Final #1

Mark all correct answers in each of the following questions.

1. Reuven and Shim'on play a dreidle game. Each spins his own dreidle at his turn. Reuven spins first. The game continues until one of them gets a 'G' (and wins all money on the table). Reuven's dreidle falls on 'G' with probability p_1 and Shim'on's with probability p_2. Let X be the number of times Reuven spins his dreidle and Y the analogous number for Shim'on. For example, if Reuven gets a 'G' at the first spin then $X = 1$ and $Y = 0$, while if the first 'G' is obtained at the second spin of Shim'on then $X = Y = 2$.

   (a) X is geometrically distributed.

   (b) $Y + 1$ is geometrically distributed.

   (c) Y is either X or $X - 1$, and in particular $\rho(X, Y) = 1$.

   (d) If $p_1 = 1/3$ and $p_2 = 3/4$, then Reuven's probability of winning the game is $2/5$.

   (e) Now suppose that the game is played 1000 times, with Reuven starting at each round. If $p_1 = 1/1000$ and $p_2 = 11/111$, then the number of rounds Reuven wins is distributed approximately $P(1)$.

   (f) Suppose the game is played 10000 times, where $p_1 = p_2 = 3/4$. Then the probability for the total number of times Reuven spins his dreidle to exceed that of Shim'on by at least 8040 is approximately 0.16.

2. A room is lit by two lamps, each using a light bulb whose lifetime in years is distributed Exp(1). We also have a spare light bulb of the same

quality, which will be used to replace the bulb to burn out first. After the second bulb burns out, the room will be lit by the surviving bulb alone while this bulb lasts. Denote by X the time until the first light bulb is burnt up, by Y the time (from the beginning) until two of the light bulbs are burnt up, and by Z the time until all light bulbs are burnt up.

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

(b) $P(X \geq x) = e^{-2x}$ for $x \geq 0$.

(c) Markov's inequality yields:

$$P(Z \geq a) \leq \frac{2}{a}, \qquad a > 0.$$

(d) Chebyshev's inequality yields:

$$P(|Z - X - E(Z - X)| \geq \varepsilon) \leq \frac{25}{16\varepsilon^2}, \qquad \varepsilon > 0.$$

(e) $0.65 \leq \rho(X, Y) \leq 0.75$.

3. The variable (X, Y) is uniformly distributed in the planar region

$$S = \{(x, y) : 0 \leq x \leq 1,\ x^3 \leq y \leq x^2\}.$$

(That is, since the area of S is $1/12$, the probability of (X, Y) to assume a value in some set $S' \subseteq S$ is the area of S' multiplied by 12.)

(a) The distribution function of X is given by:

$$F_X(x) = \begin{cases} 0, & x < 0, \\ 4x^2 - 3x^4, & 0 \leq x \leq 1, \\ 1, & 1 < x. \end{cases}$$

(b) The density function of Y is given by:

$$f_Y(y) = \begin{cases} 9y^{2/3} - 8y^{1/2}, & 0 \leq y \leq 1, \\ 0, & \text{otherwise}. \end{cases}$$

(c) $P(Y \leq 1/9 \mid X \leq 1/2) = 79/135$.

(d) $E(Y^2) = \frac{6}{35}$.

4. Let X, Y, Z be random variables.

   (a) If $X + Y, Z$ are independent, then both pairs X, Z and Y, Z are such.

   (b) Suppose $Y = \sin X$ and $Z = \cos X$. Denote $\nu = E(Y)$ and $\eta = E(Z)$. Then $\nu^2 + \eta^2 \leq 1$.

   (c) If X, Y are independent identically distributed with expectation 0 and finite variance, then for every $\varepsilon > 0$:
   $$P(|X| \geq \varepsilon) \leq P(|X + Y| \geq \varepsilon).$$

   (d) If every two of the three random variables are uncorrelated, then $E(XYZ) = E(X)E(Y)E(Z)$ (assuming all expectations exist).

   (e) The inequalities $\rho(X, Y) > 0$ and $\rho(Y, Z) > 0$ do not imply the inequality $\rho(X, Z) > 0$.

Solutions

1. Defining a success as the event that either Reuven or Shim'on get a 'G', the variable X counts the number of trials until the first success. Since the probability of success is $p_1+p_2-p_1p_2$, we have $X \sim G(p_1+p_2-p_1p_2)$.

   Since $P(Y + 1 = 1) = p_1$, if $Y + 1 \sim G(p)$ for some p, then we must have $p = 1 - p_1$. Hence $P(Y + 1 = 2)$ should be $p_1(1 - p_1)$, whereas we see easily that $P(Y + 1 = 2) = (1 - p_1)(p_2 + (1 - p_2)p_1)$.

   The correlation coefficient between two random variables is 1 iff one is a linear function of the other. Since, in our case, each of the equalities $Y = X$ and $Y = X - 1$ holds with positive probability, Y is not a linear function of X, so that $\rho(X, Y) < 1$.

   Let p be Reuven's probability of winning. Since the event that Reuven wins is made of the event that he wins at the first round, as well as the event that both he and Shim'on lose in the first round and Reuven wins later, we have

$$p = p_1 + (1 - p_1)(1 - p_2)p,$$

which yields:

$$p = \frac{p_1}{p_1 + p_2 - p_1 p_2}. \tag{1}$$

In particular, if $p_1 = 1/3$ and $p_2 = 3/4$, then $p = 2/5$.

If the game is played n times, then the number of times Reuven wins is distributed $B(n, p)$. By (1), for the data in (e) we have $p = 1/100$. Thus the number of wins of Reuven is distributed $B(1000, 1/100)$, which is approximately $P(10)$.

The number of times Reuven spins his dreidle more than Shim'on does in n games is the number of times Reuven wins. Denote this number for the data in (f) by T. Since for this data $p = 4/5$, we have $T \sim B(10000, 0.8)$. By the Central Limit Theorem

$$\begin{aligned} P(T \geq 8040) &= P\left(\tfrac{T - 10000\cdot 0.8}{\sqrt{10000\cdot 0.8\cdot 0.2}} \geq \tfrac{8040 - 10000\cdot 0.8}{\sqrt{10000\cdot 0.8\cdot 0.2}}\right) \\ &\approx P(Z \geq 1) \approx 0.16. \end{aligned}$$

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

2. The variable X is the minimum of two independent Exp(1)-distributed random variables. Hence:

$$P(X \geq x) = e^{-x} \cdot e^{-x} = e^{-2x}, \qquad x \geq 0.$$

Thus, $X \sim \text{Exp}(2)$, and in particular $E(X) = 1/2$. Similarly, and employing the lack-of-memory property of the exponential distribution, $Y - X \sim \text{Exp}(2)$ and the variables $X, Y - X$ are independent. Consequently,

$$\text{Cov}(X, Y) = \text{Cov}(X, X + (Y - X)) = V(X)$$

and

$$V(Y) = V(X + (Y - X)) = 2V(X),$$

so that

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y}{\sqrt{V(X)V(Y)}} = \frac{1}{\sqrt{2}} \approx 0.707.$$

Employing again the lack-of-memory property we obtain $Z - Y \sim \mathrm{Exp}(1)$ and, moreover, the variables $Y - X$, $Z - Y$ are independent. It follows that $E(Z) = 1/2 + 1/2 + 1 = 2$, so that Markov's inequality yields:

$$P(Z \geq a) \leq \frac{2}{a}, \qquad a > 0.$$

Also, $V(Z - X) = V(Y - X) + V(Z - Y) = (1/2)^2 + 1^2 = 5/4$, so that by Chebyshev's inequality:

$$P(|Z - X - E(Z - X)| \geq \varepsilon) \leq \frac{5}{4\varepsilon^2}, \qquad \varepsilon > 0$$

(which is stronger than the inequality in (d)).

Thus, (b), (c) and (e) are true.

3. We have

$$F_X(x) = 12\int_0^x (t^2 - t^3)dt = 12\left(\frac{x^3}{3} - \frac{x^4}{4}\right) = 4x^3 - 3x^4, \qquad 0 \leq x \leq 1.$$

Clearly, F_X vanishes on the negative half-line, and is identically 1 on $(1, \infty)$. Since the region S may be alternatively written in the form

$$S = \{(x, y) : 0 \leq y \leq 1,\ y^{1/2} \leq x \leq y^{1/3}\},$$

we have

$$F_Y(y) = 12\int_0^y (t^{1/3} - t^{1/2})dt = 9y^{4/3} - 8y^{3/2}, \qquad 0 \leq y \leq 1.$$

Therefore:

$$f_Y(y) = F'_Y(y) = 12y^{1/3} - 12y^{1/2}, \qquad 0 \leq y \leq 1.$$

Obviously, f_Y vanishes outside the interval $[0, 1]$. Hence:

$$E(Y^2) = \int_0^1 (12y^{1/3} - 12y^{1/2})y^2 dy = \frac{36}{10} - \frac{24}{7} = \frac{6}{35}\,.$$

Now

$$P(X \leq 1/2) = F_X(1/2) = 4(1/2)^3 - 3(1/2)^4 = 5/16$$

and, since $X^3 \leq Y$, so that $X \leq Y^{1/3}$:

$$\begin{aligned} P(X \leq 1/2, Y \leq 1/9) &= P(Y \leq 1/9) \\ &= 9 \cdot (1/9)^{4/3} - 8 \cdot (1/9)^{3/2} \\ &= \sqrt[3]{3}/3 - 8/27\,. \end{aligned}$$

Therefore:

$$P(Y \leq 1/9 \mid X \leq 1/2) = \frac{P(X \leq 1/2, Y \leq 1/9)}{P(X \leq 1/2)} = \frac{144\sqrt[3]{3} - 128}{135}\,.$$

Thus, only (d) is true.

4. Let W and Z be any independent random variables, X any variable, and denote $Y = W - X$. Then $X + Y, Z$ are independent, but neither X, Z nor Y, Z need be such

   Since $E(W^2) \geq E^2(W)$ for any random variable W, we have in (b):

   $$\nu^2 + \eta^2 = E^2(\sin X) + E^2(\cos X) \leq E(\sin^2 X) + E(\cos^2 X) = E(1) = 1.$$

   Let X and Y be independent and assume the values 1 and -1 with a probability $1/2$ each. Then they have 0 expectations, but

   $$P(|X| \geq 1) = 1 > 1/2 = P(|X + Y| \geq 1),$$

   which contradicts (c) for $\varepsilon = 1$. Moreover, taking $Z = XY$, we notice that both X, Z and Y, Z are independent and in particular uncorrelated. However, since $XYZ = 1$ identically:

   $$E(XYZ) = 1 \neq 0 = E(X)E(Y)E(Z).$$

   To construct X, Y, Z with $\rho(X, Y) > 0$ and $\rho(Y, Z) > 0$, but $\rho(X, Z) < 0$, we first recall that random variables with a positive correlation coefficient are analogous to vectors in a Euclidean space with an angle of less than $\pi/2$ between them. Random variables with a negative correlation coefficient are analogous to vectors with an angle larger than

$\pi/2$ between them. In the plane, the angle between $(1,0)$ and $(1,2)$ is smaller than $\pi/2$, and so is the angle between $(1,2)$ and $(-1,2)$, whereas the angle between $(1,0)$ and $(-1,2)$ is larger than $\pi/2$. Hence we start with two independent (or just uncorrelated) random variable S and T with $V(S) = V(T) > 0$ and take

$$X = S, \qquad Y = S + 2T, \qquad Z = -S + 2T.$$

Then:

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathrm{Cov}(S, S+2T) = V(S) > 0, \\
\mathrm{Cov}(Y,Z) &= \mathrm{Cov}(S+2T, -S+2T) \\
&= \mathrm{Cov}(S,-S) + \mathrm{Cov}(2T, 2T) = 3V(S) > 0, \\
\mathrm{Cov}(X,Z) &= \mathrm{Cov}(S, -S+2T) = -V(S) < 0.
\end{aligned}$$

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