Probability

Solutions to Exam A, Fall 2014

Solution Q1:

(A) First of all, if $t \leq 0$ then

$$\begin{aligned}\mathbb{P}[X \leq t] &= \mathbb{P}[X \leq t, P = 0] + \mathbb{P}[X \leq t, P > 0] \\ &= \mathbb{P}[1 \leq t, P = 0] + \mathbb{P}[\min\{U_1, \ldots, U_P\} \leq t, P > 0] \\ &\leq \sum_{n>0} \mathbb{P}[P = n, \exists\ 1 \leq j \leq n\ :\ U_j \leq 0] \leq 0,\end{aligned}$$

Because $\mathbb{P}[U_n \leq t] = 0$ for all $t \leq 0$.

If $t \geq 1$, then since $U_n \leq 1$ for all n, we have that $X \leq 1$ by definition. Thus, $\mathbb{P}[X \leq t] = 1$ for such $t \geq 1$.

Now let $0 < t < 1$. Then, using the fact that $P, (U_n)_n$ are all independent,

$$\mathbb{P}[X > t] = \sum_{n=0}^{\infty} \mathbb{P}[X > t, P = n] = \mathbb{P}[P = 0, 1 > t] + \sum_{n>0} \mathbb{P}[P = n, \min\{U_1, \ldots, U_n\} > t]$$

$$= \mathbb{P}[P = 0] + \sum_{n>0} \mathbb{P}[P = n] \cdot \mathbb{P}[\forall\ 1 \leq j \leq n\ ,\ U_j > t]$$

$$= \mathbb{P}[P = 0] + \sum_{n>0} \mathbb{P}[P = n] \cdot (1-t)^n = \sum_{n=0}^{\infty} \mathbb{P}[P = n](1-t)^n$$

$$= \sum_{n=0}^{\infty} e^{-\lambda} \frac{(1-t)^n \lambda^n}{n!} = e^{-\lambda} e^{\lambda(1-t)} = e^{-\lambda t}.$$

We conclude that:

$$F_X(t) = \begin{cases} 0 & t \leq 0 \\ 1 - e^{-\lambda t} & 0 \leq t < 1 \\ 1 & t \geq 1 \end{cases}$$

X is *not* a continuous random variable since F_X is not continuous at $t = 1$.

(B) If $t \leq 0$ then

$$\mathbb{P}[Y \leq t] \leq \mathbb{P}[E \leq t] + \mathbb{P}[1 \leq t] = 0.$$

If $t \geq 1$ then since $Y \leq 1$ by definition, $\mathbb{P}[Y \leq t] = 1$ for any such $t \geq 1$.

Let $0 < t < 1$. Then,

$$\mathbb{P}[Y > t] = \mathbb{P}[Y > t, E > 1] + \mathbb{P}[Y > t, E \leq 1] = \mathbb{P}[1 > t, E > 1] + \mathbb{P}[E > t, E \leq 1]$$
$$= \mathbb{P}[E > 1] + \mathbb{P}[t < E \leq 1] = \mathbb{P}[E > t] = e^{-\lambda t}.$$

Thus,

$$F_Y(t) = \begin{cases} 0 & t \leq 0 \\ 1 - e^{-\lambda t} & 0 \leq t < 1 \\ 1 & t \geq 1. \end{cases}$$

Note that $F_Y = F_X$.

(C) Note that $F_X = F_Y$, that is X, Y have the same distribution. Also, $X \geq 0, Y \geq 0$ so their expectation is well defined. Moreover, since $0 \leq X, Y \leq 1$ we have that $0 \leq \mathbb{E}[X] = \mathbb{E}[Y] \leq 1$. So $X - Y$ has a well defined finite expectation, and by linearity $\mathbb{E}[X - Y] = \mathbb{E}[X] - \mathbb{E}[Y] = 0$.

Solution Q2:

(A) Let's take $g(x) = \mathbf{1}_{\{x>0\}} - \mathbf{1}_{\{x\leq 0\}}, h(x) = |x|$. So g is measurable as a combination of indicators and h is measurable as a continuous function.

Let us calculate the joint distribution function of $X = g(U), Y = h(U)$:

Note that $X \in \{-1, 1\}$ by definition and $Y \in [0, 1]$ by definition. So for $t < -1$, since $F_X(t) = 0$,

$$F_{X,Y}(t, s) = \mathbb{P}[X \leq t, Y \leq s] = 0 = F_X(t) \cdot F_Y(s)$$

and for $s < 0$, since $F_Y(s) = 0$,

$$F_{X,Y}(t,s) = \mathbb{P}[X \le t, Y \le s] = 0 = F_X(t) \cdot F_Y(s).$$

Moreover, if $s \ge 1$ then since $Y \le 1$,

$$F_{X,Y}(t,s) = \mathbb{P}[X \le t, Y \le s] = \mathbb{P}[X \le t] = F_X(t) \cdot F_Y(s).$$

Similarly, if $t \ge 1$ then since $X \le 1$,

$$F_{X,Y}(t,s) = \mathbb{P}[X \le t, Y \le s] = \mathbb{P}[Y \le s] = F_X(t) \cdot F_Y(s).$$

Now, if $-1 \le t < 1, 0 \le s < 1$ then

$$\begin{aligned} F_{X,Y}(t,s) &= \mathbb{P}[X = -1, Y \le s] = \mathbb{P}[U \le 0, |U| \le s] = \mathbb{P}[-s \le U \le 0] \\ &= \tfrac{s}{2} = \tfrac{1}{2} \cdot s = \mathbb{P}[U \le 0] \cdot \mathbb{P}[-s \le U \le s] \\ &= \mathbb{P}[X = -1] \cdot \mathbb{P}[Y \le s] = F_X(t) \cdot F_Y(s). \end{aligned}$$

So we have show that for any t, s we have

$$F_{X,Y}(t,s) = F_X(t) \cdot F_Y(s).$$

That is, X, Y are independent.

(B) Take $g(x) = x, h(x) = x^2$. Then g, h are measurable as continuous functions.

We calculate $\mathrm{Cov}(g(U), h(U)) = \mathrm{Cov}(U, U^2)$:

$$\mathbb{E}[g(U)h(U)] = \mathbb{E}[U^3] = \int_{-1}^{1} t^3 \tfrac{1}{2} dt = \tfrac{1}{8} t^4 \Big|_{-1}^{1} = 0,$$

and $\mathbb{E}[g(U)] = \mathbb{E}[U] = 0$ so $\mathrm{Cov}(g(U), h(U)) = 0$ and $g(U), h(U)$ are uncorrelated.

We claim that $g(U), h(U)$ are not independent. If they were independent, then $g(U)^2, h(U)$ would be uncorrelated. That is, $0 = \mathrm{Cov}(U^2, U^2) = \mathrm{Var}[U^2]$. So $U^2 = \mathbb{E}[U^2] = \mathrm{Var}[U] = \frac{1}{3}$ a.s. However, since U is continuous,

$$\mathbb{P}[U^2 = \tfrac{1}{3}] = \mathbb{P}[U = \tfrac{1}{\sqrt{3}}] + \mathbb{P}[U = -\tfrac{1}{\sqrt{3}}] = 0,$$

contradiction!

(*) A second solution:

Let's take $g(x) = \cos(2\pi x), h(x) = \sin(2\pi x)$. Let $X = g(U), Y = h(U)$. We calculate $\mathrm{Cov}(X, Y)$:

Since U is absolutely continuous,

$$\begin{aligned}\mathbb{E}[XY] = \mathbb{E}[\cos(2\pi U)\sin(2\pi U)] &= \int_{-\infty}^{\infty} \cos(2\pi t)\sin(2\pi t) f_U(t)dt \\ &= \int_{-1}^{1} \cos(2\pi t)\sin(2\pi t)\tfrac{1}{2}dt = \tfrac{1}{4}\int_{-1}^{1} \sin(4\pi t)dt \\ &= -\tfrac{1}{16\pi}\cos(4\pi t)\Big|_{-1}^{1} = 0.\end{aligned}$$

Also,

$$\mathbb{E}[X] = \int_{-1}^{1} \cos(2\pi t)\tfrac{1}{2}dt = -\tfrac{1}{4\pi}\sin(2\pi t)\Big|_{-1}^{1} = 0.$$

So

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = 0.$$

Hence, X, Y are uncorrelated.

We claim that X, Y are *not* independent. Indeed, assume for a contradiction that they were independent. Then X^2, Y^2 are also independent. However, $X^2 + Y^2 = 1$ by definition, so

$$\mathbb{P}[X^2 \le \tfrac{1}{4}, Y^2 \le \tfrac{1}{4}] = 0 \ne \mathbb{P}[X^2 \le \tfrac{1}{4}] \cdot \mathbb{P}[Y^2 \le \tfrac{1}{4}],$$

because

$$\begin{aligned}\mathbb{P}[X^2 \le \tfrac{1}{4}] = \mathbb{P}[\cos(2\pi U) \in [-\tfrac{1}{2}, \tfrac{1}{2}]] > 0, \\ \mathbb{P}[Y^2 \le \tfrac{1}{4}] = \mathbb{P}[\sin(2\pi U) \in [-\tfrac{1}{2}, \tfrac{1}{2}]] > 0.\end{aligned}$$

(C) This is the easiest case. Just take $g(x) = h(x) = x$. If $\mathrm{Cov}(g(U), h(U)) = 0$ this implies that $\mathrm{Var}[U] = 0$ which implies that $U = \mathbb{E}[U] = 0$ a.s. Since U is not constant, this is impossible.

Solution Q3:

(A) Let X_k denote the amount the gambler wins at game k. So $(X_k)_k$ are independent, and $\mathbb{P}[X_k = 1] = p, \mathbb{P}[X_k = -1] = q$.

At game k the gambler can either win or loose one Shekel. Thus, after $2k$ games, the gambler must have an even number of Shekel, and after $2k + 1$ games he must have an odd number of Shekel.

That is

$$\sum_{j=1}^{2k} X_j \text{ is even} \qquad \sum_{j=1}^{2k+1} X_j \text{ is odd .}$$

Set $S_k = \sum_{j=1}^{k} X_j$. This is the total winnings up to game k.

Consider the event A_k that for all $j \leq k$ we have $S_{2j} = 0$. On this event, since S_{2j+1} are always odd, it must be that $S_{2j+1} \in \{-1, 1\}$ for all $j \leq k$. Thus, A_k implies that the gambler has not left after game $2k$.

On the other hand, if the gambler has not left after game $2k$, it must be that $S_{2j} = 0$ for all $j \leq k$, otherwise one of these would be -2 or 2, and he would have left.

We have shown that A_k is the event that the gambler has not left after game $2k$, which is exactly the event $\{X > 2k\}$.

Also, note that $A_k \in \sigma(X_1, \ldots, X_{2k})$. Thus, X_{2k+1}, X_{2k+2} are independent of A_k. So,

$$\begin{aligned}
\mathbb{P}[A_{k+1}] &= \mathbb{P}[S_{2(k+1)} = 0 \ , \ \forall \ j \leq k \ S_{2j} = 0] = \mathbb{P}[X_{2k+1} + X_{2k+2} = 0 \ , \ \forall \ j \leq k \ S_{2j} = 0] \\
&= \mathbb{P}[X_{2k+1} + X_{2k+2} = 0] \cdot \mathbb{P}[A_k] \\
&= \mathbb{P}[X_{2k+1} = 1, X_{2k+2} = -1] \cdot \mathbb{P}[A_k] + \mathbb{P}[X_{2k+1} = -1, X_{2k+2} = 1] \cdot \mathbb{P}[A_k] \\
&= 2pq \, \mathbb{P}[A_k].
\end{aligned}$$

Continuing inductively, we get that

$$\mathbb{P}[X > 2k] = \mathbb{P}[A_k] = (2pq)^k.$$

Note that the gambler can only leave after an even number of games, because after an odd number of games the gambler has an odd number of Shekel.

Thus,

$$R_X = \{2k \ : \ k = 1, 2, \ldots\} \qquad \text{and} \qquad \mathbb{P}[X > 2k] = (2pq)^k.$$

If we define $Y = X/2$ we get that $R_Y = \{1, 2, \ldots, \}$ and $\mathbb{P}[Y > k] = (2pq)^k$. So $Y \sim \text{Geo}(1 - 2pq)$.

Thus, for $k = 1, 2, \ldots,$

$$\mathbb{P}[X = 2k] = \mathbb{P}[Y = k] = (1 - 2pq)(2pq)^{k-1}$$

and $\mathbb{P}[X = t] = 0$ for any other t.

(*) Without recognizing the geometric distribution this is still not hard: for any positive integer k,

$$\mathbb{P}[X = 2k] = \mathbb{P}[\{X > 2(k-1)\} \setminus \{X > 2k\}] = \mathbb{P}[X > 2(k-1)] - \mathbb{P}[X > 2k]$$
$$= (2pq)^{k-1} - (2pq)^k = (1 - 2pq)(2pq)^{k-1}.$$

(B) The event that the gambler leaves with 2 Shekel, is the event that at the last two games the gambler earned 1 Shekel in each. Using the law of total probability,

$$\mathbb{P}[\text{ gambler leaves with 2 Shekel }] = \sum_{k=1}^{\infty} \mathbb{P}[X = 2k, \text{ gambler leaves with 2 Shekel }]$$
$$= \sum_{k=1}^{\infty} \mathbb{P}[X > 2(k-1), X_{2k-1} = 1, X_{2k} = 1].$$

Since $\{X > 2(k-1)\} \in \sigma(X_1, \dots, X_{2(k-1)})$ we have by independence,

$$\mathbb{P}[\text{ gambler leaves with 2 Shekel }] = \sum_{k=1}^{\infty} \mathbb{P}[X > 2(k-1)] \cdot p^2$$

$$= \sum_{k=1}^{\infty} (2pq)^{k-1} p^2 = \frac{p^2}{1-2pq}.$$