MATH/STAT 3360, Probability
FALL 2011
Toby Kenney
Midterm Examination
Model Solutions

Basic Questions

1. *How many distinct ways can the letters of the word "STATISTICS" be arranged.*

   There are a total of 10 letters, 3 'S', 3 'T', 2 'I', 1 'A' and 1 'C', so they can be rearranged in $\frac{10!}{3!3!2!}$ orders. [=50400 ways].

2. *What is the probability that a five-card poker hand is a flush (all 5 cards the same suit).*

   There are 4 suits, and for each suit there are $\binom{13}{5}$ five-card hands that are flushes in that suit. Therefore, the probability of a flush is $\frac{4\binom{13}{5}}{\binom{52}{5}}$. [=0.00198].

3. *A fair coin is tossed 7 times.*

   (a) What is the probability that the sequence HHHT occurs somewhere in the 7 tosses?

   This is the union of 4 events, namely a sequence of HHHT starting on toss i for $i = 1, 2, 3, 4$. Let these events be S_1, S_2, S_3 and S_4. Each has probability $\frac{1}{16}$, and they are mutually exclusive, so the probability that one of them occurs is $\frac{4}{16} = \frac{1}{4}$.

   (b) What is the probability that the sequence THTH occurs somewhere in the 7 tosses?

   This is the union of 4 events, namely a sequence of THTH starting on toss i for $i = 1, 2, 3, 4$. Let these events be S_1, S_2, S_3 and S_4. Each has probability $\frac{1}{16}$, and furthermore, $P(S_1 \cap S_3) = \frac{1}{64}$ and $P(S_2 \cap S_4) = \frac{1}{64}$, and it is not possible for three or more of the events to happen at once. Therefore, the probability that THTH occurs somewhere is $4\times\frac{1}{16} - 2\times\frac{1}{64} = \frac{7}{32}$.

4. *Two cards are drawn from a standard deck. Are the following events independent?*

   (i) Exactly one of the cards is a 5

   (ii) Both cards are the same suit.

   Let A be the event that one of the cards is a 5, and B be the event that both cards are the same suit. Then $P(A) = \frac{4\times 48}{\binom{52}{2}}$, $P(B) = \frac{4\times\binom{13}{2}}{\binom{52}{2}}$, and

$P(A \cap B) = \frac{4 \times 12}{\binom{52}{2}}$, so we need to determine whether $P(A \cap B) = P(A)P(B)$, i.e. whether $\frac{4 \times 12}{\binom{52}{2}} = \frac{4 \times 48}{\binom{52}{2}} \times \frac{4 \times \binom{13}{2}}{\binom{52}{2}}$. Multiplying both sides by $\binom{52}{2}^2$, this is equivalent to determining whether $4 \times 12 \times \binom{52}{2} = 4 \times 48 \times 4 \times \binom{13}{2}$. Dividing through by 24, the left hand side is 52×51, while the right-hand side is $16 \times 13 \times 12 = 52 \times 48$, so they are not equal.

5. *Suppose the number of children in a family is a Poisson random variable with parameter 2.*

   (a) What is the probability that a family has at least one child?

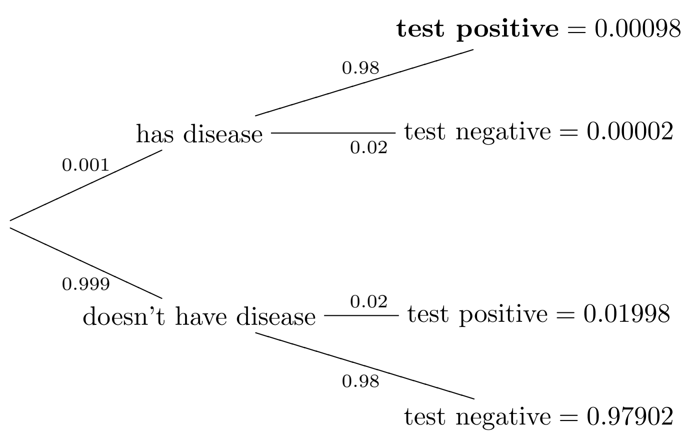
   Let N be the number of children in a family. Then $P(N \geqslant 1) = 1 - P(N = 0) = 1 - e^{-2}\frac{1}{1} = 0.865$.

   (b) Given that a family has at least one child, what is the probability that it also has a second?

   $P(N = 1) = e^{-2}\frac{2}{1} = 0.271$, so $P(N \geqslant 2) = 0.865 - 0.271 = 0.594$. The conditional probability of having a second child given that a family has one child is therefore given by $\frac{0.594}{0.865} = 0.687$.

6. *A patient is given a routine test for a rare disease. The disease affects 1 person in 1000. The test is 98% accurate, so there is only a 2% chance of giving the wrong result. The test result is positive (i.e. indicates the patient has the disease).*

   (a) What is the probability that the patient actually has the disease?

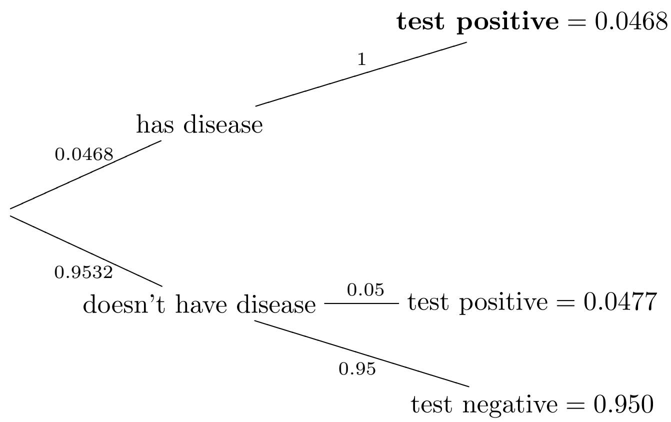


   So the total probability of a positive test result is $0.01998 + 0.00098 = 0.02096$, and the probability of having the disease given a positive test result is $\frac{0.00098}{0.02096} = 0.0468$.

   (b) The doctor sends the patient for further tests. The further test is 95% accurate if the patient does not have the disease (conditional on the first test result being positive), and 100% accurate if the patient has the disease.

If this test result is also positive, what is the probability that the patient has the disease?

Conditional on the first test being positive we have:



So the total probability of a second positive test given that the first test is positive is $0.0468 + 0.0477 = 0.944$, and the conditional probability that the patient has the disease given that the second test is positive is $\frac{0.468}{0.944} = 0.495$.

7. *The lifetime (in years) of a washing machine from a particular company is an exponential random variable with parameter $\frac{1}{5}$. The company offers to replace any machine which breaks within 6 months.*

   (a) What proportion of machines do they need to replace.

   The probabaility that they need to replace a machine is the probability that the lifetime is less than 6 months or 0.5 years. That is $1 - e^{-\frac{0.5}{5}} = 0.0952$, so the company needs to replace 9.52% of its machines.

   (b) What is the expected time until a customer needs to buy a new machine? [assume replacement machines are only replaced if they break within 6 months of the original purchase.]

   The expected time is 6 months, plus the expected lifetime of the machine that the customer has after 6 months. By the memoryless property of the exponential distribution, the remaining lifetime of the machine that the customer has after 6 months is exponentially distributed with parameter $\frac{1}{5}$, so the expected time until the customer must buy a new machine is 5 years 6 months.

8. *An insurance company sells 50,000 insurance policies. The probability of a claim for each policy is $\frac{1}{500}$, and whether each policy leads to a claim is independent of other policies.*

   (a) What is the distribution of the number of claims made?

   The number of claims made is a binomial distribution with $n = 50,000$ and $p = \frac{1}{500}$.

(b) This distribution can be approximated by a normal distribution with the same expectation and variance. Using this approximation, what is the probability that the company receives over 120 claims?

The mean of the distribution is 100, and the variance is 99.8. This means that more than 120 claims is $\frac{20.5}{\sqrt{99.8}} = 2.05$ standard deviations above the mean [remember the continuity correction, the binomial is above 120 if the normal is above 120.5.], so the probability of receiving more than this is $1 - \Phi(2.05) = 0.020$.

(c) The company wants to restrict its danger of bankrupcy to 0.1%. How many claims does it need to be able to cover in order to achieve this?

To restrict its danger of bankrupcy to less than 0.1%, we need to find x such that $\Phi(x) = 0.999$. This is $x = 3.09$. Now the company will reduce its risk of bankrupcy to 0.1% if the number of claims it can cover is at least 3.09 standard deviations above the mean, i.e. at least $100 + 3.09\sqrt{99.8} = 130.9$, so the company should be able to cover 131 claims.