Examiners' Report/
Principal Examiner Feedback

Summer 2010

GCE

Statistics S2 (6684)

Edexcel Limited. Registered in England and Wales No. 4496750
Registered Office: One90 High Holborn, London WC1V 7BH

If you have any subject specific questions about the content of this Examiners' Report that require the help of a subject specialist, you may find our Ask The Expert email service helpful.

Ask The Expert can be accessed online at the following link:

http://www.edexcel.com/Aboutus/contact-us/

Summer 2010

Publications Code UA024767

Statistics Unit S2
Specification 6684

Introduction

Candidates would appear to have had just enough time to complete this paper. There were few questions where no attempt had been made to produce an answer.

The level of work was very mixed. There were a number of candidates who had little idea about significance testing and quite a few who had problems with selecting the correct distributions.

The standard of presentation was very weak – it was often very difficult to follow the work. Presentation of working on a page was often disorganised, this was particularly the case on Q7 where candidates made many attempts to answer the question.

Report on individual questions

Question 1

This was poorly done with very few candidates scoring full marks. Those candidates who had learnt standard definitions fared better than those who used their own understanding of the terms because they were less likely to leave out vital elements of the definitions. Even those who answered parts (a) and (b) correctly were then unable to apply these definitions in context.

In part (a) a large majority of candidates omitted to mention "all", or its equivalent.

Part (b) was well answered because many candidates used a standard definition. The most common errors were using "population" instead of "sample and omitting "no unknown parameters".

In part (c) a substantial number of candidates were confused about "the population in this case". Many thought it to be the sample of 100 voters. Others were closer to the truth with "all the residents of the town", but did not earn the mark because they had failed to distinguish between registered voters and residents. The statistic was more easily identified.

Part (d) was poorly answered with many candidates having no idea what a sampling distribution was and those that did being unable to put it into context. The sampling distribution of a proportion is arguably one of the hardest to get a grip on and articulate convincingly.

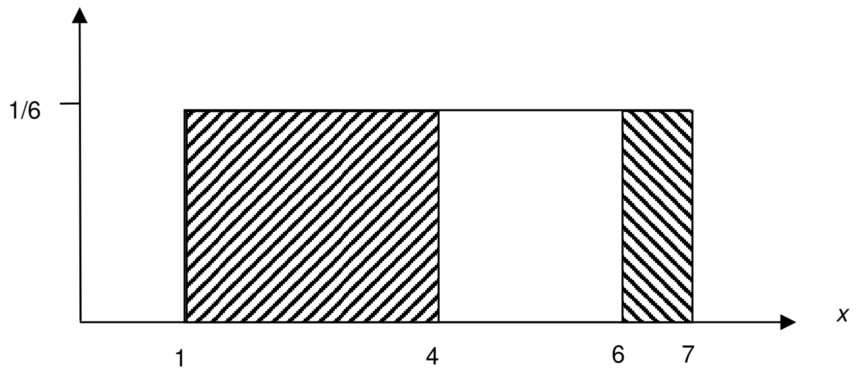
Question 2

This question was well answered by the majority of candidates with many scoring full marks. There were, of course, candidates who failed to score full marks. This was usually the result of inaccurate details, rather than lack of knowledge. In particular, manipulation of inequalities requires concentration and attention to detail. In part (a) the most common error seen was using $P(X = 3) = P(X \leq 4) - P(X \leq 3)$.

Parts (b) and (c) were usually correct. The most common error was to find $P(X \leq 3)$ rather than $P(X \leq 4)$ in part (b). A minority of candidates used the Normal as their approximation in part (d). The simple rule "*n* is large, *p* is small: use Poisson" clearly applies in this case.

Question 3

A minority of students got this completely correct and those that did often showed minimal working with only a diagram as evidence of their method. A major source of error was treating the distribution as a discrete uniform rather than a continuous uniform. Most students, however, simply worked out P(X > 6) and gave this as their answer, unaware that there was more to the question than this. The easiest and perhaps most successful solutions came from candidates who drew a diagram and realised they needed $P(X < 4) + P(X > 6)$



Question 4

Students seemed to fare better on the continuous distributions with the material spread over 2 questions rather than all together as previously.

In part (a) a large majority of candidates were able to set F(m) = 0.5 and formulate a correct equation. However then many candidates were unable to manipulate the initial equation into a more suitable form. However candidates who showed their working were able to earn credit for using a correct method on an incorrect equation. Candidates who solved the quadratic on their calculator showed no such method and therefore lost two marks instead of one for the incorrect answer. A variety of methods were then used successfully to solve the equation namely:

- 'the formula'
- 'completing the square'
- 'trial and improvement'

Part (b) was well done with only a few candidates neglecting to put "0 otherwise" in the full definition.

In part(c) the main problem was the result of confusion between discrete and continuous variables. It was not uncommon to see $P(X \geq 2) = 1 - F(1.1)$

Many candidates were able to gain the method mark in part (d). Those who didn't put $(0.6267)^4$ either repeated their answer to (c) or multiplied (c) by 4, seemingly unconcerned by a probability greater than 1.

Question 5

The majority of candidates were familiar with the technical terms in part (a), but failed to establish any context.

Part (b) was a useful source of marks for a large proportion of the candidates. The only problems were occasional errors in detail. In part (i) a few did not spot the change in time scale and used Po(4) rather than Po(8). Some were confused by the wording and calculated P(X = 8) rather than P(X = 0). The main source of error for (ii) was to find 1 – P(X≤ 4) instead of 1 – P(X ≤ 3).

In part (c) the Normal distribution was a well-rehearsed routine for many candidates with many candidates concluding the question with a clear statement in context.
The main errors were

- Some other letter (or none) in place of λ or μ
- Incorrect Normal distribution: e.g. N(60,60)
- Omission of (or an incorrect) continuity correction
- Using 48 instead of 60
- Calculation errors

A minority of candidates who used the wrong distribution (usually Poisson) were still able to earn the final two marks in the many cases when clear working was shown.

This question was generally well done with many candidates scoring full marks.

Question 6

Part (a) was well answered as no context was required.

In part (b) candidates identified the correct distribution and with much of the working being correct. However although the lower limit for the CR was identified the upper limit was often incorrect. It is disappointing to note that many candidates are still losing marks when they clearly understand the topic thoroughly and all their work is correct except for the notation in the final answer.

It cannot be overstressed that $P(X \le 6)$ is not acceptable notation for a critical region. Others gave the CR as $6 \le X \le 19$.

In part (c) the majority of candidates knew what to do and just lost the accuracy mark because of errors from part (b) carried forward.

Part (d) tested the understanding of what a critical region actually is, with candidates correctly noting that 8 was outside the critical region but then failing to make the correct deduction from it. Some were clearly conditioned to associate a claim with the alternative hypothesis rather than the null hypothesis. A substantial number of responses where candidates were confident with the language of double-negatives wrote “8 is not in the critical region so there is insufficient evidence to disprove the company’s claim”. Other candidates did not write this, but clearly understood when they said, more simply “the company is correct”.

Part (e) was generally well done with correct deductions being made and the contextual statement being made. A few worked out $P(X = 5)$ rather than $P(X \le 5)$.

Question 7

More candidates seemed to score full marks or nearly full marks than usual for this type of question. Some had problems with the concept of proof and some circular arguments were seen in part (b). There were also some problems in manipulating the algebraic fractions

In part (a) many candidates used a proof by contradiction approach rather than starting from f(y) > 0. Some wrongly thought that a probability density function cannot be 0 at any point and some thought that it can't be > 1. More attempted an explanation in words than a symbolic proof.

Part (b) saw many excellent solutions. There was a lot of detail involved. Yes, the equation to be solved was only linear, but the coefficients were potentially forbidding to those of us who only use a calculator as a last resort. There were many admirable responses, where candidates displayed persistence and excellent command of detail. The quantity of algebraic working seen varied substantially, from a few lines of genuine, succinct and accurate work, to a few pages of laboured inaccurate solutions.

Part (c) was surprisingly badly done. A few candidates were confused by the variable being y rather than the more usual x and so reflected their sketch in $y = x$. The minority who took their time over the sketch got it correct while those who just saw the squared term assumed a parabola intersecting the x-axis at 0 and 3. The mode was usually identified from their sketch although, as ever, there were those who gave the y value rather than the x value.

Grade Boundary Statistics

The table below give the lowest raw marks for the award of the stated uniform marks (UMS).

Module		Grade	A*	A	B	C	D	E
		Uniform marks	90	80	70	60	50	40
AS	6663 Core Mathematics C1			59	52	45	38	31
AS	6664 Core Mathematics C2			62	54	46	38	30
AS	6667 Further Pure Mathematics FP1			62	55	48	41	34
AS	6677 Mechanics M1			61	53	45	37	29
AS	6683 Statistics S1			55	48	41	35	29
AS	6689 Decision Maths D1			61	55	49	43	38
A2	6665 Core Mathematics C3		68	62	55	48	41	34
A2	6666 Core Mathematics C4		67	60	52	44	37	30
A2	6668 Further Pure Mathematics FP2		67	60	53	46	39	33
A2	6669 Further Pure Mathematics FP3		68	62	55	48	41	34
A2	6678 Mechanics M2		68	61	54	47	40	34
A2	6679 Mechanics M3		69	63	56	50	44	38
A2	6680 Mechanics M4		67	60	52	44	36	29
A2	6681 Mechanics M5		60	52	44	37	30	23
A2	6684 Statistics S2		68	62	54	46	38	31
A2	6691 Statistics S3		68	62	53	44	36	28
A2	6686 Statistics S4		68	62	54	46	38	30
A2	6690 Decision Maths D2		68	61	52	44	36	28

Grade A*

Grade A* is awarded at A level, but not AS to candidates cashing in from this Summer.

- For candidates cashing in for GCE Mathematics (9371), grade A* will be awarded to candidates who obtain an A grade overall (480 UMS or more) *and* 180 UMS or more on the total of their C3 (6665) and C4 (6666) units.
- For candidates cashing in for GCE Further Mathematics (9372), grade A* will be awarded to candidates who obtain an A grade overall (480 UMS or more) *and* 270 UMS or more on the total of their best three A2 units.
- For candidates cashing in for GCE Pure Mathematics (9373), grade A* will be awarded to candidates who obtain an A grade overall (480 UMS or more) *and* 270 UMS or more on the total of their A2 units.
- For candidates cashing in for GCE Further Mathematics (Additional) (9374), grade A* will be awarded to candidates who obtain an A grade overall (480 UMS or more) *and* 270 UMS or more on the total of their best three A2 units.

Further copies of this publication are available from
Edexcel Publications, Adamsway, Mansfield, Notts, NG18 4FN

Telephone 01623 467467
Fax 01623 450481

Email publications@linneydirect.com

Order Code UA024767 Summer 2010

For more information on Edexcel qualifications, please visit www.edexcel.com/quals

Edexcel Limited. Registered in England and Wales no.4496750
Registered Office: One90 High Holborn, London, WC1V 7BH