HCMC University of Technology
Dung Nguyen

Probability and Statistics

Random Variables and Random Vectors

Random Variable and its Notation

A variable that associates a number $X(u)$ with the outcome u of a random experiment is called a random variable.

$$X : \Omega \to \mathbb{R}$$
$$u \to X(u)$$

- Uppercase letters (X, Y, Z): Random variables.

Random Variable and its Notation

A variable that associates a number $X(u)$ with the outcome u of a random experiment is called a random variable.

$$X : \Omega \to \mathbb{R}$$
$$u \to X(u)$$

- Uppercase letters (X, Y, Z): Random variables.
- Lowercase letters (x, y, z): Measured values of random variables (after the experiment is conducted). Eg. $x = 70$ milliamperes.

Discrete Random Variables

A discrete random variable is a random variable with a finite or countably infinite range. Its values are obtained by counting.

- Number of scratches on a surface.

Discrete Random Variables

A discrete random variable is a random variable with a finite or countably infinite range. Its values are obtained by counting.

- Number of scratches on a surface.
- Number of defective parts among 100 tested.

Discrete Random Variables

A discrete random variable is a random variable with a finite or countably infinite range. Its values are obtained by counting.

- Number of scratches on a surface.
- Number of defective parts among 100 tested.
- Number of transmitted bits received in error.

Discrete Random Variables

A discrete random variable is a random variable with a finite or countably infinite range. Its values are obtained by counting.

- Number of scratches on a surface.
- Number of defective parts among 100 tested.
- Number of transmitted bits received in error.
- Number of common stock shares traded per day.

Continuous Random Variables

A continuous random variable is a random variable with an interval (either finite or infinite) of real numbers for its range. Its values are obtained by measuring.

- Electrical current and voltage.

Continuous Random Variables

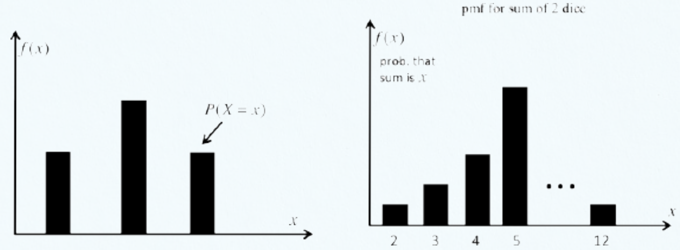
A continuous random variable is a random variable with an interval (either finite or infinite) of real numbers for its range. Its values are obtained by measuring.

- Electrical current and voltage.
- Physical measurements, e.g., length, weight, time, temperature, pressure.

Discrete Distributions

The probability mass function of X is defined by

$$f_X(u) = P(X = u)$$

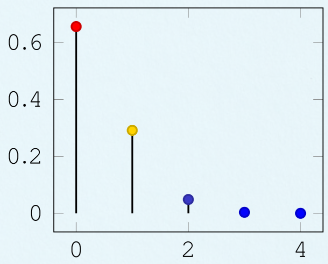


For any event A: $P(X \in A) = \sum_{u \in A} f(u)$.

Example 1 – Digital Channel

- There is a chance that a bit transmitted through a digital transmission channel is received in error.
- X:the number of bits received in error in 4 bits transmitted.

$P(X = 0)$	=	0.6561
$P(X = 1)$	=	0.2916
$P(X = 2)$	=	0.0486
$P(X = 3)$	=	0.0036
$P(X = 4)$	=	0.0001
Total		1.0000

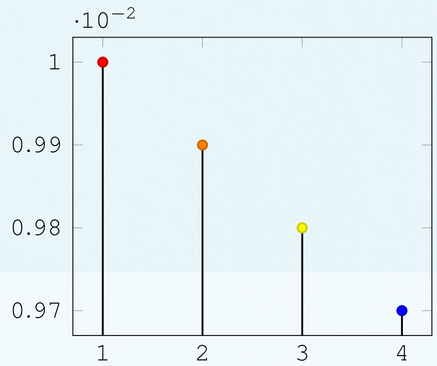


Example 2 – Wafer Contamination

X: the number of wafers that need to be analyzed to detect a large particle of contamination. Assume that the probability that a wafer contains a large particle is 0.01, and that the wafers are independent.

- $\Omega = \{p, ap, aap, aaap, \ldots\}$.
- The range of X: $\{1, 2, 3, 4, \ldots\}$.

Probability Distribution		
$P(X = 1) =$	0.01	0.01
$P(X = 2) =$	$(0.99)(0.01)$	0.0099
$P(X = 3) =$	$(0.99)^2(0.01)$	0.0098
$P(X = 4) =$	$(0.99)^3(0.01)$	0.0097
...	...	...
Total		1



Proposition (Characteristic properties)

A discrete function f is a probability mass function iff

1. $f(u) \geq 0$ *for all* u.
2. $\sum_u f(u) = 1$.

Example

3. Uniform distribution: $\Omega = \{1, 2, 3, \cdots, n\}$

$$f(k) = P(X = k) = \frac{1}{n}$$

each outcome has equal probability

Probability Mass/Density Functions Discrete random variables

Example

3. Uniform distribution: $\Omega = \{1, 2, 3, \cdots, n\}$

$$f(k) = P(X = k) = \frac{1}{n}$$

each outcome has equal probability

4. Is $f(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ $(k = 0, 1, 2 \ldots)$ a probability mass function?

Probability Mass/Density Functions Discrete random variables

Example

3. Uniform distribution: $\Omega = \{1, 2, 3, \cdots, n\}$

$$f(k) = P(X = k) = \frac{1}{n}$$

each outcome has equal probability

4. Is $f(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ $(k = 0, 1, 2 \ldots)$ a probability mass function?

5. Suppose that a random variable X has a discrete distribution with the following p.m.f.

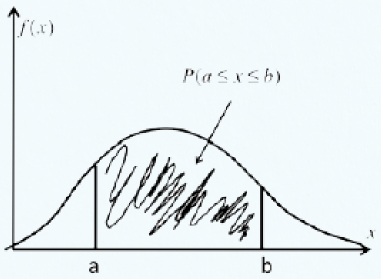
$$f(u) = \begin{cases} cu, & u = 1, \cdots, 5 \\ 0, & \text{otherwise} \end{cases}$$

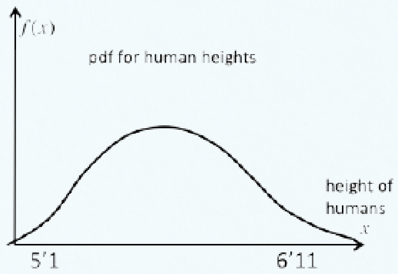
Determine the value of the constant c.

Continuous Distribution

A random variable X is continuous if $\exists f \geq 0$ such that for any $[a,b]$

$$P(a \leq X \leq b) = \int_a^b f(u)du$$





Proposition (Characteristic properties)

If f is a (probability) density function then

1. $f(u) \geq 0$
2. $\int_{-\infty}^{\infty} f(u)du = 1$

Example 6 – Current

Let X denote the current measured in a thin copper wire in milliamperes(mA). Assume that the range of X is $4.9 \leq x \leq 5.1$ and $f(x) = 5$. What is the probability that a current is

a. between 4.95mA and 5.1mA?
b. less than 5mA?

Solution

$$P(4.95 < X < 5.1) = \int_{4.95}^{5.1} f(x)dx = \int_{4.95}^{5.1} 5dx = 0.75$$

$$P(X < 5) = \int_{4.9}^{5} f(x)dx = \int_{4.9}^{5} 5dx = 0.5$$

Example

7. Uniform distribution on $[a,b]$

$$f(u) = \begin{cases} \frac{1}{b-a}, & u \in [a,b] \\ 0, & u \notin [a,b] \end{cases}$$

Example

7 Uniform distribution on $[a, b]$

$$f(u) = \begin{cases} \frac{1}{b-a}, & u \in [a,b] \\ 0, & u \notin [a,b] \end{cases}$$

8 Suppose that the p.d.f. of X is as

$$f(u) = \begin{cases} cu, & 0 < u < 4 \\ 0, & \text{otherwise} \end{cases}$$

Find c. Then determine $P(1 \leq X \leq 2)$ and $P(X > 2)$.

Probability Mass/Density Functions Continuous random variables

Example

7 Uniform distribution on $[a, b]$

$$f(u) = \begin{cases} \frac{1}{b-a}, & u \in [a,b] \\ 0, & u \notin [a,b] \end{cases}$$

8 Suppose that the p.d.f. of X is as

$$f(u) = \begin{cases} cu, & 0 < u < 4 \\ 0, & \text{otherwise} \end{cases}$$

Find c. Then determine $P(1 \leq X \leq 2)$ and $P(X > 2)$.

9 Suppose that X is a continuous random variable whose probability density function is given by

$$f(u) = \begin{cases} c(4u - 2u^2), & 0 < u < 2 \\ 0, & \text{otherwise} \end{cases}$$

Find c and $P(X > 1)$.

Example

10 The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function

$$f(u) = \begin{cases} \lambda e^{-u/100}, & u \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Compute the probability that a computer will function

Probability Mass/Density Functions Continuous random variables

Example

10 The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function

$$f(u) = \begin{cases} \lambda e^{-u/100}, & u \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Compute the probability that a computer will function

(a) between 50 and 150 hours before breaking down?

Example

10 The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function

$$f(u) = \begin{cases} \lambda\, e^{-u/100}, & u \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Compute the probability that a computer will function

(a) between 50 and 150 hours before breaking down?
(b) for fewer than 100 hours?

Example 11 – Digital channel

Consider the probability distribution for the digital channel example.

x	$P(X = x)$
0	0.6561
1	0.2916
2	0.0486
3	0.0036
4	0.0001
	1.0000

Find the probability of three or fewer bits in error.

- The event $(X \leq 3)$ is the total of the events: $(X = 0)$, $(X = 1)$, $(X = 2)$, and $(X = 3)$.

Example 11 – Digital channel

Consider the probability distribution for the digital channel example.

x	$P(X = x)$
0	0.6561
1	0.2916
2	0.0486
3	0.0036
4	0.0001
	1.0000

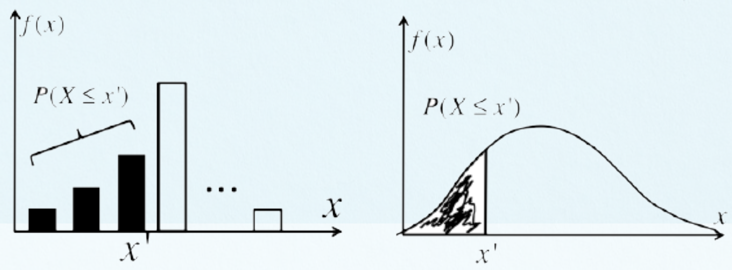
Find the probability of three or fewer bits in error.

- The event $(X \leq 3)$ is the total of the events: $(X = 0)$, $(X = 1)$, $(X = 2)$, and $(X = 3)$.
- From the table:

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 0.9999$$

Cumulative Distribution Function

$$\begin{aligned} F(u) &= P(X \leq u) \\ &= \sum_{k \leq u} f(k) \quad \text{(discrete distribution)} \\ &= \int_{-\infty}^{u} f(t)\,dt \quad \text{(continuous distribution)} \end{aligned}$$



Example 12 – Defective parts

A day's production of 850 parts contains 50 defective parts. Two parts are selected at random without replacement. Let the random variable X equal the number of defective parts in the sample. Find the cumulative distribution function of X. The probability mass function is calculated as follows:

$$P(X=0) = \frac{800}{850}\cdot\frac{799}{849} = 0.886$$

$$P(X=1) = 2\cdot\frac{800}{850}\cdot\frac{50}{849} = 0.111$$

$$P(X=2) = \frac{50}{850}\cdot\frac{49}{849} = 0.003$$

Cumulative Distribution Function

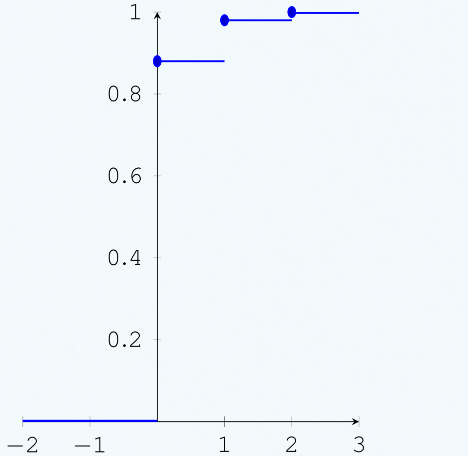
Example 12 – Defective parts

$$P(X=0) = 0.886, P(X=1) = 0.111, P(X=2) = 0.003.$$

$$F(0) = P(X \le 0) = 0.886$$
$$F(1) = P(X \le 1) = 0.997$$
$$F(2) = P(X \le 2) = 1.000.$$

$$F(-0.8) = P(X \le -0.8) = 0,$$
$$F(0.5) = P(X \le 0.5) = F(0),$$
$$F(1.4) = P(X \le 1.4) = F(1),$$
$$F(2.3) = P(X \le 2.3) = F(2).$$

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.886, & 0 \le x < 1 \\ 0.997, & 1 \le x < 2 \\ 1, & x \ge 2 \end{cases}$$

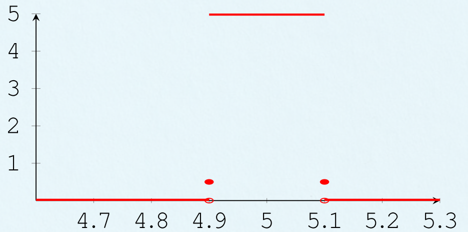


Dung Nguyen Probability and Statistics 16 / 49

Cumulative Distribution Function

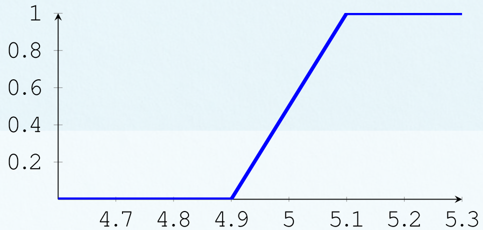
Example 13 – Electric Current

Consider the current measured in a thin copper wire in milliamperes (mA). Recall that the range of X is $4.9 \le x \le 5.1$ and $f(x) = 5$.



The cdf

$$F(x) = \begin{cases} 0, & x < 4.9 \\ 5(x-4.9), & 4.9 \le x \le 5.1 \\ 1, & 5.1 \le x \end{cases}.$$



Cumulative Distribution Function

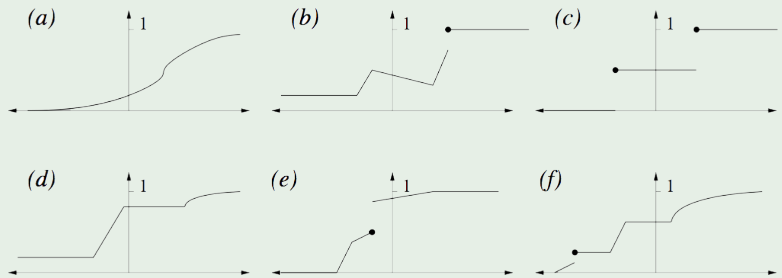
Properties

Proposition (Characteristic properties)

1. $F(u)$ is nondecreasing
2. $F(-\infty) = 0$ and $F(\infty) = 1$
3. F is right continuous: $\lim_{u \to a^+} F(u) = F(a)$

Example (14)

Which of the six functions shown are valid CDFs?



Other properties

Proposition ($F \implies$ Probability)

1. $P(X < u) = F(u^-)$ and $P(X = u) = F(u) - F(u^-)$.
2. Probability of random variable occurring within an interval

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

Proposition ($F \implies f$)

1. If X has a discrete distribution then

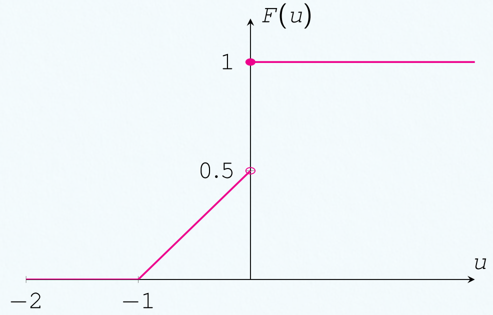
$$f(u) = F(u) - F(u^-)$$

2. If X has a continuous distribution, then F is continuous at every u and $F'(u) = f(u)$, i.e.

$$f(u) = F'(u)$$

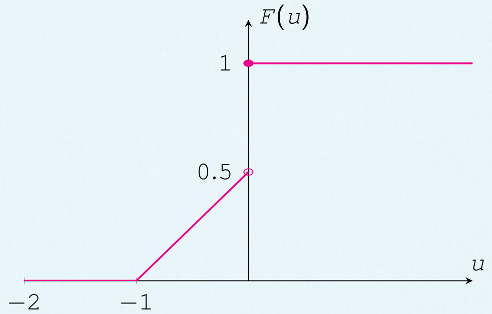
Example

15 Let X have the CDF



Example

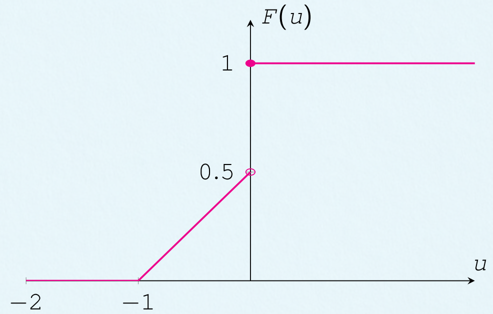
15 Let X have the CDF



(a) Determine all values of u such that $P(X = u) > 0$.

Example

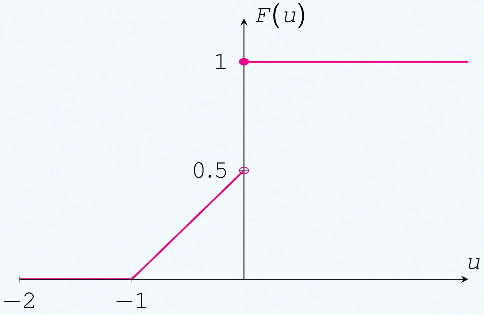
15 Let X have the CDF



(a) Determine all values of u such that $P(X = u) > 0$.
(b) Find $P(X \leq 0)$

Example

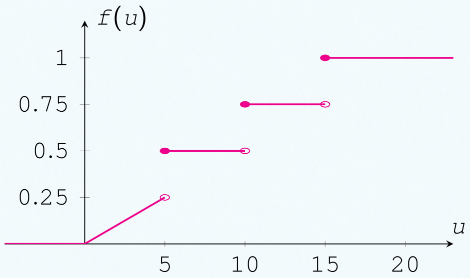
15 Let X have the CDF



ⓐ Determine all values of u such that $P(X = u) > 0$.
ⓑ Find $P(X \leq 0)$
ⓒ Find $P(X < 0)$.

Example

16 Let X have the CDF

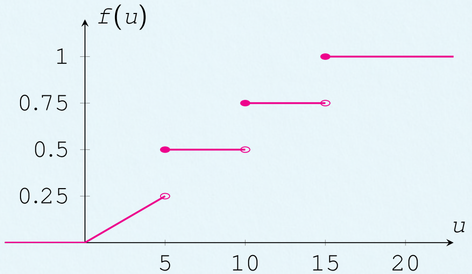


Find the numerical values of the following quantities

Cumulative Distribution Function

Example

16 Let X have the CDF



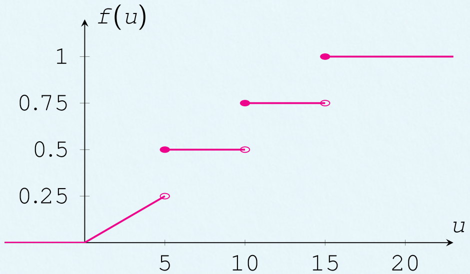
Find the numerical values of the following quantities
ⓐ $P(X \leq 1)$

Dung Nguyen Probability and Statistics21 / 49

Cumulative Distribution Function

Example

16 Let X have the CDF

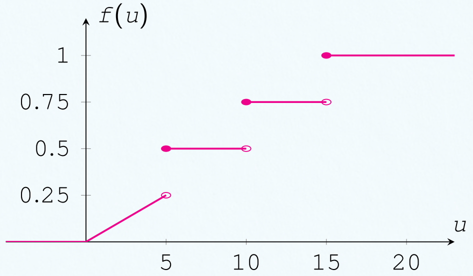


Find the numerical values of the following quantities
ⓐ $P(X \leq 1)$
ⓑ $P(X \leq 10)$

Dung Nguyen Probability and Statistics21 / 49

Example

16 Let X have the CDF

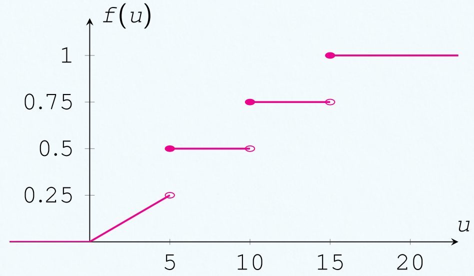


Find the numerical values of the following quantities

ⓐ $P(X \leq 1)$
ⓑ $P(X \leq 10)$
ⓒ $P(X \geq 10)$

Example

16 Let X have the CDF

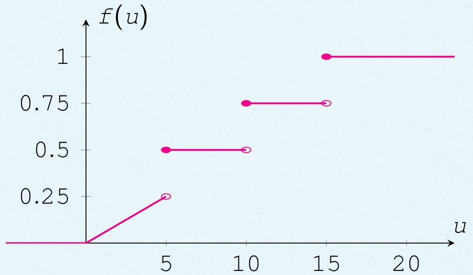


Find the numerical values of the following quantities

ⓐ $P(X \leq 1)$
ⓑ $P(X \leq 10)$
ⓒ $P(X \geq 10)$
ⓓ $P(X = 10)$

Example

16 Let X have the CDF



Find the numerical values of the following quantities

ⓐ $P(X \leq 1)$
ⓑ $P(X \leq 10)$
ⓒ $P(X \geq 10)$
ⓓ $P(X = 10)$
ⓔ $P(|X - 5| \leq 0.1)$.

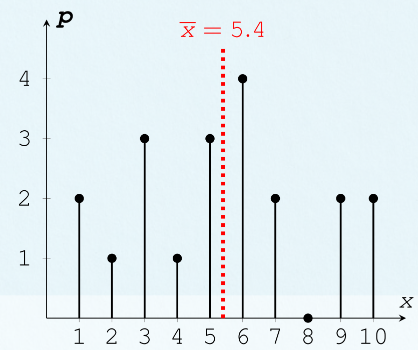
Example 17 – Introduction to expectation

Random numbers

2	9	4	3	3	10	5	7	6	6
10	6	5	6	5	7	1	3	9	1

The average value

$$\overline{x} = \frac{1}{20}(2+9+4+3+3+10+5+7+6+6 + 10+6+5+6+5+7+1+3+9+1)$$
$$= \frac{1}{20}(1\cdot 2 + 2\cdot 1 + 3\cdot 3 + 4\cdot 1 + 5\cdot 3 + 6\cdot 4 + 7\cdot 2 + 8\cdot 0 + 9\cdot 2 + 10\cdot 2)$$
$$= 1\cdot\frac{2}{20} + 2\cdot\frac{1}{20} + 3\cdot\frac{3}{20} + 4\cdot\frac{1}{20} + 5\cdot\frac{3}{20} + 6\cdot\frac{4}{20} + 7\cdot\frac{2}{20} + 8\cdot\frac{0}{20} + 9\cdot\frac{2}{20} + 10\cdot\frac{2}{20}$$
$$= 1\cdot f(1) + 2\cdot f(2) + 3\cdot f(3) + \cdots + 10\cdot f(10)$$
$$= 5.4.$$



Example 18 – Introduction to expectation

Play a game

u	-1	1
$P(X=u)$	$\frac{2}{3}$	$\frac{1}{3}$

Then

$$\begin{aligned} \mathrm{E}(X) &= \frac{(-1)\cdot 2 + (1)\cdot 1}{3} \\ &= (-1)\cdot\frac{2}{3} + (1)\cdot\frac{1}{3} \quad (= -\frac{1}{3}) \\ &= (-1)f(-1) + (1)f(1). \end{aligned}$$

Expectation

Definition

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

$$\begin{aligned} E(X) &= \sum_u u f(u) \quad (discrete) \\ E(X) &= \int_{-\infty}^{\infty} u f(u) du \quad (continuous). \end{aligned}$$

Other names: Expected value, Mean, Mean value, Average value.

Proposition (Properties)

Expectation is linear:

$$\mathrm{E}(aX+b) = a\,\mathrm{E}(X) + b$$

and

$$\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$$

Example

19 A roulette wheel has the numbers 1 through 36, as well as 0 and 00. If you bet \$1 that an odd number comes up, you win or lose \$1 according to whether that event occurs. If X denotes your net gain, find $\mathrm{E}(X)$.

Example

20 A school class of 120 students is driven in 3 buses to a symphonic performance. There are 36 students in one of the buses, 40 in another, and 44 in the third bus. When the buses arrive, one of the 120 students is randomly chosen. Let X denote the number of students on the bus of that randomly chosen student, and find $\mathrm{E}(X)$.

Example

20 A school class of 120 students is driven in 3 buses to a symphonic performance. There are 36 students in one of the buses, 40 in another, and 44 in the third bus. When the buses arrive, one of the 120 students is randomly chosen. Let X denote the number of students on the bus of that randomly chosen student, and find $\mathsf{E}(X)$.

21 Find $\mathsf{E}[X]$ when the density function of X is

$$f(u) = \begin{cases} 2u, & 0 \le u \le 1 \\ 0, & \text{otherwise} \end{cases}$$

The second moment of random variables

The expected value of a random variable X^2 is

$$\mathsf{E}(X^2) = \sum_u u^2 f(u) \quad \text{(discrete)}$$

$$\mathsf{E}(X^2) = \int_{-\infty}^{\infty} u^2 f(u) du \quad \text{(continuous)}.$$

Example

22 A roulette wheel has the numbers 1 through 36, as well as 0 and 00. If you bet \$1 that an odd number comes up, you win or lose \$1 according to whether that event occurs. If X denotes your net gain, find $\mathsf{E}(X^2)$.

Example

22 A roulette wheel has the numbers 1 through 36, as well as 0 and 00. If you bet \$1 that an odd number comes up, you win or lose \$1 according to whether that event occurs. If X denotes your net gain, find $\mathsf{E}(X^2)$.

23 Let X denote a random variable that takes on any of the values -1, 0, and 1 with respective probabilities

$$P(X = -1) = 0.2, \quad P(X = 0) = 0.5, \quad P(X = 1) = 0.3$$

Compute $\mathsf{E}(X^2)$.

Example

22 A roulette wheel has the numbers 1 through 36, as well as 0 and 00. If you bet \$1 that an odd number comes up, you win or lose \$1 according to whether that event occurs. If X denotes your net gain, find $E(X^2)$.

23 Let X denote a random variable that takes on any of the values −1, 0, and 1 with respective probabilities

$$P(X=-1)=0.2, \quad P(X=0)=0.5, \quad P(X=1)=0.3$$

Compute $E(X^2)$.

24 Let X be the current measured in mA. The PDF is $f(x)=0.05$ for $0 \le x \le 20$. What is the expected value of power when the resistance is 100 ohms?

Expectation of function of a random variable

Proposition

In general, for any function $g(u)$:

$$E\, g(X) = \sum_u g(u)f(u) \quad \text{(discrete)}$$

$$E\, g(X) = \int_{-\infty}^{\infty} g(u)f(u)du \quad \text{(continuous)}.$$

Example 25 - Digital Channel

There is a chance that a bit transmitted through a digital transmission channel is received in error. X is the number of bits received in error of the next 4 transmitted. The probabilities are

$$P(X=0)=0.6561, \quad P(X=2)=0.0486, \quad P(X=4)=0.0001,$$
$$P(X=1)=0.2916, \quad P(X=3)=0.0036,$$

What is the expected value of the cube of the number of bits in error?

Solution

Put $g(u)=u^3$.

$$E(X^3)=E(g(X))=\sum_{u=0}^{4} g(u)f(u)=\sum_{u=0}^{4} u^3 f(u)$$
$$=0^3(0.6561)+1^3(0.2916)+2^3(0.0486)+3^3(0.036)+4^3(0.0001)$$
$$=1.6588.$$

Example 26 - Expected cost

The time, in hours, it takes to locate and repair an electrical breakdown in a certain factory is a random variable X whose density function is given by

$$f(u)=\begin{cases} 1, & \text{if } 0<u<1 \\ 0, & \text{otherwise.} \end{cases}$$

If the cost involved in a breakdown of duration X is X^3, what is the expected cost of such a breakdown?

Solution

Put $h(u)=u^3$. *Then*

$$E(X^3)=E(h(X))=\int_0^1 u^3 f(u)du=\int_0^1 u^3 du=\frac{u^4}{4}\Big|_0^1=\frac{1}{4}.$$

Example

27 Suppose X is a random variable taking values in $\{-2,-1,0,1,2,3,4,5\}$, each with probability $1/8$. Let $Y = X^2$. Find $\mathsf{E}[Y]$.

Example

27 Suppose X is a random variable taking values in $\{-2,-1,0,1,2,3,4,5\}$, each with probability $1/8$. Let $Y = X^2$. Find $\mathsf{E}[Y]$.

28 Find $\mathsf{E}[e^X]$ when the density function of X is

$$f(u) = \begin{cases} 1, & 0 \le u \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Example

29 Linda is a sales associate at a large auto dealership. At her commission rate of 25% of gross profit on each vehicle she sells, Linda expects to earn \$350 for each car sold and \$400 for each truck or SUV sold. Linda motivates herself by using probability estimates of her sales. She estimates her car sales in one day as follows

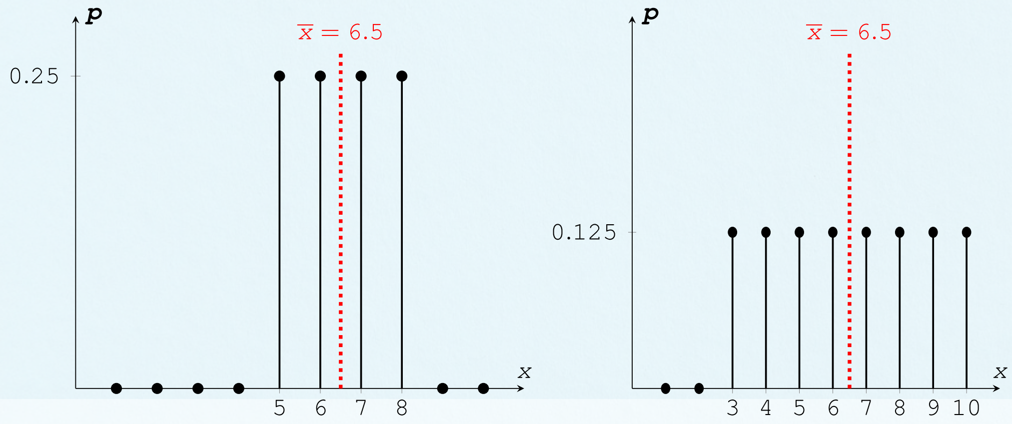
Car sales	0	1	2	3
Probability	0.3	0.4	0.2	a

and her truck or SUV sales as follows

Truck sales	0	1	2
Probability	0.3	0.5	b

What is the best estimate of Linda's earnings per day?

Example 30 - Pass or First-Class Honours?

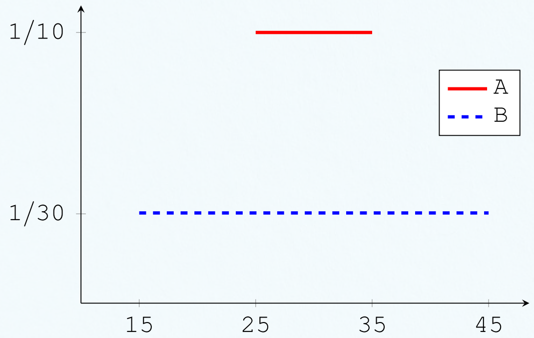


(a) Class 1 (b) Class 2

Example 31 – Stock Price Changes

Consider the prices A and B of two stocks at a time one month in the future. Assume that

- A has the uniform distribution on the interval $[25, 35]$



The distributions: the mean = 30 but different.

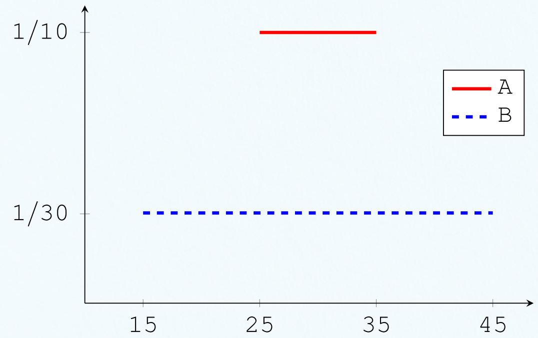
A will surely be worth at least 25 while $P(B < 25) = 1/3$.

Variance and Standard Deviation

Example 31 – Stock Price Changes

Consider the prices A and B of two stocks at a time one month in the future. Assume that

- A has the uniform distribution on the interval $[25, 35]$
- B has the uniform distribution on the interval $[15, 45]$.



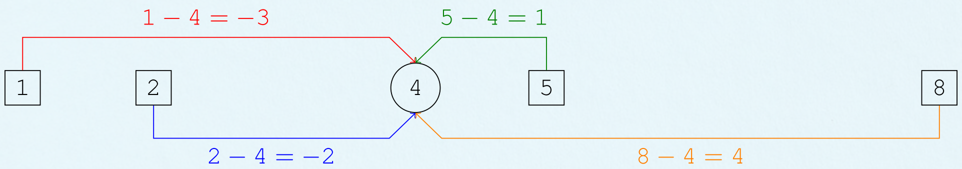
The distributions: the mean = 30 but different.

A will surely be worth at least 25 while $P(B < 25) = 1/3$.

Dung Nguyen Probability and Statistics 33 / 49

Deviations

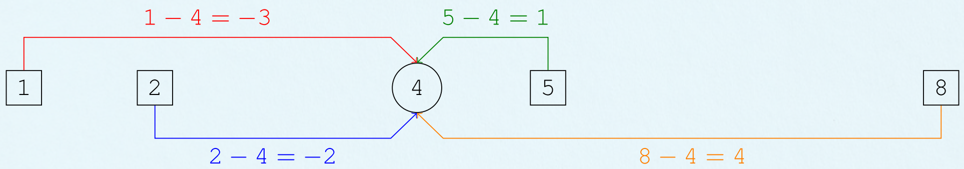
Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4.$



Variance and Standard Deviation

Deviations

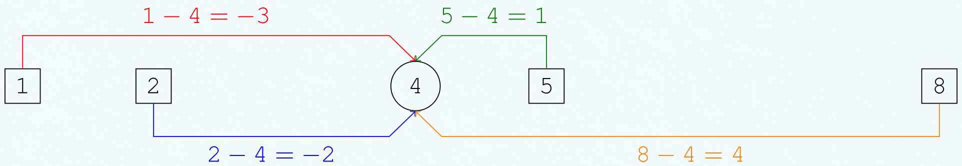
Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4.$



Dung Nguyen Probability and Statistics 34 / 49

Deviations

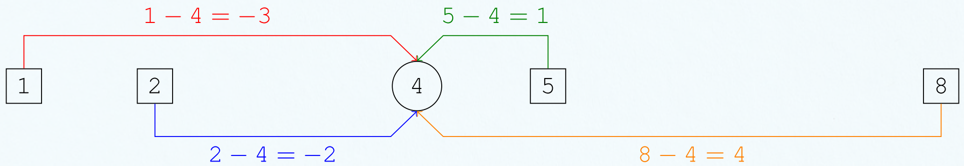
Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4.$



Variance and Standard Deviation

Deviations

Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4.$

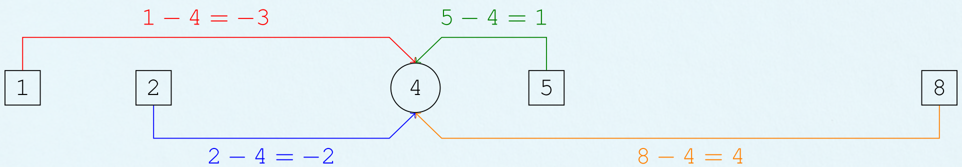


Dung Nguyen Probability and Statistics 34 / 49

Variance and Standard Deviation

Deviations

Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4.$

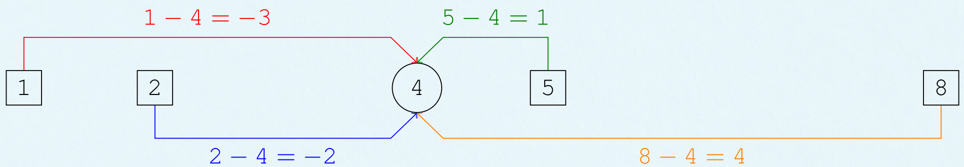


Dung Nguyen Probability and Statistics 34 / 49

Variance and Standard Deviation

Deviations

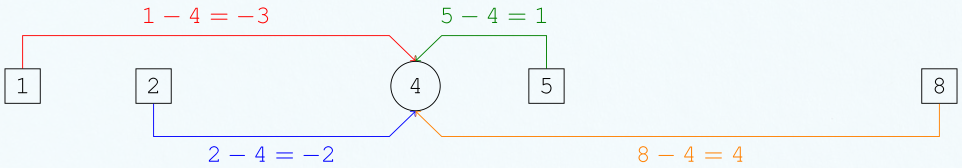
Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4.$



Dung Nguyen Probability and Statistics 34 / 49

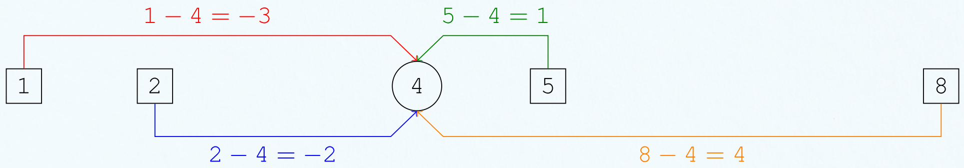
Deviations

Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4.$



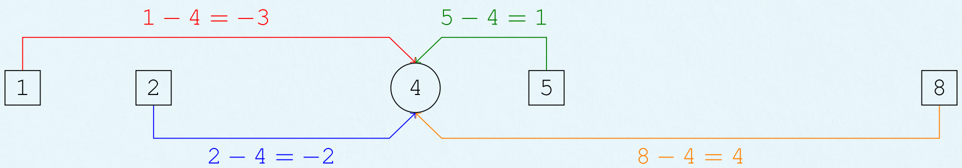
Deviations

Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4.$



Deviations

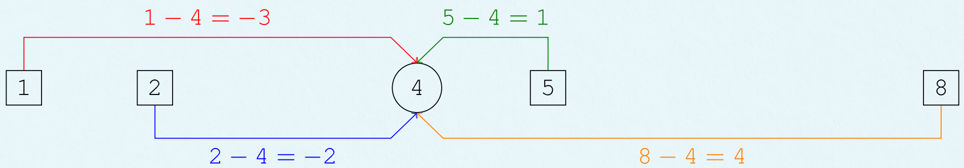
Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4.$



Sum of squares: $(1-4)^2 + (2-4)^2 + (5-4)^2 + (8-4)^2.$

Deviations

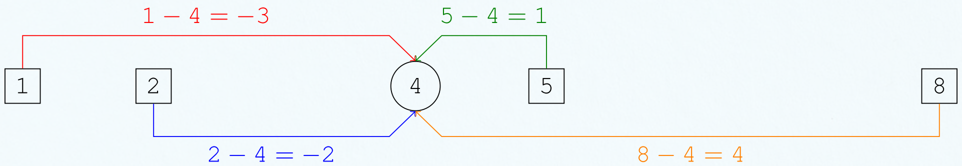
Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4.$



Average sum of squares: $\dfrac{(1-4)^2 + (2-4)^2 + (5-4)^2 + (8-4)^2}{4}.$

Deviations

Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4$.

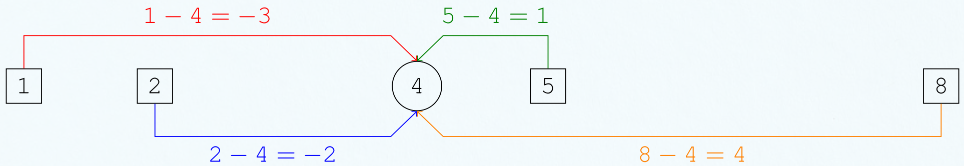


Sample variance: $\dfrac{(1-4)^2+(2-4)^2+(5-4)^2+(8-4)^2}{4-1}$.

Variance and Standard Deviation

Deviations

Observations: 1, 2, 5, 8 $\implies M = \dfrac{1+2+5+8}{4} = 4$.



Observations: $x_1, x_2, \dots, x_n \implies \overline{x} = \dfrac{x_1+x_2+\cdots+x_n}{n}$.

$$s^2 = \frac{(x_1-\overline{x})^2+(x_2-\overline{x})^2+\cdots+(x_n-\overline{x})^2}{n-1}.$$

Variance and Standard Deviation

- The variance of an r.v. X is
$$\mathsf{V}(X) = \mathsf{E}(X-\mu)^2.$$
Variance measures dispersion around the mean.

Variance and Standard Deviation

Variance and Standard Deviation

- The variance of an r.v. X is
$$\mathsf{V}(X) = \mathsf{E}(X-\mu)^2.$$
Variance measures dispersion around the mean.
    - X is discrete
$$\mathsf{V}(X) = \sum_u (u-\mu)^2 f(u).$$

Variance and Standard Deviation

- The variance of an r.v. X is
$$V(X) = E(X-\mu)^2.$$
Variance measures dispersion around the mean.
  - X is discrete
$$V(X) = \sum_u (u-\mu)^2 f(u).$$
  - X is continuous
$$V(X) = \int_{-\infty}^{\infty} (u-\mu)^2 f(u)du.$$

Variance and Standard Deviation

Variance and Standard Deviation

- The variance of an r.v. X is
$$V(X) = E(X-\mu)^2.$$
Variance measures dispersion around the mean.
  - X is discrete
$$V(X) = \sum_u (u-\mu)^2 f(u).$$
  - X is continuous
$$V(X) = \int_{-\infty}^{\infty} (u-\mu)^2 f(u)du.$$
- The standard deviation is
$$SD(X) = \sqrt{Var(X)}.$$

Properties

1. Computational formula
$$V(X) = E(X^2) - (E\,X)^2$$

Variance and Standard Deviation

Properties

1. Computational formula
$$V(X) = E(X^2) - (E\,X)^2$$
2. Variance and standard deviation are not linear
$$V(aX+b) = a^2\,V(X) \quad \text{and} \quad SD(aX+b) = a\,SD(X)$$

Properties

1. Computational formula
$$V(X) = E(X^2) - (E\,X)^2$$
2. Variance and standard deviation are not linear
$$V(aX+b) = a^2\,V(X) \quad \text{and} \quad SD(aX+b) = a\,SD(X)$$
3. If X and X are independent then
$$V(X+Y) = V(X) + V(Y)$$

Example 32 – Digital Channel

There is a chance that a bit transmitted through a digital transmission channel is received in error. X is the number of bits received in error of the next 4 transmitted. The probabilities are
$$P(X=0)=0.6561,\quad P(X=2)=0.0486,\quad P(X=4)=0.0001,$$
$$P(X=1)=0.2916,\quad P(X=3)=0.0036$$
Calculate the mean and variance.

x	$f(x)$	$xf(x)$	$(x-0.4)^2$	$(x-0.4)^2 f(x)$	$x^2 f(x)$
0	0.6561	0.0000	0.160	0.1050	0.0000
1	0.2916	0.2916	0.360	0.1050	0.2916
2	0.0486	0.0972	2.560	0.1244	0.1944
3	0.0036	0.0108	6.760	0.0243	0.0324
4	0.0001	0.0004	12.960	0.0013	0.0016
	Total	0.4000		0.3600	0.5200

Example 33 – Electric Current

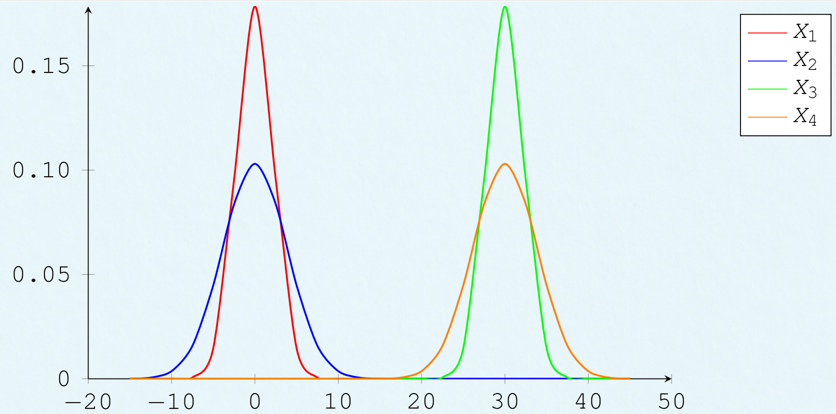
For the copper wire current measurement, the PDF is $f(u) = 0.05$ for $0 \le u \le 20$. Find the mean and variance.

Solution

$$E(X) = \int_{-\infty}^{\infty} u f(u) du = \int_0^{20} u \times (0.05) du = \frac{0.05u^2}{2}\Big|_0^{20} = 10$$
$$V(X) = \int_{-\infty}^{\infty} (u-10)^2 f(u) du = \int_0^{20} (u-10)^2 (0.05) du$$
$$= \frac{0.05(u-10)^3}{3}\Big|_0^{20} = \frac{100}{3}.$$

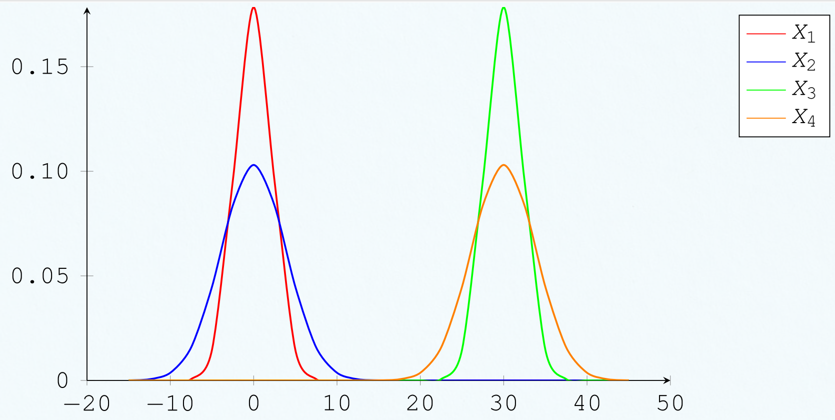
$$E(X^2) = \int_{-\infty}^{\infty} u^2 f(u) du = \int_0^{20} u^2 \times (0.05) du = \frac{0.05u^3}{3}\Big|_0^{20} = \frac{400}{3}$$
$$V(X) = E(X^2) - E(X)^2 = \frac{400}{3} - 10^2 = \frac{100}{3}.$$

Example 34 – Comparison



- $E(X_1) = E(X_2) < E(X_3) = E(X_4)$.

Example 34 – Comparison



- $\mathsf{E}(X_1) = \mathsf{E}(X_2) < \mathsf{E}(X_3) = \mathsf{E}(X_4)$.
- $\mathsf{V}(X_1) = \mathsf{E}(X_3) < \mathsf{E}(X_2) = \mathsf{E}(X_4)$.

Example

35 Suppose that X can take each of the five values $-2, 0, 1, 3, 4$ with equal probability. Determine the variance and standard deviation of X and $Y = 4X - 7$.

Example

35 Suppose that X can take each of the five values $-2, 0, 1, 3, 4$ with equal probability. Determine the variance and standard deviation of X and $Y = 4X - 7$.

36 Suppose X has the following pdf, where c is a constant to be determined

$$f(u) = \begin{cases} c(1-u^2), & -1 \le u \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Compute $\mathsf{E}(X), \mathsf{V}(X)$.

In many random experiments, more than one quantity is measured, meaning that there is more than one random variable.

Example (Cell phone flash unit)

A flash unit is chosen randomly from a production line; its recharge time X (seconds) and flash intensity Y (watt-seconds) are measured.

To make probability statements about several random variables, we need their joint probability distribution.

Joint Probability Mass Function

The joint probability mass function of the discrete random variables X and Y denoted as $f_{XY}(u,v)$ satisfies

$$f_{XY}(u,v) = P(X = u, Y = v).$$

Proposition (Characteristic properties)

1. $f_{XY}(u,v) \geq 0$ for all u, v.
2. $\sum_u \sum_v f_{XY}(u,v) = 1$

Example 37 – Signal Strength

A mobile web site is accessed from a smart phone; X is the signal strength, in number of bars, and Y is response time, to the nearest second.

y = Response time (nearest second)	x = Number of Bars of Signal Strength			
	1	2	3	**Total**
1	0.01	0.02	0.25	0.28
2	0.02	0.03	0.20	0.25
3	0.02	0.10	0.05	0.17
4	0.15	0.10	0.05	0.30
Total	0.20	0.25	0.55	1.00

Determine

a. $P(X < 3, Y \leq 2)$.
b. $P(X < 3 | Y \leq 2)$.
c. $P(Y \leq 2 | X < 3)$.

Joint Probability Density Function

The joint probability density function for the continuous random variables X and Y, denotes as $f_{XY}(u,v)$, satisfies the following properties

$$P((X,Y) \in A) = \iint_A f_{XY}(u,v)dudv.$$

Proposition (Characteristic properties)

1. $f_{XY}(u,v) \geq 0$.
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(u,v)dudv = 1$.

Example 38 – Server Access Time

Let the random variable X denote the time until a computer server connects to your machine (in milliseconds), and let Y denote the time until the server authorizes you as a valid user (in milliseconds). X and Y measure the wait from a common starting point ($u < v$). The joint probability density function for X and Y is

$$f_{XY}(u,v) = k\,\mathrm{e}^{-0.001u - 0.002v},$$

for $0 < u < v < \infty$.

a. Identify k.
b. Calculate $P(X \leq 1000, Y \leq 2000)$.

Solution

$k = 6 \times 10^{-6}$, $P(X \leq 1000, Y \leq 2000) = 0.915$

Marginal Probability Distributions (discrete)

Since X is a random variable, it also has its own probability distribution, ignoring the value of Y, called its marginal probability distribution.

The marginal probability distribution for X

$$\begin{aligned} f_X(u) &= P(X=u) \\ &= \sum_v P(X=u, Y=v) \\ &= \sum_v f_{XY}(u,v) \\ f_X(u) &= \sum_v f_{XY}(u,v). \end{aligned}$$

The marginal probability distribution for Y

$$f_Y(v) = \sum_u f_{XY}(u,v).$$

Example 39 – Signal Strength

A mobile web site is accessed from a smart phone; X is the signal strength, in number of bars, and Y is response time, to the nearest second.

$y =$ Response time (nearest second)	$x =$ Number of Bars of Signal Strength			
	1	2	3	**Marginal** $f_Y(y)$
1	0.01	0.02	0.25	0.28
2	0.02	0.03	0.20	0.25
3	0.02	0.10	0.05	0.17
4	0.15	0.10	0.05	0.30
Marginal $f_X(x)$	0.20	0.25	0.55	1.00

Marginal Probability Distributions (continuous)

If the joint probability density function of random variables X and Y is $f_{XY}(u,v)$, then

The marginal probability density functions of X:

$$f_X(u) = \int_{-\infty}^{\infty} f_{XY}(u,v)dv,$$

The marginal probability density functions of Y:

$$f_Y(v) = \int_{-\infty}^{\infty} f_{XY}(u,v)du.$$

Example 40 – Signal Strength

A mobile web site is accessed from a smart phone; X is the signal strength, in number of bars, and Y is response time, to the nearest second.

$y =$ Response time (nearest second)	$x =$ Number of Bars of Signal Strength			
	1	2	3	**Marginal** $f_Y(y)$
1	0.01	0.02	0.25	0.28
2	0.02	0.03	0.20	0.25
3	0.02	0.10	0.05	0.17
4	0.15	0.10	0.05	0.30
Marginal $f_X(x)$	0.20	0.25	0.55	1.00

Compute the mean and the variance of X and Y.