Problem 4

Solution.

a. The sample spaces consists of four outcomes, one of which is going "bust". Assume that the probability of each outcome is equal. If W denotes the number of trials before the "bust" outcome occurs, then plausibly infinitely many trials can occur and "bust" would still not have happened.

Then the range of the not-"bust" outcomes R_W is $[0, \infty)$, where $R_W(i) \in \mathbb{Z}$.

Denote i as the outcome of going "bust" on an arbitrary trial out of n total trials. Then $P(X \neq i) = (1 - p_i)^n$. And since the probability of going "bust" is $\frac{1}{4}$, we have that $P(X \neq i) = (1 - \frac{1}{4})^n = (\frac{3}{4})^n$. The PMF for the number of successful spins then is $\frac{1}{4}(\frac{3}{4})^n$ since we go bust on the $n + 1$-th trial.

b. We have that the expected winnings after one trial with the multidimensional variable X is:

$$\mathbb{E}[X_1] = \frac{\frac{1}{4}(\$100) + \frac{1}{4}(\$500) + \frac{1}{4}(\$5000) + \frac{1}{4}(\$0)}{1} = \$25 + \$125 + \$1250 + \$0 = \$1400$$

c. We are given that S_n denotes the cumulative amount of money won after n spins and want to find $\mathbb{E}[X_{n+1}|S_n]$. Going bust eliminates all previous earnings, and with a quarter chance of occurring we must remove $\frac{1}{4} \cdot S_n$ from our cumulative earnings, $(500 \cdot 0.25) + (100 \cdot 0.25) + (5000 \cdot 0.25) = \$1400$. So we have that

$$\mathbb{E}[X_{n+1}|S_n] = 1400 - \frac{S_n}{4}$$

d. We have that,

$$1400 \cdot 4 = 5600$$

and can see that

$$\mathbb{E}[X_{n+1}|5600] = 1400 - \frac{5600}{4} = 0$$

so

$$\mathbb{E}[X_{n+1}|S_n] < 0 \text{ when } S_n > 5600$$

■

Problem 5

Solution.

a. The unique value $k = 1$ makes

$$p_X(x) = \begin{cases} k(1-x), & x \geq 0 \\ k(1+x), & x < 0 \end{cases}$$

a valid density function on the interval [-1,1].

This is such because the area of this symmetric triangular density is 1, as are all probability densities, and we are given that the length of the base is 2 (being on the interval [-1,1]), so using the formula for the area of a triangle

$$a = \frac{1}{2}bh$$

where b, h, and a denote base, height, and area respectively, we have that

$$1 = \frac{1}{2}(2)h$$

$$1 = h$$

Given that $p_X(x)$ is centered at zero and $h = 1$, we know that $p_X(0) = 1$ and since $p_x(x) = k(1-x)$, $x \geq 0$

$$p_X(0) = k(1-0)$$

$$1 = k(1)$$

$$1 = k$$

b. We want to find $P(|X| > \frac{1}{2})$, which is analogous to finding the sum of $P(X > \frac{1}{2})$ and $P(X < -\frac{1}{2})$. Lines drawn at the points $\frac{1}{2}$ and $-\frac{1}{2}$ on the interval [-1,1] bisect the right and left hypotenuses of the symmetric triangular distribution (I am not going to prove this here). This means that the half of the right and left hypotenuses that lie beyond $\frac{1}{2}$ and $-\frac{1}{2}$, respectively, on the interval [-1,1] are each of length $\frac{\sqrt{2}}{2}$ (since the length of the entire hypotenuse on either side is $\sqrt{2}$). So, the height of each of these sub-triangular distributions of the larger triangular distribution on the interval [-1,1] with peak at 0 is

$$\left(\frac{1}{2}\right)^2 + h^2 = \left(\frac{\sqrt{2}}{2}\right)^2$$

$$\frac{1}{4} + h^2 = \frac{2}{4}$$

$$h^2 = \frac{1}{4}$$

$$h = \sqrt{\frac{1}{4}} = \frac{\sqrt{1}}{\sqrt{4}} = \frac{1}{2}$$

So, the area (using the formula for area of a triangle) of these two sub-triangular distributions is

$$2 \cdot \left(\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}\right) = \frac{1}{4}$$

since each base $b = \frac{1}{2}$ and each height $h = \frac{1}{2}$. There are two sub-triangular distributions, hence the 2 at the beginning. So, $P(|X| > \frac{1}{2}) = \frac{1}{4}$.

■

Problem 6

Solution.

a. We are given a circle with radius r and center (0,0); the equation for this circle is $x^2 + y^2 = r^2$. The upper half of the circle, or $y > 0$, is given by $g(x) = \sqrt{r^2 - x^2}$ where $y = g(x)$. The values of x for which $g(x)$ exists are determined by the value of $(r^2 - x^2)$. We know that $g(x) > 0$ (so $g(x) \neq 0$) and that $\sqrt{-n} \in \mathbb{C}$, where $n \in \mathbb{R}$ (so $(r^2 - x^2) > 0$) . So, $g(x)$ exists $\forall x : |x| < |r|$.

b. The area under the curve of any probability distribution is 1. This is given by

$$\int_{\omega \in \Omega} p(\omega) d\omega = 1$$

For the semi-circle given by $g(x) = \sqrt{r^2 + x^2}$ on the interval $[-r, r]$, we have that

$$p(x) = \int_{-r}^{r} k(r) g(x) dx$$

$$p(x) = k(r) \int_{-r}^{r} g(x) dx$$

$$1 = k(r) \frac{\pi r^2}{2} \text{ (area of semi-circle is area of a circle halved)}$$

$$\frac{2}{\pi r^2} = k(r)$$

■

Problem 7

Solution.

The probability density function of random variable Y with the beta distribution on the interval $[0, 1]$ is given here:

$$p_Y(y) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha - 1} (1 - y)^{\beta - 1}$$

We proceed to find the expected value of Y:

$$\mathbb{E}[Y] = \int_0^1 y p_Y dy$$

$$\mathbb{E}[Y] = \int_0^1 y \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1} dy$$

$$\mathbb{E}[Y] = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 y y^{\alpha-1}(1-y)^{\beta-1} dy$$

$$\mathbb{E}[Y] = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 y^{\alpha}(1-y)^{\beta-1} dy$$

$$\mathbb{E}[Y] = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+1+\beta)}$$

$$\mathbb{E}[Y] = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+1+\beta)}$$

$$\mathbb{E}[Y] = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+1)}{\Gamma(\alpha)\Gamma(\alpha+1+\beta)}$$

$$\mathbb{E}[Y] = \frac{\Gamma(\alpha+\beta)\alpha\Gamma(\alpha)}{\Gamma(\alpha)(\alpha+\beta)\Gamma(\alpha+\beta)}$$

$$\mathbb{E}[Y] = \frac{\alpha}{\alpha+\beta}$$

■

Problem 8

Solution.

a. We want to find $p_Z(z) := P(X+Y=z)$ and have that $R_X = R_Y = \{1,2,3,4\}$. The sum of the distributions for $p_X(x) = \sum_i P(X=i)$ and $p_Y(x) = \sum_i P(Y=i)$ forms the distribution for z, equal to (well, at least represented one way because there are other ways of representing $p_Z(z)$):

$$p_Z(z) = \sum_{i \in R_X} P(X=i \text{ and } Y=z-i)$$

since X and Y are independent this equals

$$p_Z(z) = \sum_{i \in R_X} P(X=i)P(Y=z-i)$$

this is equivalent to:

$$p_Z(z) = \sum_{i \in R_X} p_X(i) p_Y(z-i)$$

We have that $R_Z = \{2,3,4,5,6,7,8\}$ (since the ranges of X and Y, which are summed, must also be summed) and still have that $R_X = \{1,2,3,4\}$. Evaluating, we get:

$$p_Z(2) = \sum_{i \in R_X} p_X(i)p_Y(2-i) = \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot 0\right) + \left(\frac{1}{4}\cdot 0\right) + \left(\frac{1}{4}\cdot 0\right) = \frac{1}{16}$$

$$p_Z(3) = \sum_{i \in R_X} p_X(i)p_Y(3-i) = \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot 0\right) + \left(\frac{1}{4}\cdot 0\right) = \frac{1}{16} + \frac{1}{16} = \frac{1}{8}$$

$$p_Z(4) = \sum_{i \in R_X} p_X(i)p_Y(4-i) = \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot 0\right) = \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{3}{16}$$

$$p_Z(5) = \sum_{i \in R_X} p_X(i)p_Y(5-i) = \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) = \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{1}{4}$$

$$p_Z(6) = \sum_{i \in R_X} p_X(i)p_Y(6-i) = \left(\frac{1}{4}\cdot 0\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) = \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{3}{16}$$

$$p_Z(7) = \sum_{i \in R_X} p_X(i)p_Y(7-i) = \left(\frac{1}{4}\cdot 0\right) + \left(\frac{1}{4}\cdot 0\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) = \frac{1}{16} + \frac{1}{16} = \frac{1}{8}$$

$$p_Z(8) = \sum_{i \in R_X} p_X(i)p_Y(8-i) = \left(\frac{1}{4}\cdot 0\right) + \left(\frac{1}{4}\cdot 0\right) + \left(\frac{1}{4}\cdot 0\right) + \left(\frac{1}{4}\cdot\frac{1}{4}\right) = \frac{1}{16}$$

We have that:

$$p_Z(z) = \begin{cases} \frac{1}{16}, & \text{if } z = 2 \text{ or } z = 8 \\ \frac{1}{8}, & \text{if } z = 3 \text{ or } z = 7 \\ \frac{3}{16}, & \text{if } z = 4 \text{ or } z = 6 \\ \frac{1}{4}, & \text{if } z = 5 \end{cases}$$

b. Here are the results for 10000 iterations of programmatical experimentation using $R_X = R_Y = \{1,2,3,4\}$ and that $P(X+Y=z)$.
Value: 2
Number: 675
Percent Total: 6.75%

Value: 3
Number: 1252
Percent Total: 12.52%

Value: 4
Number: 1945
Percent Total: 19.45%

Value: 5
Number: 2466
Percent Total: 24.66%

Value: 6
Number: 1877
Percent Total: 18.77%

Value: 7
Number: 1191
Percent Total: 11.91%

Value: 8
Number: 594
Percent Total: 5.94%
See repository for code

■

Problem 9

Solution. We have that

$$\mathbb{E}[Y] = \sum_y yP(Y = y)$$

and that

$$\mathbb{E}[Y|X] = \sum_y yP(Y = y|X)$$

and we want to show that $\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y]$.
Proceeding, we can see that

$$\mathbb{E}[\mathbb{E}[Y|X]] = \sum_x \mathbb{E}[Y|X = x]P(X = y = x)$$

So,

$$\sum_x \mathbb{E}[Y|X = x]P(X = x) = \sum_x \left(\sum_y yP(Y = y|X = x)\right)P(X = x)$$

$$\sum_x \mathbb{E}[Y|X=x]P(X=x) = \sum_x \left(\sum_y y \frac{P(Y=y \text{ and } X=x)}{P(X=x)} \right) P(X=x)$$

$$\sum_x \mathbb{E}[Y|X=x]P(X=x) = \sum_x \left(\sum_y y \frac{P(Y=y \text{ and } X=x)}{P(X=x)} \right) P(X=x)$$

$$\sum_x \mathbb{E}[Y|X=x]P(X=x) = \sum_x \sum_y y P(Y=y \text{ and } X=x)$$

$$\sum_x \mathbb{E}[Y|X=x]P(X=x) = \sum_y y \left(\sum_x P(Y=y \text{ and } X=x) \right)$$

$$\sum_x \mathbb{E}[Y|X=x]P(X=x) = \sum_y y P(Y=y)$$

which is that

$$\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y]$$

■

We affirm that we have adhered to the Honor Code in this assignment.