Homework 2 Solutions

1. As the variance is known to be 1, the 95% CI is

$$\left[\bar{X} - \frac{z_{0.975}}{\sqrt{n}}, \bar{X} + \frac{z_{0.975}}{\sqrt{n}}\right].$$

In order for the interval length to be at most 0.5, we need

$$\frac{2z_{0.975}}{\sqrt{n}} \le 0.5 \quad \Leftrightarrow \quad n \ge 16 z_{0.975}^2 \approx 61.4,$$

which means that we need $n \ge 62$.

2. Note that

$$\begin{aligned}
\mathbb{E}(\hat{\theta} - \theta)^2 &= \mathbb{E}(\hat{\theta} - \mathbb{E}\hat{\theta} + \mathbb{E}\hat{\theta} - \theta)^2 \\
&= \mathbb{E}(\hat{\theta} - \mathbb{E}\hat{\theta})^2 + (\mathbb{E}\hat{\theta} - \theta)^2 + 2\mathbb{E}\left((\hat{\theta} - \mathbb{E}\hat{\theta})(\mathbb{E}\hat{\theta} - \theta)\right) \\
&= \mathrm{var}(\hat{\theta}) + (\mathbb{E}\hat{\theta} - \theta)^2 + 2\mathbb{E}\left((\hat{\theta} - \mathbb{E}\hat{\theta})\right) \times (\mathbb{E}\hat{\theta} - \theta) \\
&= \mathrm{var}(\hat{\theta}) + (\mathbb{E}\hat{\theta} - \theta)^2.
\end{aligned}$$

3-(a). We can view the counts in the first table as $X_1, \ldots, X_{180} \sim \text{Poisson}(10\lambda)$. Similarly, we can view the second table as $Y_1, \ldots, Y_{20} \sim \text{Poisson}(20\lambda)$. Then, given the data $X_1, \ldots, X_{180}$ and $Y_1, \ldots, Y_{20}$, the MLE of the rate λ is obtained by minimizing the log-likelihood:

$$\begin{aligned}
\ell(\lambda) &:= \log\left(\prod_{i=1}^{180} \frac{e^{-10\lambda}(10\lambda)^{X_i}}{X_i!} \times \prod_{i=1}^{20} \frac{e^{-20\lambda}(20\lambda)^{Y_i}}{Y_i!}\right) \\
&= -2200\lambda + \left(\sum_{i=1}^{180} X_i + \sum_{i=1}^{20} Y_i\right)\log(\lambda) + \text{terms independent of } \lambda.
\end{aligned}$$

Hence, we can see that

$$\hat{\lambda} = \frac{\sum_{i=1}^{180} X_i + \sum_{i=1}^{20} Y_i}{2200} \approx 0.1577.$$

3-(b). As we are assuming that the detectors are independent, we have

$$\sum_{i=1}^{180} X_i + \sum_{i=1}^{20} Y_i \sim \text{Poisson}(2200\lambda),$$

which is equivalent to the distribution of $\sum_{i=1}^{2200} Z_i$, where $Z_1, \ldots, Z_{2200}$ are i.i.d from Poisson(λ). By the CLT, we have

$$\sqrt{2200}\left(\frac{\sum_{i=1}^{2200} Z_i}{2200} - \lambda\right) \approx N(0, \lambda).$$

From this, we conclude that

$$\hat{\lambda} \stackrel{d}{=} \frac{\sum_{i=1}^{2200} Z_i}{2200} \approx N\left(\lambda, \frac{\lambda}{2200}\right).$$

4-(a). We have $\mathbb{E}(Z^2) = \text{var}(Z) = 1$. Meanwhile, $\mathbb{E}(Z^4) = 3$; see https://en.wikipedia.org/wiki/Normal_distribution.

4-(b). By definition, we have

$$\mathbb{E}(Y) = \sum_{i=1}^{n} \mathbb{E}(Z_i^2) = n.$$

Meanwhile, as $Z_1, \ldots, Z_n$ are independent, we have

$$\text{var}(Y) = \sum_{i=1}^{n} \text{var}(Z_i^2) = n(\mathbb{E}(Z_1^4) - (\mathbb{E}(Z_1^2))^2) = 2n.$$