CSE 312 HW5

Rohan Mukherjee

November 6, 2023

1. (a) Since X is the face value of a fair dice flip, with sides from 1 through 12, it is discrete and $X \sim \text{Unif}(1, 12)$, and hence $\mathbb{E}[X] = \frac{12+1}{2}$, and $\text{Var}(X) = \frac{(12-1)(12-1+2)}{12}$.

   (b) Let Y_i be the value of the ith dice. By linearity of expectation, $\mathbb{E}[Y] = \sum_{i=1}^{5} \mathbb{E}[Y_i] = \sum_{i=1}^{5} \mathbb{E}[Y_1]$ (identically distributed). We found last time that $\mathbb{E}[Y_1] = \frac{13}{2}$, so $\mathbb{E}[Y] = \frac{65}{2}$. We also know that $\text{Var}(Y) = \text{Var}(\sum_{i=1}^{5} Y_i)$ (since $Y = \sum_{i=1}^{5} Y_i$). Since the Y_i's are i.i.d. random variables, (this is precisely where I use independence), we are allowed to move the sum out of the variance, and we get that $\text{Var}(Y) = \sum_{i=1}^{5} \text{Var}(Y_i) = \sum_{i=1}^{5} \text{Var}(Y_1) = 5 \cdot \frac{11 \cdot 13}{12}$.

   (c) We do almost the same thing. Using the notation from last time, $Z = \frac{1}{5} \sum_{i=1}^{5} Y_i$, so by linearity of expectation, we have that $\mathbb{E}[Z] = \frac{1}{5} \sum_{i=1}^{5} \mathbb{E}[Y_i] = \frac{1}{5} \sum_{i=1}^{5} \mathbb{E}[Y_1] = \mathbb{E}[Y_1] = \frac{13}{2}$. We also know that $\text{Var}(Z) = \text{Var}(\frac{1}{5} \sum_{i=1}^{5} Y_i)$. First we use the property of variance that $\text{Var}(aX) = a^2\text{Var}(X)$ to get that this equals $\frac{1}{25}\text{Var}(\sum_{i=1}^{5} Y_i)$. Now, since the Y_i's are i.i.d. (this is precisely where I use independence), we can pull the sum out of the variance and then look to our answer from last question to get that this equals $\frac{1}{5} \cdot \frac{11 \cdot 13}{12}$.

2. Since the score starts off tied, what we know is that we need the last two points to be won by the same team (one can prove this inductively, the first time a team can win is after 2 turns, in which case the winning team would need to score both wins. Now suppose it's true for n. By the inductive hypothesis, the teams are at the same score, and now we can apply the base case to see one team would need to win the next two points to win). So, we may think of this as a geometric variable, with the probability in the geometric variable being the probability that some team scores two points in a row. The probability the first team wins two in a row is p^2, and the probability the second team wins two in a row is $(1-p)^2$. So our geometric variable would have probability $p^2 + (1-p)^2$. So, let $X =$ the number of 2-rounds up to an including the first time a team scores two points in a row. By our discussion above, $X \sim \text{Geo}(p^2 + (1-p)^2)$. $\mathbb{E}[X]$, from the zoo, is now $\frac{1}{p^2+(1-p)^2}$. But, this was the number of double-rounds played, so we need to double this number to get the number of rounds played, so we conclude that our answer equals $\frac{2}{p^2+(1-p)^2}$. I also did the infinite sum way and got the same answer, but I agree this was is much nicer. It took a while to fix that answer (no need to be more confusing!).

3. (a) We need

$$\int_{\mathbb{R}} f_X(x)dx = 1$$

Plugging this in gives (Note: $f_X(x) = 0$ for any $x \notin [-1/2, 1/2]$)

$$\int_{-1/2}^{1/2} r(1 - 3x^2)dx = 1$$

We see that, since $1 - 3x^2$ is an even function, and since we can pull constants out,

$$r \int_{-1/2}^{1/2} (1 - 3x^2)dx = 2r \int_0^{1/2} 1 - 3x^2 dx = 2r(x - x^3)\Big|_0^{1/2} = 2r(1/2 - 1/8) = \frac{3r}{4}$$

So we need $\frac{3r}{4} = 1$, which says that $r = \frac{4}{3}$.

(b) $F_X(k) = \int_{-\infty}^k f_X(x)dx$, so we have to evaluate this definite integral. Since $f_X(x) = 0$ for $x < -1/2$, if $k < -1/2$, then $f_X(k) = 0$ (we are integrating the 0 function in that case). We showed in part (a) that if $k > 1/2$, then $F_X(k) = 1$ (Indeed: $\int_{-\infty}^k f_X(x)dx = \int_{-1/2}^k f_X(x)dx = \int_{-1/2}^{1/2} f_X(x)dx = 1$), so we just have to evaluate this integral for $-1/2 \le k \le 1/2$:

$$\int_{-1/2}^k f_X(x)dx = \frac{4}{3} \int_{-1/2}^k (1 - 3x^2)dx = \frac{4}{3}(x - x^3)\Big|_{-1/2}^k$$

$$= \frac{4}{3}\Big[(k - k^3) - (-1/2 - (-1/2)^3)\Big]$$

$$= \frac{4}{3}\Big[(k - k^3) + 3/8\Big]$$

So, $F_X(k) = \begin{cases} 0, & x < -1/2 \\ \frac{4}{3}[(k - k^3) + 3/8], & -1/2 \le x \le 1/2 \\ 1, & x > 1/2 \end{cases}$

3

4. (a) We know that $F_X(x) = \mathbb{P}(X \leq x)$. Since the flea always lands in the ball, if $x > c$ then the fleas distance from the center is certainly less than than or equal to c, so $F_X(x) = 1$ for all $x > c$. Similarly, the flea cannot have negative distance to the center, so $F_X(x) = 0$ for $x < 0$. We are left to evaluate $\mathbb{P}(X \leq x)$ for $0 \leq x \leq c$. The flea will have distance less than or equal to x when it lands in the ball of radius x, so we can find the probability by just taking the ratio of the volumes. The volume of the ball with radius x is $\frac{4}{3}\pi x^3$, and the volume of the entire ball is $\frac{4}{3}\pi c^3$. We conclude that for $0 \leq c \leq x$,

$$F_X(x) = \frac{\frac{4}{3}\pi x^3}{\frac{4}{3}\pi c^3} = \frac{x^3}{c^3}$$

Quite a spectacular answer! So the CDF in its entirety equals

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{x^3}{c^3}, & 0 \leq x \leq c \\ 1, & c < x \end{cases}$$

(b) We can find the pdf from the cdf by just taking derivatives, so the pdf, $f_X(x)$ equals

$$f_x(X) = \begin{cases} 0, & x < 0 \\ \frac{3x^2}{c^3}, & 0 \leq x \leq c \\ 0, & c < x \end{cases}$$

(c) The expected value of X is $\mathbb{E}[X] = \int_{\mathbb{R}} x f_X(x) dx$. Plugging everything in, and using that $f_X(x) = 0$ for $x \notin [0, c]$, we see that

$$\int_{\mathbb{R}} x f_X(x) dx = \int_0^c \frac{3x^3}{c^3} dx = \frac{3}{4c^3} x^4 \Big|_0^c = \frac{3}{4}c$$

Which is quite a spectacular answer! Very intuitive, since there is more volume near the outside of the ball, so the flea on average should be closer to the outside, which is reflected by the number.

(d) The variance is equal to $\mathbb{E}[X^2] - \mathbb{E}[X]^2$. We found $\mathbb{E}[X]$, above, so we are left to evaluate $\mathbb{E}[X^2]$. By the law of an unconscious statistician, and from the same reasoning about the pdf as last time,

$$\mathbb{E}[X^2] = \int_{\mathbb{R}} x^2 f_X(x) dx = \int_0^c \frac{3x^4}{c^3} dx = \frac{3}{5c^3} x^5 \Big|_0^c = \frac{3}{5}c^2$$

We conclude that $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{3}{5}c^2 - \frac{9}{16}c^2 = \frac{3}{80}c^2$.

5. (a) I asked on the message board and they confirmed that $\mathbb{P}(X < m) = \mathbb{P}(X \leq m)$, so we are looking for $\mathbb{P}(X \leq m) = 1 - \mathbb{P}(X \leq m)$ by complimentary counting. This in turn says that $\mathbb{P}(X \leq m) = 1/2$, which, plugging in the CDF for exponential gives us $1 - e^{-\frac{3}{r}m} = \frac{1}{2}$ (We may assume that $m \geq 0$, since if $m < 0$, then $F_X(m) = 0$, i.e. $\mathbb{P}(X \leq m) = 0 \neq \frac{1}{2}$). Solving this equation gives us $-\ln(2) = -\frac{3}{r}m$, or that $m = \frac{r\ln(2)}{3}$.

(b) For $r = 12$, plugging this into the formula above gives us $m = \frac{12\ln(2)}{3} \approx 2.773$.

(c) The probability that you miss the target completely, since the target is perfectly circular with radius r, is going to be $\mathbb{P}(X > r)$. This of course equals $1 - \mathbb{P}(X \leq r)$. Using the exponential CDF, we see that $\mathbb{P}(X \leq r) = 1 - e^{-\frac{3}{r} \cdot r} = 1 - e^{-3}$. Therefore, $\mathbb{P}(X > r) = e^{-3} \approx 0.0498$, which is surprisingly high.