1. Let $X_1, X_2, \ldots$ be independent random variables with mean $\mu$ and variance $\sigma^2 < \infty$ and let $\bar{X}_n$ be the sample mean of the first $n$ observations. The Law of Large Numbers says that $\bar{X}_n \xrightarrow{p} \mu$. Write the mathematical expression implied by the statement that $\bar{X}_n \xrightarrow{p} \mu$? (Hint: The answer involves a limit.)

\begin{equation}
\lim_{n \to \infty} \Pr\left( \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right) < \varepsilon = 1 \quad \text{for every } \varepsilon > 0.
\end{equation}

2. Let $X_1, \ldots, X_{25}$ be a random sample of beta random variables with shapes 2 and 8. Approximate $\Pr\left( \frac{1}{25} \sum_{i=1}^{25} X_i > 0.24 \right)$ using $\Phi$, the cumulative distribution function of the standard normal distribution. Hint, the variance of a beta random variable with shapes $\alpha_1$ and $\alpha_2$ is $\frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2 (\alpha_1 + \alpha_2 + 1)}$.

$$
\begin{align*}
\bar{X} &= \frac{1}{n} \sum_{i=1}^{n} X_i \\
E(\bar{X}) &= E\left( \frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \frac{1}{n} n \left( \frac{\alpha_1}{\alpha_1 + \beta} \right) = \frac{\alpha_1}{\alpha_1 + \beta} = \frac{2}{2 + 8} = \frac{1}{5} \\
\text{Var}(\bar{X}) &= \text{Var}\left( \frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}(X_i) = \frac{1}{n^2} n \left( \frac{\alpha_2}{(\alpha_1 + \beta)^2 (\alpha_1 + \beta + 1)} \right) \\
&\quad \text{by independence}
\end{align*}
$$

By the C.L.T.,

$$
\bar{X} \sim N\left( \frac{1}{5}, \frac{4}{62.75} \right)
$$

$$
\Pr\left( \bar{X} > 0.24 \right) \approx \Pr\left( Z > \frac{0.24 - \frac{1}{5}}{\sqrt{4/62.75}} \right) = 1 - \Phi\left( \frac{1.6583}{\sqrt{4/62.75}} \right)
$$

$Z \sim N(0,1)$
For all the remaining parts on this exam, suppose that $X_1, \ldots, X_n$ are independent, where $X_i$ given $\lambda$ has a Poisson distribution with rate $\lambda$. The probability mass function of $X_i$ given $\lambda$ is:

$$p(x_i | \lambda) = \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \quad \text{for } x_i \in \{0, 1, \ldots\}.$$ 

Let the classical estimator of $\lambda$ be:

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Assume a gamma prior on $\lambda$. (Hint: Your calculations will be easier if you use parameterize the gamma distribution in terms of a shape and rate.) For Bayesian estimation, use squared error loss.

3. Succinctly write the model outlined above in distributional notation (e.g., $X \sim \text{Normal}(\mu, \sigma)$). Name the two parts of the model. Note that this part does not involve any densities.

$$X_i \mid \lambda \sim \text{Poisson}(\lambda)$$

$$\lambda \sim \text{Gamma}(\alpha, \beta)$$

$$\leftarrow \text{sampling model}$$

$$\leftarrow \text{prior distribution}$$

4. Write $p(x_1, \ldots, x_n | \lambda)$, the joint density of the sampling model for $X_1, \ldots, X_n$. A random sample justifies what key assumption used to find this joint density?

$$p(x_1, \ldots, x_n | \lambda) = \prod_{i=1}^{n} p(x_i | \lambda)$$

$$= \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

$$= \left( \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \right)^{\frac{1}{\lambda}} e^{-\lambda} \quad \text{for } x_i \in \{0, 1, 2, \ldots \}$$

where $\lambda > 0$.

5. For this item only, assume that $n = 3, x_1 = 3, x_2 = 5,$ and $x_3 = 4$. Write the likelihood function for $\lambda$ as simply as possible.

$$L(\lambda) = \frac{1}{3! \cdot 5! \cdot 4!} \lambda^{(3+5+4)} e^{-3\lambda}$$

$$= \frac{1}{17280} \lambda^{12} e^{-3\lambda} \quad \text{for } \lambda > 0$$
6. Write \( p(\lambda) \), the density of the prior distribution for \( \lambda \).

\[
p(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta \lambda} \quad \text{for} \quad \lambda > 0
\]

where \( \alpha > 0 \) \( \alpha \) is shape parameter
\( \beta > 0 \) \( \beta \) is rate parameter

7. Find \( p(\lambda | x_1, \ldots, x_n) \), the density of the posterior distribution of \( \lambda \).

\[
p(\lambda | x_1, \ldots, x_n) \propto p(x_1, \ldots, x_n | \lambda) \cdot p(\lambda)
\]

\[
= \left[ \frac{n}{\alpha} \lambda^\alpha \sum_{x=1}^{n} x_i \right] \lambda^{-n-\alpha} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{-1} e^{-\beta \lambda} \cdot \frac{(\alpha + \sum x_i - 1)}{\Gamma(\alpha + \sum x_i)} \lambda^{\alpha+\sum x_i - 1} e^{-(\alpha+\sum x_i) \lambda}
\]

Density of \( \Gamma(\alpha + \sum x_i, \beta + n) \)

8. What is the posterior distribution of \( \lambda \)? That is, what is the conditional distribution of \( \lambda \) given \( X_1, \ldots, X_n \)? Note that this part does not ask for a density.

\[
\lambda | x_1, x_2, \ldots, x_n \sim \Gamma(\alpha + \sum_{i=1}^{n} x_i, \beta + n)
\]
9. What is the Bayes estimator of $\lambda$?

Under squared error loss, $\hat{\lambda}_B = E(\lambda | X_1, \ldots, X_n) = \frac{\alpha + \sum_{i=1}^{n} X_i}{\beta + n}$

10. Show that the Bayes estimator of $\lambda$ is a weighted average of the prior estimate of $\lambda$ and the classical estimator of $\lambda$.

\[
\hat{\lambda}_B = \frac{\alpha + \sum_{i=1}^{n} X_i}{\beta + n} = \left( \frac{1}{\beta + n} \right) \left( \frac{\alpha}{\beta} \right) + \left( \frac{n}{\beta + n} \right) \left( \frac{\sum_{i=1}^{n} X_i}{n} \right)
\]

Note that $w_p + w_d = 1$.

11. Find the bias of the classical estimator of $\lambda$.

bias($\hat{\lambda}_c$) = $E(\hat{\lambda}_c) - \lambda$

= $E\left( \frac{1}{n} \sum_{i=1}^{n} X_i \right) - \lambda$

= $\frac{1}{n} \sum_{i=1}^{n} E(X_i) - \lambda$

= $\frac{1}{n} n \lambda - \lambda = 0$

12. Find the bias of the Bayes estimator of $\lambda$.

bias($\hat{\lambda}_B$) = $E(\hat{\lambda}_B) - \lambda$

= $E\left( \frac{\alpha + \sum X_i}{\beta + n} \right) - \lambda$

= $\frac{\alpha + \sum E(X_i)}{\beta + n} - \lambda$

= $\frac{\alpha + n \lambda}{\beta + n} - \lambda$
13. Find the mean squared error (MSE) of the classical estimator of $\lambda$.

$$\text{mse}(\hat{\lambda}_c) = \text{var}(\hat{\lambda}_c) + \text{bias}^2(\hat{\lambda}_c)$$

$$= \text{var}\left(\frac{\sum X_i}{n}\right) + 0$$

by independence

$$= \frac{1}{n^2} \sum \text{var}(X_i)$$

$$= \frac{1}{n^2} n \lambda$$

$$= \frac{\lambda}{n}$$

14. Find the mean squared error (MSE) of the Bayes estimator of $\lambda$.

$$\text{mse}(\hat{\lambda}_B) = \text{var}(\hat{\lambda}_B) + \text{bias}^2(\hat{\lambda}_B)$$

$$= \text{var}\left(\frac{\alpha + \sum X_i}{\beta + n}\right) + \left(\frac{\alpha + n \lambda}{\beta + n} - \lambda\right)^2$$

by independence

$$= \frac{1}{(\beta + n)^2} \sum_{i=1}^{n} \text{var}(X_i) + \left(\frac{\alpha + n \lambda}{\beta + n} - \lambda\right)^2$$

$$= \frac{1}{(\beta + n)^2} n \lambda + \left(\frac{\alpha + n \lambda}{\beta + n} - \lambda\right)^2$$
15. What values for the shape and rate parameters in the gamma prior are compatible with prior estimate for \( \lambda \) being 4 and effective number of observations from the prior being 10?

\[
E(\lambda) = \frac{\alpha}{\beta} = 4 \quad \Leftarrow \quad \text{prior estimate}
\]

\[
\beta = 10 \quad \Leftarrow \quad \text{effective sample size}
\]

\[
\Rightarrow \quad \alpha = 40
\]

\[
\beta = 10
\]

16. Using values from the previous problem and assuming \( n = 10 \), give a value for \( \lambda \) for which the Bayes estimator is superior in terms of mean squared error. Likewise, give a value for \( \lambda \) for which the Bayes estimator is inferior in terms of mean squared error. Justify your choices mathematically.

**Suppose \( \lambda = 4 \)**

\[
\text{mse} (\hat{\lambda}_B) = \frac{10 (4)}{(10 + 10)^2} + \left( \frac{40 + 10 (4)}{10 + 10} - 4 \right)^2
\]

\[
= \frac{40}{400} + \left( \frac{80}{20} - 4 \right)^2 = \frac{1}{10}
\]

\[
\text{mse} (\hat{\lambda}_C) = \frac{4}{10}
\]

Note that \( \text{mse} (\hat{\lambda}_B) < \text{mse} (\hat{\lambda}_C) \) \[ \frac{1}{10} < \frac{1}{40} \]

**Suppose \( \lambda = 100 \)**

\[
\text{mse} (\hat{\lambda}_B) = \frac{10 (100)}{(10 + 10)^2} + \left( \frac{40 + 10 (100)}{10 + 10} - 100 \right)^2
\]

\[
= \frac{1000}{400} + \left( \frac{1040}{20} - 100 \right)^2 = 2306.5
\]

\[
\text{mse} (\hat{\lambda}_C) = \frac{100}{10} = 10
\]

Note that \( \text{mse} (\hat{\lambda}_B) > \text{mse} (\hat{\lambda}_C) \)

\[ 2306.5 > 10 \]