Chapter 4 - Part 1

Discrete Random Variables

Random Variable - Definition

A random variable is a variable that assumes numerical values associated with the random outcomes of an experiment, where one (and only one) numerical value is assigned to each sample point.

In other words, a random variable is a function $S \to \mathbb{R}$ from a sample space of an experiment to the set of real numbers.

Examples

- **Experiment:** Tossing a coin
- **Sample Space:** $S = \{T, H\}$
- **Random Variable:** x = number of heads
- This random variable assigns $T \mapsto 0$ and $H \mapsto 1$:

$$S = \{\underbrace{T}_{x=0}, \underbrace{H}_{x=1}\}$$

- **Experiment:** Tossing a coin 3 times
- **Sample Space:** $S = \{TTT, TTH, THT, HTT, HHT, HTH, THH, HHH\}$
- **Random Variable:** x = number of heads

$$S = \{\underbrace{TTT}_{x=0}, \underbrace{TTH, THT, HTT}_{x=1}, \underbrace{HHT, HTH, THH}_{x=2}, \underbrace{HHH}_{x=3}\}$$

Definition

- Random variables that can assume a countable number of values are called discrete.
- Random variables that can assume values corresponding to any of the points contained in (at least one) interval are called continuous.

Examples

- The following are some examples of **discrete random variables**:
  - The number of heads obtained when tossing a coin three times: $x = 0, 1, 2, 3$
  - The number of students absent from a class on a given day: $x = 0, 1, 2, \ldots$
  - The number of emails a person receives in one hour: $x = 0, 1, 2, \ldots$
  - The number of defective items in a sample of 20 products: $x = 0, 1, 2, \ldots, 20$
- The following are some examples of **continuous random variables**:
  - The height of a randomly selected person (in cm)
  - The amount of time (in minutes) a student spends studying for an exam
  - The weight (in kilograms) of a bag of rice
  - The temperature (in $^\circ$C) recorded at a weather station
  - The waiting time (in minutes) for a bus to arrive

Probability Distribution Function - Definition

The probability distribution function (PDF) of a discrete random variable is a function p that specifies the probabilty associated with each possible value that the random variable can assume such that

- $0 \leq p(x) \leq 1$ for all x,
- $\sum p(x) = 1$ where the summation of $p(x)$ is over all possible values of x.

Example

Let the random variable X represent the number of heads in a 3-coin toss.

$$S = \{\underbrace{TTT}_{x=0}, \underbrace{TTH, THT, HTT}_{x=1}, \underbrace{HHT, HTH, THH}_{x=2}, \underbrace{HHH}_{x=3}\}$$

$$p(0) = \frac{1}{8}, \quad p(1) = \frac{3}{8}, \quad p(2) = \frac{3}{8}, \quad p(3) = \frac{1}{8}$$

x	0	1	2	3
p(x)	1/8	3/8	3/8	1/8

Another notation:

$$P(X = 1) = \frac{3}{8}, \quad P(X \geq 1) = \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = \frac{7}{8}, \quad P(X > 1) = \frac{3}{8} + \frac{1}{8} = \frac{1}{2}$$

Example

There are 5 marbles in a bag: 2 red and 3 blue.
3 marbles are chosen at random without replacement.
Let the random variable X represent the number of reds chosen.

$$P(X = 0) = \frac{\binom{3}{3} \cdot \binom{2}{0}}{\binom{5}{3}} = \frac{1}{10}$$

$$P(X = 1) = \frac{\binom{3}{2} \cdot \binom{2}{1}}{\binom{5}{3}} = \frac{6}{10} = \frac{3}{5}$$

$$P(X = 2) = \frac{\binom{3}{1} \cdot \binom{2}{2}}{\binom{5}{3}} = \frac{3}{10}$$

x	0	1	2
p(x)	0.1	0.6	0.3

Example

- The roulette game consists of a small ball and a wheel with 38 numbered pockets around the edge. As the wheel is spun, the ball bounces around randomly until it settles down in one of the pockets.
- Suppose random variable X represents the (monetary) outcome of a \$1 bet on a single number ("straight up" bet).
  - If the bet wins, which happens with probability $1/38$, the payoff is \$35;
  - otherwise the player loses the bet.
- PDF of X:

x	-1	35
$p(x)$	$37/38$	$1/38$

Example (cont'd)

x	-1	35
$p(x)$	$37/38$	$1/38$

- What happens in a long run?
- If we play this game 38 million times:
  - about 37 million time, the outcome will be -1 (lose \$1)
  - about 1 million time, the outcome will be 35 (win \$35)
- Which means the outcome is

$$\frac{37 \text{ million} \times (-1) + 1 \text{ million} \times 35}{38 \text{ million}} = \frac{37}{38} \cdot (-1) + \frac{1}{38} \cdot 35 = -\frac{1}{19}$$

  per game.
- We call this the expected value of the random variable X.

Expected Value - Definition

The expected value (or mean) of a discrete random variable X is

$$\mu = E(X) = \sum x \cdot p(x)$$

where the sum is over all possible values of x.

Example

Roll a die. Let X be the outcome. What is the expected value $E(X)$?

x	1	2	3	4	5	6
$p(x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

$$E(X) = \sum x \cdot p(x) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5$$

- **This doesn't mean:** If one rolls the die, we expect the outcome to be 3.5.
  - This is not true!
- **This means:** If one rolls the die n times and computes the average of the results, then as n grows, the average will be about 3.5.
  - Law of large numbers

Example

Roll a pair of dice. Let X be the sum of two dice. What is the expected value $E(X)$?

x	2	3	4	5	6	7	8	9	10	11	12
$p(x)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/26$	$1/36$

$$\begin{aligned} E(X) = \sum x \cdot p(x) &= 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{6}{36} \\ &+ 8 \cdot \frac{5}{36} + 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} \\ &= \frac{252}{36} \\ &= 7 \end{aligned}$$

Example

A weighted die with pdf

x	1	2	3	4	5	6
$p(x)$	0.4	0.2	0.15	0.1	0.1	0.05

is rolled. What is the expected value $E(X)$?

$$E(X) = \sum x \cdot p(x) = 1 \cdot 0.4 + 2 \cdot 0.2 + 3 \cdot 0.15 + 4 \cdot 0.1 + 5 \cdot 0.1 + 6 \cdot 0.05 = 2.45$$

Example

- An insurance company sells a \$10 000 one-year life insurance policy at an annual premium of \$290.
- They estimate that the probabilty of death during next year for a person of that costumer's age, sex, health, etc. is 0.001.
- What is the expected gain of this insurance company from this costumer?

S	costumer lives	costumer dies
x	290	$290 - 10\,000 = -9\,710$
$p(x)$	0.999	0.001

$$E(X) = 290 \cdot 0.999 + (-9\,710) \cdot 0.001 = 280 > 0$$

Example

- Consider the random variable X with PDF

x	0	1	2	3
p(x)	1/8	3/8	3/8	1/8

- Repeat this experiment 8 million times.
  - about 1 million time the outcome will be 0
  - about 3 million time the outcome will be 1
  - about 3 million time the outcome will be 2
  - about 1 million time the outcome will be 3
- This gives us a (population) data set:

$$\{\underbrace{0, 0, \ldots, 0}_{\text{1 million}}, \underbrace{1, 1, \ldots, 1}_{\text{3 million}}, \underbrace{2, 2, \ldots, 2}_{\text{3 million}}, \underbrace{3, 3, \ldots, 3}_{\text{1 million}}\}$$

Example (cont'd)

In this data set

$$\{\underbrace{0, 0, \ldots, 0}_{\text{1 million}}, \underbrace{1, 1, \ldots, 1}_{\text{3 million}}, \underbrace{2, 2, \ldots, 2}_{\text{3 million}}, \underbrace{3, 3, \ldots, 3}_{\text{1 million}}\}$$

- Relative frequency of 1 is

$$\frac{\text{3 million}}{\text{8 million}} = \frac{3}{8} = p(1)$$

- Mean is

$$\begin{aligned}\mu &= \frac{\text{1 million} \times 0 + \text{3 million} \times 1 + \text{3 million} \times 2 + \text{1 million} \times 3}{\text{8 million}} \\ &= \frac{1}{8} \times 0 + \frac{3}{8} \times 1 + \frac{3}{8} \times 2 + \frac{1}{8} \times 3 = 1.5 \\ &= E(X)\end{aligned}$$

Example (cont'd)

In this data set

$$\{\underbrace{0,0,\ldots,0}_{\text{1 million}},\underbrace{1,1,\ldots,1}_{\text{3 million}},\underbrace{2,2,\ldots,2}_{\text{3 million}},\underbrace{3,3,\ldots,3}_{\text{1 million}}\}$$

- Variance is

$$\begin{aligned}\sigma^2 &= \frac{(1M)\times(0-1.5)^2+(3M)\times(1-1.5)^2+(3M)\times(2-1.5)^2+(1M)\times(3-1.5)^2}{8M}\\ &= \frac{1}{8}\times(0-1.5)^2+\frac{3}{8}\times(1-1.5)^2+\frac{3}{8}\times(2-1.5)^2+\frac{1}{8}\times(3-1.5)^2\\ &= \frac{3}{4} = 0.75\end{aligned}$$

Variance - Definition

The variance of a discrete random variable X is

$$\sigma^2 = \mathrm{Var}(X) = \sum (x - \mu)^2 \cdot p(x)$$

where $\mu = E(X)$ is the expected value of x.

The standard deviation of a discrete random variable X is the square root of its variance.

$$\mathrm{StDev}(X) = \sigma = \sqrt{\sigma^2}$$

Example

Consider the pdf for the random variable X shown here:

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05

Find the expected value, variance and standard deviation of it.

- The expected value is

$$\mu = E(X) = 10 \cdot 0.1 + 15 \cdot 0.2 + 20 \cdot 0.3 + 25 \cdot 0.25 + 30 \cdot 0.1 + 35 \cdot 0.05 = 21$$

- Calculating $(x - \mu)^2$ for each x:

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05
$x - \mu$	-11	-6	-1	4	9	14
$(x - \mu)^2$	121	36	1	16	81	196

Example (cont'd)

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05
$x - \mu$	-11	-6	-1	4	9	14
$(x - \mu)^2$	121	36	1	16	81	196

- The variance is

$$\sigma^2 = \mathrm{Var}(X) = 121 \cdot 0.1 + 36 \cdot 0.2 + 1 \cdot 0.3 + 16 \cdot 0.25 + 81 \cdot 0.1 + 196 \cdot 0.05 = 41.5$$

- The standard deviation is

$$\sigma = \sqrt{\mathrm{Var}(X)} = \sqrt{41.5} \approx 6.442$$

Random Variable vs Dataset

The **exptected value** and **variance** of the random variable

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05

is the same as the **mean** and **population variance** of a dataset of the form

$$\{\underbrace{10, 10, \ldots, 10}_{10\%}, \underbrace{15, 15, \ldots, 15}_{20\%}, \underbrace{20, 20, \ldots, 20}_{30\%}, \underbrace{25, 25, \ldots, 25}_{25\%}, \underbrace{30, 30, \ldots, 30}_{10\%}, \underbrace{35, 35, \ldots, 35}_{5\%}\}$$

An example of such dataset:

$$\{\underbrace{10, 10}_{2}, \underbrace{15, 15, 15, 15}_{4}, \underbrace{20, 20, 20, 20, 20, 20}_{6}, \underbrace{25, 25, 25, 25, 25}_{5}, \underbrace{30, 30}_{2}, \underbrace{35}_{1}\}$$

Proposition

$$\sigma^2 = \mathrm{Var}(X) = \left(\sum x^2 p(x)\right) - \mu^2, \qquad \text{where } \mu = E(X).$$

Proof.

$$\begin{aligned}
\sigma^2 &= \sum (x-\mu)^2 p(x) \\
&= \sum x^2 p(x) - \sum 2x\mu p(x) + \sum \mu^2 p(x) \\
&= \sum x^2 p(x) - 2\mu \underbrace{\sum xp(x)}_{\mu} + \mu^2 \underbrace{\sum p(x)}_{1} \\
&= \sum x^2 p(x) - 2\mu^2 + \mu^2 \\
&= \sum x^2 p(x) - \mu^2
\end{aligned}$$

Example

Consider the last example:

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05
x^2	100	225	400	625	900	1225

We found $\mu = E(X) = 21$.

$$\begin{aligned}
\mathrm{Var}(X) &= \left(\sum x^2 p(x)\right) - \mu^2 \\
&= (100 \cdot 0.1 + 225 \cdot 0.2 + 400 \cdot 0.3 + 625 \cdot 0.25 + 900 \cdot 0.1 + 1225 \cdot 0.05) - (21)^2 \\
&= 482.5 - 441 \\
&= 41.5
\end{aligned}$$

Cumulative Distribution Function - Definition

Let X be a random variable with PDF P. The cumulative distribution function (CDF) of X is given by

$$F(y) = P(X \leq y).$$

Example: Consider

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05

Then,

$$\begin{aligned}
F(8) &= P(X \leq 8) = 0 \\
F(10) &= P(X \leq 10) = p(10) = 0.1 \\
F(12) &= P(X \leq 12) = p(10) = 0.1 \\
F(15) &= P(X \leq 15) = p(10) + p(15) = 0.1 + 0.2 = 0.3 \\
F(17) &= P(X \leq 17) = p(10) + p(15) = 0.1 + 0.2 = 0.3 \\
F(20) &= P(X \leq 20) = p(10) + p(15) + p(20) = 0.1 + 0.2 + 0.3 = 0.6 \\
F(21) &= P(X \leq 21) = p(10) + p(15) + p(20) = 0.1 + 0.2 + 0.3 = 0.6
\end{aligned}$$

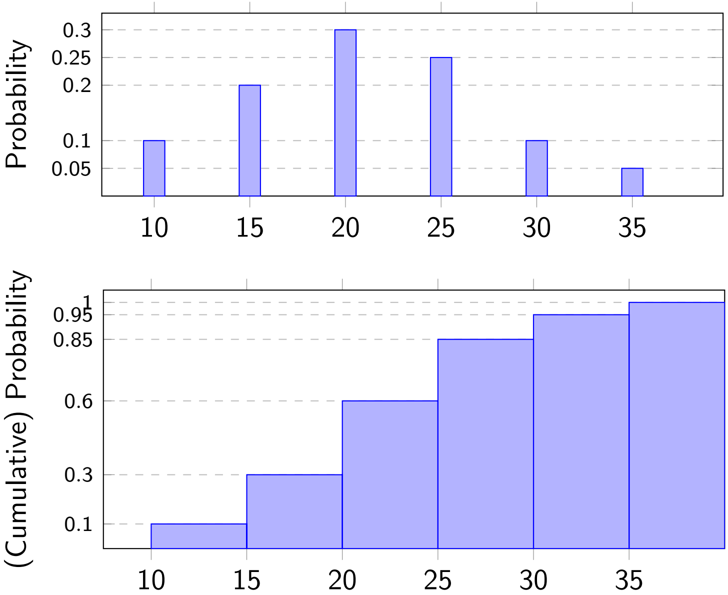
Example

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05

has CDF

$$F(x) = \begin{cases} 0 & \text{if } x < 10 \\ 0.1 & \text{if } 10 \leq x < 15 \\ 0.3 & \text{if } 15 \leq x < 20 \\ 0.6 & \text{if } 20 \leq x < 25 \\ 0.85 & \text{if } 25 \leq x < 30 \\ 0.95 & \text{if } 30 \leq x < 35 \\ 1 & \text{if } 35 \leq x \end{cases}$$

PDF vs CDF


Probability
0.3
0.25
0.2
0.1
0.05
10
15
20
25
30
35
(Cumulative) Probability
1
0.95
0.85
0.6
0.3
0.1
10
15
20
25
30
35

Example

If X has CDF

$$F(x) = \begin{cases} 0 & \text{if } x < 1 \\ 0.4 & \text{if } 1 \leq x < 3 \\ 0.7 & \text{if } 3 \leq x < 5 \\ 0.9 & \text{if } 5 \leq x < 7 \\ 1 & \text{if } 7 \leq x \end{cases}$$

then X has PDF

x	1	3	5	7
$p(x)$	0.4	0.3	0.2	0.1

Function of a Random Variable

For a random variable $X : S \to \mathbb{R}$ and a function $f : \mathbb{R} \to \mathbb{R}$, their composition $f(X) : S \to \mathbb{R}$ is also a random variable (on the same sample space) whose pdf is given by $p(f(x)) = p(x)$.

Example: Consider

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05

Applying $f(x) = x - 4$:

$y = x - 4$	6	11	16	21	26	31
$p(y) = p(x)$	0.1	0.2	0.3	0.25	0.1	0.05

Example

Consider

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05

Applying $f(x) = 2x - 3$:

$y = 2x - 3$	17	27	37	47	57	67
$p(y)$	0.1	0.2	0.3	0.25	0.1	0.05

Notice that

$$\begin{aligned} E(2X - 3) &= \sum (2x - 3)p(x) \\ &= 17 \cdot 0.1 + 27 \cdot 0.2 + 37 \cdot 0.3 + 47 \cdot 0.25 + 57 \cdot 0.1 + 67 \cdot 0.05 = 39 \end{aligned}$$

and

$$2E(X) - 3 = 2 \cdot 21 - 3 = 39$$

Proposition

Let $a, b \in \mathbb{R}$ be two constants and X be a random variable. Then,

$$E(aX + b) = aE(X) + b$$

and

$$\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$$

therefore

$$\mathrm{StDev}(aX + b) = |a|\mathrm{StDev}(X)$$

Example

Consider again

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05

Recall that $\mu = E(X) = 21$ and $\sigma = \mathrm{StDev}(X) = 6.442$. Consider $X - \mu$:

$y = x - \mu$	-11	-6	-1	4	9	14
$p(y)$	0.1	0.2	0.3	0.25	0.1	0.05

Then $E(X - \mu) = E(X) - \mu = 0$ and $\mathrm{StDev}(X - \mu) = \mathrm{StDev}(X) = 6.442$. Consider $\frac{X-\mu}{\sigma}$:

$z = (x - \mu)/\sigma$	$-11/\sigma$	$-6/\sigma$	$-1/\sigma$	$4/\sigma$	$9/\sigma$	$14/\sigma$
$p(z)$	0.1	0.2	0.3	0.25	0.1	0.05

Then $E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\underbrace{E(X-\mu)}_{0} = 0$ and $\mathrm{StDev}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\underbrace{\mathrm{StDev}(X-\mu)}_{\sigma} = 1.$

Standardization

Let X be a random variable with $E(X) = \mu$ and $\mathrm{StDev}(X) = \sigma$.
The random variable

$$Z = \frac{X - \mu}{\sigma}$$

is called the standardization of X.

We have

$$E(Z) = 0, \qquad \mathrm{StDev}(Z) = 1$$

Example

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05

Recall that $\mu = E(X) = 21$.

$y = (x-\mu)^2$	121	36	1	16	81	196
$p(y)$	0.1	0.2	0.3	0.25	0.1	0.05

$$E\left((X-\mu)^2\right) = 41.5 = \mathrm{Var}(X)$$

Proposition

Let X be a random variable with $E(X) = \mu$. Then

$$\mathrm{Var}(X) = E\left((X-\mu)^2\right).$$

Proof.

It follows from the definition of variance:

$$\mathrm{Var}(X) = \sum (x-\mu)^2 p(x) = E\left((X-\mu)^2\right)$$

□

Proposition

Let X be a random variable with $E(X) = \mu$. Then

$$\begin{aligned} \mathrm{Var}(X) &= E(X^2) - E(X)^2 \\ &= E(X^2) - \mu^2 \end{aligned}$$

Proof.

$$\mathrm{Var}(X) = \underbrace{\left(\sum x^2 p(x)\right)}_{E(X^2)} - \mu^2$$

$\square$

Example

Consider

x	10	15	20	25	30	35
$p(x)$	0.1	0.2	0.3	0.25	0.1	0.05

Applying $f(x) = x^2$:

$y = x^2$	100	225	400	625	900	1225
$p(y)$	0.1	0.2	0.3	0.25	0.1	0.05

We have

$$E(X) = 21, \qquad E(X^2) = \left(\sum x^2 p(x)\right) = 482.5$$

and

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = 482.5 - 441 = 41.5.$$