

**Question 10****(13 marks)**

A book called 'Why I Love Mathematics' is having its first print run. This is scheduled to last for one week, using four printing presses. The publisher claims that, historically, 10 books have printing errors for every two hundred that are printed.

A sample of 200 books is to be chosen to determine how many contained errors. The proposed sampling procedure is to select the first 200 books printed over a 6-hour window using the newest printing press.

- (a) (i) Identify and explain **one** possible source of bias in the proposed sampling procedure. (2 marks)

- (ii) Identify **two** changes to the sampling procedure that would reduce bias. (2 marks)

Assume that the sample was gathered using an improved procedure, and that the publisher's claim is correct.

- (b) Use the approximate normality of the distribution of sample proportions to determine the probability that the sample proportion of books with errors is less than 0.04. (2 marks)

In a different random sample of 600 books, it is found that the proportion of books containing an error is 0.1, with a margin of error of 0.024.

- (c) Determine a 95% confidence interval for the proportion of books that will have printing errors. (1 mark)
- (d) On the basis of the confidence interval determined in part (c), is the proportion of books with printing errors different from what was claimed by the publisher? (2 marks)
- (e) Suggest **two** changes that could be made in order to decrease the margin of error of the confidence interval. (2 marks)
- (f) Determine the minimum sample size that would be necessary to guarantee that the margin of error of the resulting 95% confidence interval was at most 0.02. (2 marks)