Swarthmore College Honors Exam - Spring 2009
Probability and Stochastic processes

Instructions:

- There is more here than you will have time to complete. First, look over the problems. Then, carefully select problems that you choose to answer. In addition to picking problems that you can answer well, please try to spread out your selections to cover different topics from both courses.

- There are three parts:
  (I) Exercises
  (II) Proofs
  (III) Overview of the subjects.

  This is a 3 hour exam. Please try to spend about 60 to 90 minutes on part (I), 60 to 90 minutes on part (II) and 30 minutes on part (III).

  After the appropriate time on a section, stop and move on even though you could continue.

- Write your answers clearly and concisely. Good style in writing mathematics is important. If there is part of an answer that you can't do well it is better admit the gap then to make a mistake or fudge your answer. Then move on to the parts you can do well.

- This is closed book, notes etc.

- The problems vary in length, so it is hard to predict the portion that you might answer. Do as much as you can, taking care not to ponder any one problem for too long.

- Try to pick problems that will allow you show off your mathematical skills.

Garth Isaak - Lehigh University

(I) Exercises:

The questions in this section are designed to check if you know basic results and can do elementary computations. None of the answers should be very long. Some problems require proofs, outlines of proofs or explanations. These should be fairly short. You do not need to put in every detail. Instead do enough to show the main ideas. For more computational problems at least show some of your work and give short explanations of what you are doing when appropriate.

1: Three people A, B and C play a game in which they throw coins, one after the other. A starts, then B, then C, then A again etcetera. The person who throws heads first wins the games. Construct an appropriate sample space for this game, and find the probability that A wins the game.

2: **(a)** Suppose that A and B are independent events. Show that A^c and B are independent.
(b) Prove Bonferroni's inequality $P(A \cap B) \geq P(A) + P(B) - 1$.
(c) Prove the general version of Bonferroni's inequality $P(\cap_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} P(A_i) - (n-1)$.

3: An urn contains n white and m black balls, where n and m are positive numbers. Determine the following probabilities.
(a) If k balls are randomly withdrawn (without replacement), what is the probability that exactly r of them are white?
(b) If two balls are randomly withdrawn (without replacement), what is the probability that they are the same color?
(c) If k balls are randomly drawn (with replacement, so each time a ball is selected there are n white balls and m black balls) what is the probability that exactly r of them are white?
(d) If two balls are randomly drawn (with replacement, so each time a ball is selected there are n white balls and m black balls) what is the probability that they are the same color?

4: Recall that $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ for $k = 0, 1, \ldots, n$ is the binomial distribution with parameters n and p. Prove that $E[X] = np$ in any way you like.

5: Suppose we have 10 coins such that if the i^{th} coin is flipped it will show heads with probability $i/10$ for $i = 1, 2, \ldots, 10$. A coin is randomly selected and flipped. If it shows heads what is the conditional probability that it is the 7^{th} coin?

6: Let (X, Y) have joint density $f(x, y) = e^{-y}$ for $0 < x < y$, and $f(x, y) = 0$ otherwise.
(a) Show that the marginal density of Y is $f_Y(y) = ye^{-y}$ for $y > 0$ and 0 otherwise.
(b) Show that $f_{X|Y}(x, y) = 1/y$ for $0 < x < y$
(c) Show that $E(X|Y = y) = \frac{y}{2}$ for $y > 0$.
(d) Use part (c) to compute $E(X)$.

7: Let X and Y be independent Poisson random variables with parameters λ_1 and λ_2 respectively. Use moment generating functions to show that the distribution of the sum $X + Y$ is also Poisson and determine the parameter for this sum. The moment generating function for a Poisson random variable with parameter λ is

$$\phi(t) = e^{\lambda(e^t - 1)} = \exp \lambda(e^t - 1).$$

8: Suppose that by any time t the number of people that have arrived at a train depot is a Poisson random variable with mean λt. The first train arrives at a time that is uniformly distributed over $(0, T)$ (for some T) independent of when passengers arrive. What is the mean number of passengers that enter the train?

9: The number of people that enter a store in a given hour is a Poisson random variable with parameter $\lambda = 10$. Compute the conditional probability that at most 3 men entered the store, given that 10 women entered in that hour. What assumptions have you made?

10: (a) Suppose that the weather today depends on the weather conditions of the last two days. The weather for a given day can be either wet or dry. If it rained the past two days, then it will rain today with probability 0.8. If it did not rain for the past two days then it will rain today with probability 0.3. In any other case, the weather today will, with probability 0.6 be the same as yesterday. Construct an appropriate Markov chain for this model and determine its transition matrix P.
(b) If the weather today depends on the weather of the previous a days and if there are b possible weather conditions each day, how many states are needed to analyze this with a Markov chain?

11: A small barbershop, operated by a single barber, has room for at most two customers. Customers leave if the shop is full when they arrive. Potential customers arrive at a Poisson rate of three per hour, and the successive service times are independent exponential random variables with mean 1/4 hour. What is
(a) the average number of customers in the shop?
(b) the proportion of potential customers that enter the shop?

12: Suppose that a one-celled organism can be in one of two states - either A or B. An individual in state A will change to state B at an exponential rate α; an individual in state B divides into two new individuals of type A at an exponential rate β. Define an appropriate continuous time Markov chain for a population of such organisms and determine all the appropriate parameters for this model.

13: Recall that the long term probabilities of j people being in the system for an $M/M/1$ queue are $\pi(n) = (\lambda/\mu)^n(1 - \lambda/\mu)$. Consider an $M/M/1$ queue in which arrivals finding N people in the system do not enter. What are the long term probabilities for this system?

(II) Proofs

For questions in this section show that you can write clear and concise proofs. A few of the questions are more conceptual about the ideas related to proofs. If you choose to omit some details state that you are doing this. For problems with multiple parts you may use results of previous parts even if you did not answer them.

14: A researcher wants to determine the relative efficacies of two drugs. The results (differentiated between men and women) were as follows:

women	drug I	drug II
sucess	200	10
failure	1800	190

men	drug I	drug II
sucess	19	1000
failure	1	1000

We are now faced with the question which drug is better. Here are two possible answers. (1) Drug I was given to 2020 people, of whom 219 were cured. Drug II was given to 2200 people, of whom 1010 were cured. Therefore drug II is much better. (2) Amongst women, the success rate of drug I is 1/10 and for drug II the success rate is 1/20. Amongst men, these rates are 19/20 and 1/2 respectively. In both cases drug I is better. Which of the two answers do you believe? Can you explain the paradox?

15: Answer (c) and only *one* of (a) or (b):

(a) Show that if X is normally distributed with parameters μ and σ^2 then $Y = aX + b$ is normally distributed with parameters $a\mu + b$ and $a^2\sigma^2$. Do this by showing that the cumulative distribution function of Y can be expressed in terms of that of X as $F_Y(x) = F_X\left(\frac{x-b}{a}\right)$. (Do not use integrals, just use the definition of the distribution function.) Then differentiate this equation and show clearly how the result is the density of a normal distribution with parameters $a\mu + b$ and $a^2\sigma^2$.

(b) Show the same result as (a) by first showing that the moment generating functions satisfy $M_Y(t) = e^{tb}M_X(at)$ and using the fact that $M_X(t) = e^{\mu t+\sigma^2 t^2/2}$.

(c) If X has normal distribution with mean μ and variance σ^2 then it has moment generating function $M_X(t) = e^{\mu t+\sigma^2 t^2/2}$. Use this to show that if X_1 is normal with mean μ_1 and variance σ_1^2 and X_2 is normal with mean μ_2 and variance σ_2^2 and they are independent then $Y = X_1 + X_2$ is normal with mean $\mu_1 + \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$.

16: Recall that the geometric distribution with parameter p is given by $P(X = n) = p(1 - p)^{n-1}$ for $n = 1, 2, \ldots$. Let X and Y be independent geometric random variables both with parameter p.

(a) Show $P(X > t) = (1 - p)^t$ for $t \geq 1$ analytically. That is, show that $\sum_{n=t+1}^{\infty} p(1 - p)^{n-1} = (1 - p)^t$.

(b) Show that $P(X = m + k|X > m) = P(X = k)$.

(c) The distribution of $X + Y$ is negative binomial with parameters 2 and p. That is, $P(X + Y = k) = (k - 1)p^2(1 - p)^{k-2}$ for $k = 2, 3, \ldots$. Show this directly using a formula for the distribution of $X + Y$ in terms of a sum $\sum$.

(d) Compute $E(X|X + Y = k)$ for all $k = 2, 3, \ldots$.

(e) Explain each of the results above using the interpretation of a geometric random variable as giving the probability of n trials up to and including the first success.

17: The Gamma distribution with parameters $\alpha > 0$ and $\lambda > 0$ has density $f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}$ for $x \geq 0$ and $f(x) = 0$ for $x < 0$ where $\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1}\, dy$.
(a) Show that X has moment generating function $M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha$.
(b) Show, using any method that you like, that if $X_1, X_2, \ldots, X_n$ are independent exponentially distributed random variables with the same parameter λ then $X_1 + \cdots + X_n$ has gamma distribution with parameters n and λ.

18: Let T_i, $i = 1, 2, \ldots, n$ be independent exponential random variables with parameters λ_i.
(a) Determine (with proof) $P(\min(T_1, \ldots, T_n) > t)$. Use this to show that the minimum is exponential and determine its rate.
(b) For $n = 2$, prove that $P(T_1 < T_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$.
(c) Prove that $P(T_i = \min(T_1, \ldots, T_n)) = \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_n}$.

19: Consider a Poisson process with rate $\lambda = 10$. Given that exactly 5 events occur before time $t = 12$,
(a) What is the probability that all 5 of the events occur by time $s = 9$?
(b) What is the probability that exactly 2 of the events occur in the time interval $[3, 10]$?
(c) In general, for a Poisson process and $s < t$ show that

$$P\{N(s) = k | N(t) = n\} = \binom{n}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}, \text{ for } k = 0, 1, \ldots n$$

20: Consider Gambler's ruin. On each turn you win 1 with probability p or lose 1 with probability $q = 1 - p$. You quit when you reach 0 or N.
(a) Determine with a direct proof (not using martingales) the probability of reaching N before 0 starting at x. That is, the probability of winning. If you can't do this for general p at least do it for $p = 1/2$.
(b) Writing $P(s_i = 1) = p$ and $P(s_i = -1) = 1 - p = q$ and $S_n = S_0 + s_1 + \cdots + s_n$ with $g(x) = ((1-p)/p)^x$ show that $M_n = g(S_n)$ is a Martingale.
(c) Using the stopping theorem for martingales to generalize the result of part (a), determining the probability of reaching $b > a$ before a.

(III) Overview of the subjects.
Answer all three parts here.

21: You find yourself at a graduation party sitting with your father, your grandmother and one of your professors. They ask the following questions. What are your answers?
(a) From your father: We just spent all of this money for you to study probability and stochastic processes. What is it good for?
(b) From your grandmother: You just spent a lot of time studying probability and stochastic processes. What are they?
(c) From your professor: I just spent a lot of time trying to help you understand probability and stochastic processes. What results did you find really interesting and what did you think was a waste of time?

Clarifications:
(a) What do you say to someone who asks what the 'purpose' of (this particular area of) mathematics is. This might include potential applications that you know of and/or why its worthwhile even without applications.
(b) Try to think of one or two ideas or examples that you can describe in a few minutes without using too much technical language. These should illustrate or give some hint of what is being studied in these topics.
(c) State some result that you found particularly interesting. Say something about why you liked it. Also, point out something that you did not like and explain why. This might be because it was too complicated or because for some reason you thought the problem was not interesting or perhaps something else. (Saying that you hated topic X because the week that it was covered you also had a huge history paper is probably not a good answer.)