CS 174, Fall, 2004, Final, Karp, Solution

All questions count equally. Answer all questions. If the question does not ask for a proof or an explanation you may simply write down the answer, but it is prudent to show some work in case you have made a slip in arriving at your answer.

1. Let $X_1, \ldots, X_n$ be a sequence of independent binary random variables, with each X_i being equal to 1 with probability p. A maximal consecutive sequence of 1's is called a *run*. For instance, the sequence 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0 contains 3 runs. What is the expected number of runs in the sequence $X_1, \ldots, X_n$?

(n-1)p(1-p) + p

For i = 1, 2, ..., n-1 the probability that X_i is the last element of a run is $Pr(X_i = 1 \cap X_{i+1} = 0)$ = p(1-p). The probability that X_n is the last element of a run is $Pr(X_n = 1)$ = p. The result follows by linearity of expectation.

2. If the density function of X is ce^{-2x} for $0 < x < \infty$ and 0 for $x < 0$, find c. What is $Pr(X > 3)$?

The exponential distribution with rate λ has density function $\lambda e^{-\lambda x}$ and distribution function $1 - e^{-\lambda x}$. In this case $\lambda = 2$, so

c = 2 $Pr(X \geq 3) = e^{-2*3} = e^{-6}$.

3. Customers arrive at a bank according to a Poisson process with rate λ. If exactly two customers arrived during the first hour, what is the probability that both arrived during the first 20 minuts? What property of the Poisson process are you using to arrive at this conclusion?

$$\boxed{\frac{1}{9}}$$

Conditional over a given number of arrivals, the arrival times are independent and uniformly distributed over the time span of the process.

4. A gambler starts with a fortune of i dollars, (where $1 \le i \le n$). He repeatedly tosses a fair coin ($p = \frac{1}{2}$). Each time he tosses a head he wins \$1, and each time he tosses a tail he loses \$1. He plays until he either goes broke or reaches fortune N. What is the probability that he will reach fortune N?

$$\boxed{\frac{i}{N}}$$

Because the coin is fair, the gambler's expected fortune at the end of the process is the same as his initial fortune.

Alternatively, let A_i = the probability that, starting from a fortune of i, he reaches N. Then

$$A_0 = 0,\ A_N = 1,\ A_i = \frac{A_{i-1}+A_{i+1}}{2},\ i = 1, \ldots, n-1$$

The unique solution of this recurrence is $A_i = \frac{i}{N}$

5. In the coupon collector problem with $3n$ coupons, let $T(n)$ be the expected number of trials to collect n distinct coupons. Give the limiting value of $T(n)/n$ as n tends to infinity.

$$Tn = \sum_{i=0}^{n-1} \frac{3n}{3n-i} = 3n(H(3n) - H(2n)),$$

which is asymptotic to $3n(\ln 3n - \ln 2n) = \boxed{3n \ln (3/2)}$

6. Consider a Markov chain in which state i is recurrent and state j is transient. Prove: the transition probability P_{ij} is equal to zero.

Suppose for the purpose of contradiction that $P_{ij} > 0$.

Then i is not reachable from j. For, if it were, then i and j would communicate, and both would be recurrent or both would be transient.

But if $P_{ij} > 0$ and i is not reachable from j, then the probability of returning to i when starting at i is $\leq 1 - P_{ij} < 1$, so i is not recurrent.

Thus, assuming $P_{ij} > 0$ gives a contradiction, so $P_{ij} = 0$.

7. You have a biased coin with unknown probability of heads *p*. Using it, you wish to conduct an experiment that will simulate a fair coin: i.e., its result will be 1 with probability ½ and 0 with probability ½. Your experiment will be of the following form: toss the biased coin twice and observe the outcome (head or tail) of each toss. There are four possible ordered pairs of outcomes, which can be denoted HH, HT, TH, and TT, where H denotes head and T denotes tail. Depending on the ordered pair of outcomes, give the result 1, give the result 0 or start over.

   (a) Complete the specification of the experiment by stating what happens after each ordered pair of coin tosses.
   (b) Give the expected total number of coin tosses as a function of *p*, the probability of heads for the biased coin.

(a) On HT return 0
On TH return 1
On HH or TT, repeat the experiment

(b) $\boxed{\dfrac{1}{p(1-p)}}$

The number of pairs of coin tosses is geometric with parameter $2p(1-p)$ and hence has expectation $\dfrac{1}{2p(-p)}$

8. Use Chebyshev's Inequality to prove the *weak law of large numbers*. Namely, if $X_1, X_2, \ldots$ are independently distributed with mean μ and variance σ^2 then, for any $\varepsilon > 0$, $Pr\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| > \varepsilon\right)$ tends to 0 as n tends to ∞. This result seems to capture what most people mean by the law of averages.

$X_1 + X_2 + \ldots + X_n$ has variance $n\sigma^2$

$\frac{X_1 + X_2 + \ldots + X_n}{n}$ has variance $\frac{\sigma^2}{n}$

[Using Var[ax] = a^2Var[x]]

The Chebyshev Inequality gives

$Pr\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \mu\right| \geq \varepsilon\right) \geq \frac{\sigma^2/n}{\varepsilon}$, which tends to zero as n -> ∞.

9. Give the moment generating function $E[e^{tX}]$ of the exponential distribution with rate λ and use differentiation to prove that the expectation is $1/\lambda$ and the variance is $1/\lambda^2$. Hint: if X is a continuous random variable then $E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx$. You may assume that $t < \lambda$.

$$M_x(t) = E[e^{tX}] = \int_0^{\infty} e^{tX} \lambda e^{-\lambda x}\, dx$$

$$= \frac{\lambda}{t - \lambda} e^{(t-\lambda)x} \Big]_0^{\infty} = \lambda(\lambda - t)^{-1}$$

$$M_x'(t) = \frac{\lambda}{(\lambda - t)^2} \qquad M_x'(0) = E[x] = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

$$M_x''(t) = \frac{2\lambda}{(\lambda - t)^3} \qquad M_x'(0) = E[x] = \frac{2\lambda}{\lambda^3} = \frac{2}{\lambda^2}$$

$$\mathrm{Var}[x] = E[x^2] - E[x]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

10. Thirteen men and seventeen women are seated around a circular table:
Prove: there is some set of nine consecutive seats at which three or fewer men are seated.
Hint: use the probabilistic method.

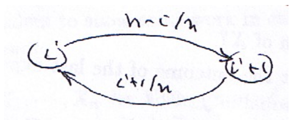
Let X be the number of men in a randomly chosen sequence of nine consecutive seats.

$$E[X] = \frac{13}{30} * 9 = \frac{117}{30} < 4.$$

Therefore, there must be some sequence of nine consecutive seats in which the number of men is < 4, and hence ≤ 3.

11. A Markov chain has state set $\{0, 1, ..., n\}$ and the following transition probabilities: $p_{i,i+1} = (n - i)/n$, $i = 0, 1, ..., n - 1$ and $p_{i,i-1} = i/n$, $i = 1, 2, ..., n$. Give the stationary distribution.

Use the method of cut sets.



The probability of leaving $\{0, 1, ..., i\}$ under the stationary distribution is equal to the probability of entering $\{0, 1, ..., i\}$ under the stationary distribution.

Hence $\frac{n - i}{n} \Pi_i = \frac{i + 1}{n} \Pi_{i+1}$

$$\Pi_{i+1} = \frac{n - i}{i + 1} \Pi_i$$

By reduction on i, $\Pi_i = \binom{n}{i} \Pi_0$

$1 = \Sigma^n_{i=0} \binom{n}{i} \Pi_0 = 2^n \Pi_0$, so $\Pi_0 = 2^n$

and $\boxed{\Pi_i = \frac{\binom{n}{i}}{2^n}}$

12. A sequence of independent experiments is conducted. Each experiment succeeds with probability p, fails with probability q, and has a neutral outcome with probability $1 - p - q$. The sequence terminates as soon as some experiment succeeds or fails. Let X be the number of experiments conducted.

(a) What is the expectation of X?
(b) Let S be the event that the outcome of the last experiment is a success. What is the probability of S?
(c) What is the conditional expectation of X given S?

(a) $\frac{1}{p + q}$

(b) $p/(p+q)$

(c) $Pr(X = i \mid \mathrm{S}) = \frac{Pr(X = i \text{ n } S)}{Pr(S)}$

$$= \frac{(1 - p - q)^{i-1}p}{p/p+q} = (1 - p - q)^{i-1}(p + q)$$

Hence the conditional distribution of X given S is geometric with parameter $p + q$, and its expectation is $\boxed{\frac{1}{p + q}}$