2 Calculation

5. Given a (fairly simple) probability mass function or probability density function of a random variable, you should be able to compute the expected value and variance of the variable. For example:

   (a) Let $0 < p < 1$. Let X be discrete with mass points $k = 1, 2, \dots$ and probability mass function $P(X = k) = p(1-p)^{k-1}$.
   (In this case, we say that X has a *geometric distribution*.)
   $E(X) = \sum_{k=1}^{\infty} kP(X = k) = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \frac{p}{(1-(1-p))^2} = \frac{1}{p}$.
   $E(X^2) = \sum_{k=1}^{\infty} k^2 P(X = k) = \sum_{k=1}^{\infty} k^2 p(1-p)^{k-1} = \frac{p(1+(1-p))}{(1-(1-p))^3} = \frac{2-p}{p^2}$, so $\text{var}(X) = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{1-p}{p^2}$.

   (b) Let X be continuous with probability density function $f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$.
   (In this case, we say that X has a *uniform distribution*.)
   $E(X) = \int_a^b x\frac{1}{b-a}\mathrm{d}x = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$.
   $E(X^2) = \int_a^b x^2\frac{1}{b-a}\mathrm{d}x = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}$, so $\text{var}(X) = E(X^2) - E(X)^2 = \frac{(b-a)^2}{12}$.

   (c) Let X be continuous with probability density function $f(x) = \begin{cases} 4x^3 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$.
   $E(X) = \int_0^1 x \cdot 4x^3\mathrm{d}x = \frac{4}{5}x^4\Big|_0^1 = \frac{4}{5}$.
   $E(X^2) = \int_0^1 x^2 \cdot 4x^3\mathrm{d}x = \frac{2}{3}x^6\Big|_0^1 = \frac{2}{3}$, so $\text{var}(X) = \frac{2}{3} - \left(\frac{4}{5}\right)^2 = \frac{2}{75}$.

3 Example Problems

6. I have 6 coins in my pocket: 3 are fair, one is double-headed, one is double-tailed, and one lands heads 75% of the time. I flip the coin three times and get heads each time. What is the probability that I am flipping the double-headed coin?

   Let H_1 be the event that the coin is double-headed, H_2 that the coin is fair, H_3 that the coin is double-tailed, and H_4 that the coin lands heads 75% of the time. Then $P(H_1) = \frac{1}{6}$, $P(H_2) = \frac{3}{6}$, $P(H_3) = \frac{1}{6}$, and $P(H_4) = \frac{1}{6}$. Let E be the event that the coin lands heads in all three flips. Using Bayes' Theorem, we calculate:

$$\begin{aligned} P(H_1|E) &= \frac{P(E|H_1)P(H_1)}{P(E|H_1)P(H_1) + P(E|H_2)P(H_2) + P(E|H_3)P(H_3) + P(E|H_4)P(H_4)} \\ &= \frac{1 \cdot \frac{1}{6}}{1 \cdot \frac{1}{6} + \frac{1}{8} \cdot \frac{3}{6} + 0 \cdot \frac{1}{6} + (0.75)^3 \cdot \frac{1}{6}} \\ &= 0.557. \end{aligned}$$

7. A point Q is chosen at random inside a sphere with radius r. What are the expected value and the standard deviation of the distance from the center of the sphere to the point Q?

   We first calculate the cumulative distribution function for X =the distance of Q from the center of the sphere. Note that for $0 \le x \le r$,

   $$P(X \le x) = \frac{\text{volume of sphere of radius } x}{\text{volume of sphere of radius } r} = \frac{\frac{4}{3}\pi x^3}{\frac{4}{3}\pi r^3} = \frac{x^3}{r^3}.$$

   So, the probability density function of X is $f(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[P(X \le x)\right] = \frac{3x^2}{r^3}$ for $0 < x < r$, and

   $$f(x) = \begin{cases} \frac{3x^2}{r^3} & \text{if } 0 < x < r \\ 0 & \text{otherwise} \end{cases}.$$

   Then, $E(X) = \int_0^r xf(x)\mathrm{d}x = \int_0^r x \cdot \frac{3x^2}{r^3}\mathrm{d}x = \frac{3r^4}{4r^3} = \frac{3}{4}r$ and $E(X^2) = \int_0^r x^2 \cdot \frac{3x^2}{r^3}\mathrm{d}x = \frac{3r^5}{5r^3} = \frac{3}{5}r^2$, so that $\mathrm{var}(X) = \frac{3}{5}r^2 - \left(\frac{3}{4}r\right)^2 = \frac{3}{80}r^2$ and $\sigma(X) = 0.1936r$.

8. Two gamblers, A and B, play the following game: each takes a turn drawing a card off the top of a well-shuffled deck until all 52 cards have been drawn. Every time a player draws a diamond, the other player gives the drawing player \$1. Let X be the number of dollars that A has won or lost after all 52 cards have been drawn. Find the expected value and standard deviation of X. (You may leave your answer unsimplified, if you wish.)

   Note that we don't care about any of the cards except the diamonds, so we can focus only on how many diamonds each gambler ends up with at the end of the game. Since each gambler draws 26 cards, each has an equal probability of getting each of the diamonds. Let Ω be the set of 13-tuples of A and B that represents which diamonds each gambler receives, so that (for example) $(A, A, A, B, A, B, B, B, B, B, B, B, A)$ means that A got the ace, two, three, five, and king of diamonds and B got the others. Note $|\Omega| = 13^2$. Let Y be the number of diamonds that A gets. Then

   $$P(Y = k) = \frac{\binom{13}{k}}{2^{13}} \text{ for } k = 0, 1, 2, \ldots, 13.$$

   If $Y = k$, then A will get k dollars from B and give $13 - k$ dollars to B, so that $X = k - (13 - k) = 2k - 13 = 2Y - 13$. Then

   $$E(X) = E(2Y - 13) = \sum_{k=0}^{13}(2k - 13)P(Y = k) = \sum_{k=0}^{13}(2k - 13)\frac{\binom{13}{k}}{2^{13}} \text{ and}$$

   $$E(X^2) = E((2Y - 13)^2) = \sum_{k=0}^{13}(2k - 13)^2\frac{\binom{13}{k}}{2^{13}},$$

   so $\mathrm{var}(X) = \sum_{k=0}^{13}(2k - 13)^2\frac{\binom{13}{k}}{2^{13}} - \left(\sum_{k=0}^{13}(2k - 13)\frac{\binom{13}{k}}{2^{13}}\right)^2$ and $\sigma(X) = \sqrt{\mathrm{var}(X)}$. Using a computer, these simplify to $E(X) = 0$ and $\sigma(X) = 25.1$.

   Fun Fact: Because $P(Y = k)$ is a probability mass function, it must be that $\sum_{k=0}^{13} P(Y = k) = 1$, which implies that $\sum_{k=0}^{13}\binom{13}{k} = 2^{13}$. More generally, $\sum_{k=0}^{n}\binom{n}{k} = 2^n$.

9. A bent coin comes up heads with probability p, where $0 < p < 1$.

   (a) You flip the coin until heads comes up. How many times do you expect to flip the coin?
   Let X be the number of times you flip the coin. If $X = k$, then the first $k-1$ flips must be tails and the kth flip must be heads, so $P(X = k) = p(1-p)^{k-1}$. This is the geometric distribution, so as we calculated in 5(a), we see that $E(X) = \frac{1}{p}$.

   (b) You flip the coin until heads comes up r times. How many times do you expect to flip the coin?
   Let Y be the number of times you flip the coin to get r heads. (We say that Y has a *negative binomial distribution* with parameters r and p.) Note that $Y = X_1 + \cdots + X_r$ where X_i is the number of times you flip the coin to go from $i-1$ heads to i heads. We calculated in (a) that $E(X_i) = \frac{1}{p}$, so $E(Y) = E(X_1) + \cdots + E(X_r) = \frac{r}{p}$.

10. Suppose a bridge player's hand of 13 cards contains an ace. What is the probability that the player has only one ace?
   Let X be the number of aces in the players hand. The probability that the player has only one ace given that the hand contains at least one ace is $P(X = 1|X \geq 1) = \frac{P(X=1)}{P(X \geq 1)} = \frac{P(X=1)}{1-P(X=0)}$. If $X = 1$, then the hand contains one of the four aces and the remaining 12 cards come from the 48 non-aces, so that $P(X = 1) = \frac{4 \cdot \binom{48}{12}}{\binom{52}{13}}$. If $X = 0$, then all 13 cards in the hand come from the 48 non-aces, so $P(X = 0) = \frac{\binom{48}{13}}{\binom{52}{13}}$. Then $P(X = 1|X \geq 1) = 0.63$.

11. Let k be a constant and consider the function

$$f(x) = \begin{cases} kx^2 & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}.$$

   (a) For what value(s) of k is f a probability density function?
   Note $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \int_0^2 kx^2 \mathrm{d}x = \frac{8k}{3}$. This is a probability density if and only if this integral equals 1, so f is a probability density function if and only if $k = \frac{3}{8}$.

   (b) For the value(s) of k you found above, suppose X is a random variable with probability density function f. Find the expected value and standard deviation of X.
   We compute $E(X) = \int_0^2 x \cdot \frac{3}{8}x^2 \mathrm{d}x = \frac{3}{2}$ and $E(X^2) = \int_0^2 x^2 \cdot \frac{3}{8}x^2 \mathrm{d}x = \frac{96}{40}$, so that $\mathrm{var}(X) = \frac{96}{40} - \left(\frac{3}{2}\right)^2 = \frac{3}{20}$, and $\sigma(X) = \sqrt{\mathrm{var}(X)} = 0.39$.

12. Suppose that U is a number chosen at random between 0 and 1 and that $X = \sqrt[3]{U}$. What is the expected value and standard deviation of X?
   First, find the cumulative distribution function of X. We compute:

$$P(X \leq x) = P(\sqrt[3]{U} \leq x) = P(U \leq x^3) = x^3 \text{ for } 0 \leq x \leq 1.$$

   Taking the derivative, the probability density function of X is $f(x) = \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$. So, $E(X) = \int_0^1 x \cdot 3x^2 \mathrm{d}x = \frac{3}{4}$, $E(X^2) = \int_0^1 x^2 \cdot 3x^2 \mathrm{d}x = \frac{3}{5}$, so $\sigma(X) = \sqrt{\frac{3}{5} - \left(\frac{3}{4}\right)^2} = 0.1936$.

13. A hand of 13 cards is dealt from a well-shuffled deck. How many four-of-a-kinds do you expect the hand to contain? How many four-of-a-kinds do you expect the 39 cards remaining in the deck to contain?

    Let X be the number of four-of-a-kinds in the 13 card hand. Then $X = X_1 + \cdots + X_{13}$ where X_i equals 1 if the hand contains the four-of-a-kind of rank i and equals 0 otherwise.

    We compute $E(X_i) = P(X_i = 1) = \frac{\binom{48}{9}}{\binom{52}{13}}$ since the hand contains the four cards of rank i and the other 9 cards in the hand can be any of the other cards in the deck. Then $E(X) = 13 \cdot \frac{\binom{48}{9}}{\binom{52}{13}} = 0.034$.

    Now, let X be the number of four-of-a-kinds in the 13 cards remaining in the deck. One way to solve this is by using exactly the same technique as above (but with a 39 card "hand" instead of a 13 card hand). Then we find that $E(X) = 13 \cdot \frac{\binom{48}{35}}{\binom{52}{13}} = 3.950$. Alternatively, you can focus on which 13 cards end up in the hand. There are $\binom{48}{13}$ ways that a 13 card hand can be drawn without taking any of the cards of rank i, so that $E(X) = 13 \cdot \frac{\binom{48}{13}}{\binom{52}{13}} = 3.950$.

14. A card is drawn from a deck of 52. Let A be the event that the card is a heart, B be the event that the card is red, and C be the event that the card is a king. Are events A and B independent? Are events A and C independent? (Justify your answer mathematically.)

    We calculate $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{2}$, $P(C) = \frac{1}{13}$, $P(AB) = \frac{1}{4}$, and $P(AC) = \frac{1}{52}$. Thus, $P(AB) \neq P(A)P(B)$ so that A and B are not independent, and $P(AC) = P(A)P(C)$, so that A and C are independent.

15. An urn contains N balls, n of which are blue. If m balls are drawn from the urn, how many of the balls drawn do you expect to be blue?

    Let X be the number of blue balls drawn. (We say that X has a *hypergeometric distribution* with parameters n, $N - n$, and m.) There are two ways to do this:

    Way 1: Let $X = X_1 + \cdots + X_m$ where X_i is 1 if the ith ball drawn is blue and 0 otherwise. Then $E(X_i) = P(X_i = 1) = \frac{n}{N}$, so that $E(X) = m \cdot \frac{n}{N} = \frac{mn}{N}$.

    Way 2: Label the blue balls $b_1, \ldots, b_n$. Let $X = Y_1 + \cdots + Y_n$ where Y_i is 1 if the ball b_i is drawn and 0 otherwise. There are $\binom{N-1}{m-1}$ ways to draw $m - 1$ balls in addition to the ball b_i, so $E(Y_i) = P(Y_i = 1) = \frac{\binom{N-1}{m-1}}{\binom{N}{m}} = \frac{m}{N}$, so that $E(X) = n \cdot \frac{m}{N} = \frac{mn}{N}$.

16. A red die is rolled and the number that comes up is observed. Then, a blue die is rolled that many times. What is the probability that the blue die will come up 6 at least once?

    $P(\text{at least one 6 on a blue die}) = 1 - P(\text{no 6 on any blue die}) = 1 - (\frac{1}{6} \cdot \left(\frac{5}{6}\right) + \frac{1}{6} \cdot \left(\frac{5}{6}\right)^2 + \cdots + \frac{1}{6} \cdot \left(\frac{5}{6}\right)^6) = 0.4457$.

17. A point Q is chosen at random from inside the triangle with vertices at $(0, 0)$, $(2, 2)$, and $(4, 0)$. Let X be the x-coordinate of Q. What are the expected value and standard deviation of X?

Note that values of x for which more y values are possible will be more likely. All x values are in the interval $[0,4]$ and the value $x = 2$ is most likely, so X has a triangular distribution (see p. 297). Instead of using those formulas, however, let's just calculate it. The density function of X is $f(x) = \begin{cases} hx & \text{if } 0 < x < 2 \\ h(4-x) & \text{if } 2 < x < 4 \\ 0 & \text{otherwise} \end{cases}$ where h is chosen so that the integral of f is equal to 1. This triangle has area 4, so $h = \frac{1}{4}$. Then $E(X) = \int_0^2 x \cdot \frac{x\mathrm{d}x}{4} + \int_2^4 x \cdot \frac{(4-x)\mathrm{d}x}{4} = 2$. We calculate that $E(X^2) = \int_0^2 x^2 \cdot \frac{x\mathrm{d}x}{4} + \int_2^4 x^2 \cdot \frac{(4-x)\mathrm{d}x}{4}$, so that $\sigma(X) = 0.82$.

18. Suppose that the time between arrivals at a certain store is exponentially distributed with parameter $\lambda = 0.2$ minutes. If it's currently 11:00 am, when would the manager expect the next customer to enter? If it turns out that no customers enter between 11:00 am and 11:35 am, when would the manager expect the next customer to enter?

    We know that the expected value of the exponential distribution is $\frac{1}{\lambda}$, so we'd expect the next customer in $\frac{1}{0.2} = 5$ minutes—that is, at 11:05. If it turns out that no customer has arrived by 11:35, then we'd expect the next customer at 11:40 since the exponential distribution is memoryless.

19. In a certain region, it rains on average once in every ten days during the summer. Rain is predicted on average for 85% of the days when rainfall actually occurs, while rain is predicted on average for 25% of the days when it does not rain. Assume that rain is predicted for tomorrow. What is the probability of rainfall actually occurring on that day?

    Let H be the event that it rains and $\overline{H}$ be the event that it doesn't rain. Let E be the event that rain is predicted. Then, by Bayes' Rule in Odds Form,

    $$\frac{P(H|E)}{P(\overline{H}|E)} = \frac{0.85}{0.25}\frac{\frac{1}{10}}{\frac{9}{10}} = 0.378.$$

    Converting back to probabilities, $P(H|E) = \frac{0.378}{1+0.378} = 0.274$. (So, because rain is uncommon in this region, it usually doesn't rain even when rain is predicted.)

4 Harder Problems (for those who find the above easy)

20. A fair coin is flipped $n+m$ times, landing heads n times and tails m times. After each flip, a "H" or a "T" is written down, so that afterward there is a record of the form "HHTHTHHTTTH." We call one or more flips coming up the same in a row a "run" (so that, for example, the record "HHTHTHHTTTH" contains 7 runs: "HH," "T," "H," "T," "HH," "TTT," and "H"). What is the expected number of runs? (*Hint:* Compute the number of heads-runs and tails-runs separately and add these together at the end. Note that the ith flip (for $i \neq 1$) begins a heads-run if the ith flip is a heads and the $(i-1)$st flip is a tails. The first flip begins a heads-run if it is a heads and begins a tails-run otherwise.)

    Let X be the number of runs and let $X = X_H + X_T$ where X_H is the number of heads-runs and X_T are the number of tails-runs. Let $X_H = H_1 + \cdots + H_{n+m}$ where H_i is 1 if the ith flip begins a heads-run and 0 otherwise. We compute $E[H_1] = \frac{n}{n+m}$ and $E[H_i] = P(\text{tails in position } i-1 \text{ and heads in position } i) = \frac{m}{n+m} \cdot \frac{n}{n+m-1}$ for $i \neq 1$.

So, $E(X_H) = \frac{n}{n+m} + (n+m-1)\frac{nm}{(n+m)(n+m-1)}$.

A similar calculation gives $E(X_T) = \frac{m}{n+m} + \frac{nm}{n+m}$, so that $E(X) = E(X_H) + E(X_T) = 1 + \frac{2nm}{n+m}$.

21. You know that bowl A has three red and two white balls inside and that bowl B has four red balls and three white balls. Without your being aware of which one it is, one of the bowls is randomly chosen and presented to you. Blindfolded, you must pick two balls out of the bowl. You may proceed according to one of the following strategies:

    (a) you will choose and replace (*i.e.*, you will replace your first ball into the bowl before choosing your second ball).

    (b) you will choose two balls without replacing any (*i.e.*, you will not replace the first ball before choosing a second).

    The blindfold is then removed and the colors of both of the balls you chose are revealed to you. Thereafter you must make a guess as to which bowl your two balls came from. For each of the two possible strategies, determine how you can make your guess depending on the colors you have been shown. Which strategy offers the higher probability for a correct guess as to which bowl the balls came from?

    Make a chance tree for each of the two strategies. Let RR be the event that both balls drawn are red, WW be the event that both are white, and RW be the event that one ball is red and one is white (in either order). Let A be the event that we are picking from bowl A.

    In the case that we replace the balls, we find from the chance tree that $P(A|RR) = 0.524$, $P(A|RW) = 0.495$, and $P(A|WW) = 0.465$, so that A is more likely if we draw two reds and less likely otherwise. So, the best strategy in this case is to guess that the bowl is A if we draw two red balls and to guess that the bowl is B otherwise. Then the probability of winning is $P(\text{guess } A \text{ correctly}) + P(\text{guess } B \text{ corectly}) = \frac{9}{50} + \frac{24+9}{98} = 0.517$.

    In the case that we don't replace the balls, we find from the chance tree that $P(A|RR) = 0.512$, $P(A|RW) = 0.512$, and $P(A|WW) = 0.412$, so in this case the best strategy is to guess B if we draw two white balls and to guess that the bowl is A otherwise. In this case, the probability of winning is $\frac{3+3+3}{20} + \frac{1}{14} = 0.521$.

    So, the best strategy is to not replace the balls and to guess bowl A unless we draw two white balls.

22. Suppose that A flips a fair coin $n+1$ times and B flips a fair coin n times. What is the probability that A will have more heads than B? (*Hint:* Condition on who has more heads after the first n flips.)

    We condition on who has more heads after the first n flips. Let B_1 be the event that A has more heads, B_2 be the event that A and B are tied, and B_3 be the event that B has more heads. Let A be the event that A has more heads than B after the $(n+1)$st flip. Interestingly, we can solve this problem without actually calculating $P(B_i)$ for any i. Note that $P(B_1) = P(B_3)$, so if we let $P(B_1) = a$ and $P(B_2) = b$, then $2a + b = 1$. Then, $P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) = 1 \cdot a + \frac{1}{2} \cdot b + 0 \cdot a = \frac{1}{2}(2a+b) = \frac{1}{2}$.

If you do want to calculate these probabilities, note that $b = P(B_2) = \sum_{k=0}^{n} P(\text{both } A \text{ and } B \text{ have exactly } k \text{ heads}) = \sum_{k=0}^{n} \left(\binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} \right) \left(\binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} \right)$. It can be shown that this last sum simplifies to $\frac{\binom{2n}{n}}{2^{2n}}$. Then $a = \frac{1-b}{2}$.