Bioassay (Biotest)

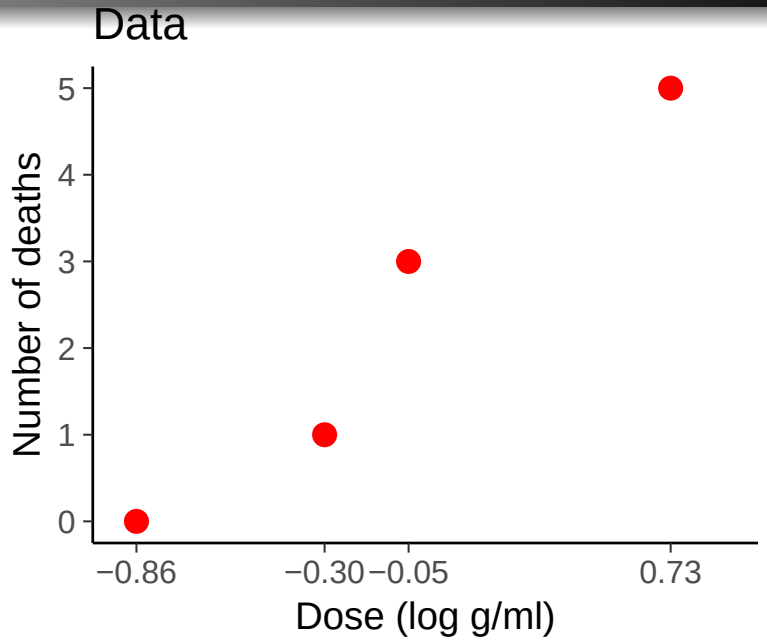
From Gelman:

> *Beyond the normal distribution, few multiparameter sampling models allow simple explicit calculation of posterior distributions. Data analysis for such models is possible using the computational methods described in Part III ...*
>
> *Here we present an example of a nonconjugate model for a bioassay experiment ...*
> *The model is a two-parameter example from the broad class of generalized linear models ...*
> *We use a particularly simple simulation approach, approximating the posterior distribution by a discrete distribution supported on a two-dimensional grid of points, that provides sufficiently accurate inferences ...*

Grid-based methods are, e.g., illustrated in Chapter 10 (Part III), section **Numerical integration** (in particular, within deterministic numerical integration methods)

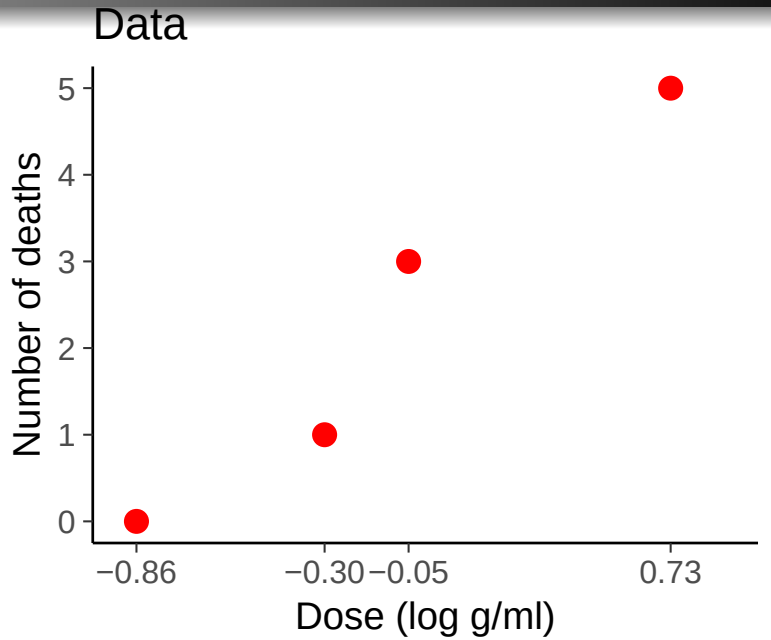
Bioassay experiment (Acute toxicity test)

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 experiment (Acute toxicity test)

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

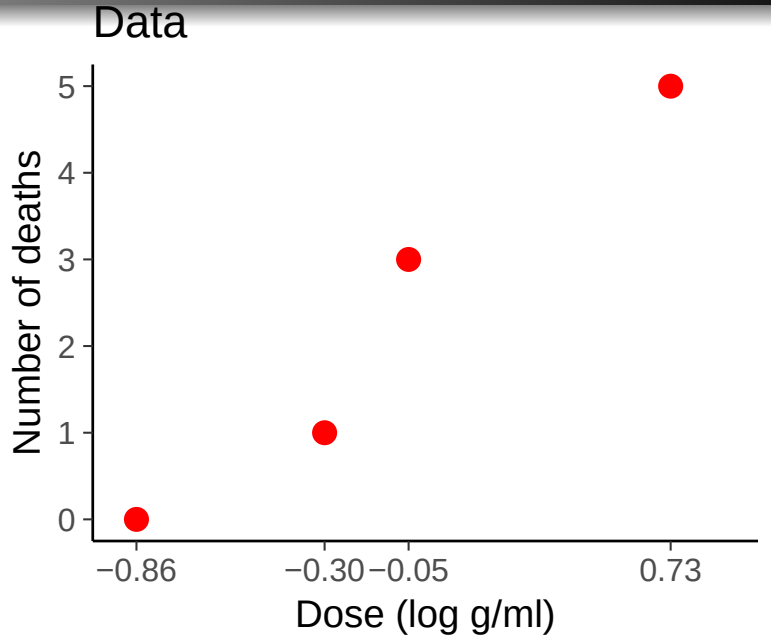


Find out lethal dose 50% (LD50), i.e. the dose level at which the probability of death is 50%.

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

Bioassay experiment (Acute toxicity test)

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), i.e. the dose level at which the probability of death is 50%.

- 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

Analisi di un esperimento biotest

- (x_i, n_i, y_i), $i = 1, \ldots, k$
  - x_i, i-esimo di k livelli di dose (spesso misurati su scala logaritmica)
  - n_i animali trattati con la dose i
  - y_i, numero di morti
- θ_i, probabilità di morte data la dose i

Analisi di un esperimento biotest

- (x_i, n_i, y_i), $i = 1, \ldots, k$
  - x_i, i-esimo di k livelli di dose (spesso misurati su scala logaritmica)
  - n_i animali trattati con la dose i
  - y_i, numero di morti
- θ_i, probabilità di morte data la dose i
- Assumiamo y_i scambiabili **entro ciascun gruppo** i, i.e. indipendenti con uguale probabilità, il che implica

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

Analisi di un esperimento biotest

- (x_i, n_i, y_i), $i = 1, \ldots, k$
  - x_i, i-esimo di k livelli di dose (spesso misurati su scala logaritmica)
  - n_i animali trattati con la dose i
  - y_i, numero di morti
- θ_i, probabilità di morte data la dose i
- Assumiamo y_i scambiabili **entro ciascun gruppo** i, i.e. indipendenti con uguale probabilità, il che implica

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

  In quale situazione il modello Binomiale e l'indipendenza non sarebbero appropriati?

Analisi di un esperimento biotest

- (x_i, n_i, y_i), $i = 1, \ldots, k$
  - x_i, i-esimo di k livelli di dose (spesso misurati su scala logaritmica)
  - n_i animali trattati con la dose i
  - y_i, numero di morti
- θ_i, probabilità di morte data la dose i
- Assumiamo y_i scambiabili **entro ciascun gruppo** i, i.e. indipendenti con uguale probabilità, il che implica

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

  In quale situazione il modello Binomiale e l'indipendenza non sarebbero appropriati?
- Infine, i risultati y_i nei quattro gruppi sono supposti indipendenti dati i parametri $\theta_1, \ldots, \theta_4$

Analisi di un esperimento biotest

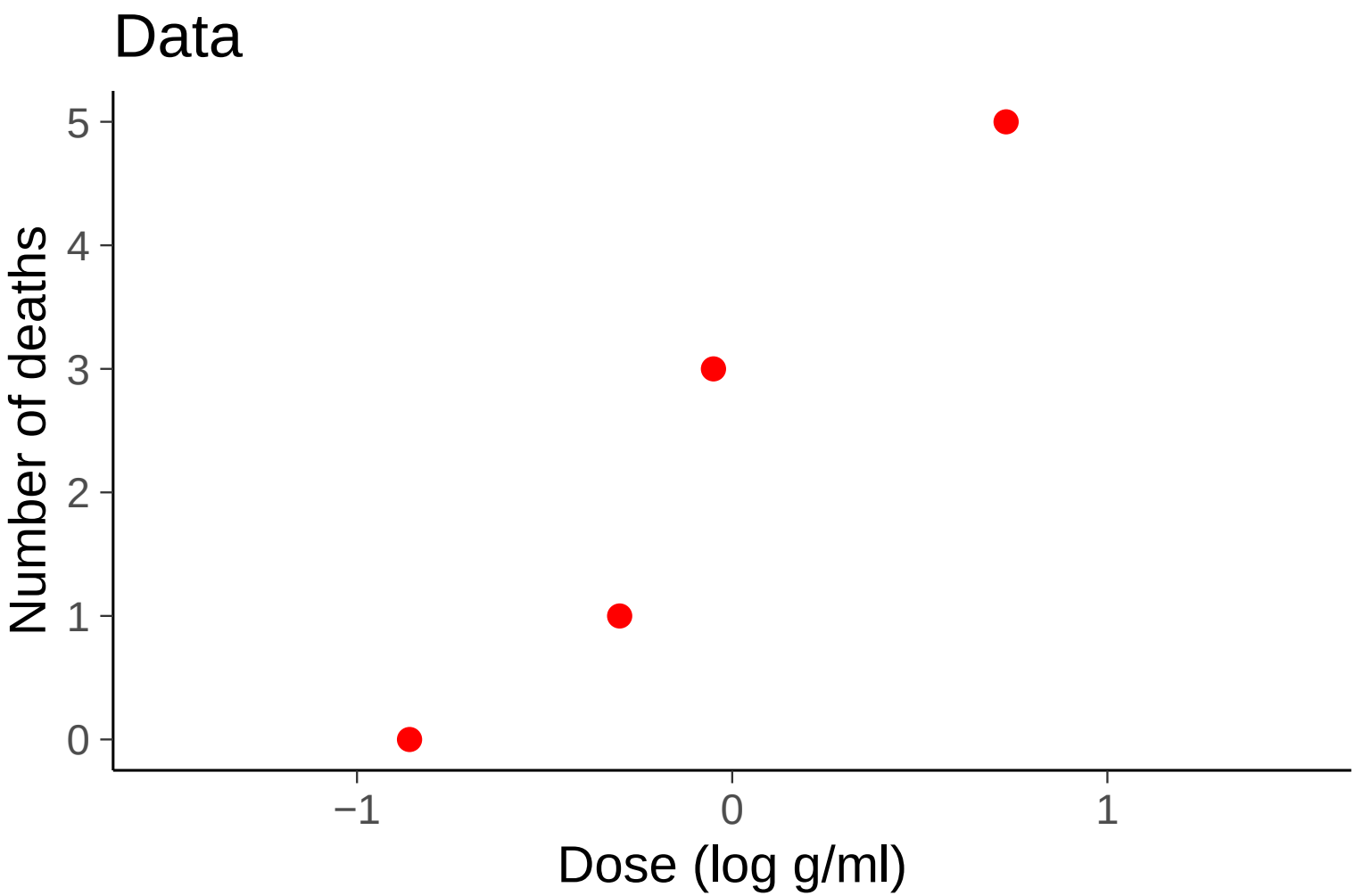
- Un'analisi più semplice tratterebbe i quattro parametri θ_i come scambiabili nella loro distribuzione a priori, e se si usasse una densità non informativa, e.g., $p(\theta_1, \ldots, \theta_4) \propto 1$, i parametri avrebbero distribuzioni a posteriori Beta indipendenti.
- La scambiabilità però tralascerebbe l'informazione data dalla conoscenza del livello di dose x_i per ogni gruppo i, mentre ci si aspetterebbe che la probabilità di morte vari sistematicamente in funzione della dose.

Analisi di un esperimento biotest

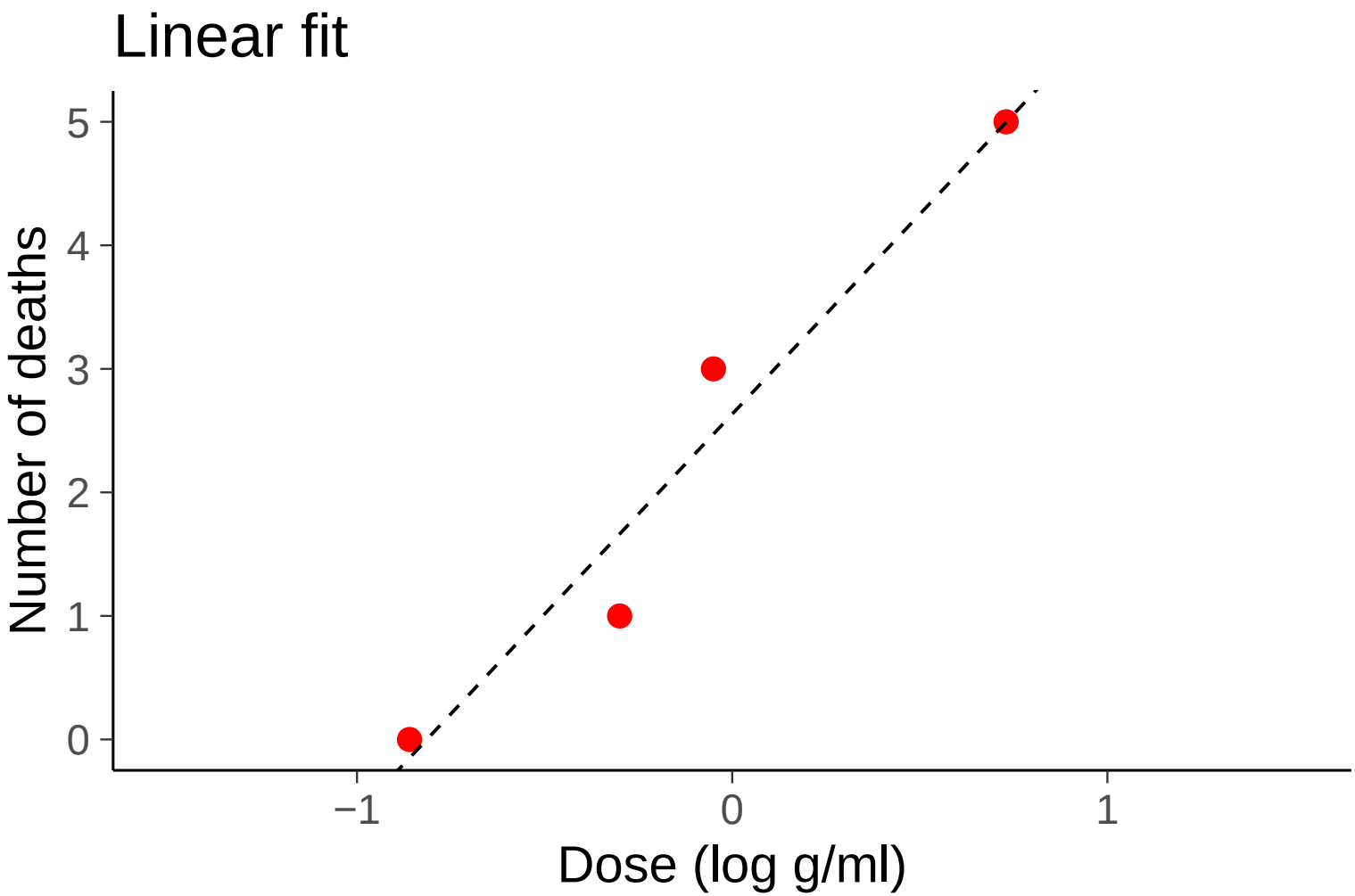
- Un'analisi più semplice tratterebbe i quattro parametri θ_i come scambiabili nella loro distribuzione a priori, e se si usasse una densità non informativa, e.g., $p(\theta_1, \ldots, \theta_4) \propto 1$, i parametri avrebbero distribuzioni a posteriori Beta indipendenti.
- La scambiabilità però tralascerebbe l'informazione data dalla conoscenza del livello di dose x_i per ogni gruppo i, mentre ci si aspetterebbe che la probabilità di morte vari sistematicamente in funzione della dose.

- Il modello più semplice della relazione dose-risposta, ovvero la relazione tra θ_i e x_i, è lineare: $\theta_i = \alpha + \beta x_i$. Purtroppo questo modello ha il difetto che a dosi basse o alte, x_i si avvicina a $\pm\infty$ (ricordiamo che la dose è misurata su scala logaritmica), mentre θ_i, essendo una probabilità, è vincolata a essere compresa tra 0 e 1.

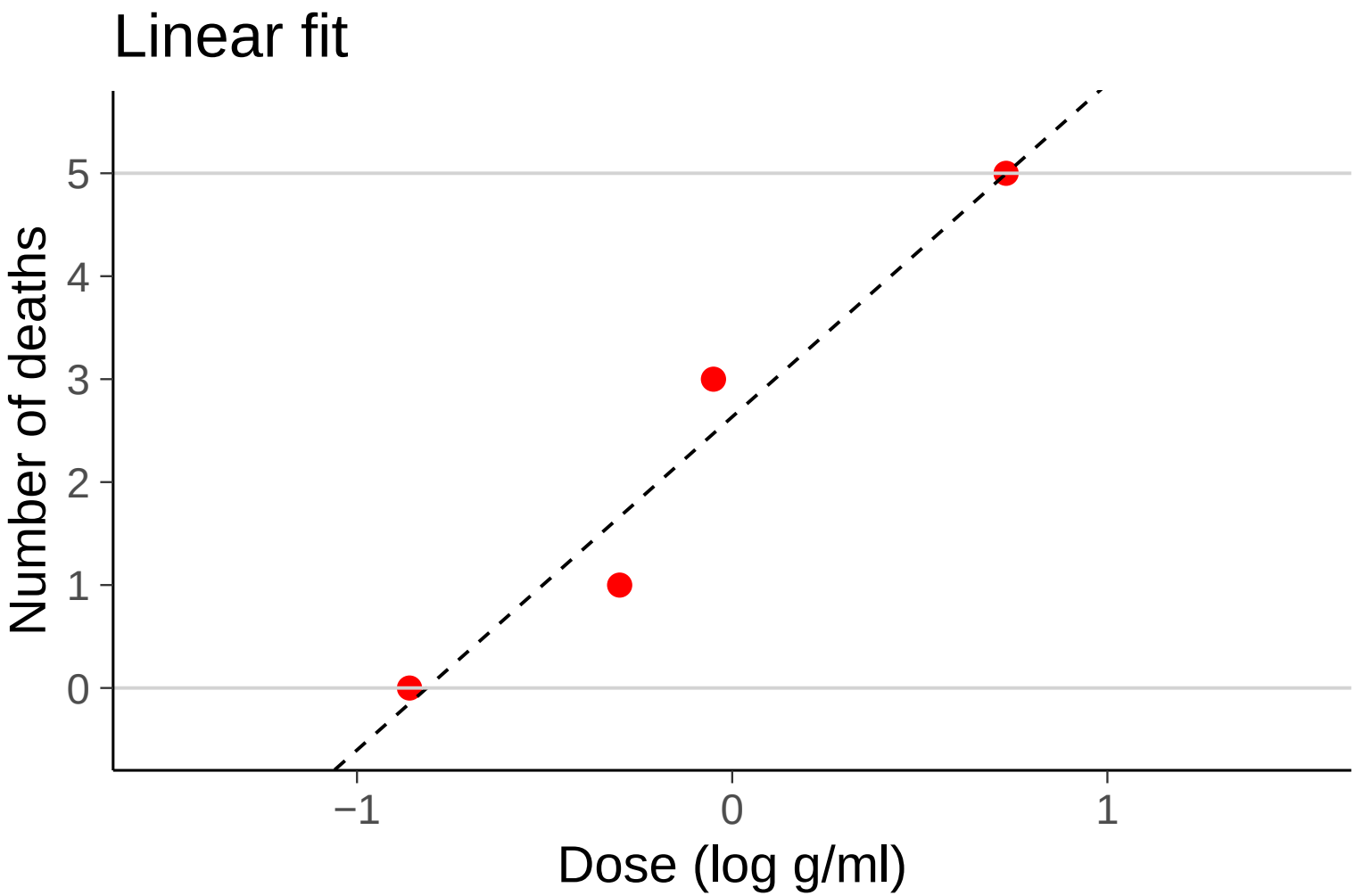
Bioassay



Bioassay



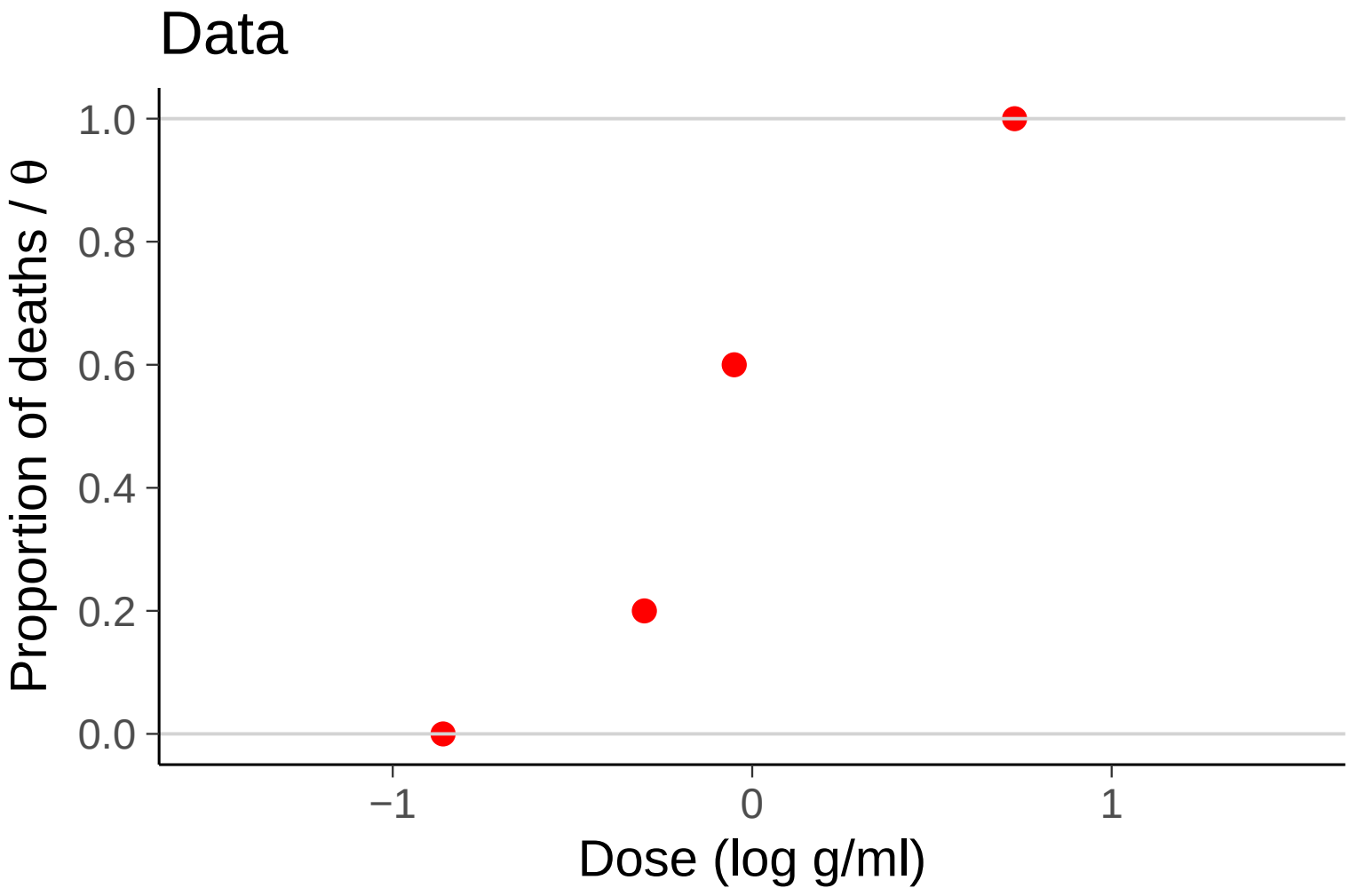
Bioassay

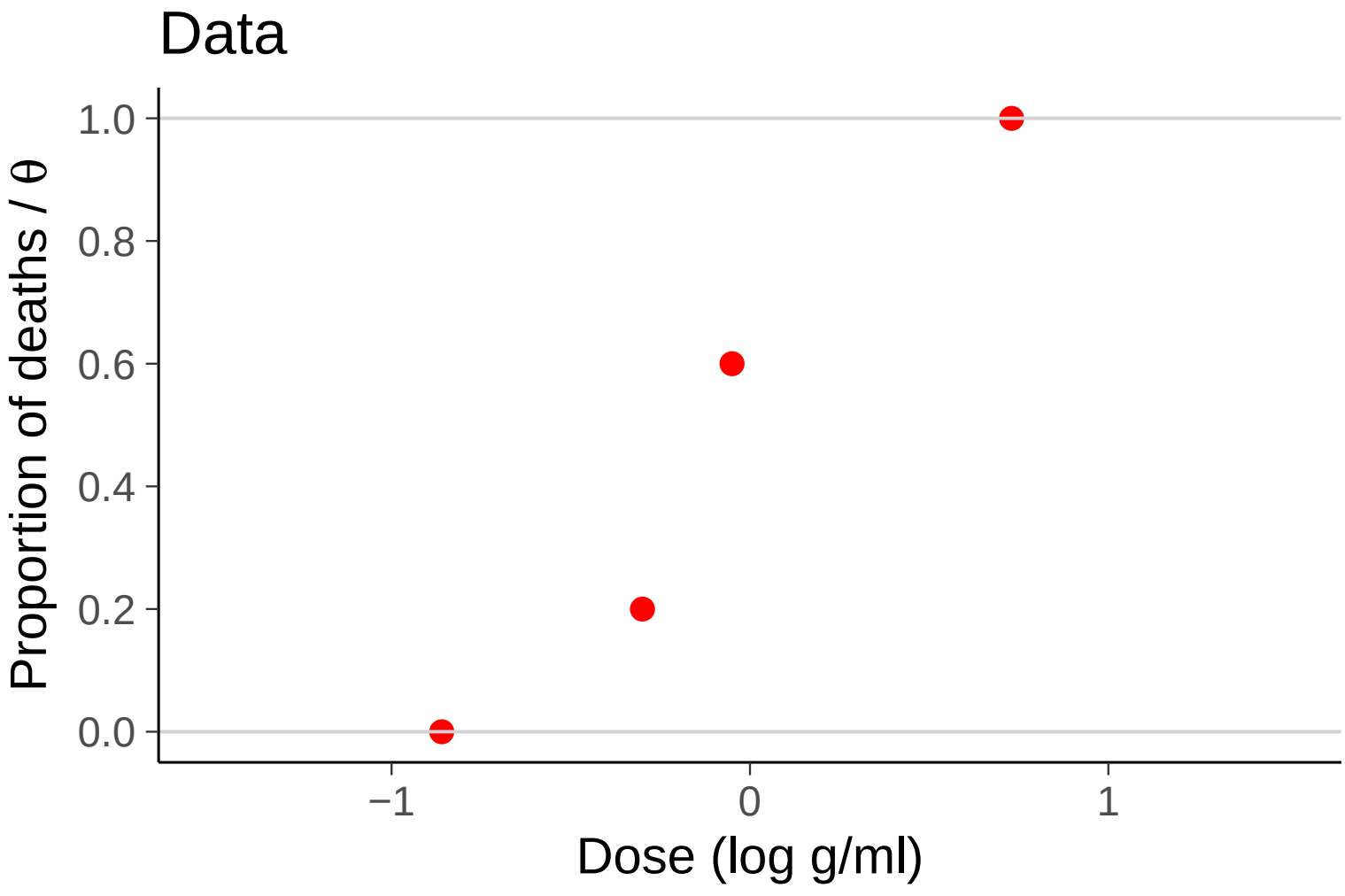


- La soluzione standard è usare una trasformazione delle θ, come il logit, nella relazione dose-risposta:

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

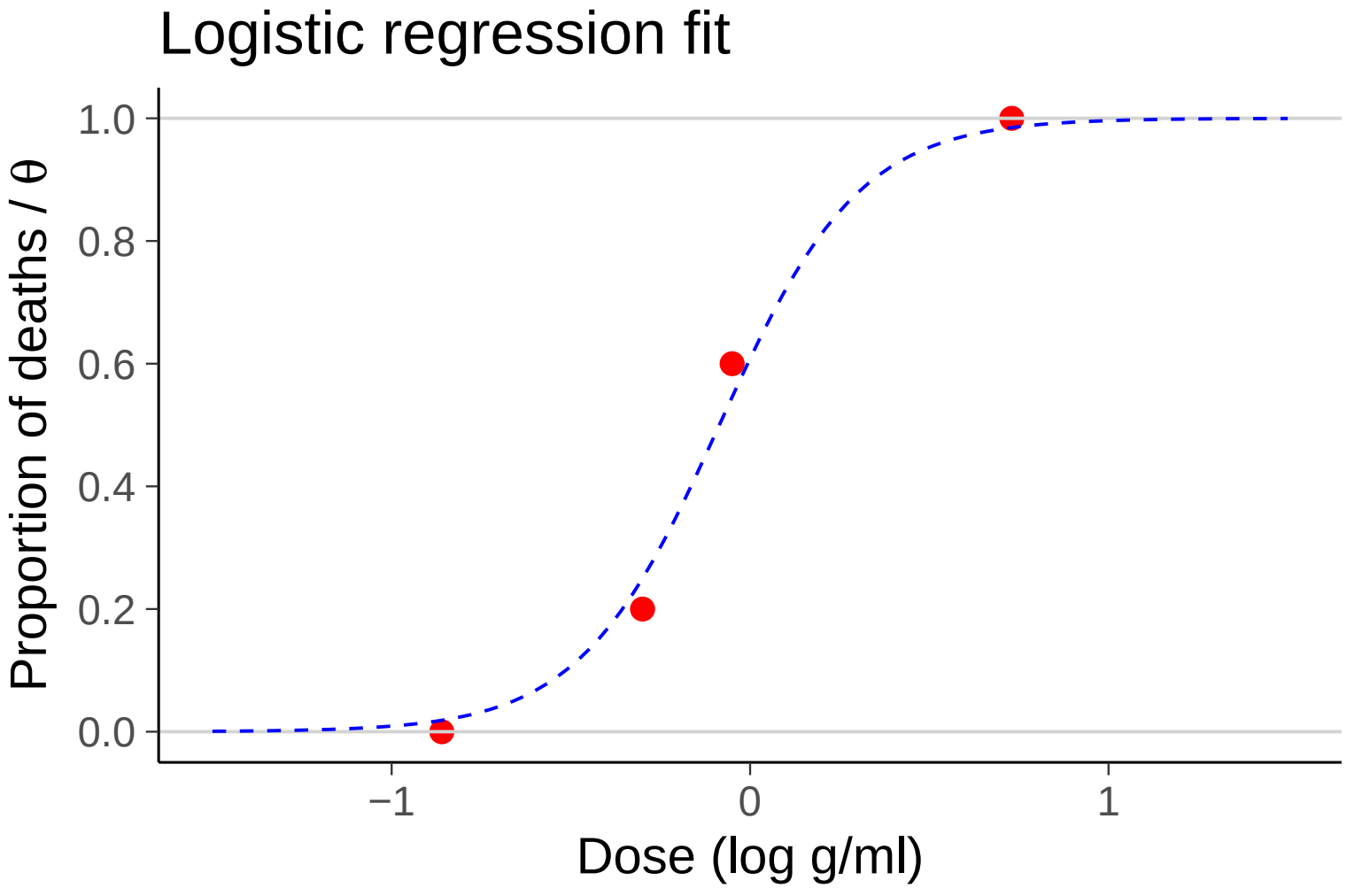
come nella regressione logistica


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



Binomial model

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



Binomial model

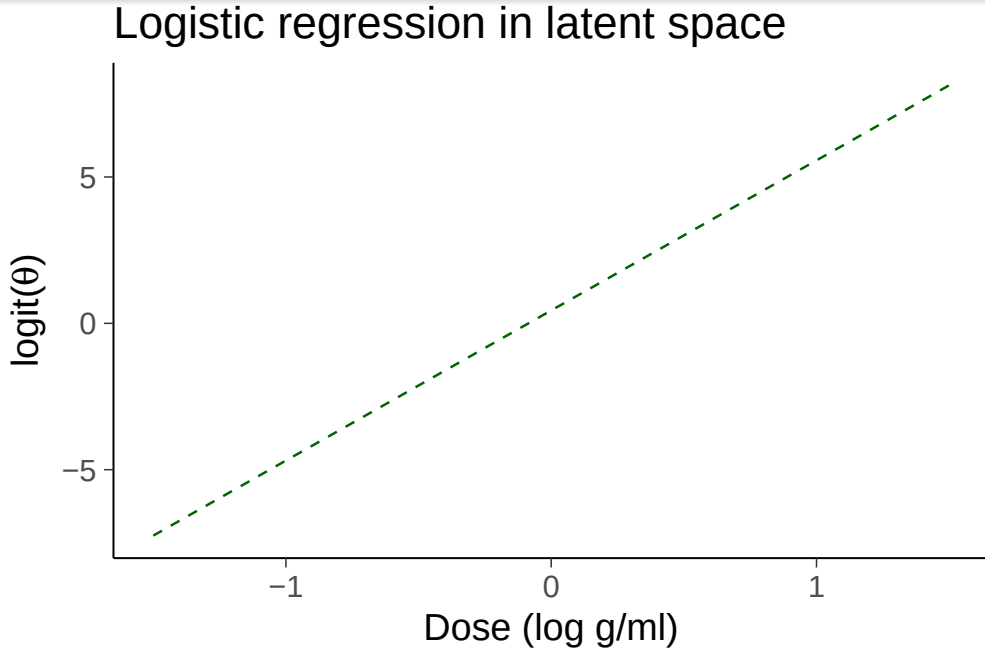
$$y_i \mid \theta_i \sim \text{Bin}(\theta_i, n_i), \quad \text{logit}(\theta_i) = \log\left(\frac{\theta_i}{1-\theta_i}\right) = \alpha + \beta x_i$$

Bioassay

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

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

$$= \alpha + \beta x_i$$

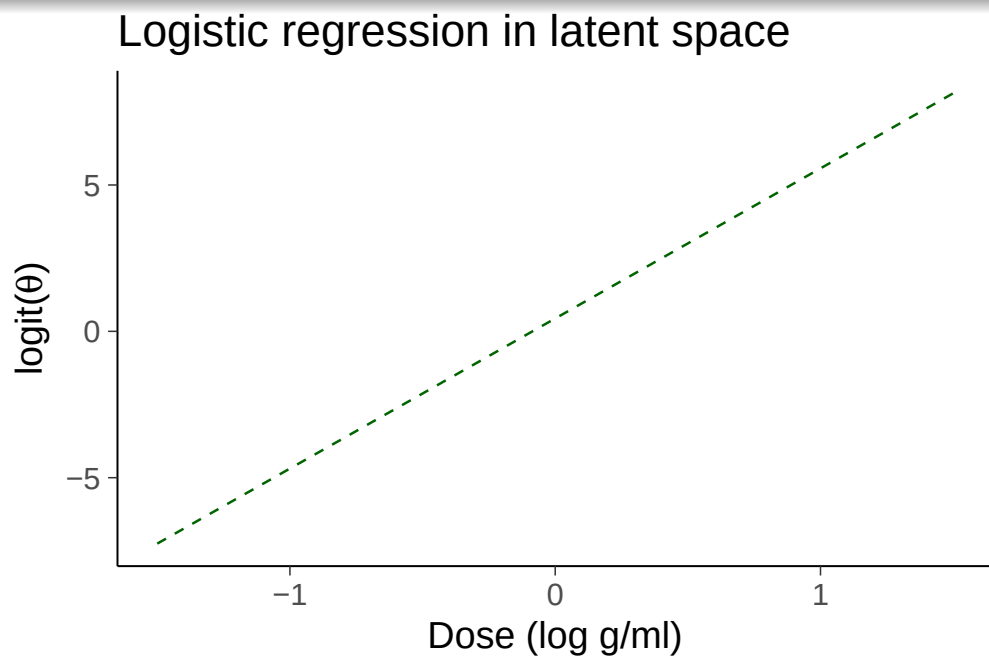


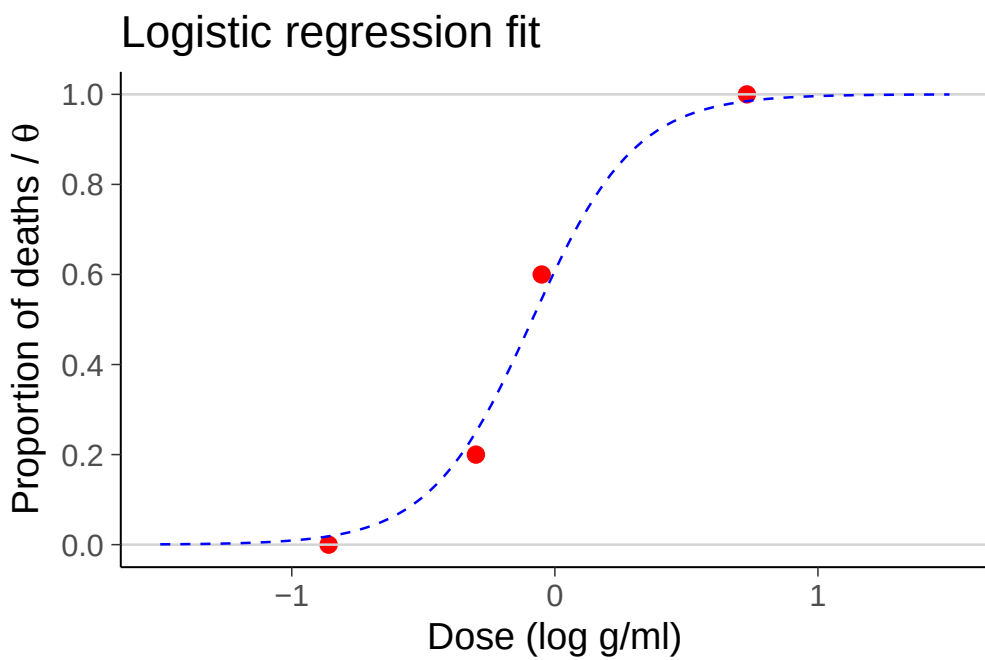
Bioassay

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

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

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





Bioassay: inferenza a posteriori su α e β

- Verosimiglianza, in termini dei parametri α e β,

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

  Nota: In questa analisi consideriamo n_i e x_i come fissati, quindi occultiamo il condizionamento a (n_i, x_i) nel prosieguo.

- A priori, indipendente e localmente uniforme per i parametri α e β,

$$p(\alpha,\beta) \propto \text{costante}$$

- A posteriori

$$p(\alpha,\beta|y) \propto p(\alpha,\beta)\prod_i p(y_i|\alpha,\beta)$$

Bioassay: a priori su α e β

From Gelman,

> *In practice, we might use*
>
> - *a uniform prior distribution if we really have no prior knowledge about the parameters, or*
> - *if we want to present a simple analysis of this experiment alone.*
>
> *If the analysis using the noninformative prior distribution is insufficiently precise, we may consider using other sources of substantive information (for example, from other bioassay experiments) to construct an informative prior distribution.*

Bioassay: calcolo della a posteriori su griglia di valori

Una stima dei parametri per impostare la griglia di valori su cui esplorare la densità a posteriori

- Stima di ML nella regressione logistica
  - $(\hat{\alpha}, \hat{\beta}) = (0.8, 7.7)$
  - $(sd(\hat{\alpha}), sd(\hat{\beta})) = (1.0, 4.9)$

La griglia di valori (α, β) su cui andiamo a calcolare la congiunta a posteriori è

- $(\alpha, \beta) \in [-5, 10] \times [10, 40]$

Possiamo quindi calcolare la densità a posteriori non normalizzata sulla griglia e disegnare le curve di livello

Bioassay: settare la griglia

From Gelman,

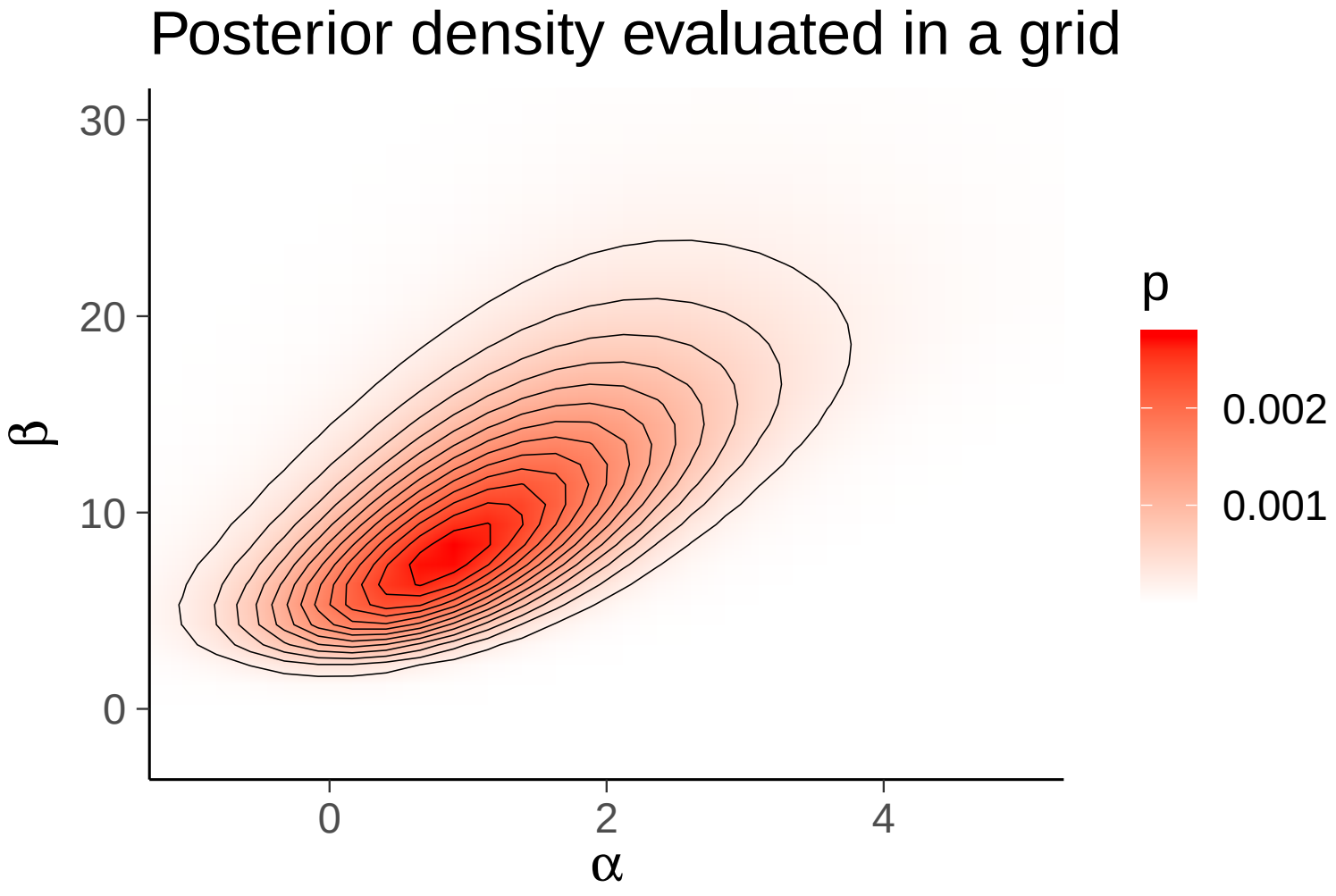
> *There can be difficulty finding the correct location and scale for the grid points.*
>
> - *A grid that is defined on too small an area may miss important features of the posterior distribution that fall outside the grid.*
> - *A grid defined on a large area with wide intervals between points can miss important features that fall between the grid points.*

Bioassay: difficoltà nel calcolo della a posteriori

From Gelman,

> *It is also important to avoid overflow and underflow operations when computing the posterior distribution.*
>
> - *It is usually a good idea to compute the logarithm of the unnormalized posterior distribution and*
> - *subtract off the maximum value before exponentiating.*
>
> *This creates an unnormalized discrete approximation with maximum value 1, which can then be normalized (by setting the total probability in the grid to 1).*

Bioassay: contour plot of the joint posterior density



Bioassay: campionamento dalla congiunta a posteriori

- Dopo aver calcolato la densità a posteriori non normalizzata su una griglia di valori che coprono il range effettivo di (α, β),
- possiamo normalizzare approssimando la distribuzione attraverso una funzione a gradini sulla griglia e imponendo che la probabilità totale sulla griglia sia 1.

Bioassay: campionamento dalla congiunta a posteriori

- Dopo aver calcolato la densità a posteriori non normalizzata su una griglia di valori che coprono il range effettivo di (α, β),
- possiamo normalizzare approssimando la distribuzione attraverso una funzione a gradini sulla griglia e imponendo che la probabilità totale sulla griglia sia 1.
- Campioniamo 1000 estrazioni casuali (α^*, β^*) dalla distribuzione a posteriori utilizzando la seguente procedura.
  - Calcoliamo la distribuzione marginale a posteriori di α (sommando numericamente su β) nella distribuzione discreta calcolata sulla griglia
  - Per $s = 1, \ldots, 1000$
    (a) Estrai α^* dalla $p(\alpha|y)$ calcolata discretamente attraverso (una versione discreta del) metodo *inverse cdf*
    (b) Estrai β^* dalla $p(\beta|\alpha, y)$ dato il valore appena estratto di α usando il metodo *inverse cdf*
    (c) Per ogni valore di α e β campionati aggiungi un *jitter* (ciò conferisce alle estrazioni una distr continua)

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

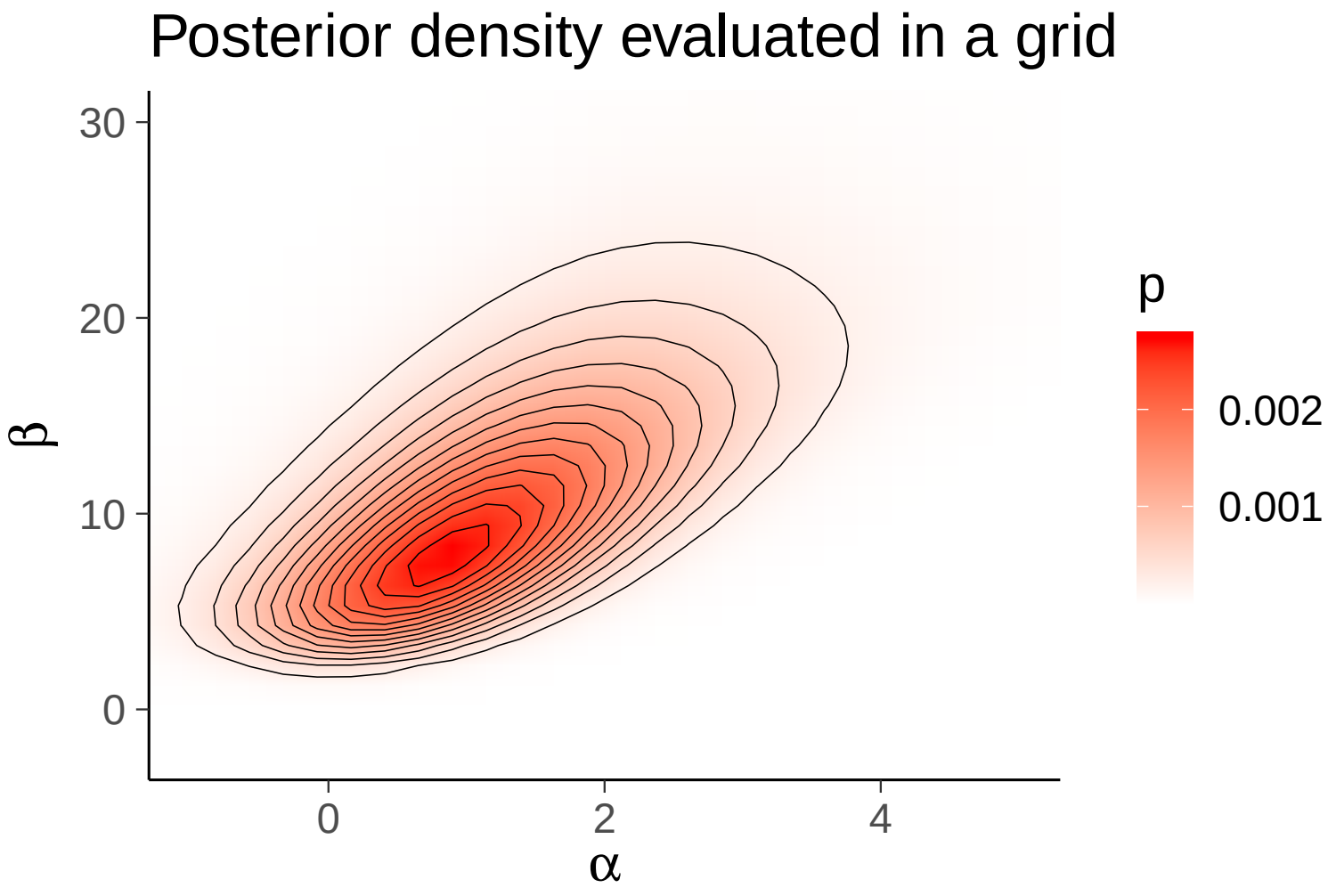
$$p(y_i \mid \alpha, \beta, n_i, x_i) \propto \theta_i^{y_i}[1 - \theta_i]^{n_i - y_i}$$

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

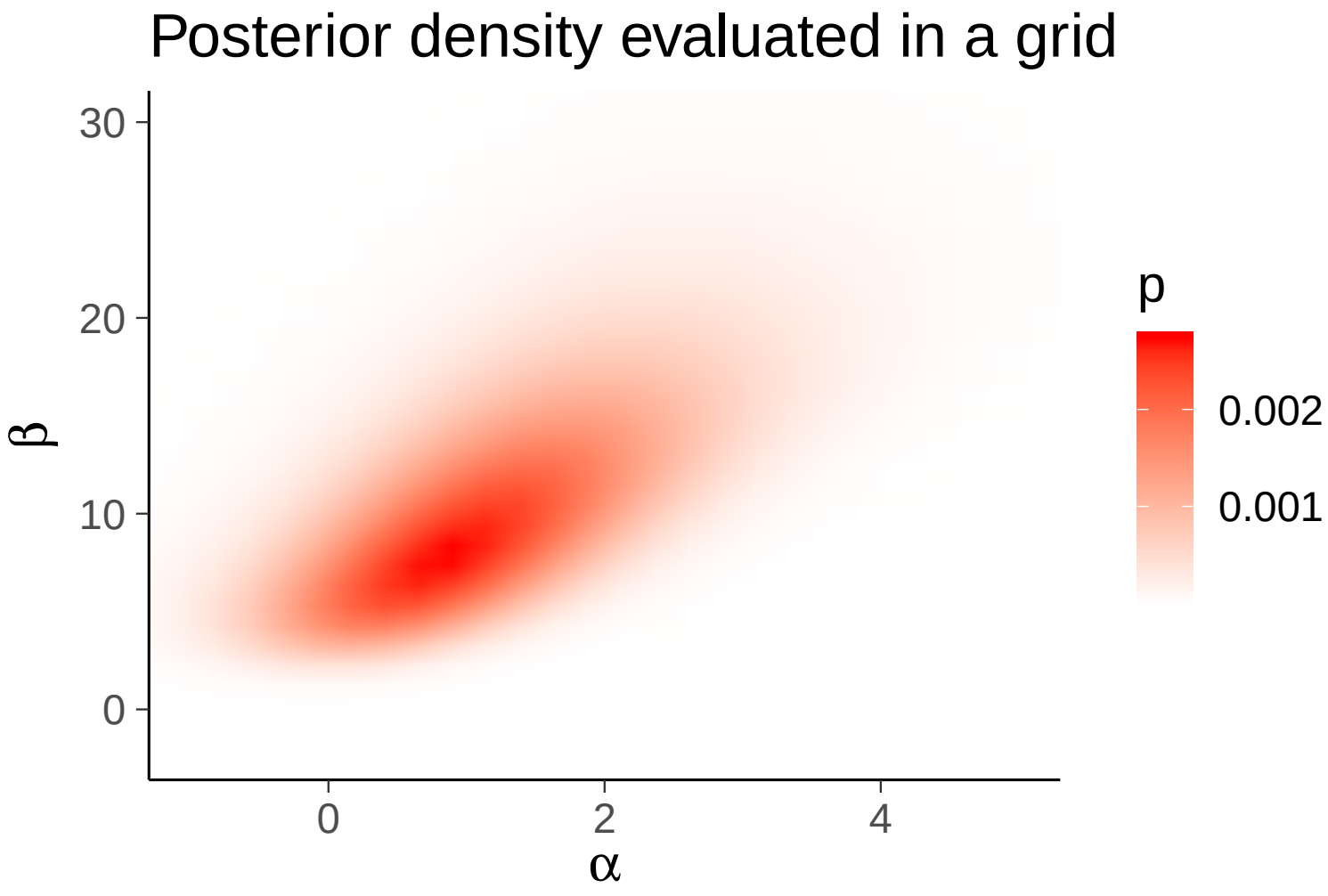
Posterior (with uniform prior on α, β)

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

Bioassay

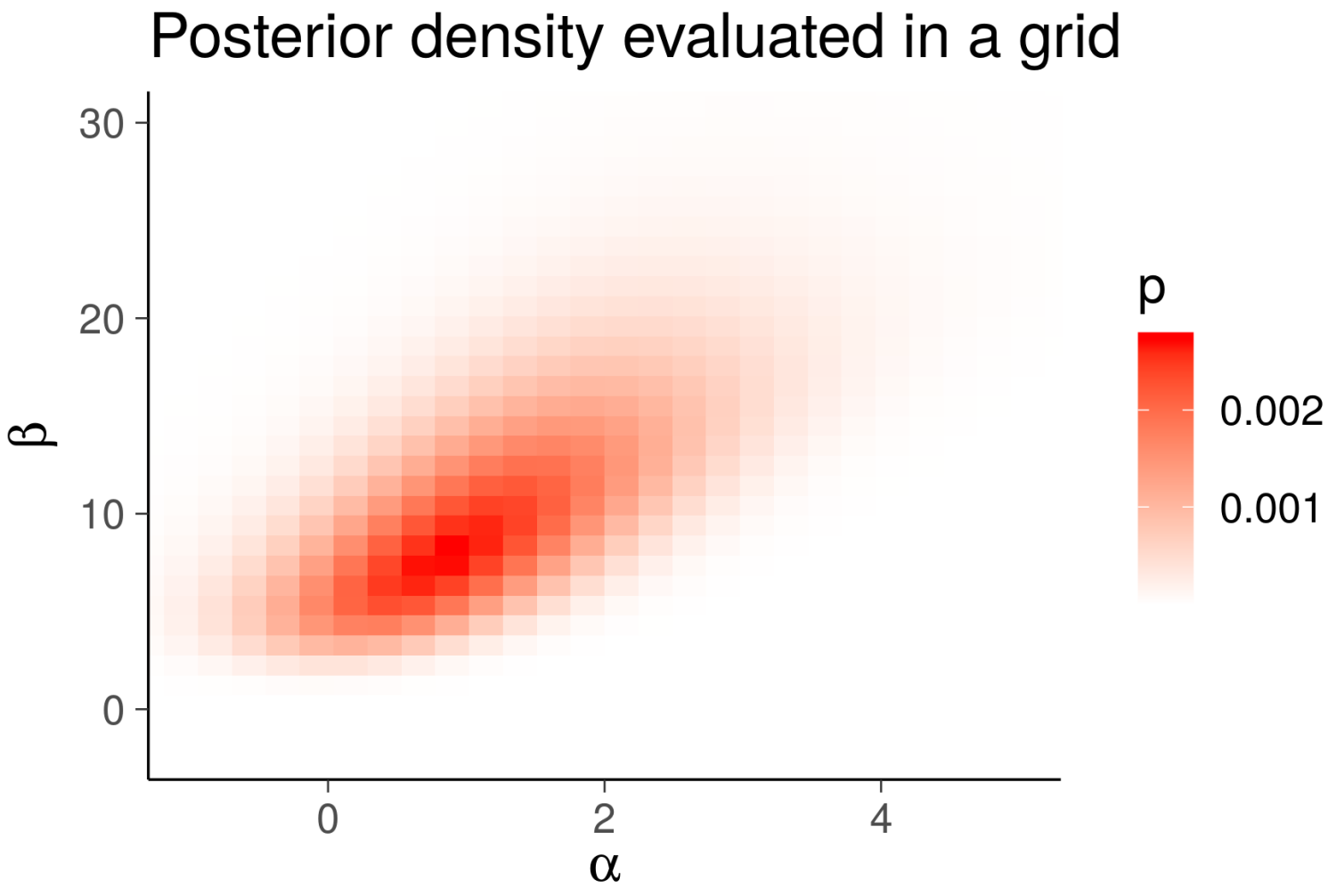

Posterior density evaluated in a grid
30
20
10
0
β
0
2
4
α
p
0.002
0.001

Bioassay



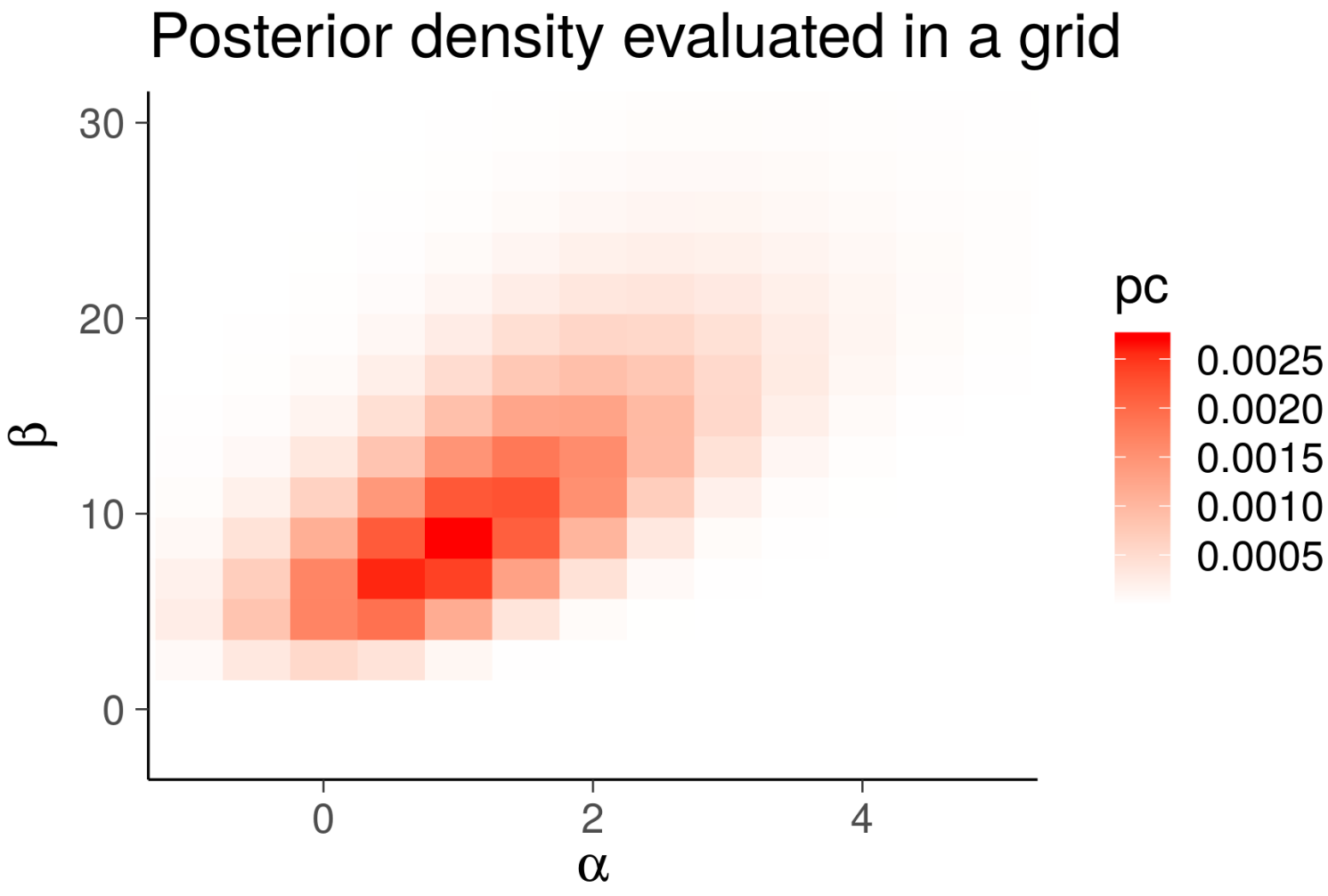
Density evaluated in grid, but plotted using interpolation

Bioassay



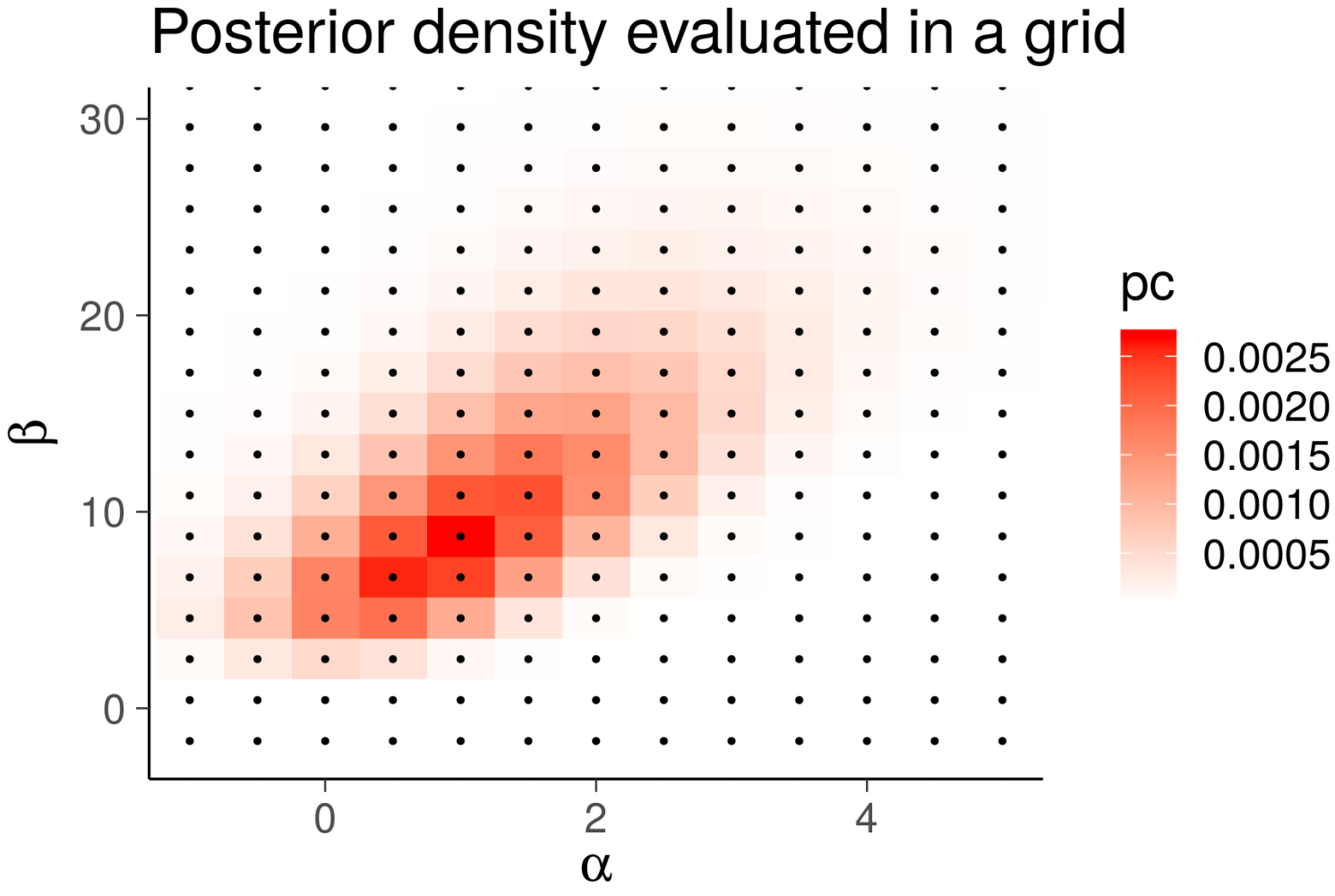
Density evaluated in grid, and plotted without interpolation

Bioassay



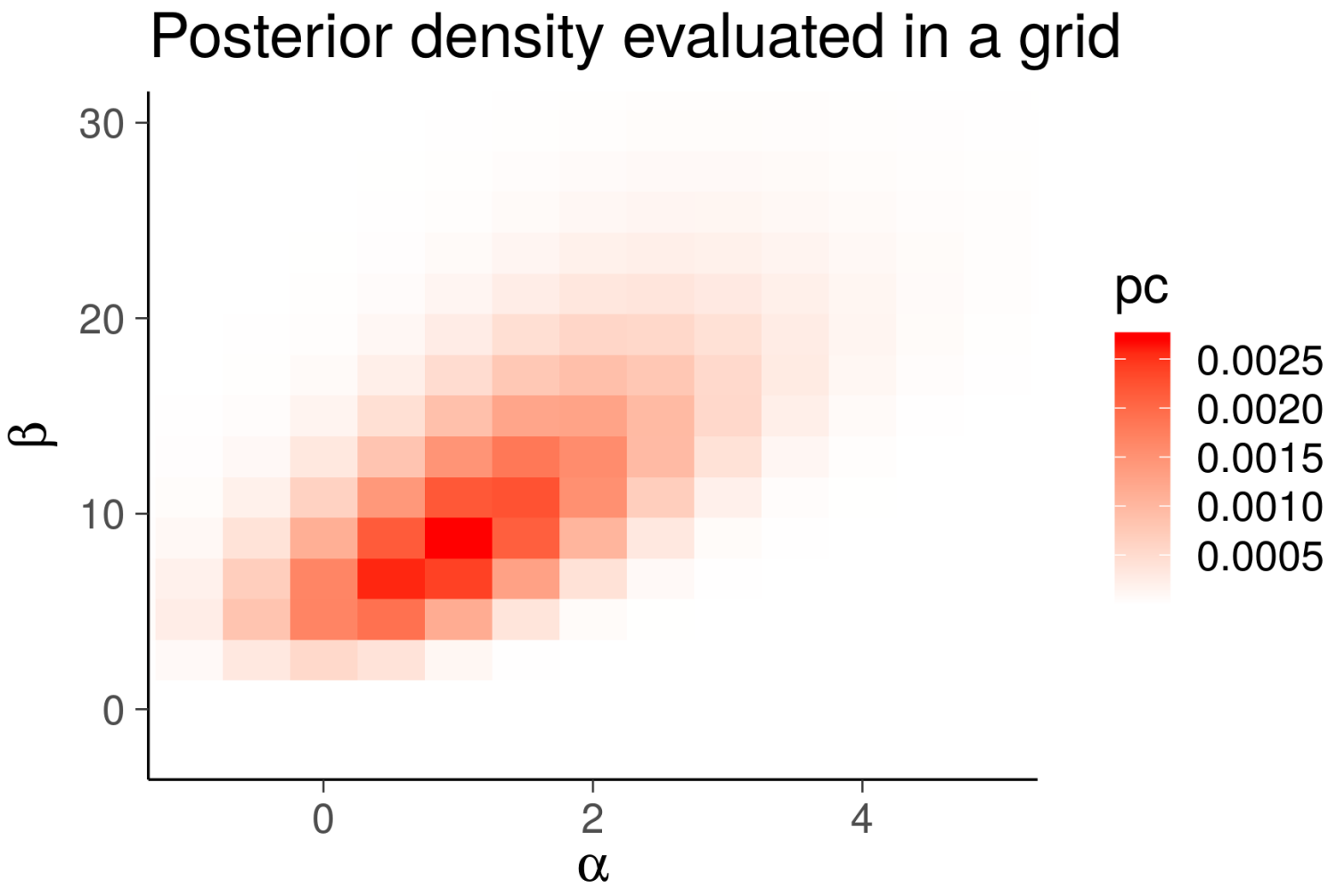
Density evaluated in a coarser grid

Bioassay



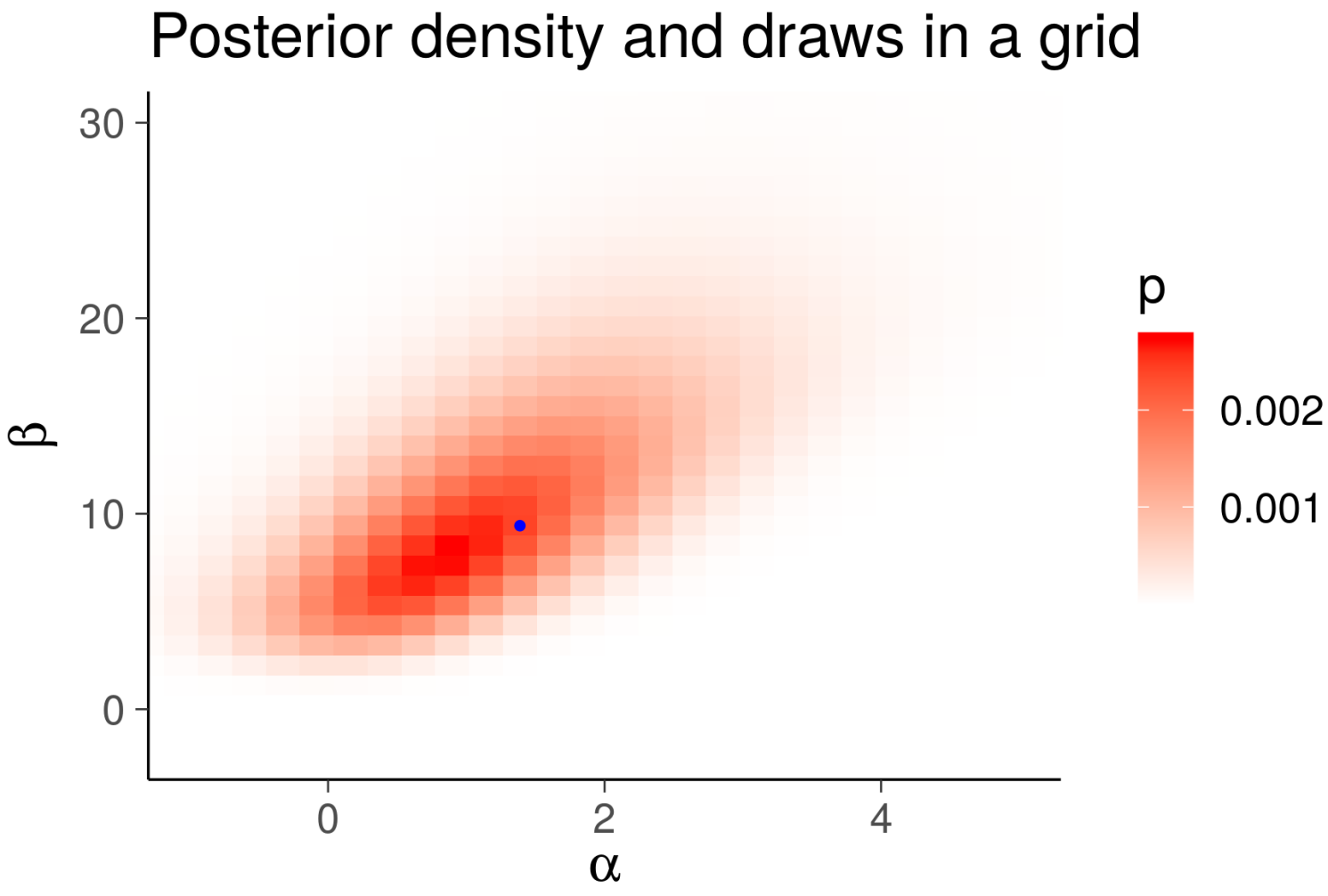
- Approximate the density as piecewise constant function
- Evaluate density in a grid over some finite region
- Density times cell area gives probability mass in each cell

Bioassay



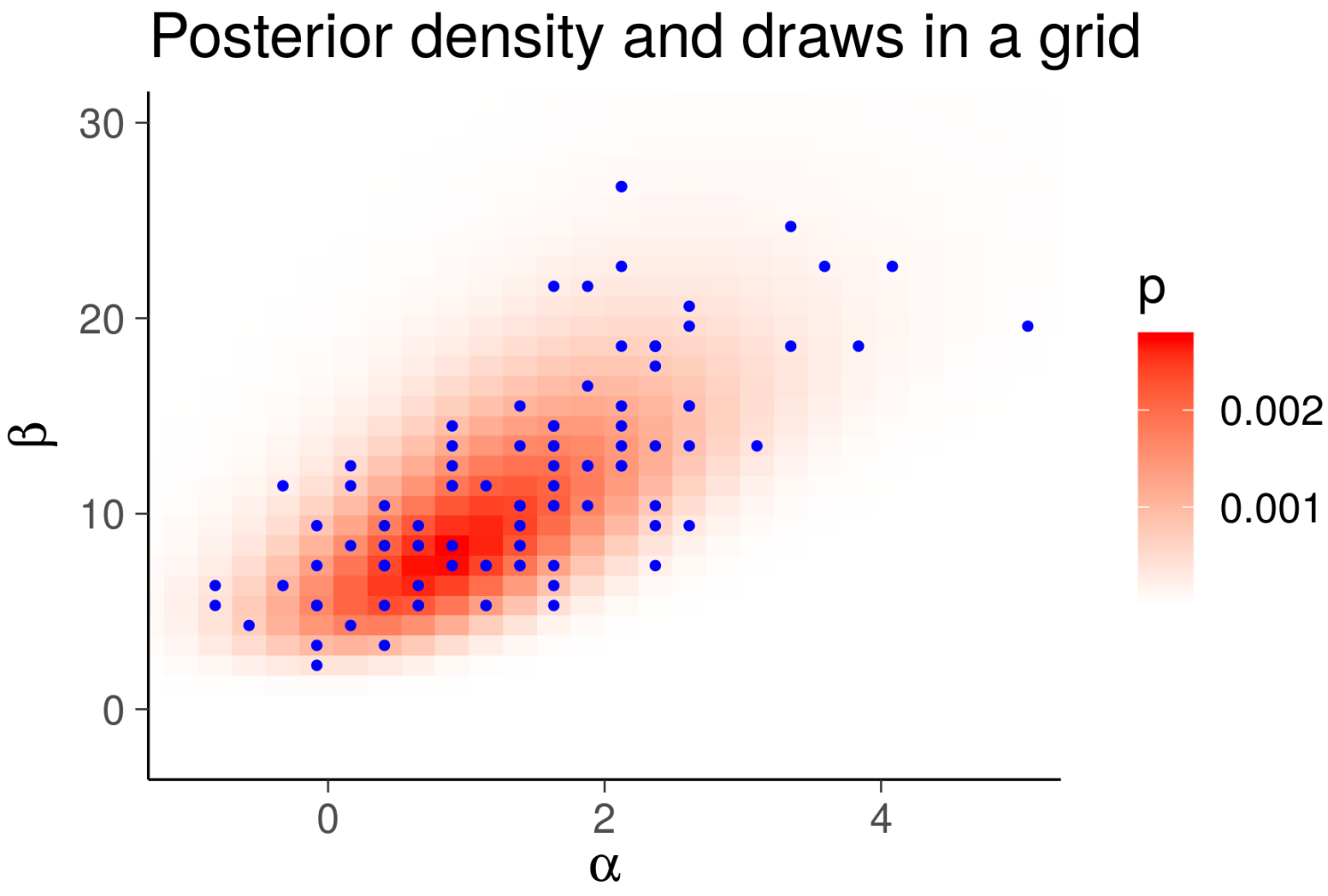
- Densities at 1, 2, and 3: 0.0027 0.0010 0.0001
- Probabilities of cells 1, 2, and 3: 0.0431 0.0166 0.0010
- Probabilities of cells sum to 1

Bioassay



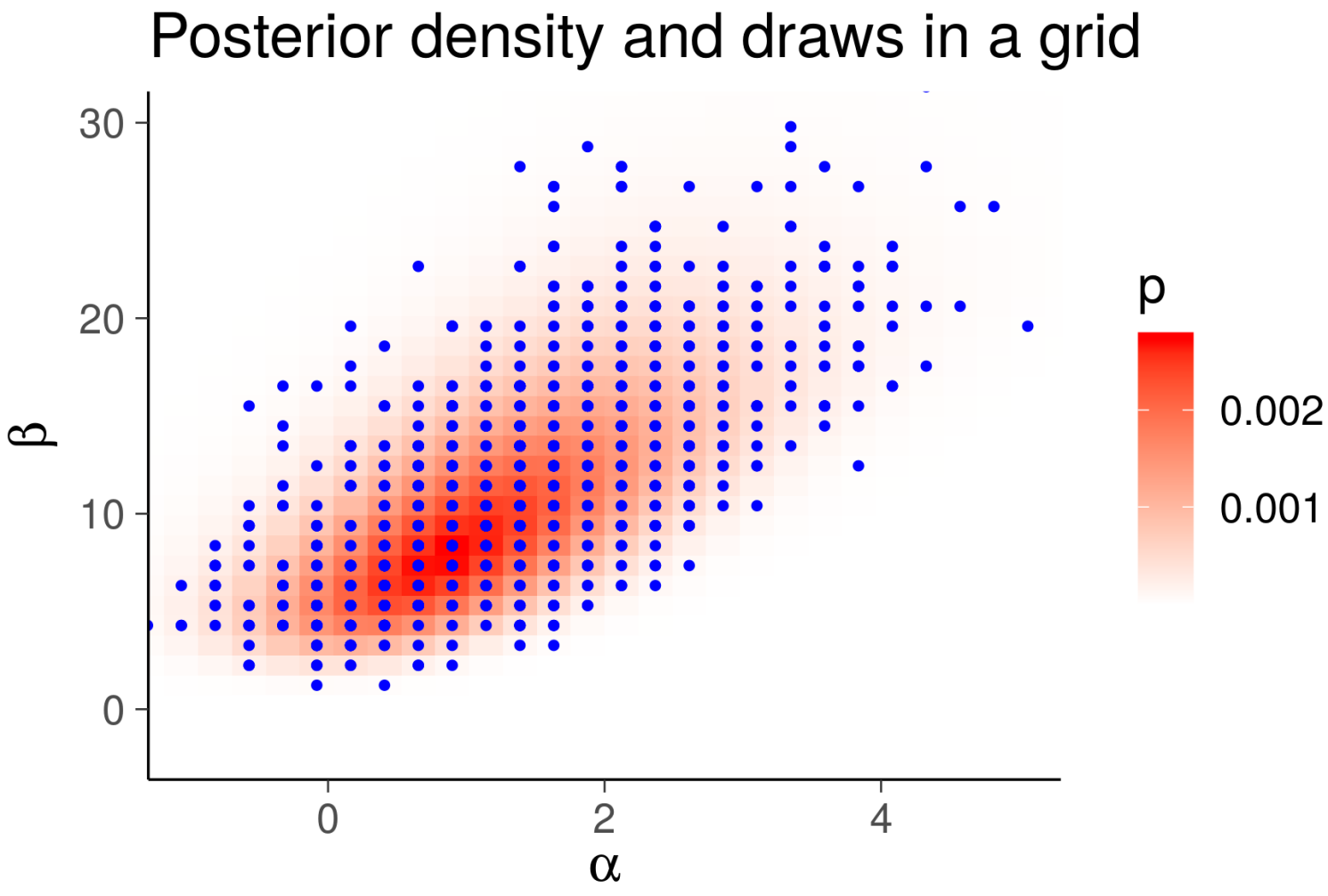
- Sample according to grid cell probabilities

Bioassay



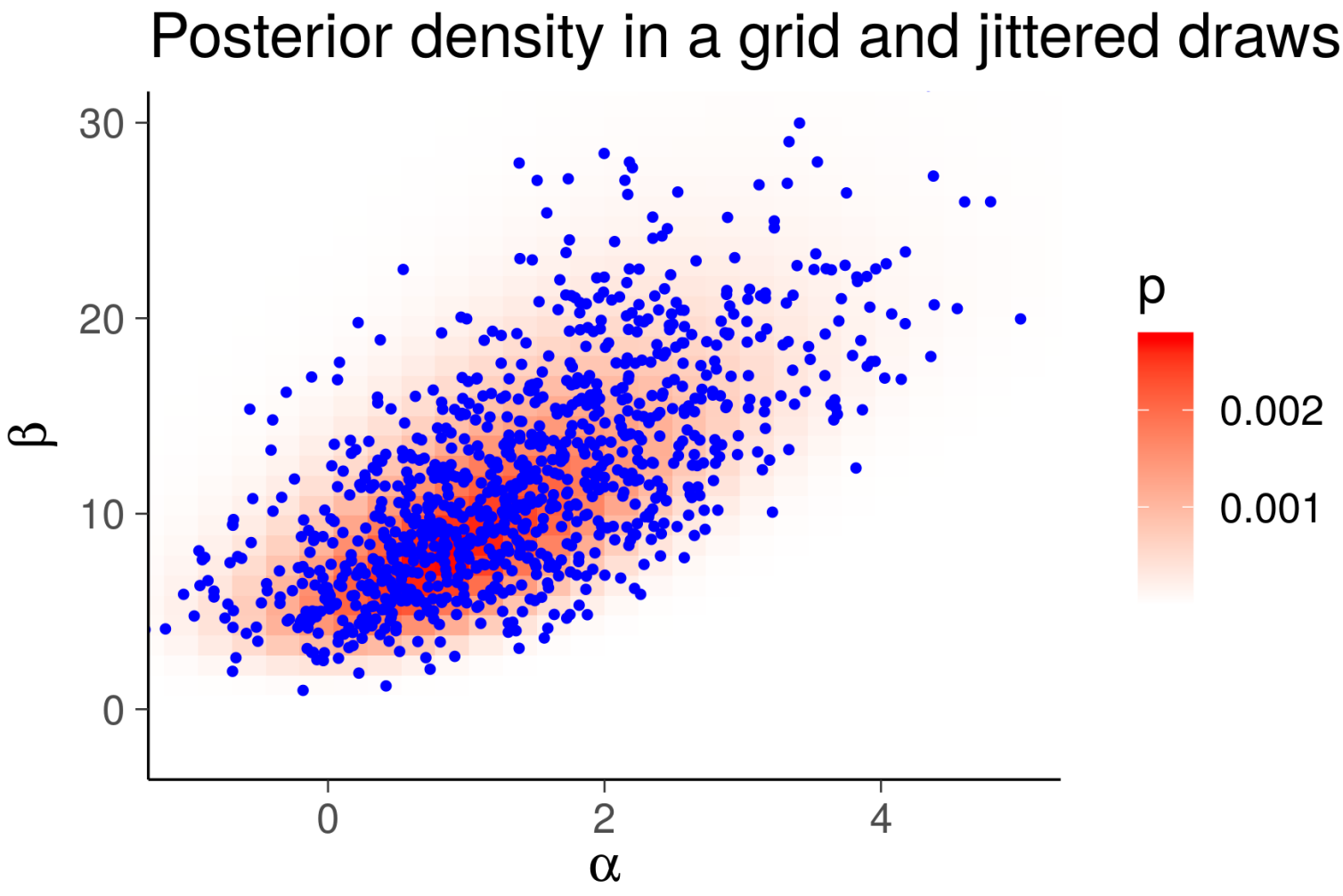
- Sample according to grid cell probabilities

Bioassay



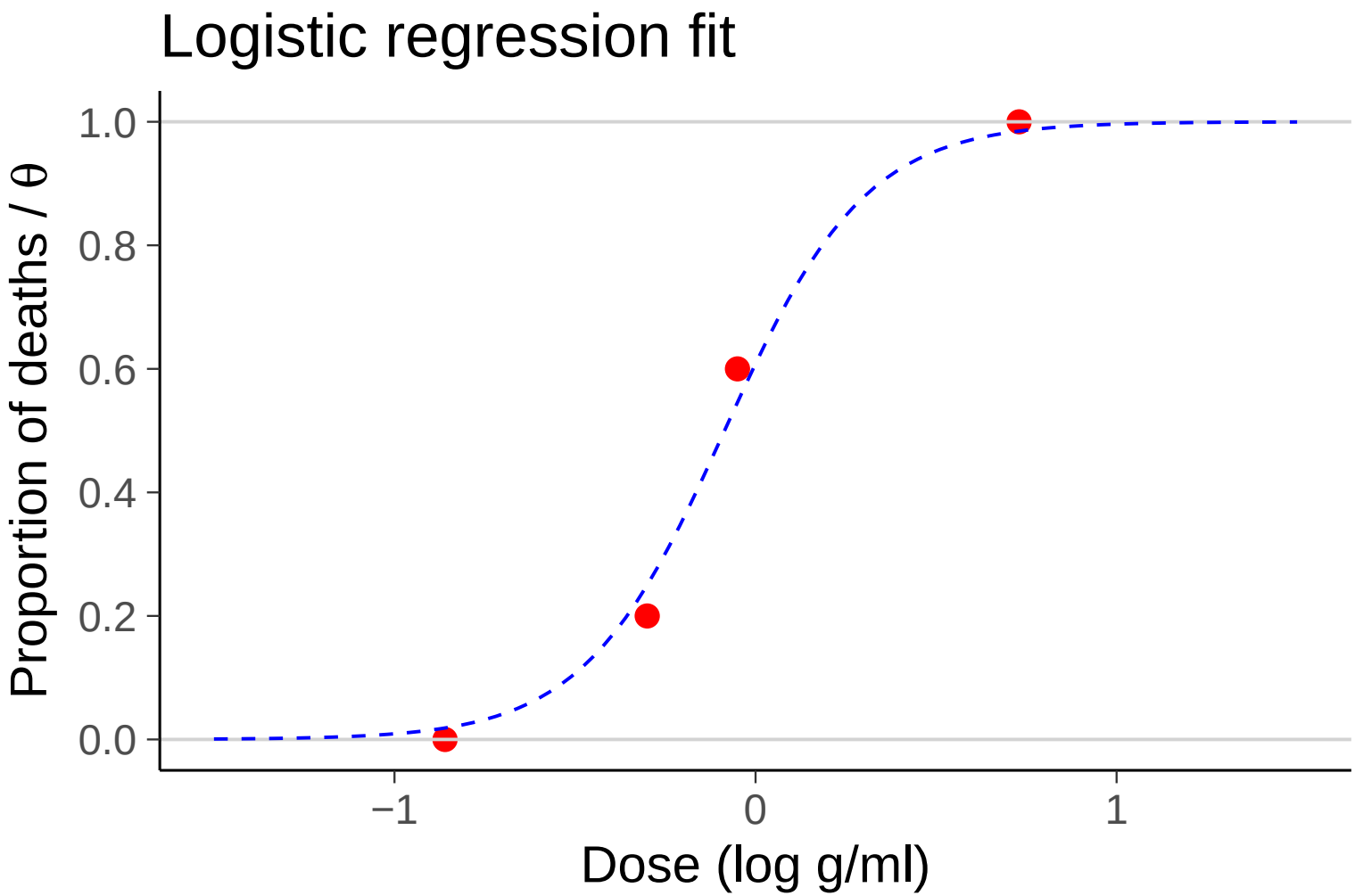
- Sample according to grid cell probabilities
- Several draws can be from the same grid cell

Bioassay

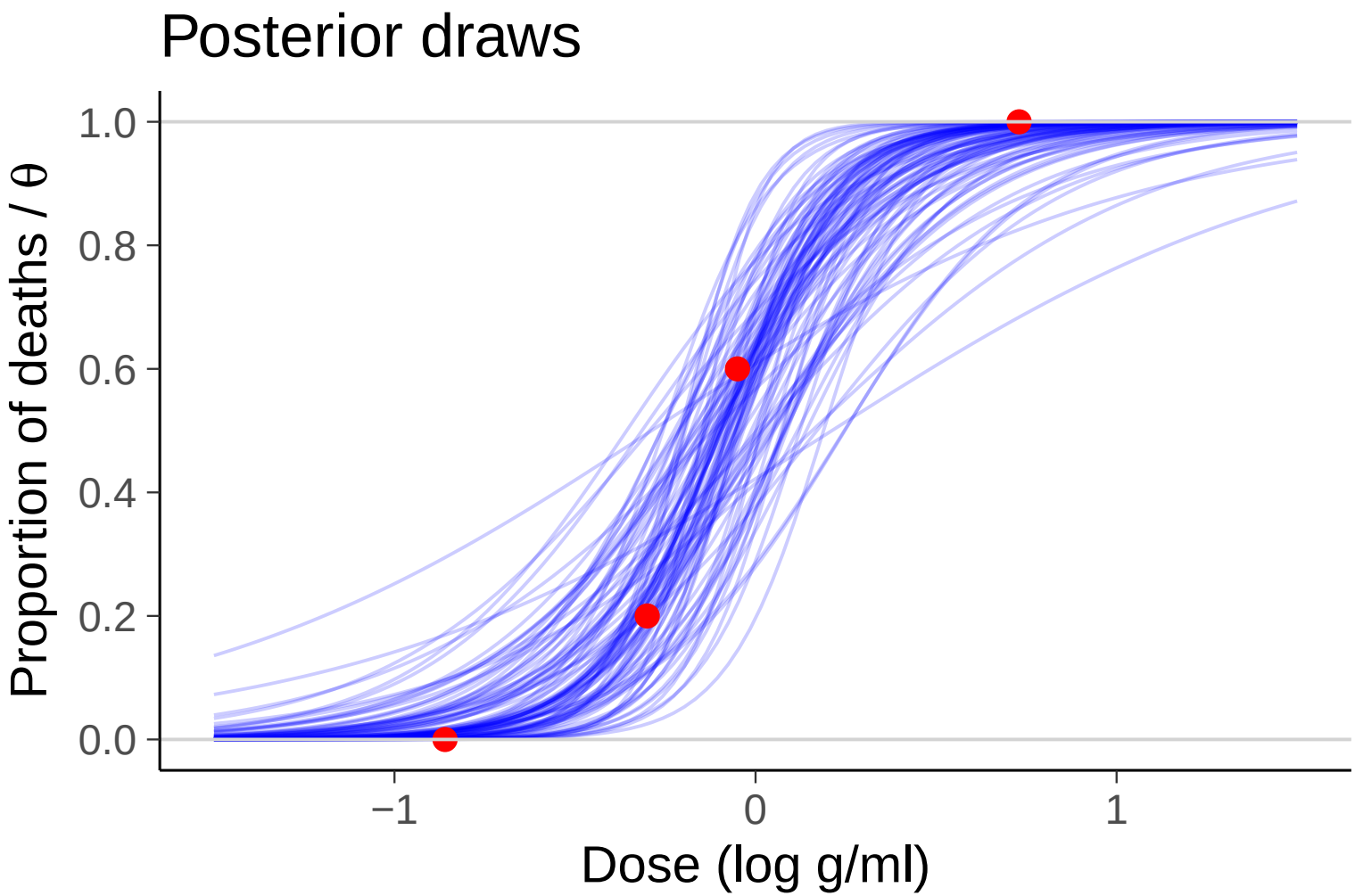


- Jitter can be added to improve visualization

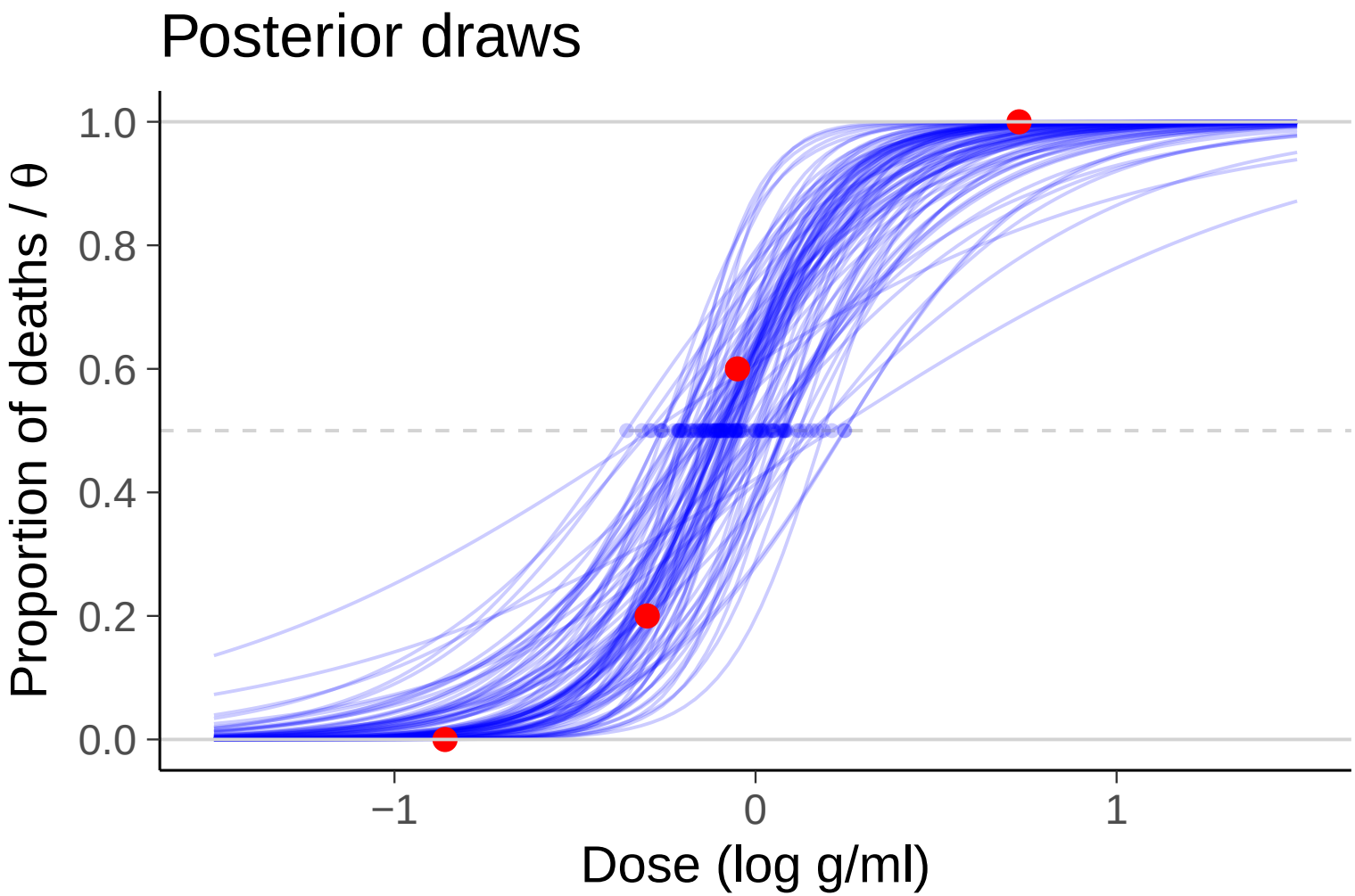
Bioassay


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

Bioassay

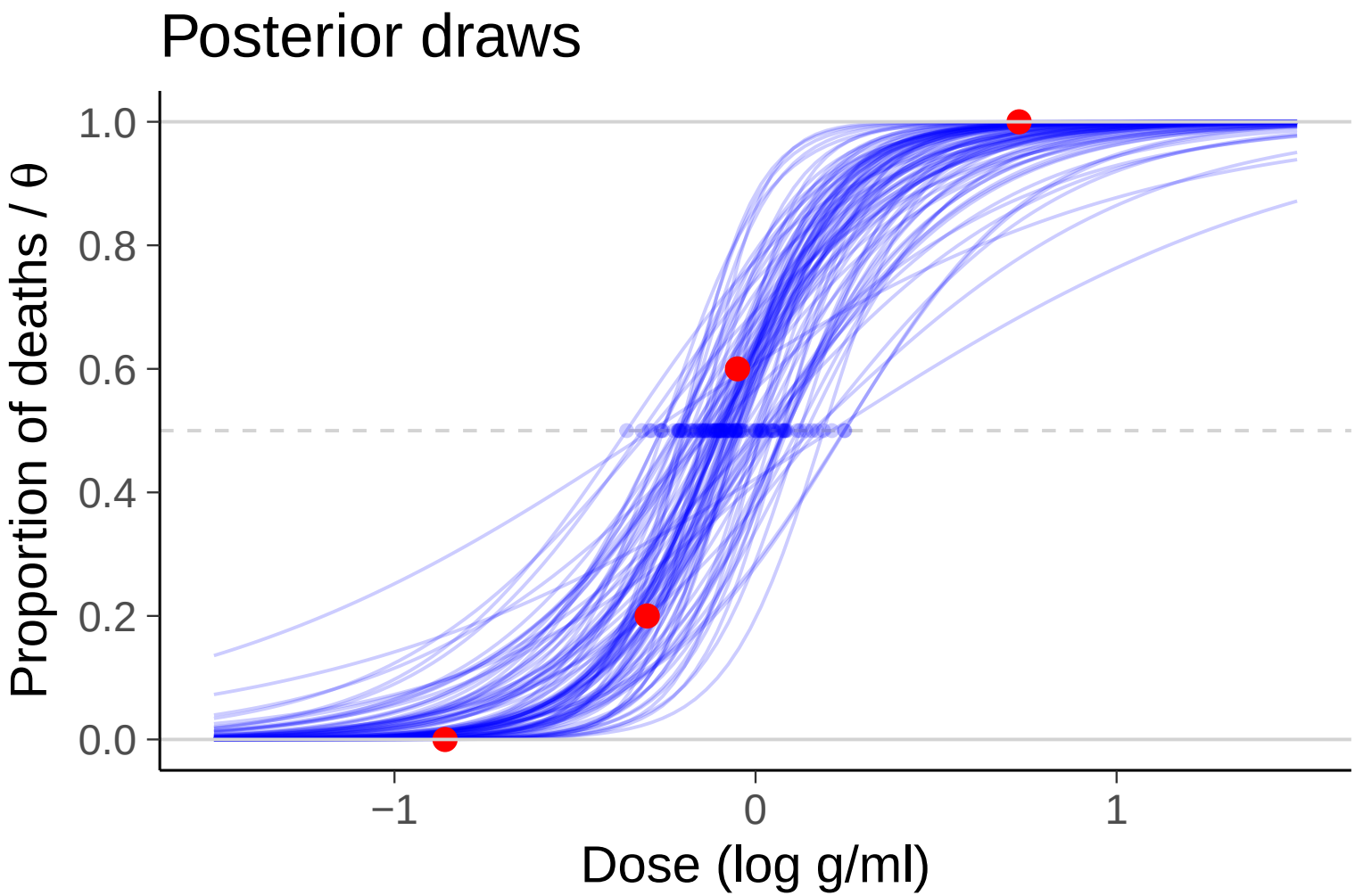

Posterior draws
Proportion of deaths / θ
1.0
0.8
0.6
0.4
0.2
0.0
−1
0
1
Dose (log g/ml)

Bioassay



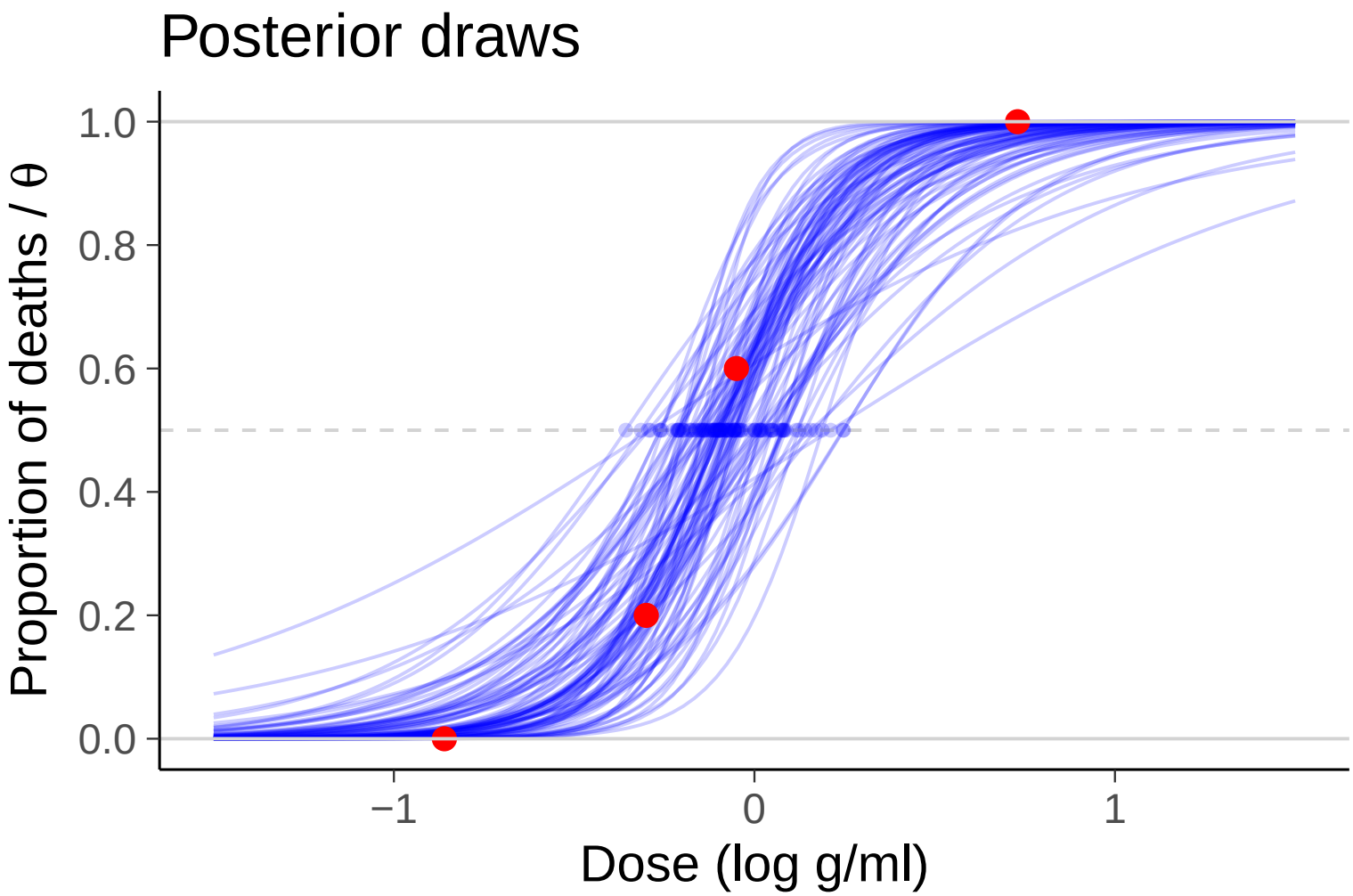
LD50: $\mathrm{E}\left(\frac{y}{n}\right) = \mathrm{logit}^{-1}(\alpha + \beta x) = 0.5$

Bioassay



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

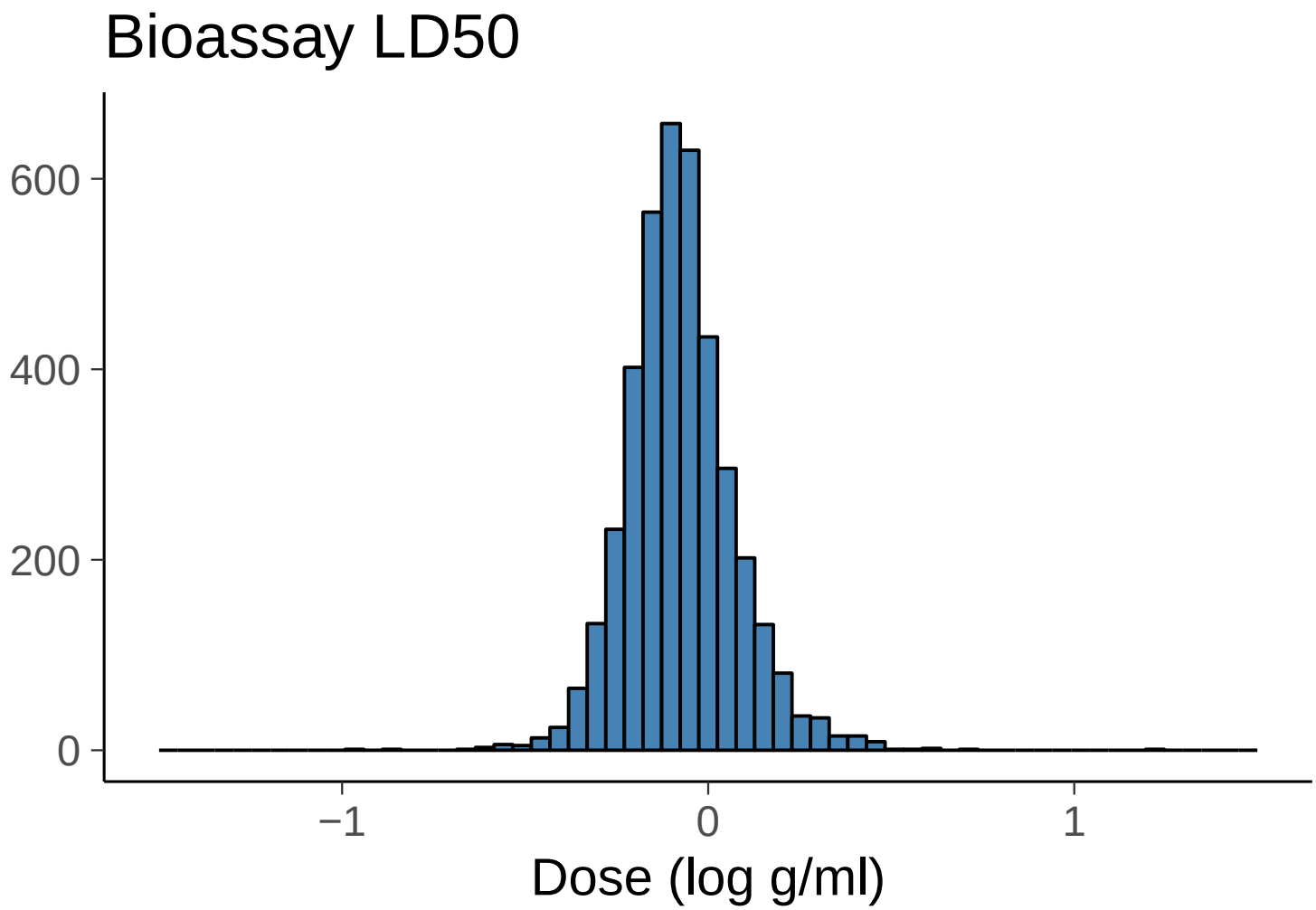
Bioassay



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

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

Bioassay



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

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

Grid sampling

- Draws can be used to estimate expectations, for example

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

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

Summary of elementary modeling and computation

From Gelman, p. 78

> *Strategy for computation of simple Bayesian posterior distributions*

Summary of elementary modeling and computation

From Gelman, p. 78

Strategy for computation of simple Bayesian posterior distributions

1. *Write the likelihood part of the model, $p(y|\theta)$, ignoring any factors that are free of θ.*

Summary of elementary modeling and computation

From Gelman, p. 78

Strategy for computation of simple Bayesian posterior distributions

1. *Write the likelihood part of the model, $p(y|\theta)$, ignoring any factors that are free of θ.*
2. *Write the posterior density, $p(\theta|y) \propto p(\theta)p(y|\theta)$.*
   - *If prior information is well-formulated, include it in $p(\theta)$.*
   - *Otherwise use a weakly informative prior distribution or temporarily set $p(\theta) \propto$ constant, with the understanding that the prior density can be altered later to include additional information or structure.*

Summary of elementary modeling and computation

From Gelman, p. 78

Strategy for computation of simple Bayesian posterior distributions

1. *Write the likelihood part of the model, $p(y|\theta)$, ignoring any factors that are free of θ.*
2. *Write the posterior density, $p(\theta|y) \propto p(\theta)p(y|\theta)$.*
   - *If prior information is well-formulated, include it in $p(\theta)$.*
   - *Otherwise use a weakly informative prior distribution or temporarily set $p(\theta) \propto$ constant, with the understanding that the prior density can be altered later to include additional information or structure.*
3. *Create a crude estimate of the parameters, θ, for use as a starting point and a comparison to the computation in the next step.*

Summary of elementary modeling and computation

1. *Draw simulations $\theta^1, \ldots, \theta^S$, from the posterior distribution. Use the sample draws to compute the posterior density of any functions of θ that may be of interest.*

Summary of elementary modeling and computation

1. *Draw simulations $\theta^1, \ldots, \theta^S$, from the posterior distribution. Use the sample draws to compute the posterior density of any functions of θ that may be of interest.*
2. *If any predictive quantities, $\tilde{y}$, are of interest, simulate $\tilde{y}^1, \ldots, \tilde{y}^S$ by drawing each $\tilde{y}^s$ from the sampling distribution conditional on the drawn value θ^s, $p(\tilde{y}|\theta^s)$. ... posterior simulations of θ and $\tilde{y}$ can be used to check the fit of the model to data and substantive knowledge.*

Summary of elementary modeling and computation

1. *Draw simulations $\theta^1, \ldots, \theta^S$, from the posterior distribution. Use the sample draws to compute the posterior density of any functions of θ that may be of interest.*
2. *If any predictive quantities, $\tilde{y}$, are of interest, simulate $\tilde{y}^1, \ldots, \tilde{y}^S$ by drawing each $\tilde{y}^s$ from the sampling distribution conditional on the drawn value θ^s, $p(\tilde{y}|\theta^s)$. ... posterior simulations of θ and $\tilde{y}$ can be used to check the fit of the model to data and substantive knowledge.*

For nonconjugate models, step 4 can be difficult. Various methods have been developed to draw posterior simulations in complicated models.

Occasionally, high-dimensional problems can be solved by combining analytical and numerical simulation methods. If θ has only one or two components, it is possible to draw simulations by computing on a grid (as for the bioassay example).