ACM/EE 116. Midterm.

- Instructor: Houman Owhadi.
- TA: Roger Donaldson and Yaniv Plan.
- Time limit: 6 hours. The honor code is in effect.
- Search engines are not allowed...Matlab and mathematical are allowed. The following sites and resources are allowed
  - *http://mathworld.wolfram.com/Probability.html*
  - *http://en.wikibooks.org/wiki/Probability*
  - *http://en.wikipedia.org/wiki/Probability_theory*

  The time limit is no more than 6 hours in front of your copy with the pen in your hand. However you are allowed to fragment that time in whichever way you want. That is to say that you can spend 2 hours in front of your copy then stop (to sleep, eat, think about something else, your problem or even read the textbooks), then spend another 3 hours in front of your copy then stop (say for 2 days) then spend the last hour in front of your copy. The exam will be accessible tomorrow for download from the ACM/EE 116 website after the class.
- Due next Tuesday (Oct 31, 2006) in class. You may ask me whatever question you may have by email (don't hesitate if you think that a problem is not clear or ambiguous) and I will answer to everyone at the same time with a broadcasted email (if the question is reasonable).

1

A tribunal is inquiring on the paternity relationship between two people. For this it gives the analysis of blood phenotypes to two independent labs which give correct results with a probability α (lab A) and β (lab B). The probability that two people taken at random have the same phenotype is τ. Write I the event: the phenotypes are identical, A+ (B+) the event: Lab A (Lab B) finds that the phenotypes are identical. Making errors, is independent of the actual state of the patient.

a Compute $\mathbb{P}(A + \cap B + | I)$.

b Compute the probability that two people have the same phenotype knowing that the results of the two labs are positive.

c Compute the probabilities in parts a) and b) for the specific case $\alpha = \beta = 0.9, \tau = 10^{-3}$.

2

Two weather stations are giving data on a climatic system which can be in two states S_1 and S_2, shifting at random from one to the other. Long observations have shown that during 30% of the time the system is in the state S_1 and 70% of the time the system in the state S_2. Station 1 gives erroneous data in 2% of cases, and station 2 in 2% of cases. Making errors is independent of the actual state of the climatic system. Each station makes its errors independently of the other.

At a given time, station 1 is communicating that the system is in the state S_1 whereas station 2 is saying that the system is in the state S_2. Which communication should be assumed to be correct?

3

It is commonly presumed that an unborn child has a 50% probability of being a female. But is it really the case? Let's look at birth statistics for the Netherlands for the years 1989, 1990 and 1991. According to the Central Bureau of Statistics, there were, in total 585609 children born during the span of those years, of which 286114 were girls. What is the estimate for the probability that a newborn child will be a girl and what is the corresponding 95% confidence interval. That is to say: writing p the probability that a newborn child will be a girl we are looking for a random variable $\hat{p}$ given as a function of our statistics and a parameter $\alpha > 0$ such that

$$\mathbb{P}\big[p \in (\hat{p} - \alpha, \hat{p} + \alpha)\big] \approx 0.95$$

where the randomness in $\hat{p}$ corresponds to the randomness in our statistics. You may use the fact that

$$\int_{-1.96}^{1.96} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \approx 0.95$$

4

Suppose for instance that you are offered a sequence of bets, each bet being a losing proposition with probability $1 - p$ and paying out f times ($f > 1$) your stake with probability p.

a Compute the expected net payoff of each bet.

b Suppose that the expected net payoff of each bet is strictly positive (you have an edge). How to gamble if you must? The idea is to bet a fixed proportion of your present bankroll. When your bankroll decreases you bet less, as it increases you bet more. Assuming that your starting bankroll is V_0, define the random variable V_n as the size of your bankroll after n bets when you bet a fixed fraction α $(0 < \alpha < 1)$ of your current bankroll each time. Here it is supposed that winnings are reinvested and that your bankroll is infinitely divisible. Find the optimal value for α. Hint: Observe that $V_n = (1 - \alpha + \alpha R_1) \times \cdots \times (1 - \alpha + \alpha R_n) V_0$ where R_k is equal to the payoff factor f if the $k - th$ bet is won and is otherwise equal to 0.

5

The waiting times at checkout desks of a supermarket are positive random variables $X_0, \ldots, X_n$ independent identically distributed with continuous density on $[0, \infty)$. Your waiting time is given by X_0, and the waiting time of the person that arrived at the same time as you at the checkout desk i is X_i. Compute the law of $N = \inf\{i \geq 1\,;\, X_i > X_0\}$. Give its expectation. Observe that N is a measure of your frustration, i.e. a large N corresponds to a large number of people right next to you having checked before you.