HW 8 Solution

7.23 CLT states that $Y_n = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ converges in distribution to a standard normal random variable, Z.

Here, $n = 100$, $\sigma = 2.5$, and the approximation is

$$P(|\bar{X} - \mu| \le 0.5) = P(-.5 \le \bar{X} - \mu \le .5)$$
$$= P\left(\frac{-.5\times 10}{2.5} \le Z \le \frac{.5 \times 10}{2.5}\right) = P(-2 \le Z \le 2)$$
$$= 1 - 2(.0228) = .9544$$

7.37 Denote X_i = the length of life for the i^{th} lamp, $i = 1, 2, \cdots, 25$.

Given X_i's are independent with $\mu = 50$ & $\sigma = 4$.

By CLT, $Y_n = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} = \frac{n(\bar{X} - \mu)}{\sqrt{n}\,\sigma} = \frac{\Sigma X_i - n\mu}{\sqrt{n}\,\sigma}$ converges to a standard normal R.V. in distribution, then

$$P\left(\sum_{i=1}^{25} X_i \ge 1300\right) = P\left(\frac{\Sigma X_i - n\mu}{\sqrt{n}\,\sigma} \ge \frac{1300 - 1250}{5 \times 4}\right) = P(Z > 2.5) = .0062$$

7.47 Let X = # of persons not showing up for a given flight

$X \sim \text{Binomial}(n, p)$, where $n = 160$ & $p = .05$

If there is to be a seat available for everyone planning to fly, then there must be at least 5 persons not showing up, i.e. $X \ge 5$.

Calculate $\mu = np = 160(.05) = 8$, $\sigma = \sqrt{npq} = 2.76$

A correction for continuity (Refer to example 7.10) is made to include the entire area under the rectangle associated with $X = 5$, and the approximation becomes $P(X \ge 4.5)$.

Correspondly, $Z = \frac{4.5 - 8}{2.76} = -1.27$

so that $P(X \ge 4.5) = P(Z \ge -1.27) = 1 - P(Z \le 1.27) = 1 - .1020$
$= .8980$

8.8 a. $Y_i \sim \text{Uniform}(\theta, \theta+1) \Rightarrow EY_i = \theta + \frac{1}{2} \quad (i = 1, 2, \cdots, n)$

$\Rightarrow E(\bar{Y}) = \theta + \frac{1}{2}$

& the bias is $B = E(\bar{Y}) - \theta = \frac{1}{2}$

b. An unbiased estimator of θ can be constructed by using $\hat{\theta} = \bar{Y} - \frac{1}{2}$, which has $E(\hat{\theta}) = \theta$

c. If $\bar{Y}$ is used as an estimator, then

$V(\bar{Y}) = \frac{1}{n} V(Y) = \frac{1}{n} \int_\theta^{\theta+1} [y - (\theta + \frac{1}{2})]^2 dy = \frac{1}{12n}$

& $MSE = V(\bar{Y}) + B^2$ (See Def'n 8.4) $= \frac{1}{12n} + \frac{1}{4}$

8.42 (a) $\hat{p} = 268/500 = .536$

Therefore, an approximate 98% confidence interval for p is

$\hat{p} \pm z_{0.01} \sqrt{\frac{\hat{p}\hat{q}}{n}}$ (see 2^{nd} line on P391)

$= .536 \pm 2.33 \sqrt{\frac{(.536)(.464)}{500}} = .536 \pm .052 = (.484, .588)$

8.46 Given $n = 75$, $\bar{y} = 4.2$, $s = 1.5$ & $\alpha = 0.05 \Rightarrow z_{0.025} = 1.96$

So, a 95% confidence interval is

$\bar{y} \pm z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) = 4.2 \pm 1.96\left(\frac{1.5}{\sqrt{75}}\right) = 4.2 \pm .34 = (3.86, 4.54)$

8.47 The 95% confidence interval is

Refer to example 8.8, $\text{var}(\bar{y}_1 - \bar{y}_2) = \text{Var}\,\bar{y}_1 + \text{Var}\,\bar{y}_2 = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}$

$\rightarrow (\bar{y}_1 - \bar{y}_2) \pm z_{0.025} \sqrt{\left(\frac{S_1}{\sqrt{n_1}}\right)^2 + \left(\frac{S_2}{\sqrt{n_2}}\right)^2}$

$= (167.1 - 140.9) \pm 1.96 \sqrt{\frac{(24.3)^2}{30} + \frac{(17.6)^2}{30}}$

$= 26.2 \pm 10.74 = (15.46, 36.94)$

8.53 The 98% confidence interval is (notice $z_{\alpha/2} = z_{0.01} = 2.33$)

$$(\hat{p}_1 - \hat{p}_2) \pm 2.33\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$= (.18 - .12) \pm 2.33\sqrt{\frac{(.18)(.82)}{100} + \frac{(.12)(.88)}{100}}$$

$$= .06 \pm .12 = (-0.06,\ 0.18)$$

B/C $(-0.06, 0.18)$ includes 0, there is no enough evidence here to suggest that one line produces a higher proportion of defectives than the other.

8.62 The 95% confidence interval for μ is

$$\hat{\mu} \pm z_{0.025}\left(\frac{s}{\sqrt{n}}\right)$$

we know $z_{0.025}\left(\frac{s}{\sqrt{n}}\right) = 1.96\left(\frac{.5}{\sqrt{n}}\right) = .1 \Rightarrow n = 97$

It's not valid if we select all our water specimens from a single rainfall since all the observations should be independent.

8.63 The parameter to be estimated is $\mu_1 - \mu_2$

Given $1-\alpha = 0.9$, $z_{0.05} = 1.645$ & $\sigma_1^2 = \sigma_2^2 = 0.25$, $n_1 = n_2 = n$.

We require $1.645\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le 0.1$

$$1.645\sqrt{\frac{0.25 + 0.25}{n}} \le 0.1 \Rightarrow n \ge 135.3$$

So, $n = 136$ samples should be selected at each location.

9.2 a. Since $EY_i = \mu, \quad i = 1, 2, \cdots, n$

$\therefore E(\hat{\mu}_1) = \frac{1}{2}[E(Y_1) + E(Y_2)] = \mu$

$E(\hat{\mu}_2) = \frac{1}{4}E(Y_1) + \frac{1}{2(n-2)}\sum_{i=2}^{n-1} E(Y_i) + \frac{1}{4}E(Y_n)$

$= \left(\frac{1}{4} + \frac{1}{2(n-2)} \cdot (n-2) + \frac{1}{4}\right)\mu = \mu$

$E(\hat{\mu}_3) = E\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} E(Y_i) = \frac{1}{n} \cdot n\mu = \mu$

b. $V(\hat{\mu}_1) = \frac{1}{4}V(Y_1) + \frac{1}{4}V(Y_2) = \frac{1}{4} \times \sigma^2 \times 2 = \frac{1}{2}\sigma^2$

$V(\hat{\mu}_2) = \frac{1}{4^2}V(Y_1) + \frac{1}{[2(n-2)]^2}\sum_{i=2}^{n-2} V(Y_i) + \frac{1}{4^2}V(Y_n)$

$= \left[\frac{1}{16} + \frac{1}{4(n-2)} + \frac{1}{16}\right]\sigma^2 = \frac{n}{8(n-2)}\sigma^2$

$V(\hat{\mu}_3) = \frac{1}{n}V(Y_i) = \frac{1}{n}\sigma^2$

$\therefore \text{eff}(\hat{\mu}_3, \hat{\mu}_1) = \frac{V(\hat{\mu}_1)}{V(\hat{\mu}_3)} = \frac{\frac{1}{2}}{\frac{1}{n}} = \frac{n}{2}$

$\text{eff}(\hat{\mu}_3, \hat{\mu}_2) = \frac{V(\hat{\mu}_2)}{V(\hat{\mu}_3)} = \frac{\frac{n}{8(n-2)}}{\frac{1}{n}} = \frac{n^2}{8(n-2)}$

9.3 For $Y_i \sim \text{Uniform}(\theta, \theta+1)$, $E(Y_i) = \theta + \frac{1}{2}$, $V(Y_i) = \frac{1}{12}$.

$E(\hat{\theta}_1) = E(\bar{Y} - \frac{1}{2}) = \theta + \frac{1}{2} - \frac{1}{2} = \theta$

& $V(\hat{\theta}_1) = V(\bar{Y}) = \frac{1}{n}V(Y_i) = \frac{1}{12n}$

Recall the density function of $Y_{(n)}$ (see Section 6.7) is

$g_n(y) = n[F(y)]^{n-1}f(y)$, where $F(y) = \int_\theta^y dt = y - \theta$

& $f(y) = 1$

$\therefore g_n(y) = n(y-\theta)^{n-1}$ for $\theta \le y \le \theta + 1$

$\therefore E(\hat{\theta}_2) = E(Y_{(n)}) - \frac{n}{n+1} = \int_\theta^{\theta+1} y \cdot n(y-\theta)^{n-1}\,dy - \left(\frac{n}{n+1}\right)$

$\overset{x=y-\theta}{=} \int_0^1 n(x+\theta)x^{n-1}\,dx - \left(\frac{n}{n+1}\right) = \left(\frac{n}{n+1} + \theta\right) - \frac{n}{n+1} = \theta$

& $V(\hat{\theta}_2) = V(Y_{(n)}) = E(Y_{(n)}^2) - [E(Y_{(n)})]^2 = \int_\theta^{\theta+1} y^2 \cdot n(y-\theta)^{n-1}\,dy$

$- \left(\frac{n}{n+1} + \theta\right)^2 \overset{x=y-\theta}{=} \int_0^1 (x+\theta)^2 \cdot nx^{n-1}\,dx - \left(\frac{n}{n+1} + \theta\right)^2 = \frac{n}{(n+2)(n+1)^2}$

$\therefore \text{eff}(\hat{\theta}_1, \hat{\theta}_2) = V(\hat{\theta}_2) \Big/ V(\hat{\theta}_1) = \frac{n}{(n+2)(n+1)^2} \Big/ \frac{1}{[illegible]} = \frac{12n^2}{[illegible]}$

9.34) $f(y|p) = p(1-p)^{y-1}$

$L(y_1, \cdots, y_n | p) = p^n (1-p)^{\Sigma y_i - n} = p^n (1-p)^{n\bar{y}-n}$

let $g(\bar{y}, p) = p^n (1-p)^{n\bar{y}-n}$ and $h(y_1, \cdots, y_n) = 1$

then by Thm 9.4, $\bar{Y}$ is sufficient for p.

9.37) The likelihood is $L(y_1, \cdots, y_n | \theta) = a(\theta)^n \prod b(y_i) e^{-c(\theta)\Sigma d(y_i)}$

let $g(\Sigma d(Y_i), \theta) = a(\theta)^n e^{-c(\theta)\Sigma d(y_i)}$ and $h(y_1, \cdots, y_n) = \prod b(y_i)$

Then by Thm 9.4. $\sum_{i=1}^{n} d(Y_i)$ is sufficient for θ.

9.44) Note that $f(y_1, y_2, \cdots, y_n | \theta) = \begin{cases} \prod_{i=1}^{n} 3y_i^2/\theta^3 & \text{for } 0 \le y_i \le \theta \\ 0 & \text{otherwise} \end{cases}$

Note also the condition that $0 \le y_i \le \theta$ for $i = 1, \cdots, n$ can be written as $0 \le y_{(1)}$ and $y_{(n)} \le \theta$ (if the minimum is greater than 0, then so must be the remainder of the y_i; the same argument applies for the maximum being less than θ) The likelihood can then be written as

$$L(y_1, \cdots, y_n | \theta) = \frac{3^n}{\theta^{3n}} \left(\prod_{i=1}^{n} y_i\right) I(0 \le y_{(1)}) I(y_{(n)} \le \theta)$$

Thus with $g(y_{(n)}, \theta) = \frac{3^n}{\theta^{3n}} I(y_{(n)} \geq \theta)$ and

$h(y_1, \cdots, y_n) = \left(\prod_{i=1}^{n} y_i\right) I(0 \le y_{(1)})$, we can see that by the factorization theorem, $Y_{(n)}$ is sufficient for θ.