Midterm

Mark all correct answers in each of the following questions.

1. A die is rolled n times. Let X_i be the result of the i-th roll, $1 \le i \le n$. Define random variables $Y_i,\ 1 \le i \le n-1$, by:

$$Y_i = \begin{cases} 1, & |X_{i+1} - X_i| = 1, \\ 0, & \text{otherwise.} \end{cases}$$

   Also put $Z_k = \sum_{i=1}^{k} X_i$ for $1 \le k \le n$ and $Y = \sum_{i=1}^{n-1} Y_i$.

   (a) $P(|X_j - X_i| = 1) = 1/3$ for $1 \le i < j \le n$.

   (b) For $n \ge 2$ we have $Y \sim B(n-1, 5/18)$.

   (c) $P(X_2 = 5 | Y_1 = 1) = 1/6$.

   (d) $\lim_{z \to 6-} F_{Z_3}(z) = 5/108$, but $\lim_{z \to 6+} F_{Z_3}(z) = 5/54$.

   (e) If $W \sim U[n, 6n]$ (discrete uniform), then the functions F_{Z_n} and F_W are not identical. However, $E(Z_n) = E(W)$ and $V(Z_n) = V(W)$.

2. Two players – A and B – play in the following game. At the first stage A tosses one coin and B tosses none, at the second stage A tosses two coins and B tosses one, and so on. That is, in general, at the k-th stage, A tosses k coins and B tosses $k-1$. The game continues until, at some stage, A has more heads than B. Let X be the total number of coin tosses throughout the game and Y the total number of tails. (For example, if at the first stage A has a tail, at the second – A has a head and a tail and B has a head, at the third – A has a head and two tails and B has two heads, and at the fourth A has three heads and one tail and B has two heads and a tail, then the game ends after the fourth stage with $X = 16$ and $Y = 6$.)

(a) $\sqrt{X} \sim G(1/2)$.

(b) Y is binomially distributed.

(c) $E(X) = 4$.

(d) The random variable $X - Y$, which counts the number of heads throughout the game, has the same distribution as Y.

(e) Suppose the game described above is played 2^{10} times. Let Z be the number of games which ended after exactly 9 stages. Then Z is distributed approximately $P(1/2)$.

3. The life-length in hours of light bulbs of a certain kind is distributed $\text{Exp}(1/100)$. We put on 100 bulbs at the same time, and leave them on until they all burn out. Let T_k be the time until the k-th of them burns out, $1 \leq k \leq 100$ (so that $0 \leq T_1 \leq T_2 \leq \ldots \leq T_{100}$).

   (a) $P(T_1 \leq 1) > 1/2$.

   (b) T_1 is exponentially distributed.

   (c) The random variables $2T_{50}$ and T_{100} are not identically distributed, but they have the same expectation.

   (d) Suppose 50 of the light bulbs are white and the other 50 are yellow. The probability that the number of white light bulbs still working is throughout the process never less than the number of yellow light bulbs still working is larger than $1/100$.

   (e) Since the life-length of a bulb is exponentially distributed, and the exponential distribution is memory-less, the events $\{T_1 \geq t_1\}$ and $\{T_2 \geq t_2\}$ are independent for any two non-negative numbers t_1 and t_2. However, for some other distributions of the life-length, the above two events could well be dependent.

Solutions

1. Out of the 36 possible outcomes for (X_i, X_j), those satisfying $|X_j - X_i| = 1$ are $(1,2), (2,3), \dots, (5,6)$, as well as the same pairs in inverse order – altogether 10 pairs. Hence $P(|X_j - X_i| = 1) = 10/36 = 5/18$. It follows that Y is indeed the number of events, out of a sequence of $n-1$ events of probability $5/18$ each, which occur. However, these events are not independent. For example, for $n = 3$, there are 18 triples (X_1, X_2, X_3) for which $Y = 2$; in fact, out of the 10 relevant values of (X_1, X_2), the values $(2,1)$ and $(5,6)$ lead to the possibilities $(2,1,2)$ and $(5,6,5)$, respectively, while each of the other 8 relevant values leads to two possibilities. As $18/6^3 \neq (5/18)^2$, the random variable Y is not distributed $B(2, 5/18)$. Now:

$$P(X_2 = 5 | Y_1 = 1) = \frac{P(X_2 = 5, Y_1 = 1)}{P(Y_1 = 1)}$$

$$= \frac{P(X_1 = 4, X_2 = 5) + P(X_1 = 6, X_2 = 5)}{5/18} = \frac{1}{5}\,.$$

The only triple (X_1, X_2, X_3) for which $Z_3 = 3$ is $(1,1,1)$, so that $P(Z_3 = 3) = 1/216$. The triples (X_1, X_2, X_3) for which $Z_3 = 4$ are $(1,1,2), (1,2,1), (2,1,1)$, and therefore $P(Z_3 = 4) = 3/216$. Similarly, $P(Z_3 = 5) = 6/216$ and $P(Z_3 = 6) = 10/216$. It follows that $F_{Z_3}(5) = 10/216 = 5/108$ and $F_{Z_3}(6) = 20/216 = 5/54$. As Z_3 assumes only integer values, this yields $\lim_{z \to 6-} F_{Z_3}(z) = 5/108$ and $\lim_{z \to 6+} F_{Z_3}(z) = 5/54$.

The variables W and Z_n both assume the values $n, n+1, \dots, 6n$, but (for $n \geq 2$) not with the same probabilities. For example, $Z_n = n$ only if the results of all rolls are 1, so that $P(Z_n = n) = 1/6^n$, whereas $P(W = n) = 1/(5n+1)$. The two variables have the same expectations, since both are symmetric with respect to the number $7n/2$, but they have distinct variances. For example, for $n = 2$ we have

$$V(W) = \frac{(12 - 2 + 1)^2 - 1}{12} = 10,$$

whereas a direct calculation yields $V(Z_2) = 35/6$.

Thus, only (d) is true.

2. As we have seen in class, at any stage the game gets to, the probability for A to have more heads than B, and thus for the game to end, is $1/2$. Hence the number of stages is distributed $G(1/2)$. Now it is easy to see (by induction, say) that X is the square of the number of stages, and hence $\sqrt{X} \sim G(1/2)$.

   To calculate $E(X)$, we shall use the fact that $\sqrt{X} \sim G(1/2)$. This implies:

$$E(X) = E\left(\sqrt{X}^2\right) = V\left(\sqrt{X}\right) + E^2\left(\sqrt{X}\right) = \frac{1/2}{(1/2)^2} + \left(\frac{1}{1/2}\right)^2 = 6.$$

   The number of tails is not bounded above, and thus Y cannot possibly be binomially distributed. Since the game ends only at a stage where A has more heads than B, the number of heads tossed in the course of the game must be strictly positive, whereas the number of tails in the game may well be 0. Hence $X - Y$ and Y have distinct distributions.

   As explained above, the number of stages in the game is distributed $G(1/2)$. Consequently, the probability for a game to end after exactly 9 stages is $1/2^9$. Hence $Z \sim B\left(2^{10}, 1/2^9\right)$. By the Poissonian approximation for the binomial distribution, Z is distributed approximately $P(2)$.

   Thus, only (a) is true.

3. For $t \geq 0$, the event $\{T_1 > t\}$ occurs if and only if all 100 light bulbs last over t hours. This event is an intersection of 100 independent events, the probability of each of which is $e^{-t/100}$. Hence:

$$P(T_1 > t) = \prod_{1 \leq i \leq 100} e^{-t/100} = e^{-t}\,.$$

   This implies $T_1 \sim \mathrm{Exp}(1)$, and in particular $P(T_1 \leq 1) = 1 - e^{-1} \approx 0.63$. One can see as above (due to the fact that the exponential distribution is memory-less) that, after i light bulbs have burnt out, the time until the next light bulb burns out is distributed $\mathrm{Exp}((n-i)/100)$. This gives $E(T_{i+1} - T_i) = 100/(n-i)$, and consequently

$$E(2T_{50}) = 2\left(\frac{100}{100} + \frac{100}{99} + \ldots + \frac{100}{51}\right),$$

whereas

$$E(T_{100}) = \frac{100}{100} + \frac{100}{99} + \ldots + \frac{100}{1}\,.$$

Obviously, this implies that $E(T_{100}) > E(2T_{50})$. (The intuitive explanation for this inequality is as follows. The more light bulbs are still working, the less time one needs to wait on the average until one of them burns out. Hence, on the average it should take less time from the beginning until the 50-th light bulb burns out than from that time point until all light bulbs burn out.)

The situation in part (d) is completely analogous to that in the ballot problem, where instead of the votes for candidate A and for candidate B we have yellow bulbs and white bulbs, respectively. Hence the required probability is 1/51.

To see that (e) is false, the simplest is to take t_1 and t_2 with $t_1 \geq t_2$. In this case $\{T_1 \geq t_1\} \subseteq \{T_2 \geq t_2\}$. As the probabilities of both events are strictly between 0 and 1, this shows that they cannot possibly be independent. (We mention in passing that it is easy to understand intuitively that the events are not independent if $t_1 < t_2$ as well.)

Thus, (a), (b) and (d) are true.