Final #2

Mark all correct answers in each of the following questions.

1. Consider the ballot problem with $m = n = 4$.

   (a) The probability that at least one of the two candidates is not behind the other throughout the entire counting process is 0.4.

   (b) Suppose that the counting process is repeated over and over independently. Let X_l be the number of times, out of the first l trials, in which candidate #1 was not behind candidate #2 throughout the entire counting process. Then $X_{2k} - X_k$ is distributed binomially.

   (c) Markov's inequality implies:

   $$P(X_{200} \geq X_{100} + 30) \leq \frac{2}{3}.$$

   (d) Chebyshev's inequality implies:

   $$P(|X_{200} - X_{100} - 20| \geq 8) \leq \frac{1}{4}.$$

   (e)
   $$P(X_{20000} \geq X_{10000} + 1960) \approx \Phi(1/\sqrt{2}),$$
   where Φ is the standard normal distribution function.

   (f)
   $$P\left(\left|\frac{X_{2l} - X_l}{l} - \frac{1}{2}\right| > \varepsilon\right) \underset{l \to \infty}{\longrightarrow} 0.$$

2. Let k, n be positive integers. An urn contains n balls out of each of k colors (altogether kn balls). The balls are drawn from the urn one by one without replacement. For $1 \le i \le k$, let X_i denote the sum of the numbers of drawings in which balls of color i are drawn. For example, if $k = 3$ and $n = 2$, and the colors of drawn balls are 3, 2, 2, 1, 3, 1, then $X_1 = 10,\ X_2 = 5,\ X_3 = 6$.

   (a) $E(X_i) = \frac{kn(kn+1)}{2}\,, \quad 1 \le i \le k$.

   (b) $\rho(X_1, X_2) \ge 0$.

   (c) If $k = 6$, then $\rho(X_1 + X_2, X_3 + X_4) = -\frac{1}{2}\,$.

   (d) Let $Y = 1$ if the first ball to be drawn is of color #1 and $Y = 0$ otherwise. Then $\text{Cov}(X_1, Y) > 0$.

   (e) Let Z be the number of indices $j,\ 1 \le j \le kn - 1$, for which the j-th ball and the $(j+1)$-st ball are of the same color. Then Z is distributed hypergeometrically.

   (f) The k-tuple $(X_1, X_2, \dots, X_k)$ is distributed multinomially.

3. Let $X \sim \text{Exp}(1)$. Put:

$$Y = X^2, \quad W = \sqrt{X}, \quad S = [X], \quad T = \{X\},$$

   where $[a]$ and $\{a\}$ denote the integer part and the fractional part of a real number a, respectively. (For example, $[2.9] = 2$ and $\{2.9\} = 0.9$.

   (a) $E(Y) = 1$.

   (b) $V(Y) = 20$.

   (c) $V(W) = \frac{\pi}{4}\,$.

   (d) $S + 1$ is geometrically distributed.

   (e) S and T are independent.

   (f) $E(T) < \frac{1}{2}\,$.

   (g) $V(S) = \frac{1}{e}\,$.

4. Let X be a random variable.

   (a) There exists a constant C such that, if X assumes only non-negative values and $E(X) = 1$, then $V(X) < C$.

   (b) If X is discrete and assumes with positive probabilities the values $x_1, x_2, \ldots$, where $|x_i - x_j| > 1$ for every $i \neq j$, then $V(X) > 1$.

   (c) If X assumes only non-negative values and $\rho(X, X^2)$ exists, then $\rho(X, X^2) \geq 0$.

   (d) Let Y be an additional random variable. Suppose that at each point of the sample space X and Y assume values of opposite signs. (Assume for simplicity that X and Y do not assume the value 0.) If $\rho(X, Y)$ exists, then $\rho(X, Y) < 0$.

   (e) If the moment generating function of X is $\psi_X(t) = \frac{e^{3t}+e^{-3t}}{2}$, then $P(-1 \leq X \leq 1) \geq \frac{1}{9}$.

   (f) It is impossible for the moment generating function of X to be $\psi_X(t) = \cos t$.

Solutions

1. The probability for a specific candidate not to be behind the other throughout the counting process is $1/(4+1) = 0.2$. Since it is impossible that each of them is never behind the other, the probability in (a) is $2 \cdot 0.2 = 0.4$.

   Defining as a success the event that candidate #1 is never behind candidate #2 throughout some specific counting process, $X_{2k} - X_k$ is the number of successes within k independent trials. Hence $X_{2k} - X_k \sim B(k, 0.2)$. For $k = 100$ we have

   $$E(X_{200} - X_{100}) = 100 \cdot 0.2 = 20,$$

   so that Markov's inequality implies:

   $$P(X_{200} \geq X_{100} + 30) = P(X_{200} - X_{100} \geq 30) \leq \frac{20}{30} = \frac{2}{3}.$$

Since

$$V(X_{200} - X_{100}) = 100 \cdot 0.2 \cdot (1 - 0.2) = 16,$$

Chebyshev's inequality implies:

$$P(|X_{200} - X_{100} - 20| \geq 8) \leq \frac{16}{8^2} = \frac{1}{4}\,.$$

By the central limit theorem

$$\begin{aligned} P(X_{20000} \geq X_{10000} + 1960) &= P(X_{20000} - X_{10000} \geq 1960) \\ &= P\left(\frac{X_{20000} - X_{10000} - 10000 \cdot 0.2}{\sqrt{10000 \cdot 0.2 \cdot (1-0.2)}} \geq \frac{1960 - 10000 \cdot 0.2}{\sqrt{10000 \cdot 0.2 \cdot (1-0.2)}}\right) \\ &\approx P(Z \geq -1), \end{aligned}$$

where Z is a standard normal random variable. Thus:

$$P(X_{20000} \geq X_{10000} + 1960) \approx \Phi(1).$$

Similarly to the proof of the weak law of large numbers for i.i.d. random variables with finite expectation and variance, we have:

$$P\left(\left|\frac{X_{2l} - X_l}{l} - 0.2\right| > \varepsilon\right) \leq \frac{0.2 \cdot (1 - 0.2)/l}{\varepsilon^2} \xrightarrow[l \to \infty]{} 0.$$

Consequently, the sequence $\left(\frac{X_{2l} - X_l}{l}\right)_{l=1}^{\infty}$ converges in probability, but its limit is 0.2 and not 0.5.

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

2. By symmetry, all $E(X_i)$'s, $1 \leq i \leq k$, are identical. Now

$$\sum_{i=1}^{k} X_i = 1 + 2 + \ldots + kn = \frac{kn(kn+1)}{2}\,,$$

and therefore

$$E(X_i) = \frac{1}{k} \cdot \frac{kn(kn+1)}{2} = \frac{n(kn+1)}{2}\,, \qquad 1 \leq i \leq k.$$

Again by symmetry, all $V(X_i)$'s, $1 \leq i \leq k$, are identical, as are all $\text{Cov}(X_i, X_j)$'s for $i \neq j$. Hence:

$$\begin{aligned} 0 &= V\left(\sum_{i=1}^k X_i\right) = \sum_{i=1}^k V(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \\ &= kV(X_1) + k(k-1)\text{Cov}(X_1, X_2). \end{aligned}$$

It follows that

$$\text{Cov}(X_1, X_2) = -\frac{V(X_1)}{k-1},$$

and therefore

$$\rho(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{V(X_1)V(X_2)}} = -\frac{1}{k-1}, \tag{1}$$

We mention in passing that the fact that distinct X_i's are negatively correlated is intuitively clear. In fact, a large value of X_i indicates that the balls of color i were drawn relatively late. This indicates that balls of other colors were drawn early, so that other X_j's should be relatively small.

If $k = 6$, we may unite colors #1 and #2, colors #3 and #4, and colors #5 and #6, and then use (1) to find the coefficient of correlation between the sum of the numbers of drawings in which balls of colors #1 and #2 together were drawn and the sum of the numbers of drawings in which balls of colors #3 and #4 together were drawn to obtain:

$$\rho(X_1 + X_2, X_3 + X_4) = -\frac{1}{3-1} = -\frac{1}{2}.$$

A large value of Y, namely $Y = 1$, means that the first ball to be drawn is of color #1, which fact tends to indicate that the sum of the numbers of drawings in which balls of color #1 are drawn is (a little) smaller than could otherwise be expected. Hence $\text{Cov}(X_1, Y) < 0$.

Z has no connection to the trial used to define the hypergeometric distribution, so that it should not be hypergeometrically distributed. (Formally, one may show that this is the case in general as follows. Take $n = k = 2$. It is easily verified that Z attains the values $0, 1, 2$ with probability $1/3$ each. Suppose $Z \sim H(m, a, b)$ for some m, a, b. Since an $H(m, a, b)$-distributed variable assumes with positive probability the integer values between $\max(0, m-b)$ and $\min(m, a)$, we must have $m \leq b$ and $\min(m, a) = 2$. If $m = 2$, then by equating the expected value $ma/(a+b)$ of an $H(m, a, b)$-distributed variable with the expected value 1 of Z, we obtain $b = a$. Since $P(Z = 0) = 1/3$, we find that

$$\frac{\binom{a}{0}\binom{b}{2}}{\binom{a+b}{2}} = \frac{1}{3}\,.$$

The solution of the equation is $b = 2$, which is impossible. If $a = 2$, then by equating expectations we get this time $b = 2m - 2$, and plugging in the probability of the event $\{Z = 0\}$ we again obtain a contradiction.)

In a multinomial distribution the (vector) values the variable may assume are all integer vectors with non-negative components of some fixed sum. The components of the values $(X_1, X_2, \dots, X_k)$ assumes are all at least $1 + 2 + \dots + n = \frac{n(n+1)}{2}$, and hence $(X_1, X_2, \dots, X_k)$ is not multinomially distributed.

Thus, only (c) is true.

3. We have seen that the moment generating function of X is $\psi_X(t) = (1-t)^{-1}$ for $t < 1$. Therefore

$$E(Y) = E(X^2) = \left[\frac{d^2}{dt^2}(1-t)^{-1}\right]_{t=0} = [2(1-t)^{-3}]_{t=0} = 2,$$

and

$$E(Y^2) = E(X^4) = \left[\frac{d^4}{dt^4}(1-t)^{-1}\right]_{t=0} = [24(1-t)^{-5}]_{t=0} = 24,$$

so that

$$V(Y) = E(Y^2) - E(Y)^2 = 20.$$

The distribution function of W is found by:

$$F_W(w) = P(W \leq w) = P(X \leq w^2) = 1 - e^{-w^2}, \qquad w \geq 0.$$

Hence

$$f_W(w) = P(W \leq w) = P(X \leq w^2) = 2we^{-w^2}, \qquad w \geq 0,$$

which gives

$$\begin{aligned} E(W) &= \int_0^\infty w \cdot 2we^{-w^2} dw = \int_{-\infty}^\infty w^2 e^{-w^2} dw \\ &= \sqrt{\pi} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi(1/\sqrt{2})^2}} w^2 e^{-\frac{w^2}{2(1/\sqrt{2})^2}} dw. \end{aligned}$$

The integral on the right hand side gives the expectation of the square of an $N(0, 1/2)$ random variable, and therefore

$$E(W) = \frac{\sqrt{\pi}}{2}.$$

It follows that:

$$V(W) = E(W^2) - E(W)^2 = E(X) - \frac{\pi}{4} = 1 - \frac{\pi}{4}.$$

To find the distribution of S we note that for a non-negative integer k:

$$P(S = k) = P(k \le X < k + 1) = e^{-k} - e^{-(k+1)} = \left(1 - \frac{1}{e}\right) e^{-k}.$$

This implies readily that $S + 1 \sim G(1 - 1/e)$. In particular:

$$V(S) = V(S + 1) = \frac{1/e}{(1 - 1/e)^2} = \frac{e}{(e - 1)^2}.$$

To verify that S and T are independent, it suffices to check that the events $\{S = k\}$ and $\{0 \le T < t\}$, with non-negative integer k and $t \in [0, 1]$, are independent. In fact,

$$\begin{aligned} P(0 \le T < t) &= \sum_{k=0}^\infty P(k \le X < k + t) \\ &= \sum_{k=0}^\infty \left((1 - e^{-(k+t)}) - (1 - e^{-k})\right) \\ &= \sum_{k=0}^\infty e^{-k}(1 - e^{-t}) = \frac{1 - e^{-t}}{1 - 1/e}, \end{aligned}$$

and therefore

$$\begin{aligned} P(S = k, 0 \le T < t) &= P(k \le X < k + t) \\ &= e^{-k} - e^{-(k+t)} = e^{-k}(1 - e^{-t}) \\ &= (1 - 1/e)e^{-k} \cdot \frac{1 - e^{-t}}{1 - 1/e} = P(S = k)P(0 \le T < t). \end{aligned}$$

To calculate the expected value of T we first find the density function of T

$$f_T(t) = \frac{e^{-t}}{1 - 1/e} = \frac{e}{e-1} \cdot e^{-t}, \qquad 0 \le t \le 1.$$

Consequently:

$$E(T) = \int_0^1 \frac{e}{e-1} \cdot te^{-t} dt = \frac{e}{e-1} \left[-te^{-t} - e^{-t} \right]_{t=0} = \frac{e-2}{e-1} < \frac{1}{2}.$$

Thus, (b), (d), (e) and (f) are true.

4. The fact that $E(X) = 1$ does not imply that $V(X)$ exists. Even if $V(X)$ does exist, it may assume an arbitrarily large value. In fact, given any $C > 0$, let X assume the two values 0 and $C + 1$ with probabilities $1 - 1/(C+1)$ and $1/(C+1)$, respectively. Then $E(X) = 1$ and $V(X) = C$.

   If X is discrete and assumes with positive probabilities infinitely many values "far" from each other ($|x_i - x_j| > 1$ for every $i \ne j$), it may still have a small variance if it assumes one value with a probability sufficiently close to 1 and all others with very small probabilities. Specifically, let $X = 2Y$, where $Y \sim G(p)$. Then all values X assumes are at a distance of at least 2 apart, yet $V(X) = 4q/p^2$, which can be made arbitrarily small by choosing p sufficiently close to 1.

   If X assumes only non-negative values, then X and X^2 take small values and large values together, so that $\rho(X, X^2)$ is positive if it exists. Formally, this may be proved using Jensen's inequality. In fact, since the functions $g(x) = x^3$ and $h(x) = x^{3/2}$ are convex on the positive x-axis, we have

$$E(X^3) \ge E(X)^3,$$

   and

$$E(X^3) = E\left((X^2)^{3/2}\right) \ge E(X^2)^{3/2}.$$

   Consequently

$$E(X^3) = E(X^3)^{1/3} E(X^3)^{2/3} \ge \left(E(X)^3\right)^{1/3} \left(E(X^2)^{3/2}\right)^{2/3} = E(X)E(X^2),$$

   which implies $\mathrm{Cov}(X, X^2) \ge 0$ and therefore $\rho(X, X^2) \ge 0$.

The fact that X and Y assume values of opposite signs at each point does not mean that they cannot be positively correlated. For example, let X assume the values 1 and 2 with probability 0.5 each, and let $Y = X - 3$. Then X is always positive and Y always negative, yet $\rho(X, Y) = 1$. The example may be strengthened so as to have X and Y assume both positive and negative values. In fact, let X assume the values 1 and 2 with probability $0.5 - \varepsilon$ each and the value -1 with probability 2ε. Let $Y = X - 3$ if $X > 0$ and $Y = 1$ if $X = -1$. Clearly, $\rho(X, Y)$ can be made arbitrarily close to 1 by taking ε sufficiently small.

If X assumes the values 3 and -3 with probability 0.5 each, then $\psi_X(t) = \frac{e^{3t}+e^{-3t}}{2}$. Since the moment generating function determines the distribution uniquely, the opposite holds as well, and hence under the assumptions of (e) we have $P(-1 \le X \le 1) = 0$.

If $\psi_X(t) = \cos t$, then

$$E(X^2) = \left[\frac{d^2}{dt^2} \cos t\right]_{t=0} = -\cos 0 = -1,$$

which is impossible.

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