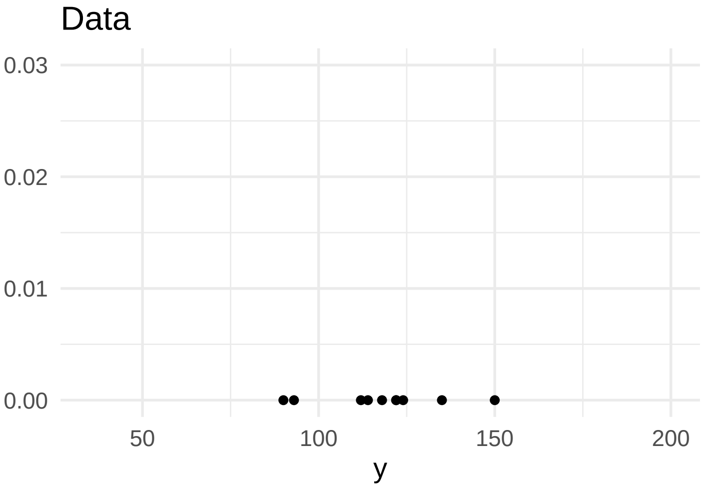
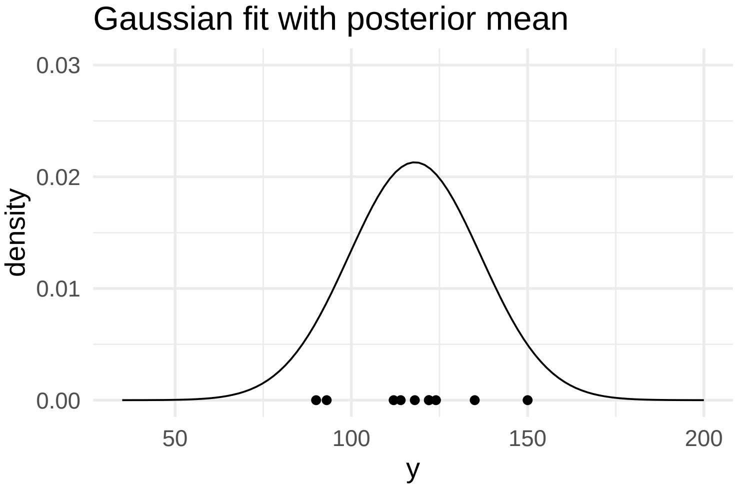
Agenda

- Gaussian Example
- Newcomb' Experiment
- Other Multiparameter Models
- Bioassay Example

Gaussian example

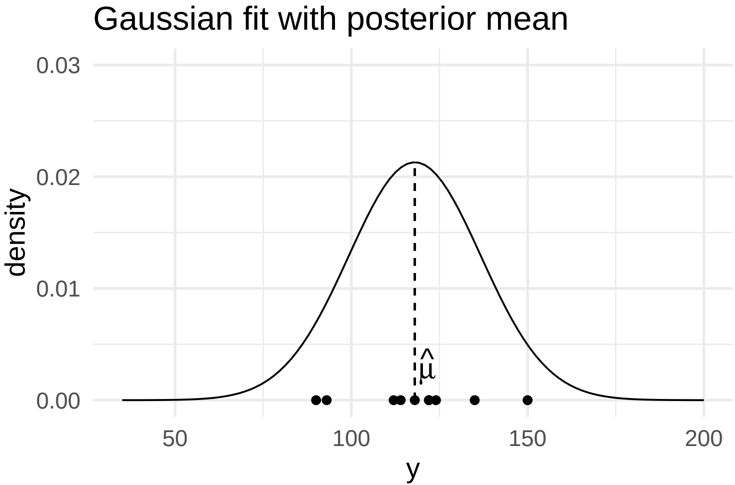


Gaussian example



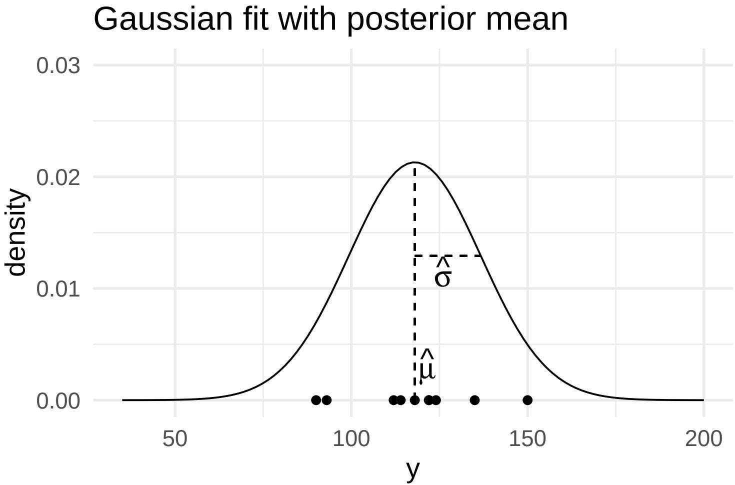
$$p({\color{red}y} \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}({\color{red}y} - \mu)^2\right)$$

Gaussian example



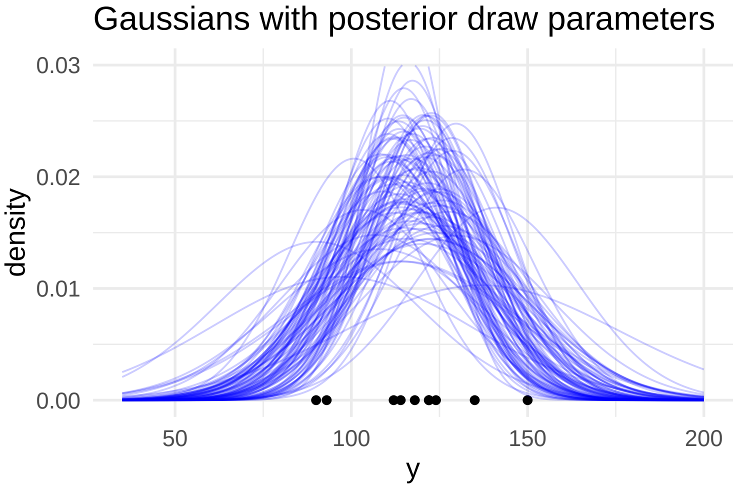
$$p(y \mid {\color{red}\mu}, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y - {\color{red}\mu})^2\right)$$

Gaussian example



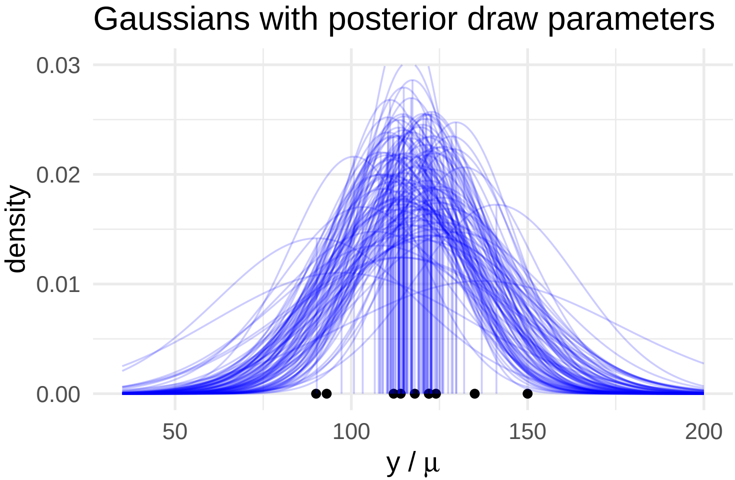
$$p(y \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right)$$

Gaussian example



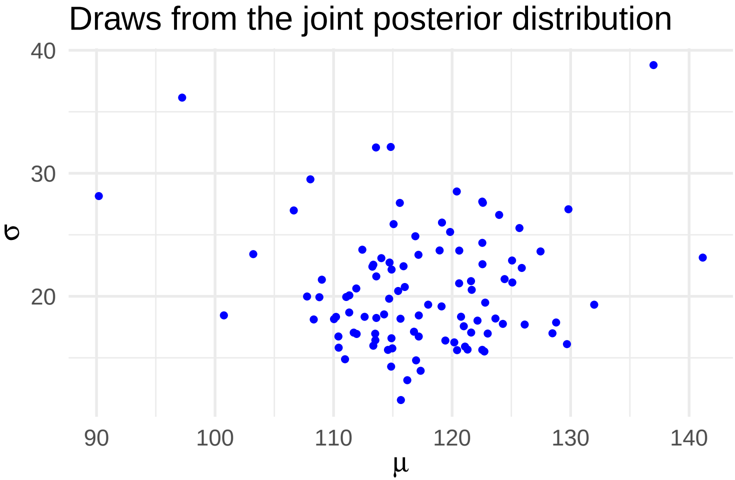
$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

Gaussian example



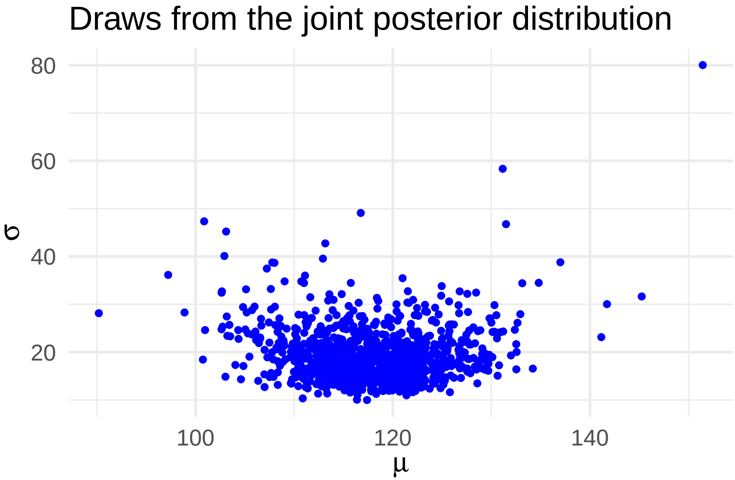
$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

Gaussian example



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

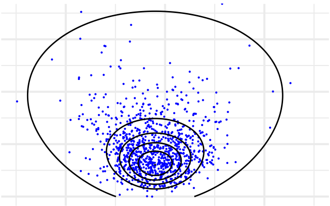
Gaussian example



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

Joint posterior

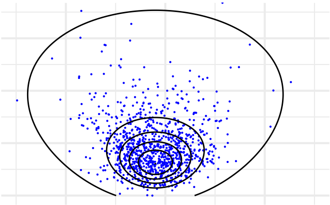
$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$



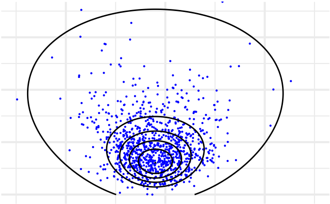
Joint posterior

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

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



Joint posterior

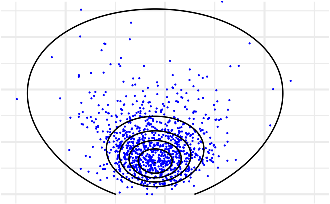


$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

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

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

Joint posterior

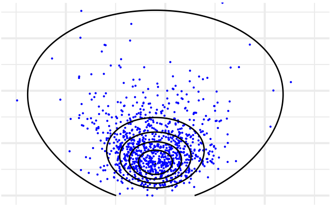


$$\color{blue}{\mu^{(s)}, \sigma^{(s)}} \sim p(\mu, \sigma \mid y)$$

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



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

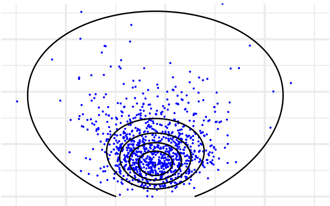
with $p(\mu, \sigma^2) \propto \sigma^{-2}$

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

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

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

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

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

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

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

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

Gaussian - non-informative prior (extended derivation)

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

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

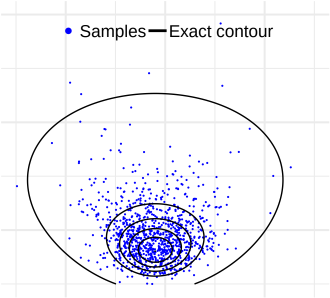
$$\sum_{i=1}^{n}(y_i^2 - 2y_i\mu + \mu^2 - \bar{y}^2 + \bar{y}^2 - 2y_i\bar{y} + 2y_i\bar{y})$$

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

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

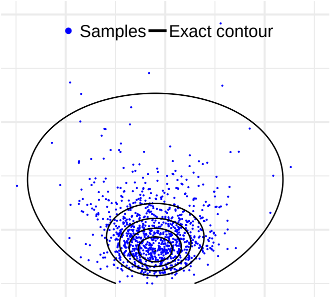
$$\sum_{i=1}^{n}(y_i - \bar{y})^2 + n(\bar{y} - \mu)^2$$

Joint posterior

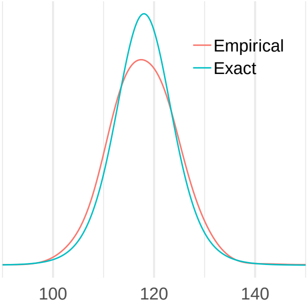


$$\color{blue}{\mu^{(s)}, \sigma^{(s)}} \sim p(\mu, \sigma \mid y)$$

Joint posterior



Marginal of mu

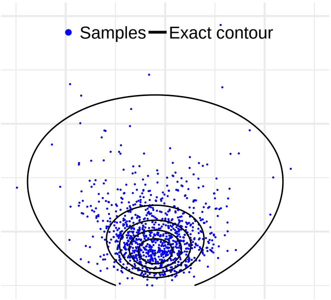


$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

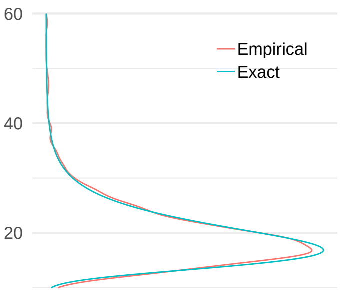
marginals

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

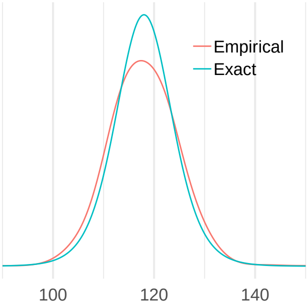
Joint posterior



Marginal of sigma



Marginal of mu



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

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

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

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

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

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

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

$$
\int \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\theta)^2\right) d\theta = 1
$$

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

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

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

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

$$\int \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(y-\theta)^2\right) d\theta = 1$$

$$
\begin{aligned}
&\propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2}(n-1)s^2\right) \sqrt{2\pi\sigma^2/n} \\
&\propto (\sigma^2)^{-(n+1)/2} \exp\left(-\frac{(n-1)s^2}{2\sigma^2}\right)
\end{aligned}
$$

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

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

Gaussian - non-informative prior

Known mean

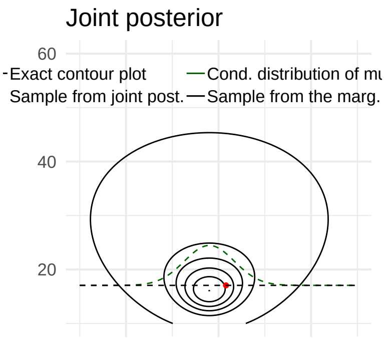
$$\sigma^2 \mid y \sim \text{Inv-}\chi^2(n, v)$$

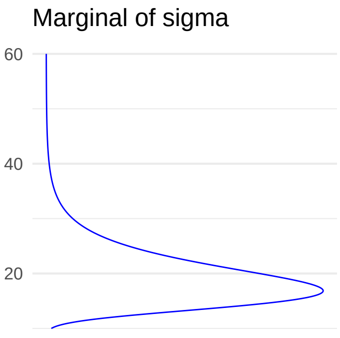
$$\text{where} \quad v = \frac{1}{n}\sum_{i=1}^{n}(y_i - \theta)^2$$

Unknown mean

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

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

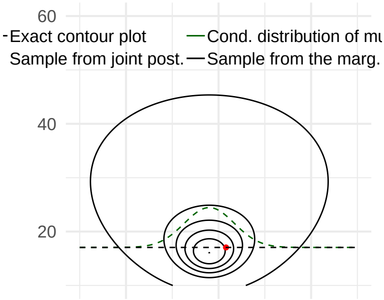




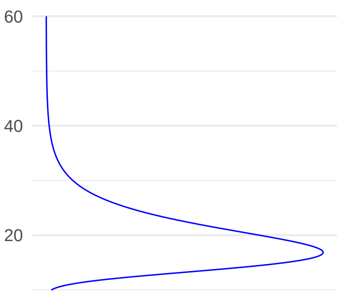
Factorization

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

Joint posterior



Marginal of sigma

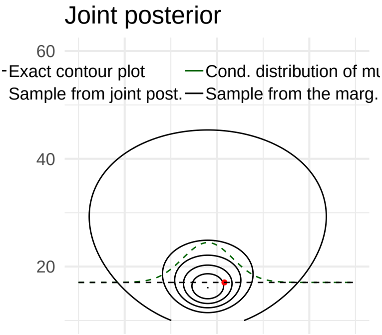


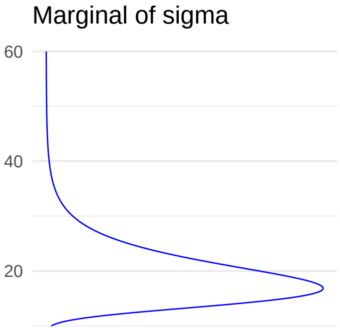
Factorization

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

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

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





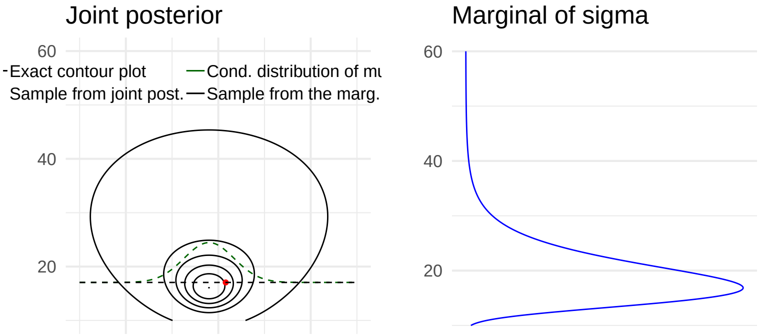
Factorization

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

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

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

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



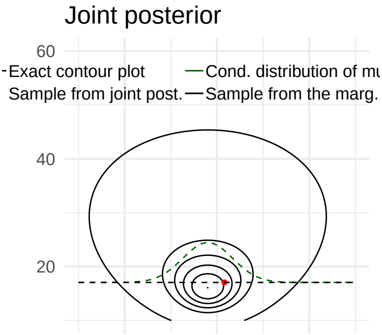
Factorization

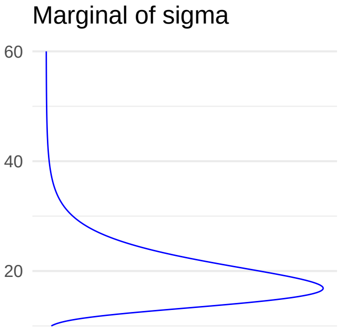
$$p(\mu, \sigma^2 \mid y) = p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y)$$

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

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

$$p(\mu \mid \sigma^2, y) = \mathrm{N}(\mu \mid \bar{y}, \sigma^2/n) \propto \exp\left(-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2\right)$$





Factorization

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

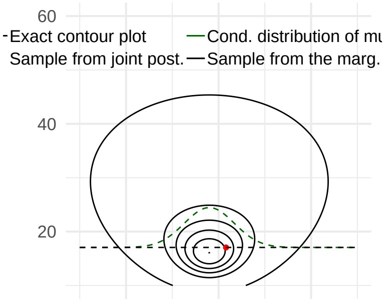
$$p(\sigma^2 \mid y) = \text{Inv-}\chi^2(\sigma^2 \mid n-1, s^2)$$

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

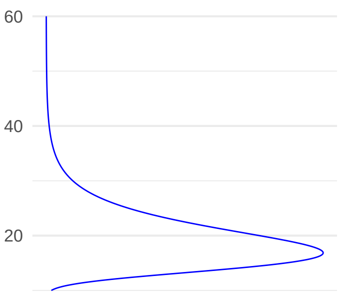
$$p(\mu \mid \sigma^2, y) = \mathrm{N}(\mu \mid \bar{y}, \sigma^2/n)$$

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

Joint posterior



Marginal of sigma



Factorization

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

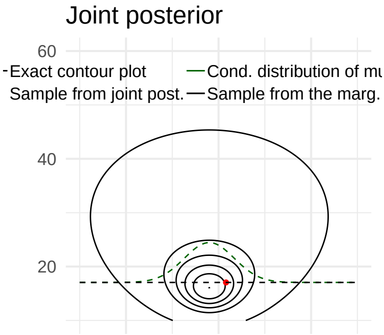
$$p(\sigma^2 \mid y) = \text{Inv-}\chi^2(\sigma^2 \mid n-1, s^2)$$

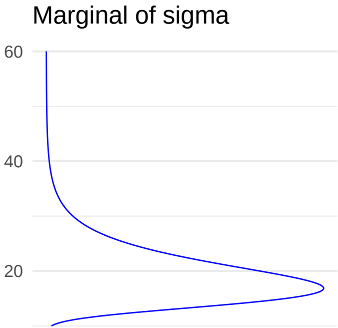
$$(\sigma^2)^{(s)} \sim p(\sigma^2 \mid y)$$

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

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

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

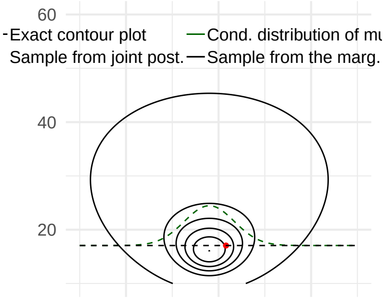




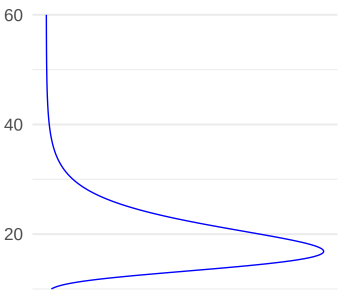
Factorization

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

Joint posterior



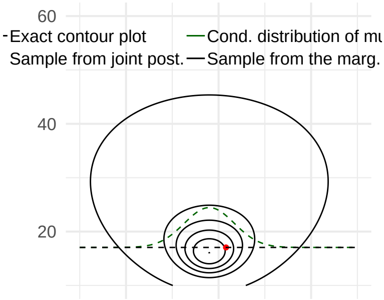
Marginal of sigma



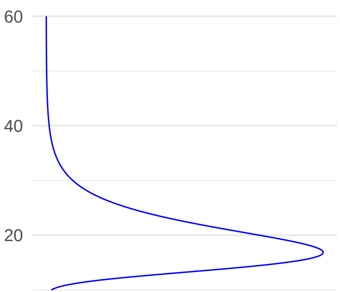
Factorization

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

Joint posterior



Marginal of sigma



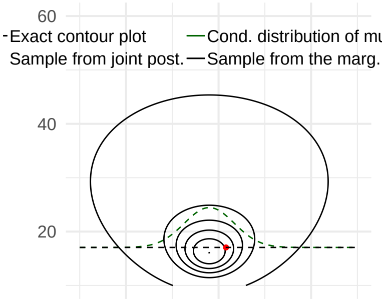
Factorization

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

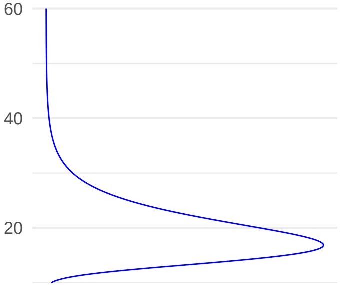
$$(\sigma^2)^{(s)} \sim p(\sigma^2 \mid y)$$

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

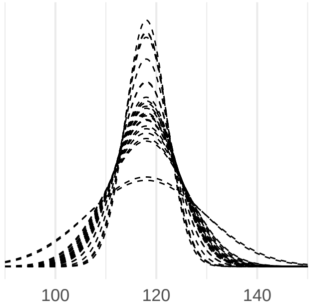
Joint posterior



Marginal of sigma



Cond distr of mu for 25 draws



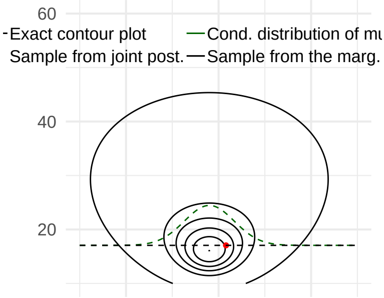
Factorization

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

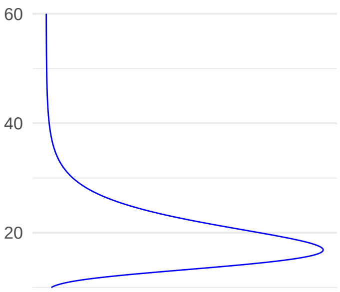
$$(\sigma^2)^{(s)} \sim {\color{blue} p(\sigma^2 \mid y)}$$

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

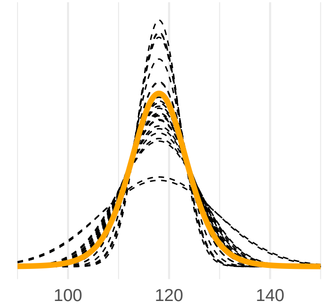
Joint posterior



Marginal of sigma



Cond distr of mu for 25 draws



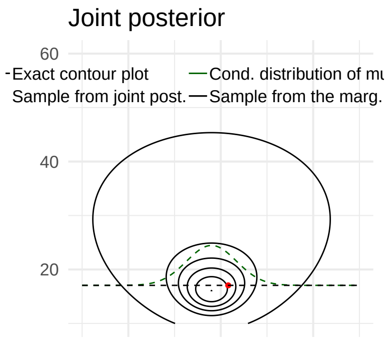
Factorization

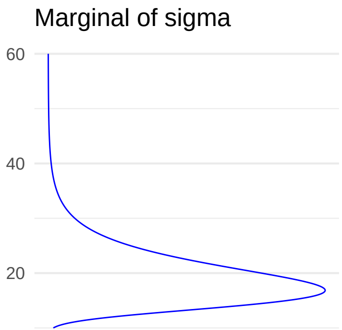
$$p(\mu, \sigma^2 \mid y) = p(\mu \mid \sigma^2, y) p(\sigma^2 \mid y)$$

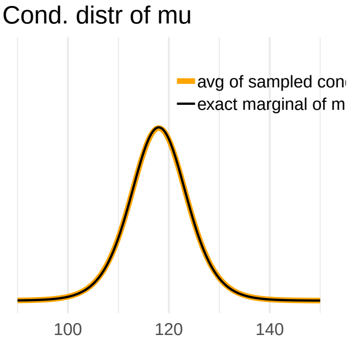
$$(\sigma^2)^{(s)} \sim p(\sigma^2 \mid y)$$

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

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







Factorization

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

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

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

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

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

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

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

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

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

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

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

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

Transformation

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

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

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

Transformation

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

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

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

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

Transformation

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

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

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

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

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

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

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

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

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

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

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

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

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

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

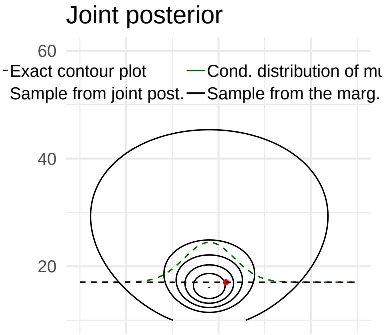
$$
A = (n-1)s^2 + n(\mu - \bar{y})^2 \quad \text{and} \quad z = \frac{A}{2\sigma^2}
$$

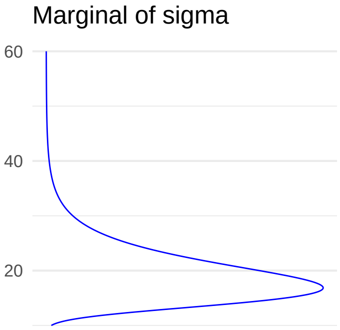
$$
p(\mu \mid y) \propto A^{-n/2} \int_0^\infty z^{(n-2)/2} \exp(-z) dz
$$

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

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

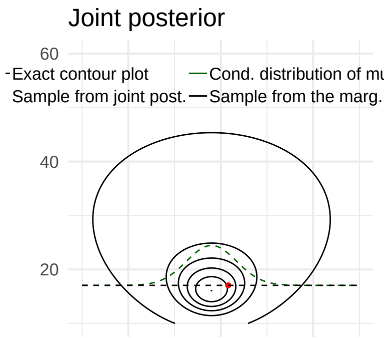
$$
p(\mu \mid y) = t_{n-1}(\mu \mid \bar{y}, s^2/n) \quad \text{Student's } t
$$

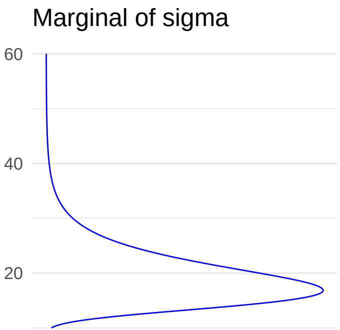




Predictive distribution for new $\tilde{y}$

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



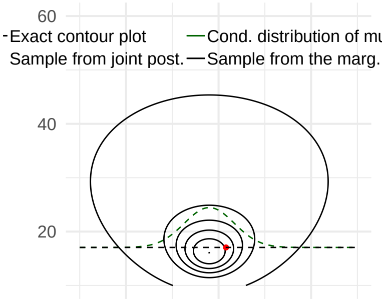


Predictive distribution for new $\tilde{y}$

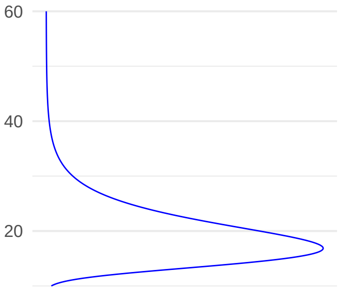
$$p(\tilde{y} \mid y) = \int p(\tilde{y} \mid \mu, \sigma) p(\mu, \sigma \mid y) d\mu\sigma$$

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

Joint posterior



Marginal of sigma



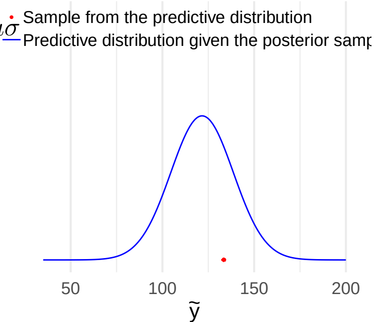
Predictive distribution for new $\tilde{y}$

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

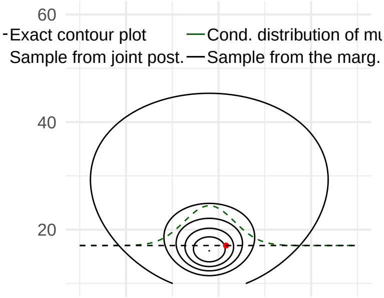
$$\color{red}{\mu^{(s)}, \sigma^{(s)}} \sim p(\mu, \sigma \mid y)$$

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

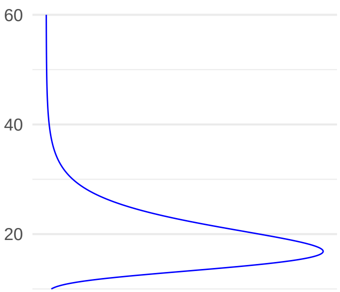
Posterior predictive distribution



Joint posterior



Marginal of sigma



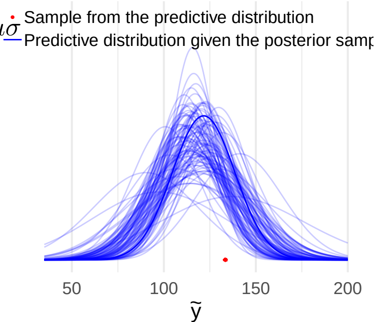
Predictive distribution for new $\tilde{y}$

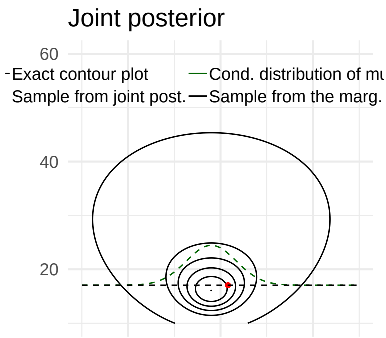
$$p(\tilde{y} \mid y) = \int p(\tilde{y} \mid \mu, \sigma) p(\mu, \sigma \mid y) d\mu\sigma$$

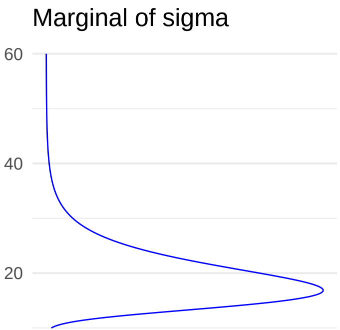
$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

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

Posterior predictive distribution





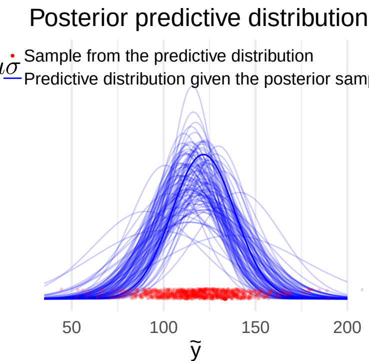


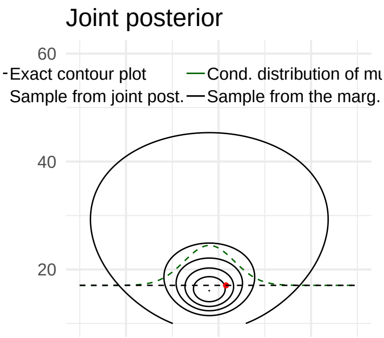
Predictive distribution for new $\tilde{y}$

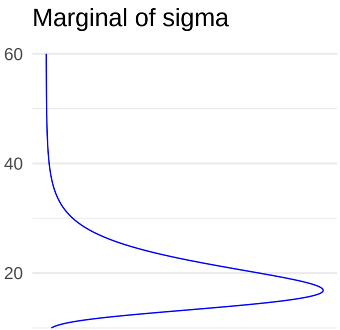
$$p(\tilde{y} \mid y) = \int p(\tilde{y} \mid \mu, \sigma) p(\mu, \sigma \mid y) d\mu\sigma$$

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

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





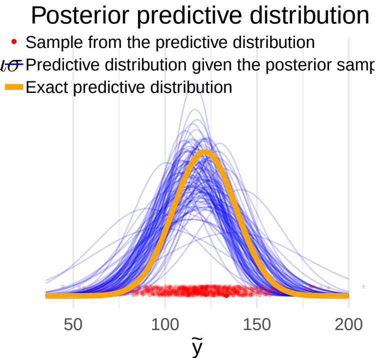


Predictive distribution for new $\tilde{y}$

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

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

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



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

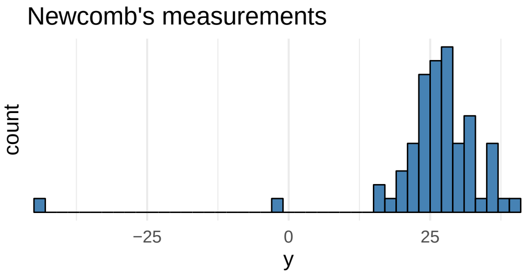
$$
\begin{aligned}
p(\tilde{y} \mid \sigma^2, y) &= \int p(\tilde{y} \mid \mu, \sigma^2) p(\mu \mid \sigma^2, y) d\mu \\
&= \int \mathrm{N}(\tilde{y} \mid \mu, \sigma^2)\, \mathrm{N}(\mu \mid \bar{y}, \sigma^2/n) d\mu \\
&= \mathrm{N}(\tilde{y} \mid \bar{y}, (1 + \tfrac{1}{n})\sigma^2)
\end{aligned}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

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

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

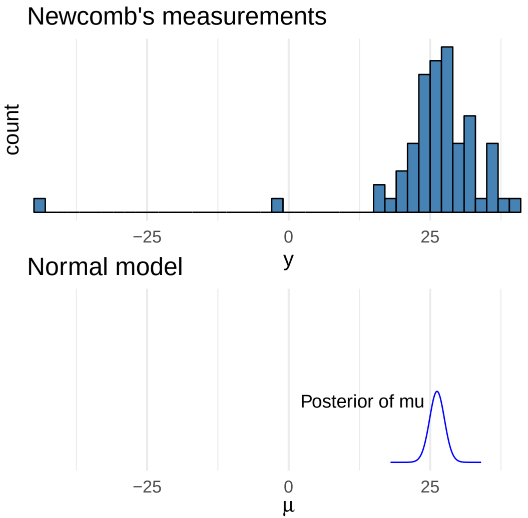
Simon Newcomb's light of speed experiment in 1882

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.



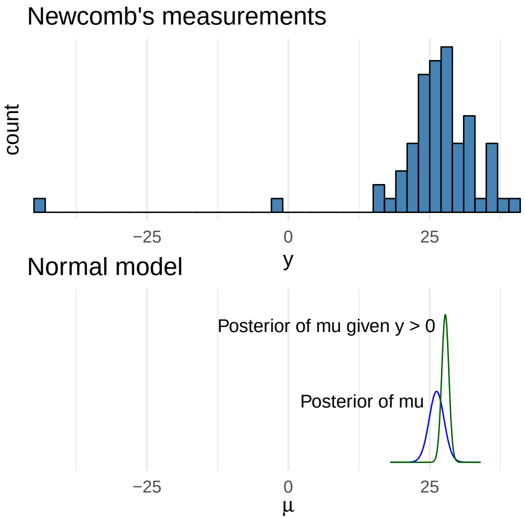
Simon Newcomb's light of speed experiment in 1882

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.



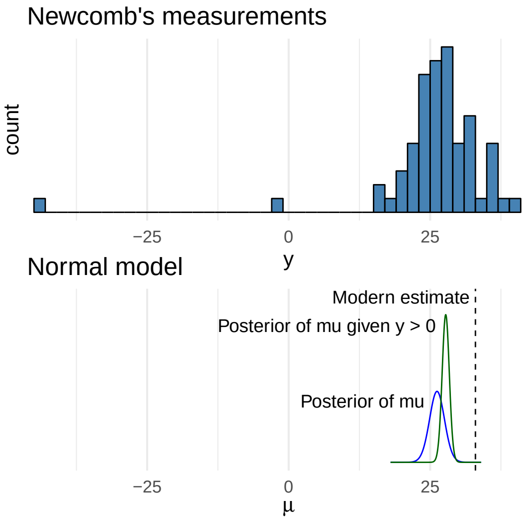
Simon Newcomb's light of speed experiment in 1882

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.



Multinomial model for categorical data

- Extension of binomial
- Observation model

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

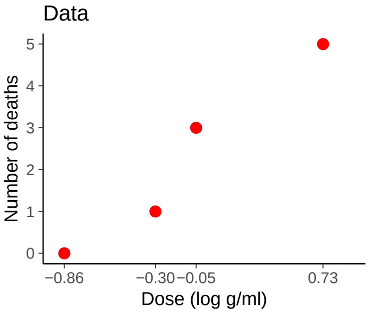
Multivariate Gaussian

- Observation model

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

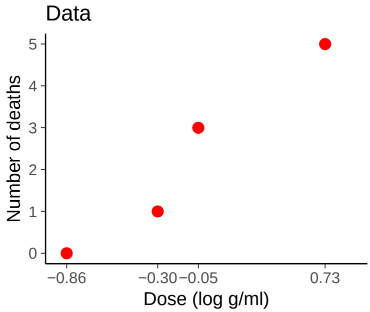
Bioassay

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

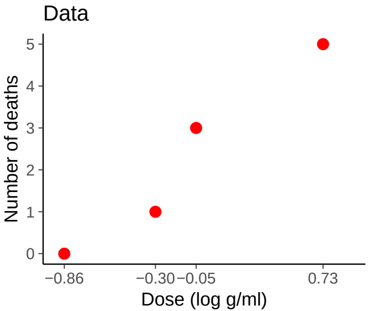


Find out lethal dose 50% (LD50)

- used to classify how hazardous chemical is

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



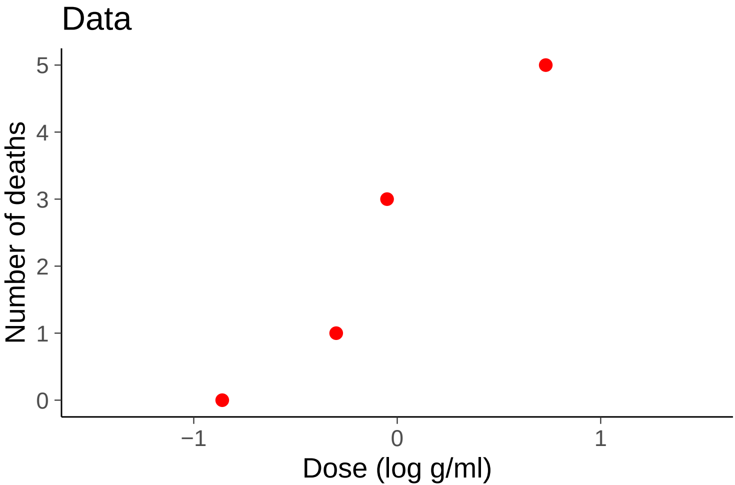
Find out lethal dose 50% (LD50)

- used to classify how hazardous chemical is

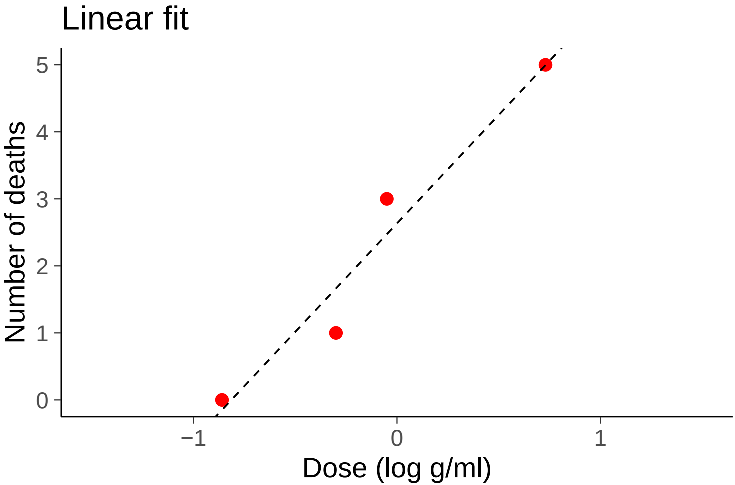
Bayesian methods help to

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

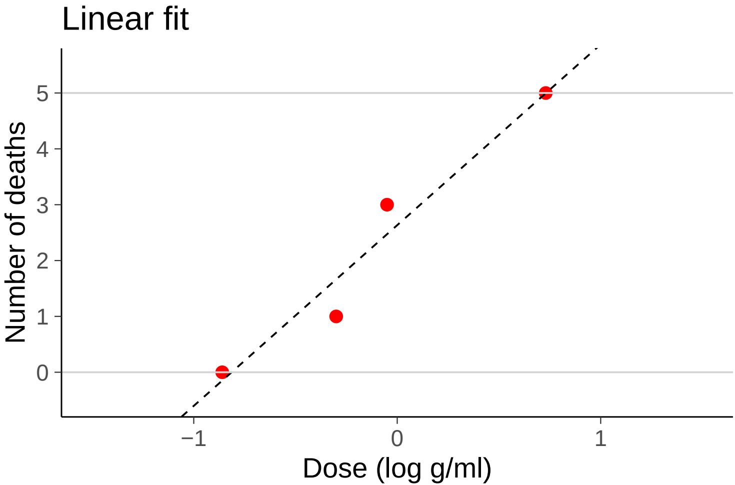
Bioassay



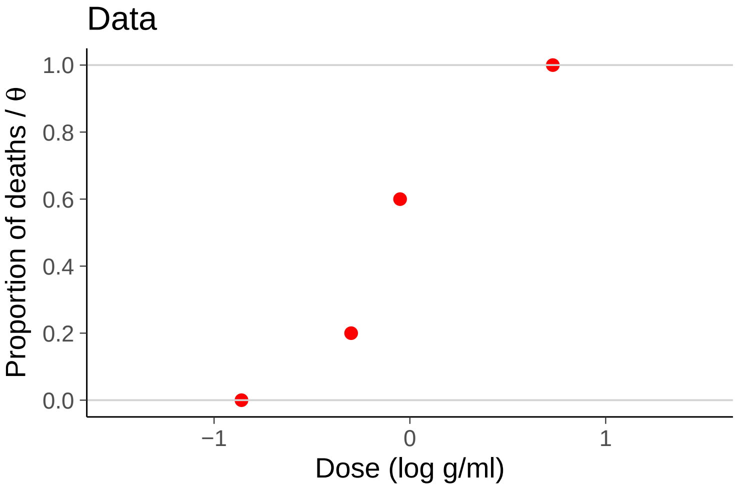
Bioassay



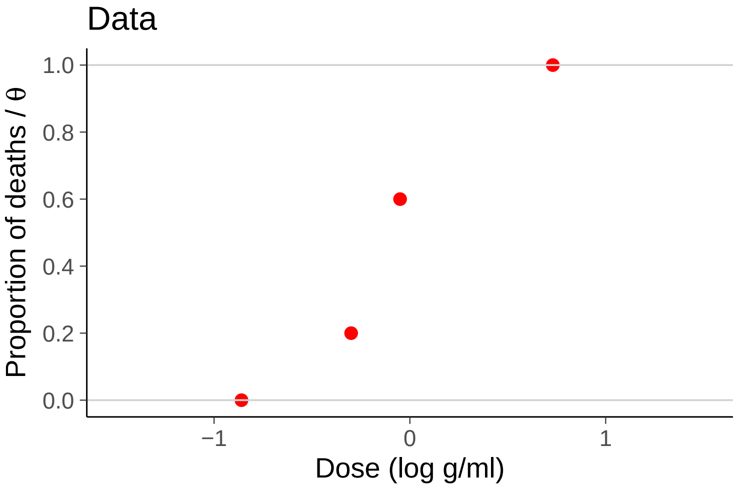
Bioassay



Bioassay



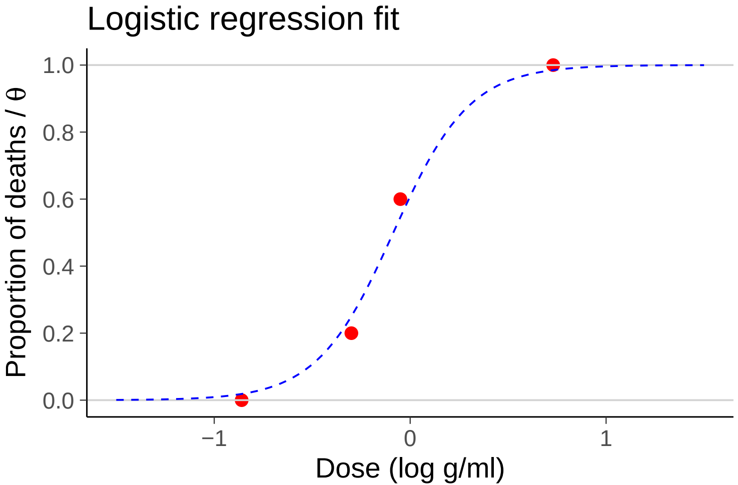
Bioassay



Binomial model

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

Bioassay



Binomial model

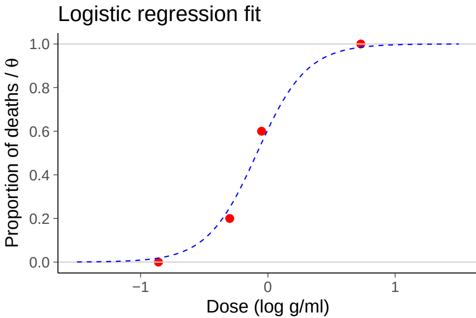
$$y_i \mid \theta_i \sim \mathrm{Bin}(n_i, \theta_i), \quad \mathrm{logit}(\theta_i) = \log\left(\frac{\theta_i}{1-\theta_i}\right) = \alpha + \beta x_i$$

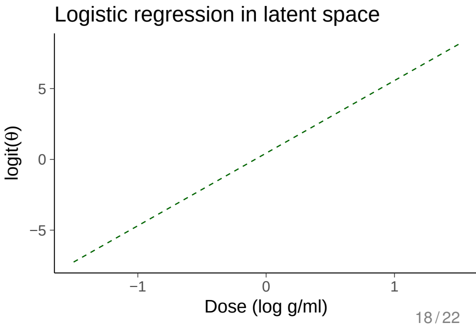
Bioassay

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

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

$$= \alpha + \beta x_i$$



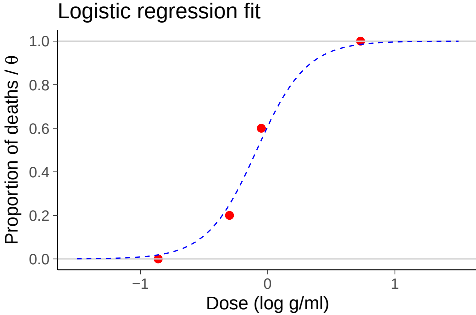


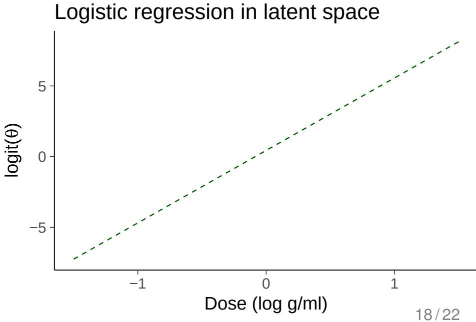
Bioassay

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

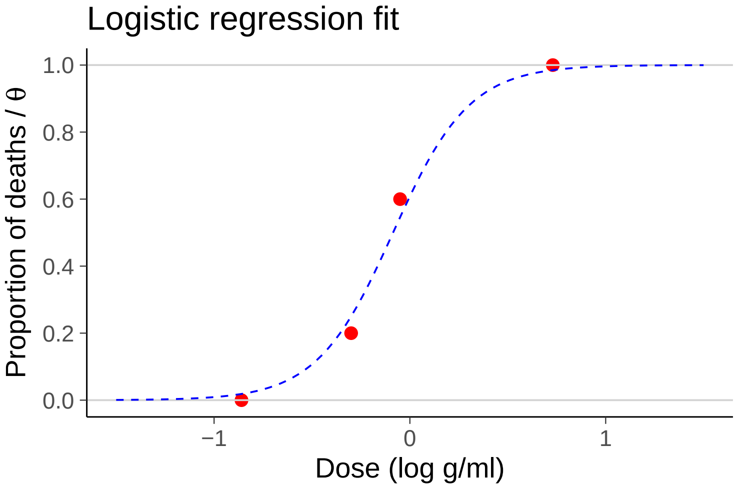
$$\text{logit}(\theta_i) = \log\left(\frac{\theta_i}{1-\theta_i}\right)$$
$$= \alpha + \beta x_i$$

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



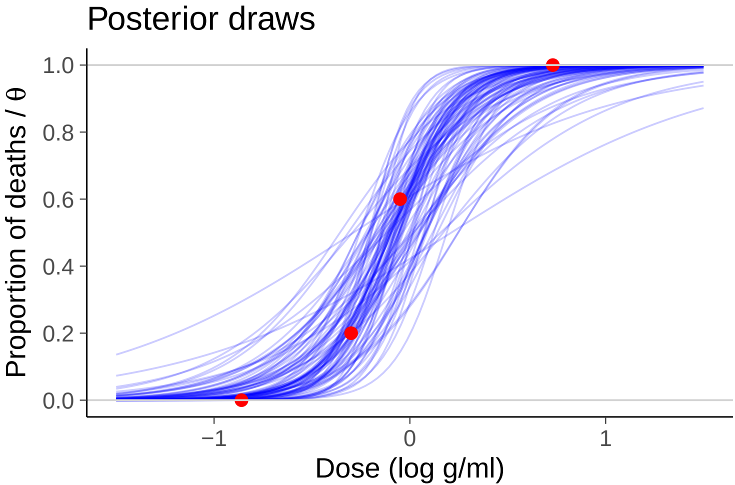


Bioassay

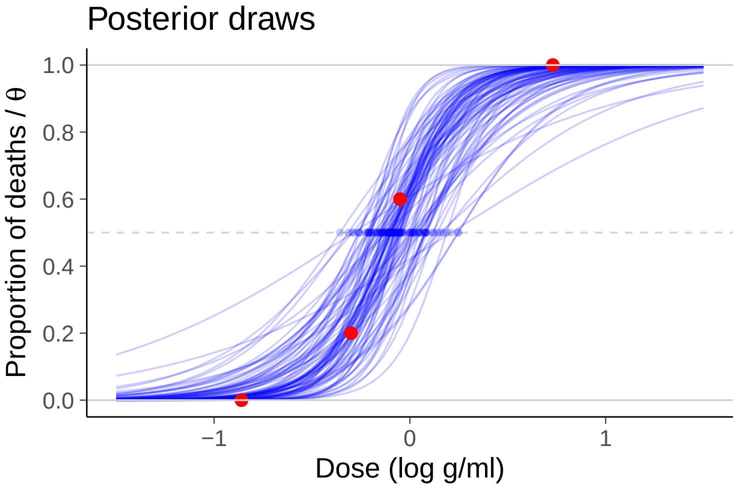


Logistic regression fit
1.0
0.8
0.6
0.4
0.2
0.0
Proportion of deaths / θ
−1
0
1
Dose (log g/ml)

Bioassay

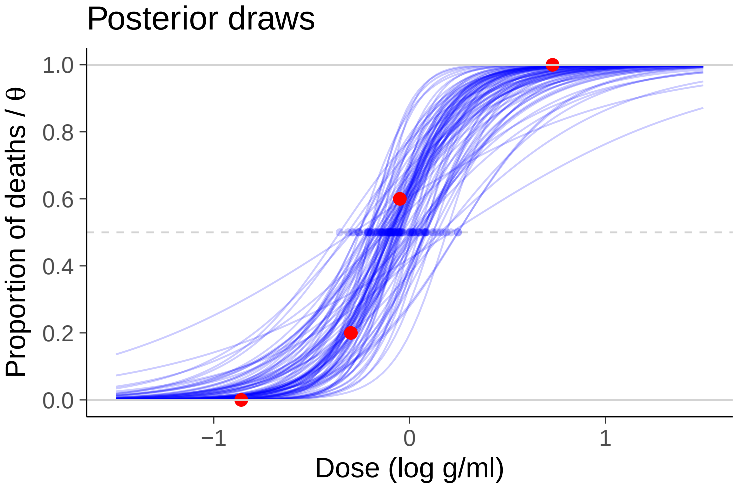

Posterior draws
1.0
0.8
0.6
0.4
0.2
0.0
Proportion of deaths / θ
−1
0
1
Dose (log g/ml)

Bioassay



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

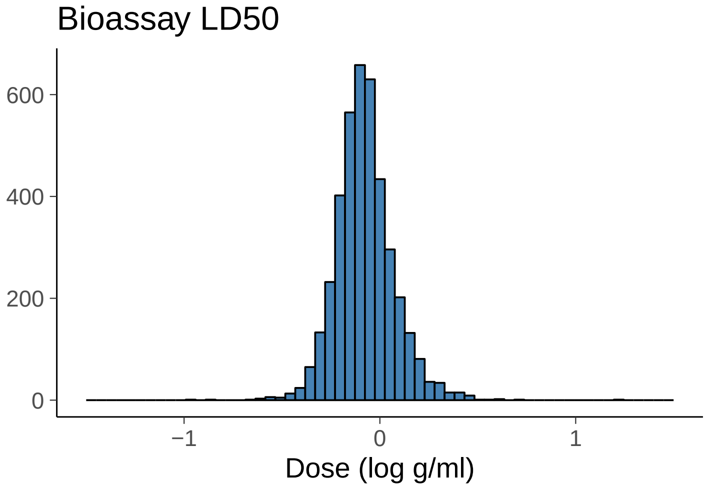
Bioassay



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

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

Bioassay



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

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

Bioassay posterior

Binomial model

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

Link function

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

Bioassay posterior

Binomial model

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

Link function

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

Likelihood

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

Bioassay posterior

Binomial model

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

Link function

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

Likelihood

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

Bioassay posterior

Binomial model

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

Link function

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

Likelihood

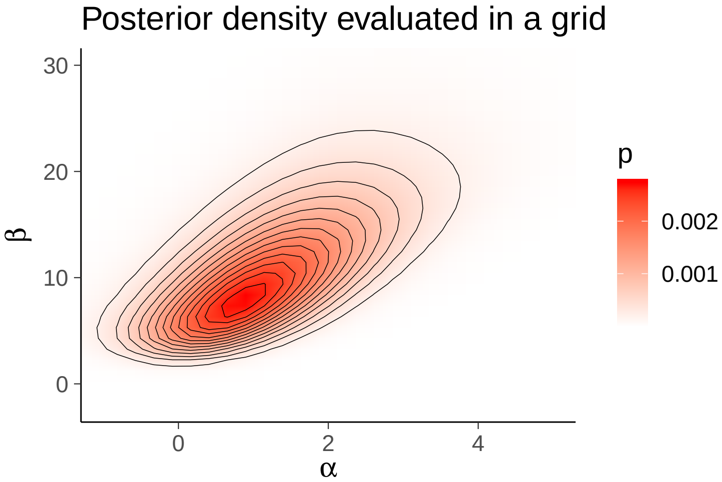
$$p(y_i \mid \alpha, \beta, n_i, x_i) \propto \theta_i^{y_i}[1 - \theta_i]^{n_i - y_i}$$

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

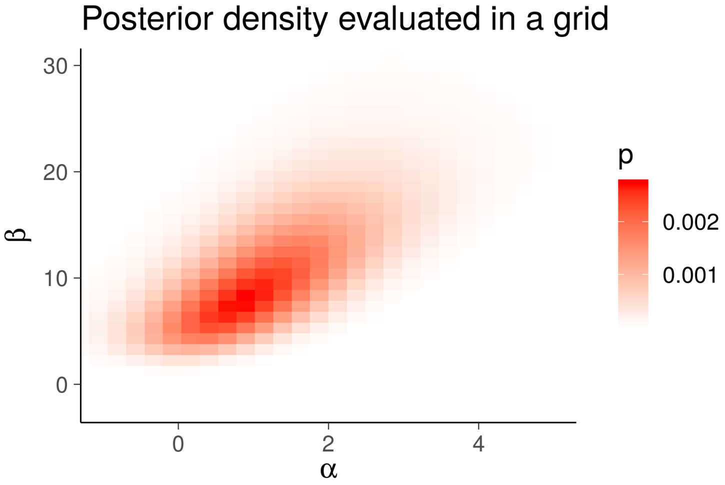
Posterior (with uniform prior on α, β)

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

Bioassay

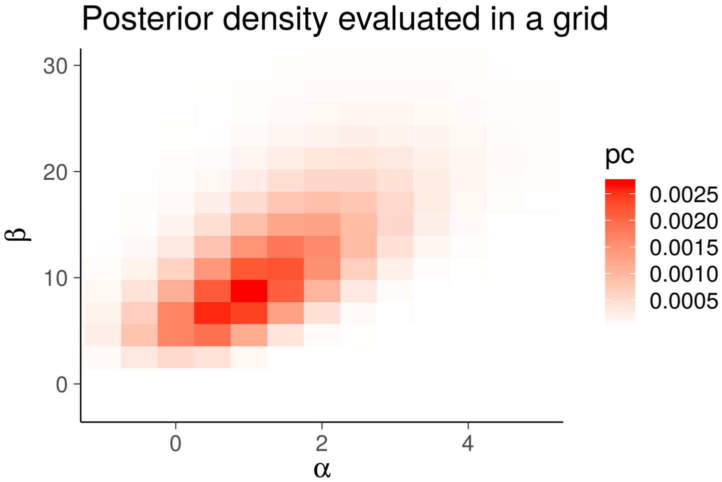

Posterior density evaluated in a grid
30
20
10
0
β
0
2
4
α
p
0.002
0.001

Bioassay



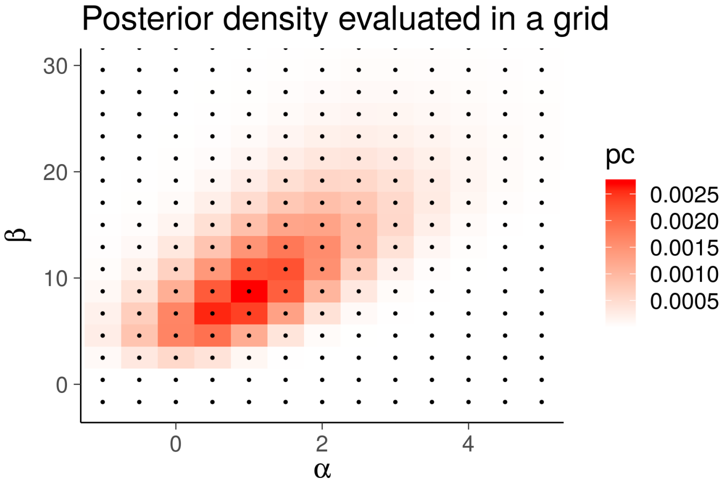
Density evaluated in grid, and plotted without interpolation

Bioassay



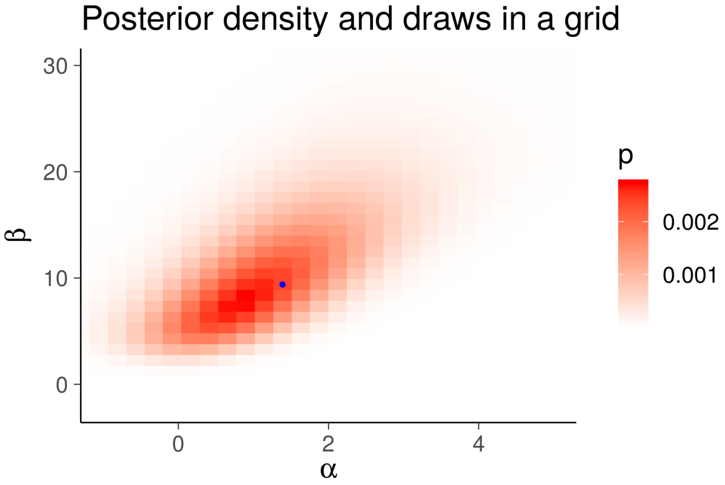
Density evaluated in a coarser grid

Bioassay



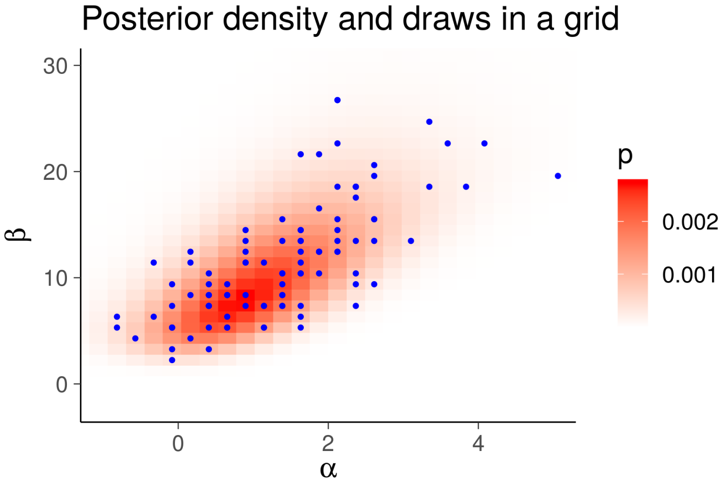
- Approximate the density as piecewise constant function
- Evaluate density in a grid over some finite region
- Density times cell area gives probability mass in each cell

Bioassay



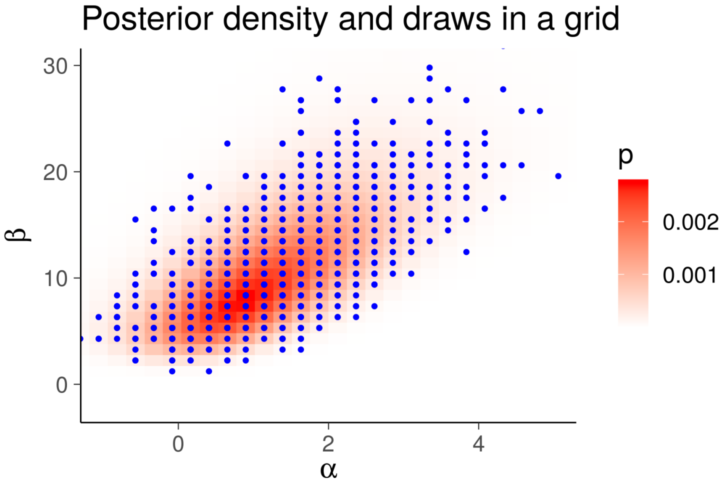
- Sample according to grid cell probabilities

Bioassay



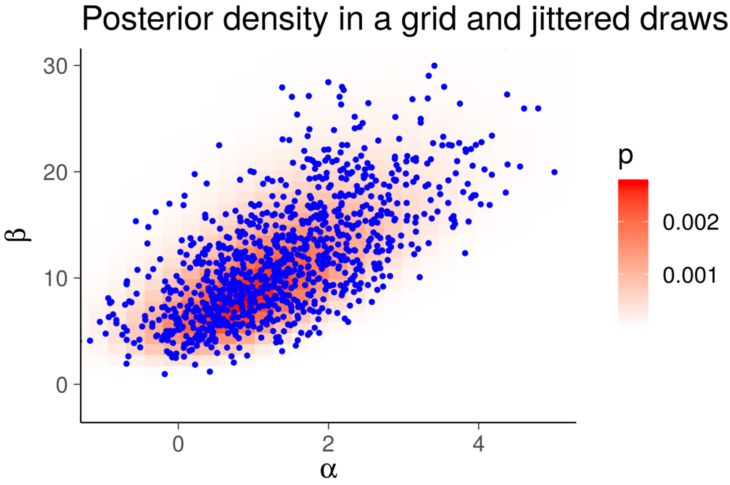
- Sample according to grid cell probabilities
- Several draws can be from the same grid cell

Bioassay



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

Bioassay



- Jitter can be added to improve visualization

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

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