

Prior-data conflict in likelihood inference

Marco Cattaneo
Department of Statistics, LMU Munich
cattaneo@stat.uni-muenchen.de

WPMSIIP 2010, Durham, UK
7 September 2010

exchangeability

- ▶ in **likelihood inference**, **prior beliefs** can be expressed as a **prior likelihood function**, which is interpreted as if it were induced from **past (virtual) data**:

$$\underbrace{lik_{posterior}(\theta)}_{P_{\theta}(\text{all data})} = \underbrace{lik_{prior}(\theta)}_{P_{\theta}(\text{virtual data})} \underbrace{lik_{data}(\theta)}_{P_{\theta}(\text{actual data})}$$

exchangeability

- ▶ in **likelihood inference**, **prior beliefs** can be expressed as a **prior likelihood function**, which is interpreted as if it were induced from **past (virtual) data**:

$$\underbrace{lik_{posterior}(\theta)}_{P_{\theta}(\text{all data})} = \underbrace{lik_{prior}(\theta)}_{P_{\theta}(\text{virtual data})} \underbrace{lik_{data}(\theta)}_{P_{\theta}(\text{actual data})}$$

- ▶ **example** with Bernoulli data: “**prior model** expects quite certainly 5 successes out of 20, **data** gives 18 successes out of 20”;

exchangeability

- ▶ in **likelihood inference**, **prior beliefs** can be expressed as a **prior likelihood function**, which is interpreted as if it were induced from **past (virtual) data**:

$$\underbrace{lik_{posterior}(\theta)}_{P_{\theta}(\text{all data})} = \underbrace{lik_{prior}(\theta)}_{P_{\theta}(\text{virtual data})} \underbrace{lik_{data}(\theta)}_{P_{\theta}(\text{actual data})}$$

- ▶ **example** with Bernoulli data: “**prior model** expects quite certainly 5 successes out of 20, **data** gives 18 successes out of 20”; for instance:

virtual data:	25	out of	100	(relative frequency: 0.25)
actual data:	18	out of	20	(relative frequency: 0.90)

exchangeability

- ▶ in **likelihood inference**, **prior beliefs** can be expressed as a **prior likelihood function**, which is interpreted as if it were induced from **past (virtual) data**:

$$\underbrace{lik_{posterior}(\theta)}_{P_{\theta}(\text{all data})} = \underbrace{lik_{prior}(\theta)}_{P_{\theta}(\text{virtual data})} \underbrace{lik_{data}(\theta)}_{P_{\theta}(\text{actual data})}$$

- ▶ **example** with Bernoulli data: “**prior model** expects quite certainly 5 successes out of 20, **data** gives 18 successes out of 20”; for instance:

virtual data: 25 out of 100 (relative frequency: 0.25)

actual data: 18 out of 20 (relative frequency: 0.90)

- ▶ if the observations are assumed to be **exchangeable**, then the above situation cannot be distinguished from the following one:

virtual data: 36 out of 100 (relative frequency: 0.36)

actual data: 7 out of 20 (relative frequency: 0.35)

exchangeability

- ▶ in **likelihood inference**, **prior beliefs** can be expressed as a **prior likelihood function**, which is interpreted as if it were induced from **past (virtual) data**:

$$\underbrace{lik_{posterior}(\theta)}_{P_{\theta}(\text{all data})} = \underbrace{lik_{prior}(\theta)}_{P_{\theta}(\text{virtual data})} \underbrace{lik_{data}(\theta)}_{P_{\theta}(\text{actual data})}$$

- ▶ **example** with Bernoulli data: “**prior model** expects quite certainly 5 successes out of 20, **data** gives 18 successes out of 20”; for instance:

virtual data: 25 out of 100 (relative frequency: 0.25)

actual data: 18 out of 20 (relative frequency: 0.90)

- ▶ if the observations are assumed to be **exchangeable**, then the above situation cannot be distinguished from the following one:

virtual data: 36 out of 100 (relative frequency: 0.36)

actual data: 7 out of 20 (relative frequency: 0.35)

- ▶ that is, likelihood inference **cannot** react to **prior-data** conflict when exchangeability is assumed

relaxed exchangeability

- ▶ however, likelihood inference **can** react to **prior-data** conflict when the exchangeability assumption is relaxed

relaxed exchangeability

- ▶ however, likelihood inference **can** react to **prior-data** conflict when the exchangeability assumption is relaxed
- ▶ in the **example** with Bernoulli data, we can assume that the probabilities of success for the **virtual** and **actual** observations are p and q , respectively, without assuming that p and q are necessarily equal:

$$\text{lik}(p, q) \propto p^{25} (1 - p)^{75} f(p, q) q^{18} (1 - q)^2 \quad \text{for all } (p, q) \in [0, 1]^2,$$

where f describes the belief about the similarity of p and q ;

relaxed exchangeability

- ▶ however, likelihood inference **can** react to **prior-data** conflict when the exchangeability assumption is relaxed
- ▶ in the **example** with Bernoulli data, we can assume that the probabilities of success for the **virtual** and **actual** observations are p and q , respectively, without assuming that p and q are necessarily equal:

$$\text{lik}(p, q) \propto p^{25} (1 - p)^{75} f(p, q) q^{18} (1 - q)^2 \quad \text{for all } (p, q) \in [0, 1]^2,$$

where f describes the belief about the similarity of p and q ; for instance:

$$f(p, q) = \exp(-\alpha |\text{logit } p - \text{logit } q|) \quad \text{for all } (p, q) \in [0, 1]^2,$$

with the constant $\alpha \in [0, +\infty]$ describing the the strength of the belief ($\alpha = 0$ means no assumption, $\alpha = +\infty$ means $p = q$)

relaxed exchangeability

- ▶ however, likelihood inference **can** react to **prior-data** conflict when the exchangeability assumption is relaxed
- ▶ in the **example** with Bernoulli data, we can assume that the probabilities of success for the **virtual** and **actual** observations are p and q , respectively, without assuming that p and q are necessarily equal:

$$\text{lik}(p, q) \propto p^{25} (1 - p)^{75} f(p, q) q^{18} (1 - q)^2 \quad \text{for all } (p, q) \in [0, 1]^2,$$

where f describes the belief about the similarity of p and q ; for instance:

$$f(p, q) = \exp(-\alpha |\text{logit } p - \text{logit } q|) \quad \text{for all } (p, q) \in [0, 1]^2,$$

with the constant $\alpha \in [0, +\infty]$ describing the the strength of the belief ($\alpha = 0$ means no assumption, $\alpha = +\infty$ means $p = q$)

- ▶ **profile likelihood** functions for $q \in [0, 1]$ resulting from the original (prior-data conflict) and modified (no prior-data conflict) situations, when $\alpha = 10$:

