Aidan Kardan Bayesian HW3

1) For a scalar parameter θ,

i) Confidence Interval, CI, Frequentist: $100(1-\alpha)\%$ Confidence Interval is constructed so that in repeated sampling the procedure yields an interval that contains the true parameter θ in $100(1-\alpha)\%$ of the samples. When using an estimator $\hat{\theta}$ whose sampling distribution is approx. normal, CI: $\left[\hat{\theta} - z_{1-\alpha/2} \cdot se(\hat{\theta}),\ \hat{\theta} + z_{1-\alpha/2} \cdot se(\hat{\theta})\right]$, where $z_{1-\alpha/2}$ is the quantile of the standard normal distribution &

$se(\hat{\theta})$: standard error

ii) Highest Posterior Density (HPD) credible interval: Given a posterior density, $p(\theta|data)$, HPD credible interval is the set C such that: 1) Total posterior probability is $P(\theta \in C | data) = 1-\alpha$

2) Every point inside C has a higher posterior density than every point outside,

i.e., if $\theta_1 \in C$ & $\theta_2 \notin C$, then $p(\theta_1|data) \geq p(\theta_2|data)$

HPD: "shortest" interval covering $1-\alpha$ of the probability mass

iii) HPD credible set: For multivariate parameters, we refer to a set instead of an interval satisfying the HPD property. In the scalar case, the HPD credible set coincides with the HPD credible interval.

iv) Equal-tailed Credible Interval: Bayesian Interval constructed such that the probability mass in each tail is equal. Mathematically, if $F(\theta|data)$ is the posterior CDF, then the interval $[a,b]$ satisfies

$$F(a|data) = \frac{\alpha}{2}, \quad F(b|data) = 1 - \frac{\alpha}{2}$$

For symmetric posteriors (i.e., normal), the equal-tailed interval coincides with the HPD interval!

For skewed posteriors, the HPD is generally narrower (shorter) than the equal-tailed interval.

Question 2 Done in separate pdf!

3) $f(x|\lambda) = \lambda e^{-\lambda x}, \quad x \geq 0$

a) likelihood for a sample $x_1, \ldots, x_n$: $L(\lambda | x_1, \ldots, x_n) = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right)$

Conjugate prior : prior chosen s.t. when combined w/ likelihood, posterior falls within same family as prior

A conjugate prior for λ is the Gamma distribution:

$\lambda \sim \text{Gamma}(\alpha_0, B_0)$ w/density: $\pi(\lambda) = \frac{B_0^{\alpha_0}}{\Gamma(\alpha_0)} \lambda^{\alpha_0 - 1} e^{-B_0 \lambda}$, multiplying the likelihood by the prior (ignoring constant factors)

yields: $\pi(\lambda | x) \propto \lambda^{\alpha_0 + n - 1} \exp\left(-\lambda\left(B_0 + \sum_{i=1}^{n} x_i\right)\right) \Rightarrow$ posterior is $\text{Gamma}\left(\alpha_0 + n, B_0 + \sum_{i=1}^{n} x_i\right)$

b) For a single observation, $\ell(\lambda | x) = \log \lambda - \lambda x$

$$\frac{\delta \ell}{\delta \lambda} = \frac{1}{\lambda} - x$$

Fisher Information: $I(\lambda) = -E\left[\frac{\delta^2 \ell}{\delta \lambda^2}\right] = \frac{1}{\lambda^2}$

$$\frac{\delta^2 \ell}{\delta \lambda^2} = -\frac{1}{\lambda^2}$$

Jeffrey's prior: $\pi_J \propto \sqrt{I(\lambda)} \propto \frac{1}{\lambda}$

$$\boxed{\pi_J \propto \frac{1}{\lambda}}$$

c), d) Done in separate pdf!

Question 4 apart from part d) done in separate pdf!

4) $\mu \sim N(\mu_0, \tau_0^2)$ $\mu_0 = 6$ $\tau_0^2 = 0.25$ $y_i \sim N(\mu, \sigma^2)$

d) Given n i.i.d. observations w/ sample mean, $\bar{y}$, likelihood function (ignoring constant factors):

Likelihood: $L(\mu) \propto \exp\left(-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2\right)$

Prior: $\pi(\mu) \propto \exp\left(-\frac{1}{2\tau_0^2}(\bar{y}-\mu_0)^2\right)$

Posterior: unnormalized posterior: $\pi(\mu|y) \propto \underbrace{\exp\left\{-\frac{1}{2}\left[\frac{n}{\sigma^2}(\mu-\bar{y})^2 + \frac{1}{\tau_0^2}(\mu-\mu_0)^2\right]\right\}}_{\text{kernel for Normal distribution}}$

$$\left.\begin{aligned} &\text{posterior variance: } \sigma_{post}^2 = \left(\frac{n}{\sigma^2} + \frac{1}{\tau_0^2}\right)^{-1} \\ &\text{posterior mean: } \mu_{post} = \sigma_{post}^2\left(\frac{n\bar{y}}{\sigma^2} + \frac{\mu_0}{\tau_0^2}\right) \end{aligned}\right\} \text{therefore, the posterior is: } \mu|y \sim N(\mu_{post}, \sigma_{post}^2)$$