UNIVERSITY OF THESSALY – SCHOOL OF ENGINEERING

DEPARTMENT OF MECHANICAL ENGINEERING

UNCERTAINTY QUANTIFICATION

Exercise 1: Consider a Gaussian di\stribution with mean μ and variance X to be the mathematical model of a physical model/system. Specifically, an output quantity of interest Y~N(μ,X) or, equivalently, the measure of the uncertainty in *y* given that $X = \sigma^2$ is given by the PDF:

$$f\left(y \middle| X, \mu, I\right) = \frac{1}{\sqrt{2\pi X}} \exp\left[-\frac{1}{2X}(y-\mu)^2\right]$$

Given a set of independent observations/data $D = \left(\hat{Y}_1, \hat{Y}_2, ..., \hat{Y}_N\right) = \left\{\hat{Y}_k\right\}_{1\to N}$ we are interested in updating the uncertainty in the variance X of the model. It is assumed that the value of the mean μ is known. Assume a uniform prior distribution for X and derive expressions for:

i. Posterior PDF $f\left(\sigma^2 \middle| \left\{\hat{Y}_k\right\}_{1\to N}, \mu, I\right)$. Note that the posterior PDF follows the inverse gamma distribution $IG(x,a,b)$. What are the values for a, b?

ii. The function $L\left(\sigma^2\right)$

iii. The MVP of $X = \sigma^2$

iv. The uncertainty of $X = \sigma^2$

v. Retain up to the quadratic terms in Taylor Series expansion of $L\left(\sigma^2\right)$ about the most probable value $\hat{\sigma}^2$ and derive the Gaussian asymptotic approximation for the posterior PDF of $f\left(\sigma^2 \middle| \left\{\hat{Y}_k\right\}_{1\to N}, \mu, I\right)$

vi. Compare the posterior PDF with the asymptotic Gaussian posterior PDF for the following values of $N = 1,2,3,4,10,100,1000$. To facilitate comparisons, plot the two posterior PDFs (exact and asymptotic) so that the maximum value of each equals unity.

i.

Prior Distribution: Assuming a uniform prior distribution for X we obtain:

$$f(X|\mu,I)=\begin{cases}\dfrac{1}{X_{\max}-X_{\min}}, & X\in[X_{\min},X_{\max}]\\ 0, & otherwise\end{cases} \quad (1)$$

Posterior Distribution: Using Bayes' theorem, the inference about the value of X, given the mean μ, the data D and the information I, is expressed by:

$$\underbrace{f\left(X\middle|\{\hat{Y}_k\}_{1\to N},\mu,I\right)}_{\text{Posterior}}\propto\underbrace{f\left(\{\hat{Y}_k\}_{1\to N}\middle|X,\mu,I\right)}_{\text{Likelihood}}\underbrace{f\left(X|\mu,I\right)}_{\text{Prior}} \quad (2)$$

Hence, to determine the Posterior PDF we need to elicit an expression for the likelihood. Therefore:

$$f\left(\{\hat{Y}_k\}_{1\to N}\middle|X,\mu,I\right)=f\left(\hat{Y}_1,\hat{Y}_2,...,\hat{Y}_N\middle|X,\mu,I\right)=f\left(\hat{Y}_1\middle|\hat{Y}_2,\hat{Y}_3,...,\hat{Y}_N,X,\mu,I\right)f\left(\hat{Y}_2,\hat{Y}_3,...,\hat{Y}_N\middle|X,\mu,I\right)\Rightarrow$$

$$f\left(\{\hat{Y}_k\}_{1\to N}\middle|X,\mu,I\right)=f\left(\hat{Y}_1\middle|\hat{Y}_2,\hat{Y}_3,...,\hat{Y}_N,X,\mu,I\right)f\left(\hat{Y}_2\middle|\hat{Y}_3,\hat{Y}_4,...,\hat{Y}_N,X,\mu,I\right)\dots f\left(\hat{Y}_N\middle|X,\mu,I\right)$$

But since observations $D=\left(\hat{Y}_1,\hat{Y}_2,...,\hat{Y}_N\right)=\{\hat{Y}_k\}_{1\to N}$ are assumed to be independent the previous expression for the likelihood can be simplified as:

$$f\left(\{\hat{Y}_k\}_{1\to N}\middle|X,\mu,I\right)=f\left(\hat{Y}_1\middle|X,\mu,I\right)f\left(\hat{Y}_2\middle|X,\mu,I\right)\cdots f\left(\hat{Y}_N\middle|X,\mu,I\right)\Rightarrow$$

$$f\left(\{\hat{Y}_k\}_{1\to N}\middle|X,\mu,I\right)=\prod_{k=1}^{N}f\left(\hat{Y}_k\middle|X,\mu,I\right)=\prod_{k=1}^{N}\frac{1}{\sqrt{2\pi X}}\exp\left[-\frac{1}{2X}\left(\hat{Y}_k-\mu\right)^2\right] \quad (3)$$

Taking into consideration equation (1) for the prior PDF as well as the expression (3) for the likelihood, we can derive the following expression for the posterior PDF:

$$f\left(X\middle|\{\hat{Y}_k\}_{1\to N},\mu,I\right)\propto\prod_{k=1}^{N}\frac{1}{\sqrt{2\pi X}}\exp\left[-\frac{1}{2X}\left(\hat{Y}_k-\mu\right)^2\right]$$

$$f\left(X\middle|\{\hat{Y}_k\}_{1\to N},\mu,I\right)\propto\left(\frac{1}{\sqrt{2\pi X}}\right)^N\exp\left[-\frac{1}{2X}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2\right]$$

$$\boxed{f\left(X\middle|\{\hat{Y}_k\}_{1\to N},\mu,I\right)\propto\left(\frac{1}{\sqrt{2\pi}}\right)^N X^{-\frac{N}{2}}\exp\left[-\frac{1}{2X}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2\right] \quad (2)}$$

We now have to show that the above expression is an Inverse Gamma Distribution for X. The Inverse Gamma Distribution is given by:

$$IG(X,a,b) = f(X,a,b) = \frac{b^a}{\Gamma(a)} X^{-a-1} \exp\left(-\frac{b}{X}\right) \quad (4)$$

By comparing (2) and (4) we can verify that the posterior PDF for X is indeed an Inverse Gamma distribution with:

$$b = \frac{1}{2}\sum_{k=1}^{N}\left(\hat{Y}_k - \mu\right)^2$$

And by comparing the exponents of X in the two expressions, $X^{-\frac{N}{2}} = X^{-a-1} \Rightarrow a = \frac{N-2}{2}$.

So to conclude, the posterior distribution of X is given by the following expression:

$$f\left(X \middle| \left\{\hat{Y}_k\right\}_{1\to N}, \mu, I\right) = \frac{\left(\sqrt{\frac{1}{2}\sum_{k=1}^{N}\left(\hat{Y}_k - \mu\right)^2}\right)^{(N-2)}}{\Gamma\left(\frac{N-2}{2}\right)} X^{-N/2} \exp\left[-\frac{1}{2X}\sum_{k=1}^{N}\left(\hat{Y}_k - \mu\right)^2\right] \quad (*)$$

ii.

We now introduce the function $L(X)$ as the minus logarithm of the posterior PDF. Thus:

$$L(X) = -\ln\left(f\left(X \middle| \left\{\hat{Y}_k\right\}_{1\to N}, \mu, I\right)\right) \Rightarrow$$

$$L(X) = -\ln\left(\frac{\left(\sqrt{\frac{1}{2}\sum_{k=1}^{N}\left(\hat{Y}_k - \mu\right)^2}\right)^{(N-2)}}{\Gamma\left(\frac{N-2}{2}\right)} X^{-N/2} \exp\left[-\frac{1}{2X}\sum_{k=1}^{N}\left(\hat{Y}_k - \mu\right)^2\right]\right) \Rightarrow$$

$$L(X) = \frac{N}{2}\ln(X) + \frac{1}{2X}\sum_{k=1}^{N}\left(\hat{Y}_k - \mu\right)^2 + \text{constant terms}$$

iii.

In order to determine the MPV of X, we need to maximize the posterior PDF with respect to X, or equivalently to minimize the formerly introduced fcn $L(X)$ with respect to X. Hence:

$$\hat{X} = X : \left\{ \left.\frac{dL}{dX}\right|_{X=\hat{X}} = 0 \ , \ \left.\frac{d^2L}{dX^2}\right|_{X=\hat{X}} < 0 \right\}$$

$$\frac{dL}{dX}=0 \Rightarrow \frac{N}{2X}-\frac{1}{2X^2}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2=0 \Rightarrow$$

$$\boxed{\hat{X}=\frac{1}{N}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2 \quad (5)}$$

$$\boxed{\hat{\sigma}=\sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2}}$$

iv.

To determine the uncertainty of X, we just have to evaluate the 2nd derivative of $L\left(X\right)$ at the most probable value. Doing so we obtain:

$$\frac{d^2L}{dX^2}=-\frac{N}{2X^2}+\frac{1}{X^3}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2$$

$$\left.\frac{d^2L}{dX^2}\right|_{X=\hat{X}}=-\frac{N}{2\left(\frac{1}{N}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2\right)^2}+\frac{1}{\left(\frac{1}{N}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2\right)^3}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2=$$

$$=\frac{N^3}{\left(\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2\right)^2}\left[1-\frac{1}{2}\right]=\frac{N^3}{2\left(\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2\right)^2}$$

And the measure of uncertainty in X is given by the square root of the inverse of the 2nd derivative of $L\left(X\right)$ evaluated at the MVP. In terms of mathematical expressions, the above statement is written:

$$\boxed{\sqrt{S}=\left(\left.\frac{d^2L}{dX^2}\right|_{X=\hat{X}}\right)^{-1/2} \Rightarrow \sqrt{S}=\sqrt{\frac{2\left(\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2\right)^2}{N^3}} \quad (6)}$$

Thus, the uncertainty in X can be quantified by using these two measures (MVP, $\sqrt{S}$):

$$X \to \hat{X} \pm \sqrt{S}$$

$$\boxed{X \to \quad \frac{1}{N}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2 \pm \sqrt{\frac{2\left(\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2\right)^2}{N^3}}}$$

v.

In order to approximate via Taylor series expansion the function $L(X)$ we recall its expression:

$$\boxed{L(X)=\frac{N}{2}\ln(X)+\frac{1}{2X}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2+\text{constant terms}}$$

$$Taylor:$$

$$L(X)=L\left(\hat{X}\right)+\frac{dL}{dX}\left(X-\hat{X}\right)+\frac{1}{2}\frac{d^2L}{dX^2}\left(X-\hat{X}\right)^2$$

Since, $L(X)=-\ln\left(f\left(X\left|\left\{\hat{Y}_k\right\}_{1\to N},\mu,I\right.\right)\right)$, we can show that

$$f\left(X\left|\left\{\hat{Y}_k\right\}_{1\to N},\mu,I\right.\right)=-\exp\left[L(X)\right]$$

We will obtain the asymptotic expression for the posterior PDF by replacing $L(X)$ in the exponent with the Taylor approximation around the MPV.

$$f\left(X\left|\left\{\hat{Y}_k\right\}_{1\to N},\mu,I\right.\right)=-\exp\left[\underbrace{L\left(\hat{X}\right)}_{\text{constant}}+\underbrace{\left.\frac{dL}{dX}\right|_{X=\hat{X}}}_{=0}\left(X-\hat{X}\right)+\frac{1}{2}\underbrace{\left.\frac{d^2L}{dX^2}\right|_{X=\hat{X}}}_{S^{-1}}\left(X-\hat{X}\right)^2\right]\Rightarrow$$

$$f\left(X\left|\left\{\hat{Y}_k\right\}_{1\to N},\mu,I\right.\right)\propto\exp\left[-\frac{1}{2S}\left(X-\hat{X}\right)^2\right]$$

And by recalling that for any PDF $\int_{-\infty}^{\infty}f\left(X\left|\left\{\hat{Y}_k\right\}_{1\to N},\mu,I\right.\right)dx=1$, it can be shown that:

$$f\left(X\left|\left\{\hat{Y}_k\right\}_{1\to N},\mu,I\right.\right)=\frac{1}{\sqrt{2\pi S}}\exp\left[-\frac{1}{2S}\left(X-\hat{X}\right)^2\right].$$

To sum up:

$$f\left(X\left|\left\{\hat{Y}_k\right\}_{1\to N},\mu,I\right.\right)=\begin{cases}\dfrac{\left(\sqrt{\dfrac{1}{2}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2}\right)^{(N-2)}}{\Gamma\left(\dfrac{N-2}{2}\right)}X^{-N/2}\exp\left[-\dfrac{1}{2X}\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2\right] & \textit{Exact}\\ \dfrac{1}{\sqrt{2\pi S}}\exp\left[-\dfrac{1}{2S}\left(X-\hat{X}\right)^2\right] & \textit{Asymptotic}\end{cases}$$

In order to obtain a better interpretation of these two relationships for the posterior PDF we need to express S and $\hat{X}$ in terms of $\hat{Y}_k$ and μ .

To do so, we assume $\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2$ to be a known constant (which is also quite reasonable) . To further exemplify our expressions we assign:

$$A=\sum_{k=1}^{N}\left(\hat{Y}_k-\mu\right)^2$$

Now it is just a matter of simple algebra to show that:

$$\hat{X}=\frac{A}{N}$$

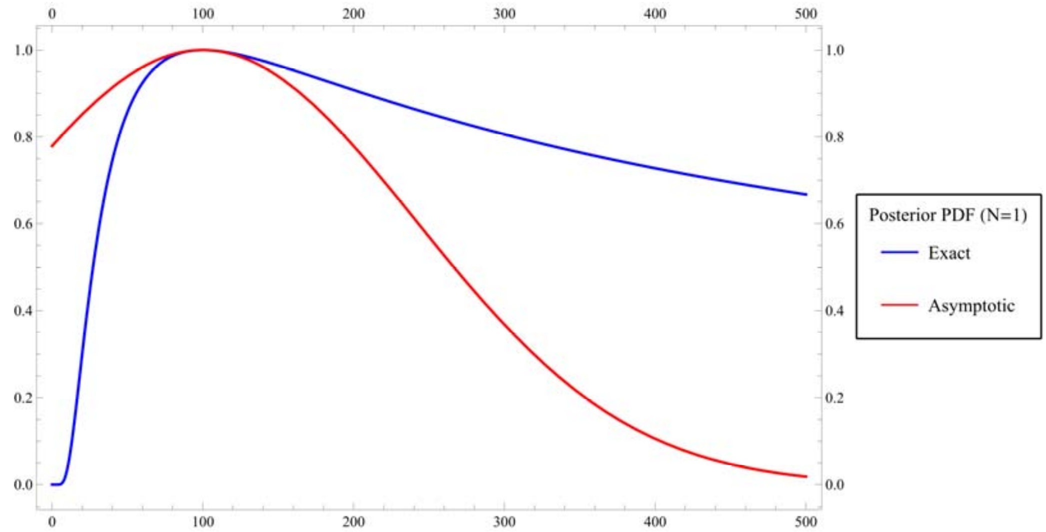
$$S=\frac{2A^2}{N^3}$$

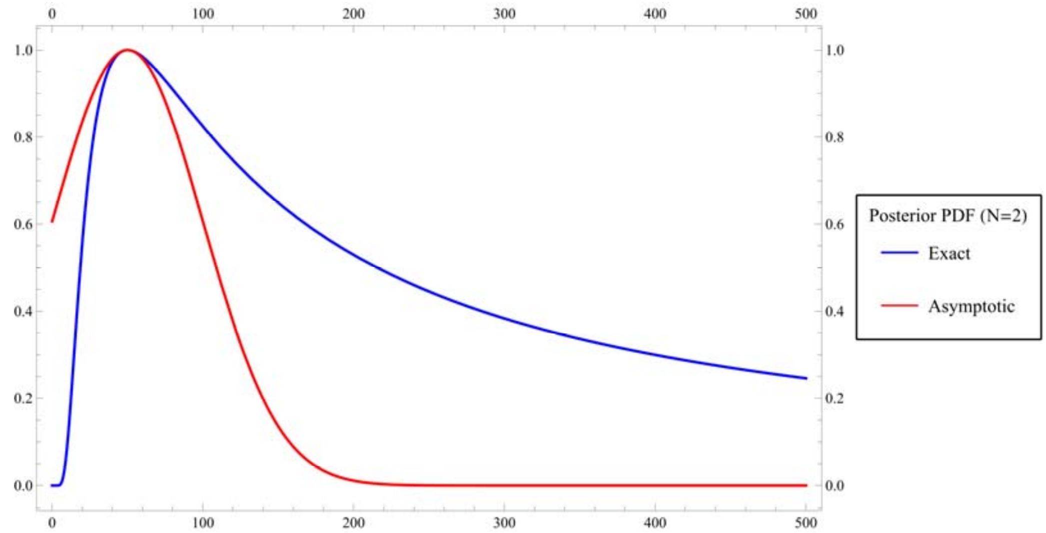
Thus, the expressions for the posterior PDF distribution as derived from the exact and asymptotic approximation can now be written in the following form:

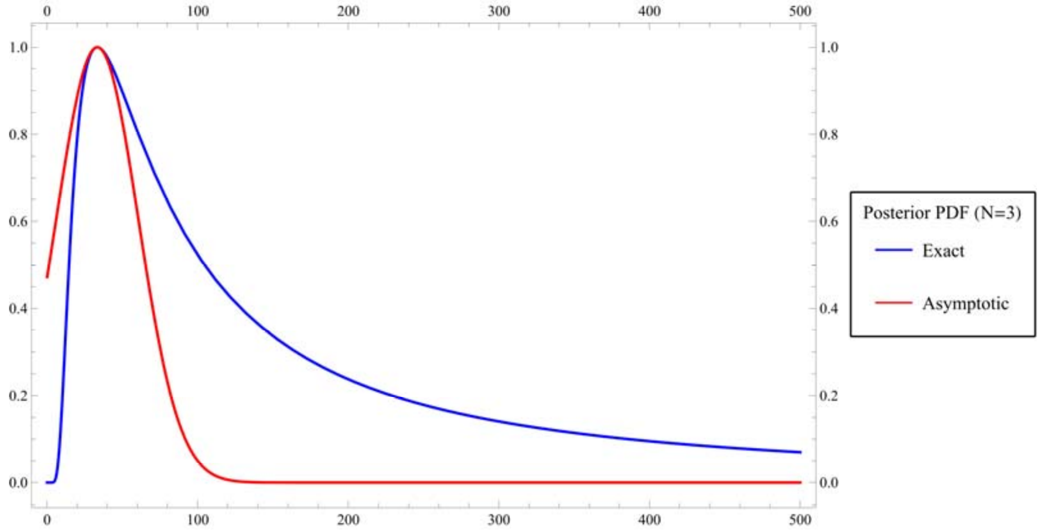
$$f\left(X\left|\left\{\hat{Y}_k\right\}_{1\to N},\mu,I\right.\right)=\begin{cases}\dfrac{\left(\sqrt{\dfrac{1}{2}A}\right)^{(N-2)}}{\Gamma\left(\dfrac{N-2}{2}\right)}X^{-N/2}\exp\left[-\dfrac{1}{2X}A\right] & \textit{Exact}\\ \dfrac{\sqrt{N^3}}{2A\sqrt{\pi}}\exp\left[-\dfrac{N^3}{4A^2}\left(X-\dfrac{A}{N}\right)^2\right] & \textit{Asymptotic}\end{cases}$$

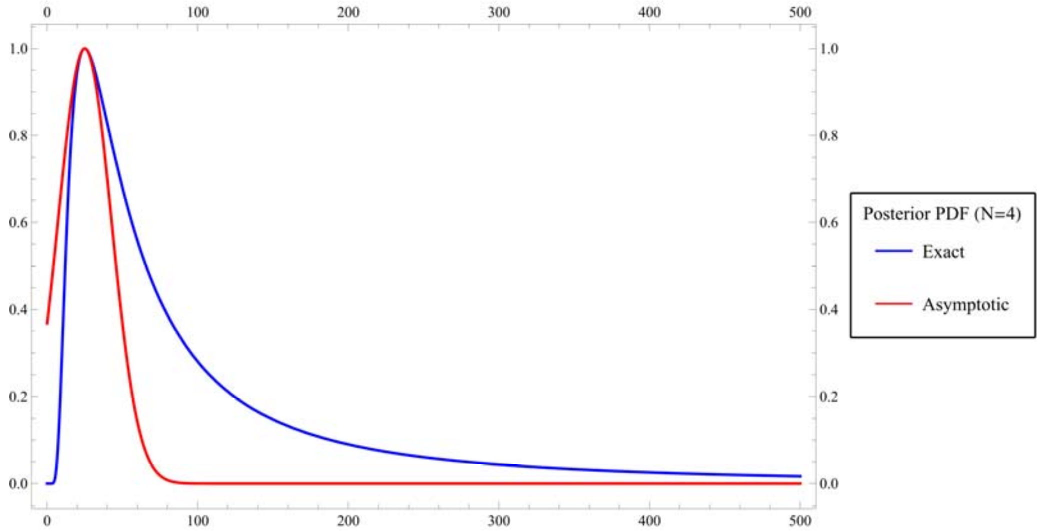
vi.

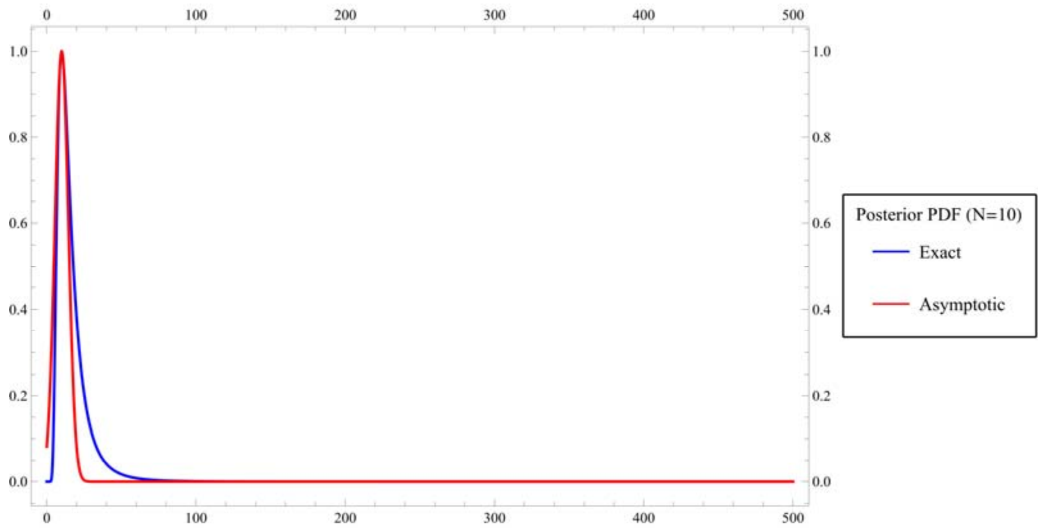
Now we can easily assign an arbitrary value to A, and plot these probability density functions together for various values of N. The resulted plots will provide a graphical representation on the effect of the number of data/measurements on the exact and asymptotic PDFs. The extracted plots are shown below:

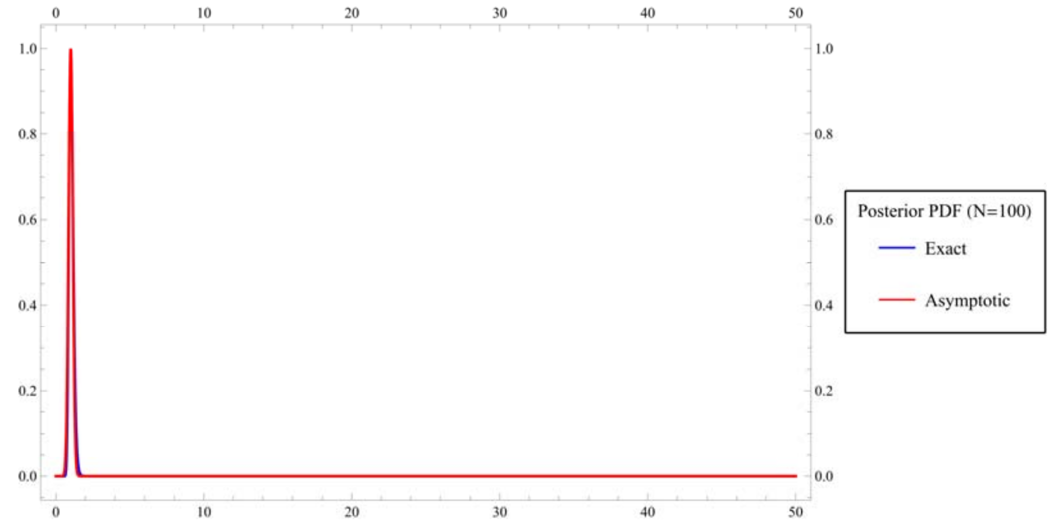

Posterior PDF (N=1)
Exact
Asymptotic

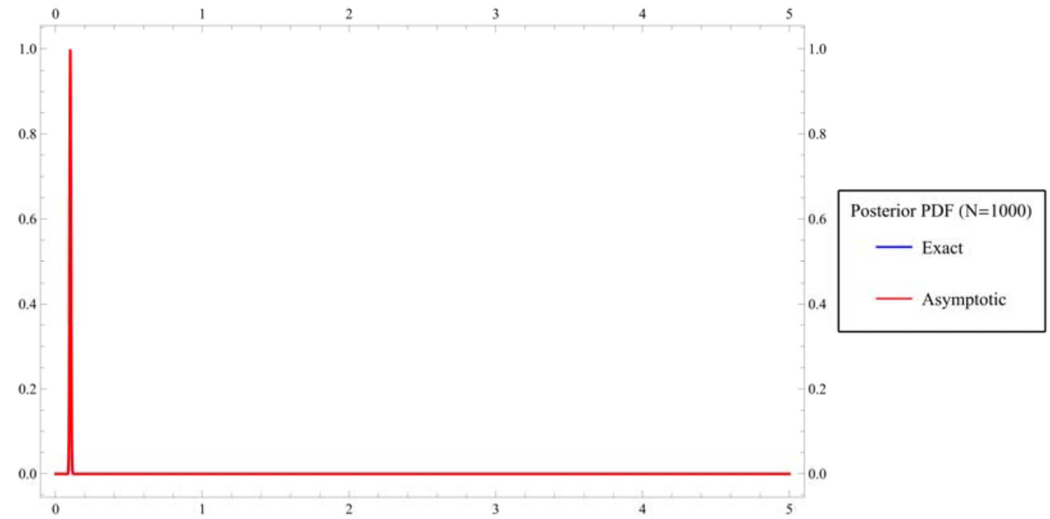

Posterior PDF (N=2)
Exact
Asymptotic


Posterior PDF (N=3)
Exact
Asymptotic


Posterior PDF (N=4)
Exact
Asymptotic


Posterior PDF (N=10)
Exact
Asymptotic


Posterior PDF (N=100)
Exact
Asymptotic
0
10
20
30
40
50
0.0
0.2
0.4
0.6
0.8
1.0


Posterior PDF (N=1000)
Exact
Asymptotic
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0