FURTHER MATHEMATICS

General Certificate of Education

Summer 2023

Advanced Subsidiary/Advanced

FURTHER STATISTICS B – A2 UNIT 5

Overview of the Unit

It was encouraging to see that the standard of mathematics, demonstrated by candidates, on this paper was the most pleasing out of all the GCE Mathematics and GCE Further Mathematics statistics papers. Candidates' performance on the 2023 paper was comparable to previous series and was a little more consistent than last summer (2022). There were a good number of candidates scoring well over 70 marks. There appear to be more better performing candidates in 2023, than in previous series. In general, candidates were adept at answering the questions on parametric tests. Calculating the confidence interval in question 3 was well done, but calculating the confidence level in question 7 was less well done. Question 5 on combinations of independent random variables was generally well done.

Comments on individual questions/sections

Q.1 This question should have been a straightforward start to the paper, and, for many candidates, it was. The error that many candidates made was not using the t-distribution in part (b). Despite having calculated an unbiased estimate for the variance in part (a), these candidates continued to use a normal distribution. The vast majority of candidates were able to give sensible answers in part (c).

Q.2 This was a lengthy question, which some candidates left until the end to answer. Part (a) was generally well done. In part (b), many candidates divided through by a without referencing $a > 0$, so did not earn all the marks available. The most common errors in part (c) were not using 10^2 as a denominator, using σ^2 for $\mathrm{Var}(\bar{X})$ and $\mathrm{Var}(\bar{Y})$, instead of $\frac{\sigma^2}{20}$ and $\frac{k\sigma^2}{25}$. As expected, $k = \frac{5}{6}$ often appeared in the absence of sufficient and convincing workings. Part (e) caused the most difficulty, with few candidates being able to find a correct expression for $\mathrm{Var}(T_3)$. Of those that did, many candidates did not differentiate their expression before setting it equal to 0. Only a handful of candidates justified that T_3 was a minimum.

Q.3 The confidence interval in part (a) was calculated correctly by the majority of candidates. There were a few incorrect z-values used and some candidates calculated the standard error incorrectly, but, on the whole, this question was answered well. Part (c) was done reasonably well, with a few omissions of $\sqrt{n}$ and incorrect values used for $\bar{x}$.

Q.4 Although this question was the most well answered on the paper, the quality and preciseness of the hypotheses was often lacking. The most succinct way to write the hypotheses was as $H_0: \eta_d = 0$, $H_1: \eta_d > 0$, where η_d is the median difference in numbers of followers before and after appearing on the show.

Q.5 Most candidates were able to answer part (a) correctly. Part (b) was also relatively successful with only a small number of candidates unable to standardise or select a correct z-value, with $z = -2.4572$ often being chosen. Unfortunately, there were several candidates who were unable to deal with $T = X_1 + X_2 + X_3 + Y_1 + Y_2 + Y_3 + Y_4$ correctly to find the variance, with $\mathrm{Var}(T) = 3^2 \times 100 + 4^2 \times 36 = 1476$ being the common error.

Q.6 Once again, this question was well answered, but, as in question 4, the quality and preciseness of the hypotheses was often lacking. The most succinct way to write the hypotheses was as $H_0: \eta_1 = \eta_2$, $H_1: \eta_1 > \eta_2$, where η_1 and η_2 are the median number of races entered by club members and non-members respectively. Another common error was to ignore the non-member who raced 0 times. This clearly arose from a misconception, where ignoring the difference of 0 in question 4 was the correct thing to do.

Q.7 This was the most poorly answered question on the paper. Many candidates, though not all, were able to correctly calculate the standard error for the difference of means. Only a few candidates were able to form an equation in k in order to find the z-value. Fewer still were able to progress further from $k = 2.333$ .... Those that did manage to find the confidence level, usually did so using their calculator. Most candidates were able to give a sensible assumption in part (b), although “riding both bikes at the same time” was not awarded a mark.