Probability theory and statistics

Academic year 2013/14, 2nd semester

1 Combinatorics

1.1 You are eating at Emile's restaurant and the waiter informs you that you have (a) two choices for appetizers: soup or juice; (b) three for the main course: a meat, fish, or vegetable dish; and (c) two for dessert: ice cream or cake. How many possible choices do you have for your complete meal?

1.2 Find the number of possible arrangements of 8 castles on the chess board in a way that they do not hit each other? What is the result if we can distinguish between the castles?

1.3 How many real four digit numbers (they can not start with zero) can be formed from digits $0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6$?

1.4 We have 12 books on the shelf. How many ways can the books be arranged on the shelf if 3 particular books must to be next to each other

a) if the order of the three books does not count?

b) if the order of the three books does count?

1.5 In how many ways can one arrange 4 math books, 3 chemistry books, 2 physics books, and 1 biology book on a bookshelf so that all the math books are together, all the chemistry books are together, and all the physics books are together?

1.6 In how many ways can 7 people be arranged around a round table?

1.7 In how many ways can 5 men and 5 women be arranged around a round table if neither two men, nor two women can sit next to each other?

1.8 In how many ways can one fill a toto coupon (14 matches, three possible results: $1,\ 2$ or X)?

1.9 Three postmen has to deliver six letters. Find the number of possible distribution of the letters.

1.10 Find the number of possible choices of four cards of four different colours from a deck of ordinary cards (4 colours, 13 cards per colour). What is the result if we require that the four cards should be off different figures?

1.11 Find the number of possible fillings of a lottery coupon (5 numbers from 90).

1.12 Find the number of possible paths from the origin to the point $(5, 3)$ if we can walk only on points with integer coordinates and we can step only upwards and right.

1.13 Starting from origin at each step we toss a coin and in case of a head we make a step to left, otherwise to right. In how many ways can we return to origin in 10 steps?

1.14 Prove the binomial theorem i.e. for all $a, b \in \mathbb{C}$ and $n \in \mathbb{N}$ we have

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

1.15 Prove that

$$\binom{n+1}{k+1} = \binom{n}{k+1} + \binom{n}{k}.$$

1.16 Prove that

$$\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{n} = 2^n.$$

1.17 Find the number of possible arrangements of n zeros of k ones $(k \leq n+1)$, if two ones can not be next to each other.

1.18 In Circus Maximus the tamer has to lead 5 lions and 4 tigers to the ring, but tigers can not follow each other because they fight. In how many ways can he do if the animals can be distinguished?

1.19 Around the round table of King Arthur 12 knights are sitting. Each of them hates his two neighbours. In how many ways can we choose five knights without having enemies among them?

1.20 A deck of ordinary cards is shuffled and 10 cards are dealt.

a) In how many cases will we have aces among the 10 cards?
b) Exactly one ace?
c) At most one ace?
d) Exactly two aces?
e) At least two aces?

1.21 In how many ways we can chose four dancing pairs from 12 girls and 15 boys?

1.22 Find the number of real six digit numbers having three odd and three even digits.

1.23 In how many ways we can distribute 14 persons into four boats with five, four, three and two seats, respectively?

1.24 In the canteen we can by four types of snacks. In how many ways can we buy 12 of them?

1.25 In how many ways we can distribute 7 apples and 9 peaches among 4 kids?

1.26 Peter is putting 10 identical balls in five different buckets. In how many ways can this be done if no bucket is allowed to be empty?

1.27 In how many ways we can choose five cards from a deck with having a club and an ace among them?

1.28 In how many ways we can choose four persons from five boys and five girls having at least two girls among them?

2 Events, operations on events

2.1 Prove the De-Morgan identities, i.e. that

$$\overline{A+B} = \overline{A} \cdot \overline{B} \qquad \text{and} \qquad \overline{A \cdot B} = \overline{A} + \overline{B}.$$

2.2 A coin is tossed. If the result is a head, it is tossed once again, otherwise it is tossed twice again. Give the sample space of the experiment.

2.3 Give the sample space of the five from ninety lottery.

2.4 A dice is thrown three times. Let A_i denote the event that the result of the ith throw is six, $i = 1, 2, 3$. What is the meaning of the following events:

$$A_1 + A_2 \quad A_1 \cdot A_2, \quad A_1 + A_2 + A_3, \quad A_1 \cdot A_2 \cdot A_3, \quad A_1 \cdot \overline{A_2}, \quad A_1 \setminus A_2 \ ?$$

2.5 In a workshop there are three machines. Let A_i denote the event that the ith machine breaks down in a year, $i = 1, 2, 3$. With the help of events A_i express the following statements:

a) only the first brakes down;
b) all three break down;
c) none of the machines breaks down;
d) the first and the second do not break down;
e) the first and the second break down, the third does not;
f) only one machine breaks down;
g) at most one machine breaks down;
h) at most two machines break down;
i) at least one machine breaks down.

2.6 What is the connection between the events A and B if

a) $A \cdot B = A$,
b) $A + B = A$,

c) $A + B = \overline{A}$,

d) $A \cdot B = \overline{A}$,

e) $A + B = A \cdot B$?

2.7 Under which conditions the following equality holds:

$$A + (B \cdot \overline{A}) = B \ ?$$

2.8 Show that the intersection of countably many σ-fields is a σ-field.

3 Classical probability space

3.1 Two fair dice are thrown. Find the probability that the sum of the numbers obtained is 8. Illustrate the sample space and the set of favourable events.

3.2 Three fair dice are thrown. Find the probability that the sum of the numbers obtained is a prime number.

3.3 A fair dice is thrown twice. Find the probability that the result of the first throw is greater than the result of the second.

3.4 Ten coins are tossed. Find the probability that all of them show head or all of them show tail.

3.5 9 balls are put randomly into 4 boxes. Find the probability that each box contains at least two balls.

3.6 A box contains n balls labelled by numbers $1, 2, \ldots, n$. One by one we draw out all the n balls. Find the probability that

a) each drawn ball but the first has a greater label than the previous one.

b) the kth drawn ball is labelled by k.

c) the kth drawn ball is labelled by k and the ℓth drawn ball is labelled by ℓ $(k \neq \ell)$.

3.7 Ten persons, 5 women and 5 men are sitting around a round table. Find the probability, that neither two women nor two men are sitting next to each other.

3.8 n persons of different heights are sitting around a round table. Find probability that the tallest and the shortest are sitting next to each other.

3.9 From a deck of cards three cards are dealt. Find the probability that there isn't any spade among them.

3.10 In a dark room we have four pairs of the same shoes mixed. Find the probability that if four shoes are chosen we have at least one pair among them.

3.11 In an urn we have three red balls. Find the minimal number of white balls to be added to have the probability of choosing a white ball be greater than 0.9.

3.12 In an urn we have 6 red and some white and black balls. The probability that a randomly chosen ball is either black or white is $3/5$; either red or black is $2/3$. Give the numbers of white and black balls contained in the urn.

3.13 20 fragile objects are packed in a box. Five of them are of value 10 euros each, four are of value 20 euros each, seven are of value 50 euros each and four are of value 100 euros each. Somebody drops the pack and breaks four objects. Assuming that the objects break independently of each other find the probability that the total loss is 100 euros.

3.14 In an urn we have 20 red and 30 white balls. 10 balls are chosen without replacement. Find the probability that

a) all the chosen balls are red.

b) 4 red, 6 white.

c) at least one red.

3.15 Solve the previous exercise under the assumption that the balls are chosen with replacement.

3.16 From 100 bananas 10 are rotten. What is the probability of having a rotten one among five randomly chosen bananas?

3.17 Find the probability that on the lottery 5 from 90 we hit at least three winning numbers.

3.18 In an urn we have 3 red, 3 white and 3 green balls. Find the probability of having all three colours among 6 randomly chosen balls.

3.19 What is the probability that two persons in a group of four have their birthdays on the same day (365 day of a year considered)?

3.20 From 40 questions a student learned just 20. On the exam he has to chose randomly two questions, but than he is free to chose one of the two to work on it. Find the probability that he passes the exam.

3.21 Find the probability that in a poker hand (5 cards out of 52) we get exactly 4 of a kind.

4 Geometric probability

4.1 On a rectangular target with sides of one meter lengths each a circle is drawn with radius of 0.5 meter. Find the probability that a random shot (given it hits the target) hits the target outside the circle.

4.2 A stick of length one meter is randomly broken into two parts. What is the probability that from the obtained parts and from a new stick of half a meter length a triangle can be formed?

4.3 A stick of length one meter is randomly broken into three parts. What is the probability that from the obtained parts a triangle can be formed?

4.4 Choose two points randomly from the $(0,1)$ interval. Find the probability that the distance of the two points is less than the distance between point 0 and the chosen point which is closer to 0.

4.5 With the help of two randomly chosen points the $(0,1)$ interval is cut into three. Find the probability that each of the obtained three sections is shorter than $1/2$.

4.6 Choose randomly two positive numbers that are less than 1. What is he probability that their sum is less than 1 and their product is less than $\frac{2}{9}$?

4.7 Choose randomly two positive numbers that are less than 1. What is he probability that the geometric mean of the two numbers is less than $1/2$?

4.8 In 24 hours time two ships arrive independently into the harbour of Chewbakka Bay, denoted by A and B, respectively. Ship A can be unloaded in an hour, while ship B in two hours. Workers start to unload a ship immediately after it's arrival and if the other ship arrives before they finish it has to wait. What is the probability that none of the ships has to wait?

4.9 Choose randomly two points from the interval $(-1,1)$ and let α and β denote their coordinates. Find the probability that equation

$$x^2 + \alpha x + \beta = 0$$

has real roots.

4.10 Two points are chosen randomly from the interval $(0,a)$. What is the probability that the sum of the squares of their coordinates is greater than a^2

4.11 Choose two points randomly from the opposite sides of a unit square. What is the probability that the distance of the two points is less than α where $1 \le \alpha < \sqrt{2}$?

5 Conditional probability, Bayes' theorem

5.1 Show that for arbitrary events A and B with $\mathsf{P}(B) > 0$

$$\mathsf{P}(\overline{A} \mid B) = 1 - \mathsf{P}(A \mid B)$$

holds.

5.2 Suppose $\mathsf{P}(B \mid A) > \mathsf{P}(B)$ and $\mathsf{P}(C \mid B) > \mathsf{P}(C)$. Do these facts imply $\mathsf{P}(C \mid A) > \mathsf{P}(C)$?

5.3 Let $\mathsf{P}(A) = 1/4$, $\mathsf{P}(A \mid B) = 1/4$ and $\mathsf{P}(B \mid A) = 1/2$. Calculate the probabilities $\mathsf{P}(A + B)$ and $\mathsf{P}(\overline{A} \mid \overline{B})$.

5.4 Two dice are rolled. Find the probability that the sum of the numbers obtained is 7 given the sum is odd.

5.5 Two dice are rolled. Find the probability that at least one of them shows six, given they show different values.

5.6 A dice is rolled until the first six appears. Find the probability that we stop after two throws given that an even number of throws is performed.

5.7 We know that at least one of the two kids in a family is a girl. Find the probability of having also a boy in the family.

5.8 Choose two points randomly from the unit interval. Find the probability that both points are closer to a previously specified end point of the interval than to each other given their distance is less than $1/2$.

5.9 From a box containing 5 red and 5 white balls 3 balls are chosen without replacement. Given the first two chosen balls are of the same colour find the probability that the third chosen ball is red.

5.10 Four cowboys, Bill, Jim, Charles and Jack are playing cards, each of them has 13 cards in hands. Given Bill haven't got any aces, find the probability that there aren't aces in Jim's hands.

5.11 From a group of n students r randomly chosen students have to write a test. Find the probability that the weakest student has to write the test, given the best student has to.

5.12 Mosquitoes are usually killed using sprayed chemicals. Scientists found that the first treatment kills 80% of the mosquitoes. However, the surviving insects become resistant, so the second treatment kills 40%, while the third kills just 20% of the mosquitoes.

a) What is the probability that a mosquito survives all three treatments?

b) Given a mosquito survived the first treatment, what is the probability that it survives two more treatments?

5.13 Tippler Joe spends $2/3$ of the day in pubs. There are five pubs in the village and as Joe does not have any preference, he visits each of them with equal probability. One day we want to find Joe. We have already checked four pubs without finding him. What is the probability that Joe is sitting in the fifths pub?

5.14 In a TV quiz show the player has to choose one from three envelopes. In the first envelope there are 5 cards saying 'Sorry, next time', 3 cards with 'You have won 100 euros' and 2 cards with 'You have won 500 euros'. The content of the second envelope: 2 cards 'Sorry, next time', 7 cards 'You have won 100 euros' and 1 card 'You have won 500 euros'. The third envelope contains only 'Sorry, next time' cards. The player chooses randomly an envelope and from the chosen envelope he chooses a card. What is the probability that the player wins 500 euros?

5.15 Alice and Bob play the following game. Alice rolls a dice and tosses two coins as many times as the dice shows. If there are two heads among the results of the tosses then Bob shall pay one euro to Alice, otherwise Alice shall pay the same amount to Bob. Who has the higher chance to win?

5.16 Humans have four different types of blood. 38% of mankind belongs to group A, 21% to group B, 8% to group AB, and 33% percent to group 0. In case of blood transfusion for a patient belonging to group A the donor can belong either to group A or to group 0. Similarly, for group B the appropriate blood types are B or 0; for group AB any type of blood is appropriate; for group 0 only the same group can be applied. Find the probability that the blood of a randomly chosen donor can be transfused to a randomly chosen patient.

5.17 There are six six-shooter revolvers laying on a table. Three of them are loaded with 1-1 bullets, two of them with 2-2 bullets and the sixth is with 3 bullets. We chose randomly a revolver and we pull the trigger. Find the probability that the chosen revolver shots.

5.18 Consider the revolvers of the previous exercise. Given a randomly chosen revolver shots, find the probability that no more bullets left in the chamber.

5.19 Rust Rider cars are produced in four factories. The the first factory produces 200 cars per day, the second 320, the third 270, while the fourth 210. The refuse ratios for the factories are 2%, 5%, 3% and 1%, respectively. We bought a Rust Rider and we found it perfect. What is the probability that it had been produced in the fourth factory?

5.20 The overseas flights of the Hornet Airways are operated with aircrafts of types D, E and F, all types fly with the same probability. On type D there are six seats per row, on type E four, on type F three (each row has two window seats) and the passengers get their seats randomly. Given you have a window seat find the probability that you are flying with type F.

5.21 We have two coins, a fair one and a loaded one where the probability of a head is double of the probability of a tail. We choose randomly a coin and we toss it. Given the result is head, find the probability that the loaded coin was tossed.

5.22 During one of his journeys Ulysses arrives to a triple turnout. The first road leads to Athens, the second to Mycenae, the third to Sparta. Athenians are merchants, the like to sham their guests and in two third of the cases they lie. Mycenaean are a bit better: they lie only in each second case. Due to their strict traditions Spartans are honest, they always tell the truth. Ulysses does not know where to go (the directions are not indicated), so he chooses a road randomly. After arriving to the city at the end of the road Ulysses asks a local man, how much is 2×2 and the answer is 4. What is the probability that Ulysses has arrived to Athens?

5.23 A professional gambler has a loaded dice where the probability of rolling a six is $2/3$, while the other four values have equal probabilities. He wants to play but he also has three fair dice in his pocket that look similar to the loaded one. The gambler chooses a dice randomly and he rolls it. Given the result is six, what is the probability that the loaded dice has been found?

5.24 In an office equipped with mechanized administration three machines classify the files. The first can process 10 files per day, the second 15, while the third 25. The average numbers of misclassified files are 0.3, 0.9 and 0.5 per day, respectively. We choose a file randomly from the daily production and we find that it has been misclassified. What is the probability that the file was processed by the first machine?

5.25 In a certain country 10% of the cabs are green and 90% are blue. The witness of an accident claimed that the wounded pedestrian had been hit by a green cab. Later the police found out that the witness mixes colours: in the case of blue and green he can find the true colour in 85% of the cases. Find the probability that the guilty cab was really green.

5.26 75% of the products produced in a factory are first class products, the rest are of second class. Later all products are double checked. The probability that a first class product is evaluated as a second class one is 0.02, while the probability that a second class product is found to be a first class one is 0.05. Find the probability that a product evaluated as a first class product is really a first class one.

5.27 On a noisy binary channel 2/5-th of the transmitted digits 0 are transformed into 1 and 1/3-rd of the transmitted digits 1 into 0. The ratio of the transmitted digits 0 and 1 is 5 : 3.

a) Given a 0 is received what is the probability that a 0 has been transmitted?

b) What is the probability of receiving a digit 1?

5.28 On a multiple-choice test three answers correspond to each question and only one answer is correct. A student taking the test knows the correct answer with probability p. Otherwise, he chooses an answer randomly, i.e. with probability $1/3$. Given his answer is correct, what is the probability that he knew the right answer and not only guessed?

6 Independence of events

6.1 A coin is tossed ten times. Let A denote the event that there are both heads and tails among the results, while B denotes the event that there is at most one tail among them. Are A and B independent?

6.2 In a box there are 2 red and 4 white balls. Four balls are chosen without replacement. Let A denote the event that the first chosen ball is white, while B denotes the event that the last chosen ball is white. Are A and B independent?

6.3 In a box we have 8 cards numbered from 1 to 8. A card is chosen randomly. Let events A, B and C denote the following:

A: the chosen number is even;

B: the chosen number is not greater than 4;

C: the chosen number is either 2 or greater than 5.

Show that

$$\mathsf{P}(A \cdot B \cdot C) = \mathsf{P}(A)\mathsf{P}(B)\mathsf{P}(C)$$

and the three events are not mutually independent.

6.4 In an urn we have 4 similar cards with three binary digits on each card. On the first we have $0,0,0,$ on the second $0,1,1,$ on the third $1,0,1,$ and on the fourth $1,1,0.$ A card is chosen randomly. Let A_i denote the event that the ith digit on the chosen card equals 1, $i = 1,2,3$. Show that the events A_i, $i = 1,2,3,$ are pairwise independent but not mutually independent.

6.5 Two lottery coupons (5 from 90) are filled independently of each other. What is the probability of winning, i.e. of hitting at least two winning numbers?

6.6 In the Csokonai Restaurant (4024 Debrecen, Kossuth street 21) together with the bill the waiter brings four dice. The guests can roll them three times and if four sixes appear at least once they win a voucher of value 2000 forints. Find the probability of this event.

6.7 Two soldiers shoot on a target in turn until the first hit. The probability that the starter hits it is $0.2,$ while the second hits it with probability $0.3.$ What is the probability that the first successful shot belongs to the soldier who started the shooting?

6.8 American Indians transfer their messages using skywriting and they forward them from tribe to tribe. Assume that the probability that a given message can be correctly decoded is 0.9. After how many forwards will the probability of the correct decode be less than 1/2?

6.9 The ten digits are written on ten separate cards. A cards is chosen randomly, the digit on it is noted and the card is replaced. How many cards should be chosen to have an even number among them with probability greater than 0.9?

6.10 5% of 500 plastic Garfield figures packed into a container have some minor errors. Before shipping the container, the quality controller randomly chooses 10 figures, checks them then replaces the figures. This procedure is repeated once again. The container can be shipped only if the all the checked Garfield figures are error free. What is the probability that the container can be shipped?

6.11 In an urn we have some red and white balls. Two balls are chosen with replacement. Show that the probability that the chosen balls have the same colour is at least 0.5.

6.12 Let A, B and C be independent events, $\mathsf{P}(A) = 0.1$, $\mathsf{P}(B) = 0.2$ and $\mathsf{P}(C) = 0.3$. What is the probability that more than one of the events A, B and C occur?

6.13 On a test exam each student has to answer 20 questions either with *yes* or with *no*. Assume that for each question a student knows the right answer with probability p, q is the probability that the student assumes that he/she knows the correct answer, but he/she is wrong, while r is the probability that the student knows that he/she doesn't know the correct answer $(p+q+r=1)$. In the latter case the student chooses either *yes* or *no* with equal probabilities. What is the probability that a given student has at least 19 correct answers?

7 Discrete random variables

7.1 Which of the following sequences form a discrete probability distribution?

a) $p^4,\ 4p^3q,\ 6p^2q^2,\ 4pq^3,\ q^4, \quad q=1-p,\ 0<p<1;$

b) $p^kq^2, \quad q=1-p,\ 0<p<1, \quad k=1,2,\ldots;$

c) $\big(k(k+1)\big)^{-1}, \quad k=1,2,\ldots;$

d) $p^{k-n}q, \quad q=1-p,\ 0<p<1, \quad k=n,n+1,\ldots.$

7.2 Two dice are rolled. What are the distributions of the minimum and of the maximum of the numbers obtained?

7.3 Two dice are rolled. What is the distribution of the distance (absolute value of the difference) of the numbers obtained?

7.4 Give the distribution of the smallest winning number of the lottery (5 from 90).

7.5 Two football players shoot penalty kicks in turn till the first goal. The probability that the starter shoots a goal is 0.5, while for the other player this probability equals 0.6. Give the distribution of the number of penalty kicks performed (inclusive the last kick).

7.6 In an experiment rats have to choose a door from four doors where they can find their lunch. After an unsuccessful trial the rat is put back to the starting point and he can try again until the right door is found. Give the distribution of the number of trials if

a) at each trial the rat chooses randomly (with equal probabilities) from the four doors;

b) at each trial the rat chooses randomly (with equal probabilities) from the previously not chosen doors;

c) at two subsequent trials the rat never chooses the same door, it chooses randomly (with equal probabilities) from the remaining doors.

7.7 In a box there are 22 cards labelled with numbers from 1 to 22. A card is chosen randomly. Let ξ denote the remainder obtained after dividing the chosen number by 2, while η denotes the remainder obtained after dividing the chosen number by 3. Give the joint distribution of (ξ,η) together with the marginal distributions.

7.8 The following table specifies the joint distribution of (ξ, η):

$\xi \backslash \eta$	-1	0	1
-1	p	$3p$	$6p$
1	$5p$	$15p$	$30p$

a) Find the value of p?

b) Are ξ and η independent?

c) Give the distributions of $\xi + \eta$ and $\xi \cdot \eta$.

7.9 Let the distribution of (ξ, η) be the distribution given in the previous example. Find the probabilities

a) $\mathsf{P}(\eta = i \,|\, \xi = -1) \quad (i = -1, 0, 1)$;

b) $\mathsf{P}(\eta < 1 \,|\, \xi = -1)$;

c) $\mathsf{P}(\eta \geq 0 \,|\, \xi = 1)$;

d) $\mathsf{P}(\xi = 1 \,|\, \eta \geq 0)$.

8 Parameters of discrete random variables

8.1 Two dice are rolled till the first six appears on one of them. What is the mean number of rolls required, inclusive the last one?

8.2 Two dice are rolled. What are the means of the minimum and of the maximum of the values obtained?

8.3 The cost of a lottery coupon 5 from 90 is 1 dollar. In case of two winning numbers hit the lottery pays 6 dollars, a triple pays 82 dollars, four winning numbers result 5000 dollars while a fiver pays 10 million dollars. Buying a single coupon what is the average gain?

8.4 A coin is tossed. If the result is a head, it is tossed one more time, otherwise it is tossed two more times. What is the mean number of heads obtained?

8.5 A blind knife thrower hits the target with probability $1/4$ and he keeps trying till the first hit. Find the mean and the standard deviation of the required trials.

8.6 Peter rolls a dice. If the result is odd, he looses 1 dollar, if it is six, he wins 4 dollars, otherwise he can roll again. If the second roll is even he wins 1 dollar, otherwise he looses 2 dollars. Find out whether this game is advantageous, fair or disadvantageous for Peter, i.e. his mean gain is positive, zero or negative, respectively. What is the variance of Peter's gain?

8.7 The nominal value of a share is one golden galleon. In a year the value either can doubled, or halved or remain the same – each of the events has the same probability. On the next year the same happens, independently of the events of the previous year. Find the distribution of the value of the share after two years. What is the mean and the variance of the value?

8.8 In basketball after certain fouls 'one plus one' penalty throws are judged. This means, a player can throw a penalty and in case of success one more. Assume that the player can realize each penalty with probability 0.6 and the penalties are independent of each other. Find the distribution, the mean and the standard deviation of successful penalties.

8.9 Four dice are rolled. Find the mean and the standard deviation of the sum of the numbers obtained.

8.10 Ten football players kick penalties. Find the mean number of goals if each player has two possibilities to kick and the probabilities of shooting goals are $p_1, p_2, \ldots, p_{10}$, respectively.

8.11 A dice is rolled one hundred times. Find the mean and the variance of the sum of the even numbers obtained.

8.12 Let ξ and η be independent binomial random variables with parameters $(10, 0.4)$ and $(6, 0.5)$, respectively. Find the mean and the variance of $\xi + \xi \cdot \eta$.

8.13 A flea performs a random walk on the x axis, i.e. at each step it jumps either one unit to the left or one unit to the right with equal probabilities. Let ξ denote the distance of the flea from the origin after four steps. Find the standard deviation of ξ.

8.14 The values of a random variable ξ are $-1,\ 0,\ 2,\ 3,$ while the corresponding probabilities are $1/12,\ 5/12,\ 1/4, 1/4,$ respectively. Find the mean and the standard deviation of ξ^2.

8.15 Let ξ be a Poisson random variable with parameter λ. Find the mean of the random variable $\eta = \frac{1}{1+\xi}$.

8.16 A fair dice is rolled 100 times. Let ξ denote the number of values 3 obtained from the first 50 rolls, while η denotes the number of even values obtained from the second 50 rolls. Find the variance of $\eta - \xi$.

8.17 In an office the number of letters arriving to the director on a particular day can be modelled by a Poisson random variable with parameter λ. The secretary of the director makes a pre-selection and a letter is forwarded to the director with probability p independently from the other letters. Find the distribution and the mean of the forwarded letters.

8.18 In a box we have 4 white, 3 red and 3 black balls. Two balls are drawn, one after the other, without replacement. Let ξ denote the result of the first draw, namely $\xi = 0$

if the first ball is black, $\xi = 1$ if it is red and $\xi = 2$ if it is white. In a similar way let η be the result of the second draw.

a) Are ξ and η independent?

b) Find the standard deviation of ξ.

8.19 In a box we have 9 cards labelled with numbers 11, 12, 12, 22, 23, 23, 31, 31, 33. A card is chosen randomly. Let ξ denote the first while η the second digit of the label of the chosen card. Are ξ and η independent? Find the distribution of $\xi \cdot \eta$ and the covariance of ξ and η.

8.20 The values of the random vector (ξ, η) are the lattice points (points with integer coordinates) of the interior of the square determined by points $(0,0)$, $(0,4)$, $(4,4)$, $(4,0)$. All points but the central have the same probability while the probability of the central point is four times the probability of any other point. Find the covariance of ξ and η. Are ξ and η independent?

8.21 On a coin 0 is written on the head side and 1 on the tail side. The coin is tossed twice. Let ξ and η denote the sum and the product of the numbers obtained, respectively. Are ξ and η independent? Find the correlation coefficient of ξ and η.

8.22 Alice and Bob play the following game. They roll a dice. If the dice shows 1 Alice wins, if it shows 6 Bob wins, otherwise none of them. Let ξ and η denote the number of games won by Alice and Bob, respectively, after two rolls.

a) Find the joint distribution of ξ and η.

b) Find the variance of $\xi + \eta$ and the correlation coefficient of ξ and η.

8.23 A fair dice is rolled n-times. Let ξ denote the number of cases when the dice shows 6 while η denotes the number of odd numbers obtained. Are ξ and η independent? Find the correlation coefficient of ξ and η.

8.24 Let ξ and η denote the values obtained after rolling a dice twice. Find the correlation coefficient of ξ and $\zeta := \max\{\xi, \eta\}$.

8.25 Let ξ and η be independent Poisson random variables with parameters λ and μ, respectively, and let $\zeta := \xi + \eta$. Find the correlation coefficient of ξ and ζ.

8.26 Let ξ_1, ξ_2 and ξ_3 be independent Poisson random variables with parameters λ_1, λ_2 and λ_3, respectively. Let $\eta_1 := \xi_1 + \xi_2$ and $\eta_2 := \xi_2 + \xi_3$. Find the covariance of η_1 and η_2.

8.27 In an urn we have 20 red and 30 white balls. 20 balls are chosen without replacement. Find thee mean and the variance of the number of red balls chosen.

9 General notion of random variables

9.1 Two dice are rolled. Find and plot the cumulative distribution function (cdf) of the sum of the numbers obtained.

9.2 In a certain zoo there are two sloths. Every day the first sleeps in half of the time while the second in one third of the time, independently of the first. Let ξ denote the number of sloths that are awaken during our visit in the zoo. Find the cdf of ξ.

9.3 Check whether the following functions are cumulative distribution functions or not.

a) $F(x) := \begin{cases} 0, & \text{if } x < 1/2, \\ \frac{x-1}{x+1}, & \text{if } x \geq 1/2; \end{cases}$

b) $F(x) := \begin{cases} 0, & \text{if } x < 1, \\ \frac{x-1}{x+1}, & \text{if } x \geq 1; \end{cases}$

c) $F(x) := \begin{cases} 0, & \text{if } x < 1, \\ \frac{2x-1}{x+1}, & \text{if } x \geq 1; \end{cases}$

d) $F(x) := \begin{cases} 0, & \text{if } x < 0, \\ \frac{x^3}{1+x^2}, & \text{if } x \geq 0; \end{cases}$

e) $F(x) := \begin{cases} 0, & \text{if } x \leq 0, \\ 1 - \frac{1-e^{-x}}{x}, & \text{if } x > 0. \end{cases}$

9.4 Find the values α and β such that

$$F(x) := e^{-\beta e^{-\alpha x}}, \qquad x \in \mathbb{R},$$

is a cdf of a random variable.

9.5 Find the cdf of the distance of two randomly chosen points of the $[0,1]$ interval. What is the probability that the distance of the points is in the interval $[1/2, 3/4]$?

9.6 Chose a point randomly on the unit square. Let ξ denote the distance of the chosen point from the closest side of the square. Find the cdf of ξ. Find the probability $\mathsf{P}(\xi \geq 1/8)$.

9.7 We shoot on a round target having unit radius. Assume that each shot hits the target and the location of the hit is uniformly distributed on the target. Let ξ denote the distance of the hit from the center of the target. Give the cdf of ξ.

9.8 We shoot on a square target having unit sides. Assume that each shot hits the target and the location of the hit is uniformly distributed on the target. Let ξ denote the distance of the hit from the lower left corner of the target. Find the cdf of ξ.

9.9 A random number is chosen from the $[0,2]$ interval. Let η denote the following random variable: if the chosen number is from the $[0,1]$ interval then η will equal this number, otherwise it will equal to 0. Find the cdf of η.

10 Absolutely continuous random variables

10.1 A stick of length two meters is randomly broken into two parts. Find the cumulative distribution function (cdf) and the probability density function (pdf) of the length of the shorter part.

10.2 A point is chosen randomly on the interval $(0, a)$. Let ξ denote the distance of the point from the center of the interval. Find the cdf and pdf of ξ.

10.3 Two points are chosen randomly on the intervals $(0, a)$ and $(a, 2a)$, respectively. Let ξ denote the distance of the chosen points.

a) Find the cdf and pdf of ξ.

b) What is the probability that the distance of the chosen points is less than $a/2$?

10.4 Check whether the following functions are probability density functions or not.

a) $f(x) := \begin{cases} \frac{sin x}{2}, & \text{if } 0 < x < 1, \\ 0, & \text{otherwise}; \end{cases}$

b) $f(x) := \begin{cases} \frac{1}{x^2}, & \text{if } x > 1, \\ 0, & \text{otherwise}; \end{cases}$

c) $f(x) := \dfrac{1}{\pi} \dfrac{1}{1 + x^2}, \qquad x \in \mathbb{R};$

d) $f(x) := \begin{cases} \frac{x}{x+1}, & \text{if } x > 0, \\ 0, & \text{otherwise}; \end{cases}$

e) $f(x) := \begin{cases} 4x^3 e^{-x^4}, & \text{if } x > 0, \\ 0, & \text{otherwise}; \end{cases}$

f) $f(x) := \dfrac{1}{2} e^{-|x|}, \qquad x \in \mathbb{R};$

g) $f(x) := \begin{cases} -\frac{e^{-x}}{x} + \frac{1-e^{-x}}{x^2}, & \text{if } x > 0, \\ 0, & \text{otherwise.} \end{cases}$

10.5 The pdf of a random variable ξ equals

$$f(x) = \begin{cases} 0, & \text{if } x < 2, \\ \frac{A}{(1-x)^2}, & \text{if } x \geq 2. \end{cases}$$

a) Find the value of A.

b) Find the probability $\mathsf{P}\big(2 < \xi < 3\big)$.

c) Find the cdf of ξ.

10.6 The pdf of a random variable ξ equals

$$f(x) = \begin{cases} A\cos\frac{x}{2}, & \text{if } 0 < x < \pi, \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the value of A.

b) Find the probability $\mathsf{P}\big(\xi > \frac{\pi}{2}\big)$.

c) Find the cdf of ξ.

10.7 The pdf of a random variable ξ equals

$$f(x) = \begin{cases} 0, & \text{if } x \le 2, \\ \frac{A}{x^3}, & \text{if } x > 2. \end{cases}$$

a) Find the value of A.

b) Find the value q such that $\mathsf{P}\big(\xi \ge q\big) = 1/2$.

c) Find the cdf of ξ.

10.8 Choose a point randomly on the $[0,1]$ interval of the x (horizontal) axis. Let ξ denote the distance of the chosen point from the $(0,1)$ point of the coordinate system. Find the cdf of ξ.

10.9 Let ξ be uniformly distributed on the (a,b) interval. Find the cdf and pdf of $\eta := 2\xi + 1$.

10.10 Let ξ be an exponential random variable with parameter λ. Find the cdf and pdf of $\eta := 2\xi + 3$.

10.11 Let ξ be uniformly distributed on the $(0,1)$ interval. Find the pdf of $\eta := \xi^2$.

10.12 Let ξ be uniformly distributed on the $(-1,1)$ interval. Find the pdf of $\eta := \xi^2$.

10.13 Let ξ be an exponential random variable with parameter λ. Find the pdf of $\eta := \xi^3$.

10.14 Let ξ be an exponential random variable with parameter λ. Find the pdf of $\eta := \sqrt{\xi}$.

11 Joint distributions, independence

11.1 The joint cdf of (ξ, η) equals:

$$F(x,y) := \begin{cases} 1 + e^{-x-y} - e^{-x} - e^{-y}, & \text{if } x > 0,\ y > 0, \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the marginal distribution functions of ξ and η.

b) Are ξ and η independent?

c) Find the probability $\mathsf{P}\big(\xi < 1,\, \eta < 1\big)$.

11.2 The joint cdf of (ξ, η) equals:

$$F(x,y) := \begin{cases} \min\big\{(1-e^{-x}), (1-e^{-y})\big\}, & \text{if } x>0,\ y>0, \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the marginal distribution functions of ξ and η.

b) Are ξ and η independent?

11.3 Let the joint pdf of (ξ, η) be:

$$f(x,y) := \begin{cases} e^{-x-y}, & x>0,\ y>0, \\ 0, & \text{otherwise.} \end{cases}$$

Find the probabilities $\mathsf{P}\big(\xi < 1,\, \eta < 1\big)$ and $\mathsf{P}\big(\xi < 1,\, \eta \geq 3/2\big)$.

11.4 Let the joint pdf of (ξ, η) be:

$$f(x,y) := \begin{cases} \frac{4}{5}(x+xy+y), & \text{if } 0<x<1,\ 0<y<1, \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the marginal density functions.

b) Are ξ and η independent?

c) Find the probabilities $\mathsf{P}\big(\xi < 1/2,\, \eta < 1/2\big)$ and $\mathsf{P}\big(\xi < 1/2,\, \eta \geq 1/4\big)$.

11.5 Let the joint pdf of (ξ, η) be:

$$f(x,y) := \begin{cases} A\Big(x+\frac{y}{2}\Big), & \text{if } 0<x<1,\ 0<y<2, \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the value of A.

b) Find the marginal density functions.

c) Are ξ and η independent?

11.6 Find the values of A such that

$$f(x,y) := x^2 + Ay^2$$

is a pdf of a random vector (ξ, η) with range $(0<x<1;\ 0<y<2)$. Find the marginal density functions of ξ and η.

11.7 Let the joint pdf of (ξ, η) be:

$$f(x,y) := \begin{cases} A, & 0 \leq x \leq 1,\ 0 \leq y \leq x, \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the value of A.

b) Find the marginal density functions.

c) Are ξ and η independent?

11.8 Let the joint pdf of (ξ, η) be:

$$f(x,y) := \begin{cases} 1, & 0 \le x \le 1,\ 0 \le y \le 2(1-x), \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the marginal density functions.

b) Are ξ and η independent?

c) Give the probability $\mathsf{P}\big(\xi < x,\ 1 < \eta < 3/2\big)$ as a function of x.

12 Parameters of absolutely continuous random variables

12.1 Find the means and standard deviations of the random variables specified by the following probability density functions.

a) $f(x) := \begin{cases} |x|, & \text{if } -1 \le x \le 1, \\ 0, & \text{otherwise;} \end{cases}$

b) $f(x) := \dfrac{1}{2} e^{-|x|}, \quad x \in \mathbb{R};$

c) $f(x) := \begin{cases} \sqrt{\frac{2}{\pi}}\, e^{-x^2/2}, & \text{if } x > 0, \\ 0, & \text{if } x \le 0; \end{cases}$

d) $f(x) := \begin{cases} \frac{3}{x^4}, & \text{if } x > 1, \\ 0, & \text{if } x \le 1. \end{cases}$

12.2 Show that the random variable specified by the pdf

$$f(x) := \begin{cases} \frac{2}{x^3}, & \text{if } x > 1, \\ 0, & \text{if } x \le 1, \end{cases}$$

has a finite mean but it does not have a finite variance.

12.3 Let ξ be uniformly distributed on the $[0,1]$ interval. Find the mean and standard deviation of $\eta = \xi^2$.

12.4 Let ξ be uniformly distributed on the $[-a, a]$ interval, where $a > 0$. Find the mean of $\eta = |\xi|$.

12.5 Let ξ be an exponentially distributed random variable with mean 2. Find the mean and standard deviation of $\eta = \xi^2$.

12.6 Two points are chosen randomly on the $[0,1]$ interval. Find the mean and standard deviation of the distance of the chosen points.

12.7 Choose a point randomly from the unit square. Let ξ denote the distance of the chosen point from the closest side of the square. Find the mean and standard deviation of ξ.

12.8 Let (ξ,η) be a random vector specified by the following pdf:

$$f(x,y) := \begin{cases} \frac{6}{5}\big(x+y^2\big), & \text{if } 0<x<1,\ 0<y<1, \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the correlation coefficient of ξ and η.

b) Are ξ and η independent?

12.9 Let (ξ,η) be a random vector specified by the following pdf:

$$f(x,y) := \begin{cases} \frac{1}{4}\big(1+xy(x^2-y^2)\big), & \text{if } |x|\le 1,\ |y|\le 1, \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the correlation coefficient of ξ and η.

b) Are ξ and η independent?

12.10 Let (ξ,η) be uniformly distributed on the triangle determined by vertices with coordinates $(0,0)$, $(2,0)$ and $(2,1)$. Find the correlation coefficient of ξ and η.

12.11 Let (ξ,η) be uniformly distributed on the triangle determined by vertices with coordinates $(0,0)$, $(0,1)$ and $(1,0)$. Find the correlation coefficient of ξ and η.

12.12 Let (ξ,η) be a random vector specified by the following pdf:

$$f(x,y) := \begin{cases} 1, & \text{if } 0\le x\le 1,\ 0\le y\le 2(1-x), \\ 0, & \text{otherwise.} \end{cases}$$

Find the correlation coefficient of ξ and η.

12.13 Let (ξ,η) be a random vector specified by the following pdf:

$$f(x,y) := \begin{cases} A\Big(x+y\Big), & \text{if } 0<x<1,\ 0<y<1, \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the value of A.

b) Give the marginal density functions.

c) Are ξ and η independent?

d) Find the correlation coefficient of ξ and η.

12.14 Let (ξ, η) be a random vector specified by the following pdf:

$$f(x,y) := \begin{cases} A\Big(x^2+y^2\Big), & \text{if } 0<x<1,\ 0<y<1, \\ 0, & \text{otherwise.} \end{cases}$$

a) Find the value of A.

b) Give the marginal density functions.

c) Are ξ and η independent?

d) Find the correlation coefficient of ξ and η.

13 Special absolutely continuous distributions

Uniform distribution

13.1 Let ξ be a uniformly distributed random variable with $\mathsf{E}\xi = \mathsf{D}^2\xi = 4$. Give the cdf of ξ.

13.2 Let the distribution of ξ be uniform on the $(a, 5)$ interval and assume

$$\mathsf{P}\big(\xi \geq \mathsf{E}(\xi^2 - 2\xi + 1)\big) = \frac{1}{6}.$$

Find $\mathsf{P}\big(\xi \leq \mathsf{E}(\xi - 1)\big)$.

13.3 Let ξ be uniformly distributed on the $(0, 1)$ interval.

a) Find the probability that the first decimal of the value of ξ equals 2.

b) Find the probability that the second decimal of the value of ξ equals 2.

c) Find the probability that the kth decimal of the value of ξ equals 2.

13.4 Let ξ be uniformly distributed on the (a, b) interval where the values a and b are unknown. However, we know that the interval $(2, 5)$ is a sub interval of (a, b) and

$$\mathsf{P}(2 \leq \xi \leq 5) = \frac{1}{3}.$$

a) Find $\mathsf{P}(3 \leq \xi \leq 5)$.

b) What is the minimal value of a and the maximal value of b?

c) Under the above conditions give an estimate of $\mathsf{P}(1 \leq \xi \leq 3)$.

Normal distribution

13.5 The air control informs the pilot of an aircraft about the altitude of the centre of the air corridor of height 100 metres where the aircraft should fly. The deviation, in metres, of the altitude of the aircraft from the given altitude is normally distributed with mean 20 and standard deviation 50. Find the probability that the aircraft flies under the air corridor, in the air corridor, above the air corridor.

13.6 An automatic packing machine fills cartons with 1 kilogram of sugar. The amount of sugar filled into a carton is a normal random variable with mean 1 kg and standard deviation 32 g. The quality control accepts a carton if its weight is between 0.95 and 1.05 kilograms.

a) Find the probability that a randomly chosen carton is accepted.

b) Find the probability that at least one of two randomly chosen cartons is accepted.

13.7 Let ξ be normally distributed with parameters $m = 3$ and $\sigma = 2$. Find the smallest value of A that ensures that the probability that ξ is in the $(2, A)$ interval is at least 1/2.

13.8 A sawmill produces boards. The length of boards, in centimetres, is normally distributed with mean 400 and standard deviation 3.

a) Find the percentage of boards that are longer than 398 cm and shorter than 401 cm.

b) Find the probability that the length of a randomly chosen board differs from the mean with at most 2.5 cm.

13.9 A cannon shoots on a target at a distance of 1200 metres. The deviation of the length of a shot around the 1200 m is a normal random variable with a standard deviation 40 m. A shot is efficient if the hit is closer to the target than 50 m. Find the percentage of inefficient shots.

13.10 The diameter, in centimetres, of bearings produced by a machine is normally distributed with mean 15 and standard deviation 0.5. Find the probability that the difference in the diameter of a bearing produced by this machine is more than 5% of the required value.

13.11 The number ξ of calls arriving to a certain break-down assistance can be considered as a normal random variable with a standard deviation $\sigma = 10$. What is the mean number of calls if $\mathsf{P}(\xi < 20) = 0.1$?

13.12 The Light Bulb Company produces fluorescent lamps. The life-span of the lamps follows normal distribution with mean 1170 hours and standard deviation 100 hours. The company gives warranty on the lamps and during the warranty period it replaces a dead lamp free of charge. How long the warranty period should last if the company wants to replace at most 5% of the sold lamps.

Exponential distribution

13.13 The probability that at a filling station one has to wait more than 6 minutes for the service is 0.1. Given the waiting time is exponentially distributed find the probability that we are served in 3 minutes after arrival?

13.14 The life-span of a plasma TV is an exponentially distributed random variable with a mean life-span of 9 years. Find the largest number K such that the probability that the TV is operable for K years is at least 0.9.

13.15 We are standing in front of a telephone box and waiting for the previous person to finish the call. The duration of the previous call is random with a pdf $\frac{1}{3}e^{-x/3}$, $x > 0$.

a) Find the probability that the duration of the call is longer than 3 minutes.

b) Find the probability that we have to wait for more than 3 minutes given the person in the phone box has been speaking fore more than 3 minutes.

c) Find the probability that the duration of the previous call is longer than $t + 3$ minutes given the person in the phone box has been speaking fore more than t minutes.

13.16 Let ξ and η be independent exponentially distributed random variables with common parameter λ. Find the pdf of $\min\{\xi, \eta\}$.

13.17 A power-loom works with 400 threads. The life-span of each thread, that is the time until it breaks, is exponentially distributed with a mean of 150 hours. Assuming that the threads break independently of each other find the probability that the loom has to be stopped for thread breaking in 3 hours after the its start.

14 Convolution of absolutely continuous distributions

14.1 Let ξ and η be independent exponentially distributed random variables with parameters $\lambda > 0$ and $\mu > 0$, respectively. Find the pdf of $\xi + \eta$.

14.2 Let ξ and η be independent random variables that are uniformly distributed on the $[0, 1]$ interval. Find the distribution of $\xi + \eta$.

14.3 Let ξ and η be independent random variables that are uniformly distributed on the $[-1/2, 1/2]$ interval. Find the pdf of $\xi + \eta$.

14.4 Let ξ and η be independent random variables that are uniformly distributed on the intervals $[0, 1]$ and $[2, 4]$, respectively. Find the distribution of $\xi + \eta$.

14.5 A point is chosen randomly from the square determined by vertices with coordinates $(0, 0)$, $(0, 1)$, $(1, 1)$ and $(1, 0)$. Let ξ denote the sum of the distances of the point from the coordinate axes. Find the cdf and pdf of ξ.

14.6 Two sticks of one meter length each are randomly broken and the two shorter parts are glued together. Find the distribution of the length of the obtained new stick.

14.7 Let ξ and η be independent random variables with common pdf $f(x) := e^{-|x|}/2$, $x \in \mathbb{R}$. Find the distribution of $\xi + \eta$.

14.8 Let ξ and η be independent standard normal random variables. Find the distribution of $\xi + \eta$.

14.9 Let ξ and η be independent standard normal random variables. Show that $\xi^2 + \eta^2$ is exponentially distributed with parameter $\lambda = \frac{1}{2}$.

15 Chebyshev's inequality, law of large numbers, central limit theorem

15.1 The number ξ of newspapers sold by a newsagent at a certain railway station follows Poisson distribution with mean $\lambda = 64$. Give a lower bound for the probability $\mathsf{P}(48 < \xi < 80)$.

15.2 In the Sweet Life Sugar Refinery the cartons are filled with an automatic packing machine. The amount of sugar, in kilograms, filled into a given carton is a random variable with a mean 5 kg and standard deviation of 10 g. Find an upper bound for the probability that the amount of sugar in a randomly chosen carton differs from the mean by more than 50 grams.

15.3 The mean length, in centimetres, of sausages produced by a certain meat-packing company is 35, the variance is 0.3. Find an upper bound for the probability that the length of a randomly chosen sausage differs from the mean by at least one centimetre.

15.4 In a match making company the matchboxes are filled with an automatic machine. The number ξ of match-sticks in a randomly chosen box is a random variable with the following distribution:

number	47	48	49	50	51	52	53
probability	0.05	0.10	0.15	0.40	0.15	0.10	0.05

a) Using Chebyshev's inequality give an estimate of $\mathsf{P}(48 < \xi < 52)$.

b) Find the exact value of the above probability.

15.5 How many times should a fair coin be tossed in order to obtain a relative frequency that approximates the probability of head with an error less than 1/20 with probability greater than or equal to 0.9?

15.6 How many times should a fair dice be rolled in order to obtain a relative frequency that approximates the probability of obtaining a six with an error less than 0.1 with probability greater than or equal to 0.8?

15.7 How many times should a loaded dice be rolled in order to obtain a relative frequency that approximates the probability of obtaining a six (that is not necessarily 1/6) with an error less than 0.1 with probability greater than or equal to 0.8?

15.8 We want to determine the proportion of alcoholics in a certain social stratum. How many observations should be performed in order to obtain a relative frequency that approximates the true proportion with an error less than 0.1 with probability greater than or equal to 0.95?

15.9 10% of plastic Garfield figures have some minor errors. Before shipping the figures the quality controller checks the containers. A container can be shipped if the proportion of defected figures is at most 12%. How many figures should be packed into a container to obtain a relative frequency that approximates the probability with an error less than 0.02 with probability greater than or equal to 0.95?

15.10 In an urn there are some white and black balls, the proportion of white balls is 70%. 1000 balls are chosen with replacement. Find an approximation of the probability that the number of chosen white balls is between 680 and 720. Solve the exercise with normal approximation, too.

15.11 A fair coin is tossed 200 times. Find an estimate of the probability that the number of heads is between 95 and 105.

15.12 A fair dice is rolled 300 times. Find the 95% confidence bounds for the number of sixes obtained.

15.13 200 shots are shot independently on a target. For each shot the probability of success (i.e. of hitting the target) is 0.4. Find the 90% confidence bounds for the number of successful shots.

15.14 At the presidential elections of the USA in year 2000 the result in Florida was extremely tight. 5 000 000 constituents had to chose between the candidates of the democrats and of the republicans. The difference between the numbers of votes obtained by the two candidates was 300. Given each constituent chooses between the two candidates with equal probabilities find the probability that the difference between the numbers of votes gathered by the two candidates is not more than 300.

15.15 A fair dice is rolled 1200 times and the sum of the even numbers obtained is calculated. Give an approximation of the probability that the sum is between 2280 and 2500.

16 Basic notions of statistics, parametric test

16.1 *Drink it* tea is sold in cartons of tea bags. The weight, in grams, of the content of each carton in any batch is supposed to be 200. A quality controller had bought five cartons and weighted their contents. The results, in grams, are the following:

$$196,\ 202,\ 198,\ 197,\ 190.$$

Find the mean and the variance of the sample and plot the empirical cumulative distribution function.

16.2 The following 8 numbers were generated by a random number generator generating uniform random numbers on the $(0,1)$ interval.

$$0.18,\ 0.57,\ 0.82,\ 0.55,\ 0.63,\ 0.12,\ 0.91,\ 0.31.$$

Find the mean and the variance of the sample and plot both the empirical and the true cumulative distribution function.

16.3 In a random sample of size five the sum of the sample values equals 155, while the sum of their squares equals 4837. Find the mean and the variance of the sample.

16.4 The volumes, in millilitres, of soft drink of a random sample of 10 bottles from a large batch were as follows:

$$499,\ 525,\ 498,\ 503,\ 501,\ 497,\ 493,\ 496,\ 500,\ 495.$$

Stating clearly your null and alternative hypotheses, test the claim that the mean volume of soft drink in bottles in the batch is 500 ml. Use a 95% significance level. Assume that, for the batch, volumes are normally distributed with a standard deviation of 3 ml.

16.5 *Drink it* tea is sold in cartons of tea bags. The weight, in grams, of the contents of the cartons in any batch is known to be normally distributed with a standard deviation of 4 g. A quality controller had bought five cartons and weighted their contents. The results, in grams, are the following:

$$196,\ 202,\ 198,\ 197,\ 190.$$

Stating clearly your null and alternative hypotheses, test the claim that the mean amount of tea in cartons is less than 200 g. Use a 98% significance level.

16.6 The diastolic blood pressure, in millimetres of mercury, of a population of healthy adults have mean 84.8 and standard deviation 12.8. The diastolic blood pressures of a random sample of members of an athletics club were measured with the following results:

$$79.2,\ 64.6,\ 86.8,\ 73.7,\ 74.9,\ 62.3.$$

a) Stating clearly your null and alternative hypotheses, test, at the 95% significance level, whether the mean diastolic blood pressure of members of the athletics club is less than 84.8. Assume that the standard deviation for members of the club is 12.8 and that the distribution is normal.

The diastolic blood pressures of a random sample of members of a chess club were also measured with the following results:

$$84.6,\ 93.2,\ 104.6,\ 106.7,\ 76.3,\ 78.2.$$

b) Stating clearly your null and alternative hypotheses, test, at the 95% significance level, whether the mean diastolic blood pressure of members of the chess club is higher then that of the members of the athletics club. Assume that the standard deviation for members of the chess club is also 12.8 and that the distribution is normal.

16.7 Given two random samples from normal distribution with a standard deviation of 0.0012. The mean of the first sample of size 9 is 0.1672, while the mean of the second one, that contains 16 sample values, equals 0.1683. Can we accept at the 92% level of significance that there is no difference between the theoretical means?

16.8 Experiments with a new variety of tomato are being conducted at an agricultural research station. The crop is grown under carefully controlled conditions on ten experimental plots and the yields, in kg per plot, are found to be as follows:

$$60.2,\ 63.4,\ 58.8,\ 63.6,\ 64.7,\ 62.5,\ 66.0,\ 59.1,\ 65.1,\ 62.0.$$

It is known that an established variety of tomato would have a mean yield of 60 kg per plot on the experimental plot. Test, at the 95% level of significance, whether the new variety has a different mean yield, stating clearly your null and alternative hypotheses. State also the assumptions underlying your analysis.

16.9 A notional allowance of 9 minutes has been given for the completion of a routine task on a production line. The operatives have complained that it appears usually to be taking slightly longer.

An inspector took a sample of 12 measurements of the time required to undertake this task. The results (in minutes) were as follows:

$$9.4,\ 8.8,\ 9.3,\ 9.1,\ 9.4,\ 8.9,\ 9.3,\ 9.2,\ 9.6,\ 9.3,\ 9.3,\ 9.1.$$

Stating clearly your null and alternative hypotheses and the assumptions underlying your analysis test, at the 99% level of significance, whether the task is indeed taking on average longer than 10 minutes.

16.10 Some athletes in *Dopeland* have complained that the samples they have given for drug testing have not been analysed correctly and that they have been accused of unduly high levels of certain substances when in fact that is not the case. The Dopelandian Athletics Federation investigated the accuracy of the laboratory which tested the samples. Eight samples each containing a known level of exactly 0.500 g/l of a banned substance were sent to the laboratory to assess the level of the substance in g/l. The results are given below:

$$0.485,\ 0.518,\ 0.460,\ 0.530,\ 0.560,\ 0.550,\ 0.490,\ 0.575.$$

Test the athletes' claim at the 95% level of significance, given the measurement error of the laboratory is normally distributed.

16.11 The Price Checking Agency is comparing prices of food in two large supermarket. The prices (in euros) of a random sample of items of food sold in both supermarkets are found to be as follows.

Item	1	2	3	4	5	6	7	8	9	10
Irobyou (euro)	23.2	7.9	18.8	5.6	4.9	4.6	1.9	3.7	3.3	1.9
Howexpensive (euro)	21.6	7.4	20.8	5.2	4.2	4.9	1.8	3.1	3.8	1.7

Given difference in prices is normally distributed, examine at the 95% level of significance whether there is a difference in the mean prizes of the two supermarkets.

16.12 The freshmen of the *Sheep's Academy* have to write two tests on probability theory. The percentage marks are as follows:

Student	A	B	C	D	E	F	G	H	I	J
I. test	57	63	67	82	45	65	53	32	51	27
II. test	53	62	63	80	46	64	44	28	50	29

Given the difference in results is normally distributed, examine at the 95% level of significance whether the two test were equally difficult.

16.13 To estimate the amount of undeclared tip of hairdressers, using information obtained from their guests, the revenue office estimated the true amount of tip earned by ten randomly chosen hairdressers in a weeks time. The obtained data are the following:

Hairdresser	A	B	C	D	E	F	G	H	I	J
Declared (1000Ft/week)	4.0	2.0	3.5	5.0	1.8	6.0	2.8	1.5	3.9	4.4
True (1000Ft/week)	9.0	5.3	6.0	9.8	4.3	10.1	5.9	4.2	9.4	10.5

Given the undeclared amount of tip follows normal distribution, test, at the the 95% level of significance, that the mean of the weakly amount of undeclared tip is more that 5000 Ft.

16.14 In comparison of the dissolving times of two types of instant coffee several experiments were made. In each case the same amount of coffee was put into 1 decilitre of boiling water and the dissolving times were recorded. The results are as follows:

Coffee	Dissolving times (seconds)							
Worst Coffee	8.2	5.0	6.8	6.7	5.8	7.3	6.4	7.8
Coffee In	5.1	4.3	3.4	3.7	6.1	4.7		

a) Given the dissolving times are normally distributed, show, at the 95% level of significance, that there is no evidence of a difference in variability between the dissolving times of the two types of coffee.

b) Hence, using the same level of significance as in part a), investigate the claim that the mean developing time of Worst Coffee exceeds that of Coffee In.

16.15 In comparison of the lengths of jumps of two different species of fleas several experiments were made. The lengths of jumps, in centimetres, of 12 fleas of species A and 10 fleas of species B were recorded.

Species	Length of jump (cm)											
A	53	59	63	67	60	57	73	65	58	68	62	71
B	61	52	47	51	58	64	60	55	49	53		

Given the jump sizes of both species are normally distributed, examine at the 99% level of significance whether fleas of species A jump longer than fleas of species B.

16.16 Parunder PLC manufacture golf equipment and believe that they can increase their market share by introducing a new longer-lasting and cut resistant golf ball. The research department has developed a new coating for golf balls and tests have been very promising. However, some concern has been expressed about the possible detrimental effect of the coating on driving distance. As a result, a comparative experiment is undertaken using an automatic driving machine so as to remove human violation. Forty-two balls are used: 26 with the current coating, 16 with the new coating. A summary of the distances driven, in yards, which for each coating may be assumed to be normally distributed, are recorded in the table below.

Coating	Sample size	Sample mean	Corrected empirical variance
Current	26	271.4	35.58
New	16	268.7	48.47

a) Show that, at the 90% significance level, the new coating has made no difference to the variability in the distance driven.

b) Hence, using the same level of significance as in part a), investigate the claim that the new coating decreases the mean distance driven.

16.17 A company undertakes investigations to compare the fuel consumption in l/100km, of two different cars, Marcides and Luxes, with a view to purchasing a number of company cars. The fuel consumption of both models is known to be normally distributed with equal variances. For a random sample of 12 Marcideses and 15 Luxeses the results are summarized as follows.

Model	Sample size	Sample mean	Corr. emp. variance
Marcides	12	8.4806	1.0703
Luxes	15	7.3799	0.8967

Stating clearly your hypotheses test, at the 95% significance level, whether there is a difference between the mean fuel consumption of the two models of car.

16.18 The variance of the lengths of a sample of 9 tent-poles produced by a machine was 63 mm^2. A second machine produced a sample of 13 tent-poles with a variance of 225 mm^2. Both these values are unbiased estimates of the population variances (corrected empirical variances).

a) Test, at the 90% level, whether there is evidence that the machines differ in variability, stating the null and alternative hypotheses.

b) State the assumptions you have made about the distribution of the populations in order to carry out the test in part a).

17 Chi-square tests

17.1 Cartons of a particular brand of mixed nuts contain five different types. It is claimed that the numbers of nuts of types A, B, C, D and E are, in percentage terms, 35, 25, 20, 10 and 10, respectively. A randomly selected carton is found to contain the following numbers of nuts of each type.

Type of nut	A	B	C	D	E
Number	184	145	100	68	63

Test the hypothesis that the numbers of nuts of different type in this carton are consistent with those claimed. Use a 90% level of significance.

17.2 A computer generates 12 observations taken at random from the uniform distribution on $[-6,6]$ and calculates their mean; this is repeated 100 times. The distribution of the 100 sample means is grouped into four classes, as follows.

	observed frequency
$(-\infty,-0.6745)$	26
$[-0.6745,0)$	21
$[0,0.6745)$	27
$[0.6745,\infty)$	26

Examine the hypothesis, at the 95% level of significance, that the sample means are equally likely to fall into any of these four classes.

17.3 A computer is programmed to generate integers from the set $1,2,\ldots,15$. A random variable X denotes the number of integers generated to first obtain an integer which is a multiple of 3. The following distribution of X is suggested.

$$\mathsf{P}(X=\ell)=p(1-p)^{\ell-1}, \qquad \ell=1,2,\ldots,$$

where p is the probability that a generated integer is a multiple of 3. The table shows the results of 160 observations on X.

Number of integers generated (x)	1	2	3	4	5	6	7	8	>8
Frequency (k)	63	34	28	13	9	7	2	4	0

a) Find the sample mean.

b) Test at the 95% level of significance that the suggested model with $p=1/3$ is appropriate for X.

c) Test at the 90% level of significance that the suggested model for X is appropriate.

17.4 A biologist was attempting to test a theory that a species of leaf insect has a lifespan whose distribution can be modelled as uniform between 0 and 20.5 days. He collected the data shown in the table.

Lifespan (in days, to the nearest whole day)	0–2	3–5	6–10	11–20
Number of insects	38	53	75	112

a) Test at the 95% level of significance whether the sample justifies the theory.

The biologist also observed that, in the sample he had taken, no insect survived for longer that 16 days. He decided to revise his theory by reducing the upper limit of the distribution to 16.5.

b) Using the new model perform a test at the 95% level of significance and comment whether this refinement seems justified.

17.5 A student of botany believed that a certain species of plants grow in random positions in grassy meadowland. He recorded the number of plants in one square metre of grassy meadow, and repeated the procedure several times. The results are given in the following table.

Number of plants	0	1	2	3	4	5	6	at least 7
Frequency	9	24	43	34	21	15	2	0

a) Find the mean number of plants in one square metre.

According to the handbooks of botany, the observations of this type can be modelled by Poisson distribution.

b) Test at the 95% level of significance whether or not the Poisson model is supported by the above data.

17.6 In a study of the possible association between the severity of a particular disease and age, an investigation is made of a random sample of 200 medical records of people suffering from the disease. The records are classified according to the severity of the disease and the age of the sufferer. The frequencies are as follows.

		Age		
		less than 40	40–60	more than 60
Severity	mild	41	34	9
	medium	25	25	12
	severe	6	33	15

Stating clearly the null and alternative hypotheses test at the 99% level of significance whether there is an association between the severity of the disease and the age of the sufferer.

17.7 A member of the Confederation of Looserlandian Industry was speaking about members' views on joining Looserland to the European Community. He claimed that the size of the business influences the view taken. A survey of a random sample of businesses discovered results as shown in the table.

	Size of business		
	Large	Medium	Small
In favour	13	24	76
Against	7	26	143

a) Test at the 99% level, whether the data support the claim in the first paragraph.

It was later found that a clerk had omitted one of the returns, and the number of medium size businesses in favour of integration should have been 25.

b) State the conclusion of the test using the the new data.