Probability III

REA1121

Mathematics for programming

Outline

- Random variables
  - Discrete random variables
  - Probability distribution
- Measures of location and dispersion
  - Expectation
  - Variance, standard deviation
- Exercises

RANDOM VARIABLES

Example 1

- 3 balls 

- Take 1 out of 3, and put it back -> label i
- Take 1 out of 3 again -> label j
- What lies in the sample space?

Solution to Example 1

Sample		i		
		1	2	3
J	1	1,1	2,1	3,1
	2	1,2	2,2	3,2
	3	1,3	2,3	3,3

Solution to Example 1

X=i+j		i		
		1	2	3
j	1	X=2	X=3	X=4
	2	X=3	X=4	X=5
	3	X=4	X=5	X=6

Solution to Example 1

- Probability

  $P(X=2) = 1/9$

  $P(X=3) = 2/9$

  $P(X=4) = 3/9$

  $P(X=5) = 2/9$

  $P(X=6) = 1/9$

Example 2

- Toss 1 coin 3 times
- Order is not important
- What lies in the sample space?

Solution to Example 2

Sample	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
X	3	2	2	2	1	1	1	0

X: the occurrence of Head

Solution to Example 2

- Probability

  $P(X=0) = 1/8$

  $P(X=1) = 3/8$

  $P(X=2) = 3/8$

  $P(X=3) = 1/8$

Random variables

- S: sample space

  $\{e_1, e_2, \cdots, e_m\}$

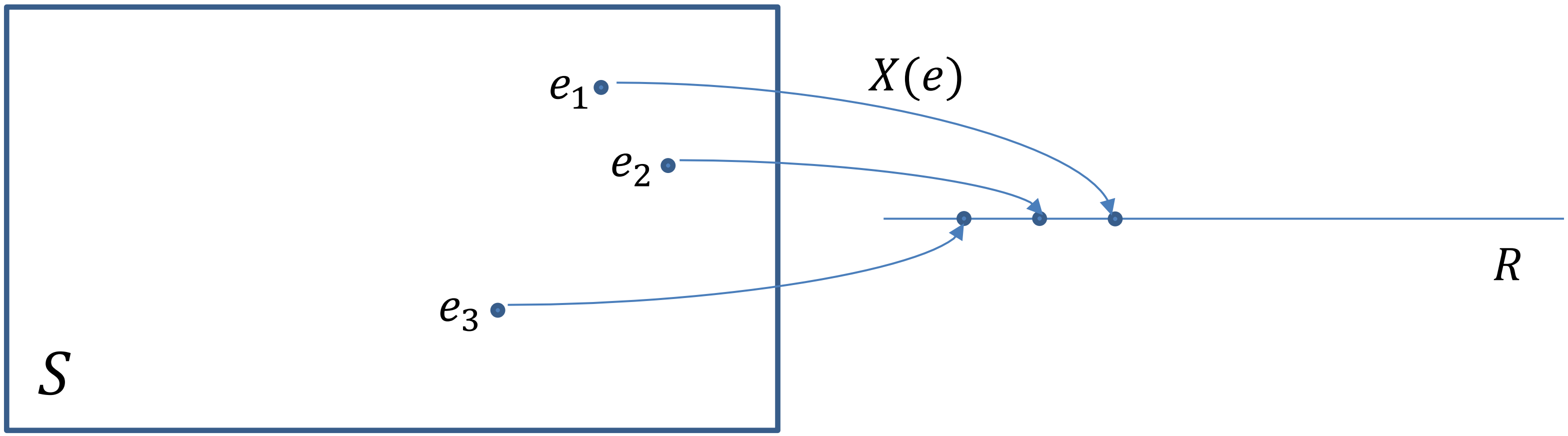
- A list of possible numerical values

  $\{x_1, x_2, \cdots, x_n\}$

- Probability

  $P(X = x_1), P(X = x_2), \cdots, P(X = x_n)$

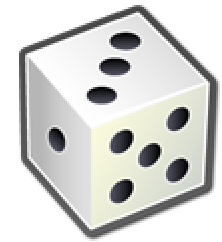
Random variables



Random variables

- A function: X
- Values are unpredictable
- Values are within a range of probability
- Distinct from typical functions

Random variables



- Examples?
  - Toss of a dice
  - Next month’s rainfall
  - A flight delay
  - The time of tomorrow’s sunrise
  - Number of attendees in the class tomorrow
  - Electricity consumed on campus per day

Random variables

- Discrete random variables
  - Number of calls a taxi call centre receives per day
  - Card draw
  - Dice roll
- Continuous random variables
  - A flight delay
  - Electricity consumed per day
  - Lifespan of a computer

Discrete random variables

- X*: a discrete random variable*
- $x_k, \mathrm{k} = 1,2,\cdots$
- $P(X = x_k) = p_k, k = 1,2,\cdots$
- p_k

  $p_k \geq 0, k = 1,2,\cdots;$

  $\sum_{k=1}^{\infty} p_k = 1.$

Probability distribution

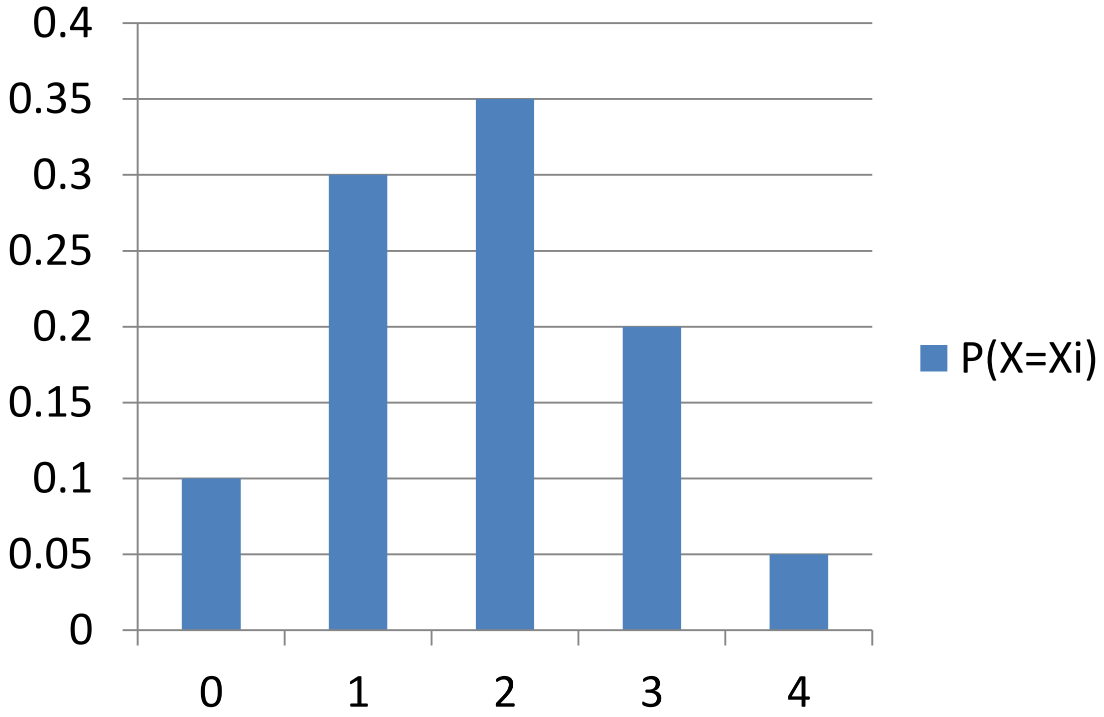
- $P(X = x_k) = p_k, k = 1,2,\cdots$
- Table
- Graph

Example 3

- The number of ships arriving at a container terminal during any one day, X, can be any integer from 0 to 4, with respective probabilities 0.1, 0.3, 0.35, 0.2, 0.05
- Describe the probability distribution of X.

Solution to Example 3

X	0	1	2	3	4
P	0.1	0.3	0.35	0.2	0.05



Example 4

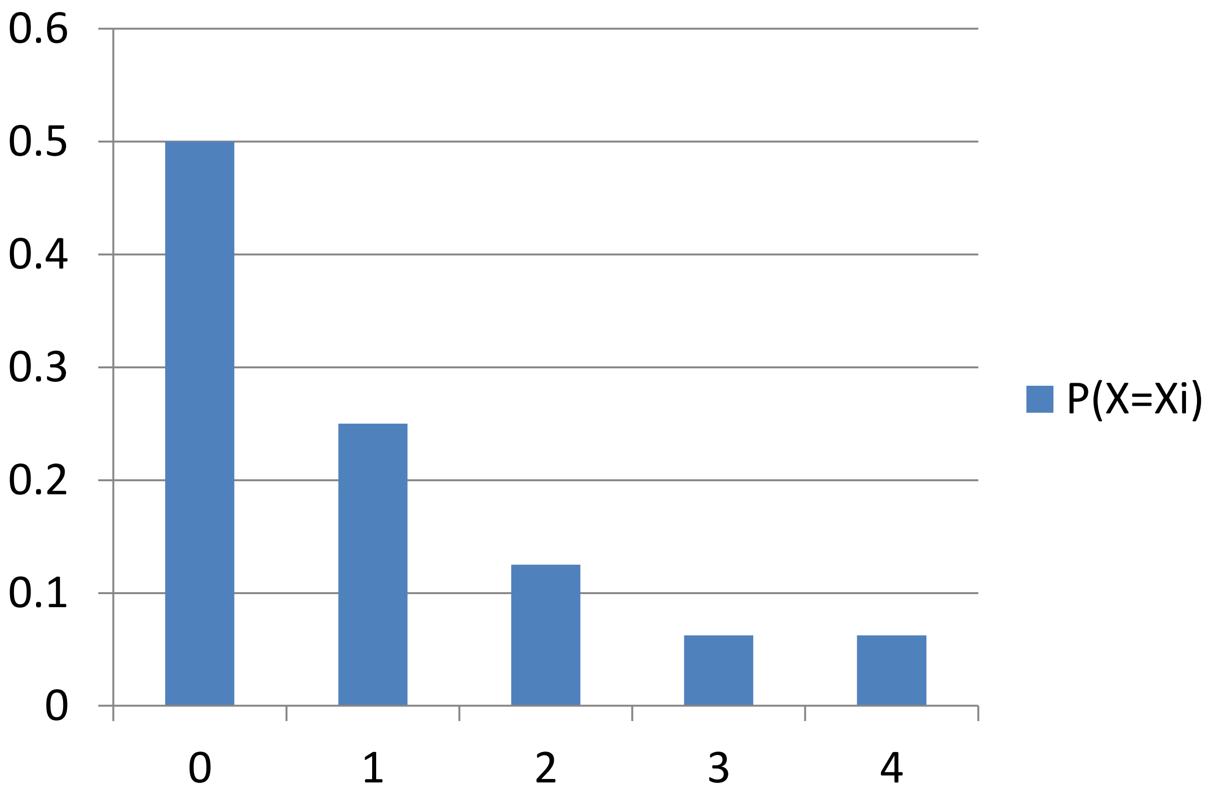
- A car running towards the destination has to pass 4 sets of traffic signal lamps
- Each set has ½ probability of "Green", and ½ of "Red".
- X: number of set of lamps the car has passed when it first stops.
- Describe the probability distribution of X.

Solution to Example 4

- Let p be the probability of "Red" Light of each set
- Then the probability of "Green" Light is (1-p) of each set

X	0	1	2	3	4
P	p	$(1-p)p$	$(1-p)^2p$	$(1-p)^3p$	$(1-p)^4$
	0.5	0.25	0.125	0.0625	0.0625

Solution to Example 4



Important probability distribution

- (0-1) distribution
- Bernoulli distribution
- Binominal distribution
- Poisson distribution

(0-1) distribution

- X: 0 or 1
- $P(X = k) = p^k(1 - p)^{1-k}, k = 0,1\ (0 < p < 1)$

X	0	1
P	1-p	p

- Examples?

Bernoulli distribution

- Generalisation of (0-1) distribution
- Bernoulli experiment
  - E*: Experiment*
  - Outcomes: $A, \bar{A}$
  - $P(A) = p,\ 0 < p < 1$
  - $P(\bar{A}) = 1 - p$

Binominal distribution

- Repeat Bernoulli experiment independently n times
  - Repeat: p remains unchanged
  - Independently: outcomes do not influence each other
- X: number of events where A occurs
- X: $\{0,1,2,\ldots,n\}$
- Suppose $\mathrm{X} = \mathrm{k}, 0 \ll k \ll n$

Binominal distribution

- $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0,1,2,\cdots,n$
- If a discrete random variable X follows binominal distribution, we write $X \sim b(n,p)$

Example 5

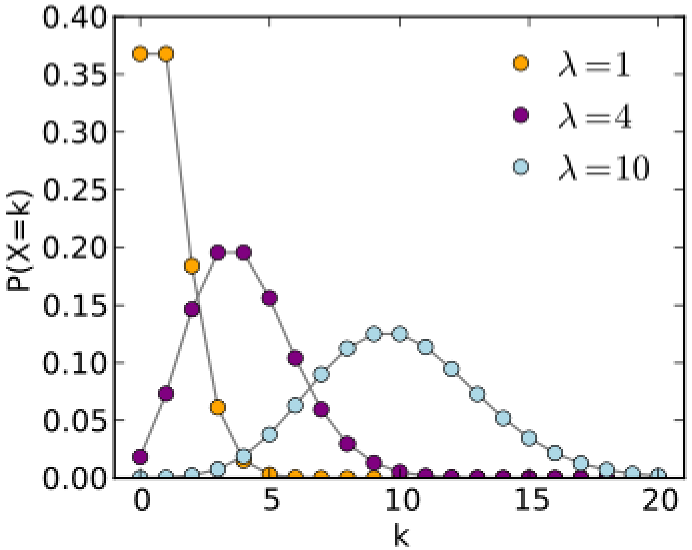
- In a shooting game
- Probability of hitting the target is 0.02 in each shooting
- Shoot 400 times
- What is the probability of hitting the target at least 2 times?

Solution to Example 5

- $P(X \geq 2) = 1 - P(X = 0) - P(X = 1)$
$= 1 - 0.98^{400} - 400 \cdot 0.02^{1} \cdot 0.98^{399}$
$= 0.9972$
- Discussions
  - $P(X \geq 2)$ is fairly large despite p is small
  - As $P(X < 2) \approx 0.003$, one may question the fact that $p = 0.02$

Poisson distribution

- $X: \{0,1,2,\cdots,\}$
- $P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}, k=0,1,2,\cdots$
- Important in many fields
  - Photon shot noise

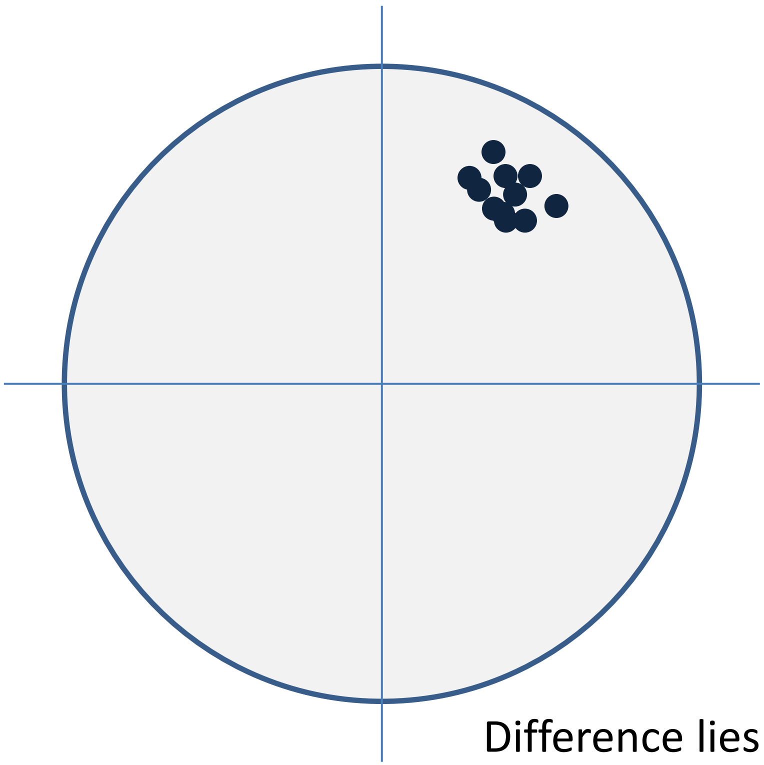


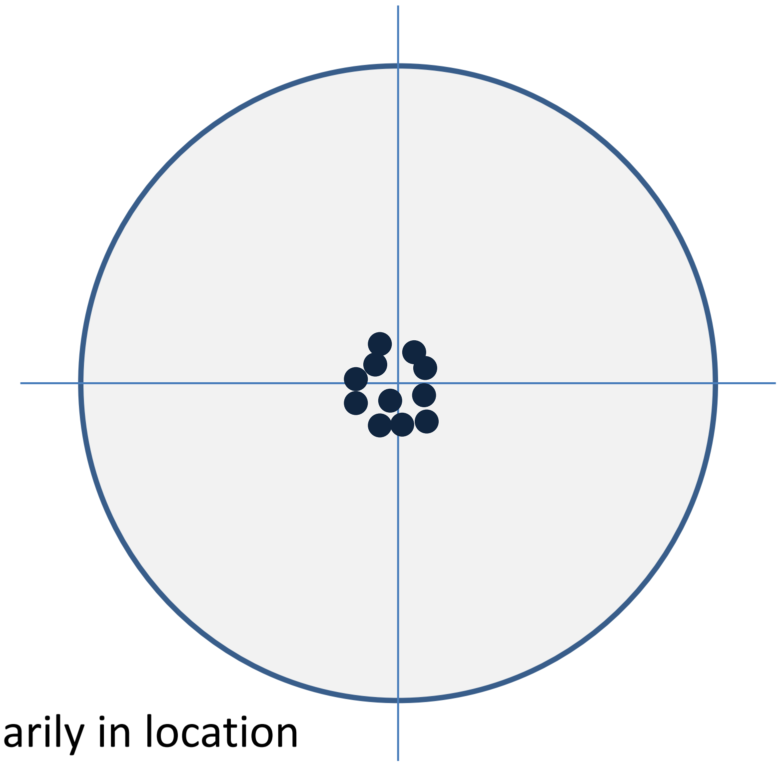
MEASURES OF LOCATION AND DISPERSION

Measures of random variable

- How do we characterise discrete random variables in an easier manner?
- Probability distribution
- Any other technique?

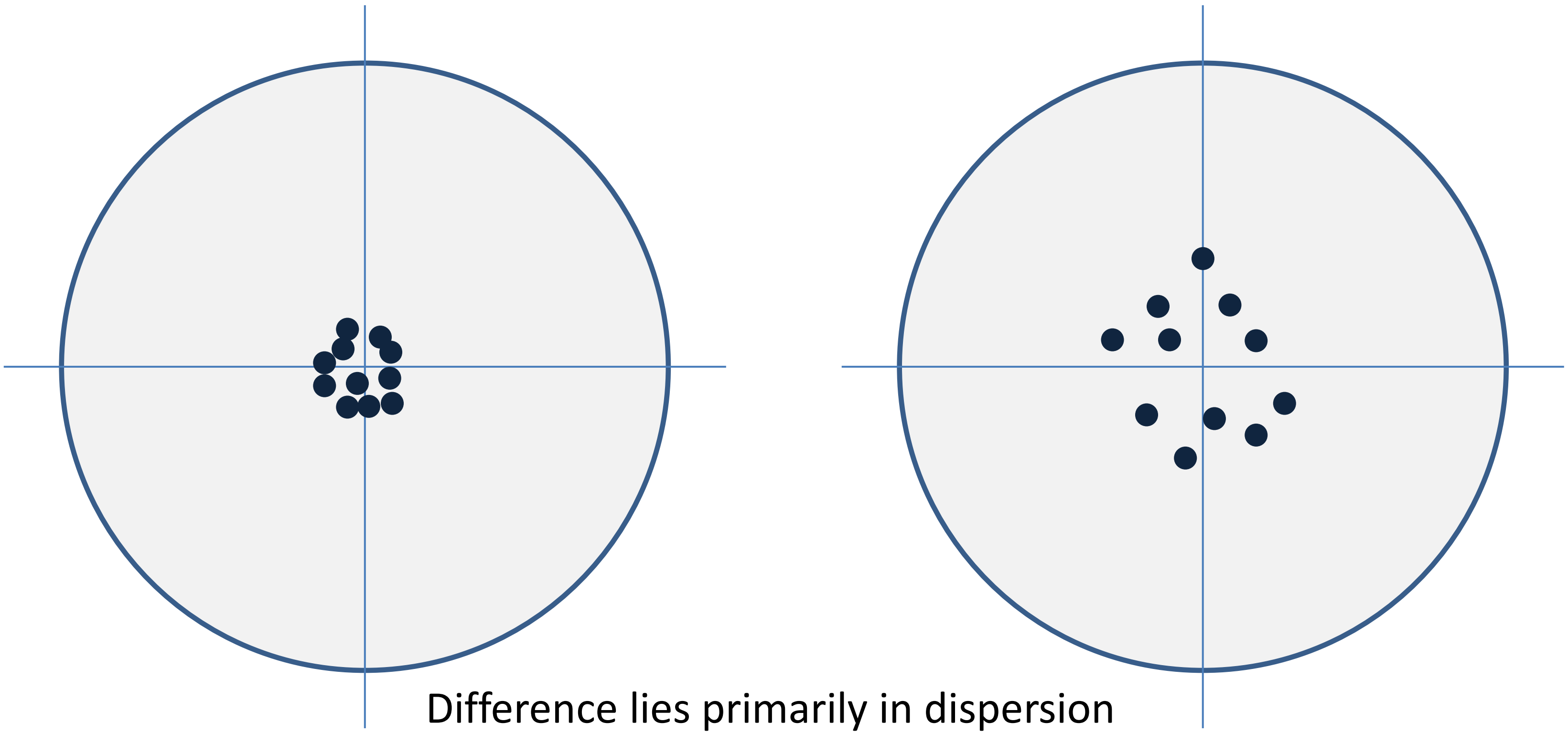
Measures of random variables





Difference lies primarily in location

Measures of random variables



Difference lies primarily in dispersion

Measures of random variables

- Mathematical expectation
  - Mean (Average)
  - Measure of location
  - $E(X) = \mu_X = \sum_{k=1}^{\infty} x_k p_k$
- Variance
  - Measure of dispersion
  - $D(X) = Var(X) = {\sigma_X}^2 = E\{[X - E(X)]^2\} = E(X^2) - [E(X)]^2$
- Standard deviation
  - $\sigma_X = \sigma(X) = \sqrt{D(X)} = \sqrt{{\sigma_X}^2}$

Example 6

- Shooting game
- 2 persons A & B
- 10 shootings
- The performance
  - A: 3 x Ring 10, 5 x Ring 8, 1 x Ring 5, 1 x Ring 4
  - B: 6 x Ring 8, 2 x Ring 7, 2 x Ring 6
- How do you compare the two?

Solution to Example 6

- For person A:
- $E(X) = \sum_{k=1}^{\infty} x_k p_k = 10 \cdot \frac{3}{10} + 8 \cdot \frac{5}{10} + 5 \cdot \frac{1}{10} + 4 \cdot \frac{1}{10} = 7.9$
- ${\sigma_X}^2 = E(X^2) - [E(X)]^2 = 10^2 \cdot \frac{3}{10} + 8^2 \cdot \frac{5}{10} + 5^2 \cdot \frac{1}{10} + 4^2 \cdot \frac{1}{10} - 7.9^2 = 3.69$
- $\sigma_X = \sqrt{{\sigma_X}^2} = \sqrt{3.69} = 1.9209$

Solution to Example 6

- For person B:
- $E(X) = \sum_{k=1}^{\infty} x_k p_k = 8 \cdot \frac{6}{10} + 7 \cdot \frac{2}{10} + 6 \cdot \frac{2}{10} = 7.4$
- ${\sigma_X}^2 = E(X^2) - [E(X)]^2 = 8^2 \cdot \frac{6}{10} + 7^2 \cdot \frac{2}{10} + 6^2 \cdot \frac{2}{10} - 7.4^2 = 0.64$
- $\sigma_X = \sqrt{{\sigma_X}^2} = \sqrt{0.64} = 0.8$

Properties of expectation

- $E(C) = C$, if C is a constant
- $E(CX) = CE(X)$, if X is a random variable, C is a constant
- $E(X + Y) = E(X) + E(Y)$, if X and Y are two random variables
- $E(XY) = E(X)E(Y)$, if X and Y are two *independent* random variables

Properties of variance

- $D(C) = 0$, if C is a constant
- $D(CX) = C^2 D(X)$, if X is a random variable, C is a constant
- $D(X+Y) = D(X) + D(Y) + 2E\{(X - E(X))(Y - E(Y))\}$, if X and Y are two random variables
- $D(X+Y) = D(X) + D(Y)$, if X and Y are two *independent* random variables
- $D(X) = 0$, if and only if $X = C$ or $P(X = C) = 1$ where C is a constant

EXERCISES

Exercise 1

- Random variable X: Age of students attending the class on the basis of the information below:

  40 students, 20 of them are 18 years old, 16 are 19 years old, 4 are 20 yeas old.

- Evaluate and present probability distribution by means of table and/or graph.
- Determine $E(X), {\sigma_X}^2, \sigma_X$

Exercise 2

- Throw a dice
  - You get 1 krone, if you get odd numbers
  - You pay 2 kroner, if you get even numbers
- What do you expect?
  - Will you gain or lose money?

Exercise 3

- A coach bus carrying 20 passengers
- There are 10 stations where the passengers may get off
- The bus stops at a station when there is a call from passengers, and otherwise does not stop
- X: number of stops in total
- Evaluate $E(X)$.
- Suppose
  - every passenger has equal probability of getting off at every station
  - passengers are independent from eath other.

Exercise 4

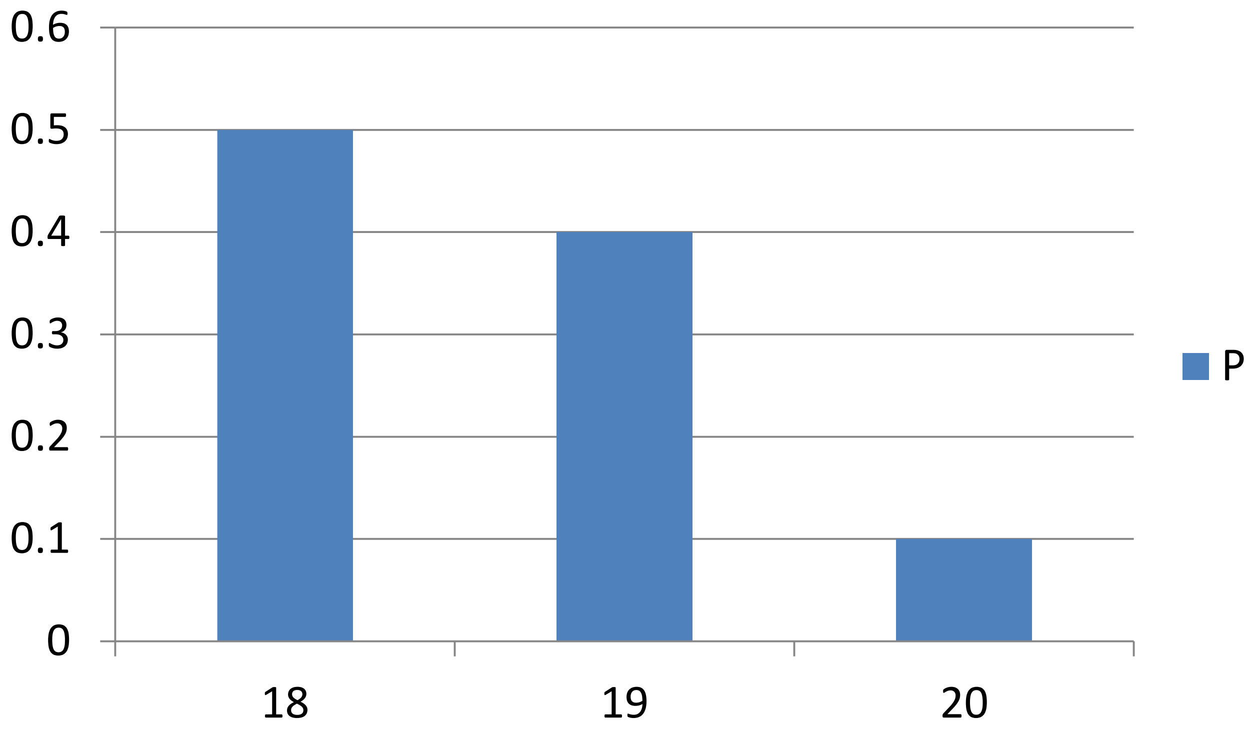
- Random variable X follows (0-1) distribution
  - $P(X = 0) = 1 - p$
  - $P(X = 1) = p$
- Determine ${\sigma_X}^2$

Exercise 5

- Random variable X follows Binominal distribution, $X \sim b(n, p)$
- Determine $E(X), {\sigma_X}^2$

Solution to Exercise 1

X	18	19	20
P	20/40	16/40	4/40



Solution to Exercise 1 (cont.)

- $E(X) = \sum_{k=1}^{\infty} x_k p_k = 18 \cdot \frac{20}{40} + 19 \cdot \frac{16}{40} + 20 \cdot \frac{4}{40} = 18.6$
- ${\sigma_X}^2 = E(X^2) - [E(X)]^2 = 18^2 \cdot \frac{20}{40} + 19^2 \cdot \frac{16}{40} + 20^2 \cdot \frac{4}{40} - 18.6^2 = 0.44$
- $\sigma_X = \sqrt{{\sigma_X}^2} = \sqrt{0.44} = 0.6633$

Solution to Exercise 2

Event	1	2	3	4	5	6
X	+1	-2	+1	-2	+1	-2
P	1/6	1/6	1/6	1/6	1/6	1/6

- $E(X) = \sum_{k=1}^{\infty} x_k p_k = 1 \cdot \frac{1}{6} + (-2) \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + (-2) \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + (-2) \cdot \frac{1}{6} = -0.5$

Solution to Exercise 3

- Let $X_i = \begin{cases} 0, nobody\ gets\ off\ at\ the\ ith\ station \\ 1, someone\ gets\ off\ at\ the\ ith\ station' \end{cases}, i = 1,2,\cdots,10$
- Then $X = X_1 + X_2 + \cdots + X_{10}$
- The probability that any passenger who does not get off at the ith station is $9/10$.
- Thus, the probability that all of 20 passengers do not get off at the ith station is $(9/10)^{20}$, and the probability that someone gets off at the ith station is $1 - (9/10)^{20}$.
- $P(X_i = 0) = (9/10)^{20}$ and $P(X_i = 1) = 1 - (9/10)^{20}$
- $E(X_i) = 0 \cdot [(9/10)^{20}] + 1 \cdot [1 - (9/10)^{20}] = 1 - (9/10)^{20}$
- $E(X) = E(X_1 + X_2 + \cdots + X_{10}) = \mathrm{E}(X_1) + E(X_2) + \cdots + E(X_{10}) = 10 \cdot \left[1 - \left(\frac{9}{10}\right)^{20}\right] = 8.7842$

Solution to Exercise 4

- $E(X) = 0 \cdot (1 - p) + 1 \cdot p = p$
- $E(X^2) = 0^2 \cdot (1 - p) + 1^2 \cdot p = p$
- ${\sigma_X}^2 = E(X^2) - [E(X)]^2 = p - p^2$

Solution to Exercise 5

- X follows binominal distribution, which means X, a random variable, refers to number of instances where event A occurs, and the probability that A occurs in each Bernoulli experiment is p.
- Let random variable

$$x_k = \begin{cases} 1, & A\ happens\ in\ the\ k^{th}\ experiment \\ 0, & A\ does\ not\ happen\ in\ the\ k^{th}\ experiment \end{cases}$$

  - ❖ $k = 1,2,\cdots,n$
- $X = X_1 + X_2 + \cdots + X_n$
- As x_k depends on the k^{th} experiment only, and each experiment is independent, $X_1, X_2, \cdots, X_n$ are independent
- x_k follows (0-1) distribution

Solution to Exercise 5 (cont.)

- As indicated by the solution to Exercise 4
  - $E(X_k) = p$
  - $D(X_k) = p(1-p), k = 1,2,\cdots,n$
- Thus
  $E(X) = E(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n} E(X_k) = np$
- As $X_1, X_2, \cdots, X_n$ are independent
  - ${\sigma_X}^2 = D(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n} D(X_k) = np(1-p)$