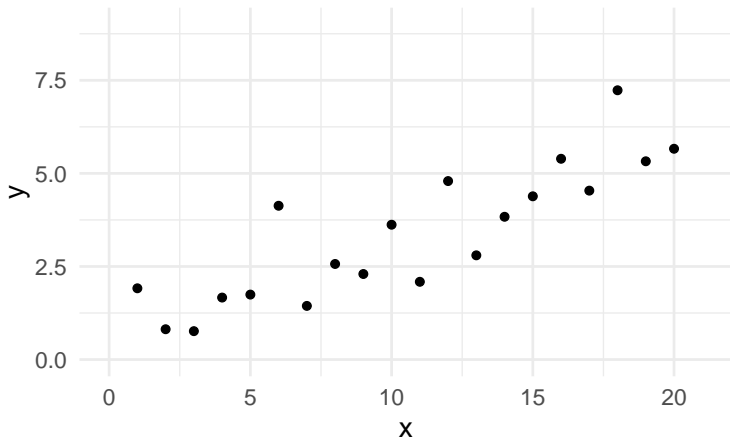


# Chapter 3

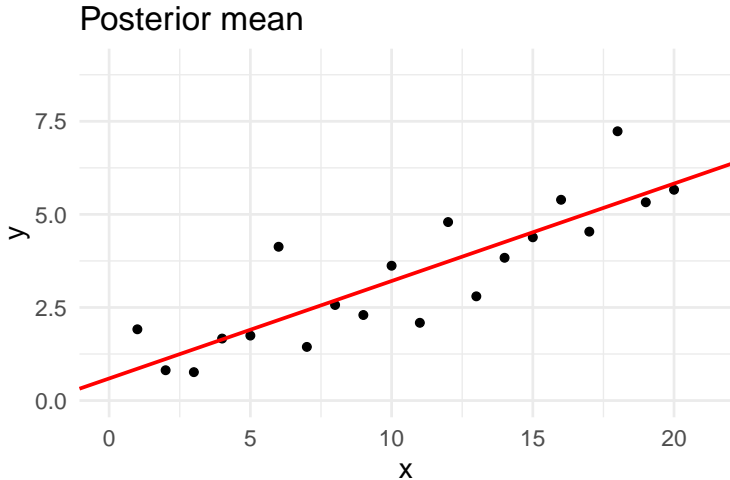
- 3.1 Marginalisation
- 3.2 Normal distribution with a noninformative prior (important)
- 3.3 Normal distribution with a conjugate prior (important)
- 3.4 Multinomial model (can be skipped)
- 3.5 Multivariate normal with known variance (useful for chapter 4)
- 3.6 Multivariate normal with unknown variance (glance through)
- 3.7 Bioassay example (very important, related to one of the exercises)
- 3.8 Summary (summary)

## Example of uncertainty in modeling

Data

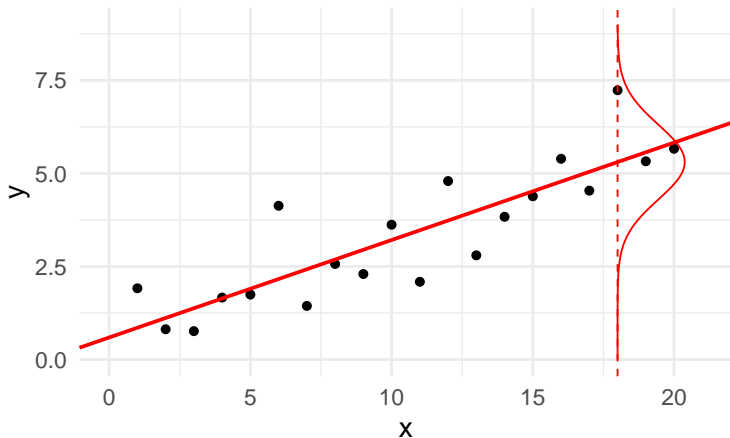


## Example of uncertainty in modeling



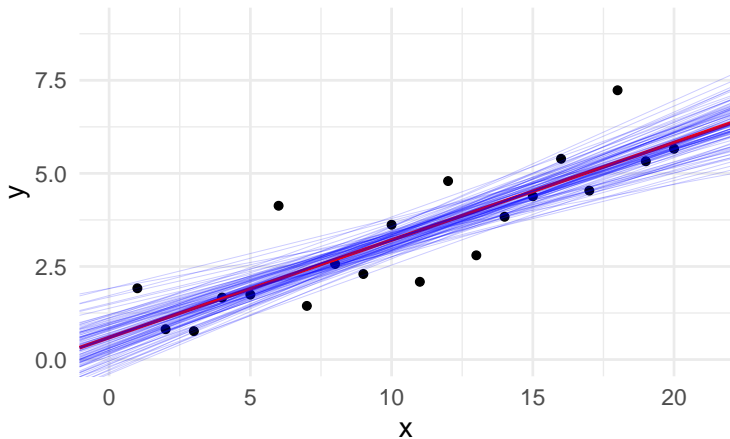
## Example of uncertainty in modeling

Predictive distribution given posterior mean



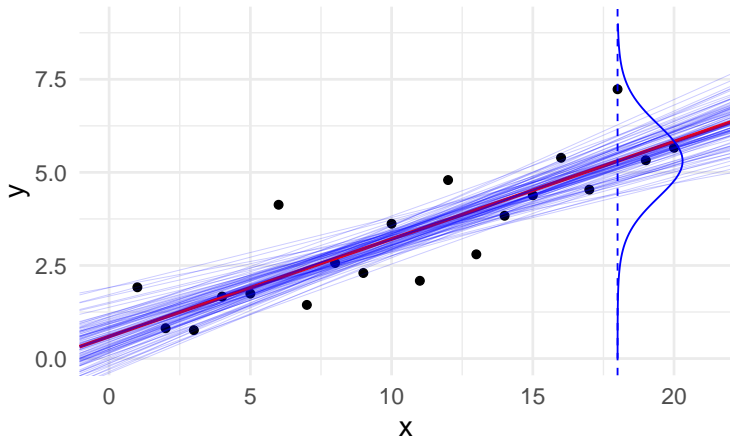
## Example of uncertainty in modeling

Posterior draws



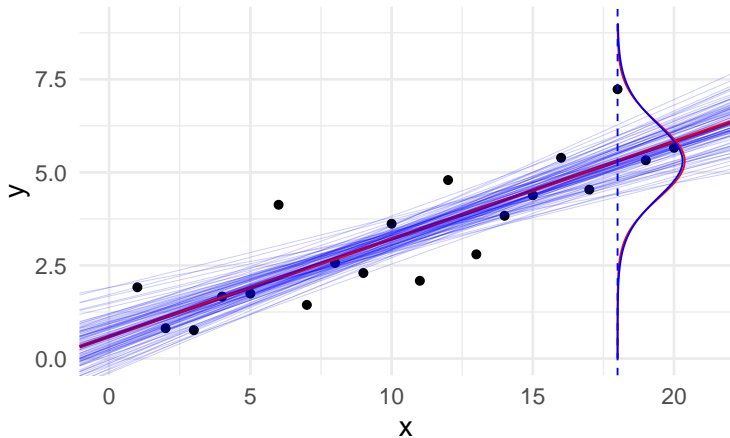
## Example of uncertainty in modeling

### Posterior draws and predictive distribution



## Example of uncertainty in modeling

### Posterior draws and predictive distribution



## Marginalization

- Joint distribution of 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



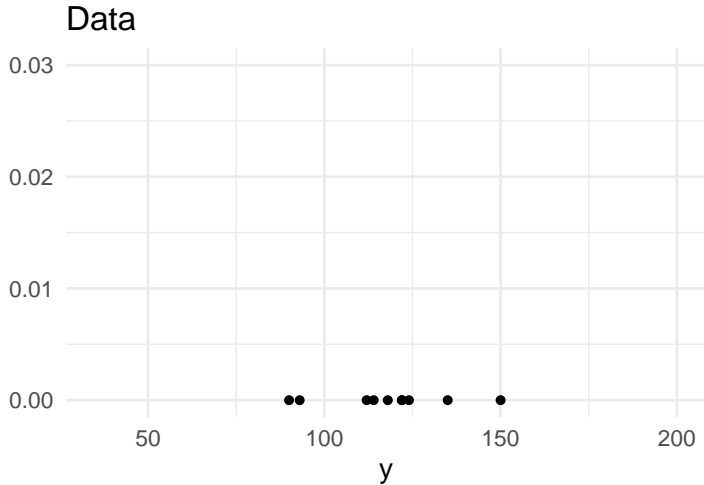
## Marginalization - predictive distribution

- Marginalization over posterior distribution

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

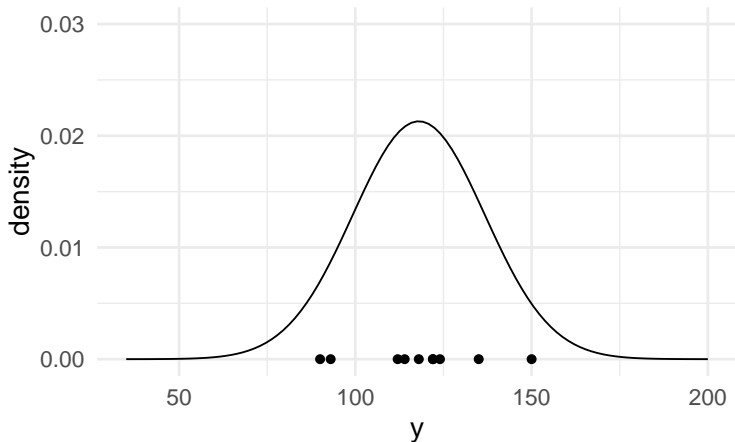
$p(\tilde{y} | y)$  is a predictive distribution

## Gaussian example



## Gaussian example

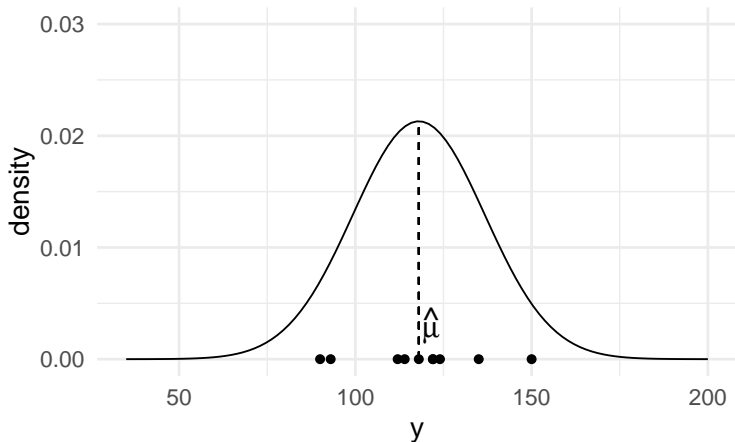
Gaussian fit with posterior mean



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

## Gaussian example

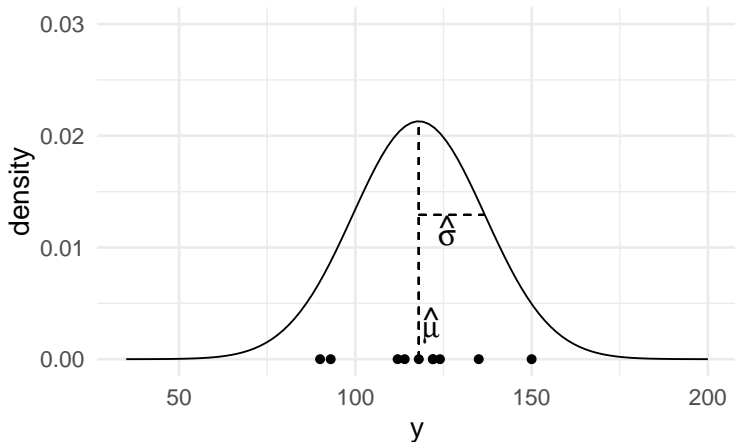
Gaussian fit with posterior mean



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

## Gaussian example

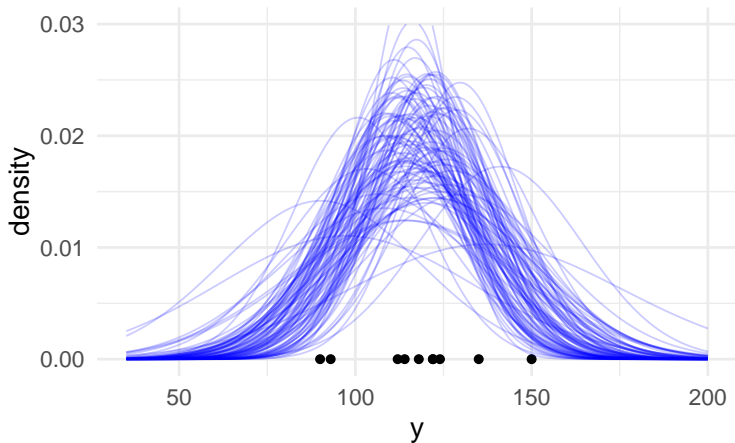
Gaussian fit with posterior mean



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

## Gaussian example

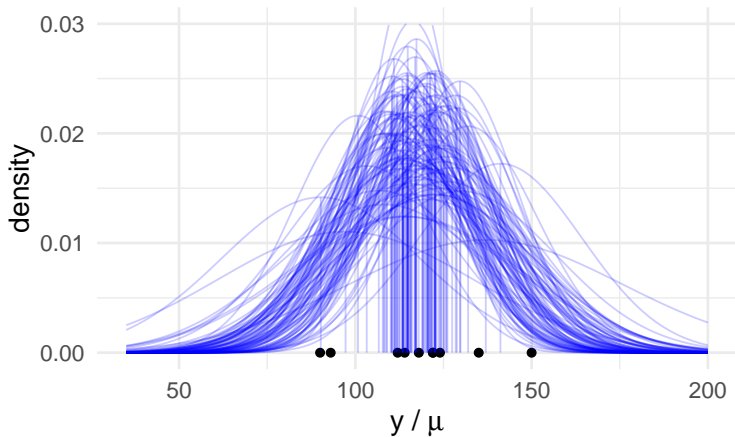
### Gaussians with posterior draw parameters



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

## Gaussian example

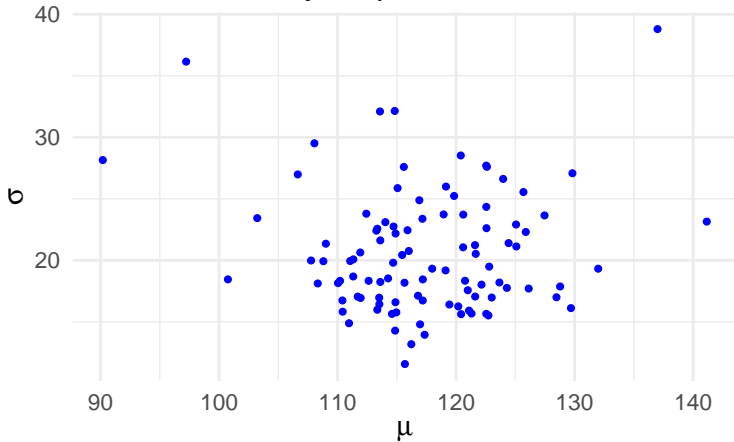
### Gaussians with posterior draw parameters



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

## Gaussian example

Draws from the joint posterior distribution

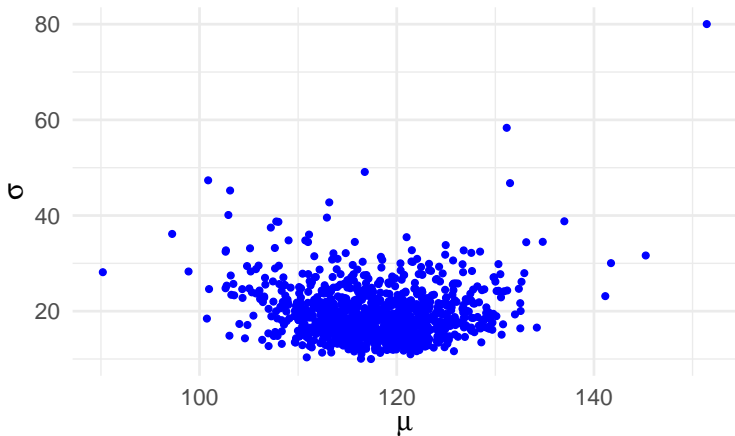


$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$



## Gaussian example

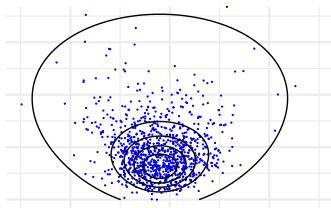
Draws from the joint posterior distribution



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

Joint posterior

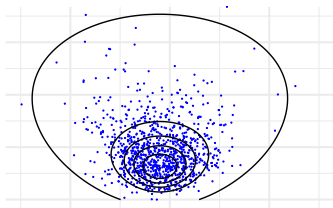
$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$



Joint posterior

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

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

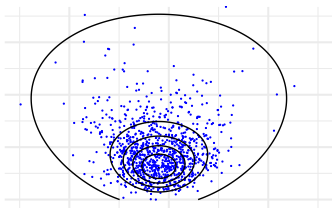


Joint posterior

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

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

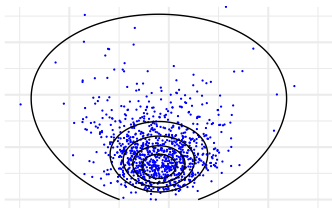
$$p(\mu, \sigma^2 \mid y) \propto \sigma^{-n-2} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \right)$$



Joint posterior

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

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



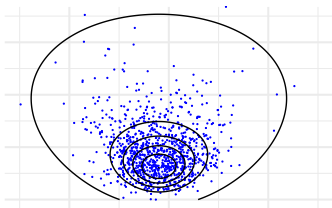
$$\begin{aligned} p(\mu, \sigma^2 \mid y) &\propto \sigma^{-n-2} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \right) \\ &= \sigma^{-n-2} \exp \left( -\frac{1}{2\sigma^2} \left[ \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - \mu)^2 \right] \right) \end{aligned}$$

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

Joint posterior

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

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



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

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

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

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

## Gaussian - non-informative prior

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

## Gaussian - non-informative prior

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

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



## Gaussian - non-informative prior

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

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

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

## Gaussian - non-informative prior

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

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

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

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

## Gaussian - non-informative prior

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

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

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

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

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

## Gaussian - non-informative prior

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

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

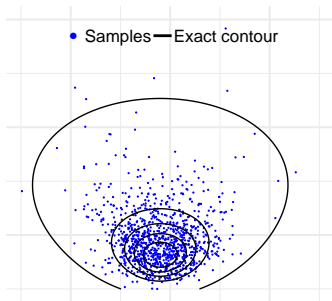
$$\sum_{i=1}^n (y_i^2 - 2y_i\mu + \mu^2 - \bar{y}^2 + \bar{y}^2 - 2y_i\bar{y} + 2y_i\bar{y})$$

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

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

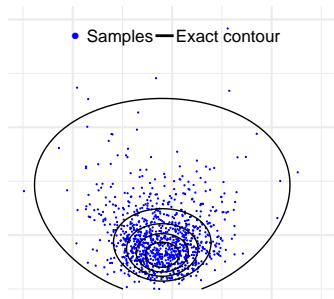
$$\sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - \mu)^2$$

## Joint posterior

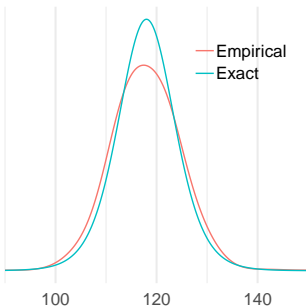


$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma \mid y)$$

## Joint posterior



## Marginal of mu

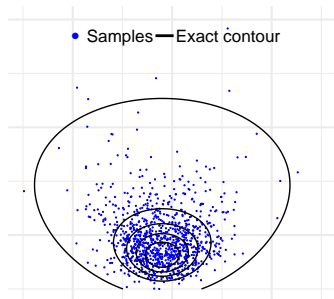


$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma | y)$$

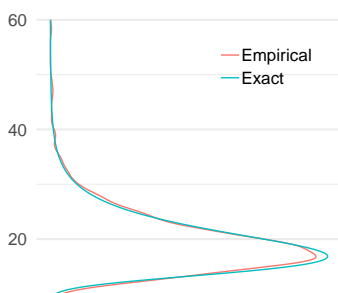
marginals

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

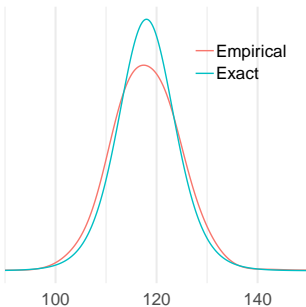
### Joint posterior



### Marginal of sigma



### Marginal of mu



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

marginals

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

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

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

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



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

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

## Marginal posterior $p(\sigma^2 | y)$ (easier for $\sigma^2$ than $\sigma$ )

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

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

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

## Marginal posterior $p(\sigma^2 | y)$ (easier for $\sigma^2$ than $\sigma$ )

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

## Marginal posterior $p(\sigma^2 | y)$ (easier for $\sigma^2$ than $\sigma$ )

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

## Marginal posterior $p(\sigma^2 | y)$ (easier for $\sigma^2$ than $\sigma$ )

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

## Gaussian - non-informative prior

Known mean

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

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

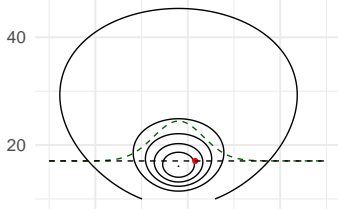
Unknown mean

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

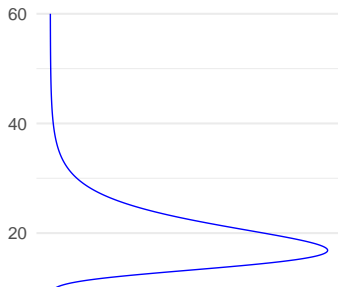
$$\text{where } s^2 = \frac{1}{n - 1} \sum_{i=1}^n (y_i - \bar{y})^2$$

## Joint posterior

60  
- Exact contour plot      — Cond. distribution of  $\mu$   
Sample from joint post. — Sample from the marg.



## Marginal of sigma



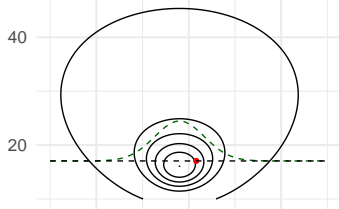
## Factorization

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

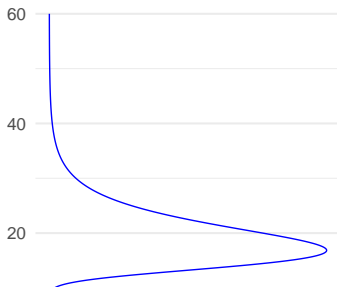


## Joint posterior

60  
 - Exact contour plot      — Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



## Factorization

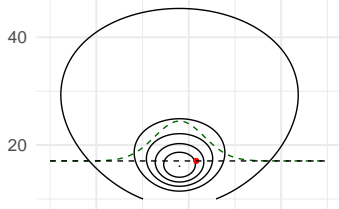
$$p(\mu, \sigma^2 | y) = p(\mu | \sigma^2, y) p(\sigma^2 | y)$$

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

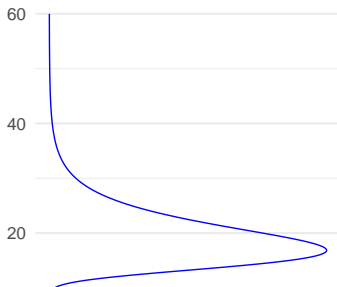
$$(\sigma^2)^{(s)} \sim p(\sigma^2 | y)$$

## Joint posterior

60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



## Factorization

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

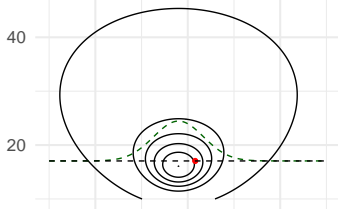
$$p(\sigma^2 | y) = \text{Inv-}\chi^2(\sigma^2 | n - 1, s^2)$$

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

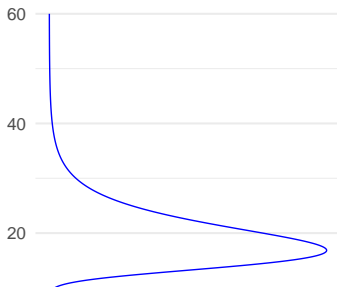
$$p(\mu | \sigma^2, y) = \text{N}(\mu | \bar{y}, \sigma^2/n)$$

## Joint posterior

60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



## Factorization

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

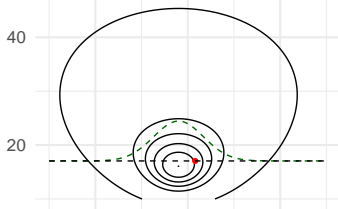
$$p(\sigma^2 | y) = \text{Inv-}\chi^2(\sigma^2 | n-1, s^2)$$

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

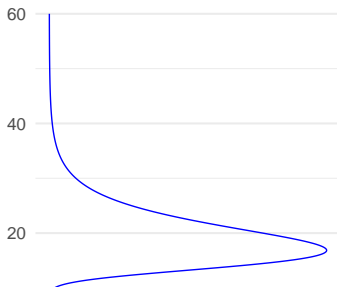
$$p(\mu | \sigma^2, y) = \text{N}(\mu | \bar{y}, \sigma^2/n) \propto \sigma^{-n-2} \exp\left(-\frac{1}{2\sigma^2} [(n-1)s^2 + n(\bar{y} - \mu)^2]\right)$$

## Joint posterior

60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



## Factorization

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

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

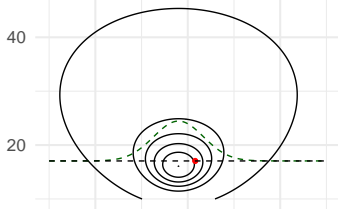
$$(\sigma^2)^{(s)} \sim p(\sigma^2 | y)$$

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

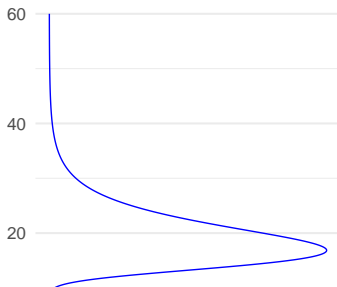
$$\mu^{(s)} \sim p(\mu | \sigma^2, y)$$

## Joint posterior

60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post.   - Sample from the marg.



## Marginal of sigma



## Factorization

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

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

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

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

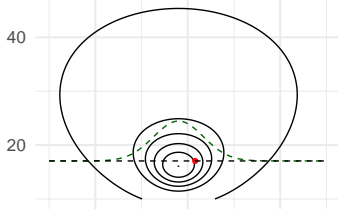
$$\mu^{(s)} \sim p(\mu | \sigma^2, y)$$

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

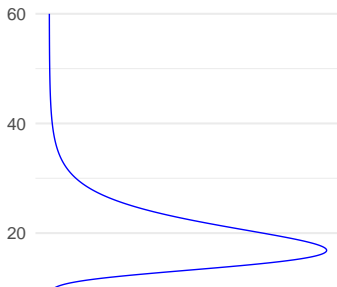
## Joint posterior

60  
40  
20

- Exact contour plot      — Cond. distribution of  $\mu$   
Sample from joint post.   — Sample from the marg.



## Marginal of sigma

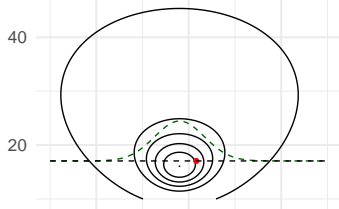


## Factorization

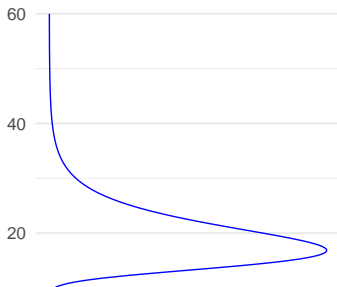
$$p(\mu, \sigma^2 | y) = p(\mu | \sigma^2, y) p(\sigma^2 | y)$$

## Joint posterior

60  
 - Exact contour plot      — Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



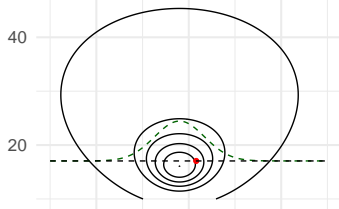
## Factorization

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

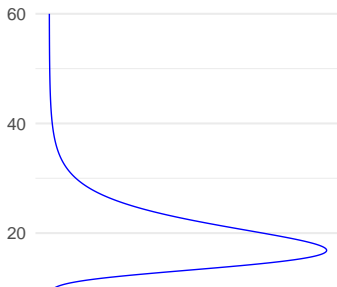
$$(\sigma^2)^{(s)} \sim p(\sigma^2 | y)$$

## Joint posterior

60  
 - Exact contour plot      — Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



## Factorization

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

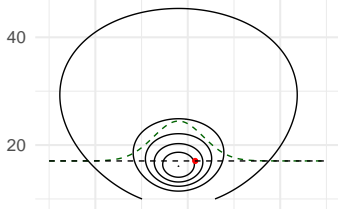
$$(\sigma^2)^{(s)} \sim p(\sigma^2 | y)$$

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

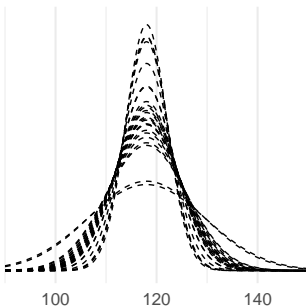


## Joint posterior

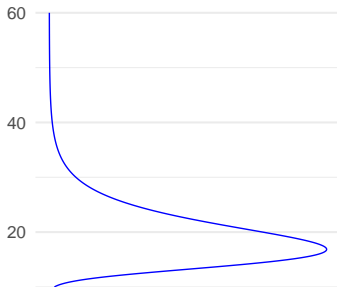
60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post.   - Sample from the marg.



## Cond distr of mu for 25 draws



## Marginal of sigma



## Factorization

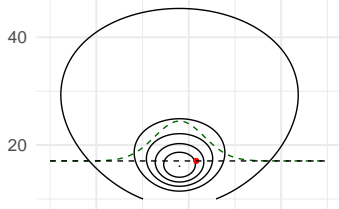
$$p(\mu, \sigma^2 | y) = p(\mu | \sigma^2, y) p(\sigma^2 | y)$$

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

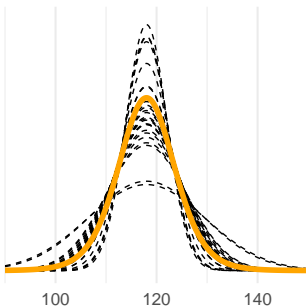
$$p(\mu | (\sigma^2)^{(s)}, y) = \text{N}(\mu | \bar{y}, (\sigma^2)^{(s)}/n)$$

## Joint posterior

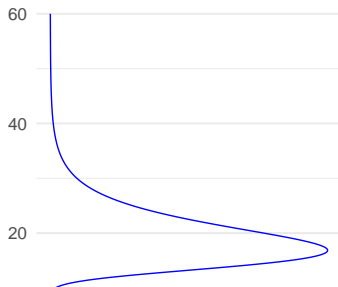
60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Cond distr of $\mu$ for 25 draws



## Marginal of sigma



## Factorization

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

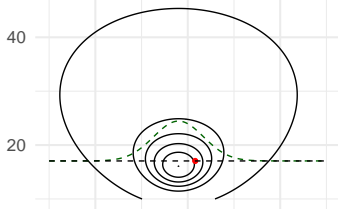
$$(\sigma^2)^{(s)} \sim p(\sigma^2 | y)$$

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

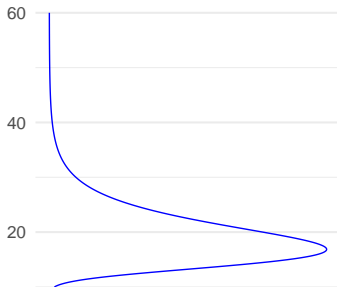
$$p(\mu | y) \approx \frac{1}{S} \sum_{s=1}^S N(\mu | \bar{y}, (\sigma^2)^{(s)}/n)$$

## Joint posterior

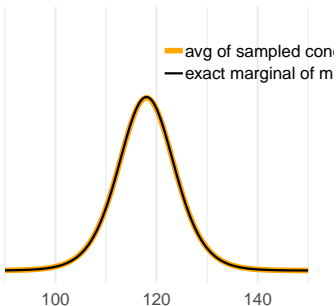
60  
- Exact contour plot      - Cond. distribution of  $\mu$   
Sample from joint post. — Sample from the marg.



## Marginal of sigma



## Cond. distr of mu



## Factorization

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

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

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

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

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 $p(\mu | y)$

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

Transformation

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

## Marginal posterior $p(\mu | y)$

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

Transformation

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

## Marginal posterior $p(\mu | y)$

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

Transformation

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

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



## Marginal posterior $p(\mu | y)$

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

Transformation

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

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

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

## Marginal posterior $p(\mu | y)$

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

Transformation

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

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

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 $p(\mu | y)$

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

Transformation

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

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

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

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

## Marginal posterior $p(\mu | y)$

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

Transformation

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

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

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

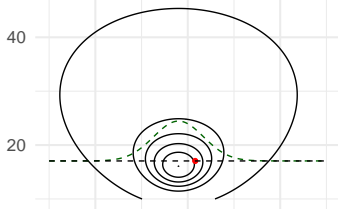
$$\propto [(n-1)s^2 + n(\mu - \bar{y})^2]^{-n/2}$$

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

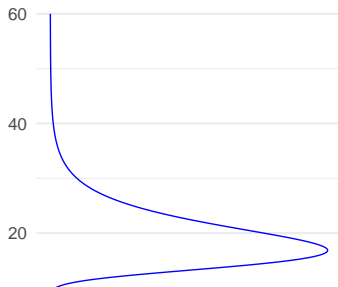
$$p(\mu | y) = t_{n-1}(\mu | \bar{y}, s^2/n) \quad \text{Student's } t$$

## Joint posterior

60  
- Exact contour plot      - Cond. distribution of  $\mu$   
Sample from joint post. — Sample from the marg.



## Marginal of sigma

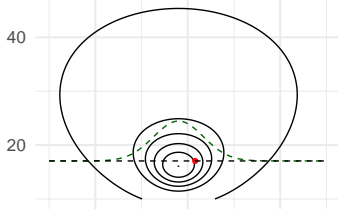


Predictive distribution for new  $\tilde{y}$

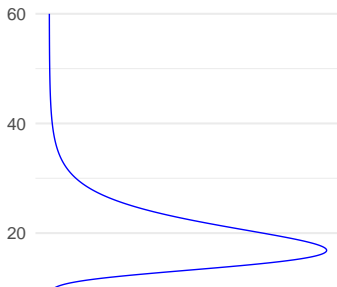
$$p(\tilde{y} | y) = \int p(\tilde{y} | \mu, \sigma) p(\mu, \sigma | y) d\mu d\sigma$$

## Joint posterior

60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



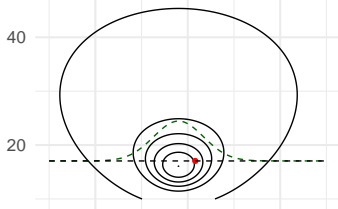
Predictive distribution for new  $\tilde{y}$

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

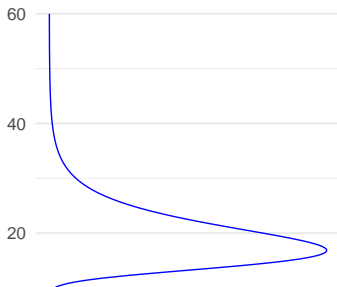
$$\mu^{(s)}, \sigma^{(s)} \sim p(\mu, \sigma | y)$$

## Joint posterior

60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



## Predictive distribution for new $\tilde{y}$

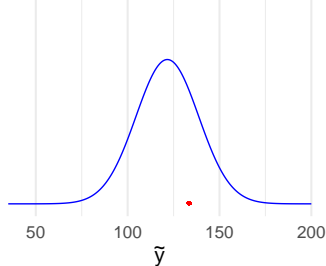
$$p(\tilde{y} | y) = \int p(\tilde{y} | \mu, \sigma) p(\mu, \sigma | y) d\mu d\sigma$$

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

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

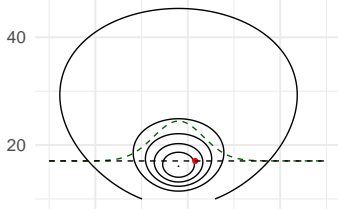
## Posterior predictive distribution

Sample from the predictive distribution  
 Predictive distribution given the posterior sample

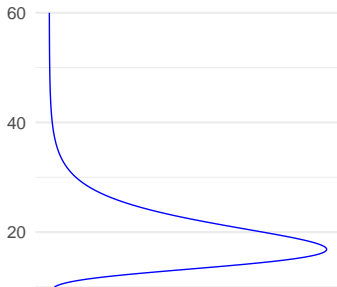


## Joint posterior

60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



## Predictive distribution for new $\tilde{y}$

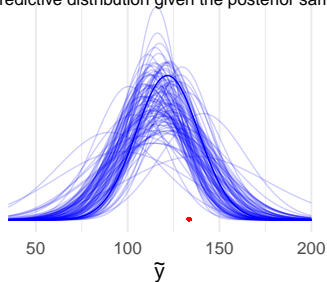
$$p(\tilde{y} | y) = \int p(\tilde{y} | \mu, \sigma) p(\mu, \sigma | y) d\mu d\sigma$$

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

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

## Posterior predictive distribution

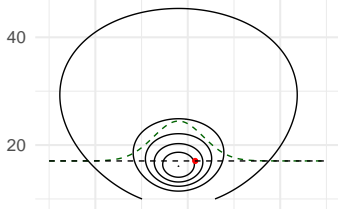
Sample from the predictive distribution  
 Predictive distribution given the posterior sample



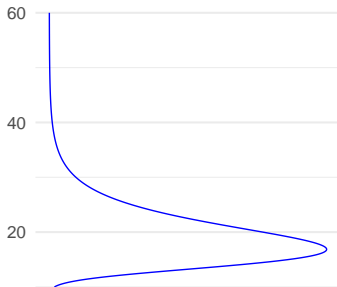


## Joint posterior

60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



## Predictive distribution for new $\tilde{y}$

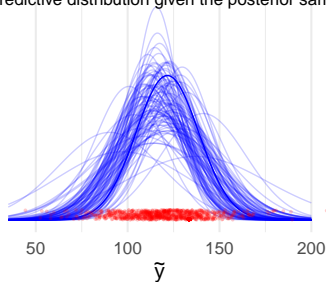
$$p(\tilde{y} | y) = \int p(\tilde{y} | \mu, \sigma) p(\mu, \sigma | y) d\mu d\sigma$$

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

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

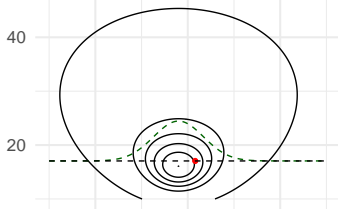
## Posterior predictive distribution

Sample from the predictive distribution  
 Predictive distribution given the posterior sample

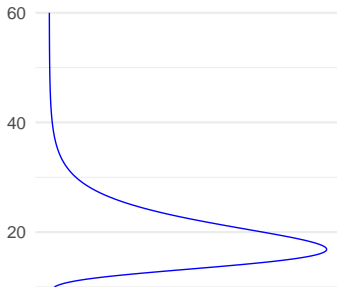


## Joint posterior

60  
 - Exact contour plot      - Cond. distribution of  $\mu$   
 Sample from joint post. — Sample from the marg.



## Marginal of sigma



Predictive distribution for new  $\tilde{y}$

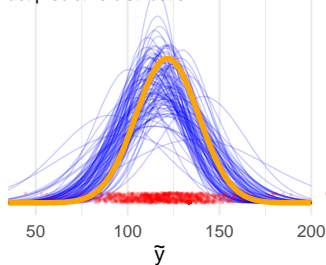
$$p(\tilde{y} | y) = \int p(\tilde{y} | \mu, \sigma) p(\mu, \sigma | y) d\mu d\sigma$$

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

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

Posterior predictive distribution

• Sample from the predictive distribution  
 - Predictive distribution given the posterior sample  
 — Exact predictive distribution



## 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

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

## Gaussian - posterior predictive distribution

Posterior predictive distribution given known variance

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

## Gaussian - posterior predictive distribution

Posterior predictive distribution given known variance

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

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

## Gaussian - posterior predictive distribution

Posterior predictive distribution given known variance

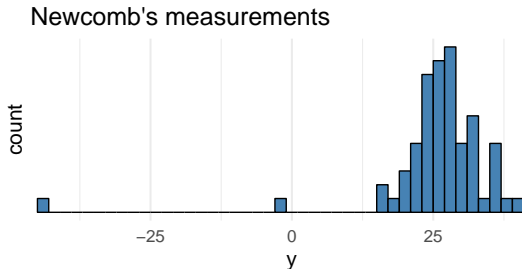
$$\begin{aligned} p(\tilde{y} \mid \sigma^2, y) &= \int p(\tilde{y} \mid \mu, \sigma^2) p(\mu \mid \sigma^2, y) d\mu \\ &= \int \mathcal{N}(\tilde{y} \mid \mu, \sigma^2) \mathcal{N}(\mu \mid \bar{y}, \sigma^2/n) d\mu \\ &= \mathcal{N}(\tilde{y} \mid \bar{y}, (1 + \frac{1}{n})\sigma^2) \end{aligned}$$

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

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

## Simon Newcomb's light of speed experiment in 1882

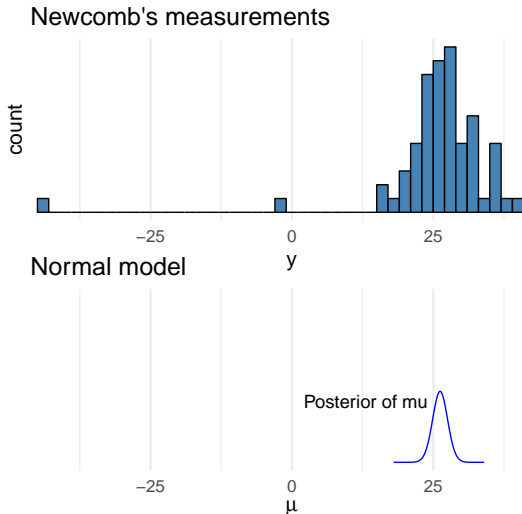
Newcomb measured ( $n = 66$ ) the time required for light to travel from his laboratory on the Potomac River to a mirror at the base of the Washington Monument and back, a total distance of 7422 meters.





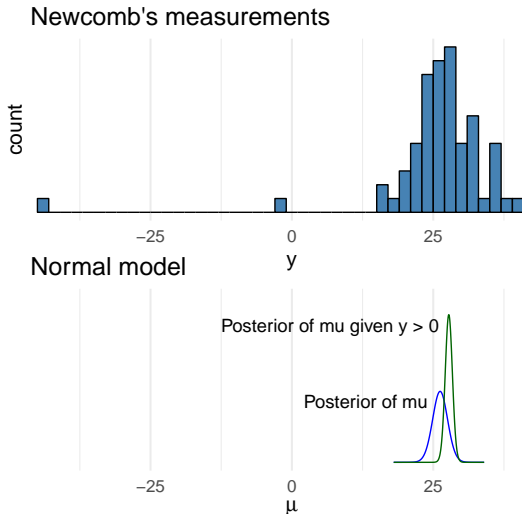
## Simon Newcomb's light of speed experiment in 1882

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.



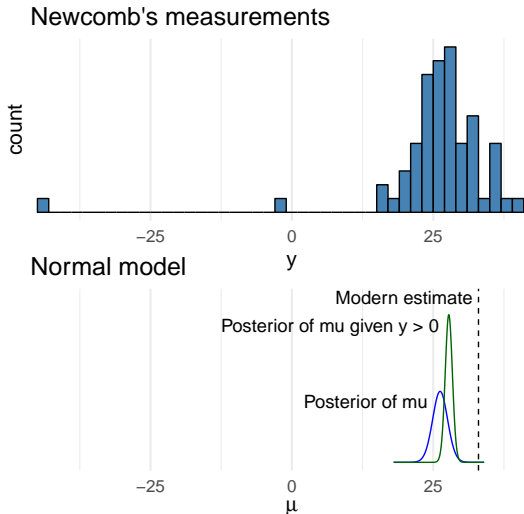
## Simon Newcomb's light of speed experiment in 1882

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  $p(\sigma^2)p(\mu | \sigma^2)$   
(see the chapter notes)

## Gaussian - conjugate prior

- Conjugate prior has to have a form  $p(\sigma^2)p(\mu \mid \sigma^2)$  (see the chapter notes)
- Handy parametrization

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

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  $p(\sigma^2)p(\mu | \sigma^2)$  (see the chapter notes)
- Handy parametrization

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

which can be written as

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

- $\mu$  and  $\sigma^2$  are a priori dependent
  - if  $\sigma^2$  is large, then  $\mu$  has wide prior

## 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

$$\mu_n = \frac{\kappa_0}{\kappa_0 + n} \mu_0 + \frac{n}{\kappa_0 + n} \bar{y}$$

$$\kappa_n = \kappa_0 + n$$

$$\nu_n = \nu_0 + n$$

$$\nu_n \sigma_n^2 = \nu_0 \sigma_0^2 + (n-1)s^2 + \frac{\kappa_0 n}{\kappa_0 + n} (\bar{y} - \mu_0)^2$$

## Multinomial model for categorical data

- Extension of binomial
- Observation model

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

- BDA3 p. 69-



## Multivariate Gaussian

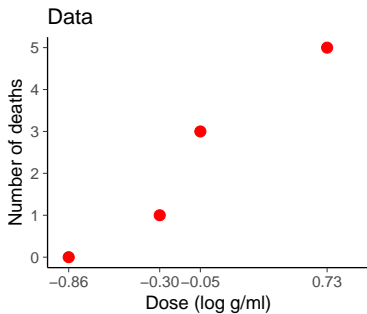
- Observation model

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

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

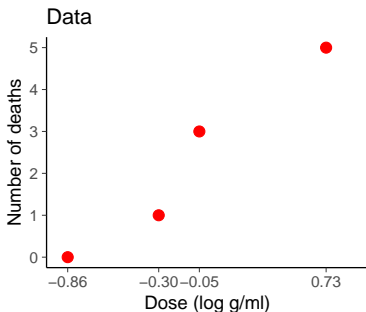
## Bioassay

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

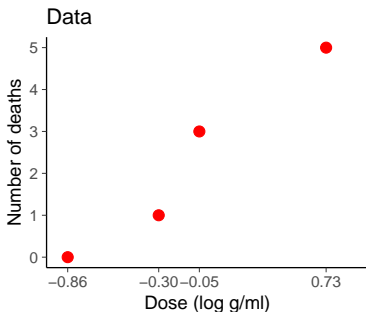


Find out lethal dose 50% (LD50)

- used to classify how hazardous chemical is
- 1984 EEC directive has 4 levels (see the chapter notes)

## 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



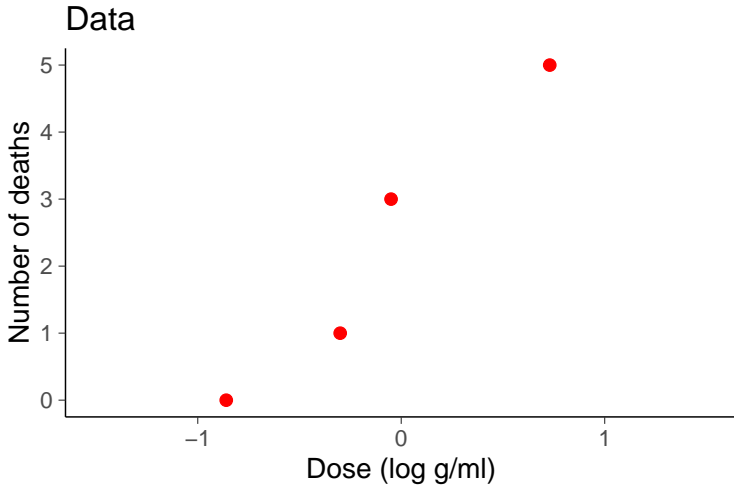
Find out lethal dose 50% (LD50)

- used to classify how hazardous chemical is
- 1984 EEC directive has 4 levels (see the chapter notes)

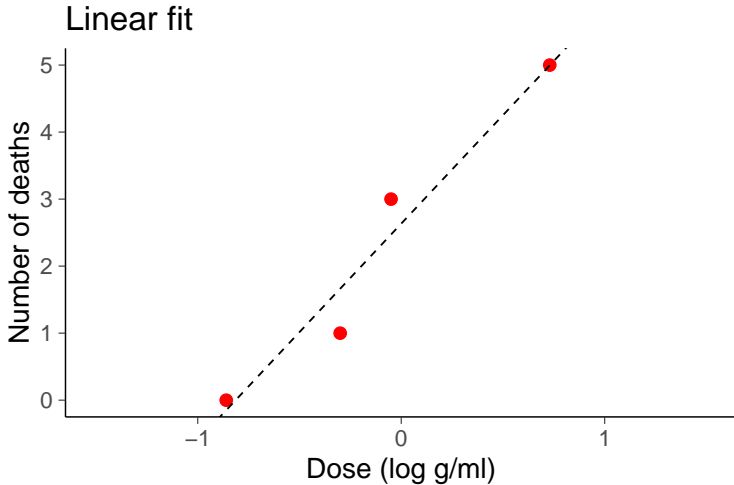
Bayesian methods help to

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

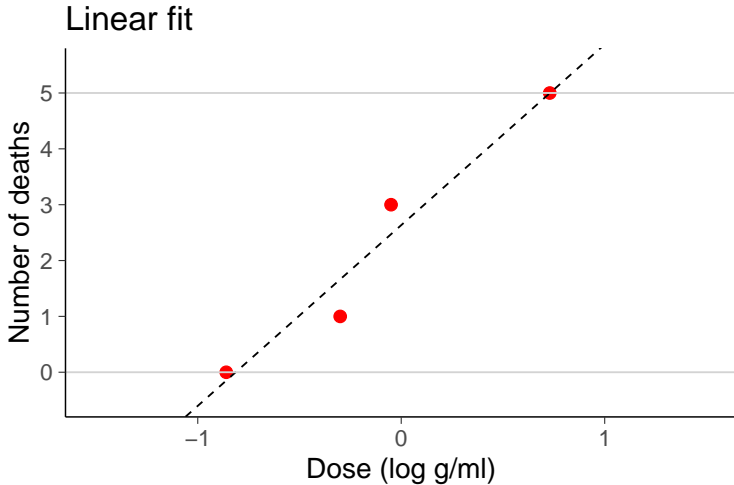
# Bioassay



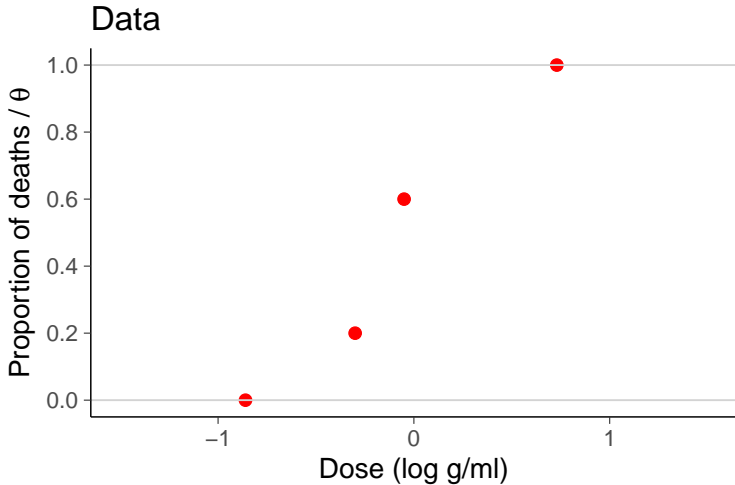
## Bioassay



## Bioassay

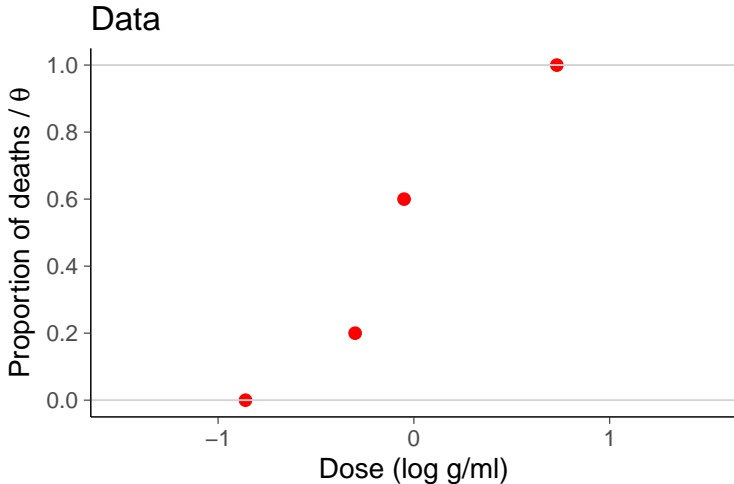


## Bioassay





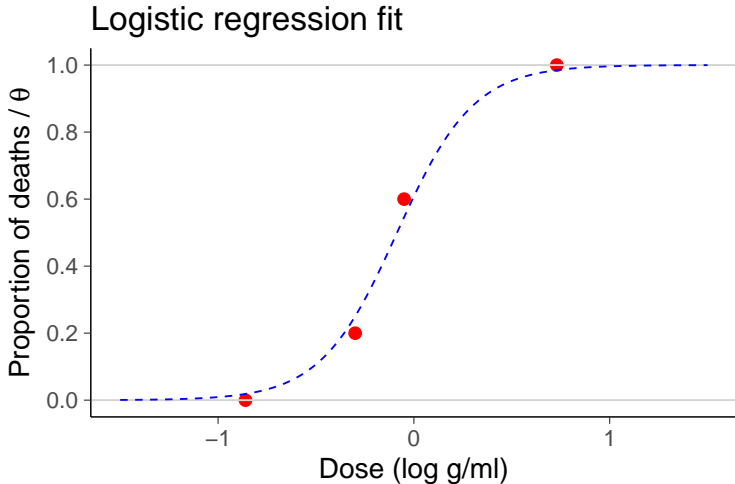
# Bioassay



Binomial model

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

## Bioassay



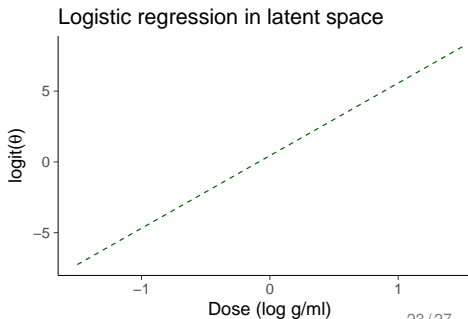
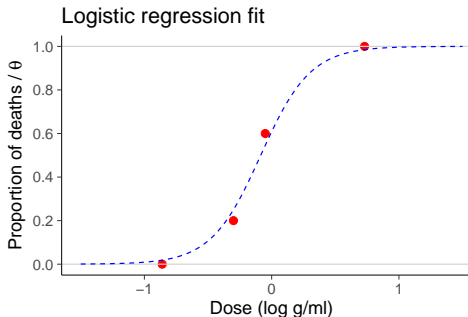
Binomial model

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

# Bioassay

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

$$\begin{aligned}\text{logit}(\theta_i) &= \log\left(\frac{\theta_i}{1 - \theta_i}\right) \\ &= \alpha + \beta \mathbf{x}_i\end{aligned}$$



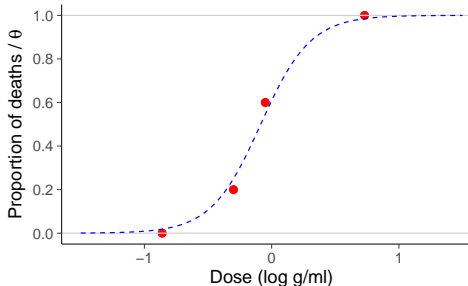
# Bioassay

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

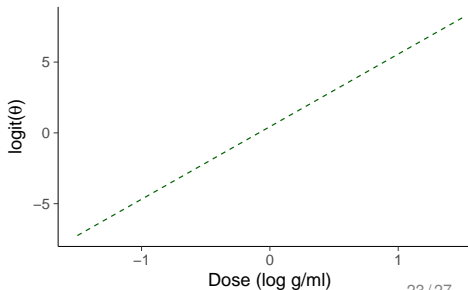
$$\begin{aligned}\text{logit}(\theta_i) &= \log\left(\frac{\theta_i}{1 - \theta_i}\right) \\ &= \alpha + \beta \mathbf{x}_i\end{aligned}$$

$$\theta = \frac{1}{1 + \exp(-(\alpha + \beta \mathbf{x}))}$$

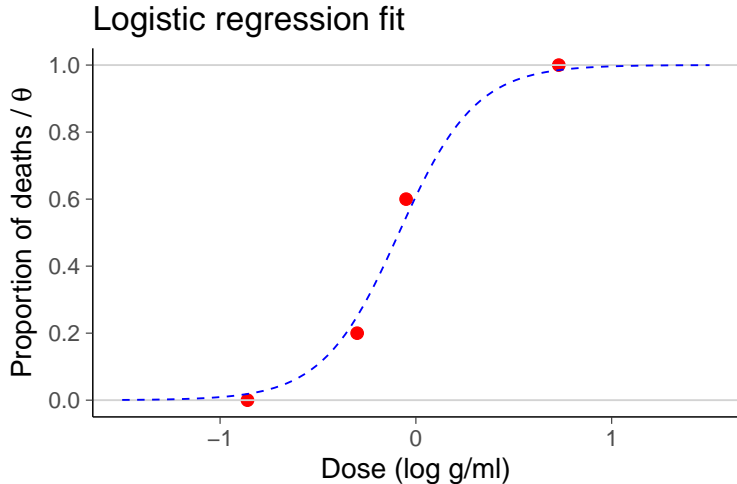
Logistic regression fit



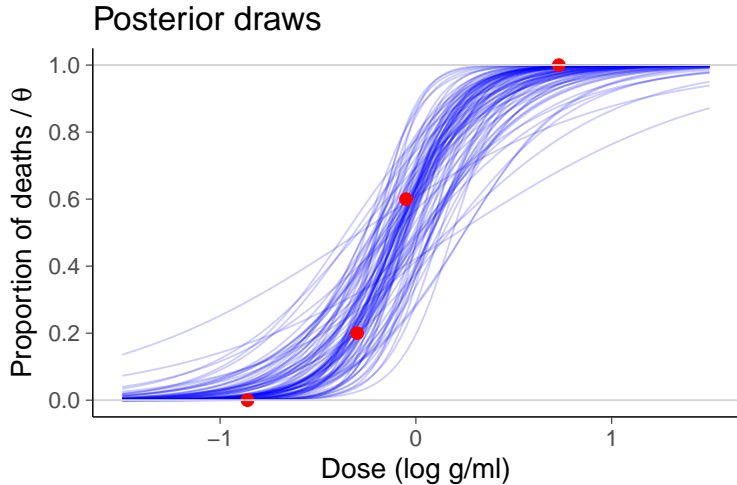
Logistic regression in latent space



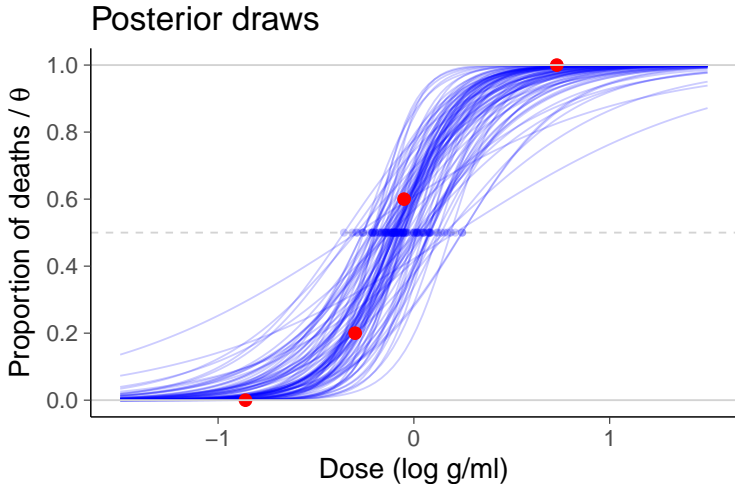
## Bioassay



## Bioassay

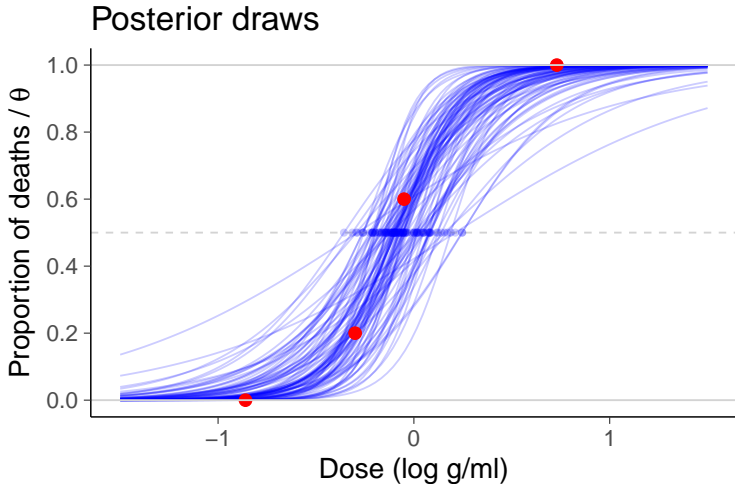


## Bioassay



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

## Bioassay

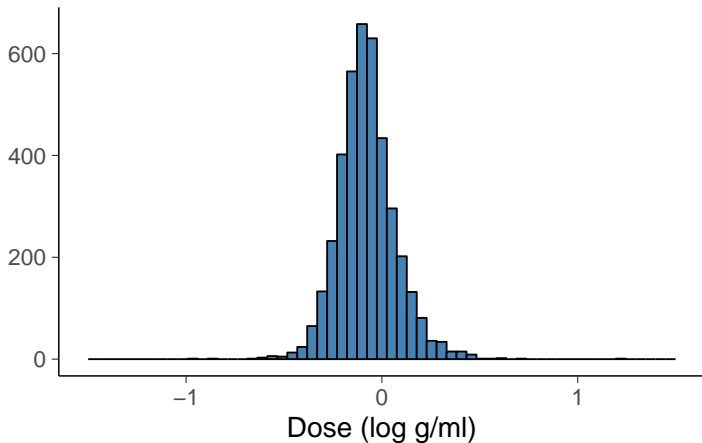


$$\text{LD50: } E\left(\frac{y}{n}\right) = \text{logit}^{-1}(\alpha + \beta x) = 0.5 \Rightarrow x_{\text{LD50}} = -\alpha/\beta$$



## Bioassay

### Bioassay LD50



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

## Bioassay posterior

### Binomial model

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

### Link function

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

# 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

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

# 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

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

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

# 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

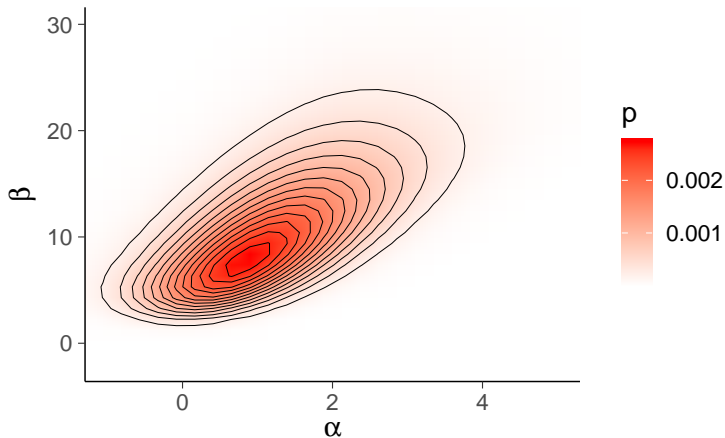
$$p(y_i \mid \alpha, \beta, n_i, x_i) \propto \theta_i^{y_i} [1 - \theta_i]^{n_i - y_i}$$

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

## Posterior (with uniform prior on $\alpha, \beta$ )

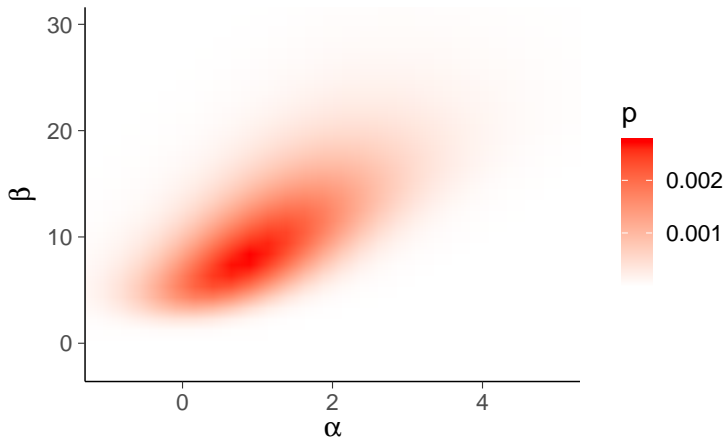
$$p(\alpha, \beta \mid y, n, x) \propto p(\alpha, \beta) \prod_{i=1}^n p(y_i \mid \alpha, \beta, n_i, x_i)$$

Posterior density evaluated in a grid



## Bioassay

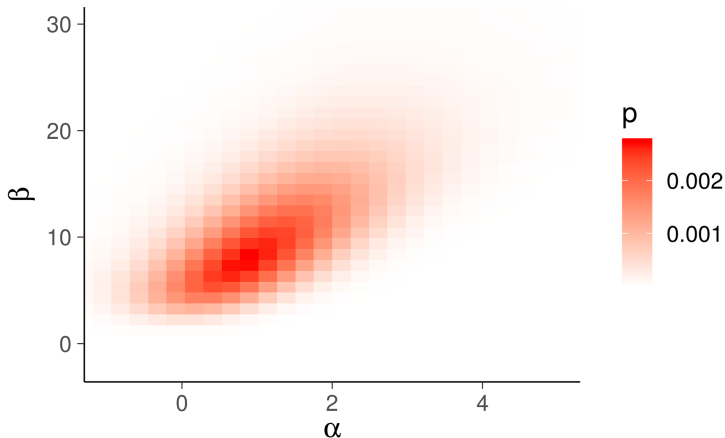
Posterior density evaluated in a grid



Density evaluated in grid, but plotted using interpolation

## Bioassay

Posterior density evaluated in a grid

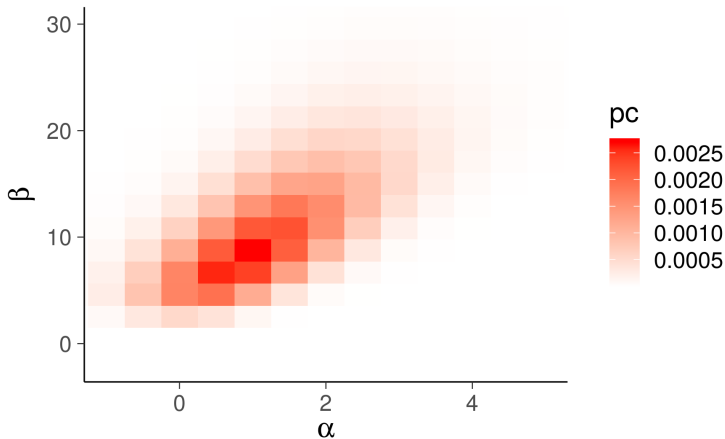


Density evaluated in grid, and plotted without interpolation



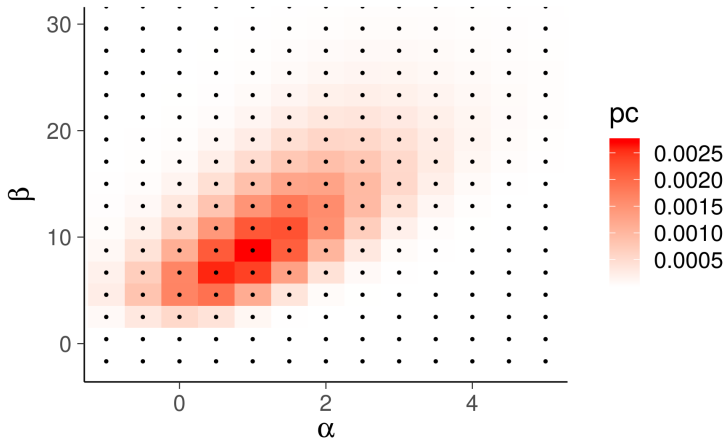
## Bioassay

Posterior density evaluated in a grid



Density evaluated in a coarser grid

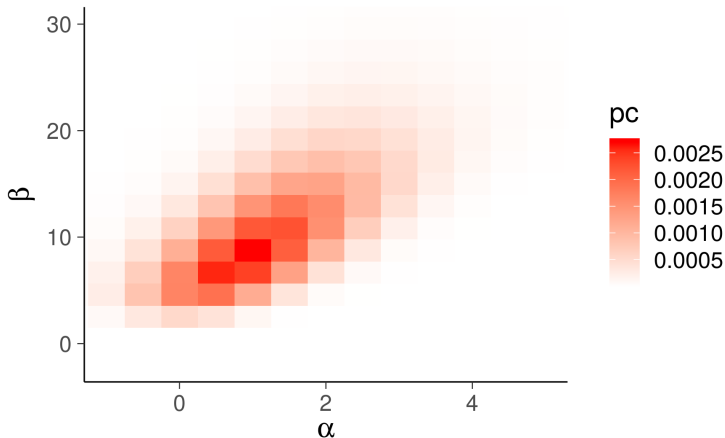
### Posterior density evaluated in a grid



- 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

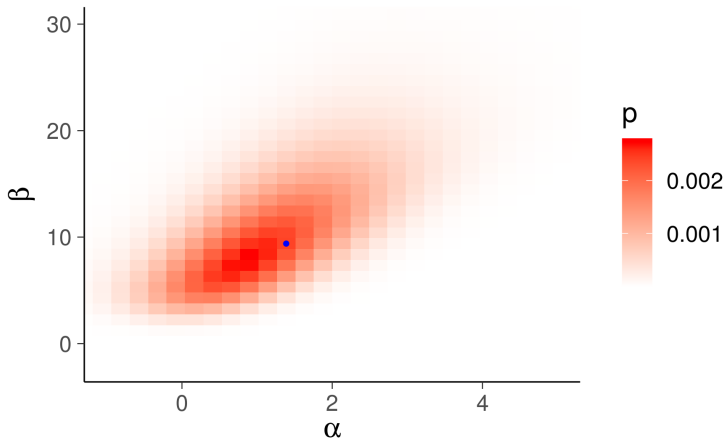
Posterior density evaluated in a grid



- 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

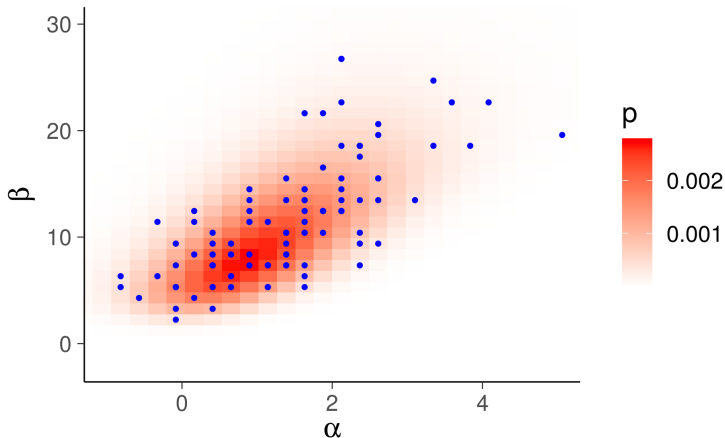
## Bioassay

### Posterior density and draws in a grid



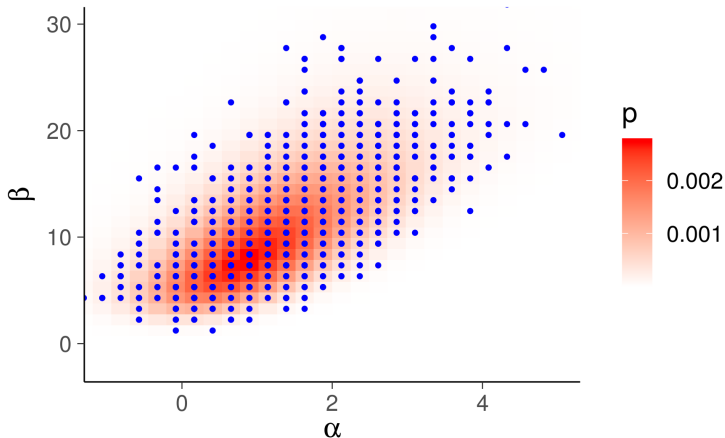
- Sample according to grid cell probabilities

### Posterior density and draws in a grid



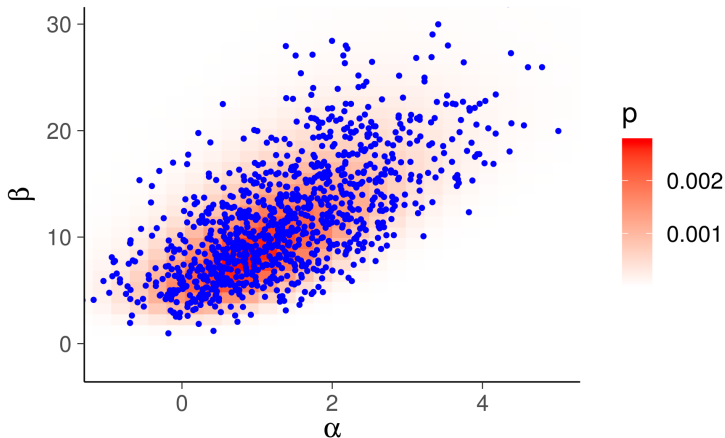
- Sample according to grid cell probabilities

### Posterior density and draws in a grid



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

### Posterior density in a grid and jittered draws



- Jitter can be added to improve visualization

## Grid sampling

- Draws can be used to estimate expectations, for example

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



## Grid sampling

- Draws can be used to estimate expectations, for example

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

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

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

where  $w_{\text{cell}}^{(t)}$  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

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

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

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

where  $w_{\text{cell}}^{(t)}$  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