Midterm

Mark all correct answers in each of the following questions.

1. The overly happy participants of a party pick, upon leaving the party, a coat and an umbrella at random. Assume the number of participants is n. Let X be the number of items, out of $2n$, picked by their rightful owners, and let Y be the number of people who pick an umbrella and a coat belonging to the same person. (For example, if $n = 4$, person #1 picks coat #1 and umbrella #4, person #2 picks coat #3 and umbrella #3, person #3 picks coat #2 and umbrella #1, and person #4 picks coat #4 and umbrella #2, then $X = 2$ and $Y = 1$.)

   (a) X is binomially distributed.

   (b) $P(X = 0) \underset{n\to\infty}{\longrightarrow} e^{-1/2}$.

   (c) $E(X) = 2$.

   (d) $P(Y = 0) \underset{n\to\infty}{\longrightarrow} e^{-1}$.

   (e) $E(Y)$ is strictly decreasing as a function of n.

   (f) $P(X = 0|Y = n) \underset{n\to\infty}{\longrightarrow} e^{-1}$.

   (g) The probability for at least one person to pick both his coat and his umbrella is:

$$\frac{1}{n} - \frac{1}{n(n-1)} + \frac{1}{n(n-1)(n-2)} - \ldots + \frac{(-1)^{n-1}}{n!} .$$

2. An urn contains two white balls and four black balls. Three balls are drawn from the urn without replacement. If one of these balls is white and the other two are black – the trial is finished. Otherwise, the balls are returned to the urn, and the trial is repeated. This continues until, for the first time, exactly one white and two black balls are drawn. Let X denote the total number of trials, Y – the number of trials in which all three balls drawn are black, and Z the number of trials in which two white and one black balls are drawn. (For example, if in the first and the second trials three black balls are drawn, in the third trial two white and one black balls are drawn, and in the fourth trial one white and two black balls are drawn, then $X = 4,\ Y = 2$ and $Z = 1$.)

   (a) X is hypergeometrically distributed.

   (b) $P(Y \geq 1) = \frac{1}{5}$.

   (c) The events $\{X \geq 1000\}$ and $\{Y \geq 500\}$ are independent.

   (d) $E(Y) = E(Z)$.

   (e) The distribution function of the random variable $Y + Z$ has infinitely many discontinuity points.

   (f) $P(Z = 1|Y = 1) = \frac{16}{125}$.

3. Consider the trial of tossing a coin 13 times. Suppose this trial is repeated $10,000$ times (so that the coin is tossed $130,000$ times in all). Let X be the number of trials (out of $10,000$) in which all 13 tosses result in a head, and Y the analogous random variable with tails instead of heads.

   (a) X is geometrically distributed.

   (b) $E(X^2) > E(X)$.

   (c) X has a finite expectation, but there exists a random variable Z, which is a function of X (that is, $Z = h(X)$ for an appropriate function $h : \mathbf{R} \longrightarrow \mathbf{R}$), such that Z does not have a finite expectation.

   (d) The distribution functions of X and Y are identical, and in particular X and Y are identical.

(e) Out of the 10,000 trials, it is most probable that each number of heads between 0 and 13 will be obtained at least once. That is, at least once all 13 tosses will result in a head, at least once we shall obtain 12 heads and a single tail, ..., at least once we shall obtain 13 tails.

(f) Suppose that, instead of repeating the trial 10,000 times, we repeat it until we obtain for the third time 13 heads in a trial. Let U be the number of times we repeat the trial. Then U is geometrically distributed.

4. A die is rolled until the outcome is the same twice in a row. Let X_i be the number of times the outcome is i, $1 \le i \le 6$. Let X be the number of times it is rolled and Y the sum of all outcomes. (For example, if the outcomes are 2, 5, 1, 2, 6, 1, 1, then $X_1 = 3,\ X_2 = 2,\ X_3 = X_4 = 0,\ X_5 = X_6 = 1,\ X = 7,\ Y = 18$.)

   (a) $P(X_6 = 0) = 1/2$.

   (b) $P(Y = 6) = 7/216$.

   (c) $P(Y = 4 | X_1 = 0) = 1/25$.

   (d) $E(X_i) = 7/6$ for $1 \le i \le 6$.

   (e) $E(X) = 13/2$.

   (f) $E(Y) = 49/2$.

Solutions

1. Obviously, the question is closely related to the negligent secretary problem discussed in class. Denote the probability of the secretary to send no letter to its right destination by p_n. Recall that $p_n \underset{n\to\infty}{\longrightarrow} e^{-1}$.

   For X to assume the value 0, it is necessary that none of the participants takes his coat and none takes his umbrella. The two events are independent, and each has probability of p_n. Therefore

$$P(X = 0) = p_n^2 \underset{n\to\infty}{\longrightarrow} e^{-2}.$$

The variable X is distributed "approximately" $B(2n, 1/n)$. In fact, we may consider X to be the number of successes out of $2n$ trials, each of success probability $1/n$. However, these trials are dependent. Therefore, X is not distributed exactly $B(2n, 1/n)$. A simple way of seeing that X is not binomially distributed is by noticing that, while it assumes values between 0 and $2n$, the value $2n - 1$ has 0 probability. However, the representation of X as a sum of $2n$ variables, each distributed $B(1, 1/n)$, does give:

$$E(X) = 2n \cdot \frac{1}{n} = 2.$$

The event $\{Y = 0\}$ may be interpreted as the event that the permutation, assigning to each coat number the number of the umbrella picked by the same person, has the property that no number is mapped to itself. Hence $P(Y = 0) = p_n$. Similarly to the calculation of $E(X)$ above, we obtain:

$$E(Y) = n \cdot \frac{1}{n} = 1.$$

If $Y = n$, then all coat-umbrella pairs go together, and we are back to the original negligent secretary problem. Consequently:

$$P(X = 0 | Y = n) = p_n.$$

The probability for some specific person to pick up both his coat and his umbrella is $1/n^2$. More generally, the probability for any specific people $i_1, i_2, \ldots, i_k$, to pick up both their coats and their umbrellas is $1/(n(n-1)\ldots(n-k+1))^2$. By inclusion-exclusion, the probability for at least one person to pick up both his coat and his umbrella is

$$n \cdot \frac{1}{n^2} - \binom{n}{2} \cdot \frac{1}{(n(n-1))^2} + \binom{n}{3} \cdot \frac{1}{(n(n-1)(n-2))^2} - \ldots + \binom{n}{n} \cdot \frac{(-1)^{n-1}}{n!}.$$

Thus, only (c), (d) and (f) are true.

2. The values X assumes with positive probability are all positive integers, so that X is not hypergeometrically distributed. In fact, defining success as obtaining exactly one white and two black balls, we see that X is the number of trials until the first success. Hence X is geometrically distributed.

   The number of white balls drawn at any particular trial is distributed $H(3,2,4)$. Hence the probability for no white balls is $\binom{2}{0}\binom{4}{3}/\binom{6}{3} = 1/5$, for one white ball $\binom{2}{1}\binom{4}{2}/\binom{6}{3} = 3/5$, and for two white balls $\binom{2}{2}\binom{4}{1}/\binom{6}{3} = 1/5$. Consequently

$$P(Y=0) = \sum_{k=0}^{\infty} \left(\frac{1}{5}\right)^k \cdot \frac{3}{5} = \frac{3}{4},$$

   and therefore

$$P(Y \geq 1) = 1 - P(Y=0) = 14\,.$$

   An alternative way of obtaining the same result is as follows. For the question of whether $Y \geq 1$ or not, trials at which two white are drawn may be ignored. Out of the other trials, the probability for three black balls is $1/4$ and for two black balls is $3/4$. The event $\{Y \geq 1\}$ is therefore of probability $1/4$. This shows also that $Y = Y' - 1$, where $Y' \sim G(3/4)$. Hence $E(Y) = 4/3 - 1 = 1/3$. Similarly, $E(Z) = 1/3$.

   The more trials at which all three balls drawn are black, the more trials we may expect altogether. Hence the events $\{X \geq 1000\}$ and $\{Y \geq 500\}$ are dependent.

   The random variable $Y + Z$ assumes each non-negative integer value with a positive probability. Hence its distribution function has discontinuities at all non-negative integer points.

   Now:

$$P(Z=1|Y=1) = \frac{P(Y=1,\ Z=1)}{P(Y=1)}.$$

   With Y' as above, we have:

$$P(Y=1) = P(Y'=2) = \left(1 - \frac{3}{4}\right) \cdot \frac{3}{4} = \frac{3}{16}.$$

   The event $\{Y = 1,\ Z = 1\}$ occurs if, out of the first two trials one results in three black balls and the other in two white balls and a black ball, and the third trial results in two black balls and a white ball. Hence:

$$P(Y=1,\ Z=1) = 2\cdot\left(\frac{1}{5}\right)^2\cdot\frac{3}{5} = \frac{6}{125}.$$

Consequently:

$$P(Z=1|Y=1) = \frac{6/125}{3/16} = \frac{32}{125}.$$

Thus, only (d) and (e) are true.

3. The values X assumes with positive probability are all integers between 0 and 10,000. Hence X is not geometrically distributed. Also, at any point ω of the sample space we have $X^2(\omega) \geq X(\omega)$. Since the values 2 and above have positive probabilities, there is a positive probability that $X^2 > X$. In particular, $E(X^2) > E(X)$.

   The variable X may assume only finitely many values, and hence it has a finite expectation. The same applies to any function of X.

   By symmetry, X and Y are identically distributed. However, they may assume distinct values, so that they are not identical (as functions from the sample space to the real line).

   The number of trials at which all 13 tosses result in a head is distributed $B(10000, 1/2^{13})$, which is approximately $P(10000/2^{13})$. Hence the probability for no trial with 13 heads is about $e^{-1.221} = 0.295$. (Moreover, the probability for no trial with 13 tails is about the same). Hence it is quite reasonable that not all numbers between 0 and 13 will appear at least once as the number of heads in some trial.

   U assumes only values from 3 and above, and in particular it is not geometrically distributed. (In fact, it is a sum of three geometric random variables.)

   Thus, only (b) is true.

4. We have

$$P(X_6 = 0) = \sum_{n=2}^{\infty} P(X = n,\ X_6 = 0) = \sum_{n=2}^{\infty} \frac{5}{6} \cdot \left(\frac{4}{6}\right)^{n-2} \cdot \frac{1}{6} = \frac{5}{12}\,.$$

The event $\{Y = 6\}$ consists of the outcome (3, 3), the outcome (4, 1, 1), and the outcome (1, 3, 1, 1), with total probability

$$\frac{1}{6^2} + \frac{1}{6^3} + \frac{1}{6^4} > \frac{7}{216}\,.$$

Now:

$$P(Y = 4 | X_1 = 0) = \frac{P(Y = 4,\ X_1 = 0)}{P(X_1 = 0)}\,.$$

The event in the numerator consists of the single outcome (2, 2), whose probability is $1/6^2$. The event in the denominator has the same probability as the one calculated in part (a). Hence:

$$P(Y = 4 | X_1 = 0) = \frac{1/6^2}{5/12} = \frac{1}{15}\,.$$

The trial cannot be over after the first toss. From the second toss on, it has a probability of $1/6$ of being over at each toss. Hence $X = 1 + X'$, where $X \sim G(1/6)$, and in particular $E(X) = 1 + 6 = 7$. Now $X = \sum_{i=1}^{6} X_i$, and by symmetry all X_i's are identically distributed, and in particular have the same expectations. Thus, $E(X_i) = 7/6$ for each i. Finally, we have

$$Y = \sum_{i=1}^{6} iX_i,$$

and consequently

$$E(Y) = \frac{7}{6} \cdot \sum_{i=1}^{6} i = \frac{49}{2}\,.$$

Thus, only (d) and (f) are true.