Identifying Best Practices for Inter-Laboratory Comparisons and Between-Method Studies

Amanda Koepke

January 2022

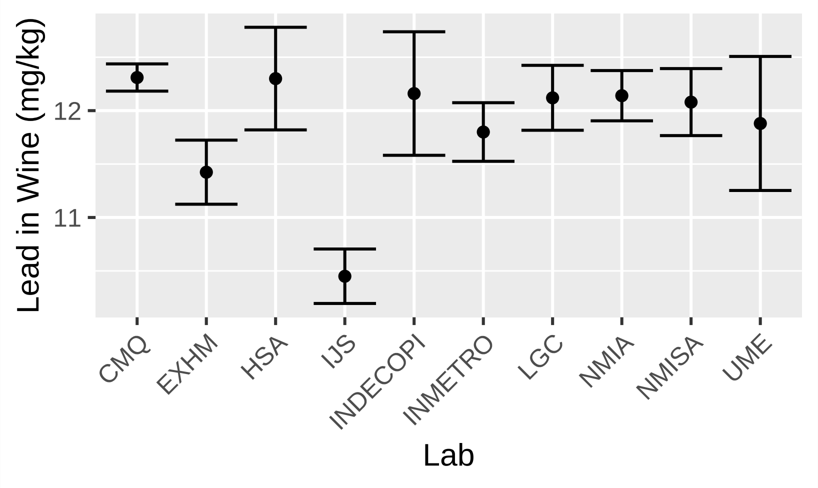
NIST

Overview

- What is an inter-laboratory comparison/between-method study?
- Methods for estimating parameters
  - Frequentist
  - Bayesian
- Performance of methods
- Guidance

What is an inter-lab comparison/between-method study?

Independent measurements from different days, labs, methods



Random effects model

$$x_i = \mu + \lambda_i + \epsilon_i$$

x_i Measured value i

μ Measurand

ϵ_i Measurement error

$\epsilon_i \sim \mathrm{N}(0, \sigma_i^2)$

λ_i Effect of method i

$\lambda_i \sim \mathrm{N}(0, \tau^2)$

or $\lambda_i \sim \mathrm{ST}(0, \tau^2, \alpha, \nu)$

Random effects model

$$x_i = \mu + \lambda_i + \epsilon_i$$

x_i Measured value i

μ Measurand

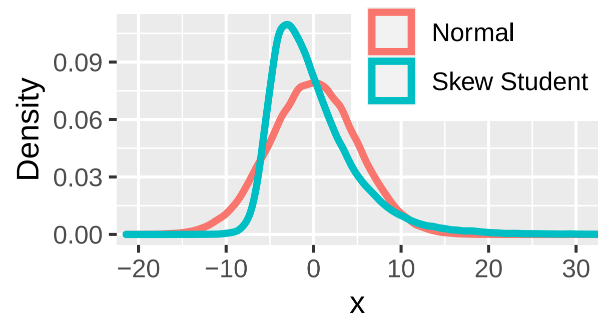
ϵ_i Measurement error

$\epsilon_i \sim \mathrm{N}(0, \sigma_i^2)$

λ_i Effect of method i

$\lambda_i \sim \mathrm{N}(0, \tau^2)$

or $\lambda_i \sim \mathrm{ST}(0, \tau^2, \alpha, \nu)$



$\tau = 5$

$\alpha = 4.5$

$\nu = 6$

Random effects model

$$x_i = \mu + \lambda_i + \epsilon_i$$

x_i Measured value i

μ Measurand

ϵ_i Measurement error

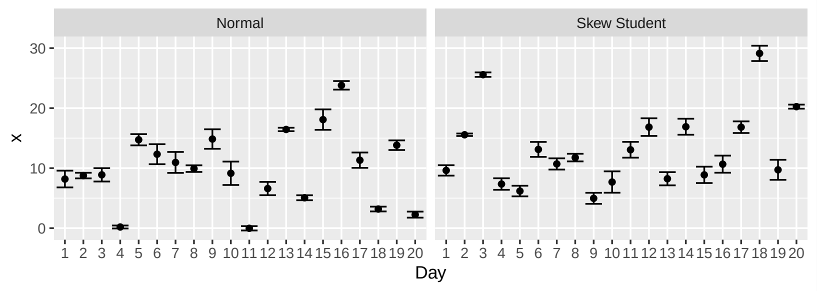
$\epsilon_i \sim \mathrm{N}(0, \sigma_i^2)$

λ_i Effect of study i

$\lambda_i \sim \mathrm{N}(0, \tau^2)$

or $\lambda_i \sim \mathrm{ST}(0, \tau^2, \alpha, \nu)$

$\mu = 10,\ \tau = 5,\ N = 20,\ \sigma_i^2 \sim \mathrm{Uniform}(0.1, 1),\ \alpha = 4.5,\ \nu = 6$



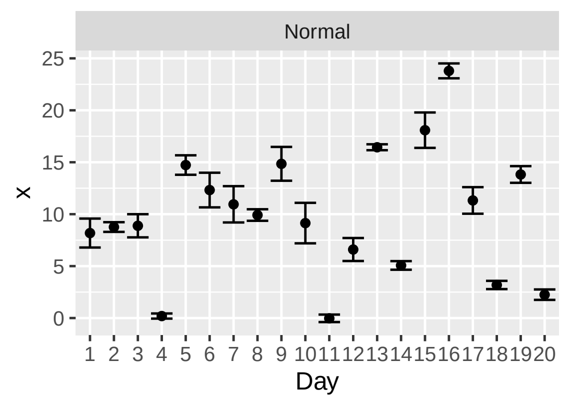
Frequentist approaches

- $\widehat{\mu} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$

  $w_i = \frac{1}{\tau^2 + \sigma_i^2}$

- $u(\mu) = \sqrt{\frac{1}{\sum_{i=1}^{n} w_i}}$

- τ is unknown



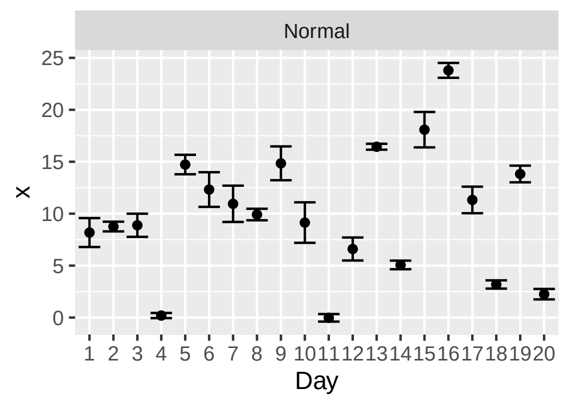
Frequentist approaches

- $\widehat{\mu} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$

  $w_i = \frac{1}{\tau^2 + \sigma_i^2}$

- $u(\mu) = \sqrt{\frac{1}{\sum_{i=1}^{n} w_i}}$

- τ is unknown



- Methods: DerSimonian-Laird, Sidik-Jonkman, restricted maximum-likelihood, and Paule-Mandel

Bayesian analysis

Posterior density: $p(\boldsymbol{\theta}|\boldsymbol{x}, \boldsymbol{\sigma}) \propto p(\boldsymbol{\theta}) \times p(\boldsymbol{x}, \boldsymbol{\sigma}|\boldsymbol{\theta})$

- Unknown: $\boldsymbol{\theta} = (\mu, \tau, \boldsymbol{\lambda})$

Bayesian analysis

Posterior density: $p(\boldsymbol{\theta}|\boldsymbol{x},\boldsymbol{\sigma}) \propto p(\boldsymbol{\theta}) \times p(\boldsymbol{x},\boldsymbol{\sigma}|\boldsymbol{\theta})$

- Unknown: $\boldsymbol{\theta} = (\mu, \tau, \boldsymbol{\lambda})$
- Prior: $p(\boldsymbol{\theta}) = p(\mu)p(\tau)\prod_{i=1}^{n} p(\lambda_i|\tau)$
  - ➠ $\mu \sim \text{Normal}(0, 10^5)$
  - ➠ $\tau \sim \text{half-t}(0, sd \in (2,4), df = 4)$
  - ➠ $\lambda_i|\tau$

Bayesian analysis

Posterior density: $p(\boldsymbol{\theta}|\boldsymbol{x},\boldsymbol{\sigma}) \propto p(\boldsymbol{\theta}) \times p(\boldsymbol{x},\boldsymbol{\sigma}|\boldsymbol{\theta})$

- Unknown: $\boldsymbol{\theta} = (\mu, \tau, \boldsymbol{\lambda})$
- Prior: $p(\boldsymbol{\theta}) = p(\mu)p(\tau)\prod_{i=1}^{n} p(\lambda_i|\tau)$
  - ➠ $\mu \sim \text{Normal}(0, 10^5)$
  - ➠ $\tau \sim \text{half-t}(0, sd \in (2,4), df = 4)$
  - ➠ $\lambda_i|\tau$
    - Gauss+Gauss: $\lambda_i|\tau \sim \text{Normal}(0, \tau^2)$
    - Skew Student+Gauss: $\lambda_i|\tau \sim \text{Skew Student}(0, \tau^2, \alpha, \nu)$
      - $\alpha \sim \text{Normal}(0, 4^2)$, 95% between (-7.8, 7.8)
      - $\nu \sim$ Gamma(shape=3, rate=0.25), 95%: (2.5, 26)
    - Laplace+Gauss: $\lambda_i|\tau \sim$ Laplace(0, variance=τ^2)

Bayesian analysis

Posterior density: $p(\boldsymbol{\theta}|\boldsymbol{x}, \boldsymbol{\sigma}) \propto p(\boldsymbol{\theta}) \times p(\boldsymbol{x}, \boldsymbol{\sigma}|\boldsymbol{\theta})$

- Unknown: $\boldsymbol{\theta} = (\mu, \tau, \boldsymbol{\lambda})$
- Prior: $p(\boldsymbol{\theta}) = p(\mu)p(\tau)\prod_{i=1}^{n} p(\lambda_i|\tau)$
  - ➠ $\mu \sim \text{Normal}(0, 10^5)$
  - ➠ $\tau \sim \text{half-t}(0, sd \in (2, 4), df = 4)$
  - ➠ $\lambda_i|\tau$
- Likelihood $p(\boldsymbol{x}, \boldsymbol{\sigma}|\boldsymbol{\theta})$: $x_i, \sigma_i|\mu, \lambda_i \sim \text{Normal}(\mu + \lambda_i, \sigma_i^2)$

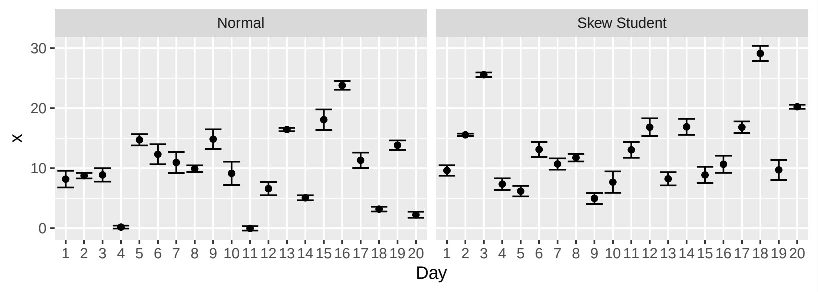
Bayesian analysis

Posterior density: $p(\boldsymbol{\theta}|\boldsymbol{x},\boldsymbol{\sigma}) \propto p(\boldsymbol{\theta}) \times p(\boldsymbol{x},\boldsymbol{\sigma}|\boldsymbol{\theta})$

- Unknown: $\boldsymbol{\theta} = (\mu, \tau, \boldsymbol{\lambda})$
- Prior: $p(\boldsymbol{\theta}) = p(\mu)p(\tau)\prod_{i=1}^{n} p(\lambda_i|\tau)$
  - ➠ $\mu \sim \text{Normal}(0, 10^5)$
  - ➠ $\tau \sim \text{half-t}(0, sd \in (2,4), df = 4)$
  - ➠ $\lambda_i|\tau$
- Likelihood $p(\boldsymbol{x},\boldsymbol{\sigma}|\boldsymbol{\theta})$: $x_i, \sigma_i|\mu, \lambda_i \sim \text{Normal}(\mu + \lambda_i, \sigma_i^2)$

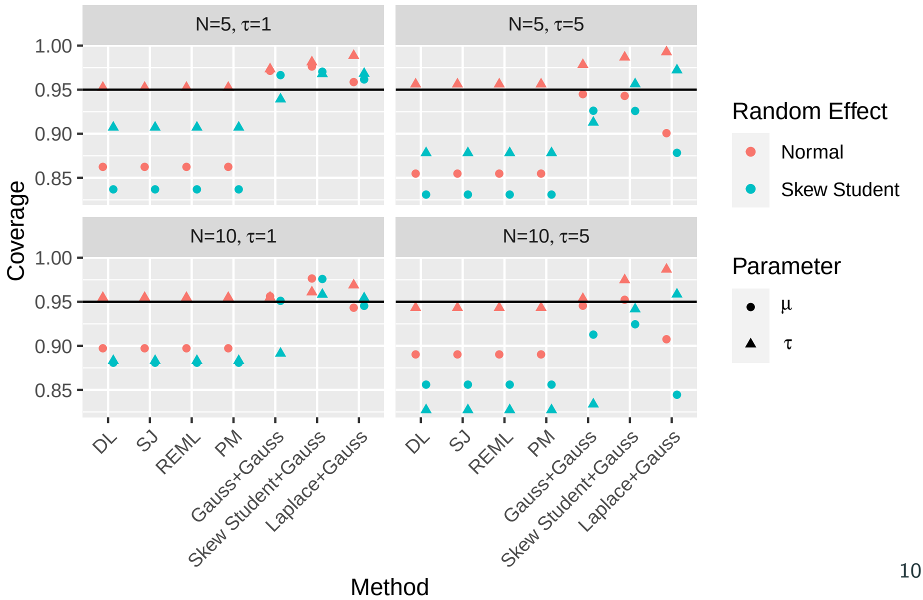
Simulate from the posterior distribution using Hamiltonian Monte Carlo.

Experiments



- $N \in (5, 10, 20, 50)$
- $\tau \in (1, 5)$
- Distribution of the random effect
- 1000 datasets each

Coverage Probability



Conclusions

- The frequentist estimators perform well when the sample size is large ($N \geq 20$)
- For more realistic settings the Bayesian estimators perform best
- Bayesian approach more flexible, allows for complex data structures

Conclusions

- The frequentist estimators perform well when the sample size is large ($N \geq 20$)
- For more realistic settings the Bayesian estimators perform best
- Bayesian approach more flexible, allows for complex data structures

Thank you! Questions? Contact: amanda.koepke@nist.gov