Recitation 4: Random Variables

1 Review

1. A discrete random variable X can be characterized by its *probability mass function* (PMF)

$$f(x) = P(X = x)$$

or by its *cumulative distribution function* (CDF)

$$F(x) = P(X \le x) = \sum_{y \le x} f(y)$$

2. The *expectation* of a discrete random variable X is

$$E[X] = \sum_{x \in S} x\, f(x)$$

where $S = \{x | P(X = x) \neq 0\}$ i.e. "every possible value X can take."

3. The *variance* of a random variable (discrete or continuous):

$$\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

which has the property that

$$\text{Var}(aX + b) = a^2 \text{Var}(X)$$

2 Exercises

1. A random variable X has the following probability mass function (PMF):

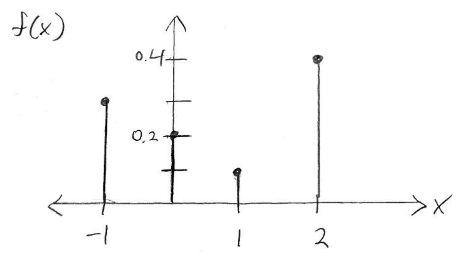
$$f(x) = \begin{cases} 0.3 & x = -1 \\ 0.2 & x = 0 \\ 0.1 & x = 1 \\ 0.4 & x = 2 \end{cases}.$$

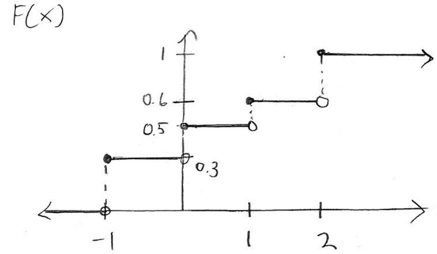
(a) Find $F(x)$, its cumulative distribution function (CDF). Remember that the CDF is defined for all real numbers x.

Using the definition of the CDF, we find that

$$F(x) = \begin{cases} 0 & x < -1 \\ 0.3 & -1 \le x < 0 \\ 0.5 & 0 \le x < 1 \\ 0.6 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

(b) Sketch both the PMF and the CDF of X.
We get the following sketches:





(c) Are the events $A = \{X \geq 0\}$ and $B = \{|X| \text{ is odd}\}$ independent? Why or why not?
Note that

$$\begin{aligned} &P(A \cap B) = P(X = 1) = 0.1; \\ &P(A) = P(X = 0) + P(X = 1) + P(X = 2) = 0.7; \\ &P(B) = P(X = -1) + P(X = 1) = 0.4. \end{aligned}$$

Since $P(A \cap B) \neq P(A)P(B)$, events A and B are not independent.

(d) Compute $E[X]$.
We have that

$$E[X] = \sum_{x=-1}^{2} xP(X = x) = -1 \cdot 0.3 + 0 \cdot 0.2 + 1 \cdot 0.1 + 2 \cdot 0.4 = 0.6.$$

2. A *binomial random variable*, denoted by $X \sim \text{Binomial}(n, p)$, represents the number of success obtained from n independent trials, where each trial is successful with probability p. The random variable X has the PMF

$$f_X(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \ldots, n$$

Values from the PMF can be computed using R's `dbinom(x, n, p)` function. For instance, `dbinom(6, 10, 0.5)` gives the probability of getting *exactly* 6 heads in 10 flips of a fair coin.

(a) A certain player in the NBA made 39.7% of his free throws last season. Compute the probability that he makes exactly 6 out of his next 10 free throws. Compute this probability by hand, then verify your answer using R.
Let X be the number of free throws he made in his next 10 trials. Then, X is binomial distributed, i.e., $X \sim \text{Binomial}(10, 0.397)$.

$$P(X = 6) = \binom{10}{6} 0.397^6 (1 - 0.397)^4 = 0.109.$$

(b) Write the probability that he makes between 4 and 9 of his next 10 free throws (inclusive) in terms of the CDF of the binomial random variable:

$$F(x) = P(X \leq x) = \sum_{y=0}^{x} \binom{n}{y} p^y (1-p)^{n-y}.$$

Do not evaluate the expression you obtain.

$$\begin{aligned} P(4 \leq X \leq 9) &= P(X \leq 9) - P(X \leq 3) \\ &= \sum_{y=0}^{9} \binom{10}{y} 0.397^y 0.603^{10-y} - \sum_{y=0}^{3} \binom{10}{y} 0.397^y 0.603^{10-y} \\ &= \sum_{y=4}^{9} \binom{10}{y} 0.397^y 0.603^{10-y}. \end{aligned}$$

(c) Use R to compute the probability from part (b). R's `pbinom(x, n, p)` may be useful here. (For example, `pbinom(6, 10, 0.5)` gives the probability of getting *at most* 6 heads in 10 flips of a fair coin.) Use R, we get $P(4 \leq X \leq 9) = 0.61$.

(d) What is the probability that he will make at least 1 shot in 5 tries? (**Hint**: What is the probability that he won't?)
Redefine X to be the number of free throws he made in 5 trials.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{5}{0} 0.397^0 (1 - 0.397)^5 = 0.92.$$

(e) How many free throws would you expect him to make in 10 attempts?
Use the definition of X in part (a),

$$E[X] = \sum_{y=0}^{10} y \binom{10}{y} 0.397^y 0.603^{10-y}$$

Use R to evaluate this expression, we obtain $E[X] = 3.97$.

Note that in R, function names for the pmfs and cdfs of random variables follow the same pattern as `dbinom` and `pbinom`. For example, you would use `dgeom` and `pgeom` for geometric random variables, and `dpois` and `ppois` for Poisson random variables.

3. Roulette is one of the most popular casino games played across the world. In this problem we will examine the probabilities associated with this "game of chance".

The game consists of spinning a wheel containing slots for the numbers 1 to 36, and the numbers 0 and 00. Each of the numbers 1 to 36 is coloured either red or black (with exactly 18 red numbers and 18 black numbers), and the other two are coloured green. A ball is tossed around the wheel and eventually lands in one of the slots, selecting that number. It is assumed (and should be practically the case if the wheel is properly manufactured) that the outcome of one spin has no impact on the outcome of any other spin, i.e. that spins are independent, and that on each spin, every number is equally likely to be selected.

Gamblers play by betting on the outcome of each spin, by placing chips on a layout representing the possible outcomes (see last page). Bets can be made on individual numbers or on certain combinations, such as strings of consecutive numbers, low or high numbers, or the colour of the number.

Bets on individual numbers from 1 to 36 pay 35 to 1, so if \$1 is bet on, say, 12, and that number comes up, the gambler keeps the \$1 and wins an additional \$35. Bets on combinations of numbers are paid in such a way that you win the same betting on the combination as you would if you bet on each of the numbers in the combination individually, with the same total bet amount. For example, a bet of \$1 on $\{17, 18\}$ wins \$17, which is the same amount you would win if you bet 50 cents on both 17 and 18.[1] Similarly, a bet on Red pays 1 to 1, which is the same payoff you would see as if you bet individually on each of the 18 red numbers.

(a) Let the random variable X denote the outcome of one spin. X will take one of the values $\{1, 2, \dots, 36, 0, 00\}$. What is the probability $P(X = 12)$? What is the probability $P(X = x)$ for any of the possible outcomes x? What (named) type of discrete distribution does this correspond to?

$P(X = x) = 1/38$ for all x in the set. This is a discrete uniform distribution.

(b) Let W denote the random amount won if a bet of \$1 is placed on the number 12 (losses are considered negative winnings). Write W explicitly as a function of X. What is the expected winning $E(W)$? What is the variance $\text{Var}(W)$?

$W = 35 \cdot \mathbf{1}_{\{X=12\}} - 1 \cdot \mathbf{1}_{\{X \neq 12\}}$
$\mathbb{E}W = 35(1/38) - 1(37/38) = -2/38 = -0.0526$
$\text{Var}(W) = \mathbb{E}(W^2) - (\mathbb{E}W)^2 = 35^2(1/38) + 1^2(37/38) - (2/38)^2 = 33.21$

(c) If you bet \$1 on a combination bet consisting of n numbers from 1 to 36, how much will you get paid if you win (find a formula in terms of n)?

This is the same as betting $\frac{1}{n}$ for each of n numbers.
Winning means that one of these n numbers came up so the player gets $\frac{35}{n}$. At the same time the player looses $\frac{1}{n}$ on the rest $(n-1)$ numbers so the final amount is $\frac{35}{n} - \frac{n-1}{n} = \frac{36-n}{n}$.

(d) Let W' be the amount won if \$1 is placed on the combination {25,26,27}. What is $E(W')$? What is the expected winning from betting \$1 on any combination of n numbers out of $\{1, 2, \dots, 36\}$?

$\mathbb{E}W' = [(36-3)/3](3/38) - 1(35/38) = -2/38 = -0.0526$
$\mathbb{E}W_n = [(36-n)/n](n/38) - 1((38-n)/38) = -2/38 = -0.0526$

(e) Recall the winning amount W' from part (d). What is its variance? What is the variance of the winning amount W from betting \$1 on Red?

$\text{Var}(W') = \mathbb{E}(W'^2) - (\mathbb{E}W')^2 = 11^2(3/38) + 1^2(35/38) - (2/38)^2 = 10.47$
$\text{Var}(W) = \mathbb{E}(W^2) - (\mathbb{E}W)^2 = 1^2(18/38) + 1^2(20/38) - (2/38)^2 = 0.997$

(f) Let X_1 be the outcome of the first spin, X_2 the outcome of the second spin, and so on. Recall that we are assuming that the outcomes of different spins are independent. What is $P(X_1 \text{ is red})$? What is $P(X_2 \text{ is red})$? What then is $P(X_2 \text{ is red} \mid X_1 \text{ is red})$? Similarly what are
$P(X_9 \text{ is red} \mid X_1 \text{ is red}, X_2 \text{ is red}, \dots, X_8 \text{ is red})$ and

[1] E.g., if 17 comes up then you win \$35/2 = \$17.50, but lose the 50 cents you bet on 18, so your total winnings (not counting the 50 cents you got back from the 17) are \$17.50 - \$0.50 = \$17. The situation's analogous if 18 comes up.

$P(X_9 \text{ is black} \,|\, X_1 \text{ is red}, X_2 \text{ is red}, \ldots, X_8 \text{ is red})$? It is a common conception that after a run of reds, we are "due" for a black, so black would be more likely on the next spin than it would be normally. What can you say about this reasoning? Would you use this idea to guide your betting strategy?

$P(X_1 \text{ is red}) = P(X_2 \text{ is red}) = P(X_2 \text{ is red} \,|\, X_1 \text{ is red}) = 18/38 = 0.474$

$P(X_9 \text{ is red} \,|\, X_1 \text{ is red}, X_2 \text{ is red}, \ldots, X_8 \text{ is red}) = 18/38$ by independence

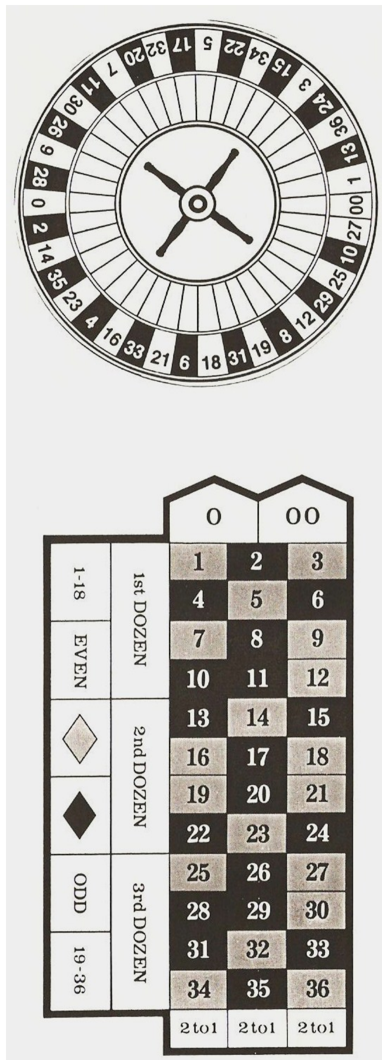
This reasoning is flawed, since the next spin is independent of the previous ones, so the next spin wouldn't "know" there was a run of reds and "try to make up for it".

(g) Assume you are making 10 bets in a row and all the bets place \$1 on red. What is the probability that at least two of these bets result in winning? What is the expected number of bets that will win?

From (g) we have $p = P(\text{spin is red}) = 0.474$

$$P(\text{at least 2 wins}) = 1 - P(\text{no wins or 1 win}) = 1 - \sum_{k=0}^{1} \binom{10}{k} p^k (1-p)^{10-k} =$$
$$= 1 - (1-0.474)^{10} - 10 \cdot 0.474 \cdot (1-0.474)^9 = 0.984$$

$\mathbb{E}(\#\text{of wins}) = np = 10 \cdot 0.474 = 4.74$


0
00
1st DOZEN
2nd DOZEN
3rd DOZEN
1-18
EVEN
ODD
19-36
2 to 1
2 to 1
2 to 1