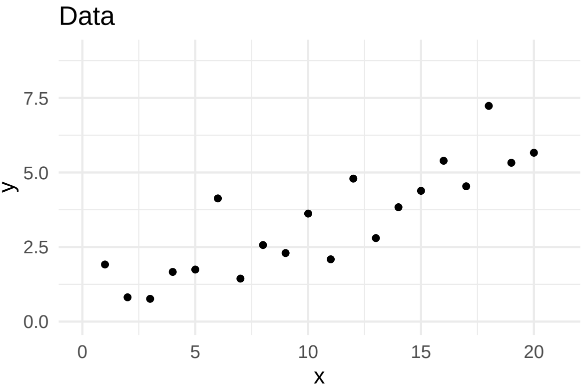
UPPSALA UNIVERSITET

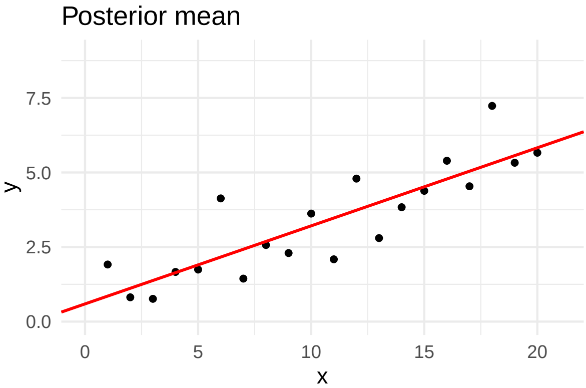
Bayesian Statistics and Data Analysis
Lecture 3

Måns Magnusson
Department of Statistics, Uppsala University
Thanks to Aki Vehtari, Aalto University

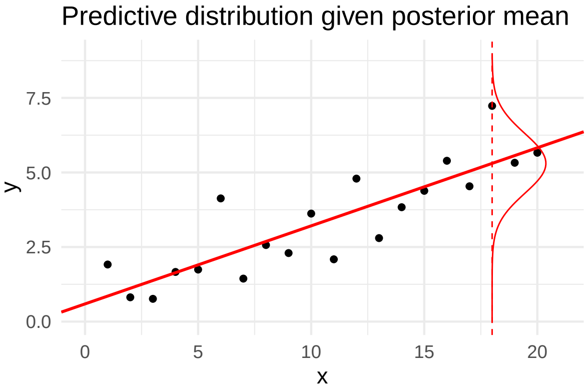
Example of uncertainty in modeling



Example of uncertainty in modeling



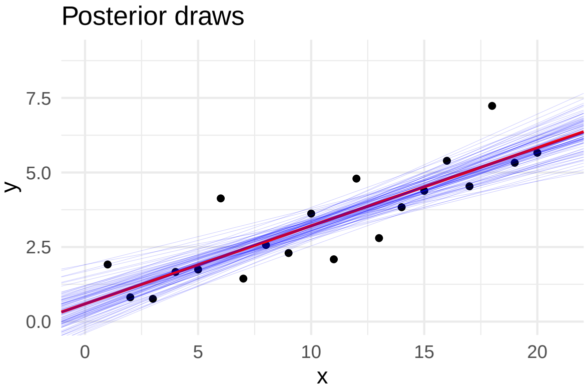
Example of uncertainty in modeling



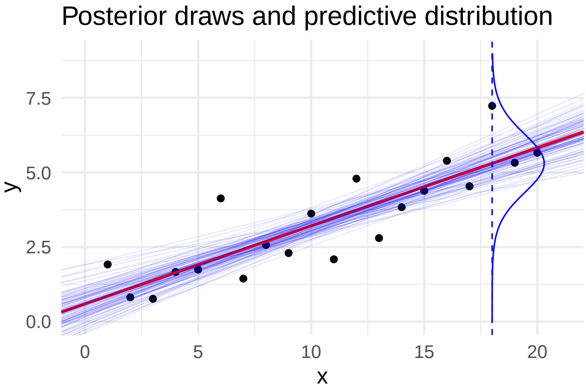
UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

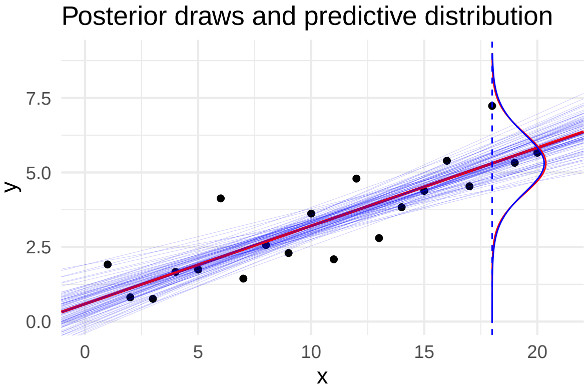
Example of uncertainty in modeling



Example of uncertainty in modeling



Example of uncertainty in modeling



UPPSALA UNIVERSITET

Monte Carlo and Posterior Draws

- Assume we can get draws from $p(\theta \mid y)$

Monte Carlo and Posterior Draws

- Assume we can get draws from $p(\theta \mid y)$
- $\theta^{(s)}$ draws from $p(\theta \mid y)$ can be used
  - for visualization
  - to approximate expectations (integrals)

$$E_{p(\theta|y)}[\theta] = \int \theta p(\theta \mid y) \approx \frac{1}{S} \sum_{s=1}^{S} \theta^{(s)}$$

Monte Carlo and Posterior Draws

- Assume we can get draws from $p(\theta \mid y)$
- $\theta^{(s)}$ draws from $p(\theta \mid y)$ can be used
  - for visualization
  - to approximate expectations (integrals)

$$E_{p(\theta|y)}[\theta] = \int \theta p(\theta \mid y) \approx \frac{1}{S} \sum_{s=1}^{S} \theta^{(s)}$$

- to approximate uncertainty intervals for θ

UPPSALA UNIVERSITET

Marginalization

- Joint posterior distribution of multiple parameters

$$p(\theta_1, \theta_2 \mid y) \propto p(y \mid \theta_1, \theta_2) p(\theta_1, \theta_2)$$

Marginalization

- Joint posterior distribution of multiple parameters

$$p(\theta_1, \theta_2 \mid y) \propto p(y \mid \theta_1, \theta_2) p(\theta_1, \theta_2)$$

- Marginalization

$$p(\theta_1 \mid y) = \int p(\theta_1, \theta_2 \mid y) d\theta_2$$

$p(\theta_1 \mid y)$ is a marginal distribution

UPPSALA UNIVERSITET

Marginalization

- Joint posterior distribution of multiple parameters

$$p(\theta_1, \theta_2 \mid y) \propto p(y \mid \theta_1, \theta_2) p(\theta_1, \theta_2)$$

- Marginalization

$$p(\theta_1 \mid y) = \int p(\theta_1, \theta_2 \mid y) d\theta_2$$

$p(\theta_1 \mid y)$ is a marginal distribution

- Goal is often to find marginal posterior of an interesting quantity
  - a parameter $p(\theta|y)$
  - a potential observation $p(\tilde{y}|y)$

Marginalization - predictive distribution

- Joint distribution of unknown future observation and parameters

$$
\begin{aligned}
p(\tilde{y}, \theta \mid y) &= p(\tilde{y} \mid \theta, y) p(\theta \mid y) \\
&= p(\tilde{y} \mid \theta) p(\theta \mid y) \qquad \text{(often)}
\end{aligned}
$$

Marginalization - predictive distribution

- Joint distribution of unknown future observation and parameters

$$\begin{aligned} p(\tilde{y}, \theta \mid y) &= p(\tilde{y} \mid \theta, y) p(\theta \mid y) \\ &= p(\tilde{y} \mid \theta) p(\theta \mid y) \qquad \text{(often)} \end{aligned}$$

- Marginalization over posterior distribution

$$\begin{aligned} p(\tilde{y} \mid y) &= \int p(\tilde{y} \mid \theta) p(\theta \mid y) d\theta \\ &= \int p(\tilde{y}, \theta \mid y) d\theta \end{aligned}$$

$p(\tilde{y} \mid y)$ is a predictive distribution

Gaussian with unknown μ and σ^2

UPPSALA UNIVERSITET

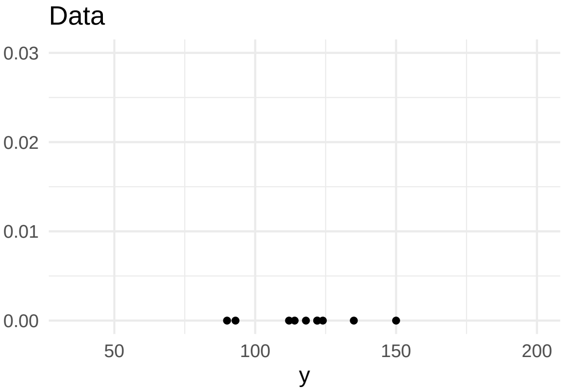
- Observation model

$$\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right)$$

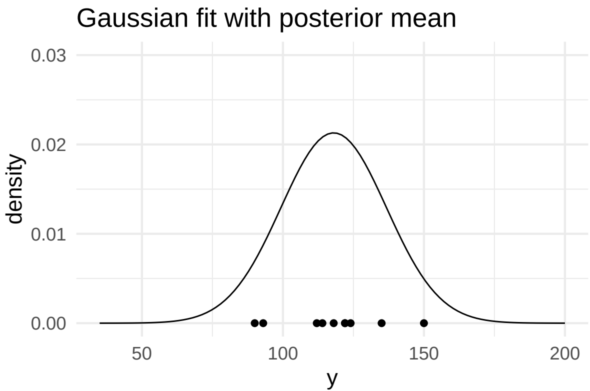
- Uninformative prior

$$p(\mu, \sigma^2) \propto \sigma^{-2}$$

Gaussian example



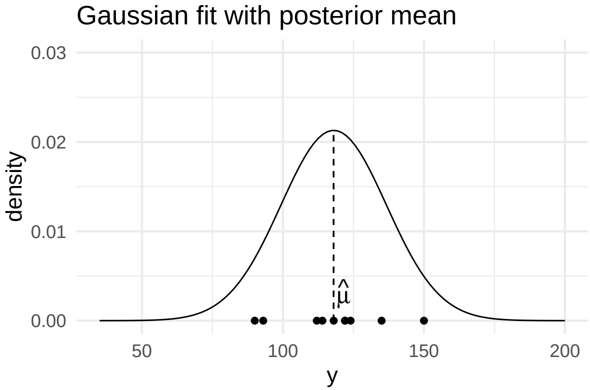
Gaussian example



$$p(y \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right)$$

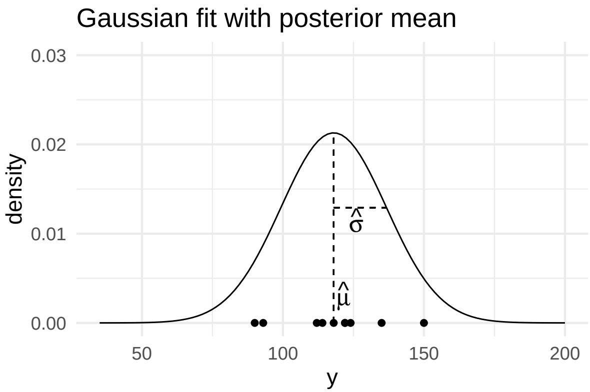
UPPSALA UNIVERSITET

Gaussian example



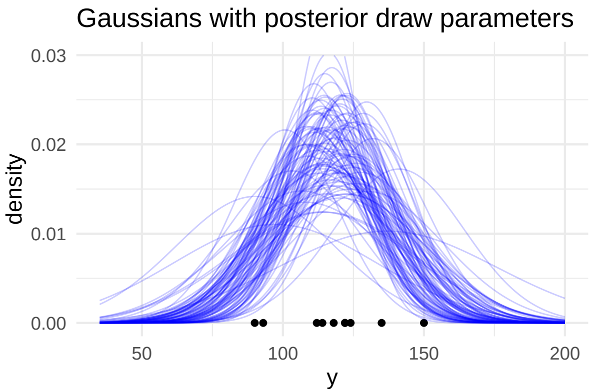
$$p(y \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right)$$

Gaussian example



$$p(y \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right)$$

Gaussian example

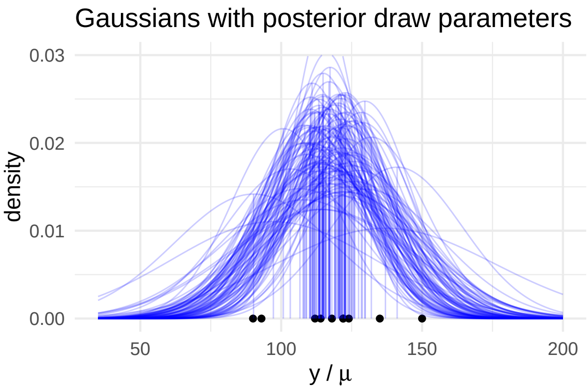


$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

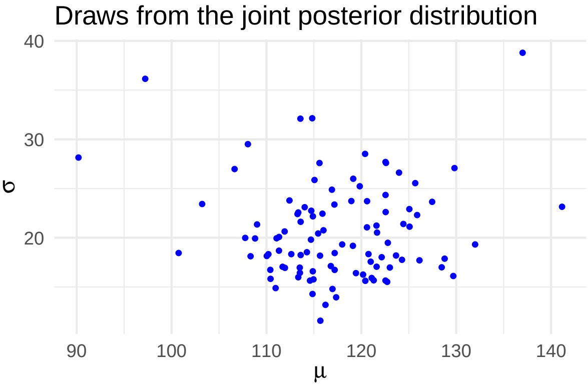
Gaussian example



$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

UPPSALA UNIVERSITET

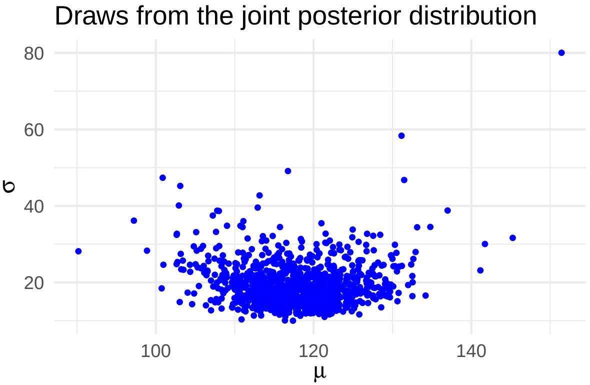
Gaussian example



$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

UPPSALA UNIVERSITET

Gaussian example

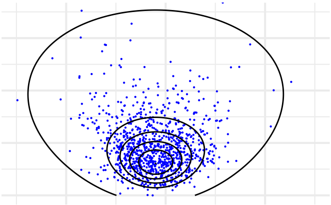


$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

Joint posterior



$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

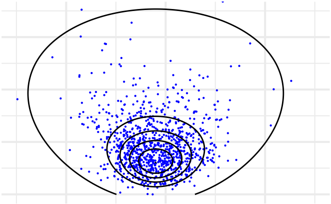
UPPSALA UNIVERSITET

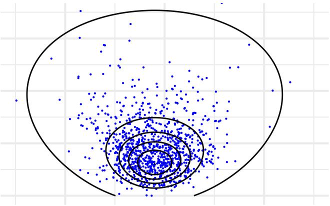
- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

Joint posterior

$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

with $p(\mu, \sigma^2) \propto \sigma^{-2}$





Joint posterior

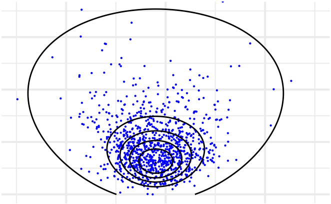
$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

with $p(\mu, \sigma^2) \propto \sigma^{-2}$

$$p(\mu, \sigma^2 \mid y) \propto \sigma^{-2} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y_i - \mu)^2\right)$$

UPPSALA UNIVERSITET

Joint posterior

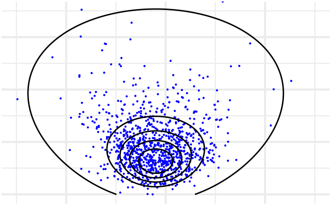


$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

with $p(\mu, \sigma^2) \propto \sigma^{-2}$

$$p(\mu, \sigma^2 \mid y) \propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2\right)$$

Joint posterior



$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

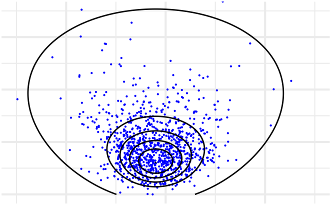
with $p(\mu, \sigma^2) \propto \sigma^{-2}$

$$p(\mu, \sigma^2 \mid y) \propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2\right)$$

$$= \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(y_i - \bar{y})^2 + n(\bar{y} - \mu)^2\right]\right)$$

$$\text{where } \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Joint posterior



$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

with $p(\mu, \sigma^2) \propto \sigma^{-2}$

$$p(\mu, \sigma^2 \mid y) \propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \mu)^2\right)$$

$$= \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2} \left[\sum_{i=1}^{n} (y_i - \bar{y})^2 + n(\bar{y} - \mu)^2\right]\right)$$

$$\text{where } \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

$$= \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2} \left[(n-1)s^2 + n(\bar{y} - \mu)^2\right]\right)$$

$$\text{where } s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Gaussian: Completing the square

$$\sum_{i=1}^{n}(y_i - \mu)^2$$

UPPSALA UNIVERSITET

Gaussian: Completing the square

$$\sum_{i=1}^{n}(y_i - \mu)^2$$

$$\sum_{i=1}^{n}(y_i^2 - 2y_i\mu + \mu^2)$$

UPPSALA UNIVERSITET

Gaussian: Completing the square

$$\sum_{i=1}^{n}(y_i-\mu)^2$$

$$\sum_{i=1}^{n}(y_i^2-2y_i\mu+\mu^2)$$

$$\sum_{i=1}^{n}(y_i^2-2y_i\mu+\mu^2-\bar{y}^2+\bar{y}^2-2y_i\bar{y}+2y_i\bar{y})$$

UPPSALA UNIVERSITET

Gaussian: Completing the square

$$\sum_{i=1}^{n}(y_i - \mu)^2$$

$$\sum_{i=1}^{n}(y_i^2 - 2y_i\mu + \mu^2)$$

$$\sum_{i=1}^{n}(y_i^2 - 2y_i\mu + \mu^2 - \bar{y}^2 + \bar{y}^2 - 2y_i\bar{y} + 2y_i\bar{y})$$

$$\sum_{i=1}^{n}(y_i^2 - 2y_i\bar{y} + \bar{y}^2) + \sum_{i=1}^{n}(\mu^2 - 2y_i\mu - \bar{y}^2 + 2y_i\bar{y})$$

Gaussian: Completing the square

$$\sum_{i=1}^{n}(y_i - \mu)^2$$

$$\sum_{i=1}^{n}(y_i^2 - 2y_i\mu + \mu^2)$$

$$\sum_{i=1}^{n}(y_i^2 - 2y_i\mu + \mu^2 - \bar{y}^2 + \bar{y}^2 - 2y_i\bar{y} + 2y_i\bar{y})$$

$$\sum_{i=1}^{n}(y_i^2 - 2y_i\bar{y} + \bar{y}^2) + \sum_{i=1}^{n}(\mu^2 - 2y_i\mu - \bar{y}^2 + 2y_i\bar{y})$$

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 + n(\mu^2 - 2\bar{y}\mu - \bar{y}^2 + 2\bar{y}\bar{y})$$

Gaussian: Completing the square

$$\sum_{i=1}^{n}(y_i-\mu)^2$$

$$\sum_{i=1}^{n}(y_i^2-2y_i\mu+\mu^2)$$

$$\sum_{i=1}^{n}(y_i^2-2y_i\mu+\mu^2-\bar{y}^2+\bar{y}^2-2y_i\bar{y}+2y_i\bar{y})$$

$$\sum_{i=1}^{n}(y_i^2-2y_i\bar{y}+\bar{y}^2)+\sum_{i=1}^{n}(\mu^2-2y_i\mu-\bar{y}^2+2y_i\bar{y})$$

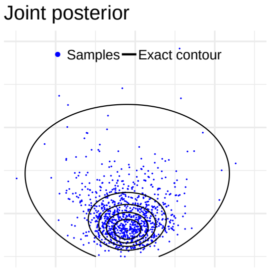
$$\sum_{i=1}^{n}(y_i-\bar{y})^2+n(\mu^2-2\bar{y}\mu-\bar{y}^2+2\bar{y}\bar{y})$$

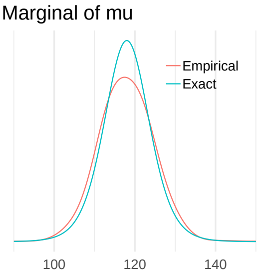
$$\sum_{i=1}^{n}(y_i-\bar{y})^2+n(\bar{y}-\mu)^2$$

UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

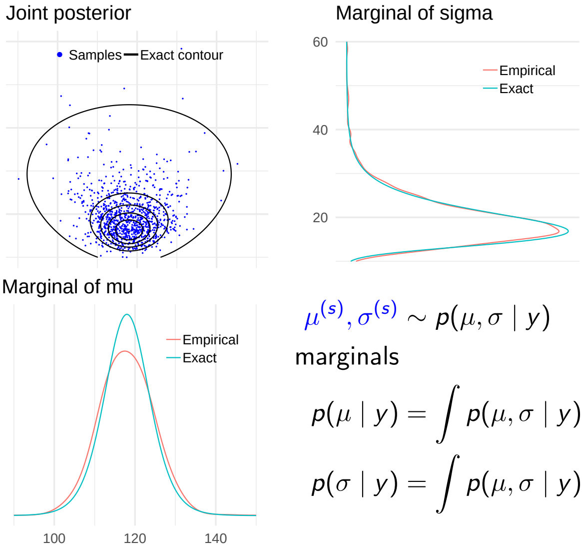
Marginal $p(\mu)$ and $p(\sigma^2)$





$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

Marginal $p(\mu)$ and $p(\sigma^2)$



$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

marginals

$$p(\mu \mid y) = \int p(\mu, \sigma \mid y) d\sigma$$

$$p(\sigma \mid y) = \int p(\mu, \sigma \mid y) d\mu$$

Marginal $p(\mu|y)$ and $p(\sigma^2|y)$

Marginal posterior $p(\sigma^2 \mid y)$ (easier for σ^2 than σ)

$$p(\sigma^2 \mid y) \quad \propto \quad \int p(\mu, \sigma^2 \mid y) d\mu$$

UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

Marginal $p(\mu|y)$ and $p(\sigma^2|y)$

Marginal posterior $p(\sigma^2 \mid y)$ (easier for σ^2 than σ)

$$\begin{aligned} p(\sigma^2 \mid y) &\propto \int p(\mu, \sigma^2 \mid y) d\mu \\ &\propto \int \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\mu \end{aligned}$$

Marginal $p(\mu|y)$ and $p(\sigma^2|y)$

Marginal posterior $p(\sigma^2 \mid y)$ (easier for σ^2 than σ)

$$
\begin{aligned}
p(\sigma^2 \mid y) &\propto \int p(\mu, \sigma^2 \mid y) d\mu \\
&\propto \int \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\mu \\
&\propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}(n-1)s^2\right) \cdot \\
&\quad \int \exp\left(-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2\right) d\mu
\end{aligned}
$$

Marginal $p(\mu|y)$ and $p(\sigma^2|y)$

Marginal posterior $p(\sigma^2 \mid y)$ (easier for σ^2 than σ)

$$
\begin{aligned}
p(\sigma^2 \mid y) &\propto \int p(\mu, \sigma^2 \mid y) d\mu \\
&\propto \int \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\mu \\
&\propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}(n-1)s^2\right) \cdot \\
&\quad \int \exp\left(-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2\right) d\mu
\end{aligned}
$$

$$
\textit{Note!} \quad \int \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right) d\mu = 1
$$

Marginal $p(\mu|y)$ and $p(\sigma^2|y)$

Marginal posterior $p(\sigma^2 \mid y)$ (easier for σ^2 than σ)

$$
\begin{aligned}
p(\sigma^2 \mid y) &\propto \int p(\mu, \sigma^2 \mid y) d\mu \\
&\propto \int \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\mu \\
&\propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}(n-1)s^2\right) \cdot \\
&\quad \int \exp\left(-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2\right) d\mu \\
\textit{Note!} &\quad \int \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right) d\mu = 1 \\
&\propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}(n-1)s^2\right) \sqrt{2\pi\sigma^2/n}
\end{aligned}
$$

UPPSALA UNIVERSITET

Marginal $p(\mu|y)$ and $p(\sigma^2|y)$

Marginal posterior $p(\sigma^2 \mid y)$ (easier for σ^2 than σ)

$$
\begin{aligned}
p(\sigma^2 \mid y) &\propto \int p(\mu, \sigma^2 \mid y) d\mu \\
&\propto \int \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\mu \\
&\propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}(n-1)s^2\right) \cdot \\
&\quad \int \exp\left(-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2\right) d\mu \\
\textit{Note!} &\quad \int \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right) d\mu = 1 \\
&\propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}(n-1)s^2\right) \sqrt{2\pi\sigma^2/n} \\
&\propto (\sigma^2)^{-(n+1)/2} \exp\left(-\frac{(n-1)s^2}{2\sigma^2}\right)
\end{aligned}
$$

Marginal $p(\mu|y)$ and $p(\sigma^2|y)$

Marginal posterior $p(\sigma^2 \mid y)$ (easier for σ^2 than σ)

$$
\begin{aligned}
p(\sigma^2 \mid y) &\propto \int p(\mu, \sigma^2 \mid y) d\mu \\
&\propto \int \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\mu \\
&\propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}(n-1)s^2\right) \cdot \\
&\qquad \int \exp\left(-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2\right) d\mu \\
\textit{Note!} &\qquad \int \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right) d\mu = 1 \\
&\propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}(n-1)s^2\right) \sqrt{2\pi\sigma^2/n} \\
&\propto (\sigma^2)^{-(n+1)/2} \exp\left(-\frac{(n-1)s^2}{2\sigma^2}\right) \\
p(\sigma^2 \mid y) &= \text{Inv-}\chi^2(\sigma^2 \mid n-1, s^2)
\end{aligned}
$$

UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

Gaussian - non-informative prior

Known mean

$$\sigma^2 \mid y \sim \text{Inv-}\chi^2(n, v)$$

$$\text{where} \quad v = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2$$

UPPSALA UNIVERSITET

Gaussian - non-informative prior

Known mean

$$\sigma^2 \mid y \sim \text{Inv-}\chi^2(n, v)$$

$$\text{where} \quad v = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2$$

Unknown mean

$$\sigma^2 \mid y \sim \text{Inv-}\chi^2(n-1, s^2)$$

$$\text{where} \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

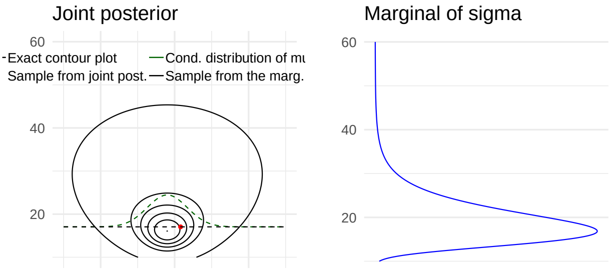
UPPSALA UNIVERSITET

Gaussian - non-informative prior

- Marginal posterior $p(\mu \mid y)$

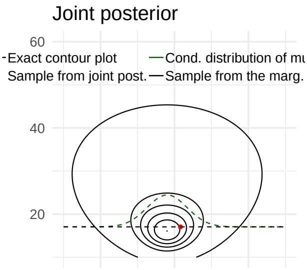
$$p(\mu \mid y) = \int_0^{\infty} p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y) d\sigma^2$$

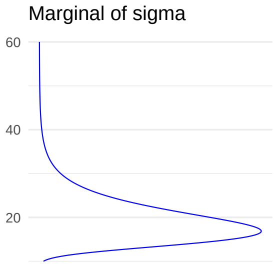
- Marginal posterior of μ, a mixture of normal distributions where mixing density is the marginal posterior of σ^2



Factorization

$$p(\mu, \sigma^2 \mid y) = {\color{darkred} p(\mu \mid \sigma^2, y)}{\color{blue} p(\sigma^2 \mid y)}$$

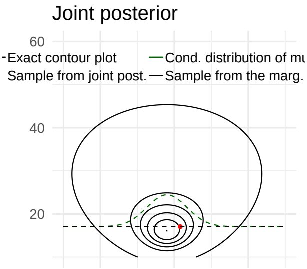


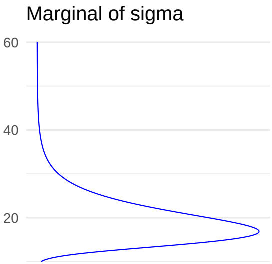


Factorization

$$p(\mu, \sigma^2 \mid y) = p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y)$$

$$(\sigma^2)^{(s)} \sim p(\sigma^2 \mid y)$$





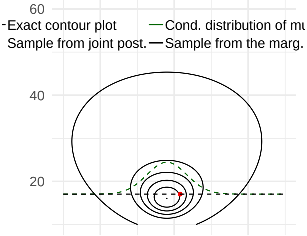
Factorization

$$p(\mu, \sigma^2 \mid y) = p(\mu \mid \sigma^2, y)p(\sigma^2 \mid y)$$

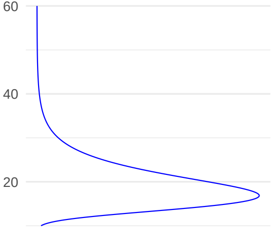
$$(\sigma^2)^{(s)} \sim p(\sigma^2 \mid y)$$

$$p(\mu \mid (\sigma^2)^{(s)}, y) = \mathcal{N}(\mu \mid \bar{y}, (\sigma^2)^{(s)}/n)$$

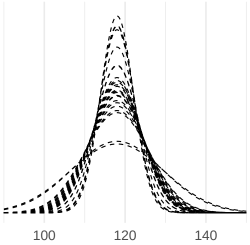
Joint posterior



Marginal of sigma



Cond distr of mu for 25 draws

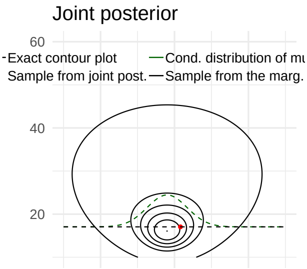


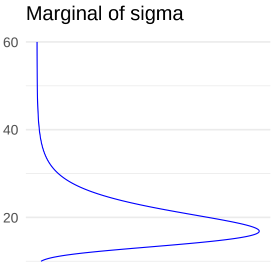
Factorization

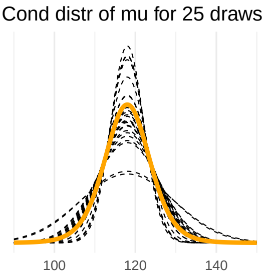
$$p(\mu, \sigma^2 \mid y) = p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y)$$

$$(\sigma^2)^{(s)} \sim p(\sigma^2 \mid y)$$

$$p(\mu \mid (\sigma^2)^{(s)}, y) = \mathcal{N}(\mu \mid \bar{y}, (\sigma^2)^{(s)}/n)$$







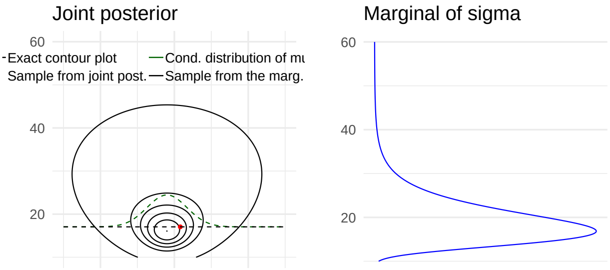
Factorization

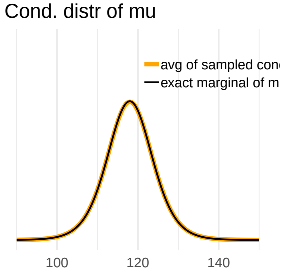
$$p(\mu, \sigma^2 \mid y) = p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y)$$

$$(\sigma^2)^{(s)} \sim p(\sigma^2 \mid y)$$

$$p(\mu \mid (\sigma^2)^{(s)}, y) = \mathcal{N}(\mu \mid \bar{y}, (\sigma^2)^{(s)}/n)$$

$$p(\mu \mid y) \approx \frac{1}{S} \sum_{s=1}^{S} \mathcal{N}(\mu \mid \bar{y}, (\sigma^2)^{(s)}$$





Factorization

$$p(\mu, \sigma^2 \mid y) = p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y)$$

$$(\sigma^2)^{(s)} \sim p(\sigma^2 \mid y)$$

$$p(\mu \mid (\sigma^2)^{(s)}, y) = \mathcal{N}(\mu \mid \bar{y}, (\sigma^2)^{(s)}/n)$$

$$p(\mu \mid y) \approx \frac{1}{S} \sum_{s=1}^{S} \mathcal{N}(\mu \mid \bar{y}, (\sigma^2)^{(s)}$$

Marginal posterior $p(\mu \mid y)$

$$p(\mu \mid y) = \int_0^\infty p(\mu, \sigma^2 \mid y) d\sigma^2$$

Marginal posterior $p(\mu \mid y)$

$$
\begin{aligned}
p(\mu \mid y) &= \int_0^\infty p(\mu, \sigma^2 \mid y) d\sigma^2 \\
&\propto \int_0^\infty \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\sigma^2
\end{aligned}
$$

UPPSALA UNIVERSITET

Marginal posterior $p(\mu \mid y)$

$$p(\mu \mid y) = \int_0^\infty p(\mu, \sigma^2 \mid y) d\sigma^2$$

$$\propto \int_0^\infty \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\sigma^2$$

Transformation (integration by substitution)

$$A = (n-1)s^2 + n(\mu - \bar{y})^2$$

Marginal posterior $p(\mu \mid y)$

$$p(\mu \mid y) = \int_0^\infty p(\mu, \sigma^2 \mid y) d\sigma^2$$

$$\propto \int_0^\infty \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\sigma^2$$

Transformation (integration by substitution)

$$A = (n-1)s^2 + n(\mu - \bar{y})^2 \quad \text{and} \quad z = \frac{A}{2\sigma^2}$$

$$dz = \left(-\frac{A}{2(\sigma^2)^2}\right) d\sigma^2$$

UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

Marginal posterior $p(\mu \mid y)$

$$p(\mu \mid y) = \int_0^\infty p(\mu, \sigma^2 \mid y) d\sigma^2$$

$$\propto \int_0^\infty \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\sigma^2$$

Transformation (integration by substitution)

$$A = (n-1)s^2 + n(\mu - \bar{y})^2 \quad \text{and} \quad z = \frac{A}{2\sigma^2}$$

$$dz = \left(-\frac{A}{2(\sigma^2)^2}\right) d\sigma^2$$

$$p(\mu \mid y) \propto A^{-n/2} \int_0^\infty z^{(n-2)/2} \exp(-z) dz$$

Marginal posterior $p(\mu \mid y)$

$$
\begin{aligned}
p(\mu \mid y) &= \int_0^\infty p(\mu, \sigma^2 \mid y) d\sigma^2 \\
&\propto \int_0^\infty \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\sigma^2
\end{aligned}
$$

Transformation (integration by substitution)

$$
A = (n-1)s^2 + n(\mu - \bar{y})^2 \quad \text{and} \quad z = \frac{A}{2\sigma^2}
$$

$$
dz = \left(-\frac{A}{2(\sigma^2)^2}\right) d\sigma^2
$$

$$
p(\mu \mid y) \propto A^{-n/2} \int_0^\infty z^{(n-2)/2} \exp(-z) dz
$$

Recognize gamma integral $\Gamma(u) = \int_0^\infty x^{u-1} \exp(-x) dx$

Marginal posterior $p(\mu \mid y)$

$$p(\mu \mid y) = \int_0^\infty p(\mu, \sigma^2 \mid y) d\sigma^2$$

$$\propto \int_0^\infty \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y}-\mu)^2\right]\right) d\sigma^2$$

Transformation (integration by substitution)

$$A = (n-1)s^2 + n(\mu - \bar{y})^2 \quad \text{and} \quad z = \frac{A}{2\sigma^2}$$

$$dz = \left(-\frac{A}{2(\sigma^2)^2}\right) d\sigma^2$$

$$p(\mu \mid y) \propto A^{-n/2} \int_0^\infty z^{(n-2)/2} \exp(-z) dz$$

Recognize gamma integral $\Gamma(u) = \int_0^\infty x^{u-1} \exp(-x) dx$

$$\propto \left[(n-1)s^2 + n(\mu - \bar{y})^2\right]^{-n/2}$$

UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

Marginal posterior $p(\mu \mid y)$

$$p(\mu|y) \propto [(n-1)s^2 + n(\mu - \bar{y})^2]^{-n/2}$$

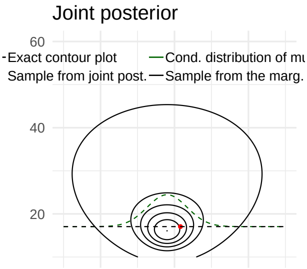
Marginal posterior $p(\mu \mid y)$

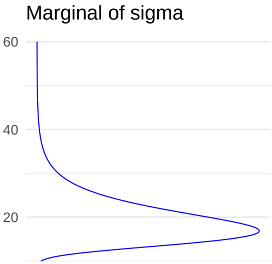
$$
\begin{aligned}
p(\mu|y) &\propto [(n-1)s^2 + n(\mu - \bar{y})^2]^{-n/2} \\
&\propto \left[1 + \frac{n(\mu - \bar{y})^2}{(n-1)s^2}\right]^{-n/2}
\end{aligned}
$$

UPPSALA UNIVERSITET

Marginal posterior $p(\mu \mid y)$

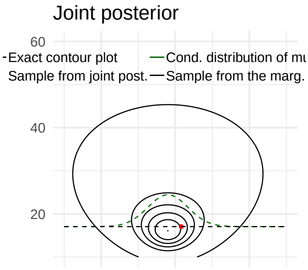
$$
\begin{aligned}
p(\mu|y) &\propto [(n-1)s^2 + n(\mu - \bar{y})^2]^{-n/2} \\
&\propto \left[1 + \frac{n(\mu - \bar{y})^2}{(n-1)s^2}\right]^{-n/2} \\
p(\mu \mid y) &= t_{n-1}(\mu \mid \bar{y}, s^2/n) \quad \text{Student's } t
\end{aligned}
$$

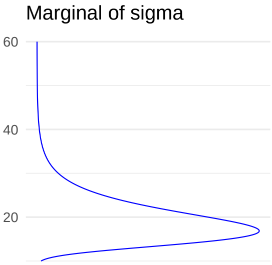




Predictive distribution for new $\tilde{y}$

$$p(\tilde{y}|y) = \int p(\tilde{y}|\mu,\sigma)p(\mu,\sigma|y)d\mu\sigma$$

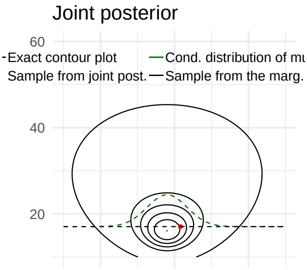


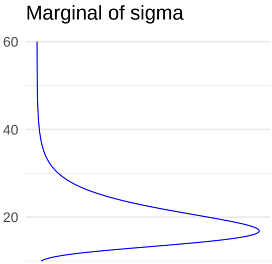


Predictive distribution for new $\tilde{y}$

$$p(\tilde{y}|y) = \int p(\tilde{y}|\mu,\sigma)p(\mu,\sigma|y)d\mu\sigma$$

$$\mu^{(s)},\sigma^{(s)} \sim p(\mu,\sigma|y)$$



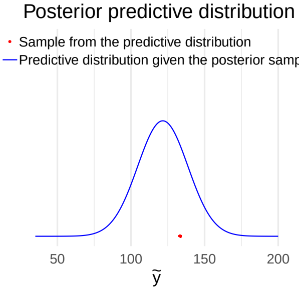


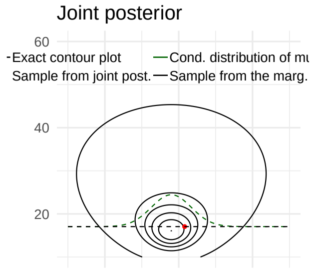
Predictive distribution for new $\tilde{y}$

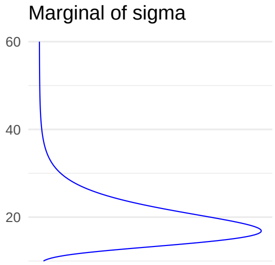
$$p(\tilde{y}|y) = \int p(\tilde{y}|\mu,\sigma)p(\mu,\sigma|y)d\mu\sigma$$

$$\mu^{(s)},\sigma^{(s)} \sim p(\mu,\sigma|y)$$

$$\tilde{y}^{(s)} \sim p(\tilde{y}|\mu^{(s)},\sigma^{(s)})$$





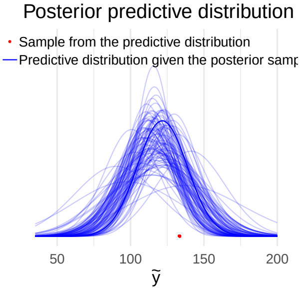


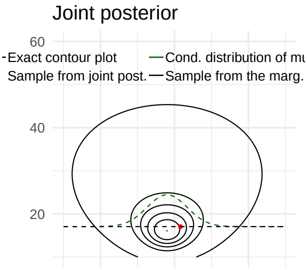
Predictive distribution for new $\tilde{y}$

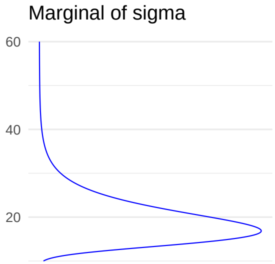
$$p(\tilde{y}|y) = \int p(\tilde{y}|\mu,\sigma)p(\mu,\sigma|y)d\mu\sigma$$

$$\color{red}{\mu^{(s)},\sigma^{(s)}} \sim p(\mu,\sigma|y)$$

$$\color{red}{\tilde{y}^{(s)}} \sim \color{blue}{p(\tilde{y}|\mu^{(s)},\sigma^{(s)})}$$





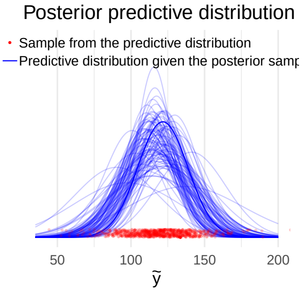


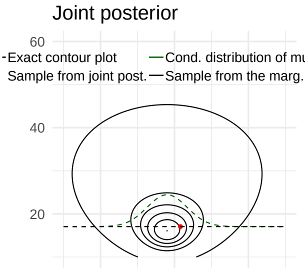
Predictive distribution for new $\tilde{y}$

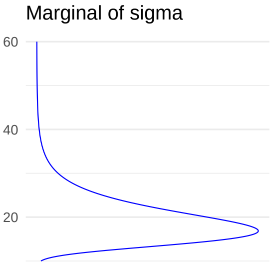
$$p(\tilde{y}|y) = \int p(\tilde{y}|\mu,\sigma)p(\mu,\sigma|y)d\mu\sigma$$

$$\mu^{(s)},\sigma^{(s)} \sim p(\mu,\sigma|y)$$

$$\tilde{y}^{(s)} \sim p(\tilde{y}|\mu^{(s)},\sigma^{(s)})$$





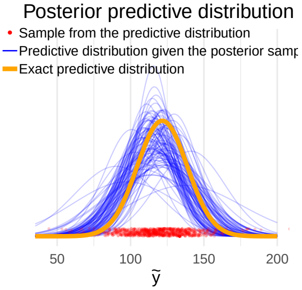


Predictive distribution for new $\tilde{y}$

$$p(\tilde{y}|y) = \int p(\tilde{y}|\mu,\sigma)p(\mu,\sigma|y)d\mu\sigma$$

$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu,\sigma|y)$$

$$\tilde{y}^{(s)} \sim p(\tilde{y}|\mu^{(s)},\sigma^{(s)})$$



Gaussian - posterior predictive distribution

Posterior predictive distribution given known variance

$$p(\tilde{y} \mid \sigma^2, y) = \int p(\tilde{y} \mid \mu, \sigma^2) p(\mu \mid \sigma^2, y) d\mu$$

Gaussian - posterior predictive distribution

Posterior predictive distribution given known variance

$$
\begin{aligned}
p(\tilde{y} \mid \sigma^2, y) &= \int p(\tilde{y} \mid \mu, \sigma^2) p(\mu \mid \sigma^2, y) d\mu \\
&= \int \mathcal{N}(\tilde{y} \mid \mu, \sigma^2) \mathcal{N}(\mu \mid \bar{y}, \sigma^2/n) d\mu
\end{aligned}
$$

UPPSALA UNIVERSITET

Gaussian - posterior predictive distribution

Posterior predictive distribution given known variance

$$\begin{aligned} p(\tilde{y} \mid \sigma^2, y) &= \int p(\tilde{y} \mid \mu, \sigma^2) p(\mu \mid \sigma^2, y) d\mu \\ &= \int \mathcal{N}(\tilde{y} \mid \mu, \sigma^2) \mathcal{N}(\mu \mid \bar{y}, \sigma^2/n) d\mu \\ &= \mathcal{N}(\tilde{y} \mid \bar{y}, (1 + \tfrac{1}{n})\sigma^2) \end{aligned}$$

UPPSALA UNIVERSITET

Gaussian - posterior predictive distribution

Posterior predictive distribution given known variance

$$\begin{aligned} p(\tilde{y} \mid \sigma^2, y) &= \int p(\tilde{y} \mid \mu, \sigma^2) p(\mu \mid \sigma^2, y) d\mu \\ &= \int \mathcal{N}(\tilde{y} \mid \mu, \sigma^2) \mathcal{N}(\mu \mid \bar{y}, \sigma^2/n) d\mu \\ &= \mathcal{N}(\tilde{y} \mid \bar{y}, (1 + \tfrac{1}{n})\sigma^2) \end{aligned}$$

this is up to scaling factor same as $p(\mu \mid \sigma^2, y)$

UPPSALA UNIVERSITET

Gaussian - posterior predictive distribution

Posterior predictive distribution given known variance

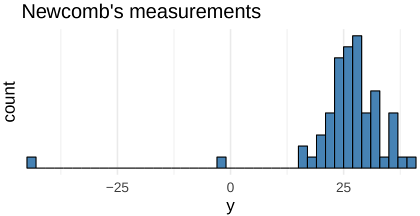
$$\begin{aligned} p(\tilde{y} \mid \sigma^2, y) &= \int p(\tilde{y} \mid \mu, \sigma^2) p(\mu \mid \sigma^2, y) d\mu \\ &= \int \mathcal{N}(\tilde{y} \mid \mu, \sigma^2) \mathcal{N}(\mu \mid \bar{y}, \sigma^2/n) d\mu \\ &= \mathcal{N}(\tilde{y} \mid \bar{y}, (1 + \tfrac{1}{n})\sigma^2) \end{aligned}$$

this is up to scaling factor same as $p(\mu \mid \sigma^2, y)$

$$p(\tilde{y} \mid y) = t_{n-1}(\tilde{y} \mid \bar{y}, (1 + \tfrac{1}{n})s^2)$$

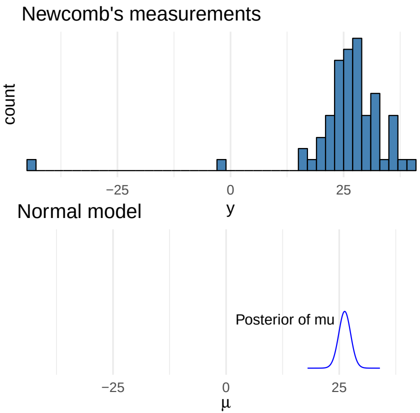
Simon Newcomb's light of speed experiment in 1882

Newcomb measured ($n = 66$) the time required for light to travel from his laboratory on the Potomac River to a mirror at the base of the Washington Monument and back, a total distance of 7422 meters.



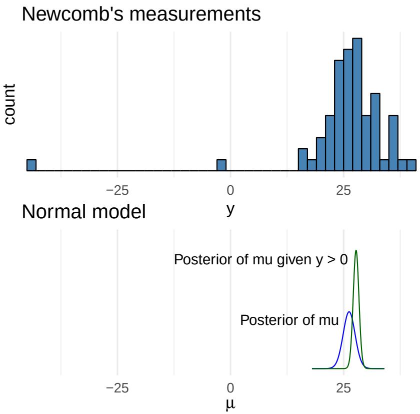
Simon Newcomb's light of speed experiment in 1882

Newcomb measured ($n = 66$) the time required for light to travel from his laboratory on the Potomac River to a mirror at the base of the Washington Monument and back, a total distance of 7422 meters.



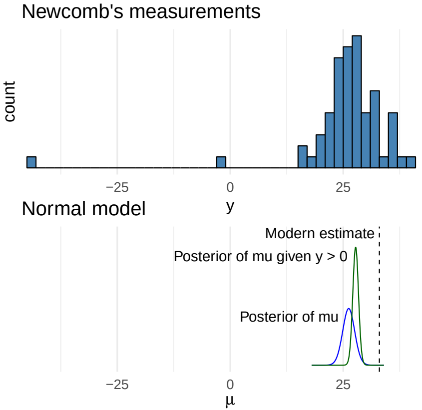
Simon Newcomb's light of speed experiment in 1882

Newcomb measured ($n = 66$) the time required for light to travel from his laboratory on the Potomac River to a mirror at the base of the Washington Monument and back, a total distance of 7422 meters.



Simon Newcomb's light of speed experiment in 1882

Newcomb measured ($n = 66$) the time required for light to travel from his laboratory on the Potomac River to a mirror at the base of the Washington Monument and back, a total distance of 7422 meters.



Gaussian - conjugate prior

- Conjugate prior has to have a form $p(\sigma^2)p(\mu \mid \sigma^2)$

Gaussian - conjugate prior

- Conjugate prior has to have a form $p(\sigma^2)p(\mu \mid \sigma^2)$
- Handy parameterization

$$\mu \mid \sigma^2 \sim \mathrm{N}(\mu_0, \sigma^2/\kappa_0)$$
$$\sigma^2 \sim \text{Inv-}\chi^2(\nu_0, \sigma_0^2)$$

which can be written as

$$p(\mu, \sigma^2) = \text{N-Inv-}\chi^2(\mu_0, \sigma_0^2/\kappa_0; \nu_0, \sigma_0^2)$$

Gaussian - conjugate prior

- Conjugate prior has to have a form $p(\sigma^2)p(\mu \mid \sigma^2)$
- Handy parameterization

$$\begin{aligned} \mu \mid \sigma^2 &\sim \mathrm{N}(\mu_0, \sigma^2/\kappa_0) \\ \sigma^2 &\sim \mathsf{Inv}\text{-}\chi^2(\nu_0, \sigma_0^2) \end{aligned}$$

which can be written as

$$p(\mu, \sigma^2) = \mathsf{N}\text{-}\mathsf{Inv}\text{-}\chi^2(\mu_0, \sigma_0^2/\kappa_0; \nu_0, \sigma_0^2)$$

- μ and σ^2 are a priori dependent
  - if σ^2 is large, then μ has wide prior

UPPSALA UNIVERSITET

Gaussian - conjugate prior

Joint posterior (exercise 3.9)

$$p(\mu, \sigma^2 \mid y) = \text{N-Inv-}\chi^2(\mu_n, \sigma_n^2/\kappa_n; \nu_n, \sigma_n^2)$$

where

$$\begin{aligned}
\mu_n &= \frac{\kappa_0}{\kappa_0 + n}\mu_0 + \frac{n}{\kappa_0 + n}\bar{y} \\
\kappa_n &= \kappa_0 + n \\
\nu_n &= \nu_0 + n \\
\nu_n\sigma_n^2 &= \nu_0\sigma_0^2 + (n-1)s^2 + \frac{\kappa_0 n}{\kappa_0 + n}(\bar{y} - \mu_0)^2
\end{aligned}$$

Gaussian - conjugate prior

- Conditional $p(\mu \mid \sigma^2, y)$

$$\begin{aligned}\mu \mid \sigma^2, y &\sim \mathrm{N}(\mu_n, \sigma^2/\kappa_n) \\ &= \mathrm{N}\left(\frac{\frac{\kappa_0}{\sigma^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{\kappa_0}{\sigma^2} + \frac{n}{\sigma^2}}, \frac{1}{\frac{\kappa_0}{\sigma^2} + \frac{n}{\sigma^2}}\right)\end{aligned}$$

Gaussian - conjugate prior

- Conditional $p(\mu \mid \sigma^2, y)$

$$
\begin{aligned}
\mu \mid \sigma^2, y &\sim \mathrm{N}(\mu_n, \sigma^2/\kappa_n) \\
&= \mathrm{N}\left(\frac{\frac{\kappa_0}{\sigma^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{\kappa_0}{\sigma^2} + \frac{n}{\sigma^2}}, \frac{1}{\frac{\kappa_0}{\sigma^2} + \frac{n}{\sigma^2}} \right)
\end{aligned}
$$

- Marginal $p(\sigma^2 \mid y)$

$$
\sigma^2 \mid y \sim \text{Inv-}\chi^2(\nu_n, \sigma_n^2)
$$

Gaussian - conjugate prior

- Conditional $p(\mu \mid \sigma^2, y)$

$$\begin{aligned}\mu \mid \sigma^2, y &\sim \mathrm{N}(\mu_n, \sigma^2/\kappa_n) \\ &= \mathrm{N}\left(\frac{\frac{\kappa_0}{\sigma^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{\kappa_0}{\sigma^2} + \frac{n}{\sigma^2}}, \frac{1}{\frac{\kappa_0}{\sigma^2} + \frac{n}{\sigma^2}}\right)\end{aligned}$$

- Marginal $p(\sigma^2 \mid y)$

$$\sigma^2 \mid y \sim \text{Inv-}\chi^2(\nu_n, \sigma_n^2)$$

- Marginal $p(\mu \mid y)$

$$\mu \mid y \sim t_{\nu_n}(\mu \mid \mu_n, \sigma_n^2/\kappa_n)$$

Multinomial model for categorical data

- Extension of binomial to K categories
- Observation model (Categorical distribution, $n = 1$)
  $y_i = (0, 1, 0, 0, 0)$ - what is K here?

Multinomial model for categorical data

- Extension of binomial to K categories
- Observation model (Categorical distribution, $n = 1$)
  $y_i = (0, 1, 0, 0, 0)$ - what is K here?

$$p(y \mid \theta) \propto \prod_{k=1}^{K} \theta_j^{y_j},$$

$$\text{where } \sum_{k}^{K} \theta_k = 1\,, \text{ and } \forall \theta_k > 0$$

- What is important when choosing the prior for θ?

UPPSALA UNIVERSITET

Multinomial model for categorical data

- Extension of binomial to K categories
- Observation model (Categorical distribution, $n = 1$)
  $y_i = (0, 1, 0, 0, 0)$ - what is K here?

$$p(y \mid \theta) \propto \prod_{k=1}^{K} \theta_j^{y_j},$$

$$\text{where } \sum_{k}^{K} \theta_k = 1\,, \text{ and } \forall \theta_k > 0$$

- What is important when choosing the prior for θ?
- Conjugate prior: The Dirichlet distribution

$$p(\theta) \propto \prod_{k=1}^{K} \theta_k^{\alpha_k - 1},$$

$$\text{where } \forall \alpha_k > 0$$

Multinomial model for categorical data: The posterior

- The posterior $p(\theta|y)$

$$p(\theta \mid y) \propto p(y \mid \theta)p(\theta)$$

Multinomial model for categorical data: The posterior

- The posterior $p(\theta|y)$

$$
\begin{aligned}
p(\theta \mid y) &\propto p(y \mid \theta)p(\theta) \\
&\propto \prod_{k=1}^{K} \theta_k^{\alpha_k - 1} \prod_{i}^{n} \prod_{k=1}^{K} \theta_k^{y_{k,i}}
\end{aligned}
$$

Multinomial model for categorical data: The posterior

UPPSALA UNIVERSITET

- The posterior $p(\theta|y)$

$$
\begin{aligned}
p(\theta \mid y) &\propto p(y \mid \theta)p(\theta) \\
&\propto \prod_{k=1}^{K} \theta_k^{\alpha_k - 1} \prod_{i}^{n} \prod_{k=1}^{K} \theta_k^{y_{k,i}} \\
&= \prod_{k=1}^{K} \theta_k^{\alpha_k - 1} \prod_{k=1}^{K} \theta_k^{\sum_i^n y_{k,i}}
\end{aligned}
$$

Multinomial model for categorical data: The posterior

- The posterior $p(\theta|y)$

$$
\begin{aligned}
p(\theta \mid y) &\propto p(y \mid \theta) p(\theta) \\
&\propto \prod_{k=1}^{K} \theta_k^{\alpha_k - 1} \prod_{i}^{n} \prod_{k=1}^{K} \theta_k^{y_{k,i}} \\
&= \prod_{k=1}^{K} \theta_k^{\alpha_k - 1} \prod_{k=1}^{K} \theta_k^{\sum_i^n y_{k,i}} \\
&= \prod_{k=1}^{K} \theta_k^{\alpha_k - 1 + \sum_i^n y_{k,i}}
\end{aligned}
$$

UPPSALA UNIVERSITET

Multinomial model for categorical data: The posterior

- The posterior $p(\theta|y)$

$$
\begin{aligned}
p(\theta \mid y) &\propto p(y \mid \theta) p(\theta) \\
&\propto \prod_{k=1}^{K} \theta_k^{\alpha_k - 1} \prod_{i}^{n} \prod_{k=1}^{K} \theta_k^{y_{k,i}} \\
&= \prod_{k=1}^{K} \theta_k^{\alpha_k - 1} \prod_{k=1}^{K} \theta_k^{\sum_i^n y_{k,i}} \\
&= \prod_{k=1}^{K} \theta_k^{\alpha_k - 1 + \sum_i^n y_{k,i}}
\end{aligned}
$$

- The posterior is $p(\theta|y) = \mathrm{Dir}(\alpha + \sum y)$

Multivariate Gaussian

- Observation model

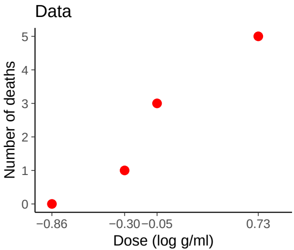
$$p(y \mid \mu, \Sigma) \propto \mid \Sigma \mid^{-1/2} \exp\left(-\frac{1}{2}(y-\mu)^T \Sigma^{-1}(y-\mu)\right),$$

where $y \in \mathcal{R}^D$

- See BDA3 p. 72–
- New recommended LKJ-prior mentioned in Appendix A, see more in Stan manual

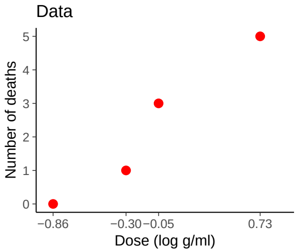
Bioassay

Dose, x_i (log g/ml)	Number of animals, n_i	Number of deaths, y_i
-0.86	5	0
-0.30	5	1
-0.05	5	3
0.73	5	5



Bioassay

Dose, x_i (log g/ml)	Number of animals, n_i	Number of deaths, y_i
-0.86	5	0
-0.30	5	1
-0.05	5	3
0.73	5	5

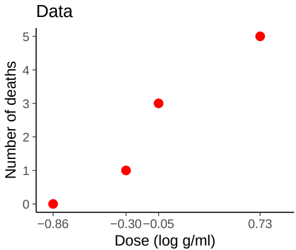


Find out lethal dose 50% (LD50)

- used to classify how hazardous chemical is
- 1984 EEC directive has 4 levels

Bioassay

Dose, x_i (log g/ml)	Number of animals, n_i	Number of deaths, y_i
-0.86	5	0
-0.30	5	1
-0.05	5	3
0.73	5	5



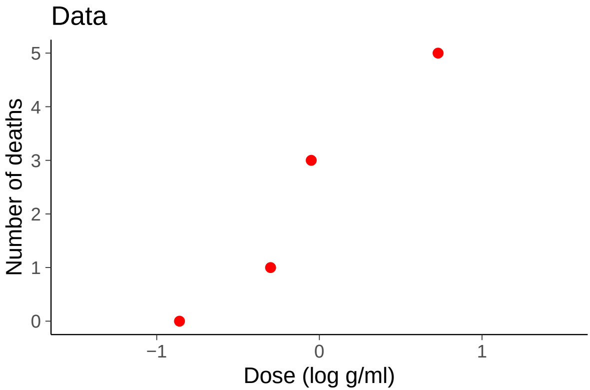
Find out lethal dose 50% (LD50)

- used to classify how hazardous chemical is
- 1984 EEC directive has 4 levels

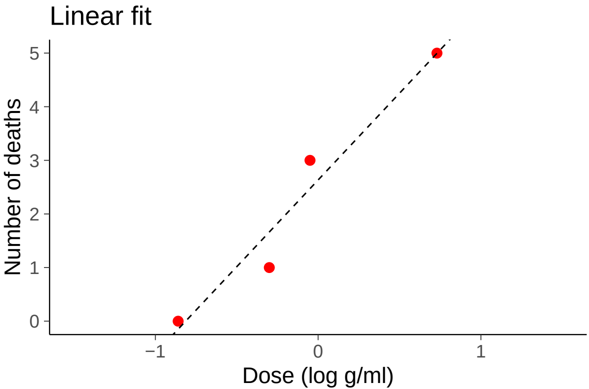
Bayesian methods help to

- reduce the number of animals needed
- easy to make sequential experiment and stop as soon as desired accuracy is obtained

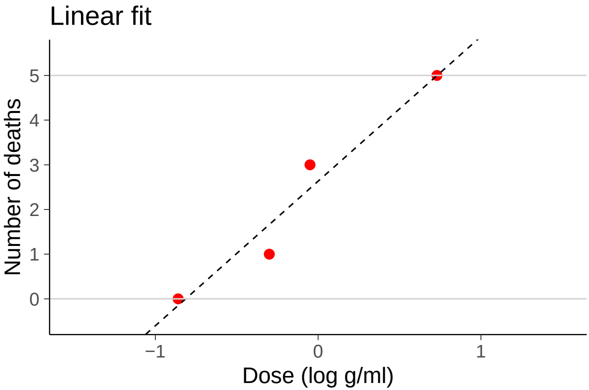
Bioassay



Bioassay

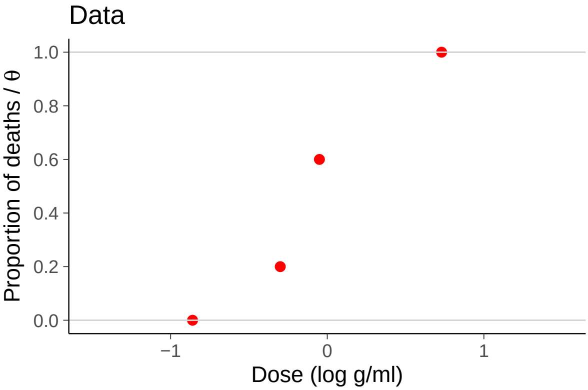


Bioassay



UPPSALA UNIVERSITET

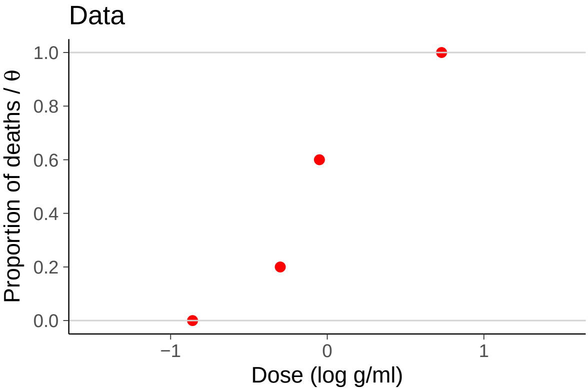
Bioassay



UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

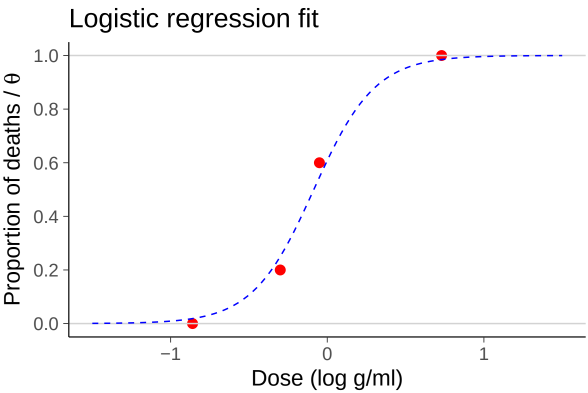
Bioassay



Binomial model

$$y_i \mid \theta_i \sim \text{Bin}(\theta_i, n_i)$$

Bioassay



Binomial model

$$y_i \mid \theta_i \sim \text{Bin}(\theta_i, n_i), \quad \text{logit}(\theta_i) = \log\left(\frac{\theta_i}{1-\theta_i}\right) = \alpha + \beta x_i$$

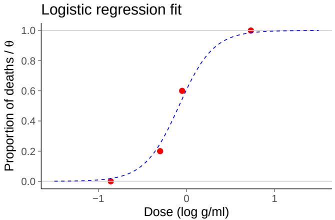
UPPSALA UNIVERSITET

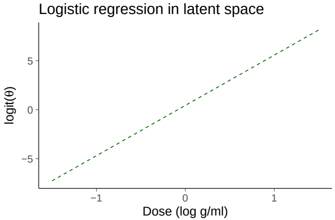
Bioassay

$$y_i \mid \theta_i \sim \text{Bin}(\theta_i, n_i)$$

$$\text{logit}(\theta_i) = \log\left(\frac{\theta_i}{1-\theta_i}\right)$$

$$= \alpha + \beta x_i$$



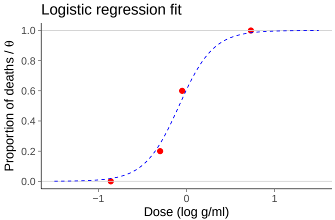


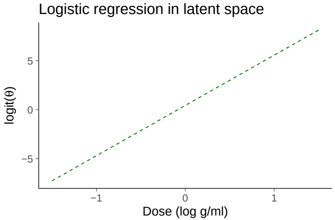
Bioassay

$$y_i \mid \theta_i \sim \mathrm{Bin}(\theta_i, n_i)$$

$$\mathrm{logit}(\theta_i) = \log\left(\frac{\theta_i}{1-\theta_i}\right) = \alpha + \beta x_i$$

$$\theta_i = \frac{1}{1+\exp(-(\alpha+\beta x_i))}$$

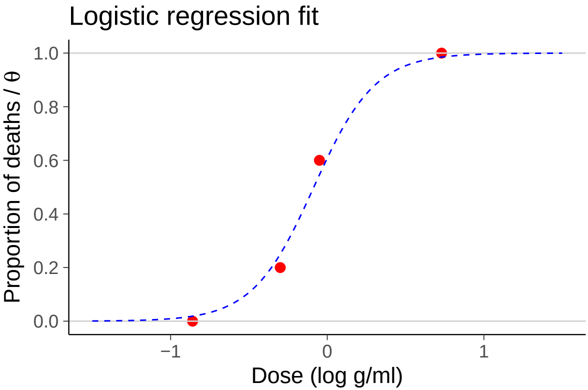




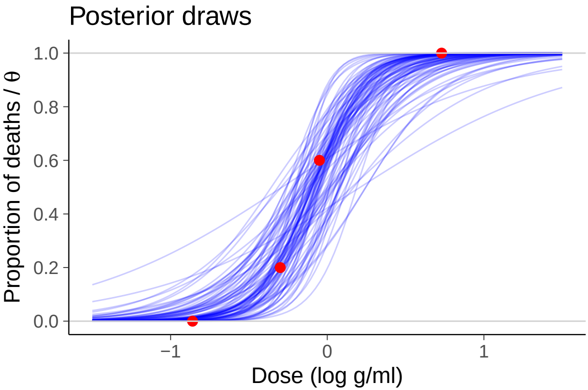
UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

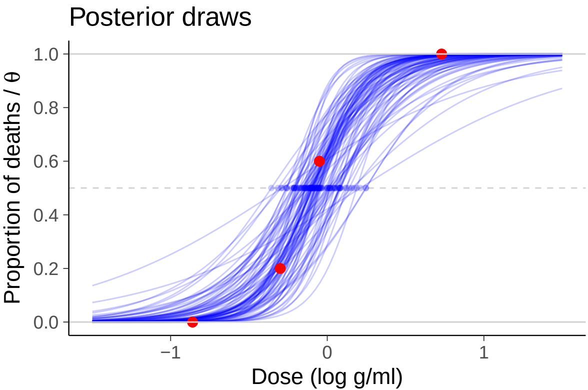
Bioassay: Lethal Dose 50%



Bioassay: Lethal Dose 50%

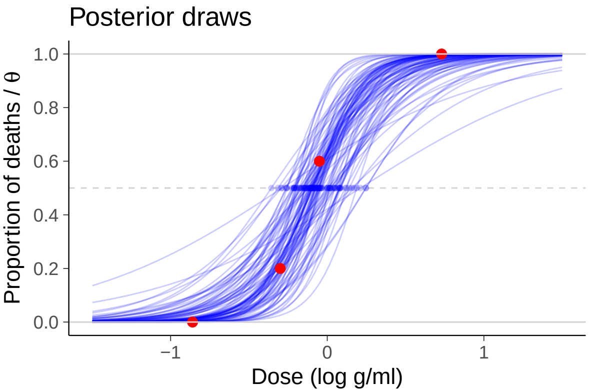


Bioassay: Lethal Dose 50%



$$\mathrm{E}\left(\frac{y}{n}\right) = \text{logit}^{-1}(\alpha + \beta x) = 0.5$$

Bioassay: Lethal Dose 50%

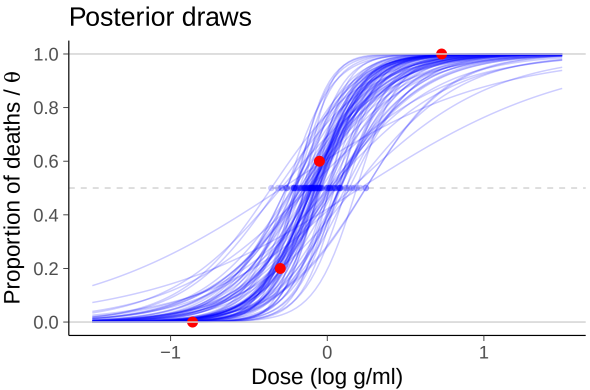


$$\mathrm{E}\left(\frac{y}{n}\right) = \text{logit}^{-1}(\alpha + \beta x) = 0.5 \quad \Rightarrow \quad x_{\mathrm{LD50}} = -\alpha/\beta$$

UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- **Bioassay example**

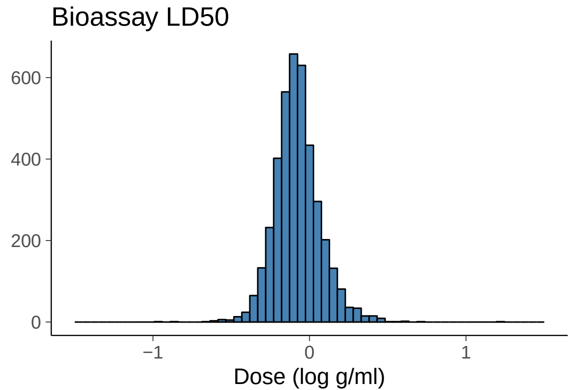
Bioassay: Lethal Dose 50%



$$\mathrm{E}\left(\frac{y}{n}\right) = \mathrm{logit}^{-1}(\alpha + \beta x) = 0.5 \quad \Rightarrow \quad x_{\mathrm{LD50}} = -\alpha/\beta$$

$$x^{(s)}_{\mathrm{LD50}} = -\alpha^{(s)}/\beta^{(s)}$$

Bioassay: Lethal Dose 50%



$$\mathrm{E}\left(\frac{y}{n}\right) = \text{logit}^{-1}(\alpha + \beta x) = 0.5 \quad \Rightarrow \quad x_{\mathrm{LD50}} = -\alpha/\beta$$

$$x^{(s)}_{\mathrm{LD50}} = -\alpha^{(s)}/\beta^{(s)}$$

Bioassay posterior

Binomial model

$$y_i \mid \theta_i \sim \mathrm{Bin}(\theta_i, n_i)$$

Link function

$$\mathrm{logit}(\theta_i) = \alpha + \beta x_i$$

Bioassay posterior

Binomial model

$$y_i \mid \theta_i \sim \text{Bin}(\theta_i, n_i)$$

Link function

$$\text{logit}(\theta_i) = \alpha + \beta x_i$$

Likelihood

$$p(y_i \mid \alpha, \beta, n_i, x_i) \propto \theta_i^{y_i}[1-\theta_i]^{n_i - y_i}$$

UPPSALA UNIVERSITET

- Introduction
- Multiple parameter models
  - Marginalization
  - Gaussian
  - Example
  - Gaussian - conjugate prior
  - Multinomial model
  - Multivariate Gaussian
- Bioassay example

Bioassay posterior

Binomial model

$$y_i \mid \theta_i \sim \text{Bin}(\theta_i, n_i)$$

Link function

$$\text{logit}(\theta_i) = \alpha + \beta x_i$$

Likelihood

$$\begin{aligned} p(y_i \mid \alpha, \beta, n_i, x_i) &\propto \theta_i^{y_i}[1-\theta_i]^{n_i-y_i} \\ &\propto [\text{logit}^{-1}(\alpha+\beta x_i)]^{y_i}[1-\text{logit}^{-1}(\alpha+\beta x_i)]^{n_i-y_i} \end{aligned}$$

Bioassay posterior

Binomial model

$$y_i \mid \theta_i \sim \text{Bin}(\theta_i, n_i)$$

Link function

$$\text{logit}(\theta_i) = \alpha + \beta x_i$$

Likelihood

$$\begin{aligned} p(y_i \mid \alpha, \beta, n_i, x_i) &\propto \theta_i^{y_i}[1-\theta_i]^{n_i - y_i} \\ &\propto [\mathrm{logit}^{-1}(\alpha + \beta x_i)]^{y_i}[1 - \mathrm{logit}^{-1}(\alpha + \beta x_i)]^{n_i - y_i} \end{aligned}$$

Posterior

$$p(\alpha, \beta \mid y, n, x) \propto p(\alpha, \beta) \prod_{i=1}^{n} p(y_i \mid \alpha, \beta, n_i, x_i)$$

No analytic posterior distribution? What can we do?

UPPSALA UNIVERSITET

Grid evaluation

1. Setup an area (can be hard) for α and β that capture most mass (here $\alpha = [-1, 5]$ and $\beta = [0, 30]$)

Grid evaluation

1. Setup an area (can be hard) for α and β that capture most mass (here $\alpha = [-1, 5]$ and $\beta = [0, 30]$)
2. Compute unnormalized $p(\alpha^{(g)}, \beta^{(g)} \mid y, n, x)$, here $\tilde{p}$, at the grid points g

Grid evaluation

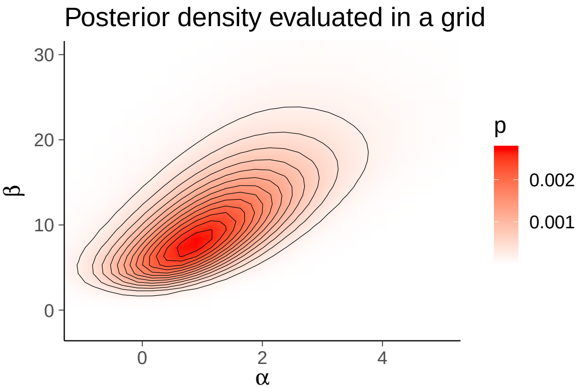
1. Setup an area (can be hard) for α and β that capture most mass (here $\alpha = [-1, 5]$ and $\beta = [0, 30]$)
2. Compute unnormalized $p(\alpha^{(g)}, \beta^{(g)} \mid y, n, x)$, here $\tilde{p}$, at the grid points g
3. Sum up $\tilde{p}$ over the whole grid (for all $g \in \{1, ..., G\}$)

Grid evaluation

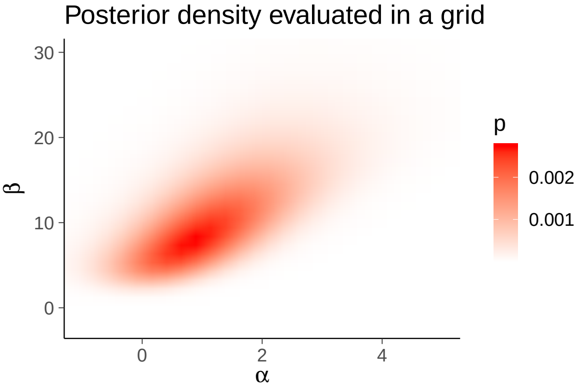
1. Setup an area (can be hard) for α and β that capture most mass (here $\alpha = [-1, 5]$ and $\beta = [0, 30]$)
2. Compute unnormalized $p(\alpha^{(g)}, \beta^{(g)} \mid y, n, x)$, here $\tilde{p}$, at the grid points g
3. Sum up $\tilde{p}$ over the whole grid (for all $g \in \{1, ..., G\}$)
4. Compute (normalize) the pmf approximation of the posterior $\hat{p}$

g	$(\alpha,\ \beta)$	$\tilde{p}$	$\hat{p}$
1	(0, -1)	0.02	0.0002
2	(0, -0.8)	0.03	0.0003
...	...	...	...
G	(30, 5)	0.001	0.00001
$\sum_g^G$	-	100	1

Bioassay (with uniform prior on α, β)

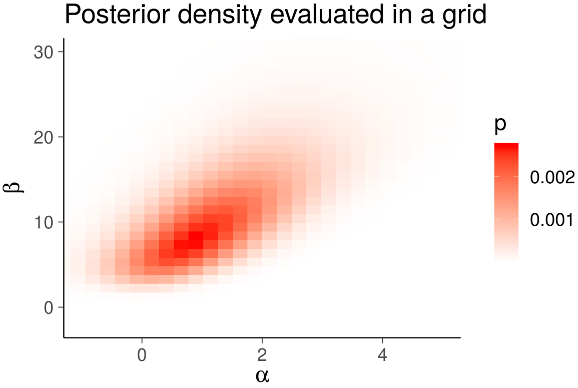


Bioassay (with uniform prior on α, β)



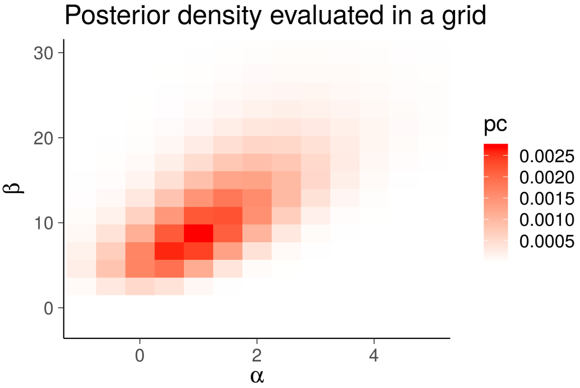
Density evaluated in grid, but plotted using interpolation

Bioassay (with uniform prior on α, β)



Density evaluated in grid, and plotted without interpolation

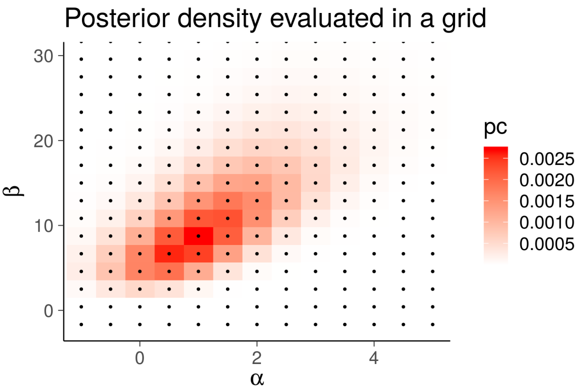
Bioassay (with uniform prior on α, β)



Density evaluated in a coarser grid

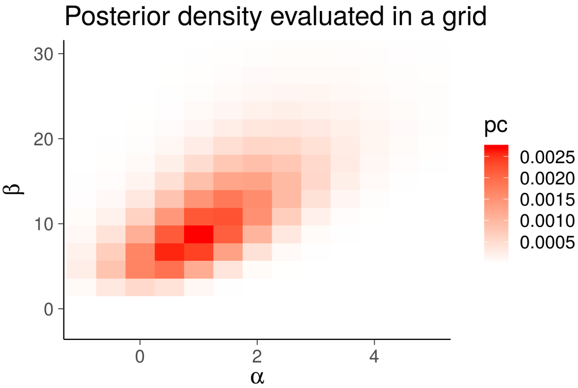
UPPSALA UNIVERSITET

Bioassay (with uniform prior on α, β)



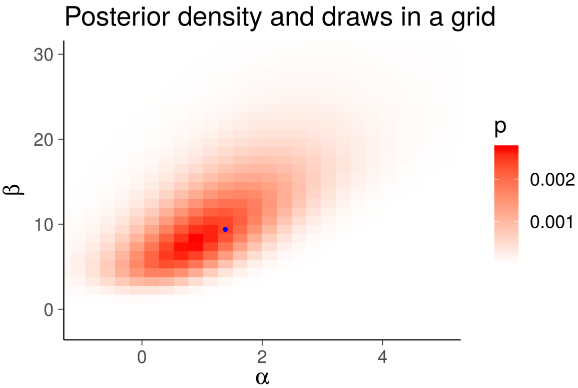
- Approximate the density as piecewise constant function
- Evaluate density in a grid over some finite region
- Density times cell area gives probability mass in each cell

Bioassay (with uniform prior on α, β)



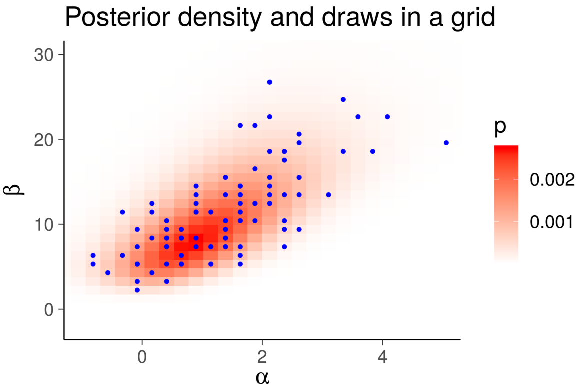
- Densities at 1, 2, and 3: 0.0027 0.0010 0.0001
- Probabilities of cells 1, 2, and 3: 0.0431 0.0166 0.0010
- Probabilities of cells sum to 1

Bioassay (with uniform prior on α, β)



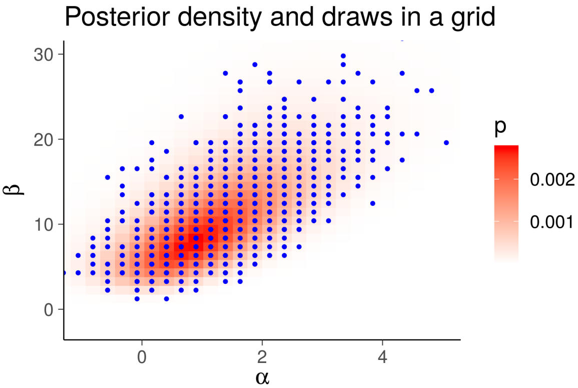
- Sample according to grid cell probabilities

Bioassay (with uniform prior on α, β)



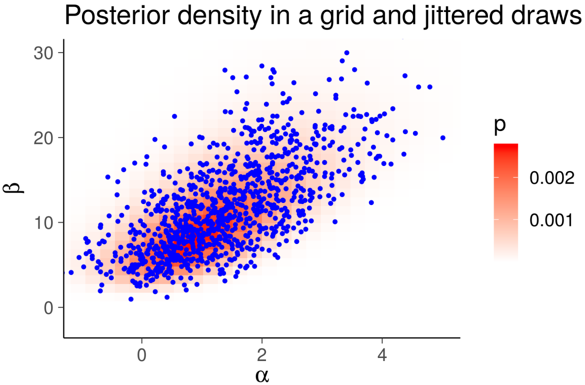
- Sample according to grid cell probabilities

Bioassay (with uniform prior on α, β)



- Sample according to grid cell probabilities
- Several draws can be from the same grid cell

Bioassay (with uniform prior on α, β)



- Jitter can be added to improve visualization

UPPSALA UNIVERSITET

Grid sampling

- Draws can be used to estimate expectations, for example

$$E[x_{\mathrm{LD50}}] = E[-\alpha/\beta] \approx \frac{1}{S}\sum_{s=1}^{S}\frac{\alpha^{(s)}}{\beta^{(s)}}$$

Grid sampling

- Draws can be used to estimate expectations, for example

$$E[x_{\mathrm{LD50}}] = E[-\alpha/\beta] \approx \frac{1}{S}\sum_{s=1}^{S}\frac{\alpha^{(s)}}{\beta^{(s)}}$$

- Instead of sampling, grid could be used to evaluate functions directly, for example

$$\mathrm{E}[-\alpha/\beta] \approx \sum_{t=1}^{T} w_{\mathrm{cell}}^{(t)}\frac{\alpha^{(t)}}{\beta^{(t)}},$$

where $w_{\mathrm{cell}}^{(t)}$ is the normalized probability of a grid cell t, and $\alpha^{(t)}$ and $\beta^{(t)}$ are center locations of grid cells

UPPSALA UNIVERSITET

Grid sampling

- Draws can be used to estimate expectations, for example

$$E[x_{\mathrm{LD50}}] = E[-\alpha/\beta] \approx \frac{1}{S}\sum_{s=1}^{S}\frac{\alpha^{(s)}}{\beta^{(s)}}$$

- Instead of sampling, grid could be used to evaluate functions directly, for example

$$\mathrm{E}[-\alpha/\beta] \approx \sum_{t=1}^{T} w_{\mathrm{cell}}^{(t)}\frac{\alpha^{(t)}}{\beta^{(t)}},$$

where $w_{\mathrm{cell}}^{(t)}$ is the normalized probability of a grid cell t, and $\alpha^{(t)}$ and $\beta^{(t)}$ are center locations of grid cells

- Grid sampling gets computationally too expensive in high dimensions