# STAT2371 Assignment

Ze Hong Zhou (46375058)

2022-08-25 (last edited 2022-08-26)

## Question 1

**Suppose that two independent binomial random variables $X_1$ and $X_2$ are observed where $X_1$ has a Binomial($n$, $p$) distribution and $X_2$ has a Binomial($2n, p$) distribution. You may assume that $n$ is known, whereas $p$ is an unknown parameter. Define two possible estimators of $p$**

$$\hat{p}_1 = \frac{1}{3n}(X_1 + X_2) \quad \textbf{and} \quad \hat{p}_2 = \frac{1}{2n}(X_1 + 0.5X_2).$$

**(a) Show that both of the estimators $\hat{p}_1$ and $\hat{p}_2$ are unbiased estimators of $p$.**

$$\begin{aligned} E(\hat{p}_1) &= \frac{1}{3n}(E(X_1) + E(X_2)) \quad \text{applying expected value linearity} \\ &= \frac{1}{3n}(n \cdot p + 2n \cdot p) \quad \text{applying the binomial expectation formula} \\ &= p \\ \text{bias}(\hat{p}_1, p) &= E(\hat{p}_1) - p = 0. \end{aligned}$$

Similarly,

$$\begin{aligned} E(\hat{p}_2) &= \frac{1}{2n}(E(X_1) + 0.5E(X_2)) \\ &= \frac{1}{2n}(n \cdot p + 0.5 \cdot 2n \cdot p) \\ &= p \\ \text{bias}(\hat{p}_2, p) &= E(\hat{p}_2) - p = 0. \end{aligned}$$

Hence $\hat{p}_1$ and $\hat{p}_2$ are unbiased estimators of $p$.

**(b) Find Var($\hat{p}_1$) and Var($\hat{p}_2$).**

$$\begin{aligned} Var(\hat{p}_1) &= \frac{1}{9n^2}(Var(X_1) + Var(X_2)) \quad \text{applying the formula for independent case} \\ &= \frac{1}{9n^2}(np(1-p) + 2np(1-p)) \quad \text{applying the binomial variance formula} \end{aligned}$$

$$= \frac{p(1-p)}{3n}.$$

Similarly,

$$\begin{aligned} Var(\hat{p}_2) &= \frac{1}{4n^2}(Var(X_1) + 0.5^2 Var(X_2)) \\ &= \frac{1}{4n^2}(np(1-p) + 0.25 \cdot 2np(1-p)) \\ &= \frac{3p(1-p)}{8n}. \end{aligned}$$

**(c) Show that both estimators are consistent estimators of $p$.**

Let $\varepsilon > 0$.

$$\begin{aligned} \lim_{n\to\infty} P(|\hat{p}_1 - p| > \varepsilon) &\leq \lim_{n\to\infty} \frac{E\left((\hat{p}_1 - p)^2\right)}{\varepsilon^2} \quad \text{applying Markov's inequality} \\ &= \lim_{n\to\infty} \frac{Var(\hat{p}_1)}{\varepsilon^2} \quad \text{since } \hat{p}_1 \text{ is unbiased} \\ &= \lim_{n\to\infty} \frac{p(1-p)}{3n\varepsilon^2} \\ &= 0 \quad \text{for all } p \in [0,1]. \end{aligned}$$

Applying the squeeze theorem, $\lim_{n\to\infty} P(|\hat{p}_1 - p| > \varepsilon) = 0$. Similarly,

$$\begin{aligned} \lim_{n\to\infty} P(|\hat{p}_2 - p| > \varepsilon) &\leq \lim_{n\to\infty} \frac{E\left((\hat{p}_2 - p)^2\right)}{\varepsilon^2} \\ &= \lim_{n\to\infty} \frac{Var(\hat{p}_2)}{\varepsilon^2} \\ &= \lim_{n\to\infty} \frac{3p(1-p)}{8n\varepsilon^2} \\ &= 0 \quad \text{for all } p \in [0,1]. \end{aligned}$$

So $\lim_{n\to\infty} P(|\hat{p}_2 - p| > \varepsilon) = 0$. Hence $\hat{p}_1$ and $\hat{p}_2$ are weakly consistent estimators of $p$.

**(d) Show that $\hat{p}_1$ is the most efficient estimator among all unbiased estimators.**

$\underset{\sim}{X} \equiv \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, $\underset{\sim}{x} \equiv \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

$$\begin{aligned} f_{\underset{\sim}{X}}(\underset{\sim}{x}, p) &= f_{X_1}(x_1, p) \cdot f_{X_2}(x_2, p) \quad \text{applying the r.v. independence definition} \\ &= \begin{cases} \binom{n}{x_1} p^{x_1}(1-p)^{n-x_1} \cdot \binom{2n}{x_2} p^{x_2}(1-p)^{2n-x_2} & \text{if } x_1 \in [0,n] \cap \mathbb{N} \text{ and } x_2 \in [0,2n] \cap \mathbb{N} \\ 0 & \text{if otherwise} \end{cases} \end{aligned}$$

$$= \begin{cases} \binom{n}{x_1}\binom{2n}{x_2} p^{x_1+x_2}(1-p)^{3n-(x_1+x_2)} & \text{if } x_1 \in [0,n] \cap \mathbb{N} \text{ and } x_2 \in [0,2n] \cap \mathbb{N} \\ 0 & \text{if otherwise} \end{cases}$$

$$= \begin{cases} \binom{n}{x_1}\binom{2n}{x_2} exp\left(\begin{bmatrix} ln(p) \\ ln(1-p) \end{bmatrix}^T \begin{bmatrix} x_1+x_2 \\ 3n-(x_1+x_2) \end{bmatrix}\right) & \text{if } x_1 \in [0,n] \cap \mathbb{N} \text{ and } x_2 \in [0,2n] \cap \mathbb{N} \\ 0 & \text{if otherwise} \end{cases}$$

Since $n$ is known and fixed, $\underset{\sim}{X}$ has a pdf in the exponential family, and any sufficient static is also complete. $X_1 + X_2$ is thus sufficient and complete by the sufficient statistic factorisation theorem. By the Lehmann-Scheffé theorem, $E(\hat{p}_1 \mid X_1 + X_2) = \hat{p}_1$ is the unique MVUE, i.e. it is the most efficient estimator among all unbiased estimators.

**(e) Derive the efficiency of the estimator $\hat{p}_1$ relative to $\hat{p}_2$.**

$$\begin{aligned} \text{eff}(\hat{p}_2, \hat{p}_1, p) &= \frac{\text{MSE}(\hat{p}_2, p)}{\text{MSE}(\hat{p}_1, p)} = \frac{\text{Var}(\hat{p}_2) + (\text{bias}(\hat{p}_2, p))^2}{\text{Var}(\hat{p}_1) + (\text{bias}(\hat{p}_1, p))^2} = \frac{\frac{3p(1-p)}{8n}}{\frac{p(1-p)}{3n}} \\ &= \frac{9}{8} \end{aligned}$$

# Question 2

The random variables $X_1, X_2, \ldots, X_{2n}$ are independent and normally distributed with common variance $\sigma^2$. However, $X_1, X_2, \ldots, X_n$ have mean 00 while $X_{n+1}, X_{n+2}, \ldots, X_{2n}$ have mean $\mu$.