MATH20802: STATISTICAL METHODS
SECOND SEMESTER
ANSWERS TO THE IN CLASS TEST

ANSWERS TO QUESTION 1

In a series of m independent Bernoulli trials there are X successes. In a further series of n trials there are Y successes. Assuming that the probability of success, p, is the same for both sets of trials, consider the estimators for p: show that

$$\widehat{p}_1 = \frac{1}{2}\left(\frac{X}{m} + \frac{Y}{n}\right) \tag{1}$$

and

$$\widehat{p}_2 = \frac{X+Y}{m+n}. \tag{2}$$

(i) The bias of the first estimator is

$$\begin{aligned}
\text{Bias}\left(\widehat{p_1}\right) &= E\left(\widehat{p_1}\right) - p \\
&= E\left[\frac{1}{2}\left(\frac{X}{m} + \frac{Y}{n}\right)\right] - p \\
&= \frac{1}{2}\left[\frac{E(X)}{m} + \frac{E(Y)}{n}\right] - p \\
&= \frac{1}{2}\left(\frac{mp}{m} + \frac{np}{n}\right) - p \\
&= \frac{1}{2}(p+p) - p \\
&= 0
\end{aligned}$$

(2 marks)

(ii) The bias of the second estimator is

$$\begin{aligned}
\text{Bias}\left(\widehat{p_2}\right) &= E\left(\widehat{p_2}\right) - p \\
&= E\left(\frac{X+Y}{m+n}\right) - p \\
&= \frac{E(X+Y)}{m+n} - p \\
&= \frac{E(X)+E(Y)}{m+n} - p \\
&= \frac{mp+np}{m+n} - p \\
&= p - p \\
&= 0.
\end{aligned}$$

(2 marks)

(iii) The MSE of the first estimator is

$$\begin{aligned}
\text{MSE}\left(\widehat{p_1}\right) &= \text{Var}\left(\widehat{p_1}\right) \\
&= \text{Var}\left(\frac{1}{2}\left(\frac{X}{m}+\frac{Y}{n}\right)\right) \\
&= \frac{1}{4}\text{Var}\left(\frac{X}{m}+\frac{Y}{n}\right) \\
&= \frac{1}{4}\left[\frac{\text{Var}(X)}{m^2}+\frac{\text{Var}(Y)}{n^2}\right] \\
&= \frac{1}{4}\left[\frac{mp(1-p)}{m^2}+\frac{np(1-p)}{n^2}\right] \\
&= \frac{1}{4}\left[\frac{p(1-p)}{m}+\frac{p(1-p)}{n}\right] \\
&= \frac{p(1-p)}{4}\left(\frac{1}{m}+\frac{1}{n}\right).
\end{aligned}$$

(2 marks)

(iv) The MSE of the second estimator is

$$\begin{aligned}
\text{MSE}\left(\widehat{p_2}\right) &= \text{Var}\left(\widehat{p_2}\right) \\
&= \text{Var}\left(\frac{X+Y}{m+n}\right) \\
&= \frac{1}{(m+n)^2}\text{Var}\left(X+Y\right) \\
&= \frac{1}{(m+n)^2}\left[\text{Var}(X)+\text{Var}(Y)\right] \\
&= \frac{1}{(m+n)^2}\left[mp(1-p)+np(1-p)\right] \\
&= \frac{p(1-p)}{m+n}.
\end{aligned}$$

(2 marks)

(v) $\widehat{p_2}$ has smaller MSE than $\widehat{p_1}$ since

$$\begin{aligned}
& \frac{p(1-p)}{m+n} \le \frac{p(1-p)}{4}\left(\frac{1}{m}+\frac{1}{n}\right) \\
\Longleftrightarrow\quad & \frac{1}{m+n} \le \frac{1}{4}\left(\frac{1}{m}+\frac{1}{n}\right) \\
\Longleftrightarrow\quad & \frac{1}{m+n} \le \frac{1}{4}\frac{m+n}{mn} \\
\Longleftrightarrow\quad & 4mn \le (m+n)^2 \\
\Longleftrightarrow\quad & 4mn \le m^2+n^2+2mn \\
\Longleftrightarrow\quad & 0 \le m^2+n^2-2mn \\
\Longleftrightarrow\quad & 0 \le (m-n)^2.
\end{aligned}$$

(2 marks)