NTHU MATH 2810 **Midterm Examination - A** Nov 22, 2016

Note. There are 6 problems in total. **For problems 2 to 6, to ensure consideration for partial/full scores, write down necessary intermediate steps.** Correct answers with inadequate or no intermediate steps may result in zero credit.

Some useful formula.

- The probability mass function (pmf) of a binomial(n, p) distribution is

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad \text{for } x = 0, 1, \ldots, n.$$

  Its mean and variance are np and $np(1-p)$, respectively. When $n = 1$, it is called a Bernoulli(p) distribution.

- The probability mass function of a negative binomial(r, p) distribution is

$$p(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad \text{for } x = r, r+1, r+2, \ldots.$$

  Its mean and variance are $\frac{r}{p}$ and $\frac{r(1-p)}{p^2}$, respectively. When $r = 1$, it is called a geometric(p) distribution.

- The probability mass function of a Poisson(λ) distribution is

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad \text{for } x = 0, 1, 2, \ldots.$$

  Its mean and variance are λ.

- The probability mass function of a hypergeometric(n, N, R) distribution is

$$p(x) = \frac{\binom{R}{x}\binom{N-R}{n-x}}{\binom{N}{n}}, \quad \text{for } x = 0, 1, 2, \ldots, n,$$

  where $\binom{s}{t} \equiv 0$ if $s < t$. Its mean and variance are $\frac{nR}{N}$ and $\frac{nR(N-R)(N-n)}{N^2(N-1)}$, respectively.

---

1. (*24 points in total*; for each, 3 points if correct; 0 points if blank; -1.5 **points if wrong**) For the following statements, please answer true or false.

   (a) Suppose that Ω is a finite sample space. For a subset A of Ω, the probability of A is $\#A/\#\Omega$, where $\#$ calculates the number of elements in a set.

   (b) For a random variable X, let $S = \{x \in \mathbb{R} : \ P(X = x) = 0\}$. Then, S can be a countable set.

   (c) If the cumulative distribution function (cdf) of a random variable X is $F_X(x)$, the cdf of $Y = e^X$ (i.e., $X = \ln Y$), denoted by $F_Y(y)$, is $F_X(\ln(y))$ for $y > 0$ and zero otherwise.

   (d) Suppose that $F_X(x)$ is a cdf. Then, for any positive integer n, there always exists a random variable Y with cdf $F_Y(y) = (F_X(y))^n$.

(e) Suppose that X and Y are two random variables mapping from Ω to $\mathbb{R}$ and have same distribution. Because it is possible that $X(\omega) \neq Y(\omega)$ for $\omega \in \Omega$, it is not necessarily true that $E(X) = E(Y)$.

(f) Suppose that X and Y are random variables with same mean and variance. Then, X and Y must have same distribution.

(g) Suppose that X is a random variable with binomial(n, 1/2) distribution. Let $Y = n - X$. Then, Y has the same distribution as X.

(h) For two events A and B with $0 < P(A), P(B) < 1$, if $P(A|B) > P(A)$, then $P(A|B^c) < P(A)$.

2. Let B be the event that a queen carries the gene for hemophilia, and suppose that the probability of B is 0.6. If the queen is a carrier, then each prince has a 50-50 chance of having hemophilia. If the queen is not a carrier, then each prince has zero chance of having hemophilia. Suppose that the queen has had four princes. Let E_1 be the event that the first three princes do not have hemophilia, and E_2 be the event that the fourth prince has hemophilia.

   (a) (*3 points*) Find the probability $P(B|E_1)$.

   (b) (*4 points*) Find the probability $P(E_2|E_1)$.

   (c) (*4 points*) Examine whether E_1 and E_2 are independent.

   (d) (*3 points*) Are E_1 and E_2 assumed conditionally independent given B or B^c in your previous calculation of probabilities? Explain.

3. The suicide rate in a certain state is 1 suicide per 100,000 inhabitants per month.

   (a) (*4 points*) Let X be the number of suicides in a city of 400,000 inhabitants within this state in a given month. Find the probability $p = P(X \geq 8)$.

   (b) (*4 points*) Let Y be the number of the months that will have 8 or more suicides during a year. Find the probability of $\{Y \geq 2\}$ and express it as a function of p.

   (c) (*4 points*) Count the present month as month number 1. Let Z be the number of months from now until the first month to have 8 or more suicides. Find the probability of $\{Z \geq 6\}$ and express it as a function of p.

   (d) (*2 points*) What assumptions are you making in the calculation of these probabilities?

4. Let N be the number of children in a randomly chosen family and let B be the number of boys in the family. Suppose that the random variable N can take any of the values $\{0, 1, 2, \ldots\}$, and for $n = 1, 2, 3, \ldots$, $P(N = n) = \alpha p^n$, where $0 < p < 1$ and $\alpha < (1-p)/p$.

   (a) (*3 points*) Find the probability mass function of N.

   (b) (*4 points*) Find the cumulative distribution function of N.

   (c) (*3 points*) If each child is equally likely to be a boy or a girl (independently of each other), identify the probability $P(B = b|N = n)$ for $b = 0, 1, 2, \ldots, n$. What is the name of the distribution of B given $N = n$?

   (d) (*4 points*) What proportion of families consists of b boys (and any number of girls) for $b = 0, 1, 2, \ldots$?

   (e) (*4 points*) Among all the families with b boys, $b \in \{0, 1, 2, \ldots\}$, what proportion of families have n children for $n = b, b + 1, b + 2, \ldots$?

5. An urn has R red and $N - R$ white balls, where $R \le N$. Balls are randomly withdrawn, *without* replacement, until a total of r, $1 \le r \le R$, *red* balls have been withdrawn. Let the random variable X_r be the *total* number of balls that are withdrawn, and let the random variable Y_r be the number of *white* balls that are withdrawn. The distribution of X_r is called *negative hypergeometric* with parameters r, N, and R.

   (a) (*2 points*) Explain how X_r differs from a negative binomial random variable.

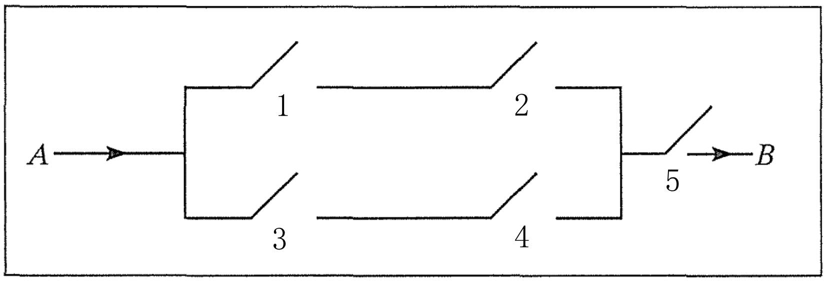
   (b) (*5 points*) Denote the probability mass function of X_r by $p_r(x)$, $x = r, r+1, \ldots, N$. Find $p_r(x)$. [**Hint**. Let Z_{x-1} be a random variable with hypergeometric($x - 1$, N, R) distribution. Identify the relationship between the probabilities $P(X_r = x)$ and $P(Z_{x-1} = r - 1)$, and use it to find $p_r(x)$.]

   (c) (*5 points*) Show that
   $$E(X_r) = r \times \frac{N+1}{R+1}.$$
   Also, find $E(Y_r)$. [**Hint**. $Y_r = X_r - r$.]

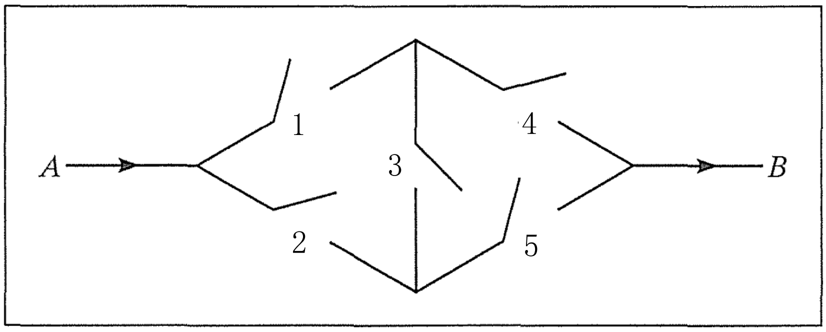
   (d) (*7 points*) Show that
   $$Var(X_r) = r \times \frac{(N-R)(N+1)(R-r+1)}{(R+1)^2(R+2)}.$$
   Also, find $Var(Y_r)$. [**Hint**. $Y_r = X_r - r$.]

6. In any of the figures given below, for $i = 1, 2, 3, 4, 5$, let T_i be the event that the ith relay closes (i.e., works properly), and let the probability of the closing of the ith relay in the circuits be $p_i = P(T_i)$. Let E be the event that a current can flow between the points A and B in the figure. Express E in terms of $T_1, \ldots, T_5$. Also, find the probability of E under the assumption that all relays function independently.

   (a) (*5 points*)



   (b) (*6 points*)



   [**Hint**. For the calculation of $P(E)$, condition on whether relay 3 closes.]