Probability & Statistics
Fall 2022
Midterm
2022/11/06
Time Limit: 100 Minutes

Name (Print): ______________________

Advisor Name ______________________

This exam contains 9 pages (including this cover page) and 8 problems. Check to see if any pages are missing. Enter all requested information on the top of this page, and put your initials on the top of every page, in case the pages become separated.

Try to answer as many problems as you can. The following rules apply:

- **Mysterious or unsupported answers will not receive full credit**. A correct answer, unsupported by calculations, explanation, or algebraic work will receive no credit; an incorrect answer supported by substantially correct calculations and explanations might still receive partial credit.

Do not write in the table to the right.

Problem	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	20	
7	10	
8	20	
Total:	100	

1. (10 points) How many number of distinct **positive** integer-valued vectors $(x_1, x_2, \ldots, x_5)$ satisfying the equation and inequalities

$x_1 \geq 8$

$$x_1 + x_2 + \cdots + x_5 = 66$$
$$x_1 > 7, x_2 > 3, x_3 > 3, x_4 \geq 5, x_5 > 4.$$

> 4

$x_1 - 7$

(-4)

2. (10 points) A family has two children. Find the probability that both children are girls, given that at least one of the two is a girl who was born in winter. Assume that the four seasons are equally likely and that gender is independent of season.

3. (10 points) A sequence of $n \geq 1$ independent trials is performed, where each trial ends in "success" or "failure" (but not both). Let p_i be the probability of success in the ith trial, $q_i = 1 - p_i$, and $b_i = q_i - 1/2$, for $i = 1, 2, \ldots, n$. Let A_n be the event that the number of successful trials is even.

   (a) (5 points) Show that for $n = 2$, $P(A_2) = 1/2 + 2b_1 b_2$.

   (b) (5 points) Find $P(A_n), n \geq 3$ by induction.

4. (10 points) There are four biased six-sided dices. For each biased dice, the corresponding roll number is 1 with probability 1/3, 2 with probability 1/6, 3 with probability 1/6, 4 with probability 1/9, 5 with probability 1/9, 6 with probability 1/9. When four biased six-sided dice are rolled, what is the probability that the sum of the total numbers will be 16?

$$E[t^X] = \sum_{i=1}^{6} P(X=i)\,t^i$$

$$= \frac{1}{3}t + \frac{1}{6}t^2 + \frac{1}{6}t^3 + \frac{1}{9}t^4 + \frac{1}{9}t^5 + \frac{1}{9}t^6$$

$$= \frac{1}{18}(6t + 3t^2 + 3t^3 + 2t^4 + 2t^5 + 2t^6)$$

$$E[t^X] = (\quad)^4$$

$$P(X=16) = \_\, t^{16}$$

$\downarrow$

找 t^{16} 系数

5. (10 points) Let a random variable X be Hypergeometric with parameters w, b, n. The PMF of X is

$$P(X = k) = \frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}},$$

for integers k satisfying $0 \leq k \leq w$ and $0 \leq n - k \leq b$, and $P(X = k) = 0$ otherwise.

(a) (5 points) Find $E\left[\binom{X}{2}\right]$

(b) (5 points) Use the result of (a) to find the variance of X.

6. (20 points) Suppose a coin with probability p for heads is tossed repeatedly, and we obtain a sequence of H and T (H denotes Head and T denotes Tail). Let N denote the number of toss to observe the first occurrence of the pattern “HT”.
   (a) (10 points) Find the recursive equation for the probability of event “N=n”.
   (b) (5 points) Find $E(N)$.
   (c) (5 points) Find $Var(N)$.

HHH.

设 $p=\frac{1}{2}$

HTH

“HT” $k \geq 3$ $\begin{cases} P(S_1=T)\,P(N=k-1) \\ P(S_1=H) \end{cases}$

$\Rightarrow \quad P_k = \frac{1}{2}P_{k-1} + \frac{1}{2^k}$

$2^k P_k = 2^{k-1}P_{k-1} + 1$

$(k \geq 3)$ 2^kP_k 等差数列，$2^2P_2 = 4 \times \frac{1}{4} = 1$ $2^3P_3 = 2^2P_2 + 1 = 2$

$2^kP_k = 2 + (k-3)\times 1 = k-1$

$P_k = \frac{k-1}{2^k}$

$E = \sum_{k=2}^{\infty} kP_k = \sum \frac{k(k-1)}{2^k}$

$g(x) = \sum_{n=1}^{\infty} k(k-1)x^k. \quad |x|<1$

$f(x) = \int_0^x \frac{g(x)}{x}dx = \int_0^x k(k-1)\cdot x^{k-1}dx = \sum_{n=1}^{\infty}(k-1)x^k$

$m(x) = \int_0^x \frac{f(x)}{x^2}dx = \int_0^x (k-1)x^{k-2}dx = \sum_{n=1}^{\infty} x^{k-1} = \frac{1}{1-x}$

$\frac{f(x)}{x^2} = \left(\frac{1}{1-x}\right)' = \frac{1}{(x-1)^2} \qquad f(x) = \frac{x^2}{(x-1)^2}$

$\frac{g(x)}{x} = \left(\frac{x^2}{(x-1)^2}\right)' = -\frac{2x^2}{(x-1)^3} + \frac{2x}{(x-1)^2} \qquad g(x) = -\frac{2x^3}{(x-1)^3} + \frac{2x^2}{(x-1)^2}$

$x = \frac{1}{2}. \quad g(x) = 4$

7. (10 points) Given two probability distributions μ and ν on the same sample space Ω, the total variation distance between μ and ν is defined as follows:

$$\| \mu - \nu \|_{TV} = \boxed{\max_{A \subset \Omega}} |\mu(A) - \nu(A)|.$$

where

$$\mu(A) = \sum_{x \in A} \mu(x), \quad \nu(A) = \sum_{x \in A} \nu(x).$$

(a) (5 points) Let $B = \{x : \mu(x) \geq \nu(x)\}$, show that

$$\| \mu - \nu \|_{TV} = \boxed{\sum_{x \in B} [\mu(x) - \nu(x)].}$$

(b) (5 points) Further to show that

$$\| \mu - \nu \|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

$D = B + B_2$

$B_2 \subset \Omega \setminus B$

$C = B \setminus B_1$

$B_1 \subset B$

$|\mu(B_1) - \nu(B_1)|$

$\leq |\mu(B) - \nu(B)|$

$\sum_{x \in C} [\mu(x) - \nu(x)]$

$= \boxed{\sum_{x \in B} [\mu(x) - \nu(x)]} - \sum_{x \in B_1} [\mu(x) - \nu(x)] \leq$

$\because B_1 \subset B$

$\boxed{\sum_{x \in B} [\mu(x) - \nu(x)]} + \sum_{x \in B_2} [\mu(x) - \nu(x)]$

$\because x \in B_1, \mu(x) \geq \nu(x)$ $\quad x \in B_2, \mu(x) < \nu(x)$

$\therefore \sum_{x \in B_1} [\mu(x) - \nu(x)] \geq 0$ $\quad \sum_{x \in B_2} [\mu(x) - \nu(x)] < 0$

$$\| \mu - \nu \|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

$$= \frac{1}{2} \sum_{x \in B} [\mu(x) - \nu(x)]$$

$$- \frac{1}{2} \sum_{x \in \Omega \setminus B} [\mu(x) - \nu(x)]$$

$\mu(\Omega) = 1$

$\mu(B) + \mu(\Omega \setminus B) = 1$

$\mu(\Omega \setminus B) = 1 - \mu(B)$

$$= \frac{1}{2} \sum_{x \in B} [\mu(x) - \nu(x)]$$

$$- \frac{1}{2} \sum_{x \in B} [(1 - \mu(x)) - (1 - \nu(x))]$$

$$= \sum_{x \in B} [\mu(x) - \nu(x)]$$

8. (20 points) Show the following Maximum-Minimum Identity. For arbitrary random variables $x_i, i = 1, \ldots, n$,

$$\max_i x_i = \sum_i x_i - \sum_{i<j} \min(x_i, x_j) + \sum_{i<j<k} \min(x_i, x_j, x_k) + \ldots + (-1)^{n+1} \min(x_1, \ldots, x_n).$$

设$x_1 \le x_2 \le \cdots \le x_n$.

考虑这样一系列集合: $\begin{cases} i \le j \\ A_i \subseteq A_j \\ A_i\text{中有}x_i\text{个元素} \end{cases}$

除max外的每个数加减次数都一样

max只出现在第一项

$n=k$　$\max_i x_i = \underbrace{\text{---}}_{n=k}$

$n=k+1$　$\max\left(\max_j x_j, x_{k+1}\right)$