Math 361 X1 **Homework 2 Solutions** **Spring 2003**

Graded problems: 1(ii); 2(c); 3(b); 6; each worth 3 pts., maximal score is 12 pts.

Problem 1.

Variation on the birthday theme. This problem is motivated by the following letter from a reader published in Marilyn vos Savant's *Parade Magazine* column (see also the Problem Sampler):

> *I started counting from the first person who came into the office. I counted until I found a matching birthday in the group. Then I started a new survey with the next person. In eight surveys, the smallest number of people it took before I found a matching birthday was 12. The largest number was only 54.*

This suggests some natural questions, such as: What is the probability that it takes exactly 12 people to find a matching birthday? What is the probability that it takes less than (or more than) 12? What are the corresponding probabilities with 54 in place of 12? More generally, let k be a given integer between 2 and 365, and solve the following two problems.
(i) What is the probability that it takes **more than** k people to find a matching birthday?
(ii) What is the probability that it takes **exactly** k people to find a matching birthday (i.e., that the k'th person walking into the office is the first one at which a matching birthday occurs).
Remarks: Part (i) is easy, if you approach the problem the right way; in particular, do **not** try to solve this using the result of part (ii). In both cases, give a general formula (involving k) for the probability in question. Numerical answers are not required. It would be easy to obtain numerical data, using a programmable calculator or software such as Mathematica; if you are curious, do that and plot (or present in table form) the probability computed in (ii) as a function of k.

Solution.

(i) This is just a disguised form of the standard birthday type problem. The key here is to realize that it takes more than k people to find a matching birthday, if and only if, *the first k people have all distinct birthdays*. Thus the probability in question is the same as the probability that there are no matching birthdays in a group of k people, which is equal to $365 \cdot 364 \cdots (365-k+1)/365^k$, by the argument used for the usual birthday problem. (Take $\Omega = \{(b_1, \dots, b_k) : b_i = 1, 2, \dots, 365\}$, $A = \{(b_1, \dots, b_k) : b_i = 1, 2, \dots, 365,\ b_i \text{ distinct}\}$, so that $\#(\Omega) = 365^k$, $\#(A) = 365 \cdot 364 \cdots (365-k+1)$ and $P(A) = \#(A)/\#(\Omega) = 365 \cdot 364 \cdots (365-k+1)/365^k$.
(ii) There are two ways to solve this problem.
Method 1: We work with an explicit outcome space Ω, and identify the event in question as a subset A of Ω, similar to the standard birthday problem: We take, as usual in these situations,

$$\Omega = \{(b_1, \dots, b_k) : b_i = 1, 2, \dots, 365\},$$

representing all possible k-tuples of birthdays of a group of k people, and the event

$$A = \{(b_1, \dots, b_k) : b_i = 1, 2, \dots, 365;\ b_i \text{ distinct for } i = 1, 2, \dots, k-1\ , \\ b_k = b_i \text{ for some } i = 1, 2, \dots, k-1\}$$

representing those k-tuples in which the last (i.e., k-th) entry is the first entry matching one of the previous entries. Obviously, $\#(\Omega) = 365^k$, but counting $\#(A)$ is harder, and is the key to the problem. Using the slot method, and counting the number of choices in each of the k available slots from left to right, we obtain

$$\#(A) = 365 \cdot 364 \cdot 363 \cdots (365-k+2) \cdot (k-1) = \frac{365!(k-1)}{(365-k+1)!}$$

where the factor $k-1$ comes from the requirement that the k-th entry should be equal to one of the previous $k-1$ entries. Hence $P(A) = \#(A)/\#(\Omega) = \dfrac{365!(k-1)}{(365-k+1)!365^k}$.

Method 2: We use the result of part (i). Denote by A_k the event considered in part (i), namely that it takes *more than* k people to find a matching birthday, and by B_k the event considered in part (ii), namely that it takes *exactly* k people to find a matching birthday. Then the events A_k and B_k are related by $(*)$ $B_k = A_{k-1} \setminus A_k$. (You should convince yourself that this is the case!) Moreover, A_k is a subset of A_{k-1} (since if it takes more than k people to find a matching birthday, then it certainly takes more than $k-1$ people). Thus, we can compute $P(B_k)$ by the relation $(*)$ and the difference rule as $P(B_k) = P(A_{k-1}) - P(A_k)$ and substitute the formula for $P(A_k)$ computed in (i). The result is the same as that obtained by Method 1.

Problem 2.

Let A and B be independent events with probabilities $P(A) = 0.2$ and $P(B) = 0.3$. Let C denote the event that *at least one* of the events A and B occurs, and let D be the event that *exactly one* of the events A and B occurs.
(a) Find $P(C)$.
(b) Find $P(D)$.
(c) Find $P(A|D)$ and $P(D|A)$.
(d) Determine (with a rigorous justification) whether the events A and D are independent.

Solution.

(a) $C = A \cup B$, so $P(C) = P(A) + P(B) - P(AB)$ by the addition formula. But, by the independence of A and B, $P(AB) = P(A)P(B)$. Hence $P(C) = 0.2 + 0.3 - (0.2)(0.3) = 0.44$.
(b) $D = (A \cup B) - AB$. so $P(D) = P(A \cup B) - P(AB)$ by the difference formula. By (a), $P(A \cup B) = 0.44$ and $P(AB) = (0.2)(0.3) = 0.06$. Hence $P(D) = 0.44 - 0.06 = 0.38$.
(c) By the definition of a conditional probability, $P(A|D) = P(AD)/P(D)$ and $P(D|A) = P(DA)/P(A)$. Now, AD is the event "A occurs, but B does not occur," which is the same as $A \setminus (AB)$. Hence $P(AD) = P(A) - P(AB) = 0.2 - 0.06 = 0.14$, by the difference formula. It follows that $P(A|D) = 0.14/0.38 = 7/19$ and $P(D|A) = 0.14/0.2 = 0.7$.
(d) **Method 1:** A and D are independent if $P(AD) = P(A)P(D)$. In (c) we calculated $P(AD) = 0.14$. On the other hand, by (c), $P(A)P(D) = (0.2)(0.38) = 0.076$ which is different from $P(AD)$. Hence A and D are not independent.
Method 2: An equivalent definition of independence of A and D is that $P(A|D) = P(A)$. Now by (c), $P(A|D) = 7/19$, which is not equal to $P(A)$. Hence A and D are not independent.

Problem 3.

The following are exercises in deriving properties and rules for probabilities from definitions or properties already derived. You may use the definitions and any of the probability rules given in class (Kolmogorov axioms, complement rule, difference formula, inclusion-exclusion, average rule, Bayes' rule, etc.). If you are using one of these rules, say so, and state which rule it is.
(a) Prove that if A and B are independent, then so are A^c and B^c.
(b) Derive a formula for $P(A^c|B^c)$ in terms of $P(A|B)$, $P(A)$, $P(B)$. (You may assume that $0 < P(B) < 1$, so that the conditional probabilities involved are well defined.)
(c) Prove that if $B_1, B_2, \ldots, B_n$ is a partition of B, then, for any A, $P(A|B) = \sum_{i=1}^n P(A|B_i)P(B_i|B)$. (Assume $P(B_i) > 0$ for each i.)

Solution.

(a) Since the sets $A^c \cap B^c$ and $A^c \cap B$ partition the set A^c, we have $P(A^c) = P(A^cB^c) + P(A^cB)$. Similarly, since the sets $A^c \cap B$ and $A \cap B$ partition B, we have $P(B) = P(AB) + P(A^cB)$. finally, since A and B are independent, $P(AB) = P(A)P(B)$. Combining these identities and solving for $P(A^cB^c)$ gives

$$
\begin{aligned}
P(A^cB^c) &= P(A^c) - P(A^cB) = P(A^c) - (P(B) - P(A)P(B)) = P(A^c) - P(B)(1 - P(A)) \\
&= P(A^c) - (1 - P(B^c))P(A^c) = P(A^c)P(B^c),
\end{aligned}
$$

which proves that A^c and B^c are independent.

(b) First, using the rule $P(A^c|B) = 1 - P(A|B)$ (proved in class) we see that $P(A^c|B^c) = 1 - P(A|B^c)$. By the definition of $P(A|B^c)$, we have $P(A|B^c) = P(AB^c)/P(B^c)$. By the complement formula, $P(B^c) = 1 - P(B)$. Also, $P(AB^c) = P(A) - P(AB)$ (proved as in (a) by partitioning A into $A \cap B^c$ and $A \cap B$). Thus,

$$P(A^c|B^c) = 1 - P(A|B^c) = 1 - \frac{P(A) - P(AB)}{1 - P(B)} = 1 - \frac{P(A) - P(A|B)P(B)}{1 - P(B)}.$$

where the last step follows by the multiplication formula. The latter expression involves only $P(A)$, $P(B)$, and $P(A|B)$, and thus is a formula for $P(A^c|B^c)$ of the desired type.

(c) Observe, by drawing a picture that, whenever $B_1, B_2, \ldots, B_n$ is a partition of B, then $A \cap B_1$, $A \cap B_2$, $\ldots$ $A \cap B_n$ is a partition of AB. Thus, by the addition formula for probabilities and the multiplication formula for conditional probabilities

$$P(AB) = \sum_{i=1}^{n} P(AB_i) = \sum_{i=1}^{n} P(A|B_i)P(B_i).$$

Now, $B_i = B_i \cap B$ (since B_i is a subset of B), and so $P(B_i) = P(B_iB) = P(B_i|B)P(B)$ by another application of the multiplication formula. Substituting this expression for $P(B_i)$ into the above formula and dividing by $P(B)$ gives the asserted formula for $P(A|B)$.

Problem 4.

[1.3:11] **Inclusion-Exclusion formula for 3 events.** Writing $A \cup B \cup C = (A \cup B) \cup C$ and applying the inclusion-exclusion formula for the union of two events, derive an analogous formula for the probability of the union of three events, $P(A \cup B \cup C)$. (The formula is stated in Problem 11 of 1.3, but try to come up with this formula on your own. Of course, stating the right formula will not earn you any credit; the key is to formally derive the formula.)

Solution.

We first apply the inclusion exclusion formula with the sets $A \cup B$ and C, to get

$$(*) \qquad P(A \cup B \cup C) = P(A \cup B) + P(C) - P((A \cup B) \cap C)$$

The term $P(A \cup B)$ equals $P(A) + P(B) - P(AB)$, by inclusion/exclusion with sets A and B. To evaluate the last term in $(*)$, we use the fact that $(A \cup B) \cap C = AC \cup BC$ (which can be seen by drawing Venn diagrams, and apply (again!) inclusion/exclusion, to get $P((A \cup B) \cap C) = P(AC \cup BC) = P(AC) + P(BC) - P((AC)(BC))$. Since $(AC)(BC) = (A \cap C) \cap (B \cap C) = A \cap B \cap C = ABC$, this simplifies to $P(AC) + P(BC) - P(ABC)$. Substituting these formulas into $(*)$ and regrouping terms gives the desired formula, namely

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC).$$

Problem 5.

A drug test has 1 % false positives (i.e., 1 % of those not taking drugs test positive), and 5 % false negatives (i.e., 5 % of those taking drugs test negative). Suppose that 2 % of those tested are taking drugs. Using Bayes' rule determine the probability that somebody who tests positive is actually taking drugs. **Make sure to introduce appropriate notation for all relevant events and express the given data and the probability asked for in terms of this notation, as in the examples worked out in class. The same applies to the next problem.**

Solution.

(a) The average rule states that $P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$ **if the B_i's form a partition of Ω.**
(b) Bayes' rule states that $P(B_i|A) = P(A|B_i)P(B_i)/P(A)$ (with $P(A)$ given by the average rule) **if the B_i's form a partition of Ω.**
(c) Let D and denote the event that a person (randomly selected from among those tested) takes drugs, and let P denote the event that a person tests positive. (Alternatively, one could define D (resp. P) as the *set* of people taking drugs (resp. testing positive) among those tested.) We are given the following data:
(i) $P(D) = 0.02$ (since 2% of those tested take drugs);
(ii) $P(P|D^c) = 0.01$ (since 1% of those not taking drugs test positive);
(iii) $P(P^c|D) = 0.05$ (since 5% of those taking drugs test negative);
We need to compute $P(D|P)$. By the average and Bayes' rules with D and D^c as the sets B_i (which obviously form a partion of Ω), we have $P(D|P) = P(P|D)P(D)/P(P)$, and $P(P) = P(P|D)P(D)+P(P|D^c)P(D^c)$. Now, by (iii) and the complement rule for conditional probabilities $P(P|D) = 1 - P(P^c|D) = 1 - 0.05 = 0.95$, and by (i), $P(D^c) = 1 - P(D) = 0.98$. Substituting these values and the values for $P(D)$ and $P(P|D^c)$ from (i) and (ii) into the above formulae, we obtain

$$P(D|P) = \frac{0.95 \cdot 0.02}{0.95 \cdot 0.02 + 0.01 \cdot 0.98} \ (= 0.659\ldots)$$

Problem 6.

[1.5:4] A digital communications system consists of a transmitter and a receiver. During each short transmission interval the transmitter sends either a zero or a one. At the end of each interval, the receiver makes its best guess at what was transmitted. Consider the events: T_0 = "Transmitter sends 0"; R_0= "Receiver receives 0"; T_1 = "Transmitter sends 1"; R_1= "Receiver receives 1". Assume that $P(R_0|T_0) = 0.99$, $P(R_1|T_1) = 0.98$, and $P(T_1) = 0.5$. Find (a) the probability of a transmission error given R_1; (b) the overall probability of a transmission error.

Solution.

Relevant events: T_0, T_1, R_0, R_1 are already defined in the problem. Note that $T_1 = T_0^c$, $R_1 = R_0^c$.
Given data: $P(R_0|T_0) = 0.99$, $P(R_1|T_1) = 0.98$, and $P(T_1) = 0.5$.
To compute: The probabilities asked for in (a) and (b) involve the concept of a "transmission error". A transmission error occurs if the digit received is not equal to the digit that was sent. Thus, if a 1 is received, a transmission error occurs if and only if a 0 was sent, so the probability asked for in (a) (i.e., the probability of a transmission error given R_1) is the same as $P(T_0|R_1)$. The probability asked for in (b), the overall probability of a transmission error, is $P(R_1T_0 \cup R_0T_1)$, which is the same as $P(R_1T_0) + P(R_0T_1)$ since the two sets R_1T_0 and T_0T_1 are disjoint.
Computations:
(a) By Bayes' rule, $P(T_0|R_1) = P(R_1|T_0)P(T_0)/P(R_1)$. Now, by the complement rule for conditional probabilities,

$$P(R_1|T_0) = P(R_0^c|T_0) = 1 - P(R_0|T_0) = 1 - 0.99 = 0.01.$$

Also, by the average rule,

$$P(R_1) = P(R_1|T_0)P(T_0) + P(R_1|T_1)P(T_1) = 0.01 \times 0.5 + 0.98 \times 0.5 = 0.495.$$

Therefore, $P(T_0|R_1) = 0.01 \times 0.5/0.495 = 1/99$.
(b) By the multiplication rule,

$$P(R_1T_0) = P(R_1|T_0)P(T_0) = 0.01 \times 0.5 = 0.005$$

and

$$P(R_0T_1) = P(R_0|T_1)P(T_1) = (1 - P(R_1|T_1))P(T_1) = (1 - 0.98) \times 0.5 = 0.01.$$

Hence,

$$P(R_1T_0 \cup R_0T_1) = P(R_1T_0) + P(R_0T_1) = 0.005 + 0.01 = 0.015.$$