ST2004 & ST2352 Assessment 1: Dice

Submission date **strictly before** 4pm Friday 20th November. Late submissions are not accepted without adequate supporting approval from your tutor.

Submit your work to the School of Computer Science and Statistics Reception Office on the first floor of the O'Reilly Building. There should be one submission per group and it should clearly state the names of the individuals in the group. At the same time, submit the peer-assessment form. Your pages should be stapled together or otherwise bound.

This is both a learning and an assessment exercise. The exam may draw on this project. It will be scaled to carry 10% of the course marks. Good students will take the opportunity to go well beyond the lectures. 75% of the marks will be awarded 'for doing a competent job'; an extra 25% will be awarded for flair, imagination, thoroughness – *i.e.* the 'extra mile'.

Q1 (10 marks) requires you to model the system (here a game) in which X_1, X_2, X_3 and X_4 denote the number of spots showing on the first, second, third, and fourth rolls of a regular die, and in which S_3, S_4 and M_3, M_4 denote the sums $S_3 = (X_1 + X_2 + X_3)$ and the maximum $M_3 = \max(X_1, X_2, X_3)$ and S_4, M_4 similarly for four dice. You are required:

(i) to determine by simulation the probability distributions for the univariate random variables S_3, S_4 and M_3, M_4, together with their long run averages and variances, the joint probability distribution of (S_3, M_3) and (S_4, M_4) with additionally their covariance, and (for three dice only) the conditional distributions of S_3 given that $M_3 = 4$ and of M_3 given that $S_3 = 10$ (4 marks);

(ii) to explain the use of probability arguments to determine these probability distributions with the Expected Values (Expected long run averages) and Variances (Expected long run variances) and Covariance (Expected long run covariance), with some discussion of $E[S_k]$ and $Var[S_k]$ (4 marks); and

(iii) to demonstrate the law of large numbers by comparing (i) and (ii) (2 marks).

To do this you can use any template resource that you think might help.

The report on Q1 should comprise no more than 6 or 7 pages, including tables and diagrams; good use of diagrams is encouraged and will be rewarded. The use of a large number of significant figures in reported computations should be avoided. The report should be prepared professionally using any appropriate software (*e.g.* Word or Latex), with the tables and diagrams being imported from EXCEL or equivalent. Word's Insert Symbol and Insert Equations facilities are adequate for the necessary Math that you need.

Q2 (5 marks) Investigate 'something' in some depth using random numbers (maximum 3 pages).

Theory

Chapters 5 and 9 and sections 11.1 and 13.1 in the course text (Tijms, Understanding Probability) provide theoretical material to complement the Lecture notes. The laws of large numbers are discussed in Ch 2; Central Confidence Intervals for simulations are discussed in Ch 5. The Central Limit Theorem in Ch 5 and also Ch 14 are relevant to a wider theoretical discussion of the distribution of S_k which some groups may wish to pursue in their extra mile.

Team composition

Groups of up to size 4 under self-allocation. If there is any problem in this regard you should contact me as soon as possible, ideally at the next laboratory class.

Team working

Your report is a group project. I propose this mainly because group learning can be valuable, but also because team-working is a useful skill. The team will receive a mark and by default all members will receive the same mark. The project work does naturally break into sections which can be conducted in parallel before being brought together.

However, as problems do arise from time to time, I must insist that <u>each team member individually</u> completes the PEER ASSESSMENT form. This will be used as necessary to re-allocate the team mark to the individual team members. If I do not receive such a completed form from you I will take it that you made no contribution and expect no mark. While learning from each other and from other groups is encouraged, plagiarism is remarkably easy to spot, especially with online resources. The same work submitted by two teams will receive a single mark; this will be divided between the teams involved, following discussion.