Pearson
Edexcel

Examiners' Report
Principal Examiner Feedback

Summer 2023

Pearson Edexcel GCE
In Statistics (9ST0)
Paper 03: Statistics in Practice

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.edexcel.com or www.btec.co.uk. Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

Summer 2023
Publications Code 9ST0_03_2306_ER

General introduction

In general candidates made good attempts at the calculations in this paper. A common area for improvement however is ensuring that explanations and assumptions for hypothesis tests and confidence intervals need to be given in the context of the question rather than simply being stated in general terms.

Question 1

Q01(a) to (d) of this question were generally well answered however many candidates were unable to accurately **show** how to calculate the mean of a discrete random variable.

Q01(e) Very few candidates gained full marks in this part of the question with the majority only gaining the mark for knowing that the central limit theorem was for use with large samples. Very few candidates could state that the distribution of sample means was normally distributed and fewer still stated the variance of the distribution of sample means.

Q01(f) Many candidates were able to make a good attempt at this question although there were some who did not recognise the significance of finding the probability of an average value.

Question 2

Q02(a) Many candidates were able to give good answers here recognising the difference between percentage point increase and percentage change of an amount. Candidates who did not obtain full marks in this question often only made a comment about one of the calculations and observed that the other one was incorrect.

Q02(b) to (d) were the best answered questions on the paper with many candidates able to identify features and good analyses from figures 3 and 4. Common errors observed were from candidates who criticised the representation of the data in figure 4 rather than commenting on the data displayed. In addition, candidates lost marks because they gave more than 3 reasons when only asked for 3, some of which were incorrect.

Question 3

Q03(a) Most candidates were able to identify this as analysis of variance however many didn't state "two factor" and therefore only obtained one mark.

Q03(b) Whilst many candidates knew the assumptions required to carry out an ANOVA test, the vast majority were not able to give the assumptions in the context of the question and therefore were not awarded full marks.

Q03(c) Many candidates were able to accurately carry out the hypothesis test however the most common errors were around stating the hypotheses and finding the correct critical value.

Q03(d) was generally well answered and credit was given to comments from candidates who incorrectly did not reject Ho in (c) for practical suggestions about which material should be chosen.

Q03(e) This part of the question was poorly answered by many candidates with many responses left blank. Many candidates correctly identified temperature as the blocking factor but then did not make use of the F-ratio already calculated in (c) to make a decision about whether it was effective or not.

Question 4

This proved to be the most challenging of the questions for many candidates and in many cases even where candidates were able to correctly identify the right models for each scenario, they were unable to articulate the reasons for their choice.

Question 5

Q05(a) to (d) were well answered by many candidates although some candidates struggled to give an interpretation for the gradient in the context of the question.

Q05(e) This part of the question proved very challenging for many candidates. Many candidates tried to circumvent the use of a database by describing how a spreadsheet could be used to solve the problem or did not read the question fully and described how a list of fastest times could be found from a list of all times.

Question 6

Q06(a) was well attempted by most candidates however common errors observed were from candidates using values from the normal distribution rather than the t-distribution as well as many candidates using one-tailed values rather than two tailed ones.

Q06(b) was poorly answered by most candidates, many of whom knew that the distributional assumption was that the data needed to be normally distributed but could not put this in the context of the question.

Q06(c) The most common error in this part of the question was that candidates concluded too definitely. Candidates should be reminded that conclusions from confidence intervals, much like for hypothesis tests, most not be too certain as confidence intervals only provide evidence about the parameters of populations.

Q06(d) was well answered by the majority of candidates.

Q06(e) Some candidates were able to give a well-reasoned argument to answer this question but many could not identify what they needed to do in order to show that $X \leq 10$ was the critical region. Some candidates attempted to use the inverse function on their calculators but very few of these were then able to write the notation or argument to show the correct critical region.

Q06(f) This question in general was well answered however many candidates misunderstood the question and tried to use a test statistic of 10 rather than 7. Many candidates used a p-value in this part of the question which was perfectly suitable if more time consuming than directly comparing the test statistic and critical value.

Q06(g) Very few candidates were able to articulate that a critical value allowed repeated tests to be carried out without need for a p-value calculation.

Q06(h) This part was generally well answered with the most common loss of marks being due to a lack of context within their explanations.

Q06(i) Although some candidates were able to show the critical region of $X \leq 10$ accurately in (e) many could not interpret their values as the probability of a type I error with the most common wrong answer being 0.05.

Q06(j) Pleasingly many candidates were able to give full solutions for this part of the question however some candidates confused the probability of a type II error with the power of the test.

Q06(k) Many candidates were not able to identify the requirements for a binomial distribution here with a common error being that candidate had not read the question fully and commented on the use of a normal distribution in (a).

Q06(l) This was generally well answered by candidates who were able to identify that since the conditions for scoring had changed, the conclusion was no longer valid.

Summary

Based on their performance on this paper, candidates should:

- Give all explanations in context.
- Be prepared to use either a critical region or a p-value method for different hypothesis tests.
- Give the number of reasons asked for in a question as they may not be awarded full marks due to additional incorrect answers.
- Be aware of the conditions needed to model values with different probability distributions
- Know when to use Student's t-distribution instead of the normal distribution in hypothesis tests and confidence intervals.

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom