Kyle McGraw

ACM 157 Set 2

1. $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{\#1}{N}$ $\quad x_i \in \{0, 1\}$

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

b)
$$s^2 = \hat{\sigma}_n^2 \frac{Nn-n}{Nn-N} = \left(1-\frac{1}{N}\right)\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$$
$$= \left(1-\frac{1}{N}\right)\frac{1}{n-1}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x}_n + \bar{x}_n^2)$$
$$= \left(1-\frac{1}{N}\right)\frac{1}{n-1}\left(\sum_{i=1}^{n}x_i - 2\bar{x}_n\sum_{i=1}^{n}x_i + \bar{x}_n^2\sum_{i=1}^{n}1\right)$$
$$= \left(1-\frac{1}{N}\right)\frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n}x_i - 2\bar{x}_n\frac{1}{n}\sum_{i=1}^{n}x_i + \bar{x}_n^2\frac{1}{n}n\right)$$
$$= \left(1-\frac{1}{N}\right)\frac{n}{n-1}\left(\bar{x}_n - 2\bar{x}_n^2 + \bar{x}_n^2\right)$$
$$= \left(1-\frac{1}{N}\right)\frac{n}{n-1}\left(\bar{x}_n - \bar{x}_n^2\right)$$
$$= \left(1-\frac{1}{N}\right)\frac{n}{n-1}\bar{x}_n(1-\bar{x}_n)$$

c) $N=300$ $\quad n=90$ $\quad \#1 = 70$ $\quad \bar{x}_n = 70/90$ $\quad 95\% = 100(1-\alpha)\%$ $\quad \alpha = 0.05$

$$I = \bar{x}_n \pm z_{1-\frac{\alpha}{2}}\, se[\bar{x}_n]$$
$$= 70/90 \pm z_{1-\frac{0.05}{2}}\frac{s}{\sqrt{n}}\sqrt{\left(1-\frac{n-1}{N-1}\right)}$$
$$= 70/90 \pm z_{0.975}\sqrt{\left(1-\frac{1}{N}\right)\frac{n}{n-1}\bar{x}_n(1-\bar{x}_n)}\frac{1}{\sqrt{n}}\sqrt{\left(1-\frac{n-1}{N-1}\right)}$$
$$= 70/90 \pm 1.96\sqrt{\left(1-\frac{1}{300}\right)\frac{90}{89}\,70/90\,(1-70/90)}\frac{1}{\sqrt{90}}\sqrt{\left(1-\frac{89}{299}\right)}$$
$$= 70/90 \pm 0.0723$$
$$[0.706, 0.850]$$

2. 95% interval length ≤ 0.04

$$2 \cdot z_{1-\frac{\alpha}{2}} \, se[\overline{X}_n] \leq 0.04$$

$$2 \cdot 1.96 \cdot \frac{s}{\sqrt{n}} \leq 0.04 \quad \text{ignore population correction}$$

$$2 \cdot 1.96 \cdot \frac{\sigma}{\sqrt{n}} \leq 0.04 \quad s \approx \sigma$$

$$2 \cdot 1.96 \cdot \frac{\sqrt{0.5(1-0.5)}}{\sqrt{n}} \leq 0.04 \quad \sigma = \sqrt{\mu(1-\mu)} = \sqrt{0.5(1-0.5)} \quad \text{worst-case variance}$$

$$n \geq 2401$$

4. $X_1, \dots, X_n$ uniform on $[0, \theta]$ $\mu = \frac{\theta}{2}$ $\sigma^2 = \frac{\theta^2}{12}$ $\mathbb{V}[\overline{X}_n] = \frac{\sigma^2}{n}$

a) $\hat{\theta} = 2\overline{X}_n$

$$\text{bias}[\hat{\theta}] = \mathbb{E}[\hat{\theta}] - \theta = \mathbb{E}[2\overline{X}_n] - \theta = 2\mathbb{E}[\overline{X}_n] - \theta$$
$$= 2\frac{\theta}{2} - \theta = \theta - \theta = 0$$
$$se[\hat{\theta}] = \sqrt{\mathbb{V}[\hat{\theta}]} = \sqrt{\mathbb{V}[2\overline{X}_n]} = \sqrt{4\mathbb{V}(\overline{X}_n)} = \sqrt{4\frac{\sigma^2}{n}}$$
$$= \sqrt{4\frac{\theta^2}{12n}} = \theta\sqrt{\frac{1}{3n}}$$
$$MSE[\hat{\theta}] = \text{bias}[\theta]^2 + se[\hat{\theta}]^2 = \frac{\theta^2}{3n}$$

b) $\hat{\theta} = X_{(n)} = \max\{X_1, \dots, X_n\}$

CDF: $F(x) = P(X_{(n)} \leq x) = P(X_1 \leq x)P(X_2 \leq x) \cdot \ldots \cdot P(X_n \leq x)$
$$= \frac{x}{\theta} \cdot \frac{x}{\theta} \cdot \ldots \cdot \frac{x}{\theta} = \left(\frac{x}{\theta}\right)^n$$

$$f(x) = \frac{nx^{n-1}}{\theta^n}$$

$$\mathbb{E}[\hat{\theta}] = \int_0^\theta x \, dF(x) = \int_0^\theta x f(x) dx = \int_0^\theta n\frac{x^n}{\theta^n} dx = \frac{n}{n+1}\frac{x^{n+1}}{\theta^n}\Big|_0^\theta$$
$$= \frac{n}{n+1}\frac{\theta^{n+1}}{\theta^n} = \frac{n}{n+1}\theta$$

$$\text{bias}[\hat{\theta}] = \mathbb{E}[\hat{\theta}] - \theta = \frac{n}{n+1}\theta - \theta = -\frac{1}{n+1}\theta$$
$$\mathbb{E}[\hat{\theta}^2] = \int_0^\theta x^2 dF(x) = \int_0^\theta x^2 f(x) dx = \int_0^\theta n\frac{x^{n+1}}{\theta^n} dx = \frac{n}{n+2}\frac{x^{n+2}}{\theta^n}\Big|_0^\theta$$
$$= \frac{n}{n+2}\frac{\theta^{n+2}}{\theta^n} = \frac{n}{n+2}\theta^2$$
$$\mathbb{V}[\hat{\theta}] = \mathbb{E}[\hat{\theta}^2] - \mathbb{E}[\hat{\theta}]^2 = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2 = \frac{n(n+1)^2}{(n+2)(n+1)^2}\theta^2 - \frac{n^2(n+2)}{(n+2)(n+1)^2}\theta^2$$
$$= \frac{1}{(n+2)(n+1)^2}\theta^2\left(n^3 + 2n^2 + n - n^3 - 2n^2\right) = \frac{n}{(n+2)(n+1)^2}\theta^2$$
$$se[\hat{\theta}] = \sqrt{\mathbb{V}[\hat{\theta}]} = \theta\sqrt{\frac{n}{(n+2)(n+1)^2}}$$

$$MSE[\hat{\theta}] = \text{bias}[\theta]^2 + se[\hat{\theta}]^2 = \left(-\frac{1}{n+1}\theta\right)^2 + \frac{n}{(n+2)(n+1)^2}\theta^2$$
$$= \frac{1}{(n+1)^2}\theta^2 + \frac{n}{(n+2)(n+1)^2}\theta^2 = \frac{2n+2}{(n+2)(n+1)^2}\theta^2 = \frac{2}{(n+2)(n+1)}\theta^2$$

c)

We can see that $X_{(n)}$ is more efficient at $n>2$ because:

$$X_{(n)} \text{ MSE} < 2\bar{X}_n \text{ MSE}$$

$$\frac{2}{(n+2)(n+1)}\theta^2 \leq \frac{\theta^2}{3n}$$

$$6n \leq (n+2)(n+1)$$

$$n^2+3n+2-6n \geq 0$$

$$n^2-3n+2 \geq 0$$

$$(n-2)(n-1) \geq 0$$

equal at $n=1,2$

greater at $n>2$