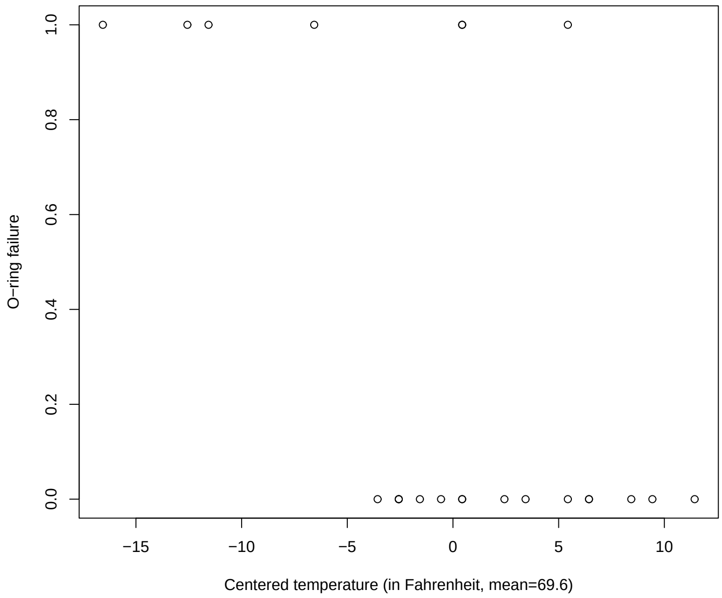
Example: O-ring failures by temperature

- Christensen (1997) and Congdon (2001) analyze 23 binary observations of O-ring failures y_i (1=failure) in relation to temperature t_i (Fahrenheit).
- $y = (1,1,1,1,0,0,0,0,0,0,0,0,1,1,0,0,0,1,0,0,0,0,0)$
- $t = (53,57,58,63,66,67,67,67,68,69,70,70,70,$
  $70,72,73,75,75,76,76,78,79,81)$



What is $Pr(\tilde{y} = 1|\tilde{x})$, for $\tilde{x} = 31, 33, \ldots, 51$?

- Bernoulli model:
$$y_i|\theta_i \sim Bern(\theta_i)$$
for $i = 1, \ldots, n = 23$.

- Logit link:
$$\log\left(\frac{\theta_i}{1-\theta_i}\right) = \alpha + \beta x_i$$
where $\alpha = -1.26$.

- $x_i = t_i - \bar{t}$ where $\bar{t} = 69.56522$.

- Likelihood for $\theta = (\theta_1, \ldots, \theta_n)$
$$l(\theta; y) = \prod_{i=1}^{n} \theta_i^{y_i}(1-\theta_i)^{1-y_i} = \left(\frac{\theta_i}{1-\theta_i}\right)^{y_i}(1-\theta_i)$$

- It is easy to see that
$$\frac{\theta_i}{1-\theta_i} = e^{\alpha+\beta x_i} \text{ and } (1-\theta_i) = \left(1 + e^{\alpha+\beta x_i}\right)^{-1}$$

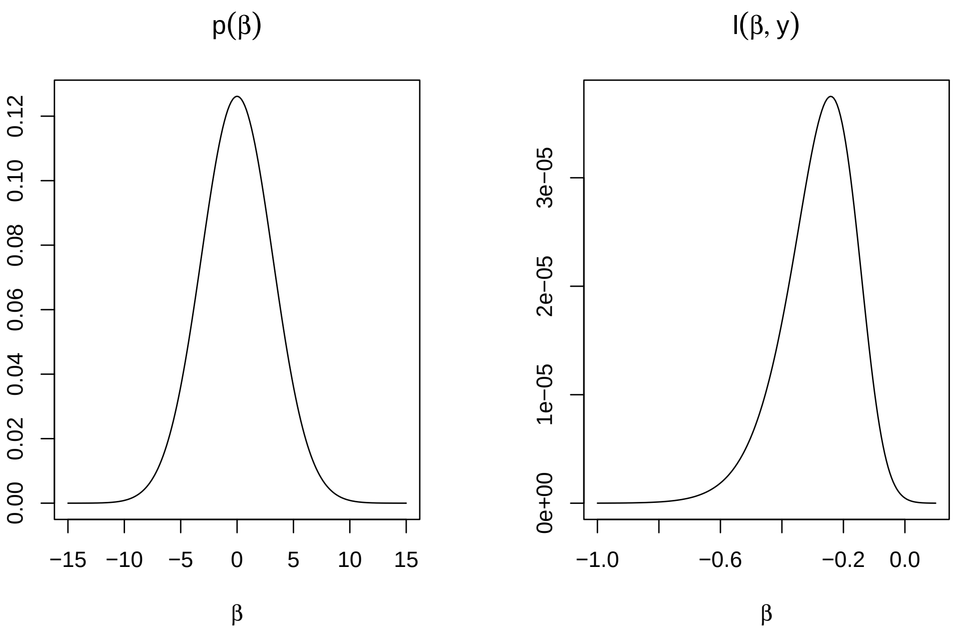
- Likelihood of β is

$$l(\beta; y) = \prod_{i=1}^{n} e^{(\alpha+\beta x_i) y_i} / (1 + e^{\alpha+\beta x_i})$$

- Prior of β is

$$p(\beta) = \frac{e^{-0.5(\beta-\beta_0)^2/V_\beta}}{\sqrt{2\pi V_\beta}}$$

for $\beta_0 = 0$ and $V_\beta = 10$.



- Kernel of the posterior of β

$$\begin{aligned}
p(\beta|y) &\propto p(\beta)l(\beta;y) \\
&\propto e^{-0.5(\beta-\beta_0)^2/V_\beta} \\
&\times \prod_{i=1}^{n} e^{(\alpha+\beta x_i)y_i}\left(1+e^{(\alpha+\beta x_i)}\right)^{-1}
\end{aligned}$$

- The posterior of β is explicitly written as

$$p(\beta|y) = \frac{p(\beta)l(\beta;y)}{p(y)}$$

- Predictive density

$$\begin{aligned}
p(y) &= \int_0^1 l(\beta;y)p(\beta)d\beta \\
&= \int_0^1 \left[\prod_{i=1}^{n} e^{(\alpha+\beta x_i)y_i}\left(1+e^{(\alpha+\beta x_i)}\right)^{-1}\right] \\
&\times p(\beta)d\beta.
\end{aligned}$$

Approximating $p(y)$ by Simpson's rule

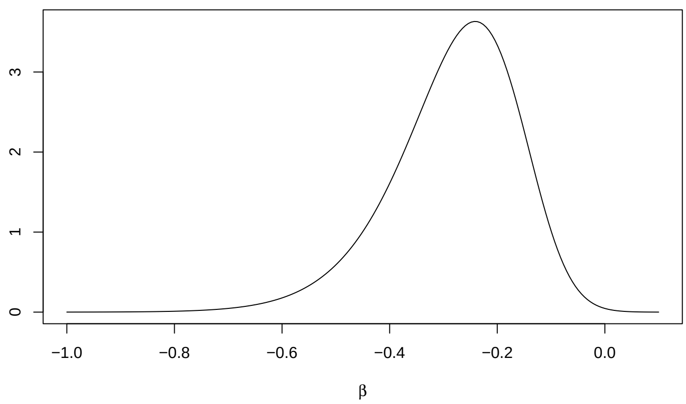
Let $\beta_1 = -1.0, \beta_2 = -0.9999, \ldots, \beta_{11001} = 0.1$.

Then $p(y)$ can be approximated by

$$\begin{aligned} \widehat{p}_{sr}(y) &= 10^{-4} \sum_{i=1}^{11001} l(\beta_i; y) p(\beta_i) \\ &= 0.000001298902. \end{aligned}$$

Therefore, $p(\beta|y)$ can be approximated by

$$\widehat{p}_{sr}(\beta|y) = \frac{l(\beta; y)p(\beta)}{\widehat{p}_{sr}(y)}.$$



Also by Simpson's rule

$$
\begin{aligned}
\widehat{E}_{sr}(\beta|y) &= \frac{1}{\widehat{p}_{sr}(y)10^4} \sum_{i=1}^{11001} \beta_i l(\beta_i; y) p(\beta_i) \\
&= -0.2806
\end{aligned}
$$

$$
\begin{aligned}
\widehat{E}_{sr}(\beta^2|y) &= \frac{1}{\widehat{p}_{sr}(y)10^4} \sum_{i=1}^{11001} \beta_i^2 l(\beta_i; y) p(\beta_i) \\
&= 0.0929 \\
\widehat{V}_{sr}(\beta|y) &= 0.0929 - (-0.2806)^2 = 0.01413
\end{aligned}
$$

Approximating $p(y)$ by MC integration

- If $\beta_1, \ldots, \beta_M$ is a sample from $p(\beta)$, then

$$\widehat{p}_{mc}(y) = \frac{1}{M} \sum_{i=1}^{M} l(\beta_i; y)$$

  is a Monte Carlo estimate of $p(y)$.

- We use $M = 50000$.

- $\widehat{p}_{mc}(y) = 0.000001270616$.

- Monte Carlo error $= 1.115089 \times 10^{-10}$.

- Monte Carlo error is 0.00877597% of $\widehat{p}_{mc}(y)$.

- Obtain a MC estimate of $E(\beta|y)$ and $V(\beta|y)$.

Approximating $p(y)$ by MC integration

- If $\beta_1, \ldots, \beta_M$ is a sample from $p(\beta)$, then

$$\widehat{p}_{mc}(y) = \frac{1}{M} \sum_{i=1}^{M} l(\beta_i; y)$$

  is a Monte Carlo estimate of $p(y)$.

- We use $M = 50000$.

- $\widehat{p}_{mc}(y) = 0.000001270616$.

- Monte Carlo error $= 1.115089 \times 10^{-10}$.

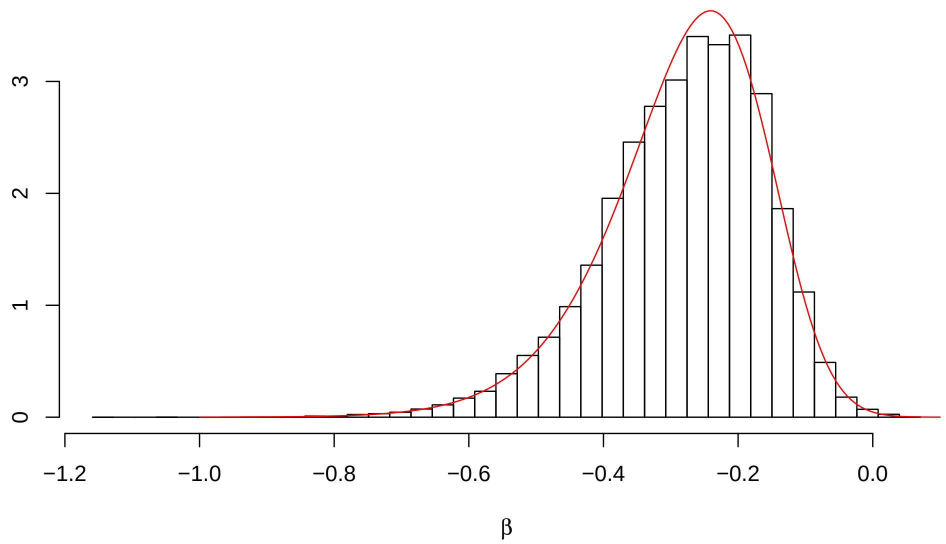
- Monte Carlo error is 0.00877597% of $\widehat{p}_{mc}(y)$.

- Obtain a MC estimate of $E(\beta|y)$ and $V(\beta|y)$.

Sampling from $p(\beta|y)$ by weighted resampling

Let $\beta_1, \ldots, \beta_M$ be (the previously used) sample from $p(\beta)$.

The posterior density of β can be approximated by a histogram based on the β_i's.



-0.281, 0.0145, -0.553, -0.268 and -0.088 are approximations to the posterior mean, variance, 2.5, 50.0 and 97.5 percentiles of $p(\beta|y)$, respectively.

Posterior predictions

$$
\begin{aligned}
Pr(\tilde{y}=1|\tilde{x}) &= E_{\beta|y}\left[Pr(\tilde{y}=1|\tilde{x},\beta)\right] \\
&= E_{\beta|y}\left(\frac{e^{\alpha+\beta\tilde{x}}}{1+e^{\alpha+\beta\tilde{x}}}\right)
\end{aligned}
$$

can be approximated (by MC integration) by

$$
\hat{P}r(\tilde{y}=1|\tilde{x}) = \frac{1}{M}\sum_{i=1}^{M}\left(\frac{e^{\alpha+\beta_i\tilde{x}}}{1+e^{\alpha+\beta_i\tilde{x}}}\right)
$$

where $\beta_1,\ldots,\beta_M$ is a sample from $p(\beta|y)$.