Examiners' commentary 2020

ST102 Elementary Statistical Theory (LT0 examination)

General remarks

Learning outcomes

By the end of this module you should:

- be able to summarise the ideas of randomness and variability, and the way in which these link to probability theory to allow the systematic and logical collection of statistical techniques of great practical importance in many applied areas
- be competent users of standard statistical operators and be familiar with a variety of well-known distributions and their moments
- understand the fundamentals of statistical inference and be able to apply these principles to choose an appropriate model and test in a number of different settings
- recognise that statistical techniques are based on assumptions and in the analysis of real problems the plausibility of these assumptions must be thoroughly checked.

Examination structure

After an initial ten minutes of reading time, you have one hour to complete this paper, which has two compulsory questions. The questions are given equal weight, and carry 50 marks each.

What are the Examiners looking for?

The Examiners are looking for you to demonstrate command of the course material. Although the final solution is 'desirable', the Examiners are more interested in how you approach each solution, as such, most marks are awarded for the 'method steps'. They want to be sure that you:

- have covered the syllabus
- know the various definitions and concepts covered throughout the year and can apply them as appropriate to examination questions
- understand and answer the questions set.

You are not expected to write long essays where explanations or descriptions are required, and note-form answers are acceptable. However, clear and accurate language, both mathematical and written, is expected and marked.

Key steps to improvement

The most important thing you can do is answer the question set! This may sound very simple, but these are some of the things that candidates did not do. Remember:

- **Always show your working.** The bulk of the marks are awarded for your approach, rather than the final answer.
- **Write legibly!**
- **Keep solutions to the same question in one place.** Avoid scattering your solutions randomly throughout the answer booklet – the Examiners will not appreciate having to spend a lot of time searching for different elements of your solutions.
- **Where appropriate, underline your final answer.**
- **Do not waste time calculating things which are not required by the Examiners!**

Using the commentary

We hope that you find the commentary useful. For each question and subquestion, it gives:

- the answers, or keys to the answers, which the Examiners were looking for
- common mistakes, as identified by the Examiners.

Dr James Abdey, ST102 Lecturer, February 2020

Examiners' commentary 2020

ST102 Elementary Statistical Theory (LT0 examination)

Specific comments on questions

Question 1

Suppose a die is thrown and the probability of getting i, for $i = 1, 2, \ldots, 6$, is given by:

$$P(D = i) = ai$$

where a is a constant, and D is the random variable denoting the outcome of the throw.

(a) **Show that $a = 1/21$.**

(5 marks)

(b) **If $D \geq 3$, then the same die is thrown again once. Otherwise, an independent fair die is thrown twice. The final outcome is recorded in the random variable S, which is the sum of all values of dice thrown.**

i. **Show that:**

$$P(S = 4) = \frac{16}{1{,}323}.$$

(20 marks)

ii. **If $S = 4$, what is the probability that $D \leq 2$?**

(10 marks)

iii. **Find $\mathrm{E}(D \mid S = 4)$.**

(15 marks)

Feedback on this question

A question on probability. Part (a) makes use of the second necessary condition of a probability function, i.e. that the probabilities sum to 1. Part (b) i. requires application of the total probability formula and makes use of independence. Part (b) ii. is a conditional probability, using the provided result of $P(S = 4)$ in the denominator. Finally, part (b) iii. is a conditional expectation – several candidates used the original probabilities in the calculation instead of the conditional probabilities.

Full solutions are as follows.

(a) We have:

$$1 = a(1 + 2 + 3 + 4 + 5 + 6) = 21a$$

and so $a = 1/21$.

(b) i. Let D' be the outcome of the throw of the biased die again, and let F_1 and F_2 be the outcomes of the two throws of the fair die. Therefore, by independence:

$$P(S = 4) = P(D = 3)\, P(D' = 1) + \sum_{i=1}^{2} P(D = i)\, P(F_1 + F_2 = 4 - i).$$

However:

$$P(F_1 + F_2 = 3) = 2 \times P(F_1 = 1)\, P(F_2 = 2) = \frac{2}{36}$$

and:

$$P(F_1 + F_2 = 2) = P(F_1 = 1)\, P(F_2 = 1) = \frac{1}{36}.$$

Hence:

$$P(S = 4) = \frac{3 \times 1}{(21)^2} + \frac{1 \times 2 + 2 \times 1}{21 \times 36} = \frac{16}{1{,}323}.$$

ii. We have:

$$P(D \leq 2 \,|\, S = 4) = \frac{\sum_{i=1}^{2} P(D = i)\, P(F_1 + F_2 = 4 - i)}{P(S = 4)} = \frac{4/(21 \times 36)}{16/1{,}323} = \frac{7}{16}.$$

iii. We have:

$$\begin{aligned} \mathrm{E}(D \,|\, S = 4) &= \frac{\sum_{i=1}^{2} i\, P(D = i)\, P(F_1 + F_2 = 4 - i) + 3 \times P(D = 3)\, P(D' = 1)}{P(S = 4)} \\ &= \frac{6/(21 \times 36) + 9/441}{16/1{,}323} \\ &= \frac{75}{32}. \end{aligned}$$

Question 2

(a) Consider two random variables X and Y, where X can take the values 0, 1 and 2 and Y can take the values 0 and 1. The joint probabilities for each pair are given by the following table.

	$X = 0$	$X = 1$	$X = 2$
$Y = 0$	0.1	0.1	0.4
$Y = 1$	0.2	0.1	0.1

Calculate:

i. $\mathrm{E}(Y)$

(6 marks)

ii. $P(Y < X \,|\, X < 2)$

(7 marks)

iii. $P(Y < X \,|\, X + Y = 2)$.

(7 marks)

(b) The random variable X has the cumulative distribution function given by:

$$F(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ 0.4x^{1.5} + 0.6x & \text{for } 0 < x \leq 1 \\ 1 & \text{for } x > 1. \end{cases}$$

i. Determine:

$$\mathrm{Cov}\left(\frac{1}{\sqrt{X}}, X\right).$$

(20 marks)

ii. Calculate:

$$P\left(X < \frac{9}{16} \,\middle|\, X > \frac{1}{4}\right).$$

(10 marks)

Feedback on this question

Part (a) is a simple discrete bivariate distribution. Part i. requires the probabilities of the marginal distribution of Y. Parts ii. and iii. are conditional probabilities although some candidates confused these with joint probabilities. For (b) part i. as the cumulative distribution function, $F(x)$, is provided, a first step is to differentiate to obtain the probability density function, $f(x)$. Once obtained, calculations of expected values of functions of X are then routine. The definition of covariance, i.e. the expectation of the product minus the product of the expectations, then allows the final answer to be determined. In part ii. the necessary probabilities could be calculated using either $F(x)$ or $f(x)$. Several candidates made arithmetic mistakes.

Full solutions are as follows.

(a) i. We have:

$$P(Y=0) = 0.1 + 0.1 + 0.4 = 0.6$$

and $P(Y=1) = 1 - 0.6 = 0.4$. Hence:

$$\mathrm{E}(Y) = 0 \times 0.6 + 1 \times 0.4 = 0.4.$$

ii. We have:

$$P(Y < X \mid X < 2) = \frac{P(Y < X < 2)}{P(X < 2)} = \frac{P(X=1, Y=0)}{1 - P(X=2)} = \frac{0.1}{1 - 0.4 - 0.1} = \frac{1}{5}.$$

iii. We have:

$$P(Y < X \mid X + Y = 2) = \frac{P(X=2, Y=0)}{P(X=2, Y=0) + P(X=1, Y=1)} = \frac{0.4}{0.4 + 0.1} = \frac{4}{5}.$$

(b) i. The probability density function is calculated by differentiating:

$$f(x) = \begin{cases} 0.6x^{0.5} + 0.6 & \text{for } 0 < x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

We then have:

$$\mathrm{E}\left(\frac{1}{\sqrt{X}}\right) = \int_0^1 x^{-0.5}(0.6x^{0.5} + 0.6)\,\mathrm{d}x = \int_0^1 (0.6 + 0.6x^{-0.5})\,\mathrm{d}x = 0.6 + \frac{0.6}{0.5} = 1.8.$$

Also:

$$\mathrm{E}(X) = \int_0^1 x\,(0.6x^{0.5} + 0.6)\,\mathrm{d}x = \int_0^1 (0.6x^{1.5} + 0.6x)\,\mathrm{d}x = \frac{0.6}{2.5} + \frac{0.6}{2} = 0.54$$

and:

$$\mathrm{E}\left(\frac{X}{\sqrt{X}}\right) = \int_0^1 x^{0.5}(0.6x^{0.5} + 0.6)\,\mathrm{d}x = \int_0^1 (0.6x + 0.6x^{0.5})\,\mathrm{d}x = \frac{0.6}{2} + \frac{0.6}{1.5} = 0.7.$$

Therefore:

$$\mathrm{Cov}\left(\frac{1}{\sqrt{X}}, X\right) = 0.7 - 1.8 \times 0.54 = -0.272.$$

ii. We have:

$$P\left(\frac{1}{4} < X < \frac{9}{16}\right) = 0.4 \times \left(\frac{9}{16}\right)^{1.5} + 0.6 \times \frac{9}{16} - 0.4 \times \left(\frac{1}{4}\right)^{1.5} - 0.6 \times \frac{1}{4} = 0.30625$$

and:

$$P\left(X > \frac{1}{4}\right) = 1 - 0.4 \times \left(\frac{1}{4}\right)^{1.5} - 0.6 \times \frac{1}{4} = 0.8.$$

Hence:

$$P\left(X < \frac{9}{16} \,\middle|\, X > \frac{1}{4}\right) = \frac{P(1/4 < X < 9/16)}{P(X > 1/4)} = \frac{0.30625}{0.8} = 0.3828.$$