The road from conditional Monte Carlo to improper priors and fiducial inference

Or trying to prove a false statement ...

G. Taraldsen B.H. Lindqvist

Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim, NORWAY

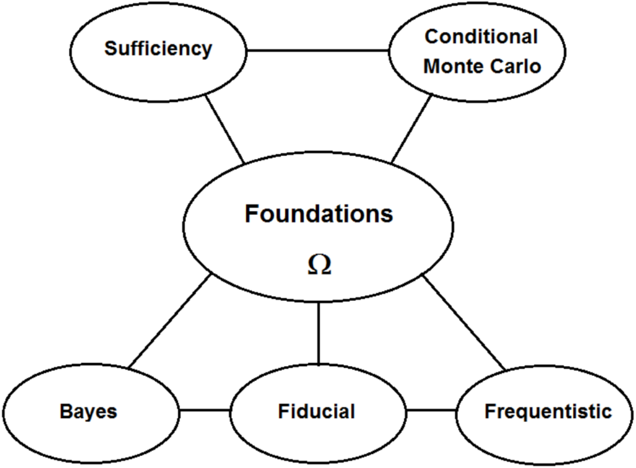
September 26th, 2016
(Inference With Confidence, 2015)

Abstract

We started out in 1997 on a project to clarify certain aspects of a conditional Monte Carlo method suggested by Engen and Lillegård. The method produces conditional samples given a sufficient statistic, and can hence be used to construct optimal inference procedures. In some cases it can also be used to eliminate nuisance parameters by conditioning. The steps in the algorithm give strong links to fiducial inference, and also to the foundations of statistics both in mathematical and philosophical terms. The talk will give an informal sketch of the initial method, and the many related foundational and practical issues.

Table of Contents

The Large Picture



From Conditional Monte Carlo to ...

- **Initial problem: Generate data X conditionally given a sufficient statistic T.**
- Tentative solution: Adjust parameter value θ for simulated data x so that the sufficient statistic is kept fixed equal to t.
- Corrected solution: The simulated data must be weighted, and the weight depends on an arbitrarily chosen improper distribution for the parameter.
- Realization after many years: Oh ... the resulting unweighted adjusted parameters are fiducial and the weighted are Bayes posterior ...
- Ingredients: Sufficiency, Optimal inference, Improper distributions, Fiducial versus Bayesian distribution, Borel paradox, Marginalization paradox, Axioms for statistics.

From Conditional Monte Carlo to ...

- Initial problem: Generate data X conditionally given a sufficient statistic T.
- **Tentative solution: Adjust parameter value θ for simulated data x so that the sufficient statistic is kept fixed equal to t.**
- Corrected solution: The simulated data must be weighted, and the weight depends on an arbitrarily chosen improper distribution for the parameter.
- Realization after many years: Oh ... the resulting unweighted adjusted parameters are fiducial and the weighted are Bayes posterior ...
- Ingredients: Sufficiency, Optimal inference, Improper distributions, Fiducial versus Bayesian distribution, Borel paradox, Marginalization paradox, Axioms for statistics.

From Conditional Monte Carlo to ...

- Initial problem: Generate data X conditionally given a sufficient statistic T.
- Tentative solution: Adjust parameter value θ for simulated data x so that the sufficient statistic is kept fixed equal to t.
- **Corrected solution: The simulated data must be weighted, and the weight depends on an arbitrarily chosen improper distribution for the parameter.**
- Realization after many years: Oh ... the resulting unweighted adjusted parameters are fiducial and the weighted are Bayes posterior ...
- Ingredients: Sufficiency, Optimal inference, Improper distributions, Fiducial versus Bayesian distribution, Borel paradox, Marginalization paradox, Axioms for statistics.

From Conditional Monte Carlo to ...

- Initial problem: Generate data X conditionally given a sufficient statistic T.
- Tentative solution: Adjust parameter value θ for simulated data x so that the sufficient statistic is kept fixed equal to t.
- Corrected solution: The simulated data must be weighted, and the weight depends on an arbitrarily chosen improper distribution for the parameter.
- **Realization after many years: Oh ... the resulting unweighted adjusted parameters are fiducial and the weighted are Bayes posterior ...**
- Ingredients: Sufficiency, Optimal inference, Improper distributions, Fiducial versus Bayesian distribution, Borel paradox, Marginalization paradox, Axioms for statistics.

From Conditional Monte Carlo to ...

- Initial problem: Generate data X conditionally given a sufficient statistic T.
- Tentative solution: Adjust parameter value θ for simulated data x so that the sufficient statistic is kept fixed equal to t.
- Corrected solution: The simulated data must be weighted, and the weight depends on an arbitrarily chosen improper distribution for the parameter.
- Realization after many years: Oh ... the resulting unweighted adjusted parameters are fiducial and the weighted are Bayes posterior ...
- **Ingredients: Sufficiency, Optimal inference, Improper distributions, Fiducial versus Bayesian distribution, Borel paradox, Marginalization paradox, Axioms for statistics.**

Table of Contents

Sufficiency with impropers

- **T is sufficient relatively to X for the parameter θ if $X \,|(T = t, \Theta = \theta)$ does not depend on θ.**
- Equivalently: X, Θ are conditionally independent given T, so $(\Theta \,|\, X, T) \sim (\Theta \,|\, T)$. If, additionally, $T = \tau(X)$, then $(\Theta \,|\, X) \sim (\Theta \,|\, T) \sim$ Bayes posterior.
- Theorem: Sufficiency implies that $[X \,|(T = t, \Theta = \theta)] \sim [X \,|\, T = t]$.
- Proof in discrete case: $E^\theta(\phi(X) \,|\, T = t) =$

$$\frac{E^\theta(\phi(X)(T = t))}{E^\theta(T = t)} = \frac{\int \pi(\theta) E^\theta(\phi(X)(T = t))\, d\theta}{\int \pi(\theta) E^\theta(T = t)\, d\theta}$$

- Beware: $X \,|\, Y = y$ does not depend on y does NOT imply that $(X \,|\, Y) \sim X$.

Sufficiency with impropers

- T is sufficient relatively to X for the parameter θ if $X \,|\, (T = t, \Theta = \theta)$ does not depend on θ.
- **Equivalently: X, Θ are conditionally independent given T, so $(\Theta \,|\, X, T) \sim (\Theta \,|\, T)$. If, additionally, $T = \tau(X)$, then $(\Theta \,|\, X) \sim (\Theta \,|\, T) \sim$ Bayes posterior.**
- Theorem: Sufficiency implies that $[X \,|\, (T = t, \Theta = \theta)] \sim [X \,|\, T = t]$.
- Proof in discrete case: $E^\theta(\phi(X) \,|\, T = t) =$

$$\frac{E^\theta(\phi(X)(T = t))}{E^\theta(T = t)} = \frac{\int \pi(\theta) E^\theta(\phi(X)(T = t))\, d\theta}{\int \pi(\theta) E^\theta(T = t)\, d\theta}$$

- Beware: $X \,|\, Y = y$ does not depend on y does NOT imply that $(X \,|\, Y) \sim X$.

Sufficiency with impropers

- T is sufficient relatively to X for the parameter θ if $X \,|(T = t, \Theta = \theta)$ does not depend on θ.
- Equivalently: X, Θ are conditionally independent given T, so $(\Theta \,|\, X, T) \sim (\Theta \,|\, T)$. If, additionally, $T = \tau(X)$, then $(\Theta \,|\, X) \sim (\Theta \,|\, T) \sim$ Bayes posterior.
- **Theorem: Sufficiency implies that $[X \,|(T = t, \Theta = \theta)] \sim [X \,|\, T = t]$.**
- Proof in discrete case: $E^\theta(\phi(X) \,|\, T = t) =$

$$\frac{E^\theta(\phi(X)(T = t))}{E^\theta(T = t)} = \frac{\int \pi(\theta) E^\theta(\phi(X)(T = t))\, d\theta}{\int \pi(\theta) E^\theta(T = t)\, d\theta}$$

- Beware: $X \,|\, Y = y$ does not depend on y does NOT imply that $(X \,|\, Y) \sim X$.

Sufficiency with impropers

- T is sufficient relatively to X for the parameter θ if $X \,|(T = t, \Theta = \theta)$ does not depend on θ.
- Equivalently: X, Θ are conditionally independent given T, so $(\Theta \,|\, X, T) \sim (\Theta \,|\, T)$. If, additionally, $T = \tau(X)$, then $(\Theta \,|\, X) \sim (\Theta \,|\, T) \sim$ Bayes posterior.
- Theorem: Sufficiency implies that $[X \,|(T = t, \Theta = \theta)] \sim [X \,|\, T = t]$.
- **Proof in discrete case:** $E^\theta(\phi(X) \,|\, T = t) =$

$$\frac{E^\theta(\phi(X)(T = t))}{E^\theta(T = t)} = \frac{\int \pi(\theta) E^\theta(\phi(X)(T = t))\, d\theta}{\int \pi(\theta) E^\theta(T = t)\, d\theta}$$

- Beware: $X \,|\, Y = y$ does not depend on y does NOT imply that $(X \,|\, Y) \sim X$.

Sufficiency with impropers

- T is sufficient relatively to X for the parameter θ if $X \,|(T = t, \Theta = \theta)$ does not depend on θ.
- Equivalently: X, Θ are conditionally independent given T, so $(\Theta \,|\, X, T) \sim (\Theta \,|\, T)$. If, additionally, $T = \tau(X)$, then $(\Theta \,|\, X) \sim (\Theta \,|\, T) \sim$ Bayes posterior.
- Theorem: Sufficiency implies that $[X \,|(T = t, \Theta = \theta)] \sim [X \,|\, T = t]$.
- Proof in discrete case: $E^\theta(\phi(X) \,|\, T = t) =$

$$\frac{E^\theta(\phi(X)(T = t))}{E^\theta(T = t)} = \frac{\int \pi(\theta) E^\theta(\phi(X)(T = t))\, d\theta}{\int \pi(\theta) E^\theta(T = t)\, d\theta}$$

- **Beware:** $X \,|\, Y = y$ **does not depend on** y **does NOT imply that** $(X \,|\, Y) \sim X$**.**

Sufficiency and optimal inference

- **Sufficiency principle: Valid inference must be based on T for all sufficient T. It can be disputed. Halmos argument relies on accepting randomized procedures.**
- If $\phi(X)$ is an estimator, then $E^\theta(\phi(X) \mid T)$ is an estimator with smaller (or equal) convex risk. If T is complete (and minimal), then it is the unique optimal estimator.
- An exact test for $(H_0 : \alpha = \alpha_0, \theta$ arbitrary) is obtainable if T is sufficient for the nuisance parameter θ. This gives exact confidence distributions. Lehmann: Optimality follows from this with additional assumptions.
- Accepting randomized procedures is equivalent to accepting construction of an improved alternative experiment. It can be disputed.

Sufficiency and optimal inference

- Sufficiency principle: Valid inference must be based on T for all sufficient T. It can be disputed. Halmos argument relies on accepting randomized procedures.
- **If $\phi(X)$ is an estimator, then $E^{\theta}(\phi(X) \mid T)$ is an estimator with smaller (or equal) convex risk. If T is complete (and minimal), then it is the unique optimal estimator.**
- An exact test for $(H_0 : \alpha = \alpha_0, \theta$ arbitrary) is obtainable if T is sufficient for the nuisance parameter θ. This gives exact confidence distributions. Lehmann: Optimality follows from this with additional assumptions.
- Accepting randomized procedures is equivalent to accepting construction of an improved alternative experiment. It can be disputed.

Sufficiency and optimal inference

- Sufficiency principle: Valid inference must be based on T for all sufficient T. It can be disputed. Halmos argument relies on accepting randomized procedures.
- If $\phi(X)$ is an estimator, then $E^\theta(\phi(X) \mid T)$ is an estimator with smaller (or equal) convex risk. If T is complete (and minimal), then it is the unique optimal estimator.
- **An exact test for $(H_0 : \alpha = \alpha_0, \theta$ arbitrary$)$ is obtainable if T is sufficient for the nuisance parameter θ. This gives exact confidence distributions. Lehmann: Optimality follows from this with additional assumptions.**
- Accepting randomized procedures is equivalent to accepting construction of an improved alternative experiment. It can be disputed.

Sufficiency and optimal inference

- Sufficiency principle: Valid inference must be based on T for all sufficient T. It can be disputed. Halmos argument relies on accepting randomized procedures.
- If $\phi(X)$ is an estimator, then $E^{\theta}(\phi(X) \mid T)$ is an estimator with smaller (or equal) convex risk. If T is complete (and minimal), then it is the unique optimal estimator.
- An exact test for $(H_0 : \alpha = \alpha_0, \theta$ arbitrary) is obtainable if T is sufficient for the nuisance parameter θ. This gives exact confidence distributions. Lehmann: Optimality follows from this with additional assumptions.
- **Accepting randomized procedures is equivalent to accepting construction of an improved alternative experiment. It can be disputed.**

An erroneous (but nice) argument

- **A joint fiducial model $X = \chi(U, \Theta)$ and $T = \tau(U, \Theta)$.**
- Assume $\tau(u, \hat{\theta}(u, t)) = t$ and $\hat{\theta}(u, \tau(u, \theta)) = \theta$.
- For any such model one may calculate (formally):
  $E^\theta(X \mid T = t) = E(\chi(U, \theta) \mid \tau(U, \theta) = t) =$
  $E(\chi(U, \theta) \mid \hat{\theta}(U, t) = \theta) = E(\chi(U, \hat{\theta}(U, t)) \mid \hat{\theta}(U, t) = \theta) =$
  $E(\chi(U, \hat{\theta}(U, t)))$
- Replacing X by $\phi(X)$ in the calculation gives that $\chi(U, \hat{\theta}(U, t))$ is distributed like $X \mid T = t$.
- What is wrong?
- Answer: There is a Borel paradox type error in the argument. The event $(\tau(U, \theta) = t) = (\hat{\theta}(U, t) = \theta)$ has zero probability. The σ-fields must be equal.

An erroneous (but nice) argument

- A joint fiducial model $X = \chi(U, \Theta)$ and $T = \tau(U, \Theta)$.
- **Assume $\tau(u, \hat{\theta}(u,t)) = t$ and $\hat{\theta}(u, \tau(u, \theta)) = \theta$.**
- For any such model one may calculate (formally): $E^\theta(X \mid T = t) = E(\chi(U,\theta) \mid \tau(U,\theta) = t) = E(\chi(U,\theta) \mid \hat{\theta}(U,t) = \theta) = E(\chi(U,\hat{\theta}(U,t)) \mid \hat{\theta}(U,t) = \theta) = E(\chi(U,\hat{\theta}(U,t)))$
- Replacing X by $\phi(X)$ in the calculation gives that $\chi(U, \hat{\theta}(U,t))$ is distributed like $X \mid T = t$.
- What is wrong?
- Answer: There is a Borel paradox type error in the argument. The event $(\tau(U,\theta) = t) = (\hat{\theta}(U,t) = \theta)$ has zero probability. The σ-fields must be equal.

An erroneous (but nice) argument

- A joint fiducial model $X = \chi(U, \Theta)$ and $T = \tau(U, \Theta)$.
- Assume $\tau(u, \hat{\theta}(u, t)) = t$ and $\hat{\theta}(u, \tau(u, \theta)) = \theta$.
- **For any such model one may calculate (formally):**
$E^\theta(X \,|\, T = t) = E(\chi(U, \theta) \,|\, \tau(U, \theta) = t) =$
$E(\chi(U, \theta) \,|\, \hat{\theta}(U, t) = \theta) = E(\chi(U, \hat{\theta}(U, t)) \,|\, \hat{\theta}(U, t) = \theta) =$
$E(\chi(U, \hat{\theta}(U, t)))$
- Replacing X by $\phi(X)$ in the calculation gives that $\chi(U, \hat{\theta}(U, t))$ is distributed like $X \,|\, T = t$.
- What is wrong?
- Answer: There is a Borel paradox type error in the argument. The event $(\tau(U, \theta) = t) = (\hat{\theta}(U, t) = \theta)$ has zero probability. The σ-fields must be equal.

An erroneous (but nice) argument

- A joint fiducial model $X = \chi(U, \Theta)$ and $T = \tau(U, \Theta)$.
- Assume $\tau(u, \hat{\theta}(u, t)) = t$ and $\hat{\theta}(u, \tau(u, \theta)) = \theta$.
- For any such model one may calculate (formally):
  $E^\theta(X \mid T = t) = E(\chi(U, \theta) \mid \tau(U, \theta) = t) =$
  $E(\chi(U, \theta) \mid \hat{\theta}(U, t) = \theta) = E(\chi(U, \hat{\theta}(U, t)) \mid \hat{\theta}(U, t) = \theta) =$
  $E(\chi(U, \hat{\theta}(U, t)))$
- **Replacing X by $\phi(X)$ in the calculation gives that $\chi(U, \hat{\theta}(U, t))$ is distributed like $X \mid T = t$.**
- What is wrong?
- Answer: There is a Borel paradox type error in the argument. The event $(\tau(U, \theta) = t) = (\hat{\theta}(U, t) = \theta)$ has zero probability. The σ-fields must be equal.

An erroneous (but nice) argument

- A joint fiducial model $X = \chi(U, \Theta)$ and $T = \tau(U, \Theta)$.
- Assume $\tau(u, \hat{\theta}(u, t)) = t$ and $\hat{\theta}(u, \tau(u, \theta)) = \theta$.
- For any such model one may calculate (formally):
  $E^\theta(X \mid T = t) = E(\chi(U, \theta) \mid \tau(U, \theta) = t) =$
  $E(\chi(U, \theta) \mid \hat{\theta}(U, t) = \theta) = E(\chi(U, \hat{\theta}(U, t)) \mid \hat{\theta}(U, t) = \theta) =$
  $E(\chi(U, \hat{\theta}(U, t)))$
- Replacing X by $\phi(X)$ in the calculation gives that $\chi(U, \hat{\theta}(U, t))$ is distributed like $X \mid T = t$.
- **What is wrong?**
- Answer: There is a Borel paradox type error in the argument. The event $(\tau(U, \theta) = t) = (\hat{\theta}(U, t) = \theta)$ has zero probability. The σ-fields must be equal.

An erroneous (but nice) argument

- A joint fiducial model $X = \chi(U, \Theta)$ and $T = \tau(U, \Theta)$.
- Assume $\tau(u, \hat{\theta}(u, t)) = t$ and $\hat{\theta}(u, \tau(u, \theta)) = \theta$.
- For any such model one may calculate (formally): $E^{\theta}(X \mid T = t) = E(\chi(U, \theta) \mid \tau(U, \theta) = t) = E(\chi(U, \theta) \mid \hat{\theta}(U, t) = \theta) = E(\chi(U, \hat{\theta}(U, t)) \mid \hat{\theta}(U, t) = \theta) = E(\chi(U, \hat{\theta}(U, t)))$
- Replacing X by $\phi(X)$ in the calculation gives that $\chi(U, \hat{\theta}(U, t))$ is distributed like $X \mid T = t$.
- What is wrong?
- **Answer: There is a Borel paradox type error in the argument. The event $(\tau(U, \theta) = t) = (\hat{\theta}(U, t) = \theta)$ has zero probability. The σ-fields must be equal.**

Three correct arguments

- **Pivotal argument: The previous calculation is correct by assuming an additional pivotal structure.**
- Discrete case argument: Elementary calculation possible, and added weight arises from added law for the parameter. A limit argument gives the continuous case.
- Sampling argument: Everything is reduced to the problem of sampling from $(U, \Theta) \,|\, T = t$ which can be done in two steps:
  - Sample u from $U \,|\, T = t$.
  - Sample θ from $\Theta \,|\, U = u, T = t$.

Three correct arguments

- Pivotal argument: The previous calculation is correct by assuming an additional pivotal structure.
- **Discrete case argument: Elementary calculation possible, and added weight arises from added law for the parameter. A limit argument gives the continuous case.**
- Sampling argument: Everything is reduced to the problem of sampling from $(U, \Theta) \mid T = t$ which can be done in two steps:
  - Sample u from $U \mid T = t$.
  - Sample θ from $\Theta \mid U = u, T = t$.

Three correct arguments

- Pivotal argument: The previous calculation is correct by assuming an additional pivotal structure.
- Discrete case argument: Elementary calculation possible, and added weight arises from added law for the parameter. A limit argument gives the continuous case.
- **Sampling argument: Everything is reduced to the problem of sampling from $(U, \Theta) \mid T = t$ which can be done in two steps:**
  - **Sample u from $U \mid T = t$.**
  - **Sample θ from $\Theta \mid U = u, T = t$.**

Three correct arguments

- Pivotal argument: The previous calculation is correct by assuming an additional pivotal structure.
- Discrete case argument: Elementary calculation possible, and added weight arises from added law for the parameter. A limit argument gives the continuous case.
- Sampling argument: Everything is reduced to the problem of sampling from $(U, \Theta) \mid T = t$ which can be done in two steps:
  - **Sample u from $U \mid T = t$.**
  - Sample θ from $\Theta \mid U = u, T = t$.

Three correct arguments

- Pivotal argument: The previous calculation is correct by assuming an additional pivotal structure.
- Discrete case argument: Elementary calculation possible, and added weight arises from added law for the parameter. A limit argument gives the continuous case.
- Sampling argument: Everything is reduced to the problem of sampling from $(U, \Theta) \,|\, T = t$ which can be done in two steps:
  - Sample u from $U \,|\, T = t$.
  - **Sample θ from $\Theta \,|\, U = u, T = t$.**

The pivotal argument

- **Assume that $\tau(u, \theta) = \theta \circ r(u)$ where $\circ$ defines a quasi-group (or loop) multiplication, and $\Omega_\Theta = \Omega_T = \Omega_R$ have been identified.**
- The previous calculation is then correct since $[\tau(U, \theta) = t] = [\theta \circ R = t] = [R = \theta^{-1} \circ t] = [t \circ R^{-1} = \theta] = [\hat{\theta}(U, t) = \theta]$, and all involved functions are one-one.
- The involved σ-fields in the conditioning are all equal.
- By-product: The quantity $R = \theta^{-1} \circ T$ is pivotal, and the law of $\hat{\theta}(U, t) = t \circ R^{-1}$ is a confidence distribution.

The pivotal argument

- Assume that $\tau(u,\theta) = \theta \circ r(u)$ where $\circ$ defines a quasi-group (or loop) multiplication, and $\Omega_\Theta = \Omega_T = \Omega_R$ have been identified.
- **The previous calculation is then correct since** $[\tau(U,\theta) = t] = [\theta \circ R = t] = [R = \theta^{-1} \circ t] = [t \circ R^{-1} = \theta] = [\hat{\theta}(U,t) = \theta]$**, and all involved functions are one-one.**
- The involved σ-fields in the conditioning are all equal.
- By-product: The quantity $R = \theta^{-1} \circ T$ is pivotal, and the law of $\hat{\theta}(U,t) = t \circ R^{-1}$ is a confidence distribution.

The pivotal argument

- Assume that $\tau(u,\theta) = \theta \circ r(u)$ where $\circ$ defines a quasi-group (or loop) multiplication, and $\Omega_\Theta = \Omega_T = \Omega_R$ have been identified.
- The previous calculation is then correct since $[\tau(U,\theta) = t] = [\theta \circ R = t] = [R = \theta^{-1} \circ t] = [t \circ R^{-1} = \theta] = [\hat{\theta}(U,t) = \theta]$, and all involved functions are one-one.
- **The involved σ-fields in the conditioning are all equal.**
- By-product: The quantity $R = \theta^{-1} \circ T$ is pivotal, and the law of $\hat{\theta}(U,t) = t \circ R^{-1}$ is a confidence distribution.

The pivotal argument

- Assume that $\tau(u, \theta) = \theta \circ r(u)$ where $\circ$ defines a quasi-group (or loop) multiplication, and $\Omega_\Theta = \Omega_T = \Omega_R$ have been identified.
- The previous calculation is then correct since $[\tau(U, \theta) = t] = [\theta \circ R = t] = [R = \theta^{-1} \circ t] = [t \circ R^{-1} = \theta] = [\hat{\theta}(U, t) = \theta]$, and all involved functions are one-one.
- The involved σ-fields in the conditioning are all equal.
- **By-product: The quantity $R = \theta^{-1} \circ T$ is pivotal, and the law of $\hat{\theta}(U, t) = t \circ R^{-1}$ is a confidence distribution.**

The discrete case argument

- $E^\theta(\phi(X) \,|\, T = t) =$

$$\frac{E^\theta(\phi(X)(T = t))}{E^\theta(T = t)} = \frac{\int \pi(\theta) E^\theta(\phi(X)(T = t))\, d\theta}{\int \pi(\theta) E^\theta(T = t)\, d\theta}$$

- . . . no time for this now.

The discrete case argument

- $E^\theta(\phi(X) \,|\, T = t) =$

$$\frac{E^\theta(\phi(X)(T = t))}{E^\theta(T = t)} = \frac{\int \pi(\theta) E^\theta(\phi(X)(T = t))\, d\theta}{\int \pi(\theta) E^\theta(T = t)\, d\theta}$$

- **. . . no time for this now.**

The sampling argument

- **Theorem:** $(U \mid T = t) \sim w_t(u)\, \mathbf{P}_U(du)$ **when** $\tau(u, \Theta) \sim w_t(u)\mu(dt)$ **is assumed.**
- Proof from change-of-variables theorem: $E(\phi(U, T)) =$

$$\int [\int \phi(u, \tau(u, \theta))\ \mathrm{P}_\Theta(d\theta)]\ \mathrm{P}_U(du) =$$

$$\int [\int \phi(u, t) w_t(u)\, \mu(dt)]\ \mathrm{P}_U(du)$$

- Conclusion: Sampling u from $U \mid T = t$ is obtained by sampling u from U with weight $w_t(u)$.
- $\tau(u, \Theta) = \Theta \circ r(u) \sim \Theta$ if P_Θ is chosen as a right-invariant measure. This does not always exist.

The sampling argument

- Theorem: $(U \mid T = t) \sim w_t(u)\, \mathrm{P}_U(du)$ when $\tau(u, \Theta) \sim w_t(u)\mu(dt)$ is assumed.
- **Proof from change-of-variables theorem:** $E(\phi(U, T)) =$

$$\int [\int \phi(u, \tau(u, \theta))\ \mathbf{P}_\Theta(d\theta)]\ \mathbf{P}_U(du) =$$

$$\int [\int \phi(u, t) w_t(u)\ \mu(dt)]\ \mathbf{P}_U(du)$$

- Conclusion: Sampling u from $U \mid T = t$ is obtained by sampling u from U with weight $w_t(u)$.
- $\tau(u, \Theta) = \Theta \circ r(u) \sim \Theta$ if P_Θ is chosen as a right-invariant measure. This does not always exist.

The sampling argument

- Theorem: $(U \mid T = t) \sim w_t(u)\, \mathrm{P}_U(du)$ when $\tau(u, \Theta) \sim w_t(u)\mu(dt)$ is assumed.
- Proof from change-of-variables theorem: $E(\phi(U, T)) =$

$$\int [\int \phi(u, \tau(u, \theta))\ \mathrm{P}_\Theta(d\theta)]\ \mathrm{P}_U(du) =$$

$$\int [\int \phi(u, t) w_t(u)\, \mu(dt)]\ \mathrm{P}_U(du)$$

- **Conclusion: Sampling u from $U \mid T = t$ is obtained by sampling u from U with weight $w_t(u)$.**
- $\tau(u, \Theta) = \Theta \circ r(u) \sim \Theta$ if P_Θ is chosen as a right-invariant measure. This does not always exist.

The sampling argument

- Theorem: $(U \mid T = t) \sim w_t(u) \, \mathrm{P}_U(du)$ when $\tau(u, \Theta) \sim w_t(u) \mu(dt)$ is assumed.
- Proof from change-of-variables theorem: $E(\phi(U, T)) =$

$$\int [\int \phi(u, \tau(u, \theta)) \; \mathrm{P}_\Theta(d\theta)] \; \mathrm{P}_U(du) =$$

$$\int [\int \phi(u, t) w_t(u) \, \mu(dt)] \; \mathrm{P}_U(du)$$

- Conclusion: Sampling u from $U \mid T = t$ is obtained by sampling u from U with weight $w_t(u)$.
- $\tau(u, \Theta) = \Theta \circ r(u) \sim \Theta$ **if** $\mathbf{P}_\Theta$ **is chosen as a right-invariant measure. This does not always exist.**

Table of Contents

Bayes versus Frequentist: No conflict

- **Different interpretations give different conclusions: No problem. Case closed.**
- Bayes: Uncertainty directly for the particular experiment at hand.
- Frequentist: Uncertainty indirectly from properties of the instrument in use.
- Fiducial: Interpretation unproblematic when identified with a special case of frequentist inference = Confidence distributions.
- Bayesian and Fiducial arguments can sometimes be used to obtain excellent (frequentistic) instruments beyond the case of confidence distributions.

Bayes versus Frequentist: No conflict

- Different interpretations give different conclusions: No problem. Case closed.
- **Bayes: Uncertainty directly for the particular experiment at hand.**
- Frequentist: Uncertainty indirectly from properties of the instrument in use.
- Fiducial: Interpretation unproblematic when identified with a special case of frequentist inference = Confidence distributions.
- Bayesian and Fiducial arguments can sometimes be used to obtain excellent (frequentistic) instruments beyond the case of confidence distributions.

Bayes versus Frequentist: No conflict

- Different interpretations give different conclusions: No problem. Case closed.
- Bayes: Uncertainty directly for the particular experiment at hand.
- **Frequentist: Uncertainty indirectly from properties of the instrument in use.**
- Fiducial: Interpretation unproblematic when identified with a special case of frequentist inference = Confidence distributions.
- Bayesian and Fiducial arguments can sometimes be used to obtain excellent (frequentistic) instruments beyond the case of confidence distributions.

Bayes versus Frequentist: No conflict

- Different interpretations give different conclusions: No problem. Case closed.
- Bayes: Uncertainty directly for the particular experiment at hand.
- Frequentist: Uncertainty indirectly from properties of the instrument in use.
- **Fiducial: Interpretation unproblematic when identified with a special case of frequentist inference = Confidence distributions.**
- Bayesian and Fiducial arguments can sometimes be used to obtain excellent (frequentistic) instruments beyond the case of confidence distributions.

Bayes versus Frequentist: No conflict

- Different interpretations give different conclusions: No problem. Case closed.
- Bayes: Uncertainty directly for the particular experiment at hand.
- Frequentist: Uncertainty indirectly from properties of the instrument in use.
- Fiducial: Interpretation unproblematic when identified with a special case of frequentist inference = Confidence distributions.
- **Bayesian and Fiducial arguments can sometimes be used to obtain excellent (frequentistic) instruments beyond the case of confidence distributions.**

Final comments on the involved ideas

And some possibly often forgotten

- **Fiducial distributions can give more than confidence intervals: Good (sometimes optimal frequentist) instruments more generally.**
- A theory with improper priors have been used repeatedly above. This is useful also more generally. It gives for instance precise limit statements involving priors and resolves marginalization type of paradoxes.
- In the above arguments the Monte Carlo law of U is a conditional law given $\Theta = \theta$ that does not depend on θ. As always $(U \mid \Theta = \theta) \not\sim U$, since U is improper.
- Fraser considers models where $U \mid \Theta = \theta$ has a distribution that depends on θ through a shape parameter.
- Much work remains regarding general theory and good examples.

Final comments on the involved ideas

And some possibly often forgotten

- Fiducial distributions can give more than confidence intervals: Good (sometimes optimal frequentist) instruments more generally.
- **A theory with improper priors have been used repeatedly above. This is useful also more generally. It gives for instance precise limit statements involving priors and resolves marginalization type of paradoxes.**
- In the above arguments the Monte Carlo law of U is a conditional law given $\Theta = \theta$ that does not depend on θ. As always $(U \mid \Theta = \theta) \not\sim U$, since U is improper.
- Fraser considers models where $U \mid \Theta = \theta$ has a distribution that depends on θ through a shape parameter.
- Much work remains regarding general theory and good

Final comments on the involved ideas

And some possibly often forgotten

- Fiducial distributions can give more than confidence intervals: Good (sometimes optimal frequentist) instruments more generally.
- A theory with improper priors have been used repeatedly above. This is useful also more generally. It gives for instance precise limit statements involving priors and resolves marginalization type of paradoxes.
- **In the above arguments the Monte Carlo law of U is a conditional law given $\Theta = \theta$ that does not depend on θ. As always $(U \mid \Theta = \theta) \not\sim U$, since U is improper.**
- Fraser considers models where $U \mid \Theta = \theta$ has a distribution that depends on θ through a shape parameter.
- Much work remains regarding general theory and good examples.

Final comments on the involved ideas

And some possibly often forgotten

- Fiducial distributions can give more than confidence intervals: Good (sometimes optimal frequentist) instruments more generally.
- A theory with improper priors have been used repeatedly above. This is useful also more generally. It gives for instance precise limit statements involving priors and resolves marginalization type of paradoxes.
- In the above arguments the Monte Carlo law of U is a conditional law given $\Theta = \theta$ that does not depend on θ. As always $(U \mid \Theta = \theta) \not\sim U$, since U is improper.
- **Fraser considers models where $U \mid \Theta = \theta$ has a distribution that depends on θ through a shape parameter.**
- Much work remains regarding general theory and good

Final comments on the involved ideas

And some possibly often forgotten

- Fiducial distributions can give more than confidence intervals: Good (sometimes optimal frequentist) instruments more generally.
- A theory with improper priors have been used repeatedly above. This is useful also more generally. It gives for instance precise limit statements involving priors and resolves marginalization type of paradoxes.
- In the above arguments the Monte Carlo law of U is a conditional law given $\Theta = \theta$ that does not depend on θ. As always $(U \mid \Theta = \theta) \not\sim U$, since U is improper.
- Fraser considers models where $U \mid \Theta = \theta$ has a distribution that depends on θ through a shape parameter.
- **Much work remains regarding general theory and good examples.**