

■ Assignment 1

- Currently being marked, almost done
- Will try to release marks/return scripts this week/early next week

■ Assignment 2

- Questions to complete: 14, 40, 41 and also 30
- Due in by 4pm, Friday 3rd May
- Associated practical: currently scheduled at 1pm, Friday 3rd May (not helpful!), so trying to reschedule (perhaps next week), check emails!

■ Mid-semester test

- Work looking good so far
- Hoping to get the work back to you very soon

■ Assignment 1

- Currently being marked, almost done
- Will try to release marks/return scripts this week/early next week

■ Assignment 2

- Questions to complete: **14, 40, 41** and also **30**
- Due in by 4pm, Friday 3rd May
- Associated practical: currently scheduled at 1pm, Friday 3rd May (not helpful!), so trying to reschedule (perhaps next week), check emails!

■ Mid-semester test

- Work looking good so far
- Hoping to get the work back to you very soon

■ Assignment 1

- Currently being marked, almost done
- Will try to release marks/return scripts this week/early next week

■ Assignment 2

- Questions to complete: **14, 40, 41** and also **30**
- Due in by 4pm, Friday 3rd May
- Associated practical: currently scheduled at 1pm, Friday 3rd May (not helpful!), so trying to reschedule (perhaps next week), check emails!

■ Mid-semester test

- Work looking good so far
- Hoping to get the work back to you very soon

■ Assignment 1

- Currently being marked, almost done
- Will try to release marks/return scripts this week/early next week

■ Assignment 2

- Questions to complete: **14, 40, 41** and also **30**
- Due in by 4pm, Friday 3rd May
- Associated practical: currently scheduled at 1pm, Friday 3rd May (not helpful!), so trying to reschedule (perhaps next week), check emails!

■ Mid-semester test

- Work looking good so far
- Hoping to get the work back to you very soon

■ Assignment 1

- Currently being marked, almost done
- Will try to release marks/return scripts this week/early next week

■ Assignment 2

- Questions to complete: **14, 40, 41** and also **30**
- Due in by 4pm, Friday 3rd May
- Associated practical: currently scheduled at 1pm, Friday 3rd May (not helpful!), so trying to reschedule (perhaps next week), check emails!

■ Mid-semester test

- Work looking good so far
- Hoping to get the work back to you very soon

Professor Richard Boys



- Sadly died on Tuesday 5th March

- Preface of your lecture notes:

"Since 1980, the number of academic staff in Mathematics & Statistics at Newcastle publishing advanced research using Bayesian methods has increased dramatically. In the 1980s, there was only one Bayesian at Newcastle. Now there are at least 12."

- "...one Bayesian at Newcastle": Professor Boys

Professor Richard Boys



- Sadly died on Tuesday 5th March
- Preface of your lecture notes:

“Since 1980, the number of academic staff in Mathematics & Statistics at Newcastle publishing advanced research using Bayesian methods has increased dramatically. In the 1980s, there was only one Bayesian at Newcastle. Now there are at least 12.”

- “...one Bayesian at Newcastle”: Professor Boys

Professor Richard Boys



- Sadly died on Tuesday 5th March
- Preface of your lecture notes:

“Since 1980, the number of academic staff in Mathematics & Statistics at Newcastle publishing advanced research using Bayesian methods has increased dramatically. In the 1980s, there was only one Bayesian at Newcastle. Now there are at least 12.”

- “...one Bayesian at Newcastle”: Professor Boys

Professor Richard Boys



- Richard was a big personality in the department, an extremely talented Statistician and a good friend and colleague to many staff
- His funeral will be taking place this coming Thursday, so our scheduled session at 2pm is **cancelled**

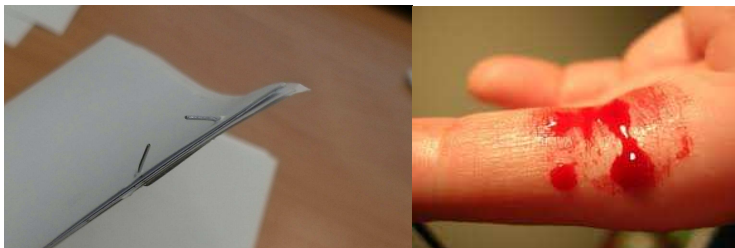
Professor Richard Boys



- Richard was a big personality in the department, an extremely talented Statistician and a good friend and colleague to many staff
- His funeral will be taking place this coming Thursday, so our scheduled session at 2pm is **cancelled**

Dangerous stapling

Several assignments were poorly stapled

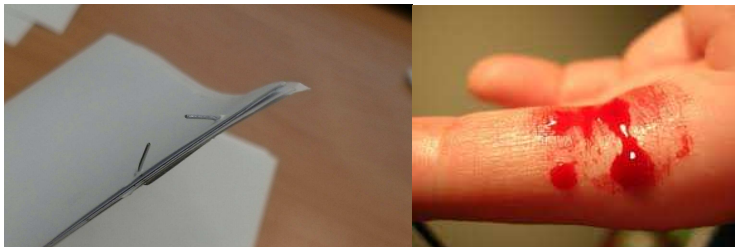


Staples like this are dangerous and can cause pages to go missing. Many of your assignments had to be re-stapled. It is your responsibility to make sure your work is held together securely and safely, by pushing the stapler down firmly.

Assignments with unsafe staples will not be marked!

Dangerous stapling

Several assignments were poorly stapled



Staples like this are dangerous and can cause pages to go missing. Many of your assignments had to be re-stapled. It is your responsibility to make sure your work is held together securely and safely, by pushing the stapler down firmly.

Assignments with unsafe staples will not be marked!

Asymptotic posterior distribution

There are many limiting results in Statistics.

The one you will probably remember is the **Central Limit Theorem**.

This concerns the distribution of $\bar{X}_n$, the mean of n independent and identically distributed random variables (each with known mean μ and known variance σ^2), as the sample size $n \rightarrow \infty$.

It is easy to show that $E(\bar{X}_n) = \mu$ and $Var(\bar{X}_n) = \sigma^2/n$, and so if we define

$$Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma},$$

then we know that

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty.$$

The following theorem gives a similar result for the posterior distribution.

Asymptotic posterior distribution

There are many limiting results in Statistics.

The one you will probably remember is the **Central Limit Theorem**.

This concerns the distribution of $\bar{X}_n$, the mean of n independent and identically distributed random variables (each with known mean μ and known variance σ^2), as the sample size $n \rightarrow \infty$.

It is easy to show that $E(\bar{X}_n) = \mu$ and $Var(\bar{X}_n) = \sigma^2/n$, and so if we define

$$Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma},$$

then we know that

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty.$$

The following theorem gives a similar result for the posterior distribution.

Asymptotic posterior distribution

There are many limiting results in Statistics.

The one you will probably remember is the **Central Limit Theorem**.

This concerns the distribution of $\bar{X}_n$, the mean of n independent and identically distributed random variables (each with known mean μ and known variance σ^2), as the sample size $n \rightarrow \infty$.

It is easy to show that $E(\bar{X}_n) = \mu$ and $Var(\bar{X}_n) = \sigma^2/n$, and so if we define

$$Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma},$$

then we know that

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty.$$

The following theorem gives a similar result for the posterior distribution.

Asymptotic posterior distribution

There are many limiting results in Statistics.

The one you will probably remember is the **Central Limit Theorem**.

This concerns the distribution of $\bar{X}_n$, the mean of n independent and identically distributed random variables (each with known mean μ and known variance σ^2), as the sample size $n \rightarrow \infty$.

It is easy to show that $E(\bar{X}_n) = \mu$ and $Var(\bar{X}_n) = \sigma^2/n$, and so if we define

$$Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma},$$

then we know that

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty.$$

The following theorem gives a similar result for the posterior distribution.

Asymptotic posterior distribution

There are many limiting results in Statistics.

The one you will probably remember is the **Central Limit Theorem**.

This concerns the distribution of $\bar{X}_n$, the mean of n independent and identically distributed random variables (each with known mean μ and known variance σ^2), as the sample size $n \rightarrow \infty$.

It is easy to show that $E(\bar{X}_n) = \mu$ and $Var(\bar{X}_n) = \sigma^2/n$, and so if we define

$$Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma},$$

then we know that

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty.$$

The following theorem gives a similar result for the posterior distribution.

Asymptotic posterior distribution

Theorem (Asymptotic posterior distribution)

Suppose we have a *statistical model* $f(\mathbf{x}|\theta)$ for data $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, together with a *prior distribution* $\pi(\theta)$ for θ . Then

$$\sqrt{J(\hat{\theta})} (\theta - \hat{\theta}) | \mathbf{x} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty,$$

where $\hat{\theta}$ is the likelihood mode and $J(\theta)$ is the *observed information*

$$J(\theta) = -\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta).$$

Using Bayes Theorem, the posterior distribution for θ is

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f(\mathbf{x}|\theta).$$

Let $\psi = \sqrt{n}(\theta - \hat{\theta})$ and

$$\ell_n(\theta) = \frac{1}{n} \log f(\mathbf{x}|\theta)$$

be the average log-likelihood per observation, in which case,

$$f(\mathbf{x}|\theta) = e^{n\ell_n(\theta)}.$$

Using Bayes Theorem, the posterior distribution for θ is

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) f(\mathbf{x}|\theta).$$

Let $\psi = \sqrt{n}(\theta - \hat{\theta})$ and

$$\ell_n(\theta) = \frac{1}{n} \log f(\mathbf{x}|\theta)$$

be the average log-likelihood per observation, in which case,

$$f(\mathbf{x}|\theta) = e^{n\ell_n(\theta)}.$$

Outline proof (2/7)

Recall Equation (3.7), which tells us about the distribution of a random variable $Y = g(X)$:

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|.$$

We want to know the distribution of $\psi = g(\theta)$, where

$$g(\theta) = \sqrt{n}(\theta - \hat{\theta}).$$

Now

$$g^{-1}(\psi) = \hat{\theta} + \frac{\psi}{\sqrt{n}} \quad \text{and} \quad \frac{d}{d\psi} g^{-1}(\psi) = \frac{1}{\sqrt{n}},$$

giving

$$\pi_{\psi}(\psi) = \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \middle| \mathbf{x} \right) \times \frac{1}{\sqrt{n}}.$$

Outline proof (2/7)

Recall Equation (3.7), which tells us about the distribution of a random variable $Y = g(X)$:

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \left| \frac{d}{dy} g^{-1}(y) \right|.$$

We want to know the distribution of $\psi = g(\theta)$, where

$$g(\theta) = \sqrt{n}(\theta - \hat{\theta}).$$

Now

$$g^{-1}(\psi) = \hat{\theta} + \frac{\psi}{\sqrt{n}} \quad \text{and} \quad \frac{d}{d\psi} g^{-1}(\psi) = \frac{1}{\sqrt{n}},$$

giving

$$\pi_{\psi}(\psi) = \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \middle| \mathbf{x} \right) \times \frac{1}{\sqrt{n}}.$$

Outline proof (2/7)

Recall Equation (3.7), which tells us about the distribution of a random variable $Y = g(X)$:

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \left| \frac{d}{dy} g^{-1}(y) \right|.$$

We want to know the distribution of $\psi = g(\theta)$, where

$$g(\theta) = \sqrt{n}(\theta - \hat{\theta}).$$

Now

$$g^{-1}(\psi) = \hat{\theta} + \frac{\psi}{\sqrt{n}} \quad \text{and} \quad \frac{d}{d\psi} g^{-1}(\psi) = \frac{1}{\sqrt{n}},$$

giving

$$\pi_{\psi}(\psi) = \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \middle| \mathbf{x} \right) \times \frac{1}{\sqrt{n}}.$$

Outline proof (2/7)

Recall Equation (3.7), which tells us about the distribution of a random variable $Y = g(X)$:

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \left| \frac{d}{dy} g^{-1}(y) \right|.$$

We want to know the distribution of $\psi = g(\theta)$, where

$$g(\theta) = \sqrt{n}(\theta - \hat{\theta}).$$

Now

$$g^{-1}(\psi) = \hat{\theta} + \frac{\psi}{\sqrt{n}} \quad \text{and} \quad \frac{d}{d\psi} g^{-1}(\psi) = \frac{1}{\sqrt{n}},$$

giving

$$\pi_{\psi}(\psi) = \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \middle| \mathbf{x} \right) \times \frac{1}{\sqrt{n}}.$$

Outline proof (2/7)

Recall Equation (3.7), which tells us about the distribution of a random variable $Y = g(X)$:

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \left| \frac{d}{dy} g^{-1}(y) \right|.$$

We want to know the distribution of $\psi = g(\theta)$, where

$$g(\theta) = \sqrt{n}(\theta - \hat{\theta}).$$

Now

$$g^{-1}(\psi) = \hat{\theta} + \frac{\psi}{\sqrt{n}} \quad \text{and} \quad \frac{d}{d\psi} g^{-1}(\psi) = \frac{1}{\sqrt{n}},$$

giving

$$\pi_{\psi}(\psi) = \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \middle| \mathbf{x} \right) \times \frac{1}{\sqrt{n}}.$$

Thus

$$\begin{aligned}\pi_{\psi}(\psi) &= \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \middle| \mathbf{x} \right) \times \frac{1}{\sqrt{n}} \\ &\propto \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \right) \exp \left\{ n \ell_n \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \right) \right\}.\end{aligned}$$

Taking Taylor series expansions about $\psi = 0$ gives

$$\begin{aligned}\pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \right) &= \pi_{\theta}(\hat{\theta}) + \pi'_{\theta}(\hat{\theta}) \frac{\psi}{\sqrt{n}} + \frac{1}{2!} \pi''_{\theta}(\hat{\theta}) \left[\frac{\psi}{\sqrt{n}} \right]^2 \\ &\quad + \frac{1}{3!} \pi'''_{\theta}(\hat{\theta}) \left[\frac{\psi}{\sqrt{n}} \right]^3 + \dots \\ &\approx \pi_{\theta}(\hat{\theta}).\end{aligned}$$

Thus

$$\begin{aligned}\pi_{\psi}(\psi) &= \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \middle| \mathbf{x} \right) \times \frac{1}{\sqrt{n}} \\ &\propto \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \right) \exp \left\{ n \ell_n \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \right) \right\}.\end{aligned}$$

Taking Taylor series expansions about $\psi = 0$ gives

$$\begin{aligned}\pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \right) &= \pi_{\theta}(\hat{\theta}) + \pi'_{\theta}(\hat{\theta}) \frac{\psi}{\sqrt{n}} + \frac{1}{2!} \pi''_{\theta}(\hat{\theta}) \left[\frac{\psi}{\sqrt{n}} \right]^2 \\ &\quad + \frac{1}{3!} \pi'''_{\theta}(\hat{\theta}) \left[\frac{\psi}{\sqrt{n}} \right]^3 + \dots \\ &\approx \pi_{\theta}(\hat{\theta}).\end{aligned}$$

Thus

$$\begin{aligned}\pi_{\psi}(\psi) &= \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \middle| \mathbf{x} \right) \times \frac{1}{\sqrt{n}} \\ &\propto \pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \right) \exp \left\{ n \ell_n \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \right) \right\}.\end{aligned}$$

Taking Taylor series expansions about $\psi = 0$ gives

$$\begin{aligned}\pi_{\theta} \left(\hat{\theta} + \frac{\psi}{\sqrt{n}} \right) &= \pi_{\theta}(\hat{\theta}) + \pi'_{\theta}(\hat{\theta}) \frac{\psi}{\sqrt{n}} + \frac{1}{2!} \pi''_{\theta}(\hat{\theta}) \left[\frac{\psi}{\sqrt{n}} \right]^2 \\ &\quad + \frac{1}{3!} \pi'''_{\theta}(\hat{\theta}) \left[\frac{\psi}{\sqrt{n}} \right]^3 + \dots \\ &\approx \pi_{\theta}(\hat{\theta}).\end{aligned}$$

Similarly,

$$\begin{aligned}n\ell_n\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right) &= n\left\{\ell_n(\hat{\theta}) + \ell'_n(\hat{\theta})\frac{\psi}{\sqrt{n}} + \frac{1}{2!}\ell''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^2\right. \\&\quad \left. + \frac{1}{3!}\ell'''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^3 + \dots\right\} \\&\approx n\ell_n(\hat{\theta}) + n\ell'_n(\hat{\theta})\frac{\psi}{\sqrt{n}} + n\frac{1}{2!}\ell''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^2 \\&\approx n\ell_n(\hat{\theta}) + \frac{1}{2}\ell''_n(\hat{\theta})\psi^2.\end{aligned}$$

Similarly,

$$\begin{aligned}n\ell_n\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right) &= n\left\{\ell_n(\hat{\theta}) + \ell'_n(\hat{\theta})\frac{\psi}{\sqrt{n}} + \frac{1}{2!}\ell''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^2\right. \\&\quad \left. + \frac{1}{3!}\ell'''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^3 + \dots\right\} \\&\approx n\ell_n(\hat{\theta}) + n\ell'_n(\hat{\theta})\frac{\psi}{\sqrt{n}} + n\frac{1}{2!}\ell''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^2 \\&\approx n\ell_n(\hat{\theta}) + \frac{1}{2}\ell''_n(\hat{\theta})\psi^2.\end{aligned}$$

Similarly,

$$\begin{aligned}n\ell_n\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right) &= n\left\{\ell_n(\hat{\theta}) + \ell'_n(\hat{\theta})\frac{\psi}{\sqrt{n}} + \frac{1}{2!}\ell''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^2\right. \\&\quad \left.+ \frac{1}{3!}\ell'''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^3 + \dots\right\} \\&\approx n\ell_n(\hat{\theta}) + n\ell'_n(\hat{\theta})\frac{\psi}{\sqrt{n}} + n\frac{1}{2!}\ell''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^2 \\&\approx n\ell_n(\hat{\theta}) + \frac{1}{2}\ell''_n(\hat{\theta})\psi^2.\end{aligned}$$

Similarly,

$$\begin{aligned}n\ell_n\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right) &= n\left\{\ell_n(\hat{\theta}) + \ell'_n(\hat{\theta})\frac{\psi}{\sqrt{n}} + \frac{1}{2!}\ell''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^2\right. \\&\quad \left.+ \frac{1}{3!}\ell'''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^3 + \dots\right\} \\&\approx n\ell_n(\hat{\theta}) + n\ell'_n(\hat{\theta})\frac{\psi}{\sqrt{n}} + n\frac{1}{2!}\ell''_n(\hat{\theta})\left[\frac{\psi}{\sqrt{n}}\right]^2 \\&\approx n\ell_n(\hat{\theta}) + \frac{1}{2}\ell''_n(\hat{\theta})\psi^2.\end{aligned}$$

This gives

$$\begin{aligned}\pi_{\psi}(\psi|\mathbf{x}) &\propto \pi_{\theta}\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right) \exp\left\{n\ell_n\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right)\right\} \\ &\approx \pi_{\theta}(\hat{\theta}) \exp\left\{n\ell_n(\hat{\theta}) + \frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{\frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{-\frac{[-\ell_n''(\hat{\theta})]}{2}(\psi - 0)^2\right\},\end{aligned}$$

showing that

$$\psi|\mathbf{x} \sim N\left(0, [-\ell_n''(\hat{\theta})]^{-1}\right) \quad \text{as } n \rightarrow \infty.$$

This gives

$$\begin{aligned}\pi_{\psi}(\psi|\mathbf{x}) &\propto \pi_{\theta}\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right) \exp\left\{n\ell_n\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right)\right\} \\ &\approx \pi_{\theta}(\hat{\theta}) \exp\left\{n\ell_n(\hat{\theta}) + \frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{\frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{-\frac{[-\ell_n''(\hat{\theta})]}{2}(\psi - 0)^2\right\},\end{aligned}$$

showing that

$$\psi|\mathbf{x} \sim N\left(0, [-\ell_n''(\hat{\theta})]^{-1}\right) \quad \text{as } n \rightarrow \infty.$$

This gives

$$\begin{aligned}\pi_{\psi}(\psi|\mathbf{x}) &\propto \pi_{\theta}\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right) \exp\left\{n\ell_n\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right)\right\} \\ &\approx \pi_{\theta}(\hat{\theta}) \exp\left\{n\ell_n(\hat{\theta}) + \frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{\frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{-\frac{[-\ell_n''(\hat{\theta})]}{2}(\psi - 0)^2\right\},\end{aligned}$$

showing that

$$\psi|\mathbf{x} \sim N\left(0, [-\ell_n''(\hat{\theta})]^{-1}\right) \quad \text{as } n \rightarrow \infty.$$

This gives

$$\begin{aligned}\pi_{\psi}(\psi|\mathbf{x}) &\propto \pi_{\theta}\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right) \exp\left\{n\ell_n\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right)\right\} \\ &\approx \pi_{\theta}(\hat{\theta}) \exp\left\{n\ell_n(\hat{\theta}) + \frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{\frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{-\frac{[-\ell_n''(\hat{\theta})]}{2}(\psi - 0)^2\right\},\end{aligned}$$

showing that

$$\psi|\mathbf{x} \sim N\left(0, [-\ell_n''(\hat{\theta})]^{-1}\right) \quad \text{as } n \rightarrow \infty.$$

This gives

$$\begin{aligned}\pi_{\psi}(\psi|\mathbf{x}) &\propto \pi_{\theta}\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right) \exp\left\{n\ell_n\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right)\right\} \\ &\approx \pi_{\theta}(\hat{\theta}) \exp\left\{n\ell_n(\hat{\theta}) + \frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{\frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{-\frac{[-\ell_n''(\hat{\theta})]}{2}(\psi - 0)^2\right\},\end{aligned}$$

showing that

$$\psi|\mathbf{x} \sim N\left(0, [-\ell_n''(\hat{\theta})]^{-1}\right) \quad \text{as } n \rightarrow \infty.$$

This gives

$$\begin{aligned}\pi_{\psi}(\psi|\mathbf{x}) &\propto \pi_{\theta}\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right) \exp\left\{n\ell_n\left(\hat{\theta} + \frac{\psi}{\sqrt{n}}\right)\right\} \\ &\approx \pi_{\theta}(\hat{\theta}) \exp\left\{n\ell_n(\hat{\theta}) + \frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{\frac{1}{2}\ell_n''(\hat{\theta})\psi^2\right\} \\ &\propto \exp\left\{-\frac{[-\ell_n''(\hat{\theta})]}{2}(\psi - 0)^2\right\},\end{aligned}$$

showing that

$$\psi|\mathbf{x} \sim N\left(0, [-\ell_n''(\hat{\theta})]^{-1}\right) \quad \text{as } n \rightarrow \infty.$$

The result on the last slide gives

$$\text{Var} \left\{ \sqrt{n}(\theta - \hat{\theta}) \right\} = \left[-\ell''_n(\hat{\theta}) \right]^{-1}.$$

Multiplying the term inside $\{\}$ by $\sqrt{-\ell''_n(\hat{\theta})}$ gives

$$\text{Var} \left\{ \sqrt{-\ell''_n(\hat{\theta})} \times \sqrt{n}(\theta - \hat{\theta}) \right\} = -\ell''_n(\hat{\theta}) \times \text{Var} \left\{ \sqrt{n}(\theta - \hat{\theta}) \right\}, \text{ i.e.}$$

$$\text{Var} \left\{ \sqrt{-n\ell''_n(\hat{\theta})}(\theta - \hat{\theta}) \right\} = -\ell''_n(\hat{\theta}) \times \left[-\ell''_n(\hat{\theta}) \right]^{-1}, \quad \text{i.e.}$$

$$\text{Var} \left\{ \sqrt{J(\hat{\theta})}(\theta - \hat{\theta}) \right\} = 1.$$

The result on the last slide gives

$$\text{Var} \left\{ \sqrt{n}(\theta - \hat{\theta}) \right\} = \left[-\ell''_n(\hat{\theta}) \right]^{-1}.$$

Multiplying the term inside $\{\}$ by $\sqrt{-\ell''_n(\hat{\theta})}$ gives

$$\text{Var} \left\{ \sqrt{-\ell''_n(\hat{\theta})} \times \sqrt{n}(\theta - \hat{\theta}) \right\} = -\ell''_n(\hat{\theta}) \times \text{Var} \left\{ \sqrt{n}(\theta - \hat{\theta}) \right\}, \text{ i.e.}$$

$$\text{Var} \left\{ \sqrt{-n\ell''_n(\hat{\theta})}(\theta - \hat{\theta}) \right\} = -\ell''_n(\hat{\theta}) \times \left[-\ell''_n(\hat{\theta}) \right]^{-1}, \quad \text{i.e.}$$

$$\text{Var} \left\{ \sqrt{J(\hat{\theta})}(\theta - \hat{\theta}) \right\} = 1.$$

The result on the last slide gives

$$\text{Var} \left\{ \sqrt{n}(\theta - \hat{\theta}) \right\} = \left[-\ell''_n(\hat{\theta}) \right]^{-1}.$$

Multiplying the term inside $\{\}$ by $\sqrt{-\ell''_n(\hat{\theta})}$ gives

$$\text{Var} \left\{ \sqrt{-\ell''_n(\hat{\theta})} \times \sqrt{n}(\theta - \hat{\theta}) \right\} = -\ell''_n(\hat{\theta}) \times \text{Var} \left\{ \sqrt{n}(\theta - \hat{\theta}) \right\}, \text{ i.e.}$$

$$\text{Var} \left\{ \sqrt{-n\ell''_n(\hat{\theta})}(\theta - \hat{\theta}) \right\} = -\ell''_n(\hat{\theta}) \times \left[-\ell''_n(\hat{\theta}) \right]^{-1}, \quad \text{i.e.}$$

$$\text{Var} \left\{ \sqrt{J(\hat{\theta})}(\theta - \hat{\theta}) \right\} = 1.$$

The result on the last slide gives

$$\text{Var} \left\{ \sqrt{n}(\theta - \hat{\theta}) \right\} = \left[-\ell''_n(\hat{\theta}) \right]^{-1}.$$

Multiplying the term inside $\{\}$ by $\sqrt{-\ell''_n(\hat{\theta})}$ gives

$$\text{Var} \left\{ \sqrt{-\ell''_n(\hat{\theta})} \times \sqrt{n}(\theta - \hat{\theta}) \right\} = -\ell''_n(\hat{\theta}) \times \text{Var} \left\{ \sqrt{n}(\theta - \hat{\theta}) \right\}, \text{ i.e.}$$

$$\text{Var} \left\{ \sqrt{-n\ell''_n(\hat{\theta})}(\theta - \hat{\theta}) \right\} = -\ell''_n(\hat{\theta}) \times \left[-\ell''_n(\hat{\theta}) \right]^{-1}, \quad \text{i.e.}$$

$$\text{Var} \left\{ \sqrt{J(\hat{\theta})}(\theta - \hat{\theta}) \right\} = 1.$$

Thus, we have the equivalent result

$$\sqrt{J(\hat{\theta})} (\theta - \hat{\theta}) | \mathbf{x} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty.$$

Dividing by $\sqrt{J(\hat{\theta})}$ and adding $\hat{\theta}$ also gives

$$\theta | \mathbf{x} \sim N(\hat{\theta}, J(\hat{\theta})^{-1}).$$

MEMORISE!

Thus, we have the equivalent result

$$\sqrt{J(\hat{\theta})} (\theta - \hat{\theta}) \big| \mathbf{x} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty.$$

Dividing by $\sqrt{J(\hat{\theta})}$ and adding $\hat{\theta}$ also gives

$$\theta | \mathbf{x} \sim N(\hat{\theta}, J(\hat{\theta})^{-1}).$$

MEMORISE!

Thus, we have the equivalent result

$$\sqrt{J(\hat{\theta})} (\theta - \hat{\theta}) \big| \mathbf{x} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty.$$

Dividing by $\sqrt{J(\hat{\theta})}$ and adding $\hat{\theta}$ also gives

$$\theta | \mathbf{x} \sim N(\hat{\theta}, J(\hat{\theta})^{-1}).$$

MEMORISE!

Thus, we have the equivalent result

$$\sqrt{J(\hat{\theta})} (\theta - \hat{\theta}) \big| \mathbf{x} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty.$$

Dividing by $\sqrt{J(\hat{\theta})}$ and adding $\hat{\theta}$ also gives

$$\theta | \mathbf{x} \sim N(\hat{\theta}, J(\hat{\theta})^{-1}).$$

MEMORISE!

Example 3.14

Suppose we have a random sample from a distribution with probability density function

$$f(x|\theta) = \frac{2x e^{-x^2/\theta}}{\theta}, \quad x > 0, \theta > 0.$$

Determine the asymptotic posterior distribution for θ . Note that from Example 3.11 we have

$$\frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i^2,$$

$$J(\theta) = -\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = -\frac{n}{\theta^2} + \frac{2}{\theta^3} \sum_{i=1}^n x_i^2 = \frac{n}{\theta^3} \left(-\theta + \frac{2}{n} \sum_{i=1}^n x_i^2 \right).$$

Solution to Example 3.14 (1/2)

The asymptotic posterior distribution is given by

$$\theta | \mathbf{x} \sim N \left(\hat{\theta}, J(\hat{\theta})^{-1} \right).$$

First, let's find $\hat{\theta}$. Now

$$\frac{\partial}{\partial \theta} \log f(\mathbf{x} | \theta) = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i^2;$$

Setting equal to zero and solving for $\theta = \hat{\theta}$ gives

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i^2 = \overline{x^2}.$$

Solution to Example 3.14 (1/2)

The asymptotic posterior distribution is given by

$$\theta | \mathbf{x} \sim N \left(\hat{\theta}, J(\hat{\theta})^{-1} \right).$$

First, let's find $\hat{\theta}$. Now

$$\frac{\partial}{\partial \theta} \log f(\mathbf{x} | \theta) = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i^2;$$

Setting equal to zero and solving for $\theta = \hat{\theta}$ gives

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i^2 = \overline{x^2}.$$

Solution to Example 3.14 (1/2)

The asymptotic posterior distribution is given by

$$\theta | \mathbf{x} \sim N \left(\hat{\theta}, J(\hat{\theta})^{-1} \right).$$

First, let's find $\hat{\theta}$. Now

$$\frac{\partial}{\partial \theta} \log f(\mathbf{x} | \theta) = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i^2;$$

Setting equal to zero and solving for $\theta = \hat{\theta}$ gives

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i^2 = \overline{x^2}.$$

Solution to Example 3.14 (2/2)

Also,

$$\begin{aligned}J(\hat{\theta}) &= \frac{n}{\hat{\theta}^3} \left(-\hat{\theta} + \frac{2}{n} \sum_{i=1}^n x_i^2 \right) \\&= \frac{n}{(\overline{x^2})^3} \left(-\overline{x^2} + 2\overline{x^2} \right) \\&= \frac{n}{(\overline{x^2})^3} \overline{x^2} = \frac{n}{(\overline{x^2})^2}.\end{aligned}$$

Therefore, for large n , the (approximate) posterior distribution for θ is

$$\theta | \mathbf{x} \sim N \left(\overline{x^2}, \frac{1}{n} (\overline{x^2})^2 \right).$$

Solution to Example 3.14 (2/2)

Also,

$$\begin{aligned}J(\hat{\theta}) &= \frac{n}{\hat{\theta}^3} \left(-\hat{\theta} + \frac{2}{n} \sum_{i=1}^n x_i^2 \right) \\&= \frac{n}{(\overline{x^2})^3} \left(-\overline{x^2} + 2\overline{x^2} \right) \\&= \frac{n}{(\overline{x^2})^3} \overline{x^2} = \frac{n}{(\overline{x^2})^2}.\end{aligned}$$

Therefore, for large n , the (approximate) posterior distribution for θ is

$$\theta | \mathbf{x} \sim N \left(\overline{x^2}, \frac{1}{n} (\overline{x^2})^2 \right).$$

Solution to Example 3.14 (2/2)

Also,

$$\begin{aligned}J(\hat{\theta}) &= \frac{n}{\hat{\theta}^3} \left(-\hat{\theta} + \frac{2}{n} \sum_{i=1}^n x_i^2 \right) \\&= \frac{n}{(\overline{x^2})^3} \left(-\overline{x^2} + 2\overline{x^2} \right) \\&= \frac{n}{(\overline{x^2})^3} \overline{x^2} = \frac{n}{(\overline{x^2})^2}.\end{aligned}$$

Therefore, for large n , the (approximate) posterior distribution for θ is

$$\theta | \mathbf{x} \sim N \left(\overline{x^2}, \frac{1}{n} (\overline{x^2})^2 \right).$$

Solution to Example 3.14 (2/2)

Also,

$$\begin{aligned}J(\hat{\theta}) &= \frac{n}{\hat{\theta}^3} \left(-\hat{\theta} + \frac{2}{n} \sum_{i=1}^n x_i^2 \right) \\&= \frac{n}{(\overline{x^2})^3} \left(-\overline{x^2} + 2\overline{x^2} \right) \\&= \frac{n}{(\overline{x^2})^3} \overline{x^2} = \frac{n}{(\overline{x^2})^2}.\end{aligned}$$

Therefore, for large n , the (approximate) posterior distribution for θ is

$$\theta | \mathbf{x} \sim N \left(\overline{x^2}, \frac{1}{n} (\overline{x^2})^2 \right).$$

Solution to Example 3.14 (2/2)

Also,

$$\begin{aligned}J(\hat{\theta}) &= \frac{n}{\hat{\theta}^3} \left(-\hat{\theta} + \frac{2}{n} \sum_{i=1}^n x_i^2 \right) \\&= \frac{n}{(\overline{x^2})^3} \left(-\overline{x^2} + 2\overline{x^2} \right) \\&= \frac{n}{(\overline{x^2})^3} \overline{x^2} = \frac{n}{(\overline{x^2})^2}.\end{aligned}$$

Therefore, for large n , the (approximate) posterior distribution for θ is

$$\theta | \mathbf{x} \sim N \left(\overline{x^2}, \frac{1}{n} (\overline{x^2})^2 \right).$$

Solution to Example 3.14 (2/2)

Also,

$$\begin{aligned}J(\hat{\theta}) &= \frac{n}{\hat{\theta}^3} \left(-\hat{\theta} + \frac{2}{n} \sum_{i=1}^n x_i^2 \right) \\&= \frac{n}{(\overline{x^2})^3} \left(-\overline{x^2} + 2\overline{x^2} \right) \\&= \frac{n}{(\overline{x^2})^3} \overline{x^2} = \frac{n}{(\overline{x^2})^2}.\end{aligned}$$

Therefore, for large n , the (approximate) posterior distribution for θ is

$$\theta|\mathbf{x} \sim N\left(\overline{x^2}, \frac{1}{n} (\overline{x^2})^2\right).$$

Example 3.15

Suppose we have a random sample from an exponential distribution, that is, $X_i|\theta \sim \text{Exp}(\theta)$, $i = 1, 2, \dots, n$ (independent).

Determine the asymptotic posterior distribution for θ .

Note that from Example 3.12 we have

$$\begin{aligned}\frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) &= \frac{n}{\theta} - n\bar{x}, \\ J(\theta) &= -\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = \frac{n}{\theta^2}.\end{aligned}$$

Solution to Example 3.15 (1/1)

We have

$$\begin{aligned}\frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = 0 &\implies \hat{\theta} = \frac{1}{\bar{x}} \\ &\implies J(\hat{\theta}) = \frac{n}{\left(\frac{1}{\bar{x}}\right)^2} = n\bar{x}^2 \\ &\implies J(\hat{\theta})^{-1} = \frac{1}{n\bar{x}^2}.\end{aligned}$$

Therefore, for large n , the (approximate) posterior distribution for θ is

$$\theta|\mathbf{x} \sim N\left(\frac{1}{\bar{x}}, \frac{1}{n\bar{x}^2}\right).$$

Solution to Example 3.15 (1/1)

We have

$$\begin{aligned}\frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = 0 &\implies \hat{\theta} = \frac{1}{\bar{x}} \\ &\implies J(\hat{\theta}) = \frac{n}{\left(\frac{1}{\bar{x}}\right)^2} = n\bar{x}^2 \\ &\implies J(\hat{\theta})^{-1} = \frac{1}{n\bar{x}^2}.\end{aligned}$$

Therefore, for large n , the (approximate) posterior distribution for θ is

$$\theta|\mathbf{x} \sim N\left(\frac{1}{\bar{x}}, \frac{1}{n\bar{x}^2}\right).$$

Example 3.15

Recall that, assuming a vague prior distribution, the posterior distribution is a $Ga(n, n\bar{x})$ distribution, with mean $1/\bar{x}$ and variance $1/(n\bar{x}^2)$.

The Central Limit Theorem tells us that, for large n , the gamma distribution tends to a normal distribution, matched, of course, for mean and variance.

Therefore, we have shown that, for large n , the asymptotic posterior distribution is the same as the posterior distribution under vague prior knowledge. Not a surprising result!

Example 3.16

Suppose we have a random sample from a $N(\mu, 1/\tau)$ distribution (with τ known). Determine the asymptotic posterior distribution for μ . Note that from Example 3.13 we have

$$\begin{aligned}\frac{\partial}{\partial \mu} \log f(\mathbf{x}|\mu) &= n\tau(\bar{x} - \mu), \\ J(\mu) &= -\frac{\partial^2}{\partial \mu^2} \log f(\mathbf{x}|\mu) = n\tau.\end{aligned}$$

Solution to Example 3.16 (1/1)

We have

$$\begin{aligned}\frac{\partial}{\partial \mu} \log f(\mathbf{x}|\mu) = 0 &\implies \hat{\mu} = \bar{x} \\ &\implies J(\hat{\mu}) = n\tau \\ &\implies J(\hat{\mu})^{-1} = \frac{1}{n\tau}.\end{aligned}$$

Therefore, for large n , the (approximate) posterior distribution for μ is

$$\mu|\mathbf{x} \sim N\left(\bar{x}, \frac{1}{n\tau}\right).$$

Example 3.16

Again, we have shown that the asymptotic posterior distribution is the same as the posterior distribution under vague prior knowledge.

Using a random sample from a $\text{Bin}(k, \theta)$ (with k known), determine the posterior distribution for θ assuming

- (i) vague prior knowledge;
- (ii) the Jeffreys prior distribution;
- (iii) a very large sample.

The **conjugate prior distribution** is a $Beta(g, h)$ distribution. Using this prior distribution, the posterior density is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \theta^{g-1}(1-\theta)^{h-1} \times \prod_{i=1}^n \theta^{x_i}(1-\theta)^{k-x_i}, \quad 0 < \theta < 1\end{aligned}$$

$$\propto \theta^{g+n\bar{x}-1}(1-\theta)^{h+nk-n\bar{x}-1}, \quad 0 < \theta < 1$$

i.e. $\theta|\mathbf{x} \sim Beta(G = g + n\bar{x}, H = h + nk - n\bar{x})$.

The **conjugate prior distribution** is a $Beta(g, h)$ distribution. Using this prior distribution, the posterior density is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \theta^{g-1}(1-\theta)^{h-1} \times \prod_{i=1}^n \theta^{x_i}(1-\theta)^{k-x_i}, \quad 0 < \theta < 1 \\ &\propto \theta^{g+n\bar{x}-1}(1-\theta)^{h+nk-n\bar{x}-1}, \quad 0 < \theta < 1\end{aligned}$$

i.e. $\theta|\mathbf{x} \sim Beta(G = g + n\bar{x}, H = h + nk - n\bar{x})$.

The **conjugate prior distribution** is a $Beta(g, h)$ distribution. Using this prior distribution, the posterior density is

$$\begin{aligned}\pi(\theta|\mathbf{x}) &\propto \pi(\theta) f(\mathbf{x}|\theta) \\ &\propto \theta^{g-1}(1-\theta)^{h-1} \times \prod_{i=1}^n \theta^{x_i}(1-\theta)^{k-x_i}, \quad 0 < \theta < 1 \\ &\propto \theta^{g+n\bar{x}-1}(1-\theta)^{h+nk-n\bar{x}-1}, \quad 0 < \theta < 1 \\ &\text{i.e. } \theta|\mathbf{x} \sim Beta(G = g + n\bar{x}, H = h + nk - n\bar{x}).\end{aligned}$$

We represent **vague prior information** by taking a conjugate prior distribution with large variance.

As the beta distribution restricts values to the range $(0,1)$, there is a finite upper limit to the variance.

Intuitively, the maximum variance is achieved when the probability density is pushed to the extremes of the range, that is, equal mass at $\theta = 0$ and $\theta = 1$ – this distribution is obtained in the limit $g \rightarrow 0$ and $h \rightarrow 0$.

Thus we will take this limit to represent vague prior information. Hence the posterior distribution under vague prior information is

$$\theta | \mathbf{x} \sim \text{Beta}(n\bar{x}, nk - n\bar{x}).$$

We represent **vague prior information** by taking a conjugate prior distribution with large variance.

As the beta distribution restricts values to the range $(0,1)$, there is a finite upper limit to the variance.

Intuitively, the maximum variance is achieved when the probability density is pushed to the extremes of the range, that is, equal mass at $\theta = 0$ and $\theta = 1$ – this distribution is obtained in the limit $g \rightarrow 0$ and $h \rightarrow 0$.

Thus we will take this limit to represent vague prior information. Hence the posterior distribution under vague prior information is

$$\theta | \mathbf{x} \sim \text{Beta}(n\bar{x}, nk - n\bar{x}).$$

We represent **vague prior information** by taking a conjugate prior distribution with large variance.

As the beta distribution restricts values to the range $(0,1)$, there is a finite upper limit to the variance.

Intuitively, the maximum variance is achieved when the probability density is pushed to the extremes of the range, that is, equal mass at $\theta = 0$ and $\theta = 1$ – this distribution is obtained in the limit $g \rightarrow 0$ and $h \rightarrow 0$.

Thus we will take this limit to represent vague prior information. Hence the posterior distribution under vague prior information is

$$\theta | \mathbf{x} \sim \text{Beta}(n\bar{x}, nk - n\bar{x}).$$

We represent **vague prior information** by taking a conjugate prior distribution with large variance.

As the beta distribution restricts values to the range $(0,1)$, there is a finite upper limit to the variance.

Intuitively, the maximum variance is achieved when the probability density is pushed to the extremes of the range, that is, equal mass at $\theta = 0$ and $\theta = 1$ – this distribution is obtained in the limit $g \rightarrow 0$ and $h \rightarrow 0$.

Thus we will take this limit to represent vague prior information. Hence the posterior distribution under vague prior information is

$$\theta | \mathbf{x} \sim \text{Beta}(n\bar{x}, nk - n\bar{x}).$$

We represent **vague prior information** by taking a conjugate prior distribution with large variance.

As the beta distribution restricts values to the range $(0,1)$, there is a finite upper limit to the variance.

Intuitively, the maximum variance is achieved when the probability density is pushed to the extremes of the range, that is, equal mass at $\theta = 0$ and $\theta = 1$ – this distribution is obtained in the limit $g \rightarrow 0$ and $h \rightarrow 0$.

Thus we will take this limit to represent vague prior information. Hence the posterior distribution under vague prior information is

$$\theta | \mathbf{x} \sim \text{Beta}(n\bar{x}, nk - n\bar{x}).$$

The **Jeffreys prior distribution** is

$$\pi(\theta) \propto \sqrt{I(\theta)}.$$

Now

$$\begin{aligned} f(\mathbf{x}|\theta) &\propto \prod_{i=1}^n \theta^{x_i} (1 - \theta)^{k - x_i} \\ &\propto \theta^{n\bar{x}} (1 - \theta)^{kn - n\bar{x}}. \end{aligned}$$

The **Jeffreys prior distribution** is

$$\pi(\theta) \propto \sqrt{I(\theta)}.$$

Now

$$\begin{aligned} f(\mathbf{x}|\theta) &\propto \prod_{i=1}^n \theta^{x_i} (1 - \theta)^{k - x_i} \\ &\propto \theta^{n\bar{x}} (1 - \theta)^{kn - n\bar{x}}. \end{aligned}$$

Therefore

$$\log f(\mathbf{x}|\theta) = \text{constant} + n\bar{x} \log \theta + n(k - \bar{x}) \log(1 - \theta)$$

$$\Rightarrow \frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = \frac{n\bar{x}}{\theta} - \frac{n(k - \bar{x})}{1 - \theta}$$

$$\Rightarrow \frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = -\frac{n\bar{x}}{\theta^2} - \frac{n(k - \bar{x})}{(1 - \theta)^2}$$

$$\begin{aligned} \Rightarrow I(\theta) &= E_{\mathbf{X}|\theta} \left[-\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{X}|\theta) \right] \\ &= \frac{nE_{\mathbf{X}|\theta}(\bar{X})}{\theta^2} + \frac{n[k - E_{\mathbf{X}|\theta}(\bar{X})]}{(1 - \theta)^2}. \end{aligned}$$

Therefore

$$\log f(\mathbf{x}|\theta) = \text{constant} + n\bar{x} \log \theta + n(k - \bar{x}) \log(1 - \theta)$$

$$\Rightarrow \frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = \frac{n\bar{x}}{\theta} - \frac{n(k - \bar{x})}{1 - \theta}$$

$$\Rightarrow \frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = -\frac{n\bar{x}}{\theta^2} - \frac{n(k - \bar{x})}{(1 - \theta)^2}$$

$$\begin{aligned} \Rightarrow I(\theta) &= E_{\mathbf{X}|\theta} \left[-\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{X}|\theta) \right] \\ &= \frac{nE_{\mathbf{X}|\theta}(\bar{X})}{\theta^2} + \frac{n[k - E_{\mathbf{X}|\theta}(\bar{X})]}{(1 - \theta)^2}. \end{aligned}$$

Therefore

$$\log f(\mathbf{x}|\theta) = \text{constant} + n\bar{x} \log \theta + n(k - \bar{x}) \log(1 - \theta)$$

$$\Rightarrow \frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = \frac{n\bar{x}}{\theta} - \frac{n(k - \bar{x})}{1 - \theta}$$

$$\Rightarrow \frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = -\frac{n\bar{x}}{\theta^2} - \frac{n(k - \bar{x})}{(1 - \theta)^2}$$

$$\begin{aligned} \Rightarrow I(\theta) &= E_{\mathbf{X}|\theta} \left[-\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{X}|\theta) \right] \\ &= \frac{nE_{\mathbf{X}|\theta}(\bar{X})}{\theta^2} + \frac{n[k - E_{\mathbf{X}|\theta}(\bar{X})]}{(1 - \theta)^2}. \end{aligned}$$

Therefore

$$\log f(\mathbf{x}|\theta) = \text{constant} + n\bar{x} \log \theta + n(k - \bar{x}) \log(1 - \theta)$$

$$\Rightarrow \frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = \frac{n\bar{x}}{\theta} - \frac{n(k - \bar{x})}{1 - \theta}$$

$$\Rightarrow \frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = -\frac{n\bar{x}}{\theta^2} - \frac{n(k - \bar{x})}{(1 - \theta)^2}$$

$$\begin{aligned} \Rightarrow I(\theta) &= E_{\mathbf{x}|\theta} \left[-\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{X}|\theta) \right] \\ &= \frac{nE_{\mathbf{x}|\theta}(\bar{X})}{\theta^2} + \frac{n[k - E_{\mathbf{x}|\theta}(\bar{X})]}{(1 - \theta)^2}. \end{aligned}$$

Therefore

$$\log f(\mathbf{x}|\theta) = \text{constant} + n\bar{x} \log \theta + n(k - \bar{x}) \log(1 - \theta)$$

$$\Rightarrow \frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = \frac{n\bar{x}}{\theta} - \frac{n(k - \bar{x})}{1 - \theta}$$

$$\Rightarrow \frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = -\frac{n\bar{x}}{\theta^2} - \frac{n(k - \bar{x})}{(1 - \theta)^2}$$

$$\begin{aligned} \Rightarrow I(\theta) &= E_{\mathbf{X}|\theta} \left[-\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{X}|\theta) \right] \\ &= \frac{nE_{\mathbf{X}|\theta}(\bar{X})}{\theta^2} + \frac{n[k - E_{\mathbf{X}|\theta}(\bar{X})]}{(1 - \theta)^2}. \end{aligned}$$

Now $E_{\mathbf{X}|\theta}(\bar{X}) = E_{X|\theta}(X) = k\theta$. Therefore

$$I(\theta) = \frac{nk}{\theta} + \frac{nk}{1-\theta} = \frac{nk}{\theta(1-\theta)}.$$

Hence the Jeffreys prior for this model is

$$\begin{aligned}\pi(\theta) &\propto \sqrt{I(\theta)} \\ &\propto \sqrt{\frac{nk}{\theta(1-\theta)}}, \quad 0 < \theta < 1 \\ &\propto \theta^{-1/2}(1-\theta)^{-1/2}, \quad 0 < \theta < 1.\end{aligned}$$

Now $E_{\mathbf{X}|\theta}(\bar{X}) = E_{X|\theta}(X) = k\theta$. Therefore

$$I(\theta) = \frac{nk}{\theta} + \frac{nk}{1-\theta} = \frac{nk}{\theta(1-\theta)}.$$

Hence the Jeffreys prior for this model is

$$\begin{aligned}\pi(\theta) &\propto \sqrt{I(\theta)} \\ &\propto \sqrt{\frac{nk}{\theta(1-\theta)}}, \quad 0 < \theta < 1 \\ &\propto \theta^{-1/2}(1-\theta)^{-1/2}, \quad 0 < \theta < 1.\end{aligned}$$

Now $E_{\mathbf{X}|\theta}(\bar{X}) = E_{X|\theta}(X) = k\theta$. Therefore

$$I(\theta) = \frac{nk}{\theta} + \frac{nk}{1-\theta} = \frac{nk}{\theta(1-\theta)}.$$

Hence the Jeffreys prior for this model is

$$\begin{aligned}\pi(\theta) &\propto \sqrt{I(\theta)} \\ &\propto \sqrt{\frac{nk}{\theta(1-\theta)}}, \quad 0 < \theta < 1 \\ &\propto \theta^{-1/2}(1-\theta)^{-1/2}, \quad 0 < \theta < 1.\end{aligned}$$

Now $E_{\mathbf{X}|\theta}(\bar{X}) = E_{X|\theta}(X) = k\theta$. Therefore

$$I(\theta) = \frac{nk}{\theta} + \frac{nk}{1-\theta} = \frac{nk}{\theta(1-\theta)}.$$

Hence the Jeffreys prior for this model is

$$\begin{aligned}\pi(\theta) &\propto \sqrt{I(\theta)} \\ &\propto \sqrt{\frac{nk}{\theta(1-\theta)}}, \quad 0 < \theta < 1 \\ &\propto \theta^{-1/2}(1-\theta)^{-1/2}, \quad 0 < \theta < 1.\end{aligned}$$

Now $E_{\mathbf{X}|\theta}(\bar{X}) = E_{X|\theta}(X) = k\theta$. Therefore

$$I(\theta) = \frac{nk}{\theta} + \frac{nk}{1-\theta} = \frac{nk}{\theta(1-\theta)}.$$

Hence the Jeffreys prior for this model is

$$\begin{aligned}\pi(\theta) &\propto \sqrt{I(\theta)} \\ &\propto \sqrt{\frac{nk}{\theta(1-\theta)}}, \quad 0 < \theta < 1 \\ &\propto \theta^{-1/2}(1-\theta)^{-1/2}, \quad 0 < \theta < 1.\end{aligned}$$

Now $E_{\mathbf{X}|\theta}(\bar{X}) = E_{X|\theta}(X) = k\theta$. Therefore

$$I(\theta) = \frac{nk}{\theta} + \frac{nk}{1-\theta} = \frac{nk}{\theta(1-\theta)}.$$

Hence the Jeffreys prior for this model is

$$\begin{aligned}\pi(\theta) &\propto \sqrt{I(\theta)} \\ &\propto \sqrt{\frac{nk}{\theta(1-\theta)}}, \quad 0 < \theta < 1 \\ &\propto \theta^{-1/2}(1-\theta)^{-1/2}, \quad 0 < \theta < 1.\end{aligned}$$

This is a $Beta(1/2, 1/2)$ prior distribution and so the resulting posterior distribution is

$$\theta|\mathbf{x} \sim Beta(1/2 + n\bar{x}, 1/2 + nk - n\bar{x}).$$

The **asymptotic posterior distribution** (as $n \rightarrow \infty$) is

$$\theta|\mathbf{x} \sim N\left(\hat{\theta}, J(\hat{\theta})^{-1}\right),$$

where

$$J(\theta) = -\frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}|\theta) = \frac{n\bar{x}}{\theta^2} + \frac{n(k - \bar{x})}{(1 - \theta)^2}.$$

Now

$$\frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = 0 \quad \Rightarrow \quad \frac{n\bar{x}}{\hat{\theta}} - \frac{n(k - \bar{x})}{1 - \hat{\theta}} = 0$$

$$\Rightarrow \quad \hat{\theta} = \frac{\bar{x}}{k}$$

$$\Rightarrow \quad J(\hat{\theta}) = \frac{nk^3}{\bar{x}(k - \bar{x})}$$

$$\Rightarrow \quad J(\hat{\theta})^{-1} = \frac{\bar{x}(k - \bar{x})}{nk^3}.$$

Now

$$\frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = 0 \quad \Rightarrow \quad \frac{n\bar{x}}{\hat{\theta}} - \frac{n(k - \bar{x})}{1 - \hat{\theta}} = 0$$

$$\Rightarrow \quad \hat{\theta} = \frac{\bar{x}}{k}$$

$$\Rightarrow \quad J(\hat{\theta}) = \frac{nk^3}{\bar{x}(k - \bar{x})}$$

$$\Rightarrow \quad J(\hat{\theta})^{-1} = \frac{\bar{x}(k - \bar{x})}{nk^3}.$$

Now

$$\frac{\partial}{\partial \theta} \log f(\mathbf{x}|\theta) = 0 \quad \implies \quad \frac{n\bar{x}}{\hat{\theta}} - \frac{n(k - \bar{x})}{1 - \hat{\theta}} = 0$$

$$\implies \quad \hat{\theta} = \frac{\bar{x}}{k}$$

$$\implies \quad J(\hat{\theta}) = \frac{nk^3}{\bar{x}(k - \bar{x})}$$

$$\implies \quad J(\hat{\theta})^{-1} = \frac{\bar{x}(k - \bar{x})}{nk^3}.$$

Therefore, for large n , the posterior distribution for θ is

$$\theta|\mathbf{x} \sim N\left(\frac{\bar{x}}{k}, \frac{\bar{x}(k - \bar{x})}{nk^3}\right) \quad \text{approximately.}$$