CP1 Feedback

Overall Feedback

There were many very good reports, and many people clearly took time to engage with the practical.

Some marks were lost for omissions that could be avoided by reading questions carefully and ensuring that they have been answered.

It seems that some students were confused about what certain questions were asking, yet did not post questions on the discussion board or email me.

Question 1

- Most students answered this correctly.
- Some students did not even provide a brief derivation, and just stated the answer; these did not receive a mark.

Question 2

- Many students answered this correctly.
- Some students used more complicated simulations than were necessary (no marks lost).
- Some students didn't indicate the theoretical value or compare it with the output of the simulation procedure; this lost the mark as it is unclear what is being compared.
- Quite a few students plotted a histogram of the sample variance of many sets of realizations of Binomial(13, 0.31) random variables, as well as a line showing the theoretical value. This is not a very convincing way to empirically verify the result, but no marks were deducted because the overall idea (using the sample variance of many realizations) was correct.
- The simplest solution is to simply simulate a large number (e.g. 100000) of realizations of Binomial(13, 0.31) random variables, divide them by 13 and calculate the sample variance. This is likely to be quite close to the theoretical variance (ultimately this is guaranteed by the WLLN).

Question 3

- Most students answered this correctly.
- Some students seemed to understand what to do, but didn't report the actual estimates.

Question 4

- Most students answered this correctly.
- Some students incorrectly thought that if two random variables U and V are independent then $\text{var}(U - V) = \text{var}(U) - \text{var}(V)$.

Question 5

- Most students realized that they should apply Chebshev's inequality to the random variable W defined in Q4.
- Quite a few students did not realize that they should use for k the realization of w (or $|w|$) corresponding to the data. This is what "such a large difference ..." refers to.
- Some students used an estimate of p to compute the unknown quantity $p(1-p)$ instead of using the upper bound $p(1-p) < 1/4$. This was awarded marks if there was some explanation of why this is reasonable, but strictly speaking it is not a valid bound.

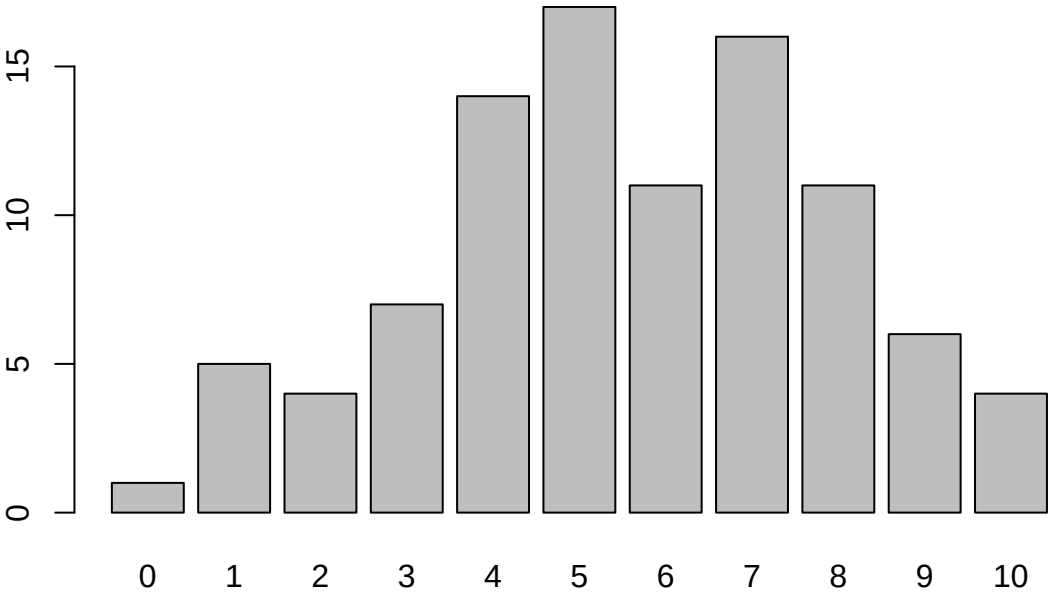
Question 6

- Most students did this well, except if they struggled with Q5 they usually had issues computing the probability bound.

Question 7

- Most students attempted this, but a large number did not quite get the log-likelihood function right and so the result of the call to `optim` was not the ML estimate for θ.
- Some students did not notice the relationship between (θ_1, θ_2) and (p_1, p_2), so they also did not notice the relationship between their maximum likelihood estimates.

Mark distribution (for the 96 people who submitted CP1)




```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   4.000   5.500   5.573   7.000  10.000
```