

This handout is meant to provide a collection of exercises that use the material from the probability and statistics portion of the course. The answers to the exercises are at the end.

Don't take the exercises as indicative of the final exam problems. In particular, some of these exercises are quite involved and are designed not so much to test your knowledge as to broaden your view of a given topic as you work through the answers. In any event, if you can come to terms with the exercises below, you will do fine with the probability and statistics part of the final exam.

- 1 Which of the following are not probability functions on  $[0, 1]$ ?
  - (a)  $\frac{\pi}{2} \sin(\pi x)$
  - (b)  $e^{-x}$
  - (c)  $\frac{\pi}{2} \cos(\pi x)$
  - (d)  $\frac{2}{(x+1)^2}$
- 2 Using the probability function  $e^{-x}$  on  $[0, \infty)$ , compute the probability that  $x$  lies in  $[1, 2] \cup [3, \infty)$ .
- 3 Using the uniform probability function  $\frac{1}{2\pi}$  on  $[0, 2\pi]$ , compute the probability that  $\sin^2(\theta)$  is less than  $\frac{1}{2}$ .
- 4 Give the sample space  $[0, \infty)$  the exponential probability function  $e^{-x}$ . Suppose that  $a$  and  $b$  are such that  $0 \leq a < b$ . Write down the probability that the random variable  $x \rightarrow x^2$  has a value between  $a$  and  $b$ .
- 5 Given that  $x$  is less than 4, what is the probability as computed using the exponential function  $e^{-x}$  on  $[0, \infty)$  that  $x$  is greater than 2?
- 6 Take the Gaussian probability function with mean 3 and standard deviation 2 for the line,  $(-\infty, \infty)$ . Are there sets  $A$  and  $B$  in  $(-\infty, \infty)$  such that  $P(A | B) = \frac{1}{3}$ ,  $P(B | A) = \frac{1}{10}$ , and  $P(B) = \frac{1}{2}$ ?
- 7 Suppose that  $N$  is a positive integer. Write down the sample space for the possible outcomes of flipping  $N$  coins simultaneously. Write a sum that gives the probability of at least three heads on the  $N$  coins if the probability of getting heads on any one coin is  $\frac{1}{4}$  and if it is assumed that the outcome on any one flip has no bearing on that of any other.
- 8 Use the exponential function  $e^{-x}$  as a probability function on  $[0, \infty)$ .
  - (a) What is the probability that the closest integer to  $x$  is odd?
  - (b) What is the conditional probability that the integer closest to  $x$  is odd given that  $x$  is larger than 1?
  - (c) Is the event that the integer closest to  $x$  is odd independent from the event that  $x$  is greater than 2?
- 9 Compute the mean and standard deviation of the random variable  $x \rightarrow x^2$  for the probability function  $\frac{6x}{(x^2+1)^4}$  on  $[0, \infty)$ .
- 10 Suppose that  $-\infty \leq a < b \leq \infty$  and that  $p(x)$  is a probability function on  $[a, b]$  with mean  $\mu$ . Explain why  $\int_a^b x^2 p(x) dx - \mu^2$  is the square of the standard deviation.
- 11 Use the probability function  $\frac{1}{2} \cos(x)$  on  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ . What are the mean and standard deviation for the random variable  $x \rightarrow \sin^2(x)$ ?
- 12 Use the probability function  $\frac{1}{2} \cos(x)$  on  $[-\frac{\pi}{2}, \frac{\pi}{2}]$ . Which of the following is the probability that the random variable  $\sin^2(x)$  has value between 0 and  $\frac{1}{5}$ ?
  - (a)  $\sqrt{\frac{1}{5}}$
  - (b)  $\frac{1}{5}$
  - (c)  $\frac{\pi}{5} \sqrt{\frac{1}{5}}$
  - (d)  $\frac{1}{15} \sqrt{\frac{1}{5}}$
  - (e)  $\frac{1}{15}$

- 13 Let  $N$  be a positive integer, and let  $S$  denote the sample space for the possible outcomes of  $N$  coin flips. Give  $S$  the probability function for which the probability of any given coin landing heads is independent from any other, and for which the probability of any given coin landing heads is  $\frac{1}{4}$ . Define  $f$  to be the random variable that multiplies the number of heads times the number of tails.
- Which of the following is the mean of  $f$ :  $N$ ,  $\frac{1}{4}(N^2 + N)$ ,  $\frac{3}{16}N^2$ , or  $\frac{3}{16}N(N - 1)$ ?
  - Assume that  $N > 2$  and is even. Is  $f$  independent of the random variable that gives 0 when the number of heads is odd and 1 when the number of heads is even?
  - Write down a sum that gives the characteristic function for the random variable that is 0 when the number of heads is odd and 1 when the number of heads is even.
- 14 Use the probability function  $\frac{1}{2\pi}$  on  $[-\pi, \pi]$ . Are the random variables  $x \rightarrow \cos(x)$  and  $x \rightarrow \sin(x)$  independent?
- 15 Suppose that the probability that I am sick on any particular day is  $\frac{1}{100}$ , the probability that I am absent any particular day is  $\frac{1}{200}$ , and the probability that I am absent given that I am sick is  $\frac{1}{10}$ . If I am absent on a given day, what is the probability that my absence is due to being sick?
- 16 Suppose that a system has four possible states, these labeled as  $\{1, 2, 3, 4\}$ . Suppose, in addition, that after any given unit of time, the probability from going from any even state to any odd state is  $\frac{1}{6}$ , the probability of going from any even state to any even state is  $\frac{1}{3}$ , the probability for going from any odd state to any even state is  $\frac{1}{3}$  and the probability from going from any odd state to any odd state is  $\frac{1}{6}$ . Let  $\vec{p}(t)$  denote the vector whose  $k$ th component is the probability of being in state  $k$  at time  $t$ .
- Write down a matrix,  $A$ , such that the equation  $\vec{p}(t + 1) = A\vec{p}(t)$  holds.
  - What is  $\lim_{t \rightarrow \infty} \vec{p}(t)$ ?
  - Denote the vector you found in part (b) by  $\vec{q}$ . Explain why  $\vec{p}(1) = \vec{q}$  as long as the entries of  $\vec{p}(0)$  sum to one.
- 17 Suppose that three coins are flipped simultaneously and we don't know the probability of heads although we do know that the probability is the same for each coin, and that the probability for any one coin is independent of that for any other. Define a function,  $f$ , on the sample space to be 1 if there is an odd number of heads and zero otherwise. Thus,  $f$  has two possible values.
- Given that the probability of heads on any coin is some number  $p \in [0, 1]$ , write down the probability that  $f = 1$  in terms of  $p$ .
  - In terms of  $p$ , what is the mean and standard deviation for  $f$ ?
  - Suppose that we now flip the three coins some large number of times and see that the average value of  $f$  is  $\frac{7}{16}$ . According to the Central Limit Theorem, which of the following is the choice for  $p$ :  $\frac{1}{16}$ ,  $\frac{1}{8}$ ,  $\frac{1}{4}$ ,  $\frac{1}{2}$ ?
  - If the fraction of times  $f$  is 1 after a large number of flips is  $\frac{7}{16}$ , what would a Bayesian guess for the probability function on the sample space for flipping three coins?
- 18 Suppose that an experiment is repeated 100 times and a certain measurement can have one of two values either 1 and 0 each time. In this regard, assume that the probability of measuring 1 is the same for each run, and that the probability for a measurement on any given run is independent of that for any other. If the probability of measuring 1 on any given run is postulated to be  $\frac{1}{10}$ , what is the probability of precisely  $n \in \{0, 1, \dots, 100\}$  of the runs giving the measurement 1?

- 19 As in Problem 18, suppose that an experiment is repeated 100 times and a certain measurement can have one of two values either 1 and 0 each time. Assume that the probability of measuring 1 is the same for each run, and that the probability for a measurement on any given run is independent of that for any other. As before, assume that the probability of measuring 1 on any given experiment is  $\frac{1}{10}$ . Let  $n$  denote the number of times that 1 is measured. Differentiate some number of times the characteristic polynomial for the associated binomial probability function to determine the mean of the random variable  $n \rightarrow f(n)$  where  $f(0) = f(1) = f(2) = 0$  and  $f(n) = n(n-1)(n-2)$  in the case that  $n \geq 3$ .
- 20 Continue here with the setup used in the previous two problems. Suppose that the measured number of occurrences of the measurement 1 is 25 after 100 runs of the experiment. Given that you will abandon your assumption of  $\frac{1}{10}$  as the probability of getting 1 on any given run if the  $P$ -value for the number of occurrences is less than  $\frac{1}{20}$ , should you do so?
- 21 Make up a scenario where you would want to use a Poisson probability function to predict the frequency of occurrences in repeated runs of a given experimental protocol.
- 22 Suppose that a given situation has possible outcomes in the set  $\{0, 1, 2, \dots\}$  of non-negative integers and is modeled using the Poisson probability function with mean 25. Explain why  $\frac{1}{9}$  is an upper bound for the probability of measuring  $n \geq 40$  and why  $\frac{1}{25}$  is an upper bound for measuring  $n \geq 50$ .
- 23 Suppose that large meteor hits on the earth are unrelated and that the average time between hits is 30 million years.
- Use an exponential probability function to determine the probability of a large meteor hitting the earth in the next 100 years (your lifetime plus change).
  - Given that the last large meteor hit on the earth was 65 million years ago, what is the probability of a large meteor hitting the earth in the next 100 years?
- 24 Suppose that a given experiment is repeated many times and a certain quantity is measured each time. Suppose that the mean of the measurements is 10 and the variance is 2. (The variance is the square of the standard deviation; that is, it is the average of the squares of the differences from the mean.)
- Write down the most likely Gaussian probability function to use to predict the fraction of measurements that lie between 6 and 8.
  - Use your answer to part (a) to write down an integral that gives the desired prediction.
- 25 Suppose that the birth weights of newborn children in Iceland distribute in a random fashion between 6 and 9 pounds. Give a Gaussian probability function that can be used to estimate the average birth weight of 1000 Icelandic babies.
- 26 Redo Problem 25 when the variations in the birth weights of newborn children in Iceland are modeled instead by the Gaussian probability function with mean 7.5 and standard deviation  $\frac{1}{2}\sqrt{3}$ .
- 27 In Las Vegas, you can gamble on the outcome of rolling a pair of six sided dice. The sides of each die are numbered from 1 to 6 and you win if the sum of the numbers showing on top are either 7 or 11. You lose otherwise. The probability of winning if the dice are unbiased is  $\frac{2}{9}$ . If you are interested in determining whether the dice used are fair, you can watch the game played many times. Suppose that you watch  $N$  games.
- Assuming the dice are fair, give an expression in terms of  $n \in \{0, 1, \dots, N\}$  for the probability of  $n$  wins in  $N$  games.

- (b) Use the Chebychev Theorem to estimate an upper bound for the probability of seeing less than 144 wins in 900 games under the assumption that the dice are fair. You needn't evaluate square roots or exponentials that appear in your answer.

28 Let  $S$  denote the sample space of pairs of the form  $(j, k)$  where  $j$  and  $k$  are integers in the set  $\{1, \dots, n\}$ . Suppose that  $S$  has a probability function, and let  $P(j | k)$  denote the conditional probability that the first entry in a pair is  $j$  given that the second entry is  $k$ .

- (a) Explain why the matrix whose entry in the  $j$ th column and  $k$ th row is  $P(j | k)$  is a Markov matrix.  
 (b) Suppose that  $A$  is an  $n \times n$  Markov matrix and  $p$  is a probability function on the set  $\{1, \dots, n\}$ . Use  $A$  and  $p$  to determine a probability function on  $S$  with the following two properties: First,  $p(k)$  is the probability that the second entry of any given pair from  $S$  is  $k$ . Second, the conditional probability  $P(j | k)$  is  $A_{jk}$ .

29 Let  $A$  denote the Markov matrix  $\begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{bmatrix}$ . Suppose that  $\vec{p}(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$  and that for any  $t \in \{1, 2, \dots\}$ , the vector  $\vec{p}(t)$  is defined to equal  $A\vec{p}(t-1)$ .

- (a) What is  $\lim_{t \rightarrow \infty} \vec{p}(t)$ ?  
 (b) How big must  $t$  be before before the difference between this limit and  $\vec{p}(t)$  is a vector whose length is less than  $\frac{1}{100}$ ?

30 Suppose that the concentration in the blood of a given medicinal drug is measured as a function of time. In particular, suppose that measurements are made at a sequence of times  $t_1 < t_2 < \dots < t_n$ , that yield the corresponding sequence of values  $\{y_1, \dots, y_n\}$ . In addition, suppose that it is believed that the function,  $t \rightarrow y(t)$ , that describes the concentration as a function of  $t$  has the form  $y(t) = ae^{-t} + be^{-2t}$  where  $a$  and  $b$  are constants. Derive an expression in terms of the data  $\{(t_k, y_k)\}$  for the pair  $(a, b)$  that minimizes  $\sum_k (y(t_k) - y_k)^2$ .

31 Let  $S = \{0, 1, 2, \dots, 9\}$  and let  $p$  denote the probability function on  $S$  that assigns the value  $\frac{1}{10}$  to each digit.

- (a) What are the mean and standard deviation for  $p$ ?  
 (b) What is the sample space for a sequence of 100 digits, each chosen from  $S$ ?  
 (c) Write down a Gaussian probability function that approximates the probabilities for the average 100 digits chosen at random from  $S$ . The integral of the latter function over any given interval of the form  $[a, b]$  should give a good approximation for the probability that the average of the 100 should be greater than or equal to  $a$  and less than or equal to  $b$ . Note that the term "at random" in this case should be taken to mean that any one of the  $10^{100}$  possible sequence of 100 digits has probability  $10^{-100}$  of appearing.

32 Suppose that on average, 1 case of a certain sort of cancer will appear per year in any population of 10,000 individuals. Suppose a town of population 100,000 sees 20 cases one year. Write an infinite sum that gives the  $P$ -value for the hypothesis that these cases are unrelated and that the appearance of this many cases is due to chance?

33 Suppose that 20% of people with a certain genetic variant develop a certain sort of skin cancer, and that 10% of the population of certain locale has this gene. Meanwhile, 5% of the people in this locale develop the cancer. Given that an individual from this locale has the cancer, what is the probability that the person has the genetic variant?

34 Suppose that a certain sort of bacteria moves along a line by sequence of flips, with each flip either one body length in the + direction, or one body length in the – direction. Suppose that the rules for deciding are as follows: The probability of flipping in a given direction if the previous flip was in that direction is  $\frac{2}{3}$ ; and the probability of flipping in a given direction if the previous flip was in the opposite direction is  $\frac{1}{3}$ . For each  $t \in \{1, 2, \dots\}$ , let  $p_+(t)$  denote the probability that the  $t$ th flip is in the + direction, and let  $p_-(t)$  denote the probability that the  $t$ th flip is in the minus direction.

- For each  $t > 1$ , give a linear equation that relates  $p_+(t)$  with  $p_+(t-1)$  and  $p_-(t-1)$ . Then, do the same for  $p_-(t)$ .
- Let  $\vec{p}(t) \in \mathbf{R}^2$  denote the vector whose top component is  $p_+(t)$  and whose bottom component is  $p_-(t)$ . Find a matrix  $A$  so that the equation  $\vec{p}(t) = A\vec{p}(t-1)$  holds for each  $t > 1$ .
- Is  $A$  a Markov matrix?
- Find the eigenvectors and eigenvalues of the matrix  $A$ .
- Find  $\vec{p}(t)$  given that  $p_+(1) = 1$  and  $p_-(1) = 0$ .
- Find  $\lim_{t \rightarrow \infty} p_+(t)$ .

35 Suppose that a survey of finds the following fact: When  $N$  is sufficiently large, roughly 51% of  $N$  births are female and so 49% are male. Said differently,  $\frac{1}{N}(\# \text{female} - \# \text{male}) = 0.02$  when  $N$  is large. Granted this data, suppose that we make the hypothesis that the probability of having a female child is  $\frac{1}{2}$  and so the probability of having a male child is also  $\frac{1}{2}$ .

- Under the assumption that the hypothesis of equal probabilities is correct, use the Central Limit Theorem to write down a probability function that gives an accurate approximation when  $N$  is large for the probability that  $\frac{1}{2}(\# \text{female} - \# \text{male})$  has its value in any given interval.
- Use the Chebychev theorem with the mean and standard deviation from your probability function from part (a) to answer the following question: How big must  $N$  be before our hypothesis of equal probabilities has  $P$ -value less than 0.05?

- 1 (b) is not normalized to have integral 1 and (c) is negative in places.
- 2  $e^{-1} - e^{-2} + e^{-3}$ , which is the sum of the integrals of  $e^{-x}$  over the two regions.
- 3 The function  $\sin^2 \theta$  is less than  $\frac{1}{2}$  on half of the circle, so the probability is  $\frac{1}{2}$ .
- 4  $e^{-\sqrt{a}} - e^{-\sqrt{b}}$ . Indeed,  $x^2$  is between  $a$  and  $b$  if, and only if,  $x$  is between  $\sqrt{a}$  and  $\sqrt{b}$ .
- 5  $(e^{-2} - e^{-4})/(1 - e^{-4})$ .
- 6 No. Here is why: Since  $P(A|B) = \frac{P(A \cap B)}{P(B)}$  and  $P(B|A) = \frac{P(A \cap B)}{P(A)}$ , one has  $P(A) = \frac{P(A|B)P(B)}{P(B|A)}$ . With the numbers given, this would give  $P(A) > 1$ .
- 7  $S$  is the set of all  $N$ -tuples of the form  $(a_1, \dots, a_N)$  where each  $a_k$  can be either  $H$  or  $T$ . The probability is  $\sum_{k=3}^{\infty} \binom{N}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{N-k} = \sum_{k=3}^{\infty} \frac{N!}{k!(N-k)!} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{N-k}$ .
- 8 (a) The closest integer to  $x$  is odd if and only if  $2k + \frac{1}{2} < x < 2k + \frac{3}{2}$  for some integer  $k$  from the set  $\{0, 1, 2, \dots\}$ . Thus, the probability is

$$e^{-1/2} - e^{-3/2} + e^{-5/2} - e^{-7/2} + \dots = e^{-1/2} \sum_{k=0}^{\infty} (-1)^k e^{-k} = e^{-1/2} \cdot \frac{1}{1 + e^{-1}}.$$

(b) This is  $(e^{-1} - e^{-3/2} + e^{-5/2} - e^{-7/2} + \dots)/e^{-1} = 1 - e^{-1/2} \cdot \frac{1}{1 + e^{-1}}$ .

(c) Yes. The probability that the closest integer to  $x$  is odd and  $x$  is greater than 2 is equal to

$$e^{-5/2} - e^{-7/2} + \dots = e^{-2} (e^{-1/2} - e^{-3/2} + e^{-5/2} - e^{-7/2} + \dots).$$

The latter is the product of the probability that  $x$  is greater than 2 (this being  $e^{-2}$ ) with the probability found from part (a) that the closest integer to  $x$  is odd. A similar argument shows that the event that the integer closest to  $x$  is odd is independent from the event that  $x$  is greater than any given even integer.

- 9 The mean is  $\frac{1}{2}$  and the standard deviation is  $\sqrt{\frac{3}{4}}$ .
- 10 As defined,  $\sigma^2 = \int_a^b (x - \mu)^2 p(x) dx$  and writing out the square finds this equal to  $\int_a^b x^2 p(x) dx - 2\mu \int_a^b x p(x) dx + \mu^2 \int_a^b p(x) dx$ . Since the integral that appears in the middle term is  $\mu$  and that in the final term is 1, this is  $\int_a^b x^2 p(x) dx - 2\mu^2 + \mu^2$ .
- 11 The mean is  $\frac{1}{3}$  and the standard deviation is  $\frac{2}{3\sqrt{5}}$ .
- 12  $\sqrt{\frac{1}{5}}$ .
- 13 (a) Let  $n$  denote the number of heads. Then  $f = n(N - n)$ . We know that  $n$  can take values in  $\{0, \dots, N\}$  and that the probability of any given value,  $k$ , from this set is given by  $P(k) = \binom{N}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{N-k}$ . Thus, the mean of  $n$  is  $\frac{1}{4}N$  and that square of the standard deviation of  $n$  is  $\frac{3N}{16}$ . Now, the square of the standard deviation is the mean of  $n^2 - \frac{1}{16}N^2$ , so the mean of  $n^2$  is  $\frac{3N}{16}N + \frac{1}{16}N^2$ . Thus, the mean of  $f$  is  $N$  times the mean of  $n$  minus the mean of  $n^2$ . Granted this, then the mean of  $f$  is therefore be equal to  $\frac{1}{4}N^2 - \frac{3}{16}N - \frac{1}{16}N^2 = \frac{3}{16}N(N - 1)$ .

(b) When  $n$  is even, then  $f$  is even; and when  $n$  is odd, then  $f$  is odd. Thus, the random variable that gives 0 when  $n$  is odd and 1 when  $n$  is even can be written as  $(1 + (-1)^f)$ . This shows that the two random variables can not be independent.

(c) Let  $\alpha = \sum_{\substack{k \text{ is odd and} \\ 1 \leq k \leq N}} \binom{N}{k} (\frac{1}{4})^k (\frac{3}{4})^{N-k}$ , thus  $\alpha$  is the probability that  $n$  is odd.

The characteristic functional is equal to  $(1 - \alpha) + \alpha t$ . Note that the sum for  $\alpha$  can be computed in closed form:  $\alpha = \frac{1}{2} + (\frac{1}{2})^{N+1}$ .

14 These two random variables are not independent. This can be seen from the fact that the absolute value of one can be deduced knowing that of the other.

15  $\frac{1}{5}$ . Use Bayes' rule:  $P(\text{Sick} | \text{Absent}) = \frac{P(\text{Absent} | \text{Sick})P(\text{Sick})}{P(\text{Absent})}$ .

16 (a) 
$$\begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$
.

(b) 
$$\vec{q} = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{3} \\ \frac{1}{6} \\ \frac{1}{3} \end{bmatrix}$$
.

(c) Since the span of the image of  $A$  is 1-dimensional (all of its columns are the same),  $A$  has three eigenvalues that are zero. Denoting these by  $\vec{v}_1, \vec{v}_2$  and  $\vec{v}_3$ , the vector  $\vec{p}(0)$  must have the form  $\vec{q} + a_1\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3$  since the entries of each  $\vec{v}_k$  sum to zero and the entries of  $\vec{q}$  sum to 1. Thus,  $A\vec{p}(0) = \vec{q}$ .

17 (a)  $p^3 + 3p(1 - p)^2 = p(4p^2 - 6p + 3)$ .

(b)  $\mu = p(4p^2 - 6p + 3)$  and  $\sigma = \sqrt{\mu - \mu^2}$ .

(c)  $\frac{1}{4}$

(d)  $P(HHH) = P(HTT) = P(THT) = P(TTH) = \frac{7}{64}$ ,  
 $P(TTT) = P(THH) = P(HTH) = P(HHT) = \frac{7}{64}$ .

To explain, the task is to guess the values for a probability function,  $P$ , on the set  $\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$ . Label these elements from 1 to 8. Let  $q_0 = \frac{9}{16}$  and  $q_1 = \frac{7}{16}$ . Then one can write  $q_0 = p_5 + p_6 + p_7 + p_8$  and likewise write  $q_1 = p_1 + p_2 + p_3 + p_4$ .

This has the form of  $q_a = \sum_{j=1}^8 B_{aj}p_j$ . The Bayesian will take  $p_j = \sum_{a=0,1} \frac{B_{aj}}{\sum_k B_{ak}} q_a$ .

18  $\binom{100}{n} (\frac{1}{10})^n (\frac{9}{10})^{100-n} = \frac{100!}{n!(100-n)!} (\frac{1}{10})^n (\frac{9}{10})^{100-n}$ .

19 The characteristic function is the polynomial in  $t$  given by  $t \rightarrow \mathcal{P}(t) = (\frac{9}{10} + \frac{1}{10}t)^{100}$ . This is equal to  $\sum_{n=0}^{100} t^n P(n)$  where  $P(n) = \binom{100}{n} (\frac{1}{10})^n (\frac{9}{10})^{100-n}$  is the probability of seeing  $n$  occurrences of 1. Differentiating the identity  $\mathcal{P}(t) = \sum_n t^n P(n)$  three times finds that its third derivative at  $t = 1$  equals

$\sum_{n \geq 3} n(n-1)(n-2)P(n)$ . This is the mean of the random variable  $f(n)$ . Meanwhile, differentiating  $(\frac{9}{10} + \frac{1}{10}t)^{100}$  three times and setting  $t = 1$  gives  $100 \cdot 99 \cdot 98 \cdot \frac{1}{10^3} = 88.2$ . Thus, the mean is 88.2.

20 Yes. The mean for the binomial probability in this case is 10 and the standard deviation is 3. As  $25 - 10 = 15 = 5\sigma$ , and the square of 5 is 25, the Chebychev inequality asserts that the probability of 25 or more occurrences of 1 is less than  $\frac{1}{25}$ .

21 The average number of large meteor hits on the earth is 1 per 30 million years. Assuming that these meteors are not traveling together (that is, not two halves of some broken comet), how likely is it for a two huge meteors to hit the earth in 2005?

22 The mean is 25 so the standard deviation is 5. Granted this, use the Chebychev inequality.

23 (a)  $1 - e^{-x}$ , where  $x = 3 \times 10^{-5}$ . This is pretty close to  $3 \times 10^{-5}$ .

(b)  $1 - e^{-x}$ , where  $x = 3 \times 10^{-5}$ .

24 (a)  $\frac{1}{2\sqrt{\pi}} e^{-(x-10)^2/8}$

(b)  $\int_6^8 \frac{1}{2\sqrt{\pi}} e^{-(x-10)^2/8} dx$

25  $\sqrt{\frac{1000}{3\pi}} e^{-2000(x-7.5)^2/3}$

26  $\sqrt{\frac{1000}{3\pi}} e^{-2000(x-7.5)^2/3}$

27 (a)  $\binom{N}{n} (\frac{2}{9})^n (\frac{7}{9})^{N-n} = \frac{N!}{n!(N-n)!} (\frac{2}{9})^n (\frac{7}{9})^{N-n}$

(b) Let  $S = \{W, L\}$  where  $W$  means win and  $L$  means lose. Define a random variable,  $f$ , on  $S$  so that  $f(W) = 1$  and  $f(L) = 0$ . The mean of  $f$  is  $\frac{2}{9}$  and the standard deviation is  $\frac{\sqrt{2}}{3}$ . Let  $\{f_1, \dots, f_{900}\}$  denote 900 identical versions of this same random variable. According to the Central Limit Theorem, the probabilities for the values of  $\bar{f} = \frac{1}{900}(f_1 + \dots + f_{900})$  are determined by the Gaussian probability function with mean  $\frac{2}{9}$  and standard deviation  $\frac{1}{\sqrt{900}} \cdot \frac{\sqrt{2}}{3} = \frac{\sqrt{2}}{90}$ . Now,  $\frac{144}{900}$  differs from  $\frac{2}{9}$  by  $\frac{1}{25}$  and this is  $R \cdot \frac{\sqrt{2}}{90}$  when  $R = \frac{18}{5\sqrt{2}}$ . The Central Limit Theorem therefore finds that the probability of less than 144 wins is less than  $\frac{1}{\sqrt{\pi}} \frac{9}{5} e^{-81/25}$ .

28 (a) Each  $P(j | k)$  is non-negative since these are probabilities, and  $P(1 | k) + \dots + P(n | k) = 1$  for each  $k$  since the sum of the conditional probabilities must be 1.

(b) The definition of the conditional probability tells us that we should set  $P(j | k)$  to equal the quotient of probabilities  $\frac{P(j,k)}{p(k)}$ . Thus,  $P(j,k) = P(j | k) p(k) = A_{jk} p(k)$ .

29 (a) This limit is  $\vec{v}_1 = \begin{bmatrix} 3 \\ 7 \\ 4 \\ 7 \end{bmatrix}$ .

(b) The other eigenvalue is  $-\frac{1}{6}$  and the eigenvector is  $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ . Thus,  $\vec{p}(0) = \vec{v}_1 - \frac{3}{7}\vec{v}_2$  and  $\vec{p}(t) - \vec{v}_1 = -\frac{3}{7}(-\frac{1}{6})^t \vec{v}_2$  has norm  $\sqrt{2} \cdot \frac{3}{7} (\frac{1}{6})^t$ . Thus, we need  $t \geq \frac{1}{\ln 6} \ln(\frac{300}{7} \sqrt{2}) \approx 2.3$ .

31 Let  $A$  denote the matrix with  $n$  rows and 2 columns whose  $k$ th row is  $(e^{t_k}, e^{-2t_k})$ . Then

$$\begin{bmatrix} a \\ b \end{bmatrix} = (A^T A)^{-1} A^T \vec{y}$$

where  $\vec{y}$  is the vector in  $\mathbf{R}^n$  whose  $k$ th entry is  $y_k$ .

- 31 (a) The mean is 4.5 and the standard deviation is  $\sqrt{8.25}$ .  
 (b) The sample space is the set of decimal fractions of the form  $0.a_1a_2\cdots a_{100}$ .  
 (c) The Central Limit Theorem finds this to be the function

$$p(x) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{100}{825}} e^{-50(x-4.5)^2/825}.$$

32 Consider the sample space that consists of the non-negative integers  $\{0, 1, 2, \dots\}$ . If the hypothesis is correct, then the probability of seeing  $n$  cases in a population of 100,000 is given by the Poisson probability function  $p(n) = e^{-10} \frac{10^n}{n!}$ . Thus, the  $P$ -value of seeing 20 cases is  $\sum_{n \geq 20} e^{-10} \frac{10^n}{n!}$ . This is smaller than 0.05 and so the  $P$ -value is significant.

33 Use  $V$  to denote the set of people with the given genetic variant, and let  $C$  denote the set of people that develop cancer. We are told that  $P(C | V) = 0.2$ ,  $P(V) = 0.1$  and that  $P(C) = 0.05$ . We are asked for  $P(V | C)$ . Bayes' theorem asserts that this last number is equal to  $\frac{0.2 \times 0.1}{0.05} = 0.4$ .

34 (a)  $p_+(t) = \frac{2}{3}p_+(t-1) + \frac{1}{3}p_-(t-1)$  and  $p_-(t) = \frac{1}{3}p_+(t-1) + \frac{1}{3}p_-(t-1)$ .

(b)  $A = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$

(c)  $A$  is Markov since the entries in each column sum to 1 and no entry is negative.

(d) The eigenvalues are 1 and  $\frac{1}{3}$  with respective eigenvectors  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  and  $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ .

(e)  $\vec{p}(t) = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \left(\frac{1}{3}\right)^t \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ .

(f)  $\lim_{t \rightarrow \infty} \vec{p}(t) = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ .

35 The sample space for  $N$  births consists of the set of  $2^N$  strings of the form  $(a_1, \dots, a_N)$ , where each  $a_k$  can be either  $F$  (for female) or  $B$  (for male). We can view the quantity  $\frac{1}{N}(\# \text{female} - \# \text{male})$  as the average of  $N$  independent, random variables where each is defined on the 2-element sample space,  $\{F, B\}$ . There, the random variable,  $x$  assigns 1 to the element  $F$  and  $-1$  to  $B$ . This random variable has mean 0 and standard deviation 1.

(a) The Central Limit Theorem then asserts that the approximate probability for the average,  $\bar{x}$ , of  $N$  such random variables should be obtained using the Gaussian with mean zero and standard deviation  $\sigma_N = \frac{1}{\sqrt{N}}$ .

(b) The survey data finds that the distance from the mean for  $\frac{1}{N}(\# \text{female} - \# \text{male})$  is 0.02 which is  $0.02\sqrt{N}\sigma_N$ . According to the Chebychev theorem, the probability of being further from the mean than  $R\sigma_N$  is less than  $\frac{1}{R^2}$ . As such, the hypothesis of equal probability has significant  $P$ -value when  $R = 0.02\sqrt{N} > \sqrt{20}$ . This occurs when  $N \geq (50)^2 20 \approx 50,000$ .