm{LEf}_{m,\overline{\tau}(x)} \propto \mathrm{CEf}_{n_0+m,\frac{n_0y_0+m\overline{\tau}(x)}{n_0+m}}.$$

- $\mathrm{CEf}_{n,y}$ is the basis for the parametric inference models.
- The predictive family of distributions is derived from the conjugate family:

$$\mathrm{PEf}_{n,y}(x) = \int_{\Psi}\mathrm{CEf}_{n,y}\mathrm{SEf}_{x} = \frac{\mathrm{c}(n,y)\mathrm{a}(x)}{\mathrm{c}(n+m,\frac{ny+m\overline{\tau}(x)}{n+m})}.$$

- $\mathrm{PEf}_{n,y}$ is the basis for the parametric inference models.

The foundations: conjugate & predictive distributions

- Basis of the prior & posterior models:

  conjugate families of distributions.

- Typical conjugate family form:

$$\mathrm{CEf}_{n,y}(\psi) = \mathrm{c}(n, y) \exp^n (\langle \psi, y \rangle - \mathrm{b}(\psi)).$$

- Updating is done using Bayes's rule:

$$\mathrm{CEf}_{n_0,y_0} \mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})} \propto \mathrm{CEf}_{n_0+m,\frac{n_0 y_0 + m\overline{\tau}(\boldsymbol{x})}{n_0+m}}.$$

- $\mathrm{CEf}_{n,y}$ is the basis for the parametric inference models.
- The predictive family of distributions is derived from the conjugate family:

$$\mathrm{PEf}_{n,y}(\boldsymbol{x}) = \int_{\Psi} \mathrm{CEf}_{n,y} \mathrm{SEf}_{\boldsymbol{x}} = \frac{\mathrm{c}(n, y)\mathrm{a}(\boldsymbol{x})}{\mathrm{c}(n + m, \frac{ny + m\overline{\tau}(\boldsymbol{x})}{n+m})}.$$

- $\mathrm{PEf}_{n,y}$ is the basis for the parametric inference models.

The foundations: conjugate & predictive distributions

- Basis of the prior & posterior models:

  conjugate families of distributions.

- Typical conjugate family form:

$$\mathrm{CEf}_{n,y}(\psi) = \mathrm{c}(n, y) \exp^{n} (\langle \psi, y \rangle - \mathrm{b}(\psi)).$$

- Updating is done using Bayes's rule:

$$\mathrm{CEf}_{n_0,y_0} \mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})} \propto \mathrm{CEf}_{n_0+m, \frac{n_0 y_0 + m\overline{\tau}(\boldsymbol{x})}{n_0+m}}.$$

- $\mathrm{CEf}_{n,y}$ is the basis for the parametric inference models.
- The predictive family of distributions is derived from the conjugate family:

$$\mathrm{PEf}_{n,y}(\boldsymbol{x}) = \int_{\Psi} \mathrm{CEf}_{n,y} \mathrm{SEf}_{\boldsymbol{x}} = \frac{\mathrm{c}(n, y)\mathrm{a}(\boldsymbol{x})}{\mathrm{c}(n + m, \frac{ny + m\overline{\tau}(\boldsymbol{x})}{n+m})}.$$

- $\mathrm{PEf}_{n,y}$ is the basis for the parametric inference models.

The foundations: conjugate & predictive distributions

- Basis of the prior & posterior models:

  conjugate families of distributions.
- Typical conjugate family form:

$$\mathrm{CEf}_{n,y}(\psi) = \mathrm{c}(n,y)\exp^{n}\left(\langle\psi,y\rangle - \mathrm{b}(\psi)\right).$$

- Updating is done using Bayes's rule:

$$\mathrm{CEf}_{n_0,y_0}\mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})} \propto \mathrm{CEf}_{n_0+m,\frac{n_0y_0+m\overline{\tau}(\boldsymbol{x})}{n_0+m}}.$$

- $\mathrm{CEf}_{n,y}$ is the basis for the parametric inference models.
- The predictive family of distributions is derived from the conjugate family:

$$\mathrm{PEf}_{n,y}(\boldsymbol{x}) = \int_{\Psi}\mathrm{CEf}_{n,y}\mathrm{SEf}_{\boldsymbol{x}} = \frac{\mathrm{c}(n,y)\mathrm{a}(\boldsymbol{x})}{\mathrm{c}(n+m,\frac{ny+m\overline{\tau}(\boldsymbol{x})}{n+m})}.$$

- $\mathrm{PEf}_{n,y}$ is the basis for the parametric inference models.

The foundations: conjugate & predictive distributions

- Basis of the prior & posterior models:

  conjugate families of distributions.
- Typical conjugate family form:

$$\mathrm{CEf}_{n,y}(\psi) = \mathrm{c}(n,y)\exp^n\left(\langle\psi,y\rangle - \mathrm{b}(\psi)\right).$$

- Updating is done using Bayes's rule:

$$\mathrm{CEf}_{n_0,y_0}\mathrm{LEf}_{m,\overline{\tau}(x)} \propto \mathrm{CEf}_{n_0+m,\frac{n_0y_0+m\overline{\tau}(x)}{n_0+m}}.$$

- $\mathrm{CEf}_{n,y}$ is the basis for the parametric inference models.
- The predictive family of distributions is derived from the conjugate family:

$$\mathrm{PEf}_{n,y}(\boldsymbol{x}) = \int_\Psi \mathrm{CEf}_{n,y}\mathrm{SEf}_{\boldsymbol{x}} = \frac{\mathrm{c}(n,y)\mathrm{a}(\boldsymbol{x})}{\mathrm{c}(n+m,\frac{ny+m\overline{\tau}(\boldsymbol{x})}{n+m})}.$$

- $\mathrm{PEf}_{n,y}$ is the basis for the parametric inference models.

The foundations: conjugate & predictive distributions

- Basis of the prior & posterior models:

  conjugate families of distributions.

- Typical conjugate family form:

$$\mathrm{CEf}_{n,y}(\psi) = \mathrm{c}(n,y)\exp^{n}\left(\langle\psi,y\rangle - \mathrm{b}(\psi)\right).$$

- Updating is done using Bayes's rule:

$$\mathrm{CEf}_{n_0,y_0}\mathrm{LEf}_{m,\overline{\tau}(\boldsymbol{x})} \propto \mathrm{CEf}_{n_0+m,\frac{n_0y_0+m\overline{\tau}(\boldsymbol{x})}{n_0+m}}.$$

- $\mathrm{CEf}_{n,y}$ is the basis for the parametric inference models.
- The predictive family of distributions is derived from the conjugate family:

$$\mathrm{PEf}_{n,y}(\boldsymbol{x}) = \int_{\Psi}\mathrm{CEf}_{n,y}\mathrm{SEf}_{\boldsymbol{x}} = \frac{\mathrm{c}(n,y)\mathrm{a}(\boldsymbol{x})}{\mathrm{c}(n+m,\frac{ny+m\overline{\tau}(\boldsymbol{x})}{n+m})}.$$

- $\mathrm{PEf}_{n,y}$ is the basis for the parametric inference models.

The 'precise' inference models I

- The inference models are linear previsions defined using a distribution:
  - The conjugate distribution is used for the parametric model,

$$P_{\mathrm{C}}(f \mid n, y) = \int_{\Psi} f \, \mathrm{CEf}_{n,y},$$

  - the predictive distribution is used for the predictive model,

$$P_{\mathrm{P}}(g \mid n, y) = \int_{\mathcal{X}_m^*} g \, \mathrm{CEf}_{n,y}.$$

- The prevision of particular gambles is easy to calculate:
  - Considering that $\nabla_{\psi} \mathrm{b}(\psi) = \int_{\mathcal{X}_m^*} \overline{\tau} \mathrm{Ef}_{\psi}$, it is a nice result for the parametric model that

$$P_{\mathrm{C}}(\nabla \mathrm{b} \mid n, y) = y.$$

  - An analogous result holds for the predictive model:

$$P_{\mathrm{P}}(\overline{\tau} \mid n, y) = y.$$

The 'precise' inference models I

- The inference models are linear previsions defined using a distribution:
  - The conjugate distribution is used for the parametric model,

$$P_{\mathrm{C}}(f \mid n, y) = \int_{\Psi} f \, \mathrm{CEf}_{n,y},$$

  - the predictive distribution is used for the predictive model,

$$P_{\mathrm{P}}(g \mid n, y) = \int_{\mathcal{X}_m^*} g \, \mathrm{CEf}_{n,y}.$$

- The prevision of particular gambles is easy to calculate:
  - Considering that $\nabla_{\psi} \mathrm{b}(\psi) = \int_{\mathcal{X}_m^*} \overline{\tau} \mathrm{Ef}_{\psi}$, it is a nice result for the parametric model that

$$P_{\mathrm{C}}(\nabla \mathrm{b} \mid n, y) = y.$$

  - An analogous result holds for the predictive model:

$$P_{\mathrm{P}}(\overline{\tau} \mid n, y) = y.$$

The ‘precise’ inference models I

- The inference models are linear previsions defined using a distribution:
  - The conjugate distribution is used for the parametric model,

$$P_{\mathrm{C}}(f \mid n, y) = \int_{\Psi} f \, \mathrm{CEf}_{n,y},$$

  - the predictive distribution is used for the predictive model,

$$P_{\mathrm{P}}(g \mid n, y) = \int_{\mathcal{X}_m^*} g \, \mathrm{CEf}_{n,y}.$$

- The prevision of particular gambles is easy to calculate:
  - Considering that $\nabla_{\psi}\mathrm{b}(\psi) = \int_{\mathcal{X}_m^*} \overline{\tau}\mathrm{Ef}_{\psi}$, it is a nice result for the parametric model that

$$P_{\mathrm{C}}(\nabla \mathrm{b} \mid n, y) = y.$$

  - An analogous result holds for the predictive model:

$$P_{\mathrm{P}}(\overline{\tau} \mid n, y) = y.$$

The 'precise' inference models II

- Problem: how to choose n_0 and y_0?
- Effect on the parameters of updating with a sequence of m samples $\boldsymbol{x}$:

$$(n_m, y_m) = (n_0 + m, \frac{n_0 y_0 + m\overline{\tau}(\boldsymbol{x})}{n_0 + m}).$$

The 'precise' inference models II

- Problem: how to choose n_0 and y_0?
- Effect on the parameters of updating with a sequence of m samples $\boldsymbol{x}$:

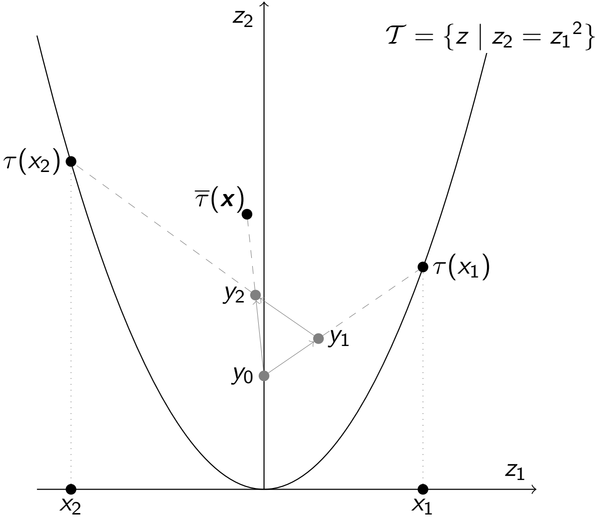
$$(n_m, y_m) = \left(n_0 + m, \frac{n_0 y_0 + m\overline{\tau}(\boldsymbol{x})}{n_0 + m}\right).$$

The ‘precise’ inference models II

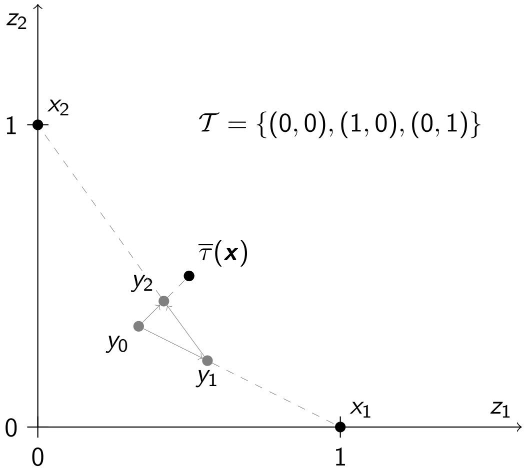
- Problem: how to choose n_0 and y_0?
- Effect on the parameters of updating with a sequence of m samples $\boldsymbol{x}$:

$$(n_m, y_m) = \left(n_0 + m, \frac{n_0 y_0 + m\overline{\tau}(\boldsymbol{x})}{n_0 + m}\right).$$

Updating the ‘precise’ inference models: normal

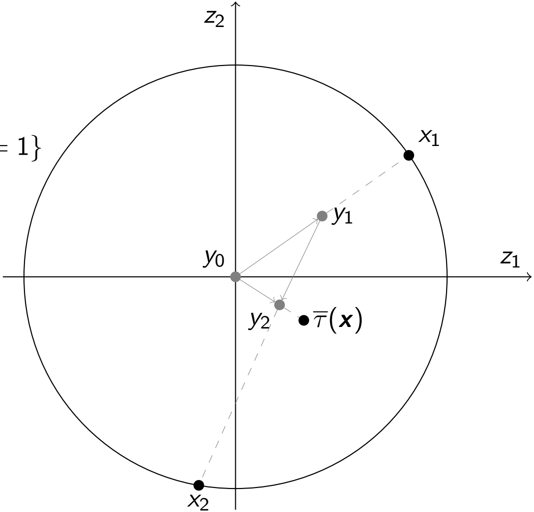


Updating the ‘precise’ inference models: Bernoulli



Updating the 'precise' inference models: von Mises

$$\mathcal{T} = \mathbb{S}_2 = \{z \mid \|z\| = 1\}$$



The 'imprecise' inference models I

- The inference models are lower previsions defined using sets of distributions:
  - Sets of conjugate distribution for the parametric model,
  $$\underline{P}_{\mathrm{C}}(f \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} P_{\mathrm{C}}(f \mid n, y),$$
  - sets of predictive distributions for the predictive model,
  $$\underline{P}_{\mathrm{P}}(g \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} P_{\mathrm{P}}(g \mid n, y).$$
- The lower prevision of some gambles is again easy to calculate:
  - The nice result for the parametric model is
  $$\underline{P}_{\mathrm{C}}(\nabla \mathrm{b} \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} y.$$
  - The analogous result for the predictive model:
  $$\underline{P}_{\mathrm{P}}(\overline{\tau} \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} y.$$

The ‘imprecise’ inference models I

- The inference models are lower previsions defined using sets of distributions:
  - Sets of conjugate distribution for the parametric model,
  $$\underline{P}_{\mathrm{C}}(f \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} P_{\mathrm{C}}(f \mid n, y),$$
  - sets of predictive distributions for the predictive model,
  $$\underline{P}_{\mathrm{P}}(g \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} P_{\mathrm{P}}(g \mid n, y).$$
- The lower prevision of some gambles is again easy to calculate:
  - The nice result for the parametric model is
  $$\underline{P}_{\mathrm{C}}(\nabla \mathrm{b} \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} y.$$
  - The analogous result for the predictive model:
  $$\underline{P}_{\mathrm{P}}(\overline{\tau} \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} y.$$

The 'imprecise' inference models I

- The inference models are lower previsions defined using sets of distributions:
  - Sets of conjugate distribution for the parametric model,

$$\underline{P}_{\mathrm{C}}(f \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} P_{\mathrm{C}}(f \mid n, y),$$

  - sets of predictive distributions for the predictive model,

$$\underline{P}_{\mathrm{P}}(g \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} P_{\mathrm{P}}(g \mid n, y).$$

- The lower prevision of some gambles is again easy to calculate:
  - The nice result for the parametric model is

$$\underline{P}_{\mathrm{C}}(\nabla \mathrm{b} \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} y.$$

  - The analogous result for the predictive model:

$$\underline{P}_{\mathrm{P}}(\overline{\tau} \mid n, \mathcal{Y}) = \inf_{y \in \mathcal{Y}} y.$$

The 'imprecise' inference models II

- Problem: how to choose n_0 and $\mathcal{Y}_0$?
- Effect on the parameters of updating with a sequence of m samples $\boldsymbol{x}$:

$$(n_m, \mathcal{Y}_m) = (n_0 + m, \frac{n_0\mathcal{Y}_0 + m\overline{\tau}(\boldsymbol{x})}{n_0 + m}).$$

The 'imprecise' inference models II

- Problem: how to choose n_0 and $\mathcal{Y}_0$?
- Effect on the parameters of updating with a sequence of m samples $\boldsymbol{x}$:

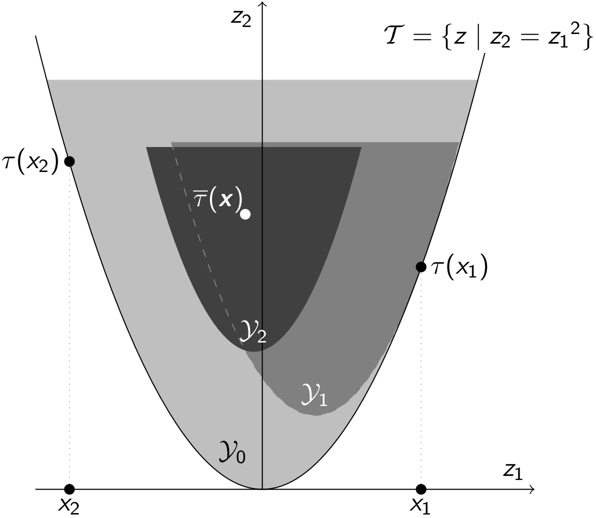
$$(n_m, \mathcal{Y}_m) = (n_0 + m, \frac{n_0\mathcal{Y}_0 + m\overline{\tau}(\boldsymbol{x})}{n_0 + m}).$$

The 'imprecise' inference models II

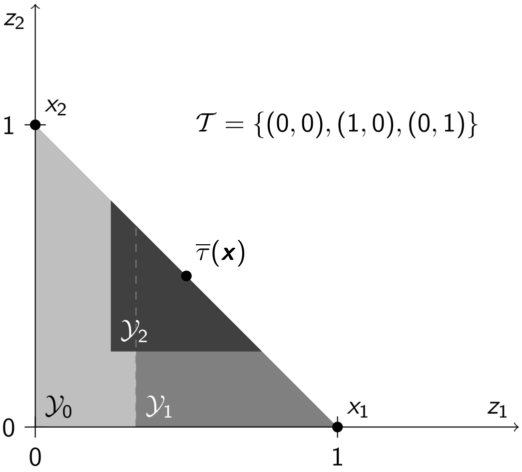
- Problem: how to choose n_0 and $\mathcal{Y}_0$?
- Effect on the parameters of updating with a sequence of m samples $\boldsymbol{x}$:

$$(n_m, \mathcal{Y}_m) = (n_0 + m, \frac{n_0\mathcal{Y}_0 + m\overline{\tau}(\boldsymbol{x})}{n_0 + m}).$$

Updating the 'imprecise' inference models: normal

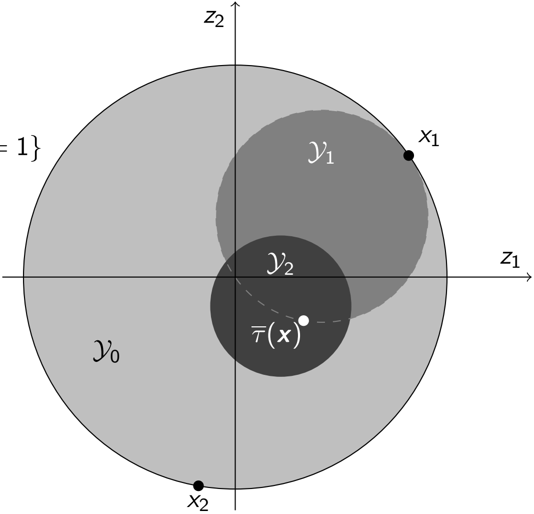


Updating the 'imprecise' inference models: Bernoulli



Updating the ‘imprecise’ inference models: von Mises

$\mathcal{T} = \mathbb{S}_2 = \{z \mid \|z\| = 1\}$



Other uses of the inference models

- Calculate lower and upper cumulative distribution functions.
- Combine inference models for multiple stochastic processes:
  - Combining independent marginals, or
  - combining marginal and conditional models.

Other uses of the inference models

- Calculate lower and upper cumulative distribution functions.
- Combine inference models for multiple stochastic processes:
  - Combining independent marginals, or
  - combining marginal and conditional models.

Other uses of the inference models

- Calculate lower and upper cumulative distribution functions.
- Combine inference models for multiple stochastic processes:
  - Combining independent marginals, or
  - combining marginal and conditional models.

Other uses of the inference models

- Calculate lower and upper cumulative distribution functions.
- Combine inference models for multiple stochastic processes:
  - Combining independent marginals, or
  - combining marginal and conditional models.

Summary

- For *all* exponential family sampling models, imprecise probabilistic parametric and predictive inference models can be defined.
- Updating the models with sample data consist of a *straightforward* modification of the model parameters.
- Some inferences can be obtained very easily.

- Open questions:
  - What other inferences can be calculated?
  - How do the inferences depend on n?
  - How do these models compare with other models?

Summary

- For *all* exponential family sampling models, imprecise probabilistic parametric and predictive inference models can be defined.
- Updating the models with sample data consist of a *straightforward* modification of the model parameters.
- Some inferences can be obtained very easily.
- Open questions:
  - What other inferences can be calculated?
  - How do the inferences depend on n?
  - How do these models compare with other models?

Summary

- For *all* exponential family sampling models, imprecise probabilistic parametric and predictive inference models can be defined.
- Updating the models with sample data consist of a *straightforward* modification of the model parameters.
- Some inferences can be obtained very easily.
- Open questions:
    - What other inferences can be calculated?
    - How do the inferences depend on n?
    - How do these models compare with other models?

Summary

- For *all* exponential family sampling models, imprecise probabilistic parametric and predictive inference models can be defined.
- Updating the models with sample data consist of a *straightforward* modification of the model parameters.
- Some inferences can be obtained very easily.

- Open questions:
  - What other inferences can be calculated?
  - How do the inferences depend on n?
  - How do these models compare with other models?

Summary

- For *all* exponential family sampling models, imprecise probabilistic parametric and predictive inference models can be defined.
- Updating the models with sample data consist of a *straightforward* modification of the model parameters.
- Some inferences can be obtained very easily.

- Open questions:
  - What other inferences can be calculated?
  - How do the inferences depend on n?
  - How do these models compare with other models?