Problem Set 3

Student handout

Counting, collisions, and derangements

100 points + 10 optional challenge points

Name: ____________________ **Due:** ____________

Instructions. Identify the outcomes, their probabilities, and the event being counted. State whether order matters and repetition is allowed. Give exact answers except where decimal approximations are requested. A calculator or short program may be used for Problems 3, 6, and 7. Throughout, D_n is the number of derangements of n objects (permutations with no fixed point), and $p_n = D_n/n!$; you may use the formula $p_n = \sum_{k=0}^{n}(-1)^k/k!$ from class. By convention $D_0 = 1$.

I. Discrete probability spaces

1. A space with infinitely many outcomes [10 points]

Flip a fair coin until the first head appears and record the number K of flips used. Take $\Omega = \{1, 2, 3, \dots\}$, let $\mathcal{F} = 2^{\Omega}$, and give outcome k the mass 2^{-k}.

(a) Explain where the mass 2^{-k} comes from, and verify that the masses sum to one. What does this say about the possibility that no head ever appears? (3)

(b) Find $\mathbb{P}(K \text{ is even})$ and $\mathbb{P}(K \text{ is divisible by } 3)$. Find $\mathbb{P}(K > n)$ in two ways: by summing masses, and by describing the event in terms of the first n flips. (4)

(c) Show that there is *no* probability measure on $\Omega = \{1, 2, 3, \dots\}$ giving every outcome the same mass. Which axiom does the argument use that a finite space never needs? (3)

2. The larger of two dice [10 points]

Roll two fair dice, one red and one green, and let M be the larger of the two values shown (the common value if they agree).

(a) Give Ω and the mass of each outcome. Show that $\mathbb{P}(M \le k) = k^2/36$ for $k = 1, \dots, 6$, and deduce that $\mathbb{P}(M = k) = (2k-1)/36$. (4)

(b) Check that your six values of $\mathbb{P}(M = k)$ sum to one. Which familiar identity about odd numbers have you verified, and how does a 6×6 grid of outcomes explain it? (3)

(c) Recording only M produces a new probability space on $\{1, \dots, 6\}$. Find $\mathbb{P}(M \text{ is even})$ in that space. Compare with the probability that a single fair die is even, and explain the direction of the difference. (3)

II. Permutations and combinations

3. Poker hands [12 points]

A hand is five cards dealt from a well-shuffled standard deck: 13 ranks in each of 4 suits. Give each probability as an exact fraction and as a decimal to four places.

(a) Describe a probability space in which a hand is an *unordered* set of cards, and count its outcomes. (2)

(b) Find the probability of a *full house*: three cards of one rank and two of another. (3)

(c) Find the probability of *exactly one pair*: two cards of one rank and three cards of three other, different ranks. (4)

(d) For *two pairs* (two cards of one rank, two of another, and a fifth card of a third rank) a student writes $13\binom{4}{2} \cdot 12\binom{4}{2} \cdot 44$. Explain the error, give the correct count, and explain why the same "first rank, then second rank" scheme was *not* an error in part (b). (3)

4. Random arrangements of a word with repeats [12 points]

The eleven letter tiles of `MISSISSIPPI` (one M, four I's, four S's, two P's) are placed in a uniformly random order.

(a) How many distinguishable arrangements are there? Explain why all of them are equally likely, even though the underlying experiment permutes eleven physical tiles. (3)

(b) Find the probability that no two S's are adjacent. (Arrange the other letters first, then choose gaps.) (4)

(c) Find the probability that the first and last letters are the same, in two ways: by counting arrangements, and by looking only at the two physical tiles that land in the first and last positions. (5)

5. Committees and Vandermonde's identity [10 points]

A five-person committee is chosen uniformly at random from a club of 6 women and 9 men.

(a) Find the probability that the committee has exactly two women. (3)

(b) Find the probability that it includes at least one woman and at least one man. (3)

(c) Prove *Vandermonde's identity* $\sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j} = \binom{m+n}{k}$ by a counting argument (with $\binom{a}{b} = 0$ when $b > a$). What does it say about the probabilities in part (a), taken over all possible numbers of women? (4)

III. Birthday collisions

6. Two honest bounds for the birthday problem [12 points]

Each of n customers independently chooses a four-digit PIN uniformly from the $d = 10^4$ possibilities. Let N_n be the event that all n PINs are different, so that $\mathbb{P}(N_n) = \prod_{k=1}^{n-1}(1 - k/d)$. In class we *approximated* $\ln(1 - x) \approx -x$. Here you will replace the approximation by inequalities.

(a) Prove that $1 - x \leq e^{-x}$ for every real x. (3)

(b) Deduce that $\mathbb{P}(\text{some two PINs agree}) \geq 1 - \exp(-n(n-1)/(2d))$: the approximation from class is a rigorous *lower* bound for the collision probability. (3)

(c) Let E_{ij} be the event that customers i and j choose the same PIN. Using $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$, show that $\mathbb{P}(\text{some two PINs agree}) \leq \binom{n}{2}/d$. (3)

(d) Use (c) to show that a collision is less likely than not when $n \leq 100$, and (b) to show that it is more likely than not when $n \geq 119$ (use $\ln 2 \approx 0.693147$). Then compute the exact collision probabilities for $n = 118$ and $n = 119$ to four decimal places. (3)

7. What kind of collision? [12 points]

(a) Four labeled people have independent birth *months*, uniform on 12 months. Classify the 12^4 outcomes by pattern: all four different; exactly one shared pair (the other two different from it and from each other); two different shared pairs; three alike and one different; all four alike. Count each pattern, and check that the counts total 12^4. (6)

(b) Now n people have independent birthdays, uniform on d days, with $2 \leq n \leq d + 1$. Show that the probability that exactly one pair shares a birthday and all other birthdays are different from it and from each other is $\binom{n}{2}\, d(d-1)\cdots(d-n+2)/d^n$. (3)

(c) For $n = 23$ and $d = 365$, compute to four decimal places the probabilities of no shared birthday, of the pattern in (b), and of any more complicated coincidence. (3)

IV. Derangements

8. A recurrence for derangements [10 points]

(a) Prove that $D_n = (n-1)\,(D_{n-1} + D_{n-2})$ for $n \geq 2$, by a counting argument. Suggested route: in a derangement guest 1 receives some hat $j \neq 1$; separate the cases in which guest j does, or does not, receive hat 1. (5)

(b) Starting from $D_0 = 1$ and $D_1 = 0$, compute $D_2, \ldots, D_6$, and check D_6 against the formula for p_6. (2)

(c) The alternating-series estimate gives $|p_n - 1/e| < 1/(n+1)!$. Deduce that for $n \geq 2$, D_n is the integer nearest to $n!/e$, and confirm this for $n = 6$. (3)

9. How many guests get the right hat? [12 points]

The hats of n guests are returned as a uniformly random permutation.

(a) Show that the probability that *exactly* k guests receive their own hats is $\binom{n}{k} D_{n-k}/n! = p_{n-k}/k!$. (3)

(b) For $n = 5$, tabulate the number of permutations with exactly k correct hats for $k = 0, 1, \ldots, 5$, and check the total. (3)

(c) Explain in words why exactly $n - 1$ correct hats is impossible, and how the formula in (a) knows this. For fixed k, find the limit of the probability in (a) as $n \to \infty$. (3)

(d) Show that the total number of correct hats, summed over all $n!$ permutations, is exactly $n!$, by counting pairs (permutation, guest with the correct hat) in a second way. Verify this with your table from (b). What is the average number of correct hats, and how does it depend on n? (3)

Optional challenge: Secret Santa [10 extra points]

In a Secret Santa, $n \geq 3$ people draw names from a hat, and the draw is repeated from scratch until nobody has drawn their own name.

(a) Explain why the final assignment is uniformly distributed over the D_n derangements. For $n = 5$, find the probability that more than three draws are needed. (3)

(b) Find the probability that persons 1 and 2 draw each other, in terms of derangement numbers, and evaluate it for $n = 6$. (3)

(c) Find the probability that person 1 draws person 2, for general n. Then use (b) to find the conditional probability that 2 drew 1, given that 1 drew 2, for $n = 6$. Is reciprocation more or less likely than the unconditional probability that 2 draws 1? (4)