Homework Assignment 3

Question 1

Consider a finite population of size N. Let $X_1, \dots, X_n$ be a random sample (without replacement) from this finite population.

1) (1 point) Show that the joint marginal distribution of any two of $X_1, \dots, X_n$ is

$$g(x_r, x_s) = \frac{1}{N(N-1)} \quad \text{for } r \neq s$$

2) (1 point) Let the population be $\{1/N, 2/N, \dots, 1\}$. Find the variance of the sample mean. (Hint: need to find the value of population variance σ^2 first.)

1) let $X_{-2} = \{x_1, \dots, x_n\} / \{x_r, x_s\}$

$$g(x_r, x_s) = \sum_{X_{-2}} f(x_1, \dots, x_n)$$

$$= \frac{1}{N(N-1)\cdots(N-n+1)} \cdot \binom{N-2}{n-2} (n-2)!$$

$$= \frac{1}{N(N-1)\cdots(N-n+1)} \cdot \frac{(N-2)!}{(N-n)!} = \frac{1}{N(N-1)}$$

2) $$Var(\bar{X}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N-1}$$

Population, $\mu = \frac{(\frac{1}{N}+1)N}{2N} = \frac{N+1}{2N}$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(c_i - \mu)^2$$

$$= \frac{1}{N}\sum_{i=1}^{N} c_i^2 - \mu^2$$

$$= \frac{1}{N^3}\frac{N(N+1)(2N+1)}{6} - \left(\frac{N+1}{2N}\right)^2$$

$$= \frac{(N+1)(N-1)}{12N^2}$$

$$Var(\bar{X}) = \frac{(N+1)(N-n)}{12N^2 n}$$

Question 2

Let $\hat{\theta}$ be an estimator of θ.

1) (1 point) Show that the mean squared error

$$\mathrm{E}[(\hat{\theta}-\theta)^2] = \mathrm{Var}(\hat{\theta}) + (\mathrm{E}(\hat{\theta})-\theta)^2$$

2) (1 point) Let X be random variable following a binomial distribution with n and $\theta = 1/2$. Consider two estimator $\hat{\theta}_1 = X/n$ and $\hat{\theta}_2 = (X+1)/(n+2)$. For what values of n is the mean square error of $\hat{\theta}_2$ less than the mean squared error of $\hat{\theta}_1$?

1) $$E\left[(\hat{\theta}-\theta)^2\right] = E\left[(\hat{\theta}-E\hat{\theta}+E\hat{\theta}-\theta)^2\right]$$

$$= E\left[(\hat{\theta}-E\hat{\theta})^2 + 2(E\hat{\theta}-\theta)(\hat{\theta}-E\hat{\theta}) + (E\hat{\theta}-\theta)^2\right]$$

$$= E(\hat{\theta}-E\hat{\theta})^2 + 2(E\hat{\theta}-\theta)E(\hat{\theta}-E\hat{\theta}) + (E\hat{\theta}-\theta)^2$$

$$= Var(\hat{\theta}) + (E\hat{\theta}-\theta)^2$$

2) $$Var\,\hat{\theta}_1 = \frac{\theta(1-\theta)}{n} \qquad E\hat{\theta}_1 = \theta = \frac{1}{2}$$

$$MSE(\hat{\theta}_1) = \frac{\theta(1-\theta)}{n} = \frac{1}{4n}$$

$$E\hat{\theta}_2 = \frac{n\theta+1}{n+2} = \frac{\frac{n}{2}+1}{n+2} = \frac{1}{2}$$

$$Var\,\hat{\theta}_2 = \frac{1}{(n+2)^2}\cdot n\theta(1-\theta) = \frac{n}{4(n+2)^2}$$

$$MSE(\hat{\theta}_2) = \frac{n}{4(n+2)^2} + \left(\frac{\frac{n}{2}+1-\frac{n}{2}-1}{n+2}\right)^2$$

$$= \frac{n}{4(n+2)^2}$$

$$MSE(\hat{\theta}_2) < MSE(\hat{\theta}_1) \iff \frac{(n+2)^2}{n} > n \iff (n+2)^2 > n^2$$

any integer

Question 3

1) (1 point) Show that sample mean $\bar{X}$ is a consistent estimator of the population mean.

2) (1 point) Let $X_1, \ldots, X_n$ be a random sample from uniform distribution from α to $\alpha + 1$. Show that

$$\hat{\alpha} = \min_{i=1,\ldots,n} X_i - \frac{1}{n+1}$$

is a consistent estimator of α.

1) $E\bar{X} = \mu$ $\quad$ $\bar{X}$ is unbiased

$Var\,\bar{X} = \frac{\sigma^2}{n}$ $\quad$ $\lim_{n\to\infty} \frac{\sigma^2}{n} \to 0$

Thus, $\bar{X}$ is a consistent estimator of the population mean

2) $P(\min_{i=1,\ldots,n} X_i \le x) = 1 - P(\min_{i=1,\ldots,n} X_i \ge x)$

$= 1 - P(X_i > x)^n$

$$= \begin{cases} 1 - (-x+\alpha+1)^n & \alpha < x < \alpha+1 \\ 1 & x \ge \alpha+1 \\ 0 & x \le \alpha \end{cases}$$

Let $Y = \min_{i=1,\ldots,n} X_i$

$F(y) = 1 - (-y+\alpha+1)^n \quad \alpha < y < \alpha+1$

$f(y) = \frac{dF(y)}{dy} = n(1+\alpha-y)^{n-1}$

$EY = \int_{\alpha}^{\alpha+1} n y (1+\alpha-y)^{n-1} dy$

$= n\int_{\alpha}^{\alpha+1} (y-1-\alpha)(1+\alpha-y)^{n-1} + (1+\alpha)(1+\alpha-y)^{n-1} dy$

$= -n\frac{1}{n+1} + n(1+\alpha)\frac{1}{n} = \alpha + \frac{1}{n+1}$

Thus. $E\hat{\alpha} = \alpha$, $\hat{\alpha}$ is an unbised estimator.

$Var\,\hat{\alpha} = Var\,Y = EY^2 - (EY)^2 = \frac{n}{(n+1)^2(n+2)}$ $\quad$ $\lim_{n\to\infty} Var\,\hat{\alpha} = 0$

Question 4

Use the factorization theorem to show that

1) (1 point) $\sum_{i=1}^{n} X_i$ is a sufficient statistic of θ, where $X_1, \ldots, X_n$ is a random sample from Bernulli distribution with parameter θ.

2) (1 point) $\bar{X}$ is a sufficient statistic of μ, where $X_1, \ldots, X_n$ is a random sample from $N(\mu, 1)$.

1) Joint Distribution of $X_1 \cdots, X_n$ is

$$f(x_1, \cdots, x_n; \theta) = \theta^{\Sigma x_i} (1-\theta)^{n - \Sigma x_i}$$

$$= g(\Sigma x_i, \theta)\, h(x_1, \cdots, x_n)$$

where $g(\Sigma x_i, \theta) = \theta^{\Sigma x_i} (1-\theta)^{n-\Sigma x_i}$

$h(x_1 \cdots x_n) = 1$

2) Joint density of $X_1, \cdots, X_n$ is

$$f(x_1, \cdots, x_n; \mu) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2}\right)$$

$$= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\bar{x}+\bar{x}-\mu)^2}{2}\right)$$

$$= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{2} - \frac{n(\bar{x}-\mu)^2}{2}\right) = g(\bar{x}; \mu)\, h(x_1 \cdots x_n)$$

where $g(\bar{x}; \mu) = \exp\left(-\frac{n(\bar{x}-\mu)^2}{2}\right)$

$$h(x_1 \cdots x_n) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{2}\right)$$