Exercise Sheet 1

1. In a random experiment, two dice are thrown simultaneously. We consider the sum S of both dice to be a random variable.

a) Which event space appears in this random experiment?

b) Which values can S attain? Compute the corresponding probabilities.

c) Sketch the probability distribution and the cumulative distribution function of the random variable S.

2. A test tube manufacturer wishes to ensure that less than 10% of the items in a large batch contain lesser-quality glass (quality level A). The company thus takes a random sample of fifty test tubes from the batch for quality control purposes. Three of these fifty are of lesser quality. The manufacturer now faces the problem of deciding – based only on this sample – whether less than 10% of the items in the batch are of this lower quality, or whether such a result came about „by chance", despite the real proportion of lesser-quality test tubes being greater than 10%.

a) Which model or distribution describes the number of lower-quality test tubes in the batch, assuming that the items are independent of each other?

b) How high is the probability of obtaining exactly three "bad" items in the sample taken if their proportion in the whole batch is in fact 10%?

c) How high is the probability of obtaining at most three "bad" items in the sample if their proportion in the whole batch is in fact 10%?

d) Formulate the manufacturer's „problem" in a concise way! (with few words)

3. In a study, water samples are investigated for signs of contamination. As only 2% of all samples are contaminated, the suggestion is made to join half of the contents of 10 samples to form a pooled sample and to initially only investigate this pooled sample. If this is uncontaminated, the investigation of all 10 individual samples is already complete; otherwise, the remaining halves of the 10 samples are tested individually.

a) What is the probability of finding no contamination in the pooled sample (assuming the independence of the individual samples)?

b) Let the random variable Y denote the number of analyses required here. What is the range of Y, and which what probability does each of numbers in its range appear?

c) What is the "average" number of analyses required for the entire investigation (i.e. what is $\mathbf{E}[Y]$)? How many analyses are saved "on average" by forming pooled samples?