MA20226 - Feedback on Practical Sheet One

Simon Shaw, s.shaw@bath.ac.uk
http://people.bath.ac.uk/masss/ma20226.html

2011/12 Semester I

1. A surprisingly large number of people who handed work in failed to do this question; I'm not sure why as those who did it did it perfectly. The only note I'd make here is an algebra saving one in that

$$E^2(M \mid \theta) = \left(1 - \frac{1}{n+1}\right)^2 \theta^2 = \left(\frac{n}{n+1}\right)^2 \theta^2 = \frac{n^2}{(n+1)^2}\theta^2$$

is easier to work with than

$$E^2(M \mid \theta) = \left(1 - \frac{1}{n+1}\right)^2 \theta^2 = \left(1 - \frac{2}{n+1} + \frac{1}{(n+1)^2}\right)\theta^2$$

2. Generally, there were no problems here.

3. Make sure that you read the question carefully and think about what information you are being given and how you are going to use it. In this question, and question 4., you were gaining "insight into the theoretical results for the sampling distribution, mean and variance" of M (as stated in the aim of the sheet). The sampling distribution, mean and variance of M depend upon θ and the number of $X_i \sim U(0, \theta)$ observed which on this sheet was $n = 15$. Thus, $f_M(m \mid \theta = 16) = 15m^{14}/16^{15}$ for $m \leq 16$, $E(M \mid \theta = 16) = 15$ and $Var(M \mid \theta = 16) = 15/17$. In this question we take 50 observations of M, stored in `obsmax50`. Theoretically, we're considering 50 iid observations $M_1, \ldots, M_{50}$ with each M_i having an expectation of 15, a variance of 15/17 and sampling distribution $15m^{14}/16^{15}$ for $m \leq 16$.

   (a) We consider M_{47}. This has expectation 15 so we'd expect to see 15. Typically, you'll see something not particularly close to 15 (and mostly something larger than 15) as you only have one observation.

   (b) Plot the observed $m_1, \ldots, m_{50}$. Most of you did this well, although you needed to ensure that the histogram was a representation of a probability density ("your histogram should have a total area of one") by setting the option `freq=FALSE` or `probability=TRUE` in your call of `hist`. The overlaying of the sampling distribution should show the curve reasonably matching the shape of the histogram. Don't be afraid to comment if you thought something should happen but it didn't on your plot (e.g. if your curve didn't appear). You'll get no marks for no comments but I do look favourably on those who know that something is wrong and tell me so.

(c) Here we utilise a result from Lecture 5 (Example 15 on p13 of the on-line notes). It's that $\overline{M} = \frac{1}{50}\sum_{r=1}^{50} M_r$ is an unbiased estimator of $E(M_i \mid \theta = 16) = 15$. Thus, we expect to see 15 and you should get something closer to 15 than in part (a) as we're averaging over 50 observations rather than one.

(d) Here we utilise a second result from Lecture 5 (Example 16 on p13 of the on-line notes). It's that $\frac{1}{49}\sum_{r=1}^{50}(M_i - \overline{M})^2$ is an unbiased estimator of $Var(M_i \mid \theta = 16) = 15/17$. (Use `help(var)` to confirm that `var` computes s^2).

4. For this question we observe 1000 M_is rather than 50 but the distribution of each M_i is unchanged. What you should see is the effect of trying to estimate with 1000 observations rather than 50: all of your results should be closer to the theoretical values. Note that if you didn't make your histogram a representation of a probability density then the scale of the default frequency plot meant that a correct sampling distribution would not be visible on the plot. Don't forget to comment: a large number neglected to do so and so dropped marks unnecessarily.