Math 152, Fall 2022
Jo Hardin
WU # 4
Thursday 9/8/22

Your name: ____________________________________

Names of people you worked with: ____________________________________

Task:
Recall the tape example. Let θ denote the average number of defects per 100 feet of tape. Assume $X_1, X_2, \ldots X_n \sim$ Poisson(θ).

Assume that you have not yet collected any data. The goal is to compare the frequentist and Bayesian **estimators** (not a comparison of estimates, but a comparison of estimators). Find

1. MSE_F for $\hat{\theta}$ (the answer should be a function of θ)
2. MSE_B for $\delta(\underline{X})$ (answer should be a function of X)
3. MSE_F for $\delta(\underline{X})$ (the answer should be a function of θ)

- Prior: Gamma(2, 10) (or (2, 1/10) depending on how you parametrize)
- Data likelihood: Poisson(θ)
- Posterior: Gamma $\left(\sum X_i + 2,\, n + 10\right)$, parameterized in such a way that $E(\theta|\underline{X}) = \frac{\sum X_i + 2}{n+10}$.

$$
\begin{aligned}
\text{(frequentist estimator)} \quad \hat{\theta} &= \frac{\sum X_i}{n} \\
\text{(Bayesian estimator)} \quad \delta(\underline{X}) &= \frac{\sum X_i + 2}{n + 10}
\end{aligned}
$$

Solution:
Because there isn't a way to directly compare MSE_F for $\hat{\theta}$ and MSE_B for $\delta(\underline{X})$, we can calculate MSE_F for $\delta(\underline{X})$ to compare the two frequentist MSE values.

1.

$$
\begin{aligned}
MSE_F(\hat{\theta}) &= var(\hat{\theta}) + bias(\hat{\theta})^2 \\
&= \frac{\sum var(X_i)}{n^2} + \left(\frac{\sum E[X_i]}{n} - \theta\right)^2 \\
&= \theta/n + 0 = \theta/n
\end{aligned}
$$

2.

$$
\begin{aligned}
MSE_B(\delta(\underline{X})) &= var(\theta|\underline{X}) \\
&= \frac{\sum X_i + 2}{(n+10)^2}
\end{aligned}
$$

3.

$$MSE_F(\delta(\underline{X})) = var(\delta(\underline{X})) + bias(\delta(\underline{X}))^2$$

$$\begin{aligned}
bias(\delta(\underline{X})) &= E\left[\frac{\sum X_i + 2}{n+10}\right] - \theta \\
&= \frac{n\theta + 2}{n+10} - \theta = \frac{n\theta + 2 - n\theta - 10\theta}{n+10} \\
&= \frac{2 - n\theta}{n+10} \\
var(\delta(\underline{X})) &= var\left[\frac{\sum X_i + 2}{n+10}\right] \\
&= \frac{1}{(n+10)^2} var\left(\sum X_i\right) \\
&= \frac{1}{(n+10)^2} n \; var(X_i) \\
&= \frac{n}{(n+10)^2}\theta
\end{aligned}$$

$$\begin{aligned}
MSE_F(\delta(\underline{X})) &= \frac{n}{(n+10)^2}\theta + \frac{(2-n\theta)^2}{(n+10)^2} \\
&= \frac{n\theta + (2-n\theta)^2}{(n+10)^2}
\end{aligned}$$

Note that we couldn't directly compare MSE_F and MSE_B (they are functions of different variables!). Because we'd have to come up with a prior to think about $MSE_B(\hat{\theta})$, it seems like we can't calculate that quantity. Instead, we take the easier route, and find $MSE_F(\delta(\underline{X}))$ in order to have a reasonable comparison of estimators.

The actual comparison of the two frequentist MSEs depend on the values of n and θ. As the researcher you usually have control of n, but you don't necessarily (usually) have control of θ.