CS 70 Discrete Mathematics and Probability Theory
Summer 2016 Dinh, Psomas, and Ye HW 5

Due Monday July 25 at 1:59PM

1. **Box of marbles (2/3/5 points)**

   You are given two boxes: one of them containing 900 red marbles and 100 blue marbles, the other one contains 500 red marbles and 500 blue marbles.

   (a) If we pick one of the boxes randomly, and pick a marble what is the probability that it is blue?

   Answer: Let B be the event that the picked marble is blue, R be the event that it is red, A_1 be the event that the marble is picked from box 1, and A_2 be the event that the marble is picked from box 2. Therefore we want to calculate $Pr(B)$. By total probability

   $$Pr(B) = Pr(B|A_1)P(A_1) + Pr(B|A_2)Pr(A_2) = 0.5 \times 0.1 + 0.5 \times 0.5 = 0.3$$

   (b) If we see that the marble is blue, what is the probability that it is chosen from box 1?

   Answer: In this part, we want to find $Pr(A_1|B)$. By Bayes' rule

   $$Pr(A_1|B) = \frac{Pr(B|A_1)Pr(A_1)}{Pr(B|A_1)Pr(A_1) + Pr(B|A_2)Pr(A_2)} = \frac{0.1 \times 0.5}{0.5 \times 0.1 + 0.5 \times 0.5} = \frac{1}{6}$$

   (c) Suppose we pick one marble from box 1 and without looking at its color we put it aside. Then we pick another marble from box 1. What is the probability that the second marble is blue?

   Answer: Let B_1 be the event that first marble is blue, R_1 be the event that the first marble is red, and B_2 be the event that second marble is blue without looking at the color of first marble. We want to find $Pr(B_2)$. By total probability,

   $$Pr(B_2) = Pr(B_2|B_1)Pr(B_1) + Pr(B_2|R_1)Pr(R_1) = \frac{99}{999} \times 0.1 + \frac{100}{999} \times 0.9 = 0.1$$

   More generally, one can see that the probability that the n-th marble picked from box 1 is blue with probability 0.1. This is clear by symmetry: all the permutations of the 1000 marbles have the same probability, so the probability that the n-th marble is blue is the same as the probability that the first marble is blue.

2. **Bayesian Inference and Pancakes (3/3/3/3/3 points)**

   Vince is making golden-brown pancakes and you are hungry!

   (a) Vince serves up a stack of 3 pancakes, but he forgot to butter the pan! Pancake A is perfect (golden-brown on both sides), pancake B is burnt on one side, and pancake C is burnt on both sides. The top of the stack is burnt. What's the probability that the other side of the top pancake is also burnt? Justify your answer.

(b) Vince agrees that a burnt pancake on top of the stack looks un-appetizing, and suggests flipping the stack over. In the same situation as before, what's the probability that the pancake side touching the plate is burnt?

(c) Suppose Vince makes a stack of n pancakes such that x pancakes are burnt on both sides and y pancakes are burnt on one side. If the top of the stack is burnt, what's the probability that the other side of the top pancake is also burnt? What if the top of the stack is golden-brown? Justify your answer.

(d) You asked for chocolate chips, so Vince adds lots of chocolate chips to the batter. He makes a stack of m pancakes next to the stack of n pancakes from before. However, the k-th pancake ($1 \leq k \leq m$) in the new stack only has a k/m chance of having chocolate chips (independent from the rest of the pancakes). If you choose a pancake randomly from either stack, what's the probability that you get chocolate chips?

(e) Vince realizes that the top few pancakes in the new stack don't really have chocolate chips in them. He shifts the top 10 pancakes from that stack (those with the smallest chance of chocolate chips) to the old stack. Given you randomly choose a pancake and it has chocolate chips, what's the probability it came from the new stack?

Answer:

Event A: A golden-brown pancake is on top of the stack.

Event B: A half-burnt pancake is on top of the stack.

Event C: A fully-burnt pancake is on top of the stack.

Event D: The top side of the stack is burnt.

Event E: The bottom side of the stack is burnt.

(a) $\Pr[A] = \Pr[B] = \Pr[C] = \frac{1}{3}$; and $\{A,B,C\}$ partition the sample space.
$\Pr[D|A] = 0, \Pr[D|B] = 0.5, \Pr[D|C] = 1$;
$\Pr[C|D] = \frac{\Pr[D|C]\Pr[C]}{\Pr[D]} = \frac{1\left(\frac{1}{3}\right)}{0\left(\frac{1}{3}\right) + 0.5\left(\frac{1}{3}\right) + 1\left(\frac{1}{3}\right)} = \frac{2}{3}$.
Alternate: 6 sides, 3 are burnt, 2 of which correspond to Pancake C.

(b)

$$\begin{aligned}
\Pr[E|D] &= \Pr[E \cap (A \cup B \cup C)|D] \\
&= \Pr[(E \cap A) \cup (E \cap B) \cup (E \cap C)|D] \\
&= \Pr[E \cap A|D] + \Pr[E \cap B|D] + \Pr[E \cap C|D] \\
&= \frac{\Pr[E \cap A \cap D]}{\Pr[D]} + \frac{\Pr[E \cap B \cap D]}{\Pr[D]} + \frac{\Pr[E \cap C \cap D]}{\Pr[D]} \\
&= 0 + \Pr[E|B \cap D]\Pr[B|D] + \Pr[E|C \cap D]\Pr[C|D] \\
&= 0 + \frac{1}{2} \times \frac{1}{3} + \frac{1}{4} \times \frac{2}{3} \\
&= \frac{1}{3}
\end{aligned}$$

(c) $\Pr[C|D] = \dfrac{\Pr[D|C]\Pr[C]}{\Pr[D]} = \dfrac{1\left(\frac{x}{n}\right)}{0\left(\frac{n-x-y}{n}\right)+0.5\left(\frac{y}{n}\right)+1\left(\frac{x}{n}\right)} = \dfrac{2x}{2x+y}$;

$\Pr[B|\overline{D}] = \dfrac{\Pr[\overline{D}|B]\Pr[B]}{\Pr[\overline{D}]} = \dfrac{0.5\left(\frac{y}{n}\right)}{1\left(\frac{n-x-y}{n}\right)+0.5\left(\frac{y}{n}\right)+0\left(\frac{x}{n}\right)} = \dfrac{y}{2n-2x-y}$.

(d) Event F: You picked from the new stack.

Event G: You got chocolate chips on your pancake.

$\Pr[F\cap G] = \sum_{k\in F}\Pr[\{k\}\cap G] = \sum_{k=1}^{m}\dfrac{1}{m+n}\cdot\dfrac{k}{m} = \dfrac{1}{m+n}\cdot\dfrac{m+1}{2}$;

$\Pr[\overline{F}\cap G] = 0$;

$\Pr[G] = \Pr[F\cap G]+\Pr[\overline{F}\cap G] = \dfrac{m+1}{2(m+n)}$.

Alternate:

$\Pr[F] = \dfrac{m}{m+n}$;

$\Pr[G|F] = \sum_{k=1}^{m}\dfrac{1}{m}\cdot\dfrac{k}{m} = \dfrac{m+1}{2m}$;

$\Pr[G|\overline{F}] = 0$;

$\Pr[G] = \Pr[G|F]\Pr[F]+\Pr[G|\overline{F}]\Pr[\overline{F}] = \dfrac{m+1}{2(m+n)}$.

(e) $\Pr[F\cap G] = \sum_{k\in F}\Pr[\{k\}\cap G] = \sum_{k=11}^{m}\dfrac{1}{m+n}\cdot\dfrac{k}{m} = \dfrac{1}{m+n}\cdot\dfrac{(m+11)(m-10)}{2m}$;

$\Pr[F|G] = \dfrac{\Pr[F\cap G]}{\Pr[G]} = \dfrac{(m+11)(m-10)}{m(m+1)} = \dfrac{m^2+m-110}{m^2+m}$.

(Below algebra not required, since $\Pr[G]$ does not change):

$\Pr[F\cap G] = \sum_{k\in F}\Pr[\{k\}\cap G] = \sum_{k=11}^{m}\dfrac{1}{m+n}\cdot\dfrac{k}{m}$;

$\Pr[\overline{F}\cap G] = \sum_{k\in\overline{F}}\Pr[\{k\}\cap G] = \sum_{k=1}^{10}\dfrac{1}{m+n}\cdot\dfrac{k}{m}$;

$\Pr[G] = \Pr[F\cap G]+\Pr[\overline{F}\cap G] = \sum_{k=1}^{m}\dfrac{1}{m+n}\cdot\dfrac{k}{m} = \dfrac{m+1}{2(m+n)}$.

Alternate:

$\Pr[F] = \dfrac{m-10}{m+n}$;

$\Pr[G|F] = \sum_{k=11}^{m}\dfrac{1}{m-10}\cdot\dfrac{k}{m} = \dfrac{m+11}{2m}$;

$\Pr[\overline{F}] = \dfrac{n+10}{m+n}$;

$\Pr[G|\overline{F}] = \sum_{k=1}^{10}\dfrac{1}{n+10}\cdot\dfrac{k}{m} = \dfrac{55}{m(n+10)}$;

$\Pr[F|G] = \frac{\Pr[F\cap G]}{\Pr[F\cap G]+\Pr[\overline{F}\cap G]} = \frac{\Pr[F]\Pr[G|F]}{\Pr[F]\Pr[G|F]+\Pr[\overline{F}]\Pr[G|\overline{F}]} = \dfrac{(m+11)(m-10)}{(m+11)(m-10)+110} = \dfrac{m^2+m-110}{m^2+m}$.

3. **Fundamentals (2/2/2/2 points)**

True or false? For the following statements, provide either a proof or a simple counterexample. Let X,Y,Z be arbitrary random variables.

(a) If (X,Y) are independent and (Y,Z) are independent, then (X,Z) are independent.

Answer: FALSE. Consider $X = Z$.
FALSE. Let X, Y be i.i.d Bernoulli$(1/2)$ random variables, and let $Z = X$. Then (X, Y) and (Y, Z) are independent by construction, but (X, Z) are not independent because $Z = X$.

(b) If (X, Y) are dependent and (Y, Z) are dependent, then (X, Z) are dependent.

Answer: FALSE. X, Z iid Bernoulli$(1/2)$. $Y = XZ$.
FALSE. Let X, Z be i.i.d Bernoulli$(1/2)$ random variables, and let $Y = XZ$. Then (X, Z) are independent by construction, but (X, Y) are not independent, since $\Pr[X = 0 \wedge Y = 1] = 0 \neq \Pr[X = 0]\Pr[Y = 1] = (1/2)(1/4)$.

(c) Assume X is discrete. If $\mathrm{Var}(X) = 0$, then X is a constant.

Answer: TRUE.
TRUE. Let $\mu = \mathbf{E}[X]$. By definition,

$$0 = \mathrm{Var}(X) = \mathbf{E}[(X-\mu)^2] = \sum_{\omega \in \Omega} \Pr[\omega](X(\omega) - \mu)^2$$

The RHS is the sum of non-negative numbers, so if the sum is 0, each term must be 0.
So $\Pr[\omega] > 0 \implies (X(\omega) - \mu)^2 = 0 \implies X(\omega) = \mu$. Therefore X is constant (equal to $\mu = \mathbf{E}[X]$).

(d) $\mathbf{E}[X]^4 \leq \mathbf{E}[X^4]$

Answer: TRUE.
TRUE. First, for an arbitrary random variable Y, we have:

$$0 \leq \mathbf{E}[(Y - \mathbf{E}[Y])^2] = \mathbf{E}[Y^2] - \mathbf{E}[Y]^2$$

So $\mathbf{E}[Y]^2 \leq \mathbf{E}[Y^2]$. Now applying this twice, once for $Y = X$ and once for $Y = X^2$:

$$\mathbf{E}[X]^4 = (\mathbf{E}[X]^2)^2 \leq (\mathbf{E}[X^2])^2 \leq \mathbf{E}[(X^2)^2] = \mathbf{E}[X^4]$$

4. **Like a Rolling Die (2/2/2/2/2/2 points)**

Suppose you roll a fair die n times. What is the expectations of each of the following random variables?

(a) A is the random variable that denotes the sum of the numbers in those rolls.

Answer: Let X_i be the value of the i-th roll. Since the die is rolled independently. All X_i follows an identical and independent distributed Uniform distribution, in the range of (1,2,...,6).

$$\mathbf{E}(A) = \sum_{i=1}^{n} \mathbf{E}(X_i) = \sum_{i=1}^{n} \frac{1+2+3+4+5+6}{6} = 3.5 \cdot n.$$

(b) B is the random variable that denotes the maximum number in the those rolls.

Answer: Since B denotes the maximum number: $B = \max(X_1, X_2, ..., X_n)$.
Since B is the maximum among variable Xs. So we can get the following:

$$\mathbf{P}(B \leq b) = \mathbf{P}(X_1 \leq b) \cdot \mathbf{P}(X_2 \leq b) ... \mathbf{P}(X_n \leq b) = (\frac{b}{6})^n.$$

Therefore,

$$\mathbf{P}(B=b)=\mathbf{P}(B\le b)-\mathbf{P}(B\le b-1)=(\frac{b}{6})^n-(\frac{b-1}{6})^n.$$

Expected value can be calculated using its definition:

$$\mathbf{E}(B)=\sum_{i=1}^{6} i\cdot P(B=i)=\sum_{i=1}^{6} i((\frac{i}{6})^n-(\frac{i-1}{6})^n).$$

Another way to solve this is to use tail sum.

$$\begin{aligned}\mathbf{E}(B)&=\sum_{i=1}^{6} i\cdot P(B=i)=\sum_{i=1}^{6} P(B\ge i)\\&=\sum_{i=1}^{6} 1-P(B<i)=\sum_{i=0}^{5} 1-(\frac{i}{6})^n\\&=6-\sum_{i=1}^{5}(\frac{i}{6})^n.\end{aligned}$$

Those two answers should give you the same results.

(c) C is the random variable that denotes the sum of the largest two numbers in the first three rolls.

Answer: Let X_1, X_2, X_3 be the value of first 3 rolls, and X_{min} be the minimum of the first 3 rolls. We can write C in terms of those four variables:

$$\mathbf{E}(C)=\mathbf{E}(X_1+X_2+X_3-X_{min})=\mathbf{E}(X_1+X_2+X_3)-\mathbf{E}(X_{min})$$

Now, we will use the same method above to find out the expectation of the minimum.

$$\mathbf{P}(X_{min}\ge m)=(\frac{6-m}{6})^3.$$

So that

$$\mathbf{P}(X_{min}=m)=\mathbf{P}(X_{min}\ge m-1)-\mathbf{P}(X_{min}\ge m)=(\frac{7-m}{6})^3-(\frac{6-m}{6})^3.$$

$$\mathbf{E}(X_{min})=\sum_{i=1}^{6} i((\frac{7-i}{6})^3-(\frac{6-i}{6})^3).$$

After you calculate the expectation of the minimum, you can get the sum of the largest two easily. You can also use tail sum to calculate it. The calculation is very similar to part b so it's omitted.

(d) D is the random variable that denotes the number of multiples of three in those rolls.

Answer: It can be seen that D follows a binomial distribution with n, $p=2/6$. So $\mathbf{E}(D)=np=n/3$.

(e) E is the random variable that denotes the number of faces which fail to appear in those rolls.

Answer: Let I_i be an indicator that takes value 1 if face i does not appear.
So that $\mathbf{E}(I_i) = \mathbf{P}(I_i = 1) = (5/6)^n$.

$$\mathbf{E}(E) = \sum_{i=1}^{6} \mathbf{E}(I_i) = \sum_{i=1}^{6} (5/6)^n = 6 \cdot (5/6)^n.$$

(f) F is the random variable that denotes the number of distinct faces that appear in those rolls.

Answer: Very similar to part (e), Let I_i be an indicator that takes value 1 if face i appears.
So that $\mathbf{E}(I_i) = \mathbf{P}(I_i = 1) = 1 - (5/6)^n$.

$$\mathbf{E}(F) = \sum_{i=1}^{6} \mathbf{E}(I_i) = \sum_{i=1}^{6} [1 - (5/6)^n] = 6 \cdot [1 - (5/6)^n].$$

5. **Linearity (4/4/4 points)**

Solve each of the following problems using linearity of expectation. Explain your methods clearly.

(a) In an arcade, you play game A 10 times and game B 20 times. Each time you play game A, you win with probability $1/3$ (independently of the other times), and if you win you get 3 tickets (redeemable for prizes), and if you lose you get 0 tickets. Game B is similar, but you win with probability $1/5$, and if you win you get 4 tickets. What is the expected total number of tickets you receive?

Answer: Let A_i be the indicator you win the i^{th} time you play game A and B_i be the same for game B. The expected value of A_i and B_i are,

$$\mathbf{E}[A_i] = 1 \cdot 1/3 + 0 \cdot 2/3 = 1/3,$$
$$\mathbf{E}[B_i] = 1 \cdot 1/5 + 0 \cdot 4/5 = 1/5.$$

Let T_A be the random variable for the number of tickets you win in game A, and T_B be the number of tickets you win in game B.

$$\begin{aligned}\mathbf{E}[T_A + T_B] &= 3\mathbf{E}[A_1] + \cdots + 3\mathbf{E}[A_{10}] + 4\mathbf{E}[B_1] + \cdots + 4\mathbf{E}[B_{20}] \\ &= 10\left(3 \cdot \frac{1}{3}\right) + 20\left(4 \cdot \frac{1}{5}\right) = 26\end{aligned}$$

□

(b) A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence "book" appears?

Answer: There are $1,000,000 - 4 + 1 = 999,997$ places where "book" can appear, each with a (non-independent) probability of $\frac{1}{26^4}$ of happening. If A is the random variable that tells how many times "book" appears, and A_i is the indicator variable that is 1 if "book" appears starting at the i^{th} letter, then

$$\begin{aligned}\mathbf{E}[A] &= \mathbf{E}[A_1 + \cdots + A_{999,997}] \\ &= \mathbf{E}[A_1] + \cdots + E[A_{999,997}] \\ &= \frac{999,997}{26^4} \approx 2.19\end{aligned}$$

times. $\square$

(c) A building has n floors numbered $1, 2, \ldots, n$, plus a ground floor G. At the ground floor, m people get on the elevator together, and each gets off at a uniformly random one of the n floors (independently of everybody else). What is the expected number of floors the elevator stops at (not counting the ground floor)?

Answer: Let A_i be the indicator that the elevator stopped at floor i.

$$\Pr[A_i = 1] = 1 - \Pr[\text{no one gets off at } i] = 1 - \left(\frac{n-1}{n}\right)^m.$$

If A is the number of floors the elevator stops at, then

$$\begin{aligned}\mathbf{E}[A] &= \mathbf{E}[A_1 + \cdots + A_n] \\ &= \mathbf{E}[A_1] + \cdots + E[A_n] = n \cdot \left(1 - \left(\frac{n-1}{n}\right)^m\right)\end{aligned}$$

$\square$

6. **Runs (5 points)**

Suppose I have a biased coin which comes up heads with probability p, and I flip it n times. A "run" is a sequence of coin flips all of the same type, which is not contained in any longer sequence of coin flips all of the same type. For example, the sequence "HHHTHH" has three runs: "HHH," "T," and "HH."

Compute the expected number of runs in a sequence of n flips.

Answer: Let X_i be the indicator variable for the event that position i is the beginning of a run. Then the total number of runs is simply $X = \sum_{i=1}^n X_i$. Now, obviously X_1 is always 1, whereas for $i \geq 2$, $X_i = 1$ if and only if the outcome of the i-th flip is different from that of the $(i-1)$-th flip. This happens with probability $p(1-p) + (1-p)p = 2p(1-p)$. By linearity of expectation,

$$\mathbf{E}[X] = \mathbf{E}[X_1] + \mathbf{E}[X_2] + \mathbf{E}[X_3] + \cdots + \mathbf{E}[X_n] = 1 + (n-1) \cdot 2p(1-p) = 1 + 2(n-1)p(1-p).$$

7. **Coupon Collection (6 points)**

Suppose you take a deck of n cards and repeatedly perform the following step: take the current top card and put it back in the deck at a uniformly random position. (I.e., the probability that the card is placed in any of the n possible positions in the deck — including back on top — is $1/n$.) Consider the card that starts off on the bottom of the deck. What is the expected number of steps until this card rises to the top of the deck? (Hint: Let T be the number of steps until the card rises to the top. We have $T = T_n + T_{n-1} + \cdots + T_2$, where the random variable T_i is the number of steps until the bottom card rises from position i to position $i-1$. Thus, for example, T_n is the number of steps until the bottom card rises off the bottom of the deck, and T_2 is the number of steps until the bottom card rises from second position to top position. What is the distribution of T_i?) (More hints: You may use the fact that $\sum_{i=1}^{n} \frac{1}{i} \approx \ln n$.)

Answer: Since a card at location i moves to location $i-1$ when the current top card is placed in any of the locations $i, i+1, \ldots, n$, it will rise with probability $p = \frac{n-i+1}{n}$. Thus, $T_i \sim \text{Geom}(p)$, and $\mathbb{E}(T_i) = \frac{1}{p} = \frac{n}{n-i+1}$. We now can see how this is exactly the coupon collector's problem, but with one fewer term (namely, without T_1). Finally, we can apply linearity of expectation to compute

$$\mathbb{E}(T) = \sum_{i=2}^{n} \mathbb{E}(T_i) = \sum_{i=2}^{n} \frac{n}{n-i+1} = n \sum_{i=2}^{n} \frac{1}{n-i+1} \approx n \ln(n-1)$$

8. **Markov's Inequality and Chebyshev's Inequality (2/2/2/2/2/2 points)**

A random variable X has variance $\text{Var}(X) = 9$ and expectation $\mathbb{E}(X) = 2$. Furthermore, the value of X is never greater than 10. Given this information, provide either a proof or a counterexample for the following statements.

(a) $\mathbb{E}(X^2) = 13$.
(b) $\Pr[X = 2] > 0$.
(c) $\Pr[X \geq 2] = \Pr[X \leq 2]$.
(d) $\Pr[X \leq 1] \leq 8/9$.
(e) $\Pr[X \geq 6] \leq 9/16$.
(f) $\Pr[X \geq 6] \leq 9/32$.

Answer:

(a) TRUE. Since $9 = \text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X^2) - 2^2$, we have $\mathbb{E}(X^2) = 9 + 4 = 13$.

(b) FALSE. Construct a random variable X that satisfies the conditions in the question but does not take on the value 2. A simple example would be a random variable that takes on 2 values, where $\Pr[X = a] = \frac{1}{2}, \Pr[X = b] = \frac{1}{2}$, and $a \neq b$. The expectation must be 2, so we have $\frac{1}{2}a + \frac{1}{2}b = 2$. The variance is 9, so $\mathbb{E}(X^2) = 13$ (from part (a)) and $\frac{1}{2}a^2 + \frac{1}{2}b^2 = 13$. Solving for a and b, we get $\Pr[X = -1] = \frac{1}{2}, \Pr[X = 5] = \frac{1}{2}$ as a counterexample.

(c) FALSE. Construct a random variable X that satisfies the conditions in the question but does not have an equal chance of being less than or greater than 2. A simple example would be a random variable that takes on 2 values, where $\Pr[X = a] = p, \Pr[X = b] = 1 - p$. Here, we use the same approach as part (b) except with a generic p, since we want $p \neq \frac{1}{2}$. The expectation must be 2, so we have $pa + (1-p)b = 2$. The variance is 9, so $\mathbb{E}(X^2) = 13$ and $pa^2 + (1-p)b^2 = 13$. Solving for a and b, we find the relation $b = 2 \pm \frac{3}{\sqrt{x}}$, where $x = \frac{1-p}{p}$. Then, we can find an example by plugging in values for x so that $a, b \leq 10$ and $p \neq \frac{1}{2}$. One such counterexample is $\Pr[X = -7] = \frac{1}{10}, \Pr[X = 3] = \frac{9}{10}$.

(d) TRUE. Let $Y = 10 - X$. Since X is never exceeds 10, Y is a non-negative random variable. By Markov's inequality,

$$\Pr[10 - X \geq a] = \Pr[Y \geq a] \leq \frac{\mathbb{E}(Y)}{a} = \frac{\mathbb{E}(10 - X)}{a} = \frac{8}{a}.$$

Setting $a = 9$, we get $\Pr[X \leq 1] = \Pr[10 - X \geq 9] \leq \frac{8}{9}$.

(e) TRUE. Chebyshev's inequality says $\Pr[|X - \mathbf{E}[X]| \geq a] \leq \frac{\mathrm{Var}(X)}{a^2}$. If we set $a = 4$, we have

$$\Pr[|X-2| \geq 4] \leq \frac{9}{16}.$$

Now we simply observe that $\Pr[X \geq 6] \leq \Pr[|X-2| \geq 4]$, because the event $X \geq 6$ is a subset of the event $|X-2| \geq 4$.

(f) FALSE. We use the same approach as in part (c), except we find a counterexample that fits the inequality $\Pr[X \geq 6] \leq 9/32$. One example is $\Pr[X = 0] = \frac{9}{13}, \Pr[X = \frac{13}{2}] = \frac{4}{13}$.

9. **Casino wins (2/2/3/3 points)**

A gambler plays 120 hands of draw poker, 60 hands of black jack, and 20 hands of stud poker per day. He wins a hand of draw poker with probability $1/6$, a hand of black jack with probability $1/2$, and a hand of stud poker with probability $1/5$. Assume the outcomes of the card games are mutually independent.

(a) What is the expected number of hands the gambler wins in a day?

(b) What is the variance in the number of hands won per day?

(c) What would the Markov bound be on the probability that the gambler will win 108 hands on a given day?

(d) What would the Chebyshev bound be on the probability that the gambler will win 108 hands on a given day?

Answer:

(a) Let R be the number of games won. By linearity of expectation:

$$\mathbf{E}[R] = 120 \cdot \frac{1}{6} + 60 \cdot \frac{1}{2} + 20 \cdot \frac{1}{5} = 54.$$

(b) The variance can also be calculated using linearity of variance, since the $120 + 60 + 20 = 200$ indicator r.v.'s, one for each hand of some game, are mutually independent. For an individual hand, the variance is $p(1-p)$ where p is the probability of winning (it's just a single "Bernoulli trial"). Therefore, $\mathrm{Var}(R) = 120(1/6)(5/6) + 60(1/2)(1/2) + 20(1/5)(4/5) = 50/3 + 15 + 16/5 = 523/15 = 3413/15$.

(c) The expected number of games won is 54, and the number of games played is non-negative, so by Markov, $\Pr[R \geq 108] \leq 54/108 = 1/2$.

(d) The Chebyshev bound yields:

$$\Pr[R - 54 \geq 54] \leq \Pr[|R - 54| \geq 54] \leq \frac{\mathrm{Var}(R)}{54^2} = \frac{523/15}{54^2} \leq 0.012$$

Note that the first inequality in this case is actually an equality, but that's irrelevant here.

10. **Those 3407 Votes (2/3/5 points)**

In the aftermath of the 2000 US Presidential Election, many people have claimed that unusually large number of votes cast for Pat Buchanan in Palm Beach County are statistically highly significant, and thus of dubious validity. In this problem, we will examine this claim from a statistical viewpoint.

The total percentage votes cast for each presidential candidate in the entire state of Florida were as follows:

Gore	Bush	Buchanan	Nader	Browne	Others
48.8%	48.9%	0.3%	1.6%	0.3%	0.1%

In Palm Beach County, the actual votes cast (before the recounts began) were as follows:

Gore	Bush	Buchanan	Nader	Browne	Others	Total
268945	152846	3407	5564	743	781	432286

To model this situation probabilistically, we need to make some assumptions. Let's model the vote cast by each voter in Palm Beach County as a random variable X_i, where X_i takes on each of the six possible values (five candidates or "Others") with probabilities corresponding to the Florida percentages. (Thus, e.g., $\Pr[X_i = \text{Gore}] = 0.488$.) There are a total of $n = 432286$ voters, and their votes are assumed to be mutually independent. Let the r.v. B denote the total votes cast for Buchanan in Palm Beach County (i.e., the number of voters i for which $X_i = \text{Buchanan}$).

(a) Compute the expectation $\mathbf{E}[B]$ and the variance $\mathrm{Var}(B)$.

Answer: Let B_i be a random variable representing whether the ith person voted for Buchanan. Then $B_i = 1$ if and only if $X_i = \text{Buchanan}$, so $B_i \sim \text{Bernoulli}(0.003)$. Note that the B_i's are independently and identically distributed, with $\mathbf{E}[B_i] = 0.003$ and $\mathrm{Var}(B_i) = 0.003 \times (1 - 0.003) = 0.002991$. Moreover, by linearity of expectation and independence, we find that $\mathbf{E}[B] = \sum_{i=1}^{n} \mathbf{E}[B_i] = 432286 \times 0.003 \approx 1297$ and $\mathrm{Var}(B) = \sum_{i=1}^{n} \mathrm{Var}(B_i) = 432286 \times 0.002991 \approx 1293$.

(b) Use Chebyshev's inequality to compute an *upper bound* b on the probability that Buchanan receives at least 3407 votes, i.e., find a number b such that

$$\Pr[B \geq 3407] \leq b.$$

Based on this result, do you think Buchanan's vote is significant?

Answer: Chebyshev's inequality says that

$$\Pr[|B - \mathbf{E}[B]| \geq a] \leq \frac{\mathrm{Var}(B)}{a^2}.$$

In our case $\mathbf{E}[B] = 1297$ and $\mathrm{Var}(B) = 1293$, so if we take $a = 2110$, we find that $\Pr[|B - 1297| \geq 2110] \leq 1293/2110^2 \approx 0.0003$. Now note that the condition $|B - 1297| < 2110$ is equivalent to the condition $-813 < B < 3407$, and since B is non-negative, we find that $\Pr[B > 3407] \leq 0.0003$ (roughly), so we can take $b \approx 0.0003$. In other words, receiving 3407 votes for Buchanan in Palm Beach County seems very unlikely to happen by chance, under this simple model. So yes, this is statistically significant.

(c) Suppose that your bound b in part (b) is exactly accurate, i.e., assume that $\Pr[X \geq 3407]$ is exactly equal to b. [*In fact the true value of this probability is much smaller*] Suppose also that all 67 counties in Florida have the same number of voters as Palm Beach County, and that all behave independently according to the same statistical model as Palm Beach County. What is the probability that in *at least one* of the counties, Buchanan receives at least 3407 votes? How would this affect your judgment as to whether the Palm Beach tally is significant?

Answer: Let p_j be the probability that the jth county does not receive 3407 votes for Buchanan. We have from part (b) that $p_j = 1 - b \approx 0.9997$. Note that the probability that no county yields at least 3407 votes for Buchanan is $p_1 \times \cdots \times p_{67}$, since the voters in each county behave independently. Thus, the probability that Buchanan does not receive 3407 votes in any county is about $(0.9997)^{67} \approx 0.98$. Consequently, the probability that Buchanan *does* receive at least 3407 votes in some county is about $1 - 0.98 \approx 0.02$. In other words, this seems unlikely to happen by chance.