FURTHER MATHEMATICS

General Certificate of Education (New)

Summer 2019

Advanced Subsidiary/Advanced

FURTHER STATISTICS A - AS UNIT 2

General Comments

The new specification continues to give a wide spread in attainment over the course of the paper. Candidates were, once again, generally very good at performing calculations using formulae in the Formula Booklet; for example, calculating Spearman's rank correlation coefficient and the equation of a regression line. Compared to last year's examination, many more candidates were able to produce appropriate hypotheses when required to. As expected, the questions which required interpretation in context were the least well answered. The question on the, previously unfamiliar, exponential distribution proved challenging.

Comments on individual questions/sections

Q.1   As stated above, the calculation for Spearman's rank correlation coefficient was successfully carried out by all but the weakest of candidates. Although the scatter diagram in part (*a*) was conceptually understood by many candidates, many simply drew a scatter diagram showing positive correlation. In part (*c*), the question was designed to demonstrate that, despite the small positive correlation, the judges disagree strongly on most of the cheeses and that one should never simply use the product moment correlation coefficient to come to any conclusion without reference to a scatter diagram. Despite the suggestion to sketch a scatter diagram, many candidates did not and therefore failed to realise the nuance of the extent of agreement and disagreement by the judges.

Q.2   This question had a familiar feel to previous questions on this topic, but with the added challenge of identifying the binomial distributions. This challenge proved a step too far for several candidates who were unable to calculate p. This led to some difficulty in the calculation of $\mathrm{Var}(XY)$. Despite this challenge, this was the second most accessible question on the paper and many candidates scored full marks.

Q.3   Candidates, on the whole, answered part (*a*) very well indeed. Part (*b*) was also well answered, with only some candidates misunderstanding which probability should be multiplied by 3. Disappointingly, however, part (*c*) was not answered at all well. Candidates seemed almost completely unprepared to tackle a question on the exponential distribution, with the vast majority of candidates having no idea how to start this part of the question.

Q.4   Although questions on this topic, continuous random variables, have appeared in the legacy specification papers, the probability density functions tended not to be piecewise. Despite this, most candidates coped well with the piecewise function and answered this question very well. However, many lost the final accuracy mark due to premature approximation. Unfortunately, some candidates were unable to access the question because they seemed unsure about how to answer questions where the probability distribution function was piecewise.

Q.5   Part (*a*) was not as well answered as part (*b*). Of the candidates that understood the hypothesis that Chris was testing, all but a small fraction were able to look at the data and conclude that there is a tendency for NHL players to have a birthday earlier in the year. Most candidates who attempted a conclusion simply stated "not uniformly distributed". Although part (*b*) was well answered, some common errors included combining the groups into 3-month intervals despite the expected frequencies all being above 5, and not stating hypotheses which are part of a hypothesis test.

Q.6   Similar to last year's question on linear regression, this was by far the best answered question on the paper, with a vast majority of candidates getting full marks in part (*b*). In part (*a*), the most common error was stating the limitations of the regression line in general terms, rather than in the context of this question with reference to the scatter diagram in the question.

Q.7   The routine calculations in parts (*a*) and (*b*) were well answered. Once again, candidates were required to consider the data in the question and to interpret it in context in order to earn the relevant marks in part (*c*). A common incorrect answer was "Ankle because 57.9551 is the biggest number". Although many candidates were able to correctly identify Hand/Fingers, few candidates were able to put this into context, with reference to the low involvement of hands in football and the high involvement of hands and fingers in basketball.

Summary of key points

- It was encouraging to see good responses to familiar questions.

- Candidates should be encouraged to engage with the data in addition to following routine calculations.

- Forming the correct hypotheses is part of a hypothesis test and candidates should be familiar with the different hypotheses that are associated with the different tests.

- Candidates should be prepared to answer questions on the exponential distribution.