Discrete Random Variables

COURSE TEACHING ASSISTANTS

Kamal Ezzo kamal.ezzo@must.edu.eg

Maryam Gebraeil maryam.gebraeil@must.edu.eg

Tasneem Abdelraouf tasneem.abdelraouf@must.edu.eg

Rules

1 Expected value (Mean) — the average value of X

$$E(X) = \mu = \sum x\,P(x)$$

TRICK Multiply each x by its probability, then add.

Keywords: *expected, average, mean*

2 Variance, Standard Deviation and CV — how spread out X is

$$V(X) = \sum x^2 P(x) - (\text{Mean})^2$$

$$\text{SD} = \sqrt{V(X)} = \sqrt{\sum x^2 P(x) - (\text{Mean})^2}$$

$$\text{CV} = \frac{\text{Standard Deviation}}{\text{Mean}} \times 100\,\%$$

TRICK Square x, multiply by $P(x)$, add, then take away the mean squared. Variance is never negative.

Bigger CV $\rightarrow$ more variation. Use CV to compare two variables.

Keywords: *variance, standard deviation, spread, coefficient of variation, more variable*

3 Probability distribution (PDF) — is the table a real distribution?

Every $P(x)$ is between 0 and 1: $0 \le P(x) \le 1$

All probabilities add up to 1: $\sum P(x) = 1$

TRICK Missing value (a, c)? Add the table and make it equal 1.

Keywords: *probability distribution, PDF, value of a*

4 Cumulative Distribution Function (CDF) $F(x)$

$$F(x) = P(X \leq x)$$

$$P(X > a) = 1 - F(a)$$

$$P(a < X \leq b) = F(b) - F(a)$$

TRICK Add the probabilities from the left up to x. The graph is steps that go up to 1.

Keywords: *CDF, cumulative, $F(x)$*

5 Rules of E and V a, b are numbers; X, Y independent

	Expected $E(X)$	**Variance** $V(X)$
Constant	$E(5) = 5$ *the same*	$V(5) = 0$ *zero*
Number with X	$E(5X) = 5\,E(X)$	$V(5X) = 5^2\,V(X)$
Both	$E(aX + b) = a\,E(X) + b$	$V(aX + b) = a^2\,V(X)$
Minus sign	$E(3X - 4) = 3\,E(X) - 4$	$V(3X - 4) = 3^2\,V(X)$
Two variables	$E(aX - bY) = aE(X) - bE(Y)$	$V(aX - bY) = a^2V(X) + b^2V(Y)$

TRICK Variance: square the number, drop the constant, minus becomes plus. Expected: keep everything as it is.

Same X twice? Simplify first: $V(5X - 2X) = V(3X) = 9\,V(X)$.

6 Words to signs

less than	$X < 2$	more than	$X > 2$
at least	$X \geq 2$	at most	$X \leq 2$
exactly	$X = 2$	zero to 2	$0 \leq X \leq 2$

TRICK At least = this number or more. At most = this number or less.

Q1

Let X be a random variable with $E(X) = 2.666$ and $V(X) = 1.8364$. Use this data to find:

1. $E(3X)$
2. $V(3X)$
3. $E(4X + 5)$
4. $V(4X + 5)$
5. $E(3X - 1X)$
6. $V(5X - 2X)$
7. $E(3X - 4)$
8. $V(3X - 10)$
9. $V(10X)$
10. $E(5X)$
11. The coefficient of variation

Solution

Given: $E(X) = 2.666, \quad V(X) = 1.8364$

1) $E(3X)$

$$E(3X) = 3\,E(X) = 3 \times 2.666 = 7.998$$

2) $V(3X)$

$$V(3X) = 3^2\,V(X) = 9 \times 1.8364 = 16.5276$$

3) $E(4X + 5)$

$$E(4X + 5) = 4\,E(X) + 5$$

$$= 4 \times 2.666 + 5 = 15.664$$

4) $V(4X+5)$

$V(4X+5) = 4^2\,V(X)$ *the constant 5 goes*

$= 16 \times 1.8364 = \boxed{29.3824}$

5) $E(3X-1X)$

$E(3X-1X) = E(2X)$ *same X: $3X-1X=2X$*

$= 2 \times 2.666 = \boxed{5.332}$

6) $V(5X-2X)$

$V(5X-2X) = V(3X)$ *same X: simplify first*

$= 9 \times 1.8364 = \boxed{16.5276}$

Not $25\,V(X)+4\,V(X)$: there is only one variable here.

7) $E(3X-4)$

$E(3X-4) = 3\,E(X) - 4$

$= 3 \times 2.666 - 4 = \boxed{3.998}$

8) $V(3X-10)$

$V(3X-10) = 3^2\,V(X)$ *the constant 10 goes*

$= 9 \times 1.8364 = \boxed{16.5276}$

9) $V(10X)$

$V(10X) = 10^2\,V(X) = 100 \times 1.8364 = \boxed{183.64}$

10) $E(5X)$

$E(5X) = 5\,E(X) = 5 \times 2.666 = \boxed{13.33}$

11) *Coefficient of variation*

$\text{SD} = \sqrt{V(X)} = \sqrt{1.8364} \approx 1.3551$

$\text{CV} = \dfrac{\text{SD}}{\text{Mean}} \times 100 = \dfrac{1.3551}{2.666} \times 100 \approx \boxed{50.83\,\%}$

Q2

The table shows the number of breakdowns per week of a machine.

Breakdowns per week (x)	**0**	**1**	**2**	**3**
Probability $P(x)$	0.15	0.2	0.35	0.3

1. Is this table a probability distribution?
2. Find the probability of the following:
   a) Zero to 2
   b) More than 1
   c) At most 2
3. Find the mean and the standard deviation.
4. Find $V(3X - 6X)$ and $E(5 - 10X)$.
5. Using the CDF, find $P(X > 2)$, and graph the CDF.
6. Draw the histogram.

Solution

1) *Is it a probability distribution?*

Every $P(x)$ is between 0 and 1, and $0.15 + 0.2 + 0.35 + 0.3 = 1$

Yes

2a) *Zero to 2*

$P(0 \leq X \leq 2) = 0.15 + 0.2 + 0.35 = 0.7$

2b) *More than 1*

$P(X > 1) = P(2) + P(3) = 0.35 + 0.3 = 0.65$

2c) *At most 2*

$P(X \leq 2) = 0.15 + 0.2 + 0.35 = 0.7$ *same as zero to 2*

3) *Mean and standard deviation*

x	$P(x)$	$x\,P(x)$	$x^2P(x)$
0	0.15	0	0
1	0.2	0.2	0.2
2	0.35	0.7	1.4
3	0.3	0.9	2.7
Total	**1**	**1.8**	**4.3**

$$\text{Mean} = \sum x\,P(x) = 1.8$$

$$V(X) = \sum x^2P(x) - (\text{Mean})^2 = 4.3 - 1.8^2 = 4.3 - 3.24 = 1.06$$

$$\text{SD} = \sqrt{1.06} \approx 1.03$$

4) $V(3X - 6X)$ *and* $E(5 - 10X)$

$$V(3X - 6X) = V(-3X) = (-3)^2\,V(X)$$ *same X: simplify first*

$$= 9 \times 1.06 = 9.54$$

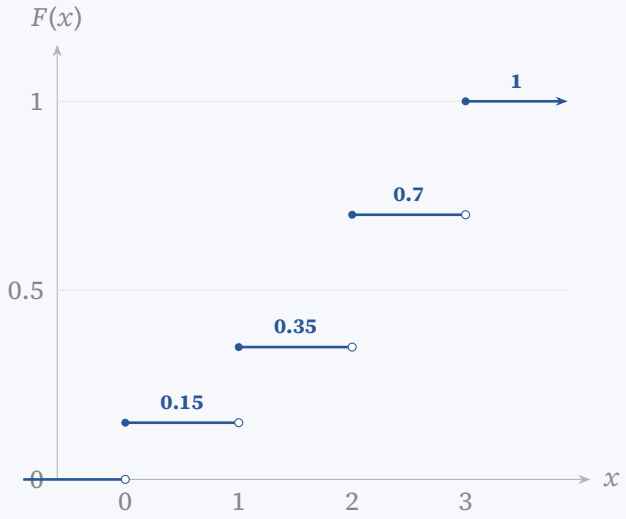
$$E(5 - 10X) = 5 - 10\,E(X)$$

$$= 5 - 10 \times 1.8 = -13$$

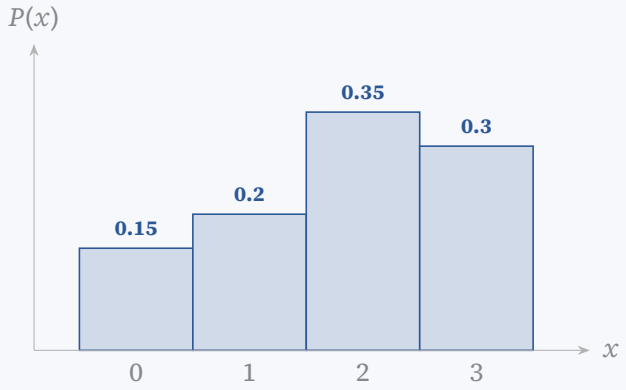
5) *CDF,* $P(X > 2)$ *and the graph*

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.15 & 0 \le x < 1 \\ 0.15 + 0.2 = 0.35 & 1 \le x < 2 \\ 0.35 + 0.35 = 0.7 & 2 \le x < 3 \\ 0.7 + 0.3 = 1 & x \ge 3 \end{cases}$$

$$P(X > 2) = 1 - F(2) = 1 - 0.7 = 0.3$$



6) *Histogram*



Q3

The following data represent the number of expatriate students in 40 schools in 6th of October City.

Class	0–20	20–40	40–60	60–80	80–100	100–120
Frequency	5	13	9	6	5	2

Find:

1. The expected value
2. The variance
3. The standard deviation
4. The coefficient of variation
5. The CDF at $X = 70$
6. $P(X > 70)$
7. $E(4X - 5)$ and $V(3X - 6)$

Solution

Given: Total frequency = 40

First, make the table: x = *midpoint of the class,* $P(x) = \dfrac{f}{40}$

Class	x	f	$P(x)$	$x\,P(x)$	$x^2P(x)$
0–20	10	5	0.125	1.25	12.5
20–40	30	13	0.325	9.75	292.5
40–60	50	9	0.225	11.25	562.5
60–80	70	6	0.15	10.5	735
80–100	90	5	0.125	11.25	1012.5
100–120	110	2	0.05	5.5	605
Total		**40**	**1**	**49.5**	**3220**

1) *Expected value*

$$E(X) = \sum x\,P(x) = 49.5$$

2) *Variance*

$$V(X) = \sum x^2P(x) - (\text{Mean})^2$$

$$= 3220 - 49.5^2 = 3220 - 2450.25 = 769.75$$

3) *Standard deviation*

$$\text{SD} = \sqrt{769.75} \approx 27.74$$

4) *Coefficient of variation*

$$\text{CV} = \frac{27.74}{49.5} \times 100 \approx 56.05\,\%$$

5) *CDF at $X = 70$*

$$F(70) = P(X \le 70) = 0.125 + 0.325 + 0.225 + 0.15 = 0.825$$

6) $P(X > 70)$

$$P(X > 70) = 1 - F(70) = 1 - 0.825 = 0.175$$

7) $E(4X - 5)$ *and* $V(3X - 6)$

$$E(4X - 5) = 4 \times 49.5 - 5 = 193$$

$$V(3X - 6) = 3^2 \times 769.75 = 6927.75$$ *the constant 6 goes*

Q4

x	1	2	3	4	5	6
$P(x)$	0.1	0.02	0.3	a	0.4	0.08

1. The value of a that makes the table a probability distribution (PDF) is
2. $P(X \leq 1)$
3. $P(X \geq 0.02)$
4. $F(X = 4)$
5. $E(X)$

Solution

1) *Value of a*

$$0.1 + 0.02 + 0.3 + a + 0.4 + 0.08 = 1 \quad \textit{the table must add up to 1}$$

$$0.9 + a = 1$$

$$a = 0.1$$

2) $P(X \leq 1)$

$P(X \leq 1) = P(1) = 0.1$ *the smallest x is 1*

3) $P(X \geq 0.02)$

$P(X \geq 0.02) = 1$ *every x (1 to 6) is at least 0.02*

4) $F(X = 4)$

$F(4) = P(X \leq 4) = 0.1 + 0.02 + 0.3 + 0.1 = 0.52$ *F = CDF*

5) $E(X)$

$$E(X) = \sum x\,P(x)$$

$$= 1(0.1) + 2(0.02) + 3(0.3) + 4(0.1) + 5(0.4) + 6(0.08)$$

$$= 0.1 + 0.04 + 0.9 + 0.4 + 2 + 0.48 = 3.92$$

Q5

Let X and Y be two independent random variables such that $E(X) = 2$, $V(X) = 3$, $E(Y) = 7$ and $V(Y) = 1$.

1. The coefficient of variation of the random variable X is
2. The coefficient of variation of the random variable Y is
3. Which random variable is more variable than the other?
4. $E(3X + 4)$
5. $V(3X + 2Y - 5)$

Solution

Given: Independent: $E(X) = 2$, $V(X) = 3$, $E(Y) = 7$, $V(Y) = 1$

1) *CV of X*

$$\text{SD}_X = \sqrt{3} \approx 1.732$$

$$\text{CV}_X = \frac{1.732}{2} \times 100 \approx 86.6\,\%$$

2) *CV of Y*

$$\text{SD}_Y = \sqrt{1} = 1$$

$$\text{CV}_Y = \frac{1}{7} \times 100 \approx 14.29\,\%$$

3) *Which is more variable?*

$86.6\,\% > 14.29\,\%$ X is more variable.

Compare the CV, not the variance.

4) $E(3X + 4)$

$E(3X + 4) = 3\,E(X) + 4 = 3 \times 2 + 4 = 10$

5) $V(3X + 2Y - 5)$

$V(3X + 2Y - 5) = 3^2\,V(X) + 2^2\,V(Y)$ *the constant 5 goes*

$= 9 \times 3 + 4 \times 1 = 31$

Q6

x	0	1	2	3	4	5
$P(x)$	0.2	0.15	$2c$	0.05	0.15	c

1. Find the value of c.
2. Find the following probabilities:
   a) $P(X > 1)$
   b) $P(X \leq 3)$
   c) Exactly 2
   d) Zero to 3
   e) More than 2
   f) At most 4
3. Find the cumulative distribution function (CDF), and graph it.
4. Using the CDF, find $P(1.1 < X < 3.5)$ and $P(X > 3)$.
5. Find the expected value, the variance and the standard deviation.
6. Find $V(3X - 3)$ and $E(5X - 10)$.
7. Draw the histogram.

Solution

1) *Value of c*

$$0.2 + 0.15 + 2c + 0.05 + 0.15 + c = 1 \quad \textit{the table must add up to 1}$$

$$0.55 + 3c = 1$$

$$3c = 0.45 \qquad c = 0.15 \qquad 2c = 0.3$$

x	0	1	2	3	4	5
$P(x)$	0.2	0.15	**0.3**	0.05	0.15	**0.15**

2a) $P(X > 1)$

$$P(X > 1) = 1 - P(0) - P(1) = 1 - 0.35 = 0.65$$

2b) $P(X \leq 3)$

$$P(X \leq 3) = 0.2 + 0.15 + 0.3 + 0.05 = 0.7$$

2c) *Exactly 2*

$$P(X = 2) = 0.3$$

2d) *Zero to 3*

$$P(0 \leq X \leq 3) = 0.2 + 0.15 + 0.3 + 0.05 = 0.7$$

2e) *More than 2*

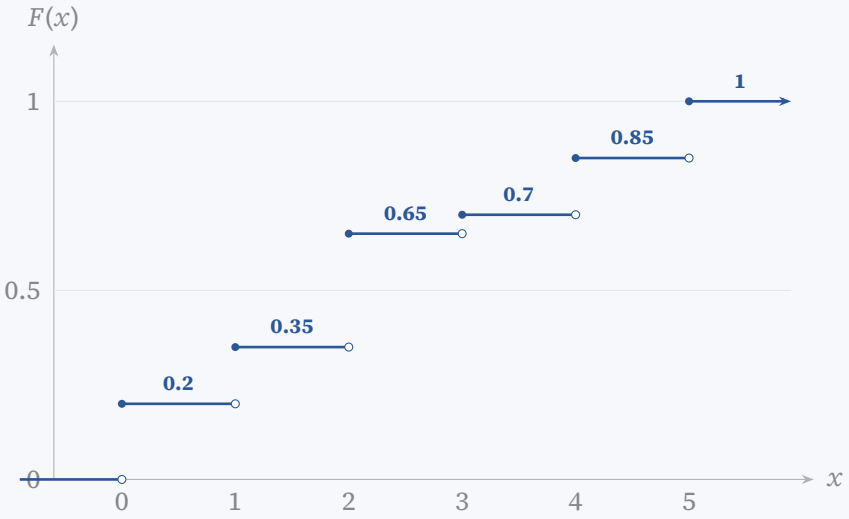
$$P(X > 2) = 0.05 + 0.15 + 0.15 = 0.35$$

2f) *At most 4*

$$P(X \leq 4) = 1 - P(5) = 1 - 0.15 = 0.85$$

3) *CDF and its graph*

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.2 & 0 \leq x < 1 \\ 0.2 + 0.15 = 0.35 & 1 \leq x < 2 \\ 0.35 + 0.3 = 0.65 & 2 \leq x < 3 \\ 0.65 + 0.05 = 0.7 & 3 \leq x < 4 \\ 0.7 + 0.15 = 0.85 & 4 \leq x < 5 \\ 0.85 + 0.15 = 1 & x \geq 5 \end{cases}$$



4) *Using the CDF*

$$P(1.1 < X < 3.5) = F(3) - F(1) \quad X \text{ can be 2 or 3}$$

$$= 0.7 - 0.35 = 0.35$$

$$P(X > 3) = 1 - F(3) = 1 - 0.7 = 0.3$$

5) *Expected value, variance and standard deviation*

x	$P(x)$	$x\,P(x)$	$x^2P(x)$
0	0.2	0	0
1	0.15	0.15	0.15
2	0.3	0.6	1.2
3	0.05	0.15	0.45
4	0.15	0.6	2.4
5	0.15	0.75	3.75
Total	**1**	**2.25**	**7.95**

$$E(X) = 2.25$$

$$V(X) = 7.95 - 2.25^2 = 7.95 - 5.0625 = 2.8875$$

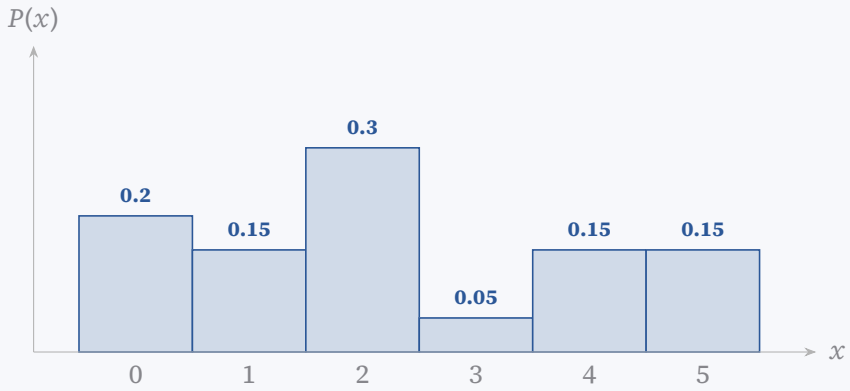
$$\text{SD} = \sqrt{2.8875} \approx 1.70$$

6) $V(3X-3)$ *and* $E(5X-10)$

$$V(3X-3) = 3^2 \times 2.8875 = 25.9875$$ *the constant 3 goes*

$$E(5X-10) = 5 \times 2.25 - 10 = 1.25$$

7) *Histogram*



Q7

A coin is tossed 3 times. Let X be the number of heads. Find:

1. The probability distribution function
2. The cumulative distribution function (CDF), and graph it
3. The expected value, the variance and the standard deviation

Solution

Given: $S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$, $n(S) = 8$

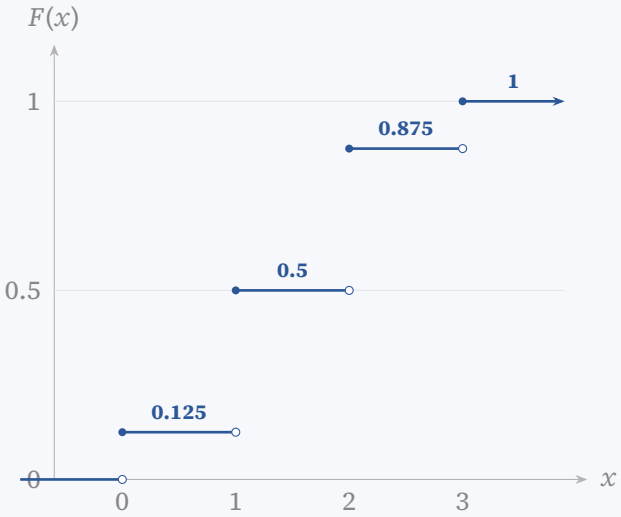
1) *Probability distribution*

0 heads: TTT	1 way
1 head: HTT, THT, TTH	3 ways
2 heads: HHT, HTH, THH	3 ways
3 heads: HHH	1 way

x	0	1	2	3
$P(x)$	1/8 = 0.125	3/8 = 0.375	3/8 = 0.375	1/8 = 0.125

2) *CDF and its graph*

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.125 & 0 \le x < 1 \\ 0.125 + 0.375 = 0.5 & 1 \le x < 2 \\ 0.5 + 0.375 = 0.875 & 2 \le x < 3 \\ 0.875 + 0.125 = 1 & x \ge 3 \end{cases}$$



3) *Expected value, variance and standard deviation*

x	$P(x)$	$x\,P(x)$	$x^2P(x)$
0	0.125	0	0
1	0.375	0.375	0.375
2	0.375	0.75	1.5
3	0.125	0.375	1.125
Total	**1**	**1.5**	**3**

$E(X) = 1.5$

$V(X) = 3 - 1.5^2 = 3 - 2.25 = 0.75$

$\text{SD} = \sqrt{0.75} \approx 0.87$

Q8

Let X be a random variable with values $\{0, 1, 2, 3\}$ and probability distribution function

$$P(x) = \frac{x^2 + 1}{18}, \qquad x = 0, 1, 2, 3$$

Find the following:

1. $P(X \leq 1)$
2. $P(0 < X < 3)$
3. The arithmetic mean and the variance

Solution

First, make the table: put each x in $\dfrac{x^2 + 1}{18}$

x	$P(x)$	$x\,P(x)$	$x^2 P(x)$
0	1/18	0	0
1	2/18	2/18	2/18
2	5/18	10/18	20/18
3	10/18	30/18	90/18
Total	18/18 = 1	42/18	112/18

1) $P(X \leq 1)$

$$P(X \leq 1) = P(0) + P(1) = \frac{1}{18} + \frac{2}{18} = \frac{3}{18} \approx 0.17$$

2) $P(0 < X < 3)$

$$P(0 < X < 3) = P(1) + P(2) = \frac{2}{18} + \frac{5}{18} = \frac{7}{18} \approx 0.39$$

0 and 3 are not included.

3) *Mean and variance*

$$\text{Mean} = \frac{42}{18} \approx 2.33$$

$$V(X) = \sum x^2 P(x) - (\text{Mean})^2$$

$$= \frac{112}{18} - 2.3333^2 \approx 6.2222 - 5.4444 \approx 0.78$$

Q9

If $P(x) = \dfrac{1}{x+3}$, $x = 0, 1, 2, 3, 4$, determine if it is a PDF or not.

Solution

Put each x in $\dfrac{1}{x+3}$

x	0	1	2	3	4
$P(x)$	1/3	1/4	1/5	1/6	1/7

Add the probabilities

$$\sum P(x) = 0.333 + 0.25 + 0.2 + 0.167 + 0.143 \approx 1.093$$

Every $P(x)$ is between 0 and 1, but the total is not 1.

Not a PDF