Miscellaneous Solutions and Comments on the Chapter 7 Exercises

7.4: For $X_1, \ldots, X_n$ iid $N(\theta, 1)$, the likelihood function may be found by setting $\mu = \theta$ and $\sigma^2 = 1$ in the expression for the joint density given in Example 6.2.7 on page 277. This leads to

$$L(\theta|\boldsymbol{x}) = c\exp(-n(\bar{x}-\theta)^2/2)$$

where

$$c = (2\pi)^{-n/2}\exp\left(-\sum_{i=1}^{n}(x_i-\bar{x})^2/2\right).$$

The log-likelihood is thus

$$\ell(\theta) = \log c - \frac{n}{2}(\bar{x}-\theta)^2 \tag{1}$$

which as a function of θ is a parabola which opens downward and achieves its maximum over **all** real values of θ at $\theta = \bar{x}$. In this problem we require $\theta \geq 0$. Let $\Theta = \{\theta : \theta \geq 0\}$. The MLE is the value of $\theta \in \Theta$ which maximizes $\ell(\theta)$. If $\bar{x} \geq 0$, then the overall maximum at $\theta = \bar{x}$ lies in Θ and is the MLE. However, if $\bar{x} < 0$, then $\theta = \bar{x}$ lies outside Θ. When $\bar{x} < 0$ the log-likelihood is a decreasing function when restricted to Θ and the maximum in Θ is achieved at the endpoint $\theta = 0$ which is the MLE in this case. (This is clear if you draw a picture of a parabola opening downward with its peak at $\bar{x} < 0$.) In summary, the MLE is $\hat{\theta} = \bar{x}$ when $\bar{x} \geq 0$ and $\hat{\theta} = 0$ otherwise.

You can reach the same conclusion without using the particular expression for $\ell(\theta)$ given in (1) above. You just need to show that $\ell'(\theta) = 0$ when $\theta = \bar{x}$ with $\ell'(\theta) > 0$ for $\theta < \bar{x}$ and $\ell'(\theta) < 0$ for $\theta > \bar{x}$. Then if $\bar{x} < 0$, you have $\ell'(\theta) < 0$ for all $\theta \geq 0$ so that the maximum in Θ is achieved at $\theta = 0$.

7.19, 7.20, 7.21 (Comparison of Variances): It is possible to show that the estimator in 7.19 has smaller variance than the estimators in 7.20 and 7.21 without any calculations or explicit use of general inequalities (such as Jensen's, etc.). By referring to the expression for the likelihood in the solution manual, we see that the joint distribution of $Y_1, \ldots, Y_n$ in 7.19 is a two-parameter exponential family (2pef) with natural sufficient statistic $T = (\sum_i Y_i^2, \sum_i x_i Y_i)$. The open set condition is satisfied so that T is also complete. The unbiased estimator found in 7.19(b) is a function of the complete sufficient statistic T (keeping in mind that $\sum_i x_i^2$ is a "constant", i.e., not a function of the data $\boldsymbol{Y}$) and is therefore best unbiased. Its variance is thus **strictly** smaller than that of the unbiased estimators in 7.20 and 7.21 (except in the special case where $x_1 = x_2 = \cdots = x_n$ in which the estimators coincide and their variances are equal).

Comparing the variances of the estimators in 7.20 and 7.21 with each other does require explicit formulas for the variances and the use of Jensen's inequality (as shown in the solutions manual). You are **NOT** responsible for knowing the Jensen's argument. However, you **ARE** responsible for being able to calculate the variances.

7.22, parts (b) and (c): The solution manual gives one approach to part (c). Another is given in the lecture notes.

The solution to part (b) in the solution manual involves some messy calculations. An alternate approach uses the following general facts:

- For any two random variables Y and Z, if $\mathcal{L}(Y|Z)$ does **not** depend on Z (i.e., is the same for all values of Z), then Y and Z are independent and $\mathcal{L}(Y)$ is the same as $\mathcal{L}(Y|Z)$.
- If Y and Z are independent with $Y \sim N(a, b)$ and $Z \sim N(c, d)$, then $Y + Z \sim N(a + c, b + d)$.

We know that $\bar{x} \,|\, \theta \sim N(\theta, \sigma^2/n)$. Therefore $(\bar{x} - \theta) \,|\, \theta \sim N(0, \sigma^2/n)$. Since this conditional distribution does not involve θ, we conclude that $\bar{x} - \theta$ and θ are independent random variables and that $\bar{x} - \theta \sim N(0, \sigma^2/n)$. We are given that $\theta \sim N(\mu, \tau^2)$. Thus $\bar{x} = (\bar{x} - \theta) + \theta$ is the sum of two independent normal random variables, and is therefore normally distributed with mean $0 + \mu$ and variance $\sigma^2/n + \tau^2$ as desired.

7.42(a): The handwritten solution involves somewhat complicated mucking about with inequalities. Here is another solution which avoids this. Let $Z_1, \ldots, Z_n$ be independent random variables with $Z_i \sim N(\theta, \sigma_i^2)$ for all i where $\sigma_1^2, \ldots, \sigma_n^2$ are known constants. For any constants $a_1, \ldots, a_n$, it is clear that $E(\sum_i a_i W_i) = E(\sum_i a_i Z_i)$ and $\text{Var}(\sum_i a_i W_i) = \text{Var}(\sum_i a_i Z_i)$. Therefore, since this problem only involves these means and variances, we may assume without loss of generality that $W_i = Z_i$ for all i. It is easy to show that the joint distribution of $Z_1, \ldots, Z_n$ forms a one-parameter exponential family (1pef) for which $T = \sum_i Z_i/\sigma_i^2$ is the natural sufficient statistic for θ. (To see this just write down and simplify the joint density.) Since the open set condition holds, T is also complete. The unbiased estimator $\hat{\theta} = (\sum_i Z_i/\sigma_i^2) \,/\, (\sum_i 1/\sigma_i^2)$ is a function of the complete sufficient statistic T, and is therefore the best unbiased estimator. Thus, any unbiased estimator of the form $\sum_i a_i Z_i$ must satisfy $\text{Var}(\sum_i a_i Z_i) \geq \text{Var}(\hat{\theta})$. (The inequality will be strict unless $\sum_i a_i Z_i = \hat{\theta}$.)

7.57: In the case $y = 1$ or 2, the final ratio is simplified incorrectly in the solution manual. The answer should be $(n+1-y)/(n+1)$ when $y = 1$ or 2. The solution manual does everything in terms of the binomial distribution. Define $Y = \sum_{i=1}^{n+1} X_i$. Following the lecture notes, another solution can be based on the fact that, conditional on $Y = y$, the values $X_1, \ldots, X_{n+1}$ are like $n + 1$ balls drawn (without replacement) from an urn containing y balls labeled 1 and $n+1-y$ balls labeled 0. Let $A = \{\sum_{i=1}^{n} X_i > X_{n+1}\}$. Using the urn model it is clear that A is impossible if $Y = 0$, that is, $P(A \,|\, Y = 0) = 0$. If $Y = 1$ or 2, then A occurs only if $X_{n+1} = 0$ (the last ball is zero) which happens with probability $(n + 1 - y)/(n + 1)$. If $Y > 2$, then clearly A must always occur (the conditional probability is one). This gives the same answer as in the solution manual.