Who is Crazier: Bayes or Fisher? A Missing (Data) Perspective on Fiducial Inference

Keli Liu and Xiao-Li Meng

Department of Statistics, Harvard University

November 15, 2013

Let's Meet the Crazies

Let's Meet the Crazies

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

Let's Meet the Crazies

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- Frequentist $1 - \alpha$ level interval: $\bar{x} \pm \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}$.

Let's Meet the Crazies

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- Frequentist $1 - \alpha$ level interval: $\bar{x} \pm \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}$.

Bayes

Let's Meet the Crazies

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- Frequentist $1 - \alpha$ level interval: $\bar{x} \pm \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}$.

Bayes

- Assume μ is random.

Let's Meet the Crazies

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- Frequentist $1 - \alpha$ level interval: $\bar{x} \pm \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}$.

Bayes

- Assume μ is random.
- Assume μ has *improper* distribution.

Let's Meet the Crazies

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- Frequentist $1 - \alpha$ level interval: $\bar{x} \pm \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}$.

Bayes

- Assume μ is random.
- Assume μ has *improper* distribution.
- $[\mu | \bar{x}, s] \sim \bar{x} - \frac{s}{\sqrt{n}} t_{n-1}$

Let's Meet the Crazies

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- Frequentist $1 - \alpha$ level interval: $\bar{x} \pm \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}$.

Bayes

- Assume μ is random.
- Assume μ has *improper* distribution.
- $[\mu|\bar{x}, s] \sim \bar{x} - \frac{s}{\sqrt{n}} t_{n-1}$

Fisher

Let's Meet the Crazies

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- Frequentist $1 - \alpha$ level interval: $\bar{x} \pm \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}$.

Bayes

- Assume μ is random.
- Assume μ has *improper* distribution.
- $[\mu|\bar{x}, s] \sim \bar{x} - \frac{s}{\sqrt{n}} t_{n-1}$

Fisher

- Fiducial Equation: $\mu = \bar{x} - \frac{s}{\sqrt{n}} t$.

Let's Meet the Crazies

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- Frequentist $1 - \alpha$ level interval: $\bar{x} \pm \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}$.

Bayes

- Assume μ is random.
- Assume μ has *improper* distribution.
- $[\mu|\bar{x}, s] \sim \bar{x} - \frac{s}{\sqrt{n}} t_{n-1}$

Fisher

- Fiducial Equation: $\mu = \bar{x} - \frac{s}{\sqrt{n}} t$.
- Assume $[t|\bar{x}, s] \sim t_{n-1}$.

Let's Meet the Crazies

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- Frequentist $1-\alpha$ level interval: $\bar{x} \pm \frac{s}{\sqrt{n}} t_{n-1,1-\alpha/2}$.

Bayes

- Assume μ is random.
- Assume μ has *improper* distribution.
- $[\mu|\bar{x}, s] \sim \bar{x} - \frac{s}{\sqrt{n}} t_{n-1}$

Fisher

- Fiducial Equation: $\mu = \bar{x} - \frac{s}{\sqrt{n}} t$.
- Assume $[t|\bar{x}, s] \sim t_{n-1}$.
- $[\mu|\bar{x}, s] \sim \bar{x} - \frac{s}{\sqrt{n}} t_{n-1}$.

Augmenting the Inferential Framework

Augmenting the Inferential Framework

Fiducial Equation (Taraldsen and Lindqvist 2013)

$$\mathbf{X} = G(\theta, \mathbf{U}) \quad \text{where} \quad \mathbf{X} \in \mathcal{X},\ \theta \in \mathbf{\Theta},\ \mathbf{U} \in \mathbb{U}$$

Augmenting the Inferential Framework

Fiducial Equation (Taraldsen and Lindqvist 2013)

$$\mathbf{X} = G(\theta, \mathbf{U}) \quad \text{where} \quad \mathbf{X} \in \mathcal{X},\ \theta \in \mathbf{\Theta},\ \mathbf{U} \in \mathbb{U}$$

- G is the structure equation. $\mathbf{U}$ is (God's) uncertainty.

Augmenting the Inferential Framework

Fiducial Equation (Taraldsen and Lindqvist 2013)

$$\mathbf{X} = G(\theta, \mathbf{U}) \quad \text{where} \quad \mathbf{X} \in \mathcal{X},\ \theta \in \mathbf{\Theta},\ \mathbf{U} \in \mathbb{U}$$

- G is the structure equation. **U** is (God's) uncertainty.
- Together, G and **U** determine the sampling distribution.

Augmenting the Inferential Framework

Fiducial Equation (Taraldsen and Lindqvist 2013)

$$\mathbf{X} = G(\theta, \mathbf{U}) \quad \text{where} \quad \mathbf{X} \in \mathcal{X},\ \theta \in \mathbf{\Theta},\ \mathbf{U} \in \mathbb{U}$$

- G is the structure equation. $\mathbf{U}$ is (God's) uncertainty.
- Together, G and $\mathbf{U}$ determine the sampling distribution.
- The sampling distribution does not determine G and $\mathbf{U}$.

Augmenting the Inferential Framework

Keli Liu and Xiao-Li Meng

Fiducial Equation (Taraldsen and Lindqvist 2013)

$$\mathbf{X} = G(\theta, \mathbf{U}) \quad \text{where} \quad \mathbf{X} \in \mathcal{X},\ \theta \in \boldsymbol{\Theta},\ \mathbf{U} \in \mathbb{U}$$

- G is the structure equation. $\mathbf{U}$ is (God's) uncertainty.
- Together, G and $\mathbf{U}$ determine the sampling distribution.
- The sampling distribution does not determine G and $\mathbf{U}$.

Method	Type of Replication	Relevant?

Augmenting the Inferential Framework

Fiducial Equation (Taraldsen and Lindqvist 2013)

$$\mathbf{X} = G(\theta, \mathbf{U}) \quad \text{where} \quad \mathbf{X} \in \mathcal{X},\ \theta \in \mathbf{\Theta},\ \mathbf{U} \in \mathbb{U}$$

- G is the structure equation. **U** is (God's) uncertainty.
- Together, G and **U** determine the sampling distribution.
- The sampling distribution does not determine G and **U**.

Method	Type of Replication		Relevant?
Frequentist	data given parameter	$\mathbf{X}\mid\theta$	×

Augmenting the Inferential Framework

Fiducial Equation (Taraldsen and Lindqvist 2013)

$$\mathbf{X} = G(\theta, \mathbf{U}) \quad \text{where} \quad \mathbf{X} \in \mathcal{X},\ \theta \in \mathbf{\Theta},\ \mathbf{U} \in \mathbb{U}$$

- G is the structure equation. $\mathbf{U}$ is (God's) uncertainty.
- Together, G and $\mathbf{U}$ determine the sampling distribution.
- The sampling distribution does not determine G and $\mathbf{U}$.

Method	**Type of Replication**		**Relevant?**
Frequentist	data given parameter	$\mathbf{X}\|\theta$	✗
Bayes	parameter given data	$\theta\|\mathbf{X}$	✓

Augmenting the Inferential Framework

Fiducial Equation (Taraldsen and Lindqvist 2013)

$$\mathbf{X} = G(\theta, \mathbf{U}) \quad \text{where} \quad \mathbf{X} \in \mathcal{X},\ \theta \in \mathbf{\Theta},\ \mathbf{U} \in \mathbb{U}$$

- G is the structure equation. $\mathbf{U}$ is (God's) uncertainty.
- Together, G and $\mathbf{U}$ determine the sampling distribution.
- The sampling distribution does not determine G and $\mathbf{U}$.

Method	Type of Replication		Relevant?
Frequentist	data given parameter	$\mathbf{X}\|\theta$	✗
Bayes	parameter given data	$\theta\|\mathbf{X}$	✓
Fiducial	uncertainty given data	$\mathbf{U}\|\mathbf{X}$	✓

Augmenting the Inferential Framework

Fiducial Equation (Taraldsen and Lindqvist 2013)

$$\mathbf{X} = G(\theta, \mathbf{U}) \quad \text{where} \quad \mathbf{X} \in \mathcal{X},\ \theta \in \mathbf{\Theta},\ \mathbf{U} \in \mathbb{U}$$

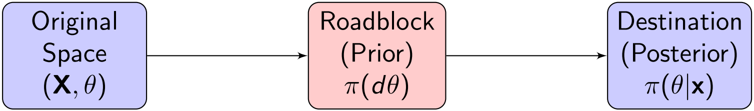
- G is the structure equation. $\mathbf{U}$ is (God's) uncertainty.
- Together, G and $\mathbf{U}$ determine the sampling distribution.
- The sampling distribution does not determine G and $\mathbf{U}$.

Method	Type of Replication		Relevant?
Frequentist	data given parameter	$\mathbf{X}\|\theta$	✗
Bayes	parameter given data	$\theta\|\mathbf{X}$	✓
Fiducial	uncertainty given data	$\mathbf{U}\|\mathbf{X}$	✓

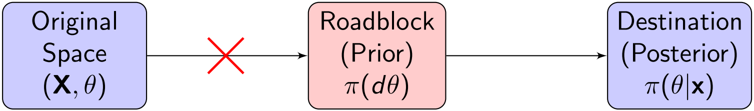
- Why should finding $\mathbf{U}|\mathbf{x}$ be any easier than finding $\theta|\mathbf{x}$?

Detour via a Fiducial Bridge?

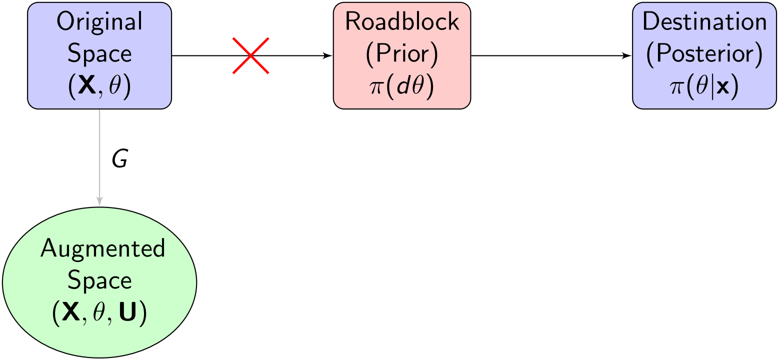
Detour via a Fiducial Bridge?



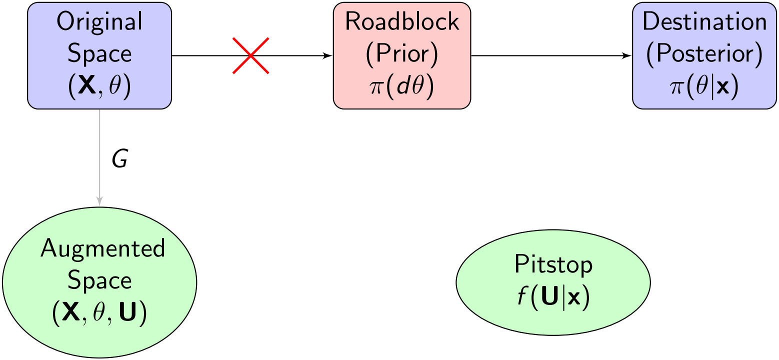
Detour via a Fiducial Bridge?



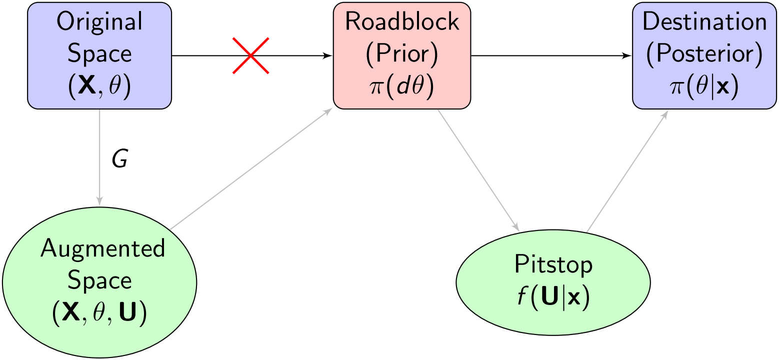
Detour via a Fiducial Bridge?



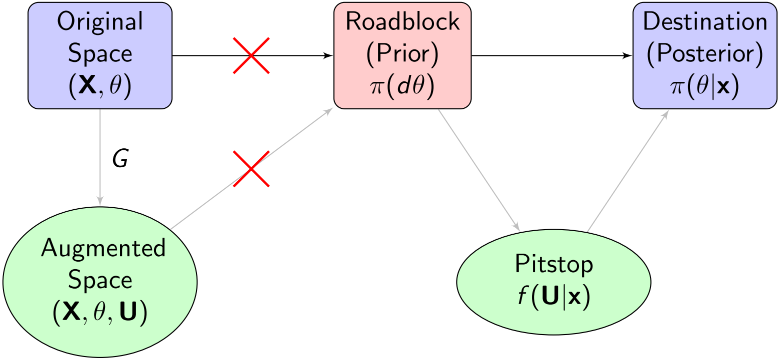
Detour via a Fiducial Bridge?



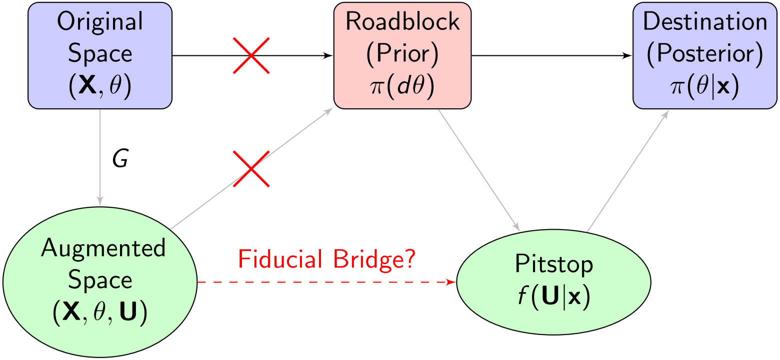
Detour via a Fiducial Bridge?



Detour via a Fiducial Bridge?



Detour via a Fiducial Bridge?



Which Smells Fishier: Bayes or Fiducial?

Which Smells Fishier: Bayes or Fiducial?

Keli Liu and Xiao-Li Meng

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

Which Smells Fishier: Bayes or Fiducial?

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- t is our missing data with prior distribution: $t \sim t_{n-1}$.

Which Smells Fishier: Bayes or Fiducial?

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- t is our missing data with prior distribution: $t \sim t_{n-1}$.

Bayes

Which Smells Fishier: Bayes or Fiducial?

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- t is our missing data with prior distribution: $t \sim t_{n-1}$.

Bayes

- *Objective* prior for μ.

Which Smells Fishier: Bayes or Fiducial?

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- t is our missing data with prior distribution: $t \sim t_{n-1}$.

Bayes

- *Objective* prior for μ.
- What does objective prior mean?

Which Smells Fishier: Bayes or Fiducial?

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- t is our missing data with prior distribution: $t \sim t_{n-1}$.

Bayes

- *Objective* prior for μ.
- What does objective prior mean?
- Ad hoc arguments give $\pi(\mu) \propto 1$.

Which Smells Fishier: Bayes or Fiducial?

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- t is our missing data with prior distribution: $t \sim t_{n-1}$.

Bayes

- *Objective* prior for μ.
- What does objective prior mean?
- Ad hoc arguments give $\pi(\mu) \propto 1$.

Fiducial

Which Smells Fishier: Bayes or Fiducial?

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- t is our missing data with prior distribution: $t \sim t_{n-1}$.

Bayes

- *Objective* prior for μ.
- What does objective prior mean?
- Ad hoc arguments give $\pi(\mu) \propto 1$.

Fiducial

- *Objective* posterior for t.

Which Smells Fishier: Bayes or Fiducial?

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- t is our missing data with prior distribution: $t \sim t_{n-1}$.

Bayes

- *Objective* prior for μ.
- What does objective prior mean?
- Ad hoc arguments give $\pi(\mu) \propto 1$.

Fiducial

- *Objective* posterior for t.
- What does objective posterior mean?

Which Smells Fishier: Bayes or Fiducial?

Inference for Mean of Normal

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} = t$$

- t is our missing data with prior distribution: $t \sim t_{n-1}$.

Bayes

- *Objective* prior for μ.
- What does objective prior mean?
- Ad hoc arguments give $\pi(\mu) \propto 1$.

Fiducial

- *Objective* posterior for t.
- What does objective posterior mean?
- **Ignore information on t in $(\bar{x}, s)$ that's tied to π.**

Which Smells Fishier: Bayes or Fiducial?

Inference for Mean of Normal

$$\frac{\bar{x}-\mu}{s/\sqrt{n}} = t$$

- t is our missing data with prior distribution: $t \sim t_{n-1}$.

Bayes

- *Objective* prior for μ.
- What does objective prior mean?
- Ad hoc arguments give $\pi(\mu) \propto 1$.

Fiducial

- *Objective* posterior for t.
- What does objective posterior mean?
- **Ignore information on t in $(\bar{x}, s)$ that's tied to π.**

- **Objective Posterior: Throw away data until we don't need a prior on μ.**

An Old Idea

An Old Idea

1. Obtain $f(\mathbf{U}|\mathbf{x})$ *without invoking* $\pi(d\theta)$.

An Old Idea

1. Obtain $f(\mathbf{U}|\mathbf{x})$ *without invoking* $\pi(d\theta)$.
2. Use structural relation $\mathbf{x} = G(\theta, \mathbf{U})$ to obtain $\pi(\theta|\mathbf{x})$ from $f(\mathbf{U}|\mathbf{x})$.

An Old Idea

1. Obtain $f(\mathbf{U}|\mathbf{x})$ *without invoking* $\pi(d\theta)$.
2. Use structural relation $\mathbf{x} = G(\theta, \mathbf{U})$ to obtain $\pi(\theta|\mathbf{x})$ from $f(\mathbf{U}|\mathbf{x})$.

A probability statement concerning $\bar{e}$ [the error] is ipso facto a probability statement concerning θ. (Fraser 1968)

An Old Idea

1. Obtain $f(\mathbf{U}|\mathbf{x})$ *without invoking* $\pi(d\theta)$.
2. Use structural relation $\mathbf{x} = G(\theta, \mathbf{U})$ to obtain $\pi(\theta|\mathbf{x})$ from $f(\mathbf{U}|\mathbf{x})$.

A probability statement concerning ē [the error] is ipso facto a probability statement concerning θ. (Fraser 1968)

One can get a random realization from the fiducial distribution of ξ by generating U and solving the structural equation for ξ. (Hannig 2009)

An Old Idea

1. Obtain $f(\mathbf{U}|\mathbf{x})$ *without invoking* $\pi(d\theta)$.
2. Use structural relation $\mathbf{x} = G(\theta, \mathbf{U})$ to obtain $\pi(\theta|\mathbf{x})$ from $f(\mathbf{U}|\mathbf{x})$.

A probability statement concerning $\bar{e}$ *[the error] is ipso facto a probability statement concerning* θ. (Fraser 1968)

One can get a random realization from the fiducial distribution of ξ *by generating* U *and solving the structural equation for* ξ. (Hannig 2009)

The key point is that knowing θ *is equivalent to knowing* $\mathbf{U}$*; in other words, inference on* θ *is equivalent to predicting the value of the unobserved* $\mathbf{U}$. (Martin, Zhang, and Liu 2010)

A Missing (Data) Perspective: Prior = Nuisance

A Missing (Data) Perspective: Prior = Nuisance

- The conditional distribution of $\mathbf{X}$ given $\mathbf{U}$ depends on π.

$$f(\mathbf{x}|\mathbf{U},\pi) = \int f(\mathbf{x}|\mathbf{U},\theta)\,\pi(d\theta) = \int 1\{\mathbf{x} = G(\theta,\mathbf{U})\}\,\pi(d\theta)$$

A Missing (Data) Perspective: Prior = Nuisance

- The conditional distribution of **X** given **U** depends on π.

$$f\left(\mathbf{x}|\mathbf{U},\pi\right)=\int f\left(\mathbf{x}|\mathbf{U},\theta\right)\pi\left(d\theta\right)=\int 1\left\{\mathbf{x}=G\left(\theta,\mathbf{U}\right)\right\}\pi\left(d\theta\right)$$

Predicting the Missing U

$$f\left(\mathbf{U}|\mathbf{x},\pi\right)\propto f\left(\mathbf{U}\right)f\left(\mathbf{x}|\mathbf{U},\pi\right).$$

A Missing (Data) Perspective: Prior = Nuisance

- The conditional distribution of $\mathbf{X}$ given $\mathbf{U}$ depends on π.

$$f(\mathbf{x}|\mathbf{U},\pi) = \int f(\mathbf{x}|\mathbf{U},\theta)\,\pi(d\theta) = \int 1\{\mathbf{x} = G(\theta,\mathbf{U})\}\,\pi(d\theta)$$

Predicting the Missing U

$$f(\mathbf{U}|\mathbf{x},\pi) \propto f(\mathbf{U})\, f(\mathbf{x}|\mathbf{U},\pi).$$

- We can treat $\pi(d\theta)$ as an **infinite dimensional nuisance parameter** in the "**U**-likelihood", $f(\mathbf{x}|\mathbf{U},\pi)$.

A Missing (Data) Perspective: Prior = Nuisance

- The conditional distribution of $\mathbf{X}$ given $\mathbf{U}$ depends on π.

$$f(\mathbf{x}|\mathbf{U},\pi) = \int f(\mathbf{x}|\mathbf{U},\theta)\,\pi(d\theta) = \int 1\{\mathbf{x} = G(\theta,\mathbf{U})\}\,\pi(d\theta)$$

Predicting the Missing U

$$f(\mathbf{U}|\mathbf{x},\pi) \propto f(\mathbf{U})\,f(\mathbf{x}|\mathbf{U},\pi).$$

- We can treat $\pi(d\theta)$ as an **infinite dimensional nuisance parameter** in the "**U**-likelihood", $f(\mathbf{x}|\mathbf{U},\pi)$.
- π can be viewed as a nuisance parameter **only if** we switch the problem from inference for θ to prediction of $\mathbf{U}$.

A Missing (Data) Perspective: Prior = Nuisance

- The conditional distribution of $\mathbf{X}$ given $\mathbf{U}$ depends on π.

$$f\left(\mathbf{x}|\mathbf{U},\pi\right) = \int f\left(\mathbf{x}|\mathbf{U},\theta\right)\pi\left(d\theta\right) = \int 1\left\{\mathbf{x} = G\left(\theta,\mathbf{U}\right)\right\}\pi\left(d\theta\right)$$

Predicting the Missing U

$$f\left(\mathbf{U}|\mathbf{x},\pi\right) \propto f\left(\mathbf{U}\right) f\left(\mathbf{x}|\mathbf{U},\pi\right).$$

- We can treat $\pi\left(d\theta\right)$ as an **infinite dimensional nuisance parameter** in the "**U**-likelihood", $f(\mathbf{x}|\mathbf{U},\pi)$.
- π can be viewed as a nuisance parameter **only if** we switch the problem from inference for θ to prediction of **U**.

Can we get to $\pi(\theta|\mathbf{x})$ without going through $\pi(\theta)$?

A Missing (Data) Perspective: Prior = Nuisance

- The conditional distribution of $\mathbf{X}$ given $\mathbf{U}$ depends on π.

$$f(\mathbf{x}|\mathbf{U},\pi) = \int f(\mathbf{x}|\mathbf{U},\theta)\,\pi(d\theta) = \int 1\{\mathbf{x} = G(\theta,\mathbf{U})\}\,\pi(d\theta)$$

Predicting the Missing U

$$f(\mathbf{U}|\mathbf{x},\pi) \propto f(\mathbf{U})\,f(\mathbf{x}|\mathbf{U},\pi).$$

- We can treat $\pi(d\theta)$ as an **infinite dimensional nuisance parameter** in the "**U**-likelihood", $f(\mathbf{x}|\mathbf{U},\pi)$.
- π can be viewed as a nuisance parameter **only if** we switch the problem from inference for θ to prediction of $\mathbf{U}$.

Can we get to $\pi(\theta|\mathbf{x})$ without going through $\pi(\theta)$?

↓

Can we predict **U** without any knowledge of the nuisance?

A Meaningful Definition of Objective

A Meaningful Definition of Objective

U-Likelihood

- $f(\mathbf{x}|\mathbf{U}, \pi)$ contains all information in $\mathbf{x}$ about $\mathbf{U}$ and π.

A Meaningful Definition of Objective

U-Likelihood

- $f(\mathbf{x}|\mathbf{U}, \pi)$ contains all information in $\mathbf{x}$ about $\mathbf{U}$ and π.

- **Goal**: Extract information on $\mathbf{U}$ not contaminated by the nuisance parameter π.

A Meaningful Definition of Objective

U-Likelihood

- $f(\mathbf{x}|\mathbf{U},\pi)$ contains all information in $\mathbf{x}$ about $\mathbf{U}$ and π.

- **Goal**: Extract information on $\mathbf{U}$ not contaminated by the nuisance parameter π.

- **Technique**: Throw away data contaminated by π. Reduce $\mathbf{x}$ to $A(\mathbf{x})$.

A Meaningful Definition of Objective

U-Likelihood

- $f(\mathbf{x}|\mathbf{U},\pi)$ contains all information in **x** about **U** and π.

- **Goal**: Extract information on **U** not contaminated by the nuisance parameter π.

- **Technique**: Throw away data contaminated by π. Reduce **x** to $A(\mathbf{x})$.

Objective Posterior for **U**

$f(\mathbf{U}|A(\mathbf{x}))$ is an objective posterior for **U** if it does not depend on the value of the nuisance parameter, π.

A Meaningful Definition of Objective

U-Likelihood

- $f(\mathbf{x}|\mathbf{U},\pi)$ contains all information in $\mathbf{x}$ about $\mathbf{U}$ and π.

- **Goal**: Extract information on $\mathbf{U}$ not contaminated by the nuisance parameter π.
- **Technique**: Throw away data contaminated by π. Reduce $\mathbf{x}$ to $A(\mathbf{x})$.

Objective Posterior for U

$f(\mathbf{U}|A(\mathbf{x}))$ is an objective posterior for $\mathbf{U}$ if it does not depend on the value of the nuisance parameter, π.

- Requires $f(A(\mathbf{x})|\mathbf{U},\pi) = f(A(\mathbf{x})|\mathbf{U})$.

Didn't You Say You Were Ancillary?

Didn't You Say You Were Ancillary?

- Does there exist a statistic, $A(\mathbf{x})$, s.t. $f(A(\mathbf{x})|\mathbf{U})$ is free of the nuisance π?

Didn't You Say You Were Ancillary?

- Does there exist a statistic, $A(\mathbf{x})$, s.t. $f(A(\mathbf{x})|\mathbf{U})$ is free of the nuisance π?
  - How about statistics ancillary for θ?

Didn't You Say You Were Ancillary?

- Does there exist a statistic, $A(\mathbf{x})$, s.t. $f(A(\mathbf{x})|\mathbf{U})$ is free of the nuisance π?
  - How about statistics ancillary for θ?

An Ancillarity Paradox from Basu (1964)

Didn't You Say You Were Ancillary?

- Does there exist a statistic, $A(\mathbf{x})$, s.t. $f(A(\mathbf{x})|\mathbf{U})$ is free of the nuisance π?
  - How about statistics ancillary for θ?

An Ancillarity Paradox from Basu (1964)

- Let (X, Y) have fiducial equation

$$X = \varepsilon_1 \text{ and } Y = \rho\varepsilon_1 + \sqrt{1-\rho}\varepsilon_2 \text{ where } \varepsilon_i \text{ i.i.d. } N(0,1)$$

Didn't You Say You Were Ancillary?

- Does there exist a statistic, $A(\mathbf{x})$, s.t. $f(A(\mathbf{x})|\mathbf{U})$ is free of the nuisance π?
  - How about statistics ancillary for θ?

An Ancillarity Paradox from Basu (1964)

- Let (X, Y) have fiducial equation

$$X = \varepsilon_1 \text{ and } Y = \rho\varepsilon_1 + \sqrt{1-\rho}\varepsilon_2 \text{ where } \varepsilon_i \text{ i.i.d. } N(0,1)$$

- X, Y are marginally ancillary but not jointly ancillary.

Didn't You Say You Were Ancillary?

- Does there exist a statistic, $A(\mathbf{x})$, s.t. $f(A(\mathbf{x})|\mathbf{U})$ is free of the nuisance π?
    - How about statistics ancillary for θ?

An Ancillarity Paradox from Basu (1964)

- Let (X, Y) have fiducial equation

$$X = \varepsilon_1 \text{ and } Y = \rho\varepsilon_1 + \sqrt{1-\rho}\varepsilon_2 \text{ where } \varepsilon_i \text{ i.i.d. } N(0,1)$$

- X, Y are marginally ancillary but not jointly ancillary.
- $P(X = x|\varepsilon, \rho) = 1\{x = \varepsilon_1\}$ is free of ρ.

Didn't You Say You Were Ancillary?

- Does there exist a statistic, $A(\mathbf{x})$, s.t. $f(A(\mathbf{x})|\mathbf{U})$ is free of the nuisance π?
  - How about statistics ancillary for θ?

An Ancillarity Paradox from Basu (1964)

- Let (X, Y) have fiducial equation

$$X = \varepsilon_1 \text{ and } Y = \rho\varepsilon_1 + \sqrt{1-\rho}\varepsilon_2 \text{ where } \varepsilon_i \text{ i.i.d. } N(0,1)$$

- X, Y are marginally ancillary but not jointly ancillary.
- $P(X = x|\varepsilon, \rho) = 1\{x = \varepsilon_1\}$ is free of ρ.
- $P(Y = y|\varepsilon, \rho) = 1\left\{y = \rho\varepsilon_1 + \sqrt{1-\rho}\varepsilon_2\right\}$ depends on ρ.

Didn't You Say You Were Ancillary?

- Does there exist a statistic, $A(\mathbf{x})$, s.t. $f(A(\mathbf{x})|\mathbf{U})$ is free of the nuisance π?
    - How about statistics ancillary for θ?

An Ancillarity Paradox from Basu (1964)

- Let (X, Y) have fiducial equation

$$X = \varepsilon_1 \text{ and } Y = \rho\varepsilon_1 + \sqrt{1-\rho}\varepsilon_2 \text{ where } \varepsilon_i \text{ i.i.d. } N(0,1)$$

- X, Y are marginally ancillary but not jointly ancillary.
- $P(X = x|\varepsilon, \rho) = 1\{x = \varepsilon_1\}$ is free of ρ.
- $P(Y = y|\varepsilon, \rho) = 1\{y = \rho\varepsilon_1 + \sqrt{1-\rho}\varepsilon_2\}$ depends on ρ.
- Situation reverses by switching the roles of X and Y in the fiducial equation.

Didn't You Say You Were Ancillary?

- Does there exist a statistic, $A(\mathbf{x})$, s.t. $f(A(\mathbf{x})|\mathbf{U})$ is free of the nuisance π?
    - How about statistics ancillary for θ?

An Ancillarity Paradox from Basu (1964)

- Let (X, Y) have fiducial equation

$$X = \varepsilon_1 \text{ and } Y = \rho\varepsilon_1 + \sqrt{1-\rho}\varepsilon_2 \text{ where } \varepsilon_i \text{ i.i.d. } N(0,1)$$

- X, Y are marginally ancillary but not jointly ancillary.
- $P(X = x|\varepsilon, \rho) = 1\{x = \varepsilon_1\}$ is free of ρ.
- $P(Y = y|\varepsilon, \rho) = 1\left\{y = \rho\varepsilon_1 + \sqrt{1-\rho}\varepsilon_2\right\}$ depends on ρ.
- Situation reverses by switching the roles of X and Y in the fiducial equation.
- Whether or not $A(\mathbf{x})$ is free of π given $\mathbf{U}$ depends on G.

Representational Ancillarity

Representational Ancillarity

Definition of R-Ancillarity

A statistic $A(\mathbf{x})$ is representationally ancillary w.r.t. G if there exists a representation A_G s.t. $A(G(\theta, \mathbf{U})) = A_G(\mathbf{U})\ \forall \theta$.

Representational Ancillarity

Definition of R-Ancillarity

A statistic $A(\mathbf{x})$ is representationally ancillary w.r.t. G if there exists a representation A_G s.t. $A(G(\theta, \mathbf{U})) = A_G(\mathbf{U})\ \forall\theta$.

- $A(\mathbf{x})$ is R-ancillary if and only if $\mathbf{U}|A(\mathbf{x}), \pi \sim \mathbf{U}|A(\mathbf{x})$.

Representational Ancillarity

Definition of R-Ancillarity

A statistic $A(\mathbf{x})$ is representationally ancillary w.r.t. G if there exists a representation A_G s.t. $A(G(\theta, \mathbf{U})) = A_G(\mathbf{U})\ \forall\theta$.

- $A(\mathbf{x})$ is R-ancillary if and only if $\mathbf{U}|A(\mathbf{x}), \pi \sim \mathbf{U}|A(\mathbf{x})$.

Lemma (also see Barnard, 1995)

If $A_1(\mathbf{x})$ and $A_2(\mathbf{x})$ are both R-ancillaries w.r.t. G, then they are jointly R-ancillary w.r.t. G.

Representational Ancillarity

Definition of R-Ancillarity

A statistic $A(\mathbf{x})$ is representationally ancillary w.r.t. G if there exists a representation A_G s.t. $A(G(\theta, \mathbf{U})) = A_G(\mathbf{U})\ \forall\theta$.

- $A(\mathbf{x})$ is R-ancillary if and only if $\mathbf{U}|A(\mathbf{x}), \pi \sim \mathbf{U}|A(\mathbf{x})$.

Lemma (also see Barnard, 1995)

If $A_1(\mathbf{x})$ and $A_2(\mathbf{x})$ are both R-ancillaries w.r.t. G, then they are jointly R-ancillary w.r.t. G.

- Write A_1, A_2 as $A_{1,G}(\mathbf{U}), A_{2,G}(\mathbf{U})$. Then $(A_{1,G}(\mathbf{U}), A_{2,G}(\mathbf{U}))$ remains a R-ancillary w.r.t. G.

Representational Ancillarity

Definition of R-Ancillarity

A statistic $A(\mathbf{x})$ is representationally ancillary w.r.t. G if there exists a representation A_G s.t. $A(G(\theta, \mathbf{U})) = A_G(\mathbf{U})\ \forall\theta$.

- $A(\mathbf{x})$ is R-ancillary if and only if $\mathbf{U}|A(\mathbf{x}), \pi \sim \mathbf{U}|A(\mathbf{x})$.

Lemma (also see Barnard, 1995)

If $A_1(\mathbf{x})$ and $A_2(\mathbf{x})$ are both R-ancillaries w.r.t. G, then they are jointly R-ancillary w.r.t. G.

- Write A_1, A_2 as $A_{1,G}(\mathbf{U}), A_{2,G}(\mathbf{U})$. Then $(A_{1,G}(\mathbf{U}), A_{2,G}(\mathbf{U}))$ remains a R-ancillary w.r.t. G.

- In Basu's paradox, X, Y are both ancillaries but they are not R-ancillaries with respect to the same G.

What Do Representations Buy Us?

What Do Representations Buy Us?

- Fiducial inference depends on G because whether a statistic is ancillary for π depends on whether it is R-ancillary for G.

What Do Representations Buy Us?

- Fiducial inference depends on G because whether a statistic is ancillary for π depends on whether it is R-ancillary for G.

- Definition of ancillarity as *distributional independence* of $A(\mathbf{x})$ from θ is insufficient.

What Do Representations Buy Us?

- Fiducial inference depends on G because whether a statistic is ancillary for π depends on whether it is R-ancillary for G.
- Definition of ancillarity as *distributional independence* of $A(\mathbf{x})$ from θ is insufficient.
- We require the notion of *representational* (or functional) *independence*.

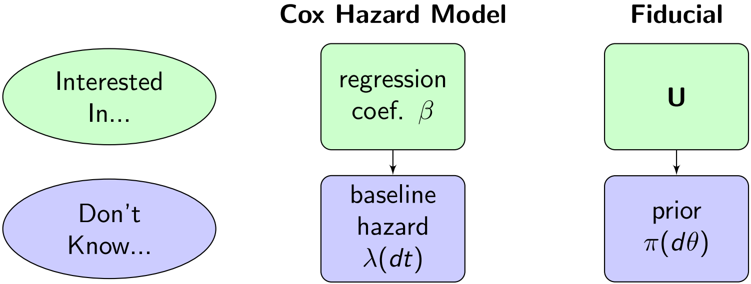
Is Fiducial Hazardous?

	Cox Hazard Model	Fiducial

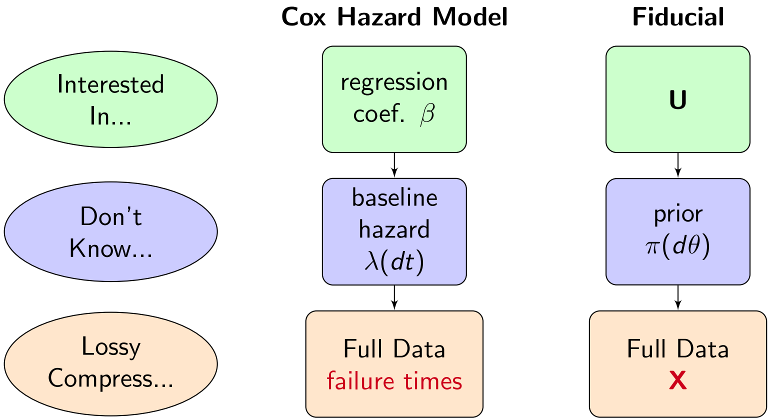
Is Fiducial Hazardous?

	Cox Hazard Model	**Fiducial**
Interested In...	regression coef. β	**U**

Is Fiducial Hazardous?



Is Fiducial Hazardous?



Is Fiducial Hazardous?

	Cox Hazard Model	**Fiducial**
Interested In...	regression coef. β	$\mathbf{U}$
Don't Know...	baseline hazard $\lambda(dt)$	prior $\pi(d\theta)$
Lossy Compress...	Full Data failure times	Full Data $\mathbf{X}$
Inference From...	λ ancillary β non-ancillary ranks	π ancillary $\mathbf{U}$ non-ancillary $A(\mathbf{x})$

Fiducial Is Just Partial Likelihood

Partial Likelihood for **U**

Fiducial Is Just Partial Likelihood

Partial Likelihood for U

- Assume the decomposition, $\mathbf{x} = (T(\mathbf{x}), A(\mathbf{x}))$ where $A(\mathbf{x})$ is R-ancillary w.r.t. G.

Fiducial Is Just Partial Likelihood

Partial Likelihood for $\mathbf{U}$

- Assume the decomposition, $\mathbf{x} = (T(\mathbf{x}), A(\mathbf{x}))$ where $A(\mathbf{x})$ is R-ancillary w.r.t. G.
- The joint $(\mathbf{U}, \pi)$-likelihood factors as

$$f(\mathbf{x}|\mathbf{U}, \pi) = f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi) f(A(\mathbf{x})|\mathbf{U}).$$

Fiducial Is Just Partial Likelihood

Partial Likelihood for U

- Assume the decomposition, $\mathbf{x} = (T(\mathbf{x}), A(\mathbf{x}))$ where $A(\mathbf{x})$ is R-ancillary w.r.t. G.
- The joint $(\mathbf{U}, \pi)$-likelihood factors as

$$f(\mathbf{x}|\mathbf{U}, \pi) = f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi) f(A(\mathbf{x})|\mathbf{U}).$$

- $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ requires prior specification.

Fiducial Is Just Partial Likelihood

Partial Likelihood for U

- Assume the decomposition, $\mathbf{x} = (T(\mathbf{x}), A(\mathbf{x}))$ where $A(\mathbf{x})$ is R-ancillary w.r.t. G.
- The joint $(\mathbf{U}, \pi)$-likelihood factors as

$$f(\mathbf{x}|\mathbf{U}, \pi) = f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi) f(A(\mathbf{x})|\mathbf{U}).$$

- $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ requires prior specification.
- $f(A(\mathbf{x})|\mathbf{U})$ is known independent of the prior, π.

Fiducial Is Just Partial Likelihood

Partial Likelihood for $\mathbf{U}$

- Assume the decomposition, $\mathbf{x} = (T(\mathbf{x}), A(\mathbf{x}))$ where $A(\mathbf{x})$ is R-ancillary w.r.t. G.
- The joint $(\mathbf{U}, \pi)$-likelihood factors as

$$f(\mathbf{x}|\mathbf{U}, \pi) = f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi) f(A(\mathbf{x})|\mathbf{U}).$$

- $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ requires prior specification.
- $f(A(\mathbf{x})|\mathbf{U})$ is known independent of the prior, π.

Fiducial Recipe for $\mathbf{U}|\mathbf{x}$

Fiducial Is Just Partial Likelihood

Partial Likelihood for $\mathbf{U}$

- Assume the decomposition, $\mathbf{x} = (T(\mathbf{x}), A(\mathbf{x}))$ where $A(\mathbf{x})$ is R-ancillary w.r.t. G.
- The joint $(\mathbf{U}, \pi)$-likelihood factors as

$$f(\mathbf{x}|\mathbf{U}, \pi) = f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi) f(A(\mathbf{x})|\mathbf{U}).$$

- $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ requires prior specification.
- $f(A(\mathbf{x})|\mathbf{U})$ is known independent of the prior, π.

Fiducial Recipe for $\mathbf{U}|\mathbf{x}$

1. Ignore factor $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ because information on $\mathbf{U}$ "cannot be disentangled" from prior.

Fiducial Is Just Partial Likelihood

Partial Likelihood for U

- Assume the decomposition, $\mathbf{x} = (T(\mathbf{x}), A(\mathbf{x}))$ where $A(\mathbf{x})$ is R-ancillary w.r.t. G.
- The joint $(\mathbf{U}, \pi)$-likelihood factors as

$$f(\mathbf{x}|\mathbf{U}, \pi) = f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi) f(A(\mathbf{x})|\mathbf{U}).$$

- $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ requires prior specification.
- $f(A(\mathbf{x})|\mathbf{U})$ is known independent of the prior, π.

Fiducial Recipe for $\mathbf{U}|\mathbf{x}$

1. Ignore factor $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ because information on $\mathbf{U}$ "cannot be disentangled" from prior.
2. Use $f(A(\mathbf{x})|\mathbf{U})$ to obtain $f(\mathbf{U}|A(\mathbf{x})) \propto f(\mathbf{U}) f(A(\mathbf{x})|\mathbf{U})$.

Fiducial Is Just Partial Likelihood

Partial Likelihood for **U**

- Assume the decomposition, $\mathbf{x} = (T(\mathbf{x}), A(\mathbf{x}))$ where $A(\mathbf{x})$ is R-ancillary w.r.t. G.
- The joint $(\mathbf{U}, \pi)$-likelihood factors as

$$f(\mathbf{x}|\mathbf{U}, \pi) = f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi) f(A(\mathbf{x})|\mathbf{U}).$$

- $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ requires prior specification.
- $f(A(\mathbf{x})|\mathbf{U})$ is known independent of the prior, π.

Fiducial Recipe for $\mathbf{U}|\mathbf{x}$

1. Ignore factor $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ because information on $\mathbf{U}$ "cannot be disentangled" from prior.
2. Use $f(A(\mathbf{x})|\mathbf{U})$ to obtain $f(\mathbf{U}|A(\mathbf{x})) \propto f(\mathbf{U}) f(A(\mathbf{x})|\mathbf{U})$.
3. Pretend $f(\mathbf{U}|A(\mathbf{x})) = f(\mathbf{U}|\mathbf{x})$.

Fiducial Is Just Partial Likelihood

Partial Likelihood for U

- Assume the decomposition, $\mathbf{x} = (T(\mathbf{x}), A(\mathbf{x}))$ where $A(\mathbf{x})$ is R-ancillary w.r.t. G.
- The joint $(\mathbf{U}, \pi)$-likelihood factors as

$$f(\mathbf{x}|\mathbf{U}, \pi) = f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi) f(A(\mathbf{x})|\mathbf{U}).$$

- $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ requires prior specification.
- $f(A(\mathbf{x})|\mathbf{U})$ is known independent of the prior, π.

Fiducial Recipe for U|x

1. Ignore factor $f(T(\mathbf{x})|A(\mathbf{x}), \mathbf{U}, \pi)$ because information on $\mathbf{U}$ "cannot be disentangled" from prior.
2. Use $f(A(\mathbf{x})|\mathbf{U})$ to obtain $f(\mathbf{U}|A(\mathbf{x})) \propto f(\mathbf{U}) f(A(\mathbf{x})|\mathbf{U})$.
3. Pretend $f(\mathbf{U}|A(\mathbf{x})) = f(\mathbf{U}|\mathbf{x})$. Refreshing or Revolting?

When the Recipe Works...

When the Recipe Works...

Exponential Hyperbola

- $V = \frac{1}{\lambda}E_1$ and $W = \lambda E_2$ where E_1, E_2 are iid exponential.

When the Recipe Works...

Exponential Hyperbola

- $V = \frac{1}{\lambda} E_1$ and $W = \lambda E_2$ where E_1, E_2 are iid exponential.
- $\hat{\lambda} = (W/V)^{1/2}$ is the MLE and $A = (VW)^{1/2}$ is an R-ancillary.

When the Recipe Works...

Exponential Hyperbola

- $V = \frac{1}{\lambda}E_1$ and $W = \lambda E_2$ where E_1, E_2 are iid exponential.
- $\hat{\lambda} = (W/V)^{1/2}$ is the MLE and $A = (VW)^{1/2}$ is an R-ancillary.
- Solve $\hat{\lambda} = \lambda\,(E_2/E_1)^{1/2}$ to obtain $\lambda = \hat{\lambda}\,(E_1/E_2)^{1/2}$.

When the Recipe Works...

Exponential Hyperbola

- $V = \frac{1}{\lambda} E_1$ and $W = \lambda E_2$ where E_1, E_2 are iid exponential.
- $\hat{\lambda} = (W/V)^{1/2}$ is the MLE and $A = (VW)^{1/2}$ is an R-ancillary.
- Solve $\hat{\lambda} = \lambda (E_2/E_1)^{1/2}$ to obtain $\lambda = \hat{\lambda} (E_1/E_2)^{1/2}$.
- Impute $(E_2/E_1)^{1/2}$ according to its prior distribution.

When the Recipe Works...

Exponential Hyperbola

- $V = \frac{1}{\lambda} E_1$ and $W = \lambda E_2$ where E_1, E_2 are iid exponential.
- $\hat{\lambda} = (W/V)^{1/2}$ is the MLE and $A = (VW)^{1/2}$ is an R-ancillary.
- Solve $\hat{\lambda} = \lambda (E_2/E_1)^{1/2}$ to obtain $\lambda = \hat{\lambda} (E_1/E_2)^{1/2}$.
- Impute $(E_2/E_1)^{1/2}$ according to its prior distribution. *Parallels using marginal likelihood $f(\hat{\lambda}|\lambda)$.*

When the Recipe Works...

Exponential Hyperbola

- $V = \frac{1}{\lambda} E_1$ and $W = \lambda E_2$ where E_1, E_2 are iid exponential.
- $\hat{\lambda} = (W/V)^{1/2}$ is the MLE and $A = (VW)^{1/2}$ is an R-ancillary.
- Solve $\hat{\lambda} = \lambda (E_2/E_1)^{1/2}$ to obtain $\lambda = \hat{\lambda} (E_1/E_2)^{1/2}$.
- Impute $(E_2/E_1)^{1/2}$ according to its prior distribution. *Parallels using marginal likelihood $f(\hat{\lambda}|\lambda)$.*
- But $\hat{\lambda}$ is not sufficient. Can we do better?

When the Recipe Works...

Exponential Hyperbola

- $V = \frac{1}{\lambda} E_1$ and $W = \lambda E_2$ where E_1, E_2 are iid exponential.
- $\hat{\lambda} = (W/V)^{1/2}$ is the MLE and $A = (VW)^{1/2}$ is an R-ancillary.
- Solve $\hat{\lambda} = \lambda (E_2/E_1)^{1/2}$ to obtain $\lambda = \hat{\lambda} (E_1/E_2)^{1/2}$.
- Impute $(E_2/E_1)^{1/2}$ according to its prior distribution. *Parallels using marginal likelihood $f(\hat{\lambda}|\lambda)$.*
- But $\hat{\lambda}$ is not sufficient. Can we do better?
- We observe $A = (E_1 E_2)^{1/2}$. Impute $(E_2/E_1)^{1/2}$ according to $f(E_1, E_2|A)$.

When the Recipe Works...

Exponential Hyperbola

- $V = \frac{1}{\lambda}E_1$ and $W = \lambda E_2$ where E_1, E_2 are iid exponential.
- $\hat{\lambda} = (W/V)^{1/2}$ is the MLE and $A = (VW)^{1/2}$ is an R-ancillary.
- Solve $\hat{\lambda} = \lambda(E_2/E_1)^{1/2}$ to obtain $\lambda = \hat{\lambda}(E_1/E_2)^{1/2}$.
- Impute $(E_2/E_1)^{1/2}$ according to its prior distribution. *Parallels using marginal likelihood $f(\hat{\lambda}|\lambda)$.*
- But $\hat{\lambda}$ is not sufficient. Can we do better?
- We observe $A = (E_1E_2)^{1/2}$. Impute $(E_2/E_1)^{1/2}$ according to $f(E_1, E_2|A)$. *Parallels using conditional likelihood $f(\hat{\lambda}|A, \lambda)$.*

When the Recipe Works...

Exponential Hyperbola

- $V = \frac{1}{\lambda}E_1$ and $W = \lambda E_2$ where E_1, E_2 are iid exponential.
- $\hat{\lambda} = (W/V)^{1/2}$ is the MLE and $A = (VW)^{1/2}$ is an R-ancillary.
- Solve $\hat{\lambda} = \lambda(E_2/E_1)^{1/2}$ to obtain $\lambda = \hat{\lambda}(E_1/E_2)^{1/2}$.
- Impute $(E_2/E_1)^{1/2}$ according to its prior distribution. *Parallels using marginal likelihood $f(\hat{\lambda}|\lambda)$.*
- But $\hat{\lambda}$ is not sufficient. Can we do better?
- We observe $A = (E_1E_2)^{1/2}$. Impute $(E_2/E_1)^{1/2}$ according to $f(E_1, E_2|A)$. *Parallels using conditional likelihood $f(\hat{\lambda}|A, \lambda)$.*

- Why does conditioning on ancillary statistics recover second order information?

When the Recipe Works...

Exponential Hyperbola

- $V = \frac{1}{\lambda}E_1$ and $W = \lambda E_2$ where E_1, E_2 are iid exponential.
- $\hat{\lambda} = (W/V)^{1/2}$ is the MLE and $A = (VW)^{1/2}$ is an R-ancillary.
- Solve $\hat{\lambda} = \lambda\,(E_2/E_1)^{1/2}$ to obtain $\lambda = \hat{\lambda}\,(E_1/E_2)^{1/2}$.
- Impute $(E_2/E_1)^{1/2}$ according to its prior distribution. *Parallels using marginal likelihood $f(\hat{\lambda}|\lambda)$.*
- But $\hat{\lambda}$ is not sufficient. Can we do better?
- We observe $A = (E_1E_2)^{1/2}$. Impute $(E_2/E_1)^{1/2}$ according to $f(E_1, E_2|A)$. *Parallels using conditional likelihood $f(\hat{\lambda}|A, \lambda)$.*

- Why does conditioning on ancillary statistics recover second order information?
- **Fiducial Answer**: It increases our efficiency of predicting $\mathbf{U}$.

An Often Forgotten Ingredient in Fiducial Cooking

An Often Forgotten Ingredient in Fiducial Cooking

- X has the following fiducial equation

$$X = G(\theta, U) = \begin{cases} 0 & 0 \leq U < \theta \\ 1 & \theta \leq U < 1 \end{cases}$$

where $U \sim$Unif$[0,1]$ and $\theta \in [1/4, 1/2]$.

An Often Forgotten Ingredient in Fiducial Cooking

- X has the following fiducial equation

$$X = G(\theta, U) = \begin{cases} 0 & 0 \leq U < \theta \\ 1 & \theta \leq U < 1 \end{cases}$$

where $U \sim$Unif$[0, 1]$ and $\theta \in [1/4, 1/2]$.

- **Key Question: Without $\pi(\theta)$, what do we know about U? What is the free information in X about U?**

An Often Forgotten Ingredient in Fiducial Cooking

- X has the following fiducial equation

$$X = G(\theta, U) = \begin{cases} 0 & 0 \le U < \theta \\ 1 & \theta \le U < 1 \end{cases}$$

where $U \sim$Unif$[0, 1]$ and $\theta \in [1/4, 1/2]$.

- **Key Question: Without $\pi(\theta)$, what do we know about U? What is the free information in X about U?**

The only additional information available to us is the fact that the value of U and $\mathbf{x}$ *must be compatible.* (Hannig 2009)

An Often Forgotten Ingredient in Fiducial Cooking

- X has the following fiducial equation

$$X = G(\theta, U) = \begin{cases} 0 & 0 \leq U < \theta \\ 1 & \theta \leq U < 1 \end{cases}$$

where $U \sim$Unif$[0,1]$ and $\theta \in [1/4, 1/2]$.

- **Key Question: Without $\pi(\theta)$, what do we know about U? What is the free information in X about U?**

The only additional information available to us is the fact that the value of U and $\mathbf{x}$ must be compatible. (Hannig 2009)

- If $X = 0$, $U \in [0, 1/2]$, shall we predict U as Unif$[0, 1/2]$?

An Often Forgotten Ingredient in Fiducial Cooking

- X has the following fiducial equation

$$X = G(\theta, U) = \begin{cases} 0 & 0 \leq U < \theta \\ 1 & \theta \leq U < 1 \end{cases}$$

where $U \sim$Unif$[0, 1]$ and $\theta \in [1/4, 1/2]$.

- **Key Question: Without $\pi(\theta)$, what do we know about U? What is the free information in X about U?**

The only additional information available to us is the fact that the value of U and $\mathbf{x}$ must be compatible. (Hannig 2009)

- If $X = 0$, $U \in [0, 1/2]$, shall we predict U as Unif$[0, 1/2]$?
If $X = 1$, $U \in [1/4, 1]$, shall we predict U as Unif$[1/4, 1]$?

An Often Forgotten Ingredient in Fiducial Cooking

- X has the following fiducial equation

$$X = G(\theta, U) = \begin{cases} 0 & 0 \leq U < \theta \\ 1 & \theta \leq U < 1 \end{cases}$$

where $U \sim$Unif$[0, 1]$ and $\theta \in [1/4, 1/2]$.

- **Key Question: Without $\pi(\theta)$, what do we know about U? What is the free information in X about U?**

  The only additional information available to us is the fact that the value of U and $\mathbf{x}$ must be compatible. (Hannig 2009)

- If $X = 0$, $U \in [0, 1/2]$, shall we predict U as Unif$[0, 1/2]$?
  If $X = 1$, $U \in [1/4, 1]$, shall we predict U as Unif$[1/4, 1]$?

- **What are we forgetting by doing this?**

We Observe and Observe that We Observe

We Observe and Observe that We Observe

x-Compatible Regions

We Observe and Observe that We Observe

$\mathbf{x}$-Compatible Regions

- If we generate data $\mathbf{x} = G(\theta_0, \mathbf{U}_0)$, we learn that $\mathbf{U}_0$ resides in

$$\mathcal{M}(G(\theta_0, \mathbf{U}_0)) = \{\mathbf{U} : \exists \theta \in \Theta \text{ s.t. } G(\theta_0, \mathbf{U}_0) = G(\theta, \mathbf{U})\}.$$

We Observe and Observe that We Observe

x-Compatible Regions

- If we generate data $\mathbf{x} = G(\theta_0, \mathbf{U}_0)$, we learn that $\mathbf{U}_0$ resides in

$$\mathcal{M}(G(\theta_0, \mathbf{U}_0)) = \{\mathbf{U} : \exists \theta \in \Theta \text{ s.t. } G(\theta_0, \mathbf{U}_0) = G(\theta, \mathbf{U})\}.$$

- **Hannig (2009) assumes that**

$$f(\mathbf{U}|\mathcal{M}(\mathbf{X}) = \mathcal{M}(\mathbf{x})) = f(\mathbf{U}|\mathbf{U} \in \mathcal{M}(\mathbf{x})),$$

hence does not depend on π.

We Observe and Observe that We Observe

x-Compatible Regions

- If we generate data $\mathbf{x} = G(\theta_0, \mathbf{U}_0)$, we learn that $\mathbf{U}_0$ resides in

$$\mathcal{M}(G(\theta_0, \mathbf{U}_0)) = \{\mathbf{U} : \exists\theta \in \Theta \text{ s.t. } G(\theta_0, \mathbf{U}_0) = G(\theta, \mathbf{U})\}.$$

- **Hannig (2009) assumes that**

$$f(\mathbf{U}|\mathcal{M}(\mathbf{X}) = \mathcal{M}(\mathbf{x})) = f(\mathbf{U}|\mathbf{U} \in \mathcal{M}(\mathbf{x})),$$

hence does not depend on π.

- The event $\{\mathcal{M}(\mathbf{X}) = \mathcal{M}(\mathbf{x})\}$ contains two pieces of information:

We Observe and Observe that We Observe

x-Compatible Regions

- If we generate data $\mathbf{x} = G(\theta_0, \mathbf{U}_0)$, we learn that $\mathbf{U}_0$ resides in

$$\mathcal{M}(G(\theta_0, \mathbf{U}_0)) = \{\mathbf{U} : \exists \theta \in \Theta \text{ s.t. } G(\theta_0, \mathbf{U}_0) = G(\theta, \mathbf{U})\}.$$

- **Hannig (2009) assumes that**

$$f(\mathbf{U}|\mathcal{M}(\mathbf{X}) = \mathcal{M}(\mathbf{x})) = f(\mathbf{U}|\mathbf{U} \in \mathcal{M}(\mathbf{x})),$$

hence does not depend on π.

- The event $\{\mathcal{M}(\mathbf{X}) = \mathcal{M}(\mathbf{x})\}$ contains two pieces of information:
  1. $\mathbf{U}_0 \in \mathcal{M}(\mathbf{x})$.

We Observe and Observe that We Observe

x-Compatible Regions

- If we generate data $\mathbf{x} = G(\theta_0, \mathbf{U}_0)$, we learn that $\mathbf{U}_0$ resides in

$$\mathcal{M}(G(\theta_0, \mathbf{U}_0)) = \{\mathbf{U} : \exists \theta \in \Theta \text{ s.t. } G(\theta_0, \mathbf{U}_0) = G(\theta, \mathbf{U})\}.$$

- **Hannig (2009) assumes that**

$$f(\mathbf{U}|\mathcal{M}(\mathbf{X}) = \mathcal{M}(\mathbf{x})) = f(\mathbf{U}|\mathbf{U} \in \mathcal{M}(\mathbf{x})),$$

hence does not depend on π.

- The event $\{\mathcal{M}(\mathbf{X}) = \mathcal{M}(\mathbf{x})\}$ contains two pieces of information:
  1. $\mathbf{U_0} \in \mathcal{M}(\mathbf{x})$.
  2. **How** we came to observe $\mathbf{U_0} \in \mathcal{M}(\mathbf{x})$.

We Observe and Observe that We Observe

x-Compatible Regions

- If we generate data $\mathbf{x} = G(\theta_0, \mathbf{U}_0)$, we learn that $\mathbf{U}_0$ resides in

$$\mathcal{M}(G(\theta_0, \mathbf{U}_0)) = \{\mathbf{U} : \exists \theta \in \Theta \text{ s.t. } G(\theta_0, \mathbf{U}_0) = G(\theta, \mathbf{U})\}.$$

- **Hannig (2009) assumes that**

$$f(\mathbf{U}|\mathcal{M}(\mathbf{X}) = \mathcal{M}(\mathbf{x})) = f(\mathbf{U}|\mathbf{U} \in \mathcal{M}(\mathbf{x})),$$

hence does not depend on π.

- The event $\{\mathcal{M}(\mathbf{X}) = \mathcal{M}(\mathbf{x})\}$ contains two pieces of information:
  1. $\mathbf{U_0} \in \mathcal{M}(\mathbf{x})$.
  2. **How** we came to observe $\mathbf{U_0} \in \mathcal{M}(\mathbf{x})$.
- To correct use the information in (1), **we need to condition on the how**, i.e., condition on the observation process.

Bernoulli with Restricted Parameter Space

Bernoulli with Restricted Parameter Space

- Reminder: $U \sim \text{Unif}[0,1]$ and $\theta \in [1/4, 1/2]$

$$X = G(\theta, U) = \begin{cases} 0 & 0 \leq U < \theta \\ 1 & \theta \leq U < 1 \end{cases}$$

Bernoulli with Restricted Parameter Space

- Reminder: $U\sim$Unif$[0,1]$ and $\theta\in[1/4,1/2]$

$$X=G\left(\theta,U\right)=\begin{cases}0 & 0\leq U<\theta\\ 1 & \theta\leq U<1\end{cases}$$

- $\mathcal{M}\left(0\right)=[0,1/2]$, $\mathcal{M}\left(1\right)=[1/4,1]$.

Bernoulli with Restricted Parameter Space

- Reminder: $U \sim$Unif$[0,1]$ and $\theta \in [1/4, 1/2]$

$$X = G(\theta, U) = \begin{cases} 0 & 0 \leq U < \theta \\ 1 & \theta \leq U < 1 \end{cases}$$

- $\mathcal{M}(0) = [0, 1/2]$, $\mathcal{M}(1) = [1/4, 1]$.

- $\mathbb{I}\{B\}$ is indicator function. Let $\mathbb{I}_x \equiv \mathbb{I}\{U \in \mathcal{M}(x)\}$ for $x = 0, 1$.

Bernoulli with Restricted Parameter Space

- Reminder: $U \sim$Unif$[0,1]$ and $\theta \in [1/4, 1/2]$

$$X = G(\theta, U) = \begin{cases} 0 & 0 \le U < \theta \\ 1 & \theta \le U < 1 \end{cases}$$

- $\mathcal{M}(0) = [0, 1/2]$, $\mathcal{M}(1) = [1/4, 1]$.

- $\mathbb{I}\{B\}$ is indicator function. Let $\mathbb{I}_x \equiv \mathbb{I}\{U \in \mathcal{M}(x)\}$ for $x = 0, 1$.
- We **do not observe** $\mathbb{I}_x$. It must be **inferred** from data.

Bernoulli with Restricted Parameter Space

- Reminder: $U \sim$Unif$[0,1]$ and $\theta \in [1/4, 1/2]$

$$X = G(\theta, U) = \begin{cases} 0 & 0 \leq U < \theta \\ 1 & \theta \leq U < 1 \end{cases}$$

- $\mathcal{M}(0) = [0, 1/2]$, $\mathcal{M}(1) = [1/4, 1]$.

- $\mathbb{I}\{B\}$ is indicator function. Let $\mathbb{I}_x \equiv \mathbb{I}\{U \in \mathcal{M}(x)\}$ for $x = 0, 1$.
- We **do not observe** $\mathbb{I}_x$. It must be **inferred** from data.

What We Actually Observe

- $\mathbb{I}^{obs} = X\mathbb{I}_1 + (1-X)\mathbb{I}_0$ (Note: $\mathbb{I}^{obs} = 1$)

Bernoulli with Restricted Parameter Space

- Reminder: $U \sim$Unif$[0,1]$ and $\theta \in [1/4, 1/2]$

$$X = G(\theta, U) = \begin{cases} 0 & 0 \leq U < \theta \\ 1 & \theta \leq U < 1 \end{cases}$$

- $\mathcal{M}(0) = [0, 1/2]$, $\mathcal{M}(1) = [1/4, 1]$.

- $\mathbb{I}\{B\}$ is indicator function. Let $\mathbb{I}_x \equiv \mathbb{I}\{U \in \mathcal{M}(x)\}$ for $x = 0, 1$.
- We **do not observe** $\mathbb{I}_x$. It must be **inferred** from data.

What We Actually Observe

- $\mathbb{I}^{obs} = X\mathbb{I}_1 + (1-X)\mathbb{I}_0$ (Note: $\mathbb{I}^{obs} = 1$)
- For $x \in \mathbb{X}$, define

$$O_x = \begin{cases} 1 & \text{if } \mathbb{I}^{obs} = \mathbb{I}_x \\ 0 & \text{otherwise} \end{cases}$$

A Sleight of Hand

A Sleight of Hand

- $\mathcal{M}(X)$ contains the information $\mathbb{I}_x$ through $\mathbb{I}^{obs}$ **and** $\{O_x\}$.

$$f\left(U|\mathcal{M}\left(X\right),\pi\right) = f\left(U|\mathbb{I}^{obs}, O_0, O_1, \pi\right)$$

A Sleight of Hand

- $\mathcal{M}(X)$ contains the information $\mathbb{I}_x$ through $\mathbb{I}^{obs}$ **and** $\{O_x\}$.

$$f\left(U|\mathcal{M}\left(X\right),\pi\right) = f\left(U|\mathbb{I}^{obs}, O_0, O_1, \pi\right)$$

- Rewrite the posterior using the law of total probability

$$\sum_{t\in\{0,1\}} f\left(U|\mathbb{I}_x = t, O_0, O_1, \pi\right) P\left(\mathbb{I}_x = t|\mathbb{I}^{obs}, O_0, O_1\right)$$

A Sleight of Hand

- $\mathcal{M}(X)$ contains the information $\mathbb{I}_x$ through $\mathbb{I}^{obs}$ **and** $\{O_x\}$.

$$f\left(U|\mathcal{M}\left(X\right),\pi\right) = f\left(U|\mathbb{I}^{obs}, O_0, O_1, \pi\right)$$

- Rewrite the posterior using the law of total probability

$$\sum_{t\in\{0,1\}} f\left(U|\mathbb{I}_x = t, O_0, O_1, \pi\right) P\left(\mathbb{I}_x = t|\mathbb{I}^{obs}, O_0, O_1\right)$$

- **Problem:** The distribution of O_0, O_1 depends on π.

A Sleight of Hand

- $\mathcal{M}(X)$ contains the information $\mathbb{I}_x$ through $\mathbb{I}^{obs}$ **and** $\{O_x\}$.

$$f\left(U|\mathcal{M}\left(X\right),\pi\right) = f\left(U|\mathbb{I}^{obs}, O_0, O_1, \pi\right)$$

- Rewrite the posterior using the law of total probability

$$\sum_{t\in\{0,1\}} f\left(U|\mathbb{I}_x = t, O_0, O_1, \pi\right) P\left(\mathbb{I}_x = t|\mathbb{I}^{obs}, O_0, O_1\right)$$

- **Problem:** The distribution of O_0, O_1 depends on π.

The Sleight of Hand

- To remove the dependence on π, make the substitution

$$f\left(U|\mathbb{I}_x = t, O_0, O_1, \pi\right) = f\left(U|\mathbb{I}_x = t\right)$$

A Sleight of Hand

- $\mathcal{M}(X)$ contains the information $\mathbb{I}_x$ through $\mathbb{I}^{obs}$ **and** $\{O_x\}$.

$$f\left(U|\mathcal{M}\left(X\right),\pi\right) = f\left(U|\mathbb{I}^{obs}, O_0, O_1, \pi\right)$$

- Rewrite the posterior using the law of total probability

$$\sum_{t\in\{0,1\}} f\left(U|\mathbb{I}_x = t, O_0, O_1, \pi\right) P\left(\mathbb{I}_x = t|\mathbb{I}^{obs}, O_0, O_1\right)$$

- **Problem:** The distribution of O_0, O_1 depends on π.

The Sleight of Hand

- To remove the dependence on π, make the substitution

$$f\left(U|\mathbb{I}_x = t, O_0, O_1, \pi\right) = f\left(U|\mathbb{I}_x = t\right)$$

- We **ignore how we learned** $\{\mathbb{I}_x = t\}$.

$\mathcal{O}$, We Just Ignored You!

O, We Just Ignored You!

A Mismatch

- Using the sleight of hand, rewrite the now π-free posterior

O, We Just Ignored You!

A Mismatch

- Using the sleight of hand, rewrite the now π-free posterior

$$\sum_{t \in \{0,1\}} f\left(U | \mathbb{I}_x = t\right) P\left(\mathbb{I}_x = t | \mathbb{I}^{obs}, O_0, O_1\right)$$

O, We Just Ignored You!

A Mismatch

- Using the sleight of hand, rewrite the now π-free posterior

$$\sum_{t \in \{0,1\}} f\left(U | \mathbb{I}_x = t\right) P\left(\mathbb{I}_x = t | \mathbb{I}^{obs}, O_0, O_1\right)$$

- We ignored O_0, O_1 in the first term **BUT** continue to condition on it in the second term—**incoherent**.

O, We Just Ignored You!

A Mismatch

- Using the sleight of hand, rewrite the now π-free posterior

$$\sum_{t\in\{0,1\}} f\left(U|\mathbb{I}_x = t\right) P\left(\mathbb{I}_x = t|\mathbb{I}^{obs}, O_0, O_1\right)$$

- We ignored O_0, O_1 in the first term **BUT** continue to condition on it in the second term—**incoherent**.
- It is this incoherence which leads to the incorrect assumption

$$f(U|\mathcal{M}(X) = \mathcal{M}(x), \pi) = f(U|U \in \mathcal{M}(x))$$

O, We Just Ignored You!

A Mismatch

- Using the sleight of hand, rewrite the now π-free posterior

$$\sum_{t\in\{0,1\}} f\left(U|\mathbb{I}_x = t\right) P\left(\mathbb{I}_x = t|\mathbb{I}^{obs}, O_0, O_1\right)$$

- We ignored O_0, O_1 in the first term **BUT** continue to condition on it in the second term—**incoherent**.
- It is this incoherence which leads to the incorrect assumption

$$f(U|\mathcal{M}(X) = \mathcal{M}(x), \pi) = f(U|U \in \mathcal{M}(x))$$

Confidence Validity (Rubin 1976)

- The posterior $f(U|U \in \mathcal{M}(x))$ leads to valid confidence regions for U if the observation process is ignorable

$$P\left(\mathbb{I}_x = 1|O_0 = o_0, O_1 = o_1, U, \pi\right) = P\left(\mathbb{I}_x = 1|U\right)$$

A Free Lunch? Take a Second Look.

A Free Lunch? Take a Second Look.

- Suppose we observe $\mathbb{I}_0 = 1$ and $\mathbb{I}_1$ is missing.

A Free Lunch? Take a Second Look.

- Suppose we observe $\mathbb{I}_0 = 1$ and $\mathbb{I}_1$ is missing.
- Condition for confidence validity does not hold:

A Free Lunch? Take a Second Look.

- Suppose we observe $\mathbb{I}_0 = 1$ and $\mathbb{I}_1$ is missing.
- Condition for confidence validity does not hold:

$$P\left(\mathbb{I}_0 = 1 | U, \theta\right) = \mathbb{I}\left\{U \leq 1/2\right\}$$

$$P\left(\mathbb{I}_0 = 1 | O_0 = 1, O_1 = 0, U, \theta\right) = \mathbb{I}\{U \leq \theta\}$$

A Free Lunch? Take a Second Look.

- Suppose we observe $\mathbb{I}_0 = 1$ and $\mathbb{I}_1$ is missing.
- Condition for confidence validity does not hold:

$$P\left(\mathbb{I}_0 = 1 | U, \theta\right) = \mathbb{I}\left\{U \leq 1/2\right\}$$

$$P\left(\mathbb{I}_0 = 1 | O_0 = 1, O_1 = 0, U, \theta\right) = \mathbb{I}\{U \leq \theta\}$$

- **Confidence valid inferences for U requires modeling O_0, O_1.**

A Free Lunch? Take a Second Look.

- Suppose we observe $\mathbb{I}_0 = 1$ and $\mathbb{I}_1$ is missing.
- Condition for confidence validity does not hold:

$$P\left(\mathbb{I}_0 = 1 | U, \theta\right) = \mathbb{I}\left\{U \leq 1/2\right\}$$

$$P\left(\mathbb{I}_0 = 1 | O_0 = 1, O_1 = 0, U, \theta\right) = \mathbb{I}\{U \leq \theta\}$$

- **Confidence valid inferences for U requires modeling O_0, O_1.**
- The distribution of (O_0, O_1) depends on π: the reduction $X \to \mathcal{M}(X)$ does not throw away enough information.

A Free Lunch? Take a Second Look.

- Suppose we observe $\mathbb{I}_0 = 1$ and $\mathbb{I}_1$ is missing.
- Condition for confidence validity does not hold:

$$P\left(\mathbb{I}_0 = 1 | U, \theta\right) = \mathbb{I}\left\{U \leq 1/2\right\}$$

$$P\left(\mathbb{I}_0 = 1 | O_0 = 1, O_1 = 0, U, \theta\right) = \mathbb{I}\{U \leq \theta\}$$

- **Confidence valid inferences for U requires modeling O_0, O_1.**
- The distribution of (O_0, O_1) depends on π: the reduction $X \to \mathcal{M}(X)$ does not throw away enough information.
- The information $U \in \mathcal{M}(x)$ seems *free*. But **to use it correctly** requires paying the price of a prior on θ.

The Consequences of Being Cheap

The Consequences of Being Cheap

Actual Posterior, $X = 0$

$$\mathbb{E}^{\pi(\theta|x=0)}[\mathbb{I}\left\{U \leq \theta\right\}/\theta]$$

The Consequences of Being Cheap

Actual Posterior, $X = 0$

$$\mathbb{E}^{\pi(\theta|x=0)}[\mathbb{I}\{U \leq \theta\}/\theta]$$

Naive Posterior, $X = 0$

$$2\mathbb{I}\{U \leq 1/2\}$$

The Consequences of Being Cheap

Actual Posterior, $X = 0$	**Naive Posterior, $X = 0$**
$\mathbb{E}^{\pi(\theta \mid x=0)}[\mathbb{I}\{U \le \theta\}/\theta]$	$2\mathbb{I}\{U \le 1/2\}$

- Ignoring $\{O_0, O_1\}$ equivalent to assuming point mass prior at $\theta = 1/2$—most dogmatic of all!

The Consequences of Being Cheap

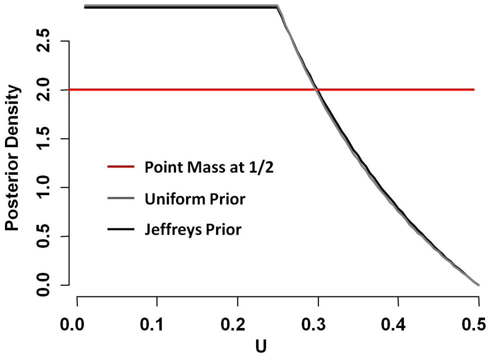
Actual Posterior, $X = 0$

$$\mathbb{E}^{\pi(\theta|x=0)}[\mathbb{I}\{U \le \theta\}/\theta]$$

Naive Posterior, $X = 0$

$$2\mathbb{I}\{U \le 1/2\}$$

- Ignoring $\{O_0, O_1\}$ equivalent to assuming point mass prior at $\theta = 1/2$—most dogmatic of all!



Know the Price

Know the Price

- **Takeaway:** The information $\mathbf{U} \in \mathcal{M}(\mathbf{x})$ is *not free*—may require assuming π.

Know the Price

- **Takeaway:** The information $\mathbf{U} \in \mathcal{M}(\mathbf{x})$ is *not free*—may require assuming π.

- **Question 1:** So when is this information free? When is $\mathbf{U}|\mathcal{M}(\mathbf{x})$ objective—free of π?

Know the Price

- **Takeaway:** The information $\mathbf{U} \in \mathcal{M}(\mathbf{x})$ is *not free*—may require assuming π.

- **Question 1:** So when is this information free? When is $\mathbf{U}|\mathcal{M}(\mathbf{x})$ objective—free of π?

- **Question 2:** What is maximal amount of free information about $\mathbf{U}$? What is the "best" objective posterior for $\mathbf{U}$?

How Small Can You Go?

How Small Can You Go?

- $\mathbf{U}|\mathcal{M}(\mathbf{x})$ is an objective posterior for $\mathbf{U}$ if and only if an R-ancillary statistic captures the information in $\mathcal{M}(\mathbf{x})$.

How Small Can You Go?

- $\mathbf{U}|\mathcal{M}(\mathbf{x})$ is an objective posterior for $\mathbf{U}$ if and only if an R-ancillary statistic captures the information in $\mathcal{M}(\mathbf{x})$.

R-Ancillary Regions

How Small Can You Go?

- $\mathbf{U}|\mathcal{M}(\mathbf{x})$ is an objective posterior for $\mathbf{U}$ if and only if an R-ancillary statistic captures the information in $\mathcal{M}(\mathbf{x})$.

R-Ancillary Regions

- If $A(\mathbf{x})$ is R-ancillary s.t. $A(G(\theta_0, \mathbf{U_0})) = A_G(\mathbf{U_0})$, the R-ancillary region defined by A is the set

$$\mathcal{A}_G\left(\mathbf{U_0}\right) = \left\{\mathbf{U} : A_G\left(\mathbf{U}\right) = A_G\left(\mathbf{U}_0\right)\right\}$$

How Small Can You Go?

- $\mathbf{U}|\mathcal{M}(\mathbf{x})$ is an objective posterior for $\mathbf{U}$ if and only if an R-ancillary statistic captures the information in $\mathcal{M}(\mathbf{x})$.

R-Ancillary Regions

- If $A(\mathbf{x})$ is R-ancillary s.t. $A(G(\theta_0, \mathbf{U_0})) = A_G(\mathbf{U_0})$, the R-ancillary region defined by A is the set

$$\mathcal{A}_G(\mathbf{U_0}) = \{\mathbf{U} : A_G(\mathbf{U}) = A_G(\mathbf{U}_0)\}$$

- The smaller the R-ancillary region, the more informative A is.

How Small Can You Go?

Keli Liu and Xiao-Li Meng

- $\mathbf{U}|\mathcal{M}(\mathbf{x})$ is an objective posterior for $\mathbf{U}$ if and only if an R-ancillary statistic captures the information in $\mathcal{M}(\mathbf{x})$.

R-Ancillary Regions

- If $A(\mathbf{x})$ is R-ancillary s.t. $A(G(\theta_0, \mathbf{U_0})) = A_G(\mathbf{U_0})$, the R-ancillary region defined by A is the set

$$\mathcal{A}_G(\mathbf{U_0}) = \{\mathbf{U} : A_G(\mathbf{U}) = A_G(\mathbf{U}_0)\}$$

- The smaller the R-ancillary region, the more informative A is.
- Any R-ancillary region is rougher than the $\mathbf{x}$ compatible region.

$$\mathcal{M}(G(\theta_0, \mathbf{U_0})) \subset \mathcal{A}_G(\mathbf{U_0}) \quad \forall\theta_0, \forall\mathcal{A}_G$$

How Small Can You Go?

- $\mathbf{U}|\mathcal{M}(\mathbf{x})$ is an objective posterior for $\mathbf{U}$ if and only if an R-ancillary statistic captures the information in $\mathcal{M}(\mathbf{x})$.

R-Ancillary Regions

- If $A(\mathbf{x})$ is R-ancillary s.t. $A(G(\theta_0, \mathbf{U_0})) = A_G(\mathbf{U_0})$, the R-ancillary region defined by A is the set

$$\mathcal{A}_G\left(\mathbf{U_0}\right) = \{\mathbf{U} : A_G\left(\mathbf{U}\right) = A_G\left(\mathbf{U}_0\right)\}$$

- The smaller the R-ancillary region, the more informative A is.
- Any R-ancillary region is rougher than the $\mathbf{x}$ compatible region.

$$\mathcal{M}(G(\theta_0, \mathbf{U_0})) \subset \mathcal{A}_G\left(\mathbf{U_0}\right) \quad \forall \theta_0, \forall \mathcal{A}_G$$

- **What is the smallest we can make $\mathcal{A}_G$? Does a smallest region even exist?**

From the Cave to Utopia?

From the Cave to Utopia?

- R-ancillary statistics are jointly R-ancillary. **Intersection of R-ancillary regions is an R-ancillary region.**

From the Cave to Utopia?

- R-ancillary statistics are jointly R-ancillary. **Intersection of R-ancillary regions is an R-ancillary region.**

Utopia ⊂ Cave

$$\mathcal{U}_G(\mathbf{U}_0) = \bigcup_{\theta_0 \in \Theta} \mathcal{M}(G(\theta_0, \mathbf{U}_0)) \subset \bigcap_{A:\ R\text{-ancillary for } G} \mathcal{A}_G(\mathbf{U}_0) = \mathcal{C}_G(\mathbf{U}_0)$$

From the Cave to Utopia?

- R-ancillary statistics are jointly R-ancillary. **Intersection of R-ancillary regions is an R-ancillary region.**

Utopia ⊂ Cave

$$\mathcal{U}_G(\mathbf{U}_0) = \bigcup_{\theta_0 \in \Theta} \mathcal{M}(G(\theta_0, \mathbf{U}_0)) \subset \bigcap_{A:\ R\text{-ancillary for } G} \mathcal{A}_G(\mathbf{U}_0) = \mathcal{C}_G(\mathbf{U}_0)$$

- The maximal R-ancillary, A_{max}, restricts $\mathbf{U_0}$ to the Cave.

From the Cave to Utopia?

- R-ancillary statistics are jointly R-ancillary. **Intersection of R-ancillary regions is an R-ancillary region.**

Utopia $\subset$ Cave

$$\mathcal{U}_G(\mathbf{U}_0) = \bigcup_{\theta_0 \in \Theta} \mathcal{M}(G(\theta_0, \mathbf{U}_0)) \subset \bigcap_{A:\ R\text{-ancillary for } G} \mathcal{A}_G(\mathbf{U}_0) = \mathcal{C}_G(\mathbf{U}_0)$$

- The maximal R-ancillary, $A_{\max}$, restricts $\mathbf{U_0}$ to the Cave.
- Ultimately, we hope to restrict $\mathbf{U_0}$ to Utopia, which is the universal **"Cramer-Rao lower bound for conditioning"**.

From the Cave to Utopia?

- R-ancillary statistics are jointly R-ancillary. **Intersection of R-ancillary regions is an R-ancillary region.**

Utopia ⊂ Cave

$$\mathcal{U}_G(\mathbf{U}_0) = \bigcup_{\theta_0 \in \Theta} \mathcal{M}(G(\theta_0, \mathbf{U}_0)) \subset \bigcap_{A:\ R\text{-ancillary for } G} \mathcal{A}_G(\mathbf{U}_0) = \mathcal{C}_G(\mathbf{U}_0)$$

- The maximal R-ancillary, $A_{\max}$, restricts $\mathbf{U_0}$ to the Cave.
- Ultimately, we hope to restrict $\mathbf{U_0}$ to Utopia, which is the universal **"Cramer-Rao lower bound for conditioning"**.
- Can we achieve Utopia?

Can Statisticians Achieve Utopia?

Can Statisticians Achieve Utopia?

Despite all the cynicism in the world, we happily report...

Can Statisticians Achieve Utopia?

Despite all the cynicism in the world, we happily report...

Utopia Can Always Be Achieved

$$\textbf{Utopia} = \bigcup_{\theta_0 \in \Theta} \mathcal{M}\left(G\left(\theta_0, \mathbf{U}_0\right)\right) = \bigcap_{A:\ R\text{-ancillary for } G} \mathcal{A}_G\left(\mathbf{U}_0\right) = \textbf{Cave}$$

Can Statisticians Achieve Utopia?

Despite all the cynicism in the world, we happily report...

Utopia Can Always Be Achieved

$$\textbf{Utopia} = \bigcup_{\theta_0 \in \Theta} \mathcal{M}\left(G\left(\theta_0, \mathbf{U}_0\right)\right) = \bigcap_{A:\ R\text{-ancillary for } G} \mathcal{A}_G\left(\mathbf{U}_0\right) = \textbf{Cave}$$

1. The collection of Utopia sets, $\{\mathcal{U}_G(\mathbf{U_0})\}_{\mathbf{U_0} \in \mathbb{U}}$, partitions the pivot space into equivalence classes.

Can Statisticians Achieve Utopia?

Despite all the cynicism in the world, we happily report...

Utopia Can Always Be Achieved

$$\textbf{Utopia} = \bigcup_{\theta_0 \in \Theta} \mathcal{M}\left(G\left(\theta_0, \mathbf{U}_0\right)\right) = \bigcap_{A:\ R\text{-ancillary for } G} \mathcal{A}_G\left(\mathbf{U}_0\right) = \textbf{Cave}$$

1. The collection of Utopia sets, $\{\mathcal{U}_G(\mathbf{U_0})\}_{\mathbf{U_0} \in \mathbb{U}}$, partitions the pivot space into equivalence classes.
2. The set of pivot space equivalence classes corresponds to a set of sample space equivalence classes.

Can Statisticians Achieve Utopia?

Despite all the cynicism in the world, we happily report...

Utopia Can Always Be Achieved

$$\textbf{Utopia} = \bigcup_{\theta_0 \in \Theta} \mathcal{M}\left(G\left(\theta_0, \mathbf{U}_0\right)\right) = \bigcap_{A:\ R\text{-ancillary for } G} \mathcal{A}_G\left(\mathbf{U}_0\right) = \textbf{Cave}$$

1. The collection of Utopia sets, $\{\mathcal{U}_G(\mathbf{U_0})\}_{\mathbf{U_0} \in \mathbb{U}}$, partitions the pivot space into equivalence classes.
2. The set of pivot space equivalence classes corresponds to a set of sample space equivalence classes.
3. The index for the sample space equivalence class is observed and is R-ancillary.

Can Statisticians Achieve Utopia?

Despite all the cynicism in the world, we happily report...

Utopia Can Always Be Achieved

$$\textbf{Utopia} = \bigcup_{\theta_0 \in \Theta} \mathcal{M}\left(G\left(\theta_0, \mathbf{U}_0\right)\right) = \bigcap_{A:\ R\text{-ancillary for } G} \mathcal{A}_G\left(\mathbf{U}_0\right) = \textbf{Cave}$$

1. The collection of Utopia sets, $\{\mathcal{U}_G(\mathbf{U_0})\}_{\mathbf{U_0} \in \mathbb{U}}$, partitions the pivot space into equivalence classes.
2. The set of pivot space equivalence classes corresponds to a set of sample space equivalence classes.
3. The index for the sample space equivalence class is observed and is R-ancillary.

- **Utopia represents an upper bound on the informativeness of an R-ancillary. It is always achieved.**

Making Relevant Subsets Relevant Again

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*.

Making Relevant Subsets Relevant Again

Keli Liu and Xiao-Li Meng

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

- Subsets of the **sample space**, $\mathbb{X}$.

Keli Liu and Xiao-Li Meng

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

- Subsets of the **sample space**, $\mathbb{X}$.
- Level-sets of ancillary statistics.

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

- Subsets of the **sample space**, $\mathbb{X}$.
- Level-sets of ancillary statistics.
- Paradoxes: Which ancillary statistics? Existence?

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

- Subsets of the **sample space**, $\mathbb{X}$.
- Level-sets of ancillary statistics.
- Paradoxes: Which ancillary statistics? Existence?

New Way

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

- Subsets of the **sample space**, $\mathbb{X}$.
- Level-sets of ancillary statistics.
- Paradoxes: Which ancillary statistics? Existence?

New Way

- Subsets of the **pivot space**, $\mathbb{U}$.

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

- Subsets of the **sample space**, $\mathbb{X}$.
- Level-sets of ancillary statistics.
- Paradoxes: Which ancillary statistics? Existence?

New Way

- Subsets of the **pivot space**, $\mathbb{U}$.
- Level-sets of R-ancillary statistics.

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

- Subsets of the **sample space**, $\mathbb{X}$.
- Level-sets of ancillary statistics.
- Paradoxes: Which ancillary statistics? Existence?

New Way

- Subsets of the **pivot space**, $\mathbb{U}$.
- Level-sets of R-ancillary statistics.
- Unique "most relevant" subset—Utopia.

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

- Subsets of the **sample space**, $\mathbb{X}$.
- Level-sets of ancillary statistics.
- Paradoxes: Which ancillary statistics? Existence?

New Way

- Subsets of the **pivot space**, $\mathbb{U}$.
- Level-sets of R-ancillary statistics.
- Unique "most relevant" subset—Utopia.

- **U**, the uncertainty, dictates how hard the inference problem is.

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

- Subsets of the **sample space**, $\mathbb{X}$.
- Level-sets of ancillary statistics.
- Paradoxes: Which ancillary statistics? Existence?

New Way

- Subsets of the **pivot space**, $\mathbb{U}$.
- Level-sets of R-ancillary statistics.
- Unique "most relevant" subset—Utopia.

- $\mathbf{U}$, the uncertainty, dictates how hard the inference problem is.
- Condition on $\mathbf{U}$ $\Leftrightarrow$ Condition on difficulty of inference.

Making Relevant Subsets Relevant Again

- Fisher (1934) argued that we should condition on *relevant subsets*. **Of what space?**

Classical Way

- Subsets of the **sample space**, $\mathbb{X}$.
- Level-sets of ancillary statistics.
- Paradoxes: Which ancillary statistics? Existence?

New Way

- Subsets of the **pivot space**, $\mathbb{U}$.
- Level-sets of R-ancillary statistics.
- Unique "most relevant" subset—Utopia.

- $\mathbf{U}$, the uncertainty, dictates how hard the inference problem is.
- Condition on $\mathbf{U}$ $\Leftrightarrow$ Condition on difficulty of inference.
- **Want to give same effort for all difficulties**.

The Best for **U**

The Best for $\mathbf{U}$

- The **optimal achievable objective** (free of π) posterior for $\mathbf{U}$: $f(\mathbf{U}|A_{\max})$.

The Best for $\mathbf{U}$

- The **optimal achievable objective** (free of π) posterior for $\mathbf{U}$: $f(\mathbf{U}|A_{\max})$.

"Heaven Is Possible" If and Only If...

With respect to fixed G,

$$f\left(\mathbf{U}|A_{\max}\left(\mathbf{x}\right)\right) = f\left(\mathbf{U}|\mathcal{M}\left(\mathbf{x}\right)\right), \ \forall \mathbf{x} \in \mathbb{X}$$

if and only if $\forall \mathbf{U} \in \mathbb{U}$, $\forall \theta \in \Theta$, $\mathcal{M}\left(G\left(\theta, \mathbf{U}\right)\right) = \mathcal{U}_G\left(\mathbf{U}\right)$ where $\mathcal{U}_G\left(\mathbf{U}\right)$ is the Utopia set containing $\mathbf{U}$.

The Best for **U**

- The **optimal achievable objective** (free of π) posterior for $\mathbf{U}$: $f(\mathbf{U}|A_{\max})$.

"Heaven Is Possible" If and Only If...

With respect to fixed G,

$$f\left(\mathbf{U}|A_{\max}\left(\mathbf{x}\right)\right) = f\left(\mathbf{U}|\mathcal{M}\left(\mathbf{x}\right)\right), \ \forall \mathbf{x} \in \mathbb{X}$$

if and only if $\forall \mathbf{U} \in \mathbb{U}$, $\forall \theta \in \Theta$, $\mathcal{M}\left(G\left(\theta, \mathbf{U}\right)\right) = \mathcal{U}_G\left(\mathbf{U}\right)$ where $\mathcal{U}_G\left(\mathbf{U}\right)$ is the Utopia set containing $\mathbf{U}$.

Interpretation

- $\mathcal{M}(G(\theta, \mathbf{U}))$ is the information we get back about $\mathbf{U}$ if data are generated using θ.

The Best for U

- The **optimal achievable objective** (free of π) posterior for $\mathbf{U}$: $f(\mathbf{U}|A_{\max})$.

"Heaven Is Possible" If and Only If...

With respect to fixed G,

$$f\left(\mathbf{U}|A_{\max}\left(\mathbf{x}\right)\right) = f\left(\mathbf{U}|\mathcal{M}\left(\mathbf{x}\right)\right), \ \forall \mathbf{x} \in \mathbb{X}$$

if and only if $\forall \mathbf{U} \in \mathbb{U}$, $\forall \theta \in \Theta$, $\mathcal{M}\left(G\left(\theta, \mathbf{U}\right)\right) = \mathcal{U}_G\left(\mathbf{U}\right)$ where $\mathcal{U}_G\left(\mathbf{U}\right)$ is the Utopia set containing $\mathbf{U}$.

Interpretation

- $\mathcal{M}(G(\theta, \mathbf{U}))$ is the information we get back about $\mathbf{U}$ if data are generated using θ.
- Invariance of $\mathcal{M}(G(\theta, \mathbf{U}))$ to θ implies that information content is **independent** of the true parameter value.

When Is the Best for **U** Good Enough for θ?

When Is the Best for **U** Good Enough for θ?

An Objective Posterior for θ

1. $f(\mathbf{U}|A_{\max}) = f(\mathbf{U}|\mathcal{M}(\mathbf{x}))$.

When Is the Best for $\mathbf{U}$ Good Enough for θ?

An Objective Posterior for θ

1. $f(\mathbf{U}|A_{\mathrm{max}}) = f(\mathbf{U}|\mathcal{M}(\mathbf{x}))$.
2. $\mathbf{x} = G(\theta, \mathbf{U})$ can be solved for θ, yielding a function, $\theta(\mathbf{U}; \mathbf{x})$, from $\mathcal{M}(\mathbf{x})$ to Θ.

When Is the Best for $\mathbf{U}$ Good Enough for θ?

An Objective Posterior for θ

1. $f(\mathbf{U}|A_{\max}) = f(\mathbf{U}|\mathcal{M}(\mathbf{x}))$.
2. $\mathbf{x} = G(\theta, \mathbf{U})$ can be solved for θ, yielding a function, $\theta(\mathbf{U}; \mathbf{x})$, from $\mathcal{M}(\mathbf{x})$ to Θ.

- Then $\theta(\mathbf{U}; \mathbf{x})$ as a function of $\mathbf{U} \sim f(\mathbf{U}|A_{\max})$ induces a **Frequentist calibrated** posterior distribution for θ.

When Is the Best for $\mathbf{U}$ Good Enough for θ?

An Objective Posterior for θ

1. $f(\mathbf{U}|A_{\max}) = f(\mathbf{U}|\mathcal{M}(\mathbf{x}))$.
2. $\mathbf{x} = G(\theta, \mathbf{U})$ can be solved for θ, yielding a function, $\theta(\mathbf{U}; \mathbf{x})$, from $\mathcal{M}(\mathbf{x})$ to Θ.

- Then $\theta(\mathbf{U}; \mathbf{x})$ as a function of $\mathbf{U} \sim f(\mathbf{U}|A_{\max})$ induces a **Frequentist calibrated** posterior distribution for θ.

Most Relevant Confidence Regions

- Let C_U be a $1 - \alpha$ probability set w.r.t. $f(\mathbf{U}|A_{\max})$.

When Is the Best for $\mathbf{U}$ Good Enough for θ?

An Objective Posterior for θ

1. $f(\mathbf{U}|A_{\max}) = f(\mathbf{U}|\mathcal{M}(\mathbf{x}))$.
2. $\mathbf{x} = G(\theta, \mathbf{U})$ can be solved for θ, yielding a function, $\theta(\mathbf{U}; \mathbf{x})$, from $\mathcal{M}(\mathbf{x})$ to Θ.

- Then $\theta(\mathbf{U}; \mathbf{x})$ as a function of $\mathbf{U} \sim f(\mathbf{U}|A_{\max})$ induces a **Frequentist calibrated** posterior distribution for θ.

Most Relevant Confidence Regions

- Let C_U be a $1 - \alpha$ probability set w.r.t. $f(\mathbf{U}|A_{\max})$.
- For fixed G, C_U is a **most relevant** confidence region for $\mathbf{U}$.

When Is the Best for $\mathbf{U}$ Good Enough for θ?

An Objective Posterior for θ

1. $f(\mathbf{U}|A_{\max}) = f(\mathbf{U}|\mathcal{M}(\mathbf{x}))$.
2. $\mathbf{x} = G(\theta, \mathbf{U})$ can be solved for θ, yielding a function, $\theta(\mathbf{U}; \mathbf{x})$, from $\mathcal{M}(\mathbf{x})$ to Θ.

- Then $\theta(\mathbf{U}; \mathbf{x})$ as a function of $\mathbf{U} \sim f(\mathbf{U}|A_{\max})$ induces a **Frequentist calibrated** posterior distribution for θ.

Most Relevant Confidence Regions

- Let C_U be a $1-\alpha$ probability set w.r.t. $f(\mathbf{U}|A_{\max})$.
- For fixed G, C_U is a **most relevant** confidence region for $\mathbf{U}$.
- The mapping $\theta(\mathbf{U}, \mathbf{x})$ (not necessarily function) converts C_U into a most relevant $1-\alpha$ level confidence region, C_θ for θ.

When Is the Best for $\mathbf{U}$ Good Enough for θ?

An Objective Posterior for θ

1. $f(\mathbf{U}|A_{\max}) = f(\mathbf{U}|\mathcal{M}(\mathbf{x}))$.
2. $\mathbf{x} = G(\theta, \mathbf{U})$ can be solved for θ, yielding a function, $\theta(\mathbf{U}; \mathbf{x})$, from $\mathcal{M}(\mathbf{x})$ to Θ.

- Then $\theta(\mathbf{U}; \mathbf{x})$ as a function of $\mathbf{U} \sim f(\mathbf{U}|A_{\max})$ induces a **Frequentist calibrated** posterior distribution for θ.

Most Relevant Confidence Regions

- Let C_U be a $1-\alpha$ probability set w.r.t. $f(\mathbf{U}|A_{\max})$.
- For fixed G, C_U is a **most relevant** confidence region for $\mathbf{U}$.
- The mapping $\theta(\mathbf{U}, \mathbf{x})$ (not necessarily function) converts C_U into a most relevant $1-\alpha$ level confidence region, C_θ for θ.
- C_θ attains posterior probability $1-\alpha$ if (1) and (2) hold.

Posterior Reflections: From $\mathbf{U}$ to θ Through G

Posterior Reflections: From $\mathbf{U}$ to θ Through G

Invert $\mathbf{x} = G(\theta, \mathbf{U})$ to

$$\theta^*(\mathbf{U}; \mathbf{x}) = \{\theta : G(\theta, \mathbf{U}) = \mathbf{x}\}$$

Posterior Reflections: From $\mathbf{U}$ to θ Through G

Invert $\mathbf{x} = G(\theta, \mathbf{U})$ to $\theta^*(\mathbf{U}; \mathbf{x}) = \{\theta : G(\theta, \mathbf{U}) = \mathbf{x}\}$

$$\mathcal{M}(\mathbf{x}) = \mathcal{U}_G(\mathbf{U}_0)$$

Posterior Reflections: From $\mathbf{U}$ to θ Through G

Invert $\mathbf{x} = G(\theta, \mathbf{U})$ to $\theta^*(\mathbf{U}; \mathbf{x}) = \{\theta : G(\theta, \mathbf{U}) = \mathbf{x}\}$

$\mathcal{M}(\mathbf{x}) = \mathcal{U}_G(\mathbf{U}_0)$

$\theta^*(\mathbf{U}; \mathbf{x})$ is singleton.

Posterior Reflections: From $\mathbf{U}$ to θ Through G

Invert $\mathbf{x} = G(\theta, \mathbf{U})$ to $\theta^*(\mathbf{U}; \mathbf{x}) = \{\theta : G(\theta, \mathbf{U}) = \mathbf{x}\}$

$\mathcal{M}(\mathbf{x}) = \mathcal{U}_G(\mathbf{U}_0)$

$\theta^*(\mathbf{U}; \mathbf{x})$ is singleton.

Fiducial posterior is $\theta^*(\mathbf{U}; \mathbf{x}) | A_{max}$

Posterior Reflections: From $\mathbf{U}$ to θ Through G

Invert $\mathbf{x} = G(\theta, \mathbf{U})$ to $\theta^*(\mathbf{U}; \mathbf{x}) = \{\theta : G(\theta, \mathbf{U}) = \mathbf{x}\}$

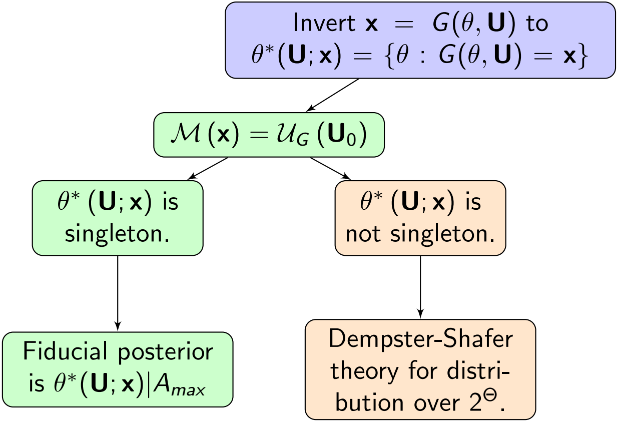
$\mathcal{M}(\mathbf{x}) = \mathcal{U}_G(\mathbf{U}_0)$

$\theta^*(\mathbf{U}; \mathbf{x})$ is singleton.

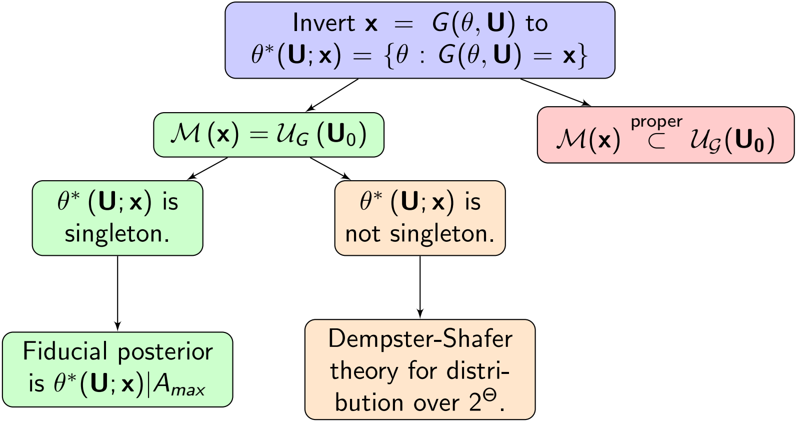
$\theta^*(\mathbf{U}; \mathbf{x})$ is not singleton.

Fiducial posterior is $\theta^*(\mathbf{U}; \mathbf{x})|A_{max}$

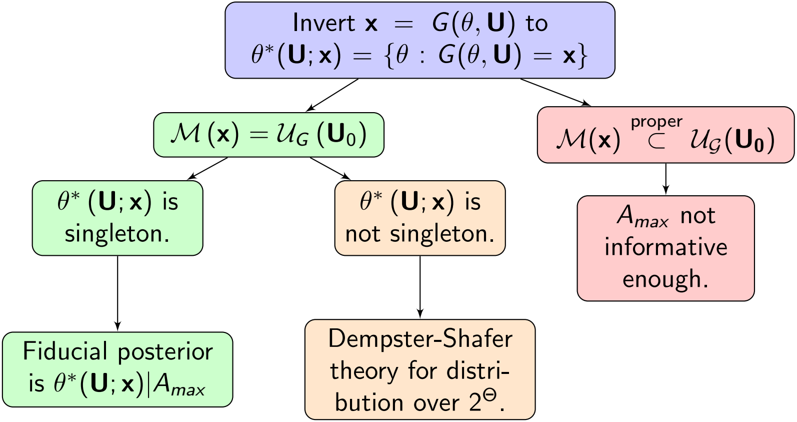
Posterior Reflections: From $\mathbf{U}$ to θ Through G



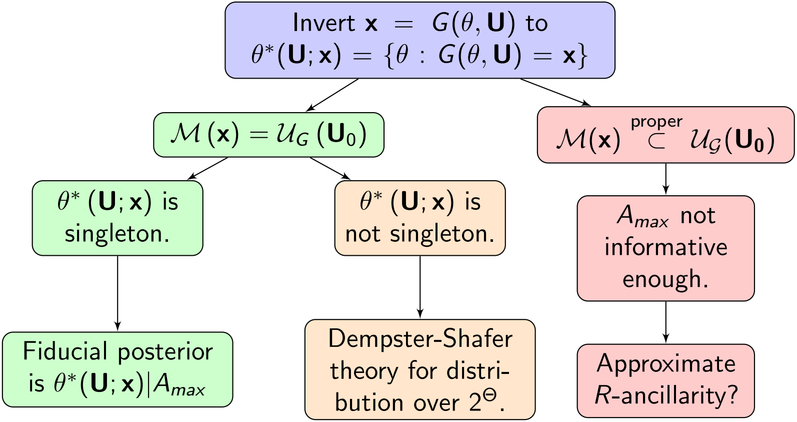
Posterior Reflections: From $\mathbf{U}$ to θ Through G



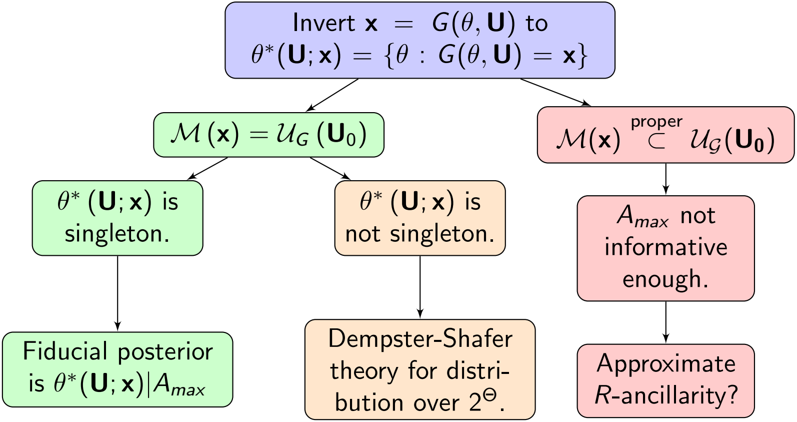
Posterior Reflections: From $\mathbf{U}$ to θ Through G



Posterior Reflections: From $\mathbf{U}$ to θ Through G



Posterior Reflections: From $\mathbf{U}$ to θ Through G



- **Uncongeniality:** Inconsistent use of the full data, $\mathbf{x}$, and the partial data, A_{max}, in different phases of the analysis.

Posterior Reflections: From $\mathbf{U}$ to θ Through G

Invert $\mathbf{x} = G(\theta, \mathbf{U})$ to $\theta^*(\mathbf{U}; \mathbf{x}) = \{\theta : G(\theta, \mathbf{U}) = \mathbf{x}\}$

- $\mathcal{M}(\mathbf{x}) = \mathcal{U}_G(\mathbf{U}_0)$
  - $\theta^*(\mathbf{U}; \mathbf{x})$ is singleton. → Fiducial posterior is $\theta^*(\mathbf{U}; \mathbf{x}) | A_{max}$
  - $\theta^*(\mathbf{U}; \mathbf{x})$ is not singleton. → Dempster-Shafer theory for distribution over 2^{Θ}.
- $\mathcal{M}(\mathbf{x}) \overset{\text{proper}}{\subset} \mathcal{U}_{\mathcal{G}}(\mathbf{U_0})$ → A_{max} not informative enough. → Approximate R-ancillarity?

- **Uncongeniality:** Inconsistent use of the full data, $\mathbf{x}$, and the partial data, A_{max}, in different phases of the analysis.
- **Non-uniqueness**: How does one choose G?

Bibliography

- Barnard, G. A. (1995). Pivotal Models and the Fiducial Argument. *International Statistical Review*, 63(3): 309-323.
- Basu, D. (1964). Recovery of ancillary information. *Sankhya*, 26: 3-16.
- Dawid, A.P. and M. Stone (1982). The Function-Model Basis of Fiducial Inference. *Annals of Statistics*, 10(4): 1054-1067.
- Dempster, A.P. (2008). The Dempster-Shafer calculus for statisticians. *International Journal of Approximate Reasoning*, 48: 365-377.
- Fraser, D.A.S. (1968). *The Structure of Inference.* John Wiley & Sons, New York-London-Sydney.

Bibliography

- Hannig, Jan. (2009). On Generalized Fiducial Inference. *Statistica Sinica*, 19: 491-544.
- Martin, R., J. Zhang, and C. Liu (2010). Dempster-Shafer Theory and Statistical Inference with Weak Beliefs. *Statistical Science*, 20(1): 72-87.
- Rubin, Donald B. (1976). Inference and missing data. *Biometrika*, 63(3): 581-592.
- Taraldsen, G. and B. Lindqvsit (2013). Fiducial Theory and Optimal Inference. *Annals of Statistics*, 41(1): 323-341.