**Solution to Problem 8.58**

---

**8.58.** If we refer to $AA, Aa$ and $aa$ as categories $1, 2$ and 3, then

$$\begin{aligned} p_1(\theta) &= (1-\theta)^2 \\ p_2(\theta) &= 2\theta(1-\theta) \\ p_3(\theta) &= \theta^2. \end{aligned}$$

Note that $p_1(\theta) + p_2(\theta) + p_3(\theta) = 1$ and all are legitimate probabilities ($0 \le p_i(\theta) \le 1$) provided $0 \le \theta \le 1$.

a) This is a multinomial sampling situation where the probabilities depend upon the unknown parameter $\theta$. If the observed frequencies for the categories are denoted $x_1, x_2$ and $x_3$ ($x_1+x_2+x_3 = n$), the log-likelihood is given by (see page 273)

$$l(\theta) = \log L(\theta) = \log n! - \sum_{i=1}^{m} \log x_i! + \sum_{i=1}^{m} x_i \log p_i(\theta).$$

Substituting in the above expressions for the probabilities $p_i(\theta)$, we obtain

$$l(\theta) = \log n! - \sum_{i=1}^{3} \log x_i! + x_1 \log((1-\theta)^2) + x_2 \log(2\theta(1-\theta)) + x_3 \log(\theta^2),$$

and, after a bit of simplification, this becomes

$$l(\theta) \propto (2x_1 + x_2) \log(1-\theta) + (2x_3 + x_2) \log \theta.$$

Taking the derivative with respect to $\theta$ yields

$$l'(\theta) = -\frac{2x_1 + x_2}{(1-\theta)} + \frac{x_2 + 2x_3}{\theta}.$$

Setting $l'(\theta)$ to zero, and solving for $\theta$ gives

$$\theta = \frac{1}{2} \frac{x_2 + 2x_3}{x_1 + x_2 + x_3} = \frac{x_2 + 2x_3}{2n}.$$

Because $l''(\theta) < 0$ at this value of $\theta$ (in fact, at all permitted values of $\theta$), this root corresponds to a local maximum of the log-likelihood. But $L(\theta) > 0$ for $0 < \theta < 1$ and $L(\theta) \to 0$ as $\theta \to 0$ (provided $x_2 > 0$ or $x_3 > 0$) and as $\theta \to 1$ (provided $x_1 > 0$ or $x_2 > 0$), so this is a global maximum; that is,

$$\hat{\theta}_{ML} = \frac{X_2 + 2X_3}{2n}.$$

Alternately, you could check that $l'(\theta) > 0 \Leftrightarrow \theta < (x_2 + 2x_3)/2n$, leading to the same conclusion.

Substituting in the observed values, the maximum likelihood estimate for this data set (the value of the ML estimator) becomes $\hat{\theta}_{ML} = (68 + 2 \times 112)/(2 \times 190) \cong 0.768$.

b) The asymptotic variance of $\hat{\theta}_{ML}$ is given by $1/nI(\theta)$, where $nI(\theta)$, the Fisher Information for $\theta$ in the sample, can be evaluated as

$$nI(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right].$$

The second derivative is given by

$$\begin{aligned} l''(\theta) &= -\left(\frac{2x_1 + x_2}{(1-\theta)^2} + \frac{x_2 + 2x_3}{\theta^2}\right) \\ \Rightarrow nI(\theta) &= \frac{2EX_1 + EX_2}{(1-\theta)^2} + \frac{EX_2 + 2EX_3}{\theta^2} \end{aligned}$$

But each $X_i$ is binomial with probability $p_i(\theta)$, so we have

$$\begin{aligned} EX_1 &= n(1-\theta)^2 \\ EX_2 &= 2n\theta(1-\theta) \\ EX_3 &= n\theta^2 \\ \Rightarrow nI(\theta) &= \frac{2n}{\theta(1-\theta)}. \end{aligned}$$

Of course, the true value of $\theta$ is unknown, so we plug-in the value of $\hat{\theta}_{ML} \cong 0.768$ to obtain the estimated asymptotic variance of $\hat{\text{Var}}(\hat{\theta}_{ML}) = 1/nI(\hat{\theta}_{ML}) \approx 0.00047$. Alternately, we have $\hat{SE}(\hat{\theta}_{ML}) \cong \sqrt{0.00047} \cong 0.0216 \approx 0.022$.

**Note:** As in problem 8.4, you can determine the exact expression for the variance of $\hat{\theta}_{ML}$ using the properties of the multinomial distribution. You should check that you get exactly the same expression as above; that is, $\text{Var}(\hat{\theta}_{ML}) = \theta(1-\theta)/2n$.

c) Recall that, for large values of $n$, $\sqrt{nI(\theta)}(\hat{\theta}_{ML} - \theta)$ has approximately a standard normal distribution. Putting this another way, for large values of $n$,

$$\hat{\theta}_{ML} \approx N\left(\theta,\, SE^2(\hat{\theta}_{ML})\right),$$

where $SE^2(\hat{\theta}_{ML}) = 1/nI(\theta)$. So the form of the approximate $1-\alpha$ confidence interval (CI) for $\theta$ is given by

$$\hat{\theta}_{ML} \pm z(\alpha/2)\, \hat{SE}^2(\hat{\theta}_{ML}),$$

where we plugged $\hat{\theta}_{ML}$ into the expression for $SE^2(\hat{\theta}_{ML})$ to obtain $\hat{SE}^2(\hat{\theta}_{ML})$. As $z(0.01/2) \cong 2.58$, the approximate 99% CI for $\theta$ becomes $0.768 \pm 2.58 \times 0.0216$, or the interval $(0.7126,\ 0.8243) \approx (0.71,\ 0.82)$.

d) We generate bootstrap samples (of size $n = 190$) from a multinomial distribution with probabilities:

$$\begin{aligned} p_1 &= (1-\hat{\theta}_{ML})^2 \\ p_2 &= 2\hat{\theta}_{ML}(1-\hat{\theta}_{ML}) \\ p_3 &= \hat{\theta}_{ML}^2 \end{aligned}$$

or $0.0536, 0.3559$ and $0.5905$, respectively. For each of these bootstrap samples, we calculate the MLE. Doing this $B = 1000$ times, and taking the standard deviation of these 1000 bootstrap MLE's gives $\hat{SE}(\hat{\theta}_{ML}) \cong 0.02088 \approx 0.021$, essentially the same value obtained for $\hat{SE}(\hat{\theta}_{ML})$ in part b). Of course, you will get a slightly different answer because your bootstrap samples will be different (you would also get a different answer if you generated a different number of bootstrap samples) but your answer should be pretty close to the value obtained in part b) as long as you use a fairly large value of $B\,(B > 200$ say).

A piece of R code to do this is:

```
mlOuts <- rep(NA, 1000)
for (i in 1:1000){
   x <- rmultinom(1, 190, prob=c(0.0536, 0.3559, 0.5905))
   mlOuts[i] <- 0.5*(x[2]+2*x[3])/190)
}
sd(mlOuts)
```

e) Using the notation on page 284 of the text, we have $\hat{\theta} = 0.768$, and we use the bootstrap samples as in d) to generate estimates $\theta^*$. The distribution of $\hat{\theta} - \theta$ is then approximated by the histogram of $\theta^* - \hat{\theta}$ to obtain estimates of the desired quantiles of the distribution of $\hat{\theta} - \theta$. Carrying this out leads to an approximate 99% CI for $\theta$ of $(0.714, 0.827)$. This interval is very similar to the approximate 99% CI obtained in c).

A piece of R code to do this is:

```
mlDiffs <- rep(NA, 1000)
for (i in 1:1000){
   x <- rmultinom(1, 190, prob=c(0.0536, 0.3559, 0.5905))
   mlOuts <- 0.5*(x[2]+2*x[3])/190)
   mlDiffs[i] <- mlOuts - 0.768
}
0.768 - quantile(mlDiffs, c(0.005, 0.995))
```