Department of Computer Science and Engineering
M.Tech. I Computer Science and Engineering
Mathematical Foundations of Computer Science CSE 601
Tutorial 5

1. Suppose we have a stream of tuples with the schema: Grades (university, courseID, studentID, grade) Assume universities are unique, but a courseID is unique only within a university (i.e., different universities may have different courses with the same ID, e.g., "CS101") and likewise, studentID's are unique only within a university (different universities may assign the same ID to different students).
   Suppose we want to answer certain queries approximately from a 1/20th sample of the data. For each of the queries below, indicate how you would construct the sample. That is, tell what the key attributes should be.
   (a) For each university, estimate the average number of students in a course.
   (b) Estimate the fraction of students who have a GPA of 3.5 or more.
   (c) Estimate the fraction of courses where at least half the students got "A."
2. A computer store has purchased three computers of a certain type at $500 apiece. The computers then are sold for $1000 each. The manufacturer has agreed to repurchase any computers that remain unsold after one month for $200 each. Let X be the random variable that denotes the number of computers sold. Suppose the probabilities for selling i computers for i = 0, 1, 2, 3 are p(O) = 0.1; p(l) = 0.2; p(2) = 0.3; and p(3) = 0.4. Let h(x) denote the profit from selling X units. Find the expected value of h as well as the standard deviation and the variance.
3. Student workers find that 75% of all help desk inquiries involve programs with syntax errors. Let X be the random variable that counts the number of programs with syntax errors in 10 randomly chosen consultations. Find the expected value, the variance, and the standard deviation of this random variable.
4. A chain of home entertainment stores sells three different brands of DVD players. Fifty percent of its sales are brand 1, 30% are brand 2, and 20% are brand 3. Each manufacturer offers a one-year warranty on parts and labor. It is known that 25% of brand l's DVD players require warranty repair work, whereas the corresponding percentages for brands 2 and 3 are 20% and 10%, respectively.
   (a) What is the probability that a randomly selected purchaser has bought a brand 1 DVD player that will need repair under warranty?
   (b) What is the probability that a randomly selected purchaser has a DVD player that will need repair while under warranty?
   (c) If a customer returns to the store with a DVD player that needs warranty repair work, what is the probability that it is a brand 1 DVD player? A brand 2 DVD player? A brand 3 DVD player?
5. Four individuals have responded to a request by a blood bank for donations. None of the four has donated before, so each person's blood type is unknown. Suppose that only type A positive is desired and that only one of the four actually has this type. If the potential donors are selected in random order for blood typing, what is the probability that at least three individuals must be typed to find a donor of type A positive?
6. Three automatic machines produce similar automobile parts. Machine A produces 40% of the total, machine B 25%, and machine C 35%. On average, 10% of the parts turned out by machine A do not conform to specifications, and for machines B and C, the corresponding percentages are 5% and 1%, respectively. If one part is selected at random from the combined output and does not conform to the specifications, what is the probability that it was produced by machine A?
7. Consider the following darts game: The target consists of a bull's-eye, which is a circle of radius 1, surrounded by a middle ring of outer radius 3 and inner radius 1; this region is in turn surrounded by another ring of outer radius 5 and inner radius 3. If you hit the bull's-eye, you win $10 plus the opportunity to throw again. If you hit the middle ring, you lose $2, and if you hit the outer ring, you lose $5. You must stop throwing as soon as you make a losing toss or hit three bull's-eyes. Suppose the probability that you hit a region is proportional to its area.
   (a) Set up an underlying sample space Ω and its probability density p.
   (b) Define a random variable X on Ω.
   (c) Define a sample space Ω_x and a probability distribution p_x on Ω_x.