ISyE8843
Brani Vidakovic
Labor Day **Name:**____

Take Home Quiz 2

Mushrooms. The unhappy outcome of uninformed mushroom-picking is poisoning. In many cases such poisoning is due to ignorance or a superficial approach to identification. The most dangerous fungi are Death Cap (*Amanita Phaloides*) and two species akin to it *Amanita Verna* and Destroying Angel (*Amanita Virosa*). These three toadstools cause the majority of fatal poisoning.

Figure 1: Death Cap (*Amanita Phaloides).*

One of the keys for mushroom identification is the spore deposit. Spores of *Amanita Phaloides* are colorless, almost spherical, and smooth. Measurements in $m\mu$ of 14 spores are given below:

7.96	6.73	8.52	8.68	7.25	9.59	9.58
8.36	8.72	8.57	8.21	9.12	7.81	10.58

The proposed model for the measurements is $[X|\theta,\sigma^2] \sim \mathcal{N}(\theta,\sigma^2)$ with both θ and σ^2 unknown. We are interested in parameter θ.

(a) Find the frequentist 95% confidence interval for the unknown mean.
Assume noninformative prior on $[\theta,\sigma^2]$, $\pi(\theta,\sigma^2) \propto \frac{1}{\sigma^2}$.
(b) Find the 95% credible set for θ if the prior is noninformative. Compare the solutions in (a) and (b).
(c) Assume now that the prior is normal-inverse-chi square, $\mathcal{N} - Inv\chi^2$,

$$\begin{aligned}\theta|\sigma^2 &\sim \mathcal{N}(8.35, \sigma^2/10),\\ \sigma^2 &\sim Inv\chi^2(4, 1.5).\end{aligned}$$

(Use notation from the Text, page 78)
Find the 95% credible set for θ and compare it with interval estimators in (a) and (b).

Solution: [By T. DasGupta] From the data, the following frequentist estimators of θ and σ^2 are obtained,

$$\hat{\theta} = \bar{X} = 8.5486, \qquad s^2 = 0.9758.$$

(a) The frequentist 95% confidence interval for θ is given by

$$\bar{X} \pm t_{n-1,0.025}\frac{s}{\sqrt{n}},$$

for $n = 14$ and $t_{n-1,0.025} = t_{13,0.025} = 2.16$. Thus, the confidence interval is,

$$8.5486 \pm 2.16\frac{0.9878}{\sqrt{14}} \quad \equiv \quad [7.9784, 9.1188].$$

(b) The prior on (θ, σ^2) is given by $\pi(\theta, \sigma^2) \propto \frac{1}{\sigma^2}$. Since, given $x_1, \ldots, x_n$ the likelihood is

$$\frac{1}{(\sqrt{2\pi})^n \sigma^n} \exp\left\{-1/(2\sigma^2)\sum_{i=1}^{n}(x_i - \theta)^2\right\},$$

the joint posterior is

$$\begin{aligned}
\pi(\theta, \sigma^2 | x_1, \ldots, x_n) &\propto \frac{1}{\sigma^{n+2}} \exp\left\{-1/(2\sigma^2)((n-1)s^2 + n(\bar{x} - \theta)^2)\right\} \\
&\propto \frac{1}{\sigma^{n+2}} e^{\frac{A}{2\sigma^2}}, \text{ where } A = (n-1)s^2 + n(\bar{x} - \theta)^2.
\end{aligned}$$

By integrating out σ^2 we get

$$\begin{aligned}
\pi(\theta|x) &\propto \int_0^\infty \frac{1}{\sigma^{n+2}} \exp\{-A/(2\sigma^2)\} d\sigma^2 \\
&\propto \int_\infty^0 z^{(n+2)/2} e^{-Az/2}(-\frac{1}{z^2})dz \qquad [\text{ after } 1/\sigma^2 = z] \\
&\propto \int_0^\infty z^{n/2-1} e^{-Az/2} dz \\
&\propto \frac{\Gamma(n/2)}{(A/2)^{n/2}} \propto A^{-n/2} \\
&\propto \{(n-1)s^2 + n(\bar{x} - \theta)^2\}^{-n/2} \\
&\propto \left\{1 + \left(\frac{\bar{x} - \theta}{s/\sqrt{n}}\right)^2 /(n-1)\right\}^{-\frac{(n-1)+1}{2}},
\end{aligned}$$

which gives

$$\frac{\theta - \bar{x}}{s/\sqrt{n}} \sim t_{n-1}.$$

Thus, the 95% credible set for θ is given by

$$\bar{X} \pm t_{n-1,0.025}\frac{s}{\sqrt{n}},$$

which coincides with the frequentist confidence interval from (a).

(c) Now, we have

$$\begin{aligned}
\theta|\sigma^2 &\sim \mathcal{N}(8.35, \sigma^2/10), \\
\sigma^2 &\sim inv\chi^2(4, 1.5).
\end{aligned}$$

Using the notation from the text [GCSR-BDA], pages 78-79, we get

$$\mu_0 = 8.35, \kappa_0 = 10, \nu_0 = 4, \sigma_0^2 = 1.5.$$

The conditional posterior density $\pi(\theta|\sigma^2, \bar{x})$ is given by

$$\pi(\theta|\sigma^2, \bar{x}) \sim \mathcal{N}(\mu_n, \sigma^2/\kappa_n),$$

where

$$\mu_n = \frac{\kappa_0}{\kappa_0 + n}\mu_0 + \frac{n}{\kappa_0 + n}\bar{x} = \frac{10}{24} \times 8.35 + \frac{14}{24} \times 8.5486 = 8.46585, \quad \text{and } \kappa_n = \kappa_0 + n = 10 + 14 = 24.$$

Also, $\nu_n = \nu_0 + n = 4 + 14 = 18$. Since, $18\sigma_n^2 = 4 \times 1.5 + 13 \times (0.9879)^2 + \frac{10\times 14}{10+14}(8.5486 - 8.35)^2$, it follows $\sigma^n = 1.0251$.

From $\sqrt{\kappa_k}\frac{\theta-\mu_n}{\sigma_n} \sim t_{\nu_n}$, and $t_{18,0.025} = 2.101$, the 95% credible set is

$$\mu_n \pm t_{\nu_n,0.025}\frac{\sigma_n}{\sqrt{\kappa_n}} \qquad \equiv [8.0263, 8.9054].$$

This credible set is shorter than one in (b) or (a), because the prior distribution is informative.