Name:______________________________________ **Date:**____________________

PSTAT 5A: Discussion Worksheet 04

Spring 2023, with *Ethan P. Marzban*

1. In a particular textbook, it is found that each page contains a typo with probability 10%, independently across pages. Suppose 120 pages are selected (assume selection is done with replacement), and the number of pages that have typos is recorded.

   (a) Define the random variable of interest, and call it X.

   > **Solution:** Let X denote the number of pages in our sample of 120 that contain typos.

   (b) Does X follow the Binomial distribution? If so, state its parameters.

   > **Solution:** We check the four binomial criteria:
   >
   > (1) Independent trials? Yes, because we are sampling with replacement.
   >
   > (2) Fixed number of trials? Yes; we have $n = 120$ trials (since a 'trial' corresponds to picking a page and recording whether there is a typo or not).
   >
   > (3) Well-defined notion of success? Yes; a 'success' on any given trial is 'observing a typo on the selected page' (and a 'failure' on any given trial is 'not observing a typo on the selected page').
   >
   > (4) Constant probability of success? Yes; there is a $p = 0.1$ chance that any given page contains a typo.
   >
   > Since all four critera are satisfied, we conclude that X does follow the Binomial distribution; specifically, with parameters $n = 120$ and $p = 0.1$:
   >
   > $$X \sim \mathrm{Bin}(120, 0.1)$$

   (c) What is the probability that the textbook contains exactly 17 pages with typos?

   > **Solution:** We seek $\mathbb{P}(X = 17)$. Since $X \sim \mathrm{Bin}(120, 0.1)$, we have
   >
   > $$\mathbb{P}(X = k) = \binom{120}{k}(0.1)^k(1 - 0.1)^{120-k}$$

and so

$$\mathbb{P}(X = 17) = \binom{120}{17}(0.1)^{17}(0.9)^{103} \approx 0.03677 = 3.677\%$$

(d) What is the probability that the textbook contains between 16 and 18 (inclusive on both ends) pages with typos?

Solution: We seek $\mathbb{P}(16 \leq X \leq 18)$, which can be decomposed in the following manner:

$$\begin{aligned}
\mathbb{P}(16 \leq X \leq 18) &= \mathbb{P}(X = 16) + \mathbb{P}(X = 17) + \mathbb{P}(X = 18) \\
&= \sum_{k=16}^{18} \binom{120}{k}(0.1)^k(1-0.1)^{120-k} \\
&= \left[\binom{120}{16}(0.1)^{16}(1-0.1)^{120-16}\right] + \left[\binom{120}{17}(0.1)^{17}(1-0.1)^{120-17}\right] \\
&\qquad + \left[\binom{120}{18}(0.1)^{18}(1-0.1)^{120-18}\right] \\
&= 0.0541 + 0.0368 + 0.0234 = 0.1143 = 11.43\%
\end{aligned}$$

(e) What is the expected number of pages in the textbook that contain typos?

Solution: In general, if $X \sim \mathrm{Bin}(n, p)$ we know that $\mathbb{E}[X] = np$. Here, we have $n = 120$ and $p = 0.1$ meaning

$$\mathbb{E}[X] = (120) \cdot (0.1) = 12 \text{ pages}$$

(f) What is the variance of the number of pages in the textbook that contain typos?

Solution: In general, if $X \sim \mathrm{Bin}(n, p)$ we know that $\mathrm{Var}(X) = np(1-p)$. Here, we have $n = 120$ and $p = 0.1$ meaning

$$\mathbb{E}[X] = (120) \cdot (0.1) \cdot (0.9) = 10.8 \text{ pages}^2$$

2. Consider the following game: a fair six-sided die is rolled. If the number showing is 1 or 2, you win a dollar; if the number showing is 3, 4, or 5 you win 2 dollars; if the number showing is 6, you lose 1 dollar. Let W denote your net winnings after playing this game once.

(a) Write down the state space S_W of W.

Solution: The only possibilities are: we lose a dollar (i.e we gain $\$-1$), we gain a dollar, or we gain 2 dollars. Hence, since W denotes our net winnings,

$$S_W = \{-1,\ 1,\ 2\}$$

(b) Find the p.m.f. of W.

Solution: We examine each point in the state space separately:

- $\mathbb{P}(W = -1)$: we only lose a dollar if the die lands on 6. Because the die is fair, the probability that it lands on 6 is simply $1/6$; hence

$$\mathbb{P}(W = -1) = \mathbb{P}(\text{die lands on } 6) = \frac{1}{6}$$

- $\mathbb{P}(W = 1)$: we only gain a dollar if the die lands on either 1 or 2. Because the die is fair, the probability that it lands on 1 or 2 is simply $2/6 = 1/3$; hence

$$\mathbb{P}(W = 1) = \mathbb{P}(\text{die lands on 1 or 2}) = \frac{1}{3}$$

- $\mathbb{P}(W = 2)$: we only gain 2 dollars if the die lands on 3, 4, or 5. Because the die is fair, the probability that it lands on 3, 4, or 5 is simply $3/6 = 1/2$; hence

$$\mathbb{P}(W = 2) = \mathbb{P}(\text{die lands on 3, 4, or 5}) = \frac{1}{2}$$

Hence, putting everything together, the p.m.f. of W is

k	-1	1	2
$\mathbb{P}(W = k)$	$1/6$	$1/3$	$1/2$

As a quick check; $(1/6) + (1/3) + (1/2) = 1$, as we expected.

(c) What are your expected winnings after one round of the game?

Solution: We seek $\mathbb{E}[W]$, which we compute using the definition of Expected Value:

$$\mathbb{E}[W] = \sum_k k \cdot \mathbb{P}(W = k) = (-1) \cdot \mathbb{P}(W = -1) + (1) \cdot \mathbb{P}(W = 1) + (2) \cdot \mathbb{P}(W = 2)$$

$$= (-1) \cdot \left(\frac{1}{6} \right) + (1) \cdot \left(\frac{1}{3} \right) + (2) \cdot \left(\frac{1}{2} \right) = \boxed{\frac{7}{6} = \$1.1\overline{6}}$$

3. Let X be a random variable with the following p.m.f. (probability mass function):

$\boldsymbol{k}$	-3	-2	0.5	3	4
$\mathbb{P}(\boldsymbol{X = k})$	0.12	0.08	0.43	$\boldsymbol{a}$	0.27

(a) Find the value of a.

Solution: The probabilities in the p.m.f. must sum to 1; i.e.

$$(0.12) + (0.08) + (0.43) + (a) + (0.27) = 1$$

which means

$$a = 1 - (0.12 + 0.08 + 0.43 + 0.27) = \boxed{0.1}$$

(b) Compute $\mathbb{P}(X = 1)$.

Solution: Because $1 \notin S_X$, we know $\boxed{\mathbb{P}(X = 1) = 0}$.

(c) Compute $\mathbb{P}(X \leq 0.5)$.

Solution: $\mathbb{P}(X \leq 0.5) = \mathbb{P}(X = -3) + \mathbb{P}(X = -2) + \mathbb{P}(X = 0.5) = \boxed{0.63}$

Or, if we wanted to use the Complement Rule,

$$\mathbb{P}(X \leq 0.5) = 1 - \mathbb{P}(X > 0.5) = 1 - [\mathbb{P}(X = 3) + \mathbb{P}(X = 4)] = 0.63$$

(d) Compute $\mathbb{P}(X \leq 1)$.

Solution: $\mathbb{P}(X \leq 1) = \mathbb{P}(X = -3) + \mathbb{P}(X = -2) = 0.12 + 0.08 = \boxed{0.2}$

(e) Compute $\mathbb{E}[X]$.

Solution:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_k k \cdot \mathbb{P}(X = k) \\
&= (-3) \cdot \mathbb{P}(X = -3) + (-2) \cdot \mathbb{P}(X = -2) + (0.5) \cdot \mathbb{P}(X = 0.5) \\
&\qquad + (3) \cdot \mathbb{P}(X = 3) + (4) \cdot \mathbb{P}(X = 4) \\
&= (-3) \cdot (0.12) + (-2) \cdot (0.08) + (0.5) \cdot (0.43) + (3) \cdot (0.1) + (4) \cdot (0.27) \\
&= \boxed{1.075}
\end{aligned}$$

(f) Compute $\text{SD}(X)$.

Solution: If we use the second formula for variance, we first compute

$$\begin{aligned}
\sum_k k^2 \cdot \mathbb{P}(X = k) &= (-3)^2 \cdot \mathbb{P}(X = -3) + (-2)^2 \cdot \mathbb{P}(X = -2) + (0.5)^2 \cdot \mathbb{P}(X = 0.5) \\
&\qquad + (3)^2 \cdot \mathbb{P}(X = 3) + (4)^2 \cdot \mathbb{P}(X = 4) \\
&= (-3)^2 \cdot (0.12) + (-2)^2 \cdot (0.08) + (0.5)^2 \cdot (0.43) + (3)^2 \cdot (0.1) \\
&\qquad + (4)^2 \cdot (0.27) = 6.7275
\end{aligned}$$

Thus,

$$\text{Var}(X) = \left(\sum_k k^2 \cdot \mathbb{P}(X = k) \right) - (\mathbb{E}[X])^2 = 6.7275 - (1.075)^2 \approx 5.5719$$

which means $\text{SD}(X) = \sqrt{5.5719} \approx \boxed{2.3604}$. Using the first formula for variance would have also resulted in the same answer (and you are encouraged to try it out on your own!)

Name:______________________________ Date:________________

4. (**The Geometric Distribution**) Consider the following situation: suppose we toss a p−coin repeatedly, and we let X denote the number of tosses (including our final toss) until we observe our first heads. So, for example, the outcome

$$(T, T, T, H)$$

would correspond to a value of $X = 4$. If we assume independence across trials, then X is said to follow the **Geometric Distribution with parameter p**, notated $X \sim \text{Geom}(p)$.

(a) What is the state space of X? **Hint:** is there an upper limit to the values X can attain? What about a lower limit?

> **Solution:** To answer the second question in the hint: the lower limit is 1, as the "best-case" scenario is that our first toss lands heads and we are done with the experiment. On the other end, however, we could continue to get tails after tails, meaning the upper limit of S_X is actually ∞; as such, we write
>
> $$S_X = \{1, 2, 3, \cdots\}$$
>
> where the "$\cdots$" implies that the set goes to infinity.

(b) The event $\{X = k\}$ means "we observed our first heads on the k^{th} toss". Find $\mathbb{P}(X = k)$.

> **Solution:** As the problem says, the event $\{X = k\}$ means "we observed our first heads on the k^{th} toss" which is equivalent to "we had $(k - 1)$ tails followed by a heads on the k^{th} toss"; i.e.
>
> $$(\underbrace{T, T, \cdots, T}_{k \text{ tails}}, H)$$
>
> which has a probability $(1 - p)^{k-1} \cdot p$ of occurring. Hence,
>
> $$\mathbb{P}(X = k) = (1 - p)^{k-1} \cdot p$$

(c) It turns out that $\mathbb{E}[X] = 1/p$. If we toss a fair coin, what is the expected number of tosses (including the final toss) needed to observe our first heads?

> **Solution:** Letting Y denote the number of times we need to toss a fair (i.e. $p = 0.5$) coin in order to observe our first heads, we see that $Y \sim \text{Geom}(0.5)$ meaning
>
> $$\mathbb{E}[Y] = \frac{1}{p} = \frac{1}{0.5} = \text{2 tosses}$$