MA20226 - Feedback on Question Sheet Four

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

2011/12 Semester I

Overall, this was another sheet that was well done with most people performing strongly on questions 1. and 2. though there was a little more uncertainty about the use of a pivot in question 3.

1. In calculating s^2 you have to be careful about accuracy as the calculation suffers from ill-conditioning. The range of the data is three so, given the size of the observations, the variance is going to be small. This means that when calculating s^2 using the formula $\frac{1}{n-1}\left\{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right\}$, $\sum_{i=1}^{n} x_i^2$ will be close to $n\overline{x}^2$ and so you need to be careful about the accuracy of $\overline{x}$. For example, if you round the sample mean to two decimal places, so use $\overline{x} = 60.69$, then you will find

$$s^2 = \frac{1}{14}\left\{55254.35 - 15(60.69)^2\right\} = \frac{5.2085}{14} = 0.3720$$

rather than the correct value obtained using $\overline{x} = 910.3/15$ which is

$$s^2 = \frac{1}{14}\left\{55254.35 - 15\left(\frac{910.3}{15}\right)^2\right\} = \frac{11.2773}{14} = 0.8055.$$

2. Part (c) is an illustration of the duality of confidence intervals and hypothesis testing. The equivalent hypothesis test is a t-test which we'll explore in Lecture 18. As 100 is not in the 95% confidence interval for μ then we would reject H_0 in the test of

$$H_0 : \mu = 100 \quad \text{versus} \quad H_1 : \mu \neq 100$$

at the 5% level. On the solution sheet I additionally calculate a 98% confidence interval for μ which also doesn't contain 100 and a 99% confidence interval which does contain 100. What this tells you is the p-value of the hypothesis test is somewhere between 0.01 and 0.02 (it's actually 0.01737).

3. In part (b) people either knew how to use a pivot or they didn't. Remember that a pivot is just a function of the data and the parameter whose distribution, given the value of the parameter, does not depend upon the parameter. (See Definition 9, p20 of the on-line notes.) Here, we have $Y = \frac{2n}{\theta}\overline{X}$ is a pivot for θ as the distribution of Y, given θ, is χ^2_{2n} which does not depend upon θ. Even if I don't know θ, I know everything about the distribution χ^2_{2n}. In particular, I can find constants c_1 and c_2 such that $P(c_1 < \chi^2_{2n} < c_2) = 1 - \alpha$ which is the starting point of constructing a confidence interval; the general construction approach can be found on p21 of the on-line notes.