

## Exercise Sheet Lecture 4 — The Binomial distribution

1. State the condition under which the Binomial distribution arises?
2. If  $X \sim \text{Bin}(12, 0.7)$  calculate
  - (a)  $P(X = 9)$
  - (b)  $P(X > 10)$
  - (c)  $P(X \leq 11)$
3. A lecturer uses a laptop to give a series of 8 lectures. There is a 5% chance that the laptop causes problems in any given lecture. What is the probability of observing 2 out of 8 lectures in which the laptop exhibits ‘technical difficulties’.
4. Two groups of twelve children are taught two different methods of arithmetic. (Assume that a child in group one is matched in terms of their arithmetic ability with a child in group 2 before the start of the study). What is the probability that at least 9 children from one of the groups will obtain higher scores than the other group? What other assumptions have you made?
5. Two teams, A and B, play a series of games. If team A has probability 0.4 of winning each game, is it to its advantage to play the best three out of five games or the best four out of seven? Assume outcomes of successive games are independent.
6. One of the most important developments in epidemiology was establishing that smoking causes cancer, and generally higher mortality. The argument was made, that this was a result of “confounding”: There is some genetic factor that predisposes some people to smoke, and this also predisposes them to various diseases. In order to test this argument, a study was conducted, in which monozygotic twins were found, who were “discordant for smoking” — that is, one smokes and one doesn’t. These twins were followed for 12 years, during which time some died. If smoking really has no direct effect of causing disease and death, what would that say about the probability of one or the other twin dying first?

In the course of 12 years of this study, there were 22 pairs of which at least one of the twins died. The data are given in the table below (copied from D. Freedman, “From Association to Causation: Some Remarks on the History of Statistics,” *Statistical Science* 1999, **14**:3, 243–58.) These show, for each cause of death, how many of each category of twins died first. Thus, 17 of the smokers died before their nonsmoking twin, and 5 of the nonsmokers died before their smoking twin.

Cause	Smoker	Nonsmoker
All	17	5
Coronary Heart Disease	9	0
Lung Cancer	2	0

Use these data to test the hypotheses: a) Smoking does not cause early death from all causes; b) Smoking does not cause early death from Coronary Heart Disease; c) Smoking does not cause early death from lung cancer. What conclusions would you draw?

**Exam Paper Questions** The table below lists past exam paper *question numbers* that involve the Binomial distribution (useful for revision). NB. bracketed question numbers indicate that only part of the question involves the Binomial distribution.

	Human Sciences	Psychology		
Year	TT	MT	HT	TT
2001	3		3	3
2000	3	(3)	4	3
1999	(4)(6)	3	(3)	(4)(6)
1998			3	
1997	(9)		(4)	(9)