Mathematics for Business Decisions, Part I

Homework Set 8: Introduction to Random Variables, Expected Value, and Variance

Solutions

NOTE: For more practice problems with solutions, see my practice problem sets and my class notes handout.

Recall: Let \( X \) be a random variable with range \( \{x_1, x_2, \ldots, x_n\} \), where \( x_1 < x_2 < \cdots < x_n \). Then

\[
\mu_X = E(X) = \sum_{i=1}^{n} x_i f_X(x_i) \quad \text{(Expected Value formula)}
\]

\[
\sigma_X^2 = V(X) = \sum_{i=1}^{n} (x_i - \mu_X)^2 f_X(x_i) \quad \text{(Variance formula);} \quad \sigma_X = \sqrt{V(X)} \quad \text{(Standard Deviation)}
\]

The standard deviation is a weighted average, and can be thought of as the average distance from the mean. Let \( X \) be a random variable. If \( \sigma_X \) is small (much less that one), then we are likely to observe \( X \)-values close to the mean \( \mu_X \). Conversely, if \( \sigma_X \) is large (much greater that one), then we are likely to observe \( X \)-values far from the mean \( \mu_X \).

Elementary-Level Problems

Problems 1-2 Consider the experiment of flipping a coin and then rolling a die. To each outcome in the sample space we assign the number

\[ X = \text{number of heads times the face value of the die} \]

Notice that \( X \) is a random variable.

1. Write down the sample space \( \Omega \) (the set of all possible outcomes of the experiment).

Solution: \( \Omega = \{(H,1), (H,2), (H,3), (H,4), (H,5), (H,6), (T,1), (T,2), (T,3), (T,4), (T,5), (T,6)\} \)

2. Write down all the possible events \( X = x \), and give the probabilities associated with each of the events.

Solution: To find the range of \( X \), it helps to explicitly write out the formula for \( X \) in the case of heads and tails separately. If the coin toss results in heads, then \( X((H,i)) = 1 \cdot i \), for \( i = 1,2,3,4,5,6 \). If the coin toss results in tails, then \( X((T,i)) = 0 \cdot i = 0 \), for \( i = 1,2,3,4,5,6 \). Thus the range of \( X \) is \( R_X = \{0,1,2,3,4,5,6\} \). We now write out the associated probabilities:

\[
P(X = 0) = P(\{(T,1), (T,2), (T,3), (T,4), (T,5), (T,6)\}) = \frac{6}{12} = \frac{1}{2}.
\]

\[
P(X = i) = P((H,i)) = \frac{1}{12} \quad \text{for } i = 1, 2, 3, 4, 5, 6.
\]

Notice that the sum of all of the probabilities is one: \( \sum_{i=0}^{6} P(X = i) = 1 \).
Problems 3-6 Consider the experiment of flipping a coin two times. Let \( H \) be the event that a toss turns up heads and \( T \) be the event that a toss turns up tails. To each outcome in the sample space we assign the number

\[
X = \text{number of heads in two tosses} - \text{number of tails in two tosses}.
\]

Notice that \( X \) is a random variable.

3. Write down the sample space \( \Omega \) (the set of the four possible outcomes of the experiment).

\textbf{Solution:} \( \Omega = \{(H,H), (H,T), (T,H), (T,T)\} \)

4. Write down all the possible events \( X = x \), and give the probabilities associated with each of the events.

\textbf{Solution:} To find the range of \( X \) it helps to explicitly write out the formula for \( X \). If the coin toss results in two heads, then \( X((H,H)) = 2 - 0 = 2 \). If the coin toss results in one head and one tail in either order, then \( X((H,T)) = X((T,H)) = 1 - 1 = 0 \). If the coin toss results in two tails, then \( X((T,T)) = 0 - 2 = -2 \). Thus the range of \( X \) is \( R_X = \{-2, 0, 2\} \). We now write out the associated probabilities:

\[
f_X(-2) = P(X = -2) = P((T,T)) = \frac{1}{4}.
\]

\[
f_X(0) = P(X = 0) = P((H,T), (T,H)) = \frac{2}{4} = \frac{1}{2}.
\]

\[
f_X(2) = P(X = 2) = P((H,H)) = \frac{1}{4}
\]

Notice that the sum of all of the probabilities is one: \( \sum_{i=1}^{3} P(X = x_i) = 1 \), where by convention we take \( x_1 = -2 \), \( x_2 = 0 \), and \( x_3 = 2 \).

5. Compute the expected value of \( X \).

(a) -2 (b) -1 (c) 0 (d) 1 (e) 2 (f) none of these

\textbf{Solution:} \( \mu_X \equiv E(X) = \sum_{i=1}^{n} x_i f_X(x_i) = (-2) \frac{1}{4} + (0) \frac{1}{2} + (2) \frac{1}{4} = 0 \).

6. Compute the standard deviation of \( X \).

(a) -2 (b) \( 3\sqrt{2} \) (c) 0 (d) 2 (e) \( \sqrt{2} \) (f) none of these

\textbf{Solution:} \( \sigma_X^2 \equiv V(X) = \sum_{i=1}^{n} (x_i - \mu_X)^2 f_X(x_i) = (-2)^2 \frac{1}{4} + (0)^2 \frac{1}{2} + (2)^2 \frac{1}{4} = 1 + 1 = 2 \). The standard deviation is \( \sigma_X = \sqrt{2} \).
7-12. (Random Variables and Expected Values) \( X \) is a random variable that can only assume the values given in the table below. The corresponding probabilities are also listed in the table.

<table>
<thead>
<tr>
<th>( X )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_X(x) = P(X = x) )</td>
<td>.1</td>
<td>.1</td>
<td>.3</td>
<td>.3</td>
<td>.1</td>
<td>.1</td>
</tr>
</tbody>
</table>

7. What is the range (the allowable values of \( X \)) of the random variable \( X \)?

Solution: We can read the range of \( X \) straight off of the table above: \( R_X = \{0,1,2,3,4,10\} \). By convention we let \( x_1 = 0 \), \( x_2 = 1 \), \( x_3 = 2 \), \( x_4 = 3 \), \( x_5 = 4 \), and \( x_6 = 10 \). We will use this in the formulas below.

8. Find \( P(3 \leq X < 7) \).

(a) 0 (c) .2 (e) .4 (g) .6 (i) .8 (k) 1
(b) .1 (d) .3 (f) .5 (h) .7 (j) .9 (l) none of the these

Solution: \( P(3 \leq X < 7) = P(X = 3) + P(X = 4) = .3 + .1 = .4. \)

9. Find the probability that \( X \) is at least 6.

(b) 0 (c) .2 (e) .4 (g) .6 (i) .8 (k) 1
(b) .1 (d) .3 (f) .5 (h) .7 (j) .9 (l) none of the these

Solution: \( P(X \geq 6) = P(X = 10) = .1 \)

10. Find the probability that \( X \) is at most 6.

(a) 0 (c) .2 (e) .4 (g) .6 (i) .8 (k) 1
(b) .1 (d) .3 (f) .5 (h) .7 (j) .9 (l) none of the these

Solution: \( P(X \leq 6) = 1 - P(X > 6) = 1 - P(X = 10) = 1 - .1 = .9. \)

11. Find \( E(X) \). The formula for expected value is \( E(X) = \sum_{i=1}^{n} x_i P(X = x_i) \).

Solution:
\[
E(X) = \sum_{i=1}^{n} x_i P(X = x_i) = 0(.1) + 1(.1) + 2(.3) + 3(.3) + 4(.1) + 10(.1) = .1 + .6 + .9 + .4 + 1 = 3
\]

12. Find \( V(X) \). The formula for variance is \( V(X) = \sum_{i=1}^{n} (x_i - \mu_X)^2 P(X = x_i) \).

Solution:
\[
V(X) = \sum_{i=1}^{n} (x_i - \mu_X)^2 P(X = x_i) = 0^2(.1) + 1^2(.1) + 2^2(.3) + 3^2(.3) + 4^2(.1) + 10^2(.1) \\
= .1 + 1.2 + 2.7 + 1.6 + 10 = 15.6
\]
Problem 13. (Expected Value) You have been given the opportunity to invest in two different companies: company A and company B. Both companies are reputable, and each is working on potentially important scientific projects. If you invest in company A there is a 30% chance that you lose $30,000, a 50% chance that you break even, and a 20% chance that you make $70,000. If you invest in company B there is a 20% chance that you lose $75,000, a 70% chance that you break even, and a 10% chance that you make $150,000. Based on the expected value of each, which investment should you make?

Solution:

Step 1: Define the events: let

\[ X_A \] = return on investment in company A (profit, which could be negative in the case of a loss)
\[ X_B \] = return on investment in company B

Step 2: Write down the given information:

- Range of \( X_A \) = \{-30,000, 0, 70,000\}
- Range of \( X_B \) = \{-75,000, 0, 150,000\}
- The probabilities associated with each of these values are:
  - \( P(X_A = -30000) = .3 \)
  - \( P(X_A = 0) = .5 \)
  - \( P(X_A = 70000) = .2 \)
  - \( P(X_B = -75000) = .2 \)
  - \( P(X_B = 0) = .7 \)
  - \( P(X_B = 150000) = .1 \)

Step 3: Write down what you are trying to solve for:

Compute the expected returns:

\[
E[X_A] = -30000 \cdot P(X_A = -30000) + 0 \cdot P(X_A = 0) + 70000 \cdot P(X_A = 70000)
\]

\[
= -30000 \cdot (.3) + 0 \cdot (.5) + 70000 \cdot (.2)
\]

\[
= -9000 + 14000 = 5000
\]

\[
E[X_B] = -75000 \cdot P(X_A = -75000) + 0 \cdot P(X_A = 0) + 150000 \cdot P(X_A = 150000)
\]

\[
= -75000 \cdot (.2) + 0 \cdot (.7) + 150000 \cdot (.1)
\]

\[
= 0
\]

Based on the expected returns, you should invest in project A.
Problem 14. (Expected Value) The Fair & Selfless insurance company charges a 40-year-old man $200 for a 1-year term life insurance policy that will pay the beneficiary $10,000 if the man dies within the year. Assuming the probability that a 40-year-old man will die during the next year is known to be .001, and that the $200 premium paid at the beginning of the year will be invested over the period of the entire year at 10% compounded monthly, determine the expected profit at the end of the year, per policy, if the insurance company sells many such policies.

Solution: Let $X$ be the random variable representing the net profit to the insurance company. Let $R$ be the revenue (the return) on a single policy and $C$ be the cost of a single policy to the insurance company. The cost is the payout by the insurance company to the policy holder. The return is a constant, namely 200 dollars plus interest made over the year from the investment, but the cost is a random variable. A policy can cost the company $0 or $10,000, depending on the outcome: either there is no cost to the insurance company if the individual lives through the year, or there is a payout of 10,000 dollars if the individual dies. Thus,

Range of $C = \{0, 10,000\}$,
with the associated probabilities

\[
P(C = 0) = 1 -.001 = .999, \quad P(C = 10,000) = .001.
\]

The expected cost to the company is

\[
E(C) = \sum_{i=1}^{2} x_i P(C = x_i) = 0(.999) + .001(10,000) = 10
\]

The net profit can then be expressed as

\[
X = R - C = P\left(1 + \frac{r}{n}\right)^n - C,
\]

where $P = 200$, $r = .10$, $n = 12$, and $t = 1$ year. Substituting in the numbers yields

\[
X = 200\left(1 + \frac{1}{12}\right)^{12} - C \approx 221 - C.
\]

Fact: If $X$ is a random variables and $\alpha$ is a constant, then

\[
E(X + \alpha) = E(X) + E(\alpha) = E(X) + \alpha E(1) = E(X) + \alpha.
\]

Using this formula, we see that the expected profit to the company is

\[
E(X) = E(221 - C) = 221 - E(C) = 221 - 10 = 211.
\]

If you don't like using this formula, there is a second way to compute the result. Notice that the range of $X$ can be expressed as

Range of $X = 221 - \text{Range of } C = \{221 - 0, 221 - 10,000\} = \{221, -9779\}$

with the associated probabilities:

\[
P(X = 221) = 1 -.001 = .999, \quad P(X = -9779) = .001.
\]

We can now compute the expected return directly.

\[
E(X) = \sum_{i=1}^{2} x_i P(X = x_i) = 221(.999) - 9779(.001) = 211.
\]
Problems 15: (Expected Value) Let \( X \) be the number of children in a household in Manhattan. In a certain year the U.S. Census reported that \( X \) has the probability distribution given in the table below:

<table>
<thead>
<tr>
<th>( x )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4 and up</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(X=x) )</td>
<td>0.3</td>
<td>0.4</td>
<td>0.2</td>
<td>0.099</td>
<td>0.001</td>
</tr>
</tbody>
</table>

15. What is the approximate average number of children in a household in Manhattan?

**Solution:** \( E(X) \approx (0)(.3) + (1)(.4) + (2)(.2) + (3)(.099) + (4)(.001) = .4 + .4 + 3(.1) + .001 = 1.101 \)

Note: You might be wondering why I keep using the word “approximate” with expected value and standard deviation. The reason is that in all of the calculations, I have had to replace the phrase “4 and up” by a number. Given the incomplete data set, I have chosen to assume that most of the families that fall under the category of “4 and up” are 4-child families. However, this assumption could be way off, it maybe that most of the families in the category of “4 and up” are 5-child families, or even worse: 10-child families! However, since the probability of such a large family is so small, the error that I make by assuming all of the “4 and up” families are just 4-child families will be small.

Problem 16: (Expected Value) Hurricanes are classified by categories that are based on wind speeds, where a category 1 hurricane has the lowest wind speeds, and a category 5 hurricane has the highest wind speeds. Let \( X \) be the category of a hurricane that has hit the U.S. mainland. According to USA Today’s Weather Almanac, \( X \) has the probability distribution given in the table below:

<table>
<thead>
<tr>
<th>( x )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P(X=x) )</td>
<td>0.372</td>
<td>0.229</td>
<td>0.288</td>
<td>0.098</td>
<td>0.013</td>
</tr>
</tbody>
</table>

What is \( E(X) \)?

**Solution:** \( E(X) = (1)(.372) + (2)(.229) + (3)(.288) + (4)(.098) + (5)(.013) = 2.151 \)

Problem 17: (Expected Value) A ticket for a certain state lottery costs \$1. Each ticket has a probability of 0.01 of winning \$10, a probability of 0.001 of winning \$100, and a probability of 0.0000001 of winning \$1,000,000. If 10,000,000 tickets are sold, what profit can the state expect to make? (Hint: Look at the expected profit on a single ticket.)

**Solution:** Let \( X \) be the random variable that gives the state’s net profit on a single ticket. Let \( R \) be the revenue (the return) on a single ticket and \( C \) be the cost to the state of a single ticket. The cost is the payout by the state to the ticket holder. The return is a constant, namely 1 dollar, but the cost is a random variable. The payout on a ticket can cost the state 0, 10, 100, 1,000,000 dollars. Then \( X = R - C = 1 - C \). Thus

Range of \( C = \{0, 10, 100, 1,000,000\} \). Subtracting off the value of \( C \) from the ticket price of 1 dollar gives the range of the net profit random variable.
The expected profit on a single ticket is then

\[ E(X) = \sum_{i=1}^{4} x_i P(X = x_i) = 1(.9889999) - 9(.01) - 99(.001) - 999,999(.0000001) = .70. \]

Note: If we let \( X \) be the random variable that gives the net profit on a single ticket to the ticket holder, and let \( R \) be the random variable that gives the revenue (the return) on a single ticket to the ticket holder, then \( X = R - C \), where \( C \) is the cost of a single ticket to the ticket holder. A ticket can return \{0, 10, 100, 1,000,000\} dollars. The reader should verify that \( E(X) = -E(X) \).

Let \( P_{\text{state}} \) = net profit to the state from all of the ticket sales. If the state sells \( n \) tickets, then the expected net profit from all of the lottery tickets is \( E(P_{\text{state}}) = E(nX) = nE(X) \). In our case the state should expect to bring in 10,000,000 (.7) = 7,000,000 dollars.

Problem 18: (Expected Value) An entrepreneur is planning to open a business either in the city or in a suburb. If he locates in the city, there is a 20% chance that he will suffer a loss of $20,000; a 50% chance that he will break even; and a 30% chance that he will make a profit of $100,000. In the suburbs, he has a 30% chance of losing $70,000; a 60% chance of breaking even; and a 10% chance of making a profit of $150,000. Where should he locate his business to maximize the expected value of his profit?

Solution:

Let \( X_C \) = return from the business if it is located in the city.
Let \( X_S \) = return from the business if it is located in the suburb.

Then

\begin{align*}
\text{Range of } & X_C = \{-20,000, 0, 100,000\}, \\
\text{Range of } & X_S = \{-70,000, 0, 150,000\},
\end{align*}

with associated probabilities:

\begin{align*}
P(X_C = -20,000) &= .2 & P(X_S = -70,000) &= .3 \\
P(X_C = 0) &= .5 & P(X_S = 0) &= .6 \\
P(X_C = 100,000) &= .3 & P(X_S = 150,000) &= .1
\end{align*}
Next, we compute the potential expected returns from each investment:

\[
E(X_c) = \sum_{i=1}^{3} x_i P(X_c = x_i) = (-20,000)(.2) + 0(.5) + 100,000(.3) = 26,000
\]

\[
E(X_s) = \sum_{i=1}^{3} x_i P(X_s = x_i) = (-70,000)(.3) + 0(.6) + 150,000(.1) = -6,000
\]

Since the expected return if the business is located in the city is much higher than the expected return from the suburbs (which is actually a loss), the entrepreneur would do best to open his business in the city.

**Problem 19-21: (Expected Value and Variance)** Below is the graph of the probability distribution of the random variable \(X\). Notice that the range of \(X = \{-3, -2, -1, 0, 1, 2, 3\}\).

**19.** Explicitly write out a table for the probability distribution: \(f_X(x_i) = P(X = x_i)\).

<table>
<thead>
<tr>
<th>(x)</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f_X(x))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Solution:**

<table>
<thead>
<tr>
<th>(x)</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f_X(x))</td>
<td>.1</td>
<td>.1</td>
<td>.2</td>
<td>.2</td>
<td>.2</td>
<td>.1</td>
<td>.1</td>
</tr>
</tbody>
</table>

NOTE: the sum of all of the probabilities adds to one, as they must.

**20.** Compute the mean of the random variable \(X\).

**Solution:** \(E(X) = (-3)(.1) + (-2)(.1) + (-1)(.2) + (0)(.2) + (1)(.2) + (2)(.1) + (3)(.1) = 0\)
21. Compute the variance and standard deviation of the random variable $X$.

**Solution:** First compute the variance.

\[
V(X) = (-3)^2(.1) + (-2)^2(.1) + (-1)^2(.2) + (0)^2(.2) + (1)^2(.2) + (2)^2(.1) + (3)^2(.1)
\]
\[
= 2[(1)^2(.2) + (2)^2(.1) + (3)^2(.1)]
\]
\[
= 2[.2 + .4 + .9] = 3
\]

The standard deviation is

\[
\sigma_X = \sqrt{V(X)} = \sqrt{3}.
\]
22-23. Below are two graphs of the probability distributions of two random variables $X$ (graph A) and $Y$ (graph B). Notice that both graphs are symmetric about the origin.

22. By inspection (no computations), determine the expected values of two distributions. RECALL: The expected value is the center of mass of the probability distribution.

**Solution:** Clearly the graphs of the distributions of each of the random variables are centered about their respective origins. The center of mass is at the origin. Thus $E(X) = E(Y) = 0$.

23. By inspection (no computations), determine which random variable has the larger variance, and hence standard deviation. RECALL: The standard deviation is the average distance from the mean.

**Solution:** Clearly the graph of the random variable $Y$ has a larger variance since its probability mass is further from the mean. (It has a larger moment of inertia about the axis $Y = \mu_Y$).
24. (Expected Value) The range of a random variable $X$ is \{0, 1, 2, 3, 4, 5\}. Without knowing any of the probabilities associated with the events $X = x$, determine which of the following values are possible expected values for the random variable $X$. Circle your answer/answers (there may be more than one). **HINT:** Remember the center of mass analogy for expected value (the weights hanging from the stick). You can’t do any simple calculations here. This is a problem where you know it, or you don’t.

(a.) –13  (b.) 5.1  (c.) –0.1  (d.) 10  (e) 3.2

**Solution:** The expected value of $X$ must lie somewhere between the minimum range value and the maximum range value. If

Range of $X = \{x_1, x_2, \ldots, x_n\}$, where $x_1 < x_2 < \cdots < x_n$, then $x_1 \leq E(X) \leq x_n$.

In our case, $x_1 = 0 \leq E(X) \leq 5 = x_6$. The only possibility is (e).

25. (Expected Value) Find the missing probability in the table below and use it to compute the expected value of $X$.

<table>
<thead>
<tr>
<th>$x$</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X=x)$</td>
<td>0.2</td>
<td>0.1</td>
<td>.3</td>
<td></td>
<td>.2</td>
</tr>
</tbody>
</table>

**Solution:** Recall: The sum of all of the probabilities must be one. So if the range of $X$ is \{x_1, x_2, \ldots, x_n\}, then $\sum_{i=1}^{n} P(X = x_i) = 1$. Thus $0.2 + 0.1 + 0.3 + P(X = 1) + 0.2 = 1$. Solving for $P(X = 1)$ yields $P(X = 1) = 1 - 0.8 = 0.2$. The expected value is $E(X) = -1(0.1) + 1(0.2) = 0.1$. 