GIZAPEDIA

gizapedia.hirusta.io

STATISTICS FOR BUSINESS

Workbook

Author: Josemari Sarasola

1 Probability calculus

1.1 Combinatorics

1. How many ways are there to complete 3 tasks one after the other? Make the list of those ways. Idem with 4, 5 and 10 tasks.
2. How many ways are there to schedule 4 work days and 2 holidays in the next 6 days?
3. There are 6 workers in a shop. You have to choose 2 of them to work on Saturday. In how many ways can you do it?
4. Among 8 people you have to choose the president, the vice president and the secretary. How many ways are there to do it?
5. 6 men work in a factory. 4 tasks have to be completed. If men are able to complete one task or more, how many ways are there to assign the tasks? Remark: Guys are distinguishable.
6. A firm packs delicatessen products into gift boxes. A box must be filled with 2 big products and 2 small products. The customer can choose among 4 big products and 6 small products. How many ways are there to fill a box?
7. You have to choose a commission of 2 students between 10 students. Order is not relevant. Two students refuse to be together in the commission. How many ways are there to choose the commission. Idem if the commission comprises 4 students.
8. To list the reference codes of the product from a catalog you have to choose 3 letters. If the catalog includes 3200 products, at least how many letters must you choose from?

1.2 Laplace rule

9. You have to deliver 4 packages to 4 customers, but as you have lost their addresses, you deliver the packages randomly. What is the probability of delivering them in the correct way?
10. A postman has to deliver 4 letters. She chooses randomly the delivery order.
    (a) What is the probability of choosing first the closest recipient?
    (b) What is the probability of delivering A letter immediately before B letter?
11. **Urn problem**: In an urn there are 12 faultless and 4 faulty items. You extract 4 items at one time (or without devolution).
    (a) What is the probability of all of them being faultless?
    (b) What is the probability of being 3 faultless items?
    (c) Which is the event with the bigger probability? Why?
12. **Urn problem**: There are 6 women and 8 men in a group. 4 people have to be drawn randomly.
    (a) What is the probability of all of them being women?
    (b) What is the probability of Anne and Mary being drawn together?
    (c) What is the probability of drawing 2 men and 2 women?
    (d) What is the probability of drawing 2 men and 2 women, if we add 6 more men to the main group?
13. We have 6 A, 8 B and 10 C type pieces in a box. We draw at the same time 6 pieces randomly.
    (a) What is the probability of drawing 3 A, 2 B and 1 C type pieces?
    (b) What is the probability of drawing 5 A type pieces?
14. We have received only one night bookings in our hostel for the next week, but we don't know the exact day. What is the probability of all of them being in different days?
15. There are 5 queues in a supermarket. Customers choose randomly their queue. What is the probability of being among 3 customers of being 2 of them in A queue?

1.3 Event algebra

16. There are 345 people in a village. Among those. 123 people gets bread at Anne's bakery. There are 102 people who gets bread at Laura's bakery. Everybody else makes bread at home. Formulate using event algebra symbols these events and calculate their probability:

    (a) A person gets brad at a bakery.

    (b) A person makes bread at home.

17. 60 people are attending a congress. 40 of them speak only Spanish, 20 of them French. What is the probability of two of them, randomly chosen, understand each other.

18. We have asked some people about the film they saw last week (A or B). With the data we have drawn this contingency table:

The film you have seen	A yes	A no	Total
B yes	61	84	145
B no	109	112	221
Total	170	196	366

    What is the probability of a person seeing *at least* one film?

19. Tomorrow the probability of raining is 0.8. The day after tomorrow the probability of raining is 0.6. The probability of raining during the following two days is 0.5.

    (a) What is the probability of raining *at least* one day?

    (b) What is the probability of no raining during the following two days?

20. A product may have A and B failures, with a probability of 0.04 and 0.11 respectively. Both failures appear with a 0.01 probability.

    (a) What is the probability of a product not having any failure?

    (b) What is the probability of a product having only one failure?

21. Let A and B events. Applying event algebra symbols, quote these complex events:

    (a) A and B don't happen at the same time,

    (b) at least one happens,

    (c) only one happens,

    Let A, B and C events. Applying event algebra symbols, quote these complex events:

    (d) only A happens,

    (e) at least one happens,

    (f) the three events happen at the same time,

    (g) A and B happen, but not C

    (h) at least two events happen,

    (i) no event happen.

22. In a urn there are 12 faulty items, and 46 faultless. We draw 4 pieces randomly at the same time. What is the probability of drawing at least 2 faultless items?

23. We have collected these data from a classroom:

Students	Math	Geography	Biology	Language
A	5.6	3.6	4.0	8.2
B	3.2	6.4	5.6	4.6
C	2.2	3.4	3.4	4.0
D	8.2	8.6	9.7	9.2
E	7.2	4.5	6.4	7.0
F	2.2	5.4	6.6	4.3
G	6.8	5.3	3.9	4.1
H	8.2	7.6	4.4	8.3
I	4.2	5.5	7.5	6.8
J	7.6	8.0	3.3	5.4
K	6.2	8.5	3.5	4.8
L	3.6	3.9	6.5	6.9
M	4.2	3.5	4.5	5.2
N	4.2	5.5	2.3	4.0

Apply this code: MP: math passed; LN: language not passed, and so on. Calculate these probabilities:

(a) P(MP and BP);

(b) P(MN or LP), applying inclusion-exclusion rule;

(c) P(MP or GP or BN), applying inclusion-exclusion rule. Which is the student who doesn't gather in that probability?

24. In a bag you have 10 A type pieces, 20 B type eta 30 C type. You draw randomly 2 pieces, which is the probability of all of them being of the same type?

1.4 Conditional probability. Dependence and independence. Multiplication rule.

25. In a urn we have 22 faultless and 7 faulty pieces. We draw randomly 4 pieces, without devolution.

(a) What is the probability of all being faultless?

(b) What is the probability of the first three being faultless and the last one faulty?

(c) What is the probability of drawing 3 faultless and 1 faulty pieces?

(d) What is the probability of drawing 2 faultless and 2 faulty pieces?

26. In a urn we have 12 faultless and 4 faulty pieces. We draw randomly 6 pieces, without devolution (or at the same time). What is the probability of having 2 faulty pieces or less?

27. 200 men eta 300 women live in a village. We draw a sample of 5 people.

(a) What is the probability of drawing the number of men and women that matches the population proportion? Calculate the probability with devolution as well as without devolution.

(b) What is the probability of drawing one man in the sample drawn without devolution? Compare this result with the former one.

(c) What is the probability of drawing at least one man? Calculate the probability with devolution as well as without devolution.

28. In a factory we produce with total independence day by day. We can produce 1, 2, 3 or 4 items with 0.1, 0.2, 0.3 eta 0.4 probabilities respectively. Within five days, calculate the probability of producing:

(a) on the first two days 2 or more, and on the last two days less than 2,

(b) 1 only on each of 2 days,

(c) always producing the same number of items,

(d) always producing 3 or less,

(e) 4 on at least one day.

29. A share price increases with a 0.4 probability, when the day before increased, and with 0.7 probability, when the day before the price decreased.

(a) What is the probability of the share price increasing on the first two days and decreasing on the next two days?

(b) Within the next two days, what is the probability of the share price increasing on one day and decreasing on the other?

30. You have a 4 pieces batch, but you don't know which of the are faulty and faultless. Trying to improve the quality of the batch, you draw pieces one by one: if the piece you draw is faulty, you withdraw that piece and put into the batch 2 faultless pieces; if the piece you draw is faultless, you return it into the batch. You draw 3 pieces. If you have 2 faulty and 2 faultless pieces in the batch, calculate the probabilities of all the possible outcomes, taking the order into account. If the three pieces you have drawn are oxo, what should you deduce about the number of faulty pieces in the urn at the beginning?

1.5 Probability trees

31. 600 men and 400 women have signed up for an insurance policy with a company. Among these, 200 had an accident: 150 men and 50 women. What is the probability of a person having an accident? Give the probability directly as well as by means of the law of total probability, taking into account the sex.

32. Possible daily sales in an auto dealership are 1 to 4, with 0.1 to 0.4 probabilities respectively, with total independence. What is the probability of selling 6 or more autos in two days?

33. A student takes two partial exams in a course. The probabilities of passing these exams are 0.5 and 0.6 respectively. Partial exams are saved and in the final exam the student must take only the failed partial exams. The probabilities of passing these final partial exams are 0.7 and 0.8 respectively. What is the probability of finally passing the two exams?

34. Because of budget restictions, a company that sells a type of machine can attend only two exhibitions this year: A and B exhibitions, or alternatively, C adn D exhibitions. Here you have the probabilities for selling the possible numbers of machines:

Numbers →	0	1	2
A	0.10	0.20	0.30
B	0.35	0.30	0.20
C	0.20	0.20	0.30
D	0.30	0.30	0.20

The goal is to maximize the probability of selling 3 machines al least. Which exhibitions should the company attend?

35. We have three applicants for a job. Two exams will be hold through the recruitment; the probability for passing the first exam is 0.6, and 0.3 for the second one. What is the probability of remaining 0, 1, 2 and 3 applicants after the exams?

36. Our stock price increases with a 0.4 probability, if it increased the day before; and decreases with a 0.7 probability, under the same circumstances. Through the next four days, what is the probability of the stock price increasing over two days and decreasing over the other two?

1.6 Bayes theorem

37. In a population, the percentage of people being affected by a disease is 15%. In order to diagnosticate the disease we have performed a test, but not perfectly: among people with desease, the test performs well in the 80% of cases; among people woth no desease, the test says the person is ill in the 10% of cases. For Lucy, the test was positive. What is the probability of Lucy of being ill? And what if the test was negative?

38. A, B and C machines produced 2000, 5000 and 3000 items respectively last year. A machines produces faulty items with a 5% probability. B and C machines produce them with 10% and 2% probability respectively. A customer made a complaint for a faulty item. What should you say about the machine it comes from?

39. A factory makes boxes of chocolate:
    - the red box contains 4 A type chocolates, 6 from B type eta 10 from C type;
    - the blue box contains 10 A type chocolates, 4 from B type eta 6 from C type;
    - the yellow box contains 8 A type chocolates, 8 from B type eta 4 from C type.

    We sent to a shop 200 red boxes, 100 blue boxes and 300 yellow boxes. The shopkeeper found two A type chocolate wrapping papers outside the batch. So, he thought somebody opened a chocolate box, picked up two chocolates, ate them and left the papers out there. He must find the box missing 2 A type chocolates. Which box type should he begin searching from?

40. In a test each question has 4 possible answers. The probability of knowing the answer, and therefore, giving the correct answer is 0.7. We know that 10% of student leave the question blank, and the rest of them answer it randomly. A student answered correctly. What is the probability of really knowing the answer? How mahy options sholuld we give in each question in order to, supposed the answer is correct, the probability of knowing it, is 0.96? Idem, with a 0.99 probability.

41. Broadly speaking, the customers of a new product will be 10%, 20%, 30% and 40% of the total customers in the market. There is total uncertainty about the probability of those values. A survey took 10 people and among them 6 claimed they would purchase that product. How should we change the a priori probabilites about the percentage of customers?

42. In the customer database of an insurance company, we have these data: 100 policy holder had an accident last year, 200 policy holder didn't have any accident last year, 400 policy holders didn't have any accident last two years, and 500 policy holders didn't have any accident last thrre years. According to our estimations, a person that had accident last year has a 0.22 probability of having another accident next year, and that probability declines by 0.04 by year without accident. We have just received a claim about an accident, but we know nothing about the policy holder. Into which policy holder group should we enter him?

2 Random variables

2.1 Discrete random variables

43. We have 6 faulty washing machines in stock, and 12 faultless. We sold 8 of them. As we don't know which of these are faulty and faultless, we delivered them randomly. We expect the customer who takes a faulty machine will claim.
    (a) Give the probability mass function and the cumulative distribution function of the number of claims.
    (b) We need a spare part for each claim. How many spare parts do we need in order to satisfy all the claims with a probability of 0.8

44. The time lasting until we get a new customer in a web page is a r.v. given by this function:
$$P[X = x] = 0.1 + \frac{k}{x}\ ;\ x = 1, 2, 3, 4$$
    (a) Calculate k, $P[X = x]$ to be a mass function.
    (b) Give mass function and cumulative distribution function as a table.

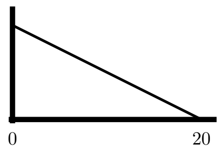
45. The distribution of the X random variable is given by:
$$P[X = x] = \frac{1}{k}\ ;\ x = 1, 2, \dots, k$$
Calculate k to be a mass function.

46. The distribution of the X random variable is given by:
$$F(x) = 1 - \left(\frac{1}{2}\right)^x\ ;\ x = 1, 2, \dots$$
Give $P[X = 3]$, $P[X = 2]$, $P[X > 4]$ and $P[X < 6]$.

2.2 Continuous random variables

47. The percentage of students that pass an exam in given in this way:
$$f(x) = 2 - 2x\ ;\ 0 < x < 1$$
    (a) Draw the density function and interpret it.
    (b) What is the probability of being more than 10% of students passing the exam.
    (c) What is the probability of students passing the exam exceeding 50%.
    (d) Calvulate the probability of the percentage being in the %20-%30 interval. Warning: calculate for the closed and open intervals.
    (e) Give the distribution function and calculate by its means the previous probabilities.
    (f) Proof tha given functions are true pdf and cdf.
    (g) If you want the 20% of students to pass the exam, which should be the pass note?

48. Sales in a store (thousands euros) follow this distribution:



   (a) Give the density function.
   (b) Give the cdf
   (c) Calculate the probability of the salens benig more than 10.000 euros, by means of both the pdf and the cdf.
   (d) Supposed independence between sales in different days, calculate the probability of sales exceeding everyday 10.000 euros over 5 consecutive days. Provided that really happened, which conclusion should be drawn?

49. A random variable follows this distribution:

$$f(x) = k - x \ ; \ 0 < x < k$$

   (a) Calculate k the previous function to be a pdf.
   (b) Give the cdf.
   (c) Calculate $P[0.5 < X < 1]$, by means of both the pdf and the cdf.

50. Daily production (kg) is rv, depending on the number of machines (k):

$$f(x) = \frac{2x}{100k^2} \ ; \ 0 < x < 10k$$

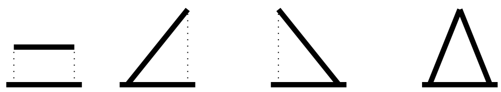
   (a) Proof that k is a parameter.
   (b) With 4 machines, what is he probability of production exceeding 30?
   (c) Give cdf, depending on k, and based on that, proof k is a parameter.
   (d) If product is packed by the kilogram, how many packs do we need, with 6 machines, in order to be the probability of packing all the production 0.8? And to be 0.9?
   (e) How many machines do we need in order to the probability of the production being more than 60 reaching 0.5?

51. The time to complete a task is given by the following cdf:

$$F(x) = \frac{x^3 - 8}{19} \ ; \ \ 2 < x < 3$$

   (a) Proof it's a cdf.
   (b) Give the pdf.
   (c) Calculate $P[X \geq 2.5]$ by means of both the pdf and the cdf.

52. Draw in a approximative way the cdfs corresponding to these pdfs:



53. The weight of an *clementina* orange follows this distribution:

$$f(x) = \frac{2}{10000}(x - 100) \ ; \ 100 < x < 200$$

   Calculate the probability of the weight being exactly 110 gr, both in a theoretical way and a practical way, known than balance gives 100-110-120-130-... weights. And 130 gr or less?

54. The number of customers entering a store follows this distribution:

$$f(x) = \frac{1}{100} \ ; \quad 200 < x < 300$$

   What is the probability of entering exactly 250 customers? And 250-260 customers (closed interval)?

2.3 Expected value, variance and other moments

55. Sales in a month (units) distribute in this way:

Number	Probability
0	0.05
1	0.2
2	0.25
3	0.25
4	0.15
5	0.1
	1

Give the expected value and interpret it.

56. In a urn we have 20 faulty items, and 80 faultless. We extract 5 items randomly and at the same time. Give the expected value of the number of faulty items and propose a formula for calculating that value in general cases.

57. The number of items produced by a machine (thousands) distribute in this way:

$$f(x) = \frac{x}{2} \ ; \ 0 < x < 2$$

Graphic the density function and approximate it at first sight. Calculate it.

58. Monthly sales may be 1 and 2, with 0.4 and 0.6 probabilities respectively.
    (a) Which is the expected sales value for two months?
    (b) Variable cost and selling price per item are 100€and 200€, respectively. Which the expected gain value for one month, given that fixed costs are 50 €?
    (c) Through the last three months we sold 2, 1 adn 2 items, which is the average gain per month. Why is not the same as the expected value?

59. The tenperature in a freezer in any moment has the following distribution:

$$f(x) = \frac{1}{a} \ ; \ 0 < x < a$$

    (a) Plot the distribution and intepret it about the mean value.
    (b) Calculate the mean value and the standard deviation.
    (c) Calculate the 3rd central moment.

60. The number of faults in an item follows this distribution:

Number of faults	Probability
0	0.5
1	0.4
2	0.1
	1

    (a) Calculate the 3rd raw moment.
    (b) Calculate the expected value and the standard deviation.
    (c) Calculate the 2nd central moment.
    (d) Calculate the mean value and standard deviation of faults in 100 items, both in the case of independence and dependence betwen items about the numbers of faults.

61. The numbers of items produced in a factory may be 2 and 3 with 0.3 and 0.7 probabilities respectively. The next day, where the production the day before was 2, the production follows the same distribution; where the prudction was 3, the production may be 3 and 4 wtih 0.3 and 0.7 probabilities respectively. And so on the following days: whenever the maximum production happens, the next day we can produce that maximum and one more item; otherwise, the production is the same as the day before.
    (a) What is the probability of producing 12 or more items in the next 4 days?
    (b) Give the mean value of the production through the next 4 days.
    (c) If items are produced uniformly throughoute the time, how much time do we need to produce 8 pieces on average?

62. The is just one worker in a factory. His production may be 1 or 2 pieces , with 0.4 and 0.6 probabilities. When the production is 1, the nex day we take another worker, which could produce 1 or 2 pieces, just like the firts worker and with the same probabilities, but independently with him. When the production of the first worker in the first day is 2, we don't take the second worker and productions follow the same distribution as the fist day.

    (a) Give the expected production for those two days.

    (b) Give the expectet time we need to produce 2 finished items.

2.4 Expected value and variance as criterions for decision

63. A firm has four investment choices. The gains of each investment are random and distribute in the following manner (negative gains are losses):

Gains	A investment	B investment	C investment	D investment
-2	0.05	0	0	0.05
-1	0.25	0.20	0.15	0.10
0	0.30	0.50	0.40	0.35
1	0.20	0.25	0.30	0.40
2	0.20	0.05	0.15	0.10

If needed, use this utility function:

$$U = \frac{\mu}{0.5\sigma} - 3P[loss]$$

    (a) Show that the utility function is correct.

    (b) Comparing two investments at once, discuss the preference for all the investments.

    (c) Sort the investments about the expected value, the risk and probability of loss and discuss which is the best investment.

64. We have to choose betwen two stock options, whose gains distribute in the following way:

    - A stock option:

$$f(x) = \frac{1}{2} - \frac{x}{8} \ ; \ 0 < x < 4$$

    - B stock option:

$$f(x) = \frac{1}{5}; \ 0 < x < 5$$

If needed, you should use this utility function:

$$U = \frac{\mu}{\sigma}$$

Discuss which is the best choice, in the long term as in the short term.

2.5 Chebysev's unequality

65. We don't know the exact distribution of the daily production, but we know that the mean is 100 items, and the standard deviation 20. What is the probability of the production being in the 60-140 interval?

66. The standard length for a given piece is 1000 mm. On average we comply the standard, with a 100 mm standard deviation. We accept the piece if the deviation from the standard value is less than 300mm. What is the probability of not accepting the piece?

67. Daily sales in a shop are 2000€ in a normal day, with a 200€ deviation. Whenever we predict sales will be more than 2600€, we will hire more workers. What is the probability of this happening?

68. The mean number of sandwiches requested in a day in a restaurant is 100. Deviation is 10. We need a bun for each sandwich. How many buns do we need in order of to be the probability of having enough buns to make all requested buns 0.9? Idem, with a 0.99 probability. Idem, when the deviation is 20 (and the probability 0.99).

69. Daily sales in a shop are 100€, on average, with a 200€ deviation. Bound the probability of being the sales 0f 3 days less than 2800.

3 Binomial distribution

70. The probability of a student passing an exam is 0.7.
    (a) In a group pf 10 students, what is the probabilitaty of 6 students passing the exam?
    (b) And the probability of 6 students passing the exam?
    (c) And the probability of x students passing the exam?
    (d) Write in a simplified notation the distribution of the number of students passing the exam.
71. In a factory, we must produce 12 items for a customer. The probability of each item being faulty is 0.12.
    (a) Give the distribution of the faulty items among the 12 items.
    (b) Give the distribution of the faultless items among the 12 items.
    (c) What is the probability of having 4 faulty items?
    (d) What is the probability of having 3 faultless items?
    (e) What is the probability of having 2 faulty items or less?
    (f) What is the probability of having less than 3 faulty items?
    (g) What is the probability of having more than 8 faulty items?
    (h) What is the probability of having 10 faulty items or more?
    (i) What is the probability of the number of faulty events being between 4 and 6, both included?
    (j) Solve (c)-(i) questions with R software.
    (k) What is the mean number of faulty items.
    (l) Solve by R: which is the most probable number of faulty items?
    (m) Solve by R: how many items can we assure with a probability of at least 0.9?
72. 300 women eta 200 men live in a village.
    (a) We choose 15 personn radomly and with devolution. Give the distribution of the number of women.
    (b) What is the probability of the number of women beimg just proportional to the number of women in the village?
    (c) What is the meaning of the number of women in the latter question?
    (d) Give the probability of the number of women being 9 ± 2?
    (e) Interpret the results in questions (b), (c) and (d).
    (f) Give the probability of question (d), supposed you only know μ and σ and no that the number of women follows a binomial distribution.
73. In a given place the probability of raining is 0.4 and it's assumed total independence among different days. In order to construct a rooftop we need 7 days without rain. For how many day should we rent a crane to construct the rooftopwtih a 0.9 probability?
74. We sell packages of 20 pieces. One of our customers inspects all the pieces and at the end of the year if he finds one faulty pice or more into more than 10% of the packages, he will cacel the contract with us. Calculate the probability of producing faulty pieces for the purpose of keeping the contract.
75. In a flight 25% of tickets are cancelled or become vacnat. For a given flight we have 12 seats. How many tickets should we sell in order to be the probability of having overbooking at most 0.15?

3.1 Return periods in binomial distributions

76. The return period of having more than 200 mm rain in a day is 8 years. Which is more probable through 6 years: to have a rain of that magnitude of not to have it?
77. A bridge has been finished this year. The contractors think that a flood that would wreck the bridge should happen only once in 1000 years.
    (a) What is the probability of the bridge standing in the next 100 years?
    (b) For how many years will the bridge stand with a 0.95 probability?

3.2 Statistical testing related with binomial distribution

78. Normally 2% pf pieces are faulty in a factory. Among the last 10 pieces 2 faulty pieces have been found. Should we decide the production process is wrong? Significance level: 1%.
79. There are 20 questions in a test, each with 5 choices. A student has answered correctly 6 questions. Should we determine that he knew some of the questions or answered them randomly? Significance level: 10%.
80. 5 sellers (one of them is Peter) sold 15 machines in a exhibition. Peters sold 10 of them. Should we conclude that Peter is better than the other sellers? Significance level: 1%.
81. Last year there were 15 accicents in a road amd 10 of them happened on Sundays and bank holidays. Should we conclude accidents become more frequent on tohse days? Remark: there were 41 Sundays and bank hlidays last year. Significance level: 1%.
82. There are 4 workers in a factory and each of them produces 6 pieces a day. Among the 24 pieces produced in a day we found 6 faulty pieces and only one of them was proced bt the the elder worker. Sholud we decide he produces better that his colleagues? Significance level: 1%.

3.3 Geometric distribution. Negative binomial distribution.

83. The probability of a faulty piece being produced is 0.12.
    (a) Which is the distribution of the number of produced faultless pieces before the first produced faulty piece?
    (b) What is the probability of having 6 faultless pieces before the first faulty piece?
    (c) And the probability of having 2 faultless pieces or less?
    (d) What is the expected number of faultless pieces before the first faulty piece?
    (e) What is the probability of having 2 faulty pieces before the first faultless piece?
    (f) What is the expected number of faulty pieces before the first faultless piece?
84. The probability of a faulty piece being produced is 0.12.
    (a) Which is the distribution of the number of produced faultless pieces before the third produced faulty piece?
    (b) What is the probability of having 6 faultless pieces before the third faulty piece?
    (c) And the probability of having 2 faultless pieces or less?
    (d) What is the expected number of faultless pieces before the third faulty piece?
    (e) What is the probability of having 2 faulty pieces before the fourth faultless piece?
    (f) What is the expected number of faulty pieces before the fourth faultless piece?

4 Poisson distribution

85. In a machine 4.2 failure happen on average randomly.
    (a) What is the probability of having no failures in 2 hours?
    (b) Calculate the probability of having respectively 2, 3, 4 eta 28 failures in one hour. Give the interpretation of the results.
    (c) We must make a task through 22 minutes. What is the probability of completing the task without any failure?
    (d) What is the probability of having 3 failures or less in 2 hours?
    (e) What is the probability of having 8 failures or more in 2 hours?
    (f) What is the probability of having 10 to 15 failures in 2 hours, both included?
    (g) What is the probability of the time between 2 failures being al least 2 hours?
    (h) How many failures will be in an hour with a 0.9 probability? And with a 0.99 probability? Give the answer from both the seller's and the production manager's point of view.
86. Hepatitis-C infected patients admitted in a hospital are 8.6 a week on average, randomly and independently one from another. We need one dose per patient to treat the illness.
    (a) How many doses must we prepare in order to treat all the patients with a 0.9 probability?

(b) Idem, with a 0.99 probability.

(c) In two weeks 24 patients were treated. Should we conclude that the prevalence of the illness has grown? Significance level: 0.01.

(d) As health managers are worried about the growth of the number of patients, thay have prepared a prevention campaign. Following the campaign, in the next 4 weeks, there were 14 patients. Should we state that campaign was successful? Significance level: 0.01.

87. On average 6 customers enter a shop, randomly and with independence from one to another. Each customer needs 2 minutes to pay. How many cashiers should we hire in order to have a 0.9 probability of not having a queue?

4.1 Poisson distribution as limit of the binomial distribution

88. In a industrial process, the probability of producing a faulty item is 0.0022. We a have bought a batch of 4000 items, baina we'll return it whenever we find more han 6 faulty items.
    (a) What is the probability of returning the batch? Calculate both by means of the binomial distribution and the Poisson distribution.
    (b) Among 1000 items, how many faukty items can we ensure with a 0.9 probability?

89. The probability of a worker having a accident in a year is 0.0012.
    (a) On average, how many accidents happen among 1000 workers?
    (b) In a factory there are 3400 workers and there were 7 accidents last year. Shoud we conclude that prevention is not as successful as in the other factories? Significance level: 0.05.

4.2 Return periods in Poisson distributions

90. The return period of having more than 200 mm rain in a day is 8 years. Which is more probable through 6 years: to have a rain of that magnitude of not to have it?

91. A bridge has been finished this year. The contractors think that a flood that would wreck the bridge should happen only once in 1000 years.
    (a) What is the probability of the bridge standing in the next 100 years?
    (b) For how many years will the bridge stand with a 0.95 probability?

5 Exponential distribution

92. A machine stops twice an hour, randomly adn independently.
    (a) What is the probability of the time between consecutive stops being larger than 20 minutes? Calculate both in terms of minutes and of hours.
    (b) What is the probability of the time to the next stop being larger than 20 minutes?
    (c) What is the probability of that time being 20 to 30 minutes?

93. The mean time between failures in a machine is 10 minutes, and those failues happen randomly and independently.
    (a) What is the probability of the time between consecutive stops being larger than 5 minutes?
    (b) Given that it is 60 minutes that the last failure happened, what is the probability of the time to the next stop being larger than 5 minutes?
    (c) Interpret the previous results.

94. On average 4 customers visit a website in an hour. The website shuts down for 40 minutes. What is the probability of losing a customer in that interval? Caculate it both by means of the number of customers and the time between customers.

95. Customers come randomly and independently to a queue, with a mean time till the next customer of 5 minutes. It turns out that the next customer has come in 20 minutes. Should we reject the 5 minutes average? Significance level: %1.

6 Uniform distribution

96. We just know the price growth of an item will be 0 to 10 in the next year.
    (a) What is the probability of the price gowth being bigger than 6 (give the answer both by reasoning and calculation).
    (b) Which is the expected price growth? Give the answer both by reasoning and calculation.
    (c) Which is the variance of the price growth?

7 Normal distribution

97. Tenperature into a freezer follows a $N(0,1)$ standard normal distribution. Calculate these probabilities both by means of the statistical table and R programming language commands:
    (a) $P[Z < 1.42]$
    (b) $P[Z < 3.98]$
    (c) $P[Z < 5.6]$
    (d) $P[Z > 2.75]$
    (e) $P[Z < -0.68]$
    (f) $P[Z > -1.02]$
    (g) $P[0.48 < Z < 1.92]$
    (h) $P[-1.24 < Z < -0.98]$
    (i) $P[-2.19 < Z < 0.55]$
    (j) Give the value in the standard normal distribution that accumulates below it a 95% probability.
    (k) Give the value in the standard normal distribution that accumulates beyond it a 20% probability.
    (l) Give the value in the standard normal distribution that accumulates below it a 10% probability.
98. Daily production (E) in a factory follows a $N(68kg, 4kg)$ distribution. Calculate these probabilities:
    (a) $P[E < 77]$
    (b) $P[E > 62]$
    (c) $P[64 < E < 72]$
    (d) Calculate the maximum production in 99% of the days.
    (e) How much production can we ensure with a 0.9 probability?
    (f) Known that production in one day has been 54 kg, should we conclude that production has dropped?
99. Grade in a test follows the normal distribution, with an average of 110 poins and a deviation of 5 points. In order to pass the test, the students must achieve 120 points.
    (a) What is the probability of passing the test? Zenbat da proba gainditzeko probabilitatea?
    (b) We exclude students below 95 points. How many students do we exclude as a percentage?
    (c) Calculate the number of points to exclude 30% of the students.
    (d) Give a 90% symmetric confidence interval for the number of points.
100. Daily production in a factory follows the normal distribution, with a 8 ton average and a deviation of 1 ton.
    (a) What is the probability of producing less than 30 ton in 4 days?
    (b) How much production can we ensure with a 90% probability?
    (c) How many days do we need to fulfill a batch of 60 tons, if we want the probability of not fulfilling the deadline to be 15%?
    (d) Idem, with a 1% probability.
    (e) Taxes for producing will be 40.000€, with a tax allowance of 1.000€per tom. What is the probability of paying for taxes more than 10.000€?

101. 12 people get into an elevator, and the maximum weight allowed is 900 kg. Which should the average weight of a person in order to not exceeding it being 0.1? Remark: standard deviation of the weight of a person is 10 kg.

7.1 De Moivre-Laplace theorem

102. In a production process the probability of a faulty item is 0.05. We a 100 item batch.
    (a) What is the probability of having 80 faulty items or less?
    (b) How many faulty items are expected to be?
    (c) What is the probability of having exactly the number of faulty items expected?
    (d) How many faulty items can we ensure with a probability of 90%?
103. 1000 people are attending a recruitment process. The probability of passing the first exam is 0.6. How many chairs should we get if we want the probability of all the candidates in the second exam being sitting to be at least 0.99?
104. Daily production follows a N(100 kg, 10 kg) distribution. What is the probability of having in year (365 days) at least 67 dayes with a production of at most 110 kg?

7.2 Normal approximation of the Poisson distribution

105. The expected daily number of failures in a machine is 0.16, following the Poisson distribution. We need one piece to fix every failure.
    (a) How many pieces do we need in order to fix all the failures in a year with a 0.99 probability?
    (b) With 100 pieces, for how many days can we ensure that we have enough pieces to fix all the failures with a 0.95 probability?
106. Following a Poisson process, 1.4 batteries are exhausted on average each day in a machine.
    (a) Througout 40 days, what is the probability of exhausting more than 50 batteries?
    (b) How many batteries should we get in order to have enough energy for 80 days?

7.3 Central limit theorem

107. The share price possible daly increments are +1, 0 eta -1€ with respectively 0.2, 0.5 and 0.3 probabilities.
    (a) After 100 days, what is the probability of not losing?
    (b) How much money should we have after 100 dayes in order to be able to pay the losses with a 0.99 probability?
108. Substance consumed in a day follows an uniform distribution, in a 5 to 10 liters interval.
    (a) What is the probability of consuming less than 420 liters in 40 days?
    (b) How many liters must we get for 40 days in order to being the probability of having enough substance 0.99?
    (c) Idem, if the consumption follows a U(0,15) distribution.
    (d) Provided we have 500 liters, for how many days do we have enough substance with a 0.99 probability?
109. 1.4 batteries are exhausted each day on average following the Poisson distribution.
    (a) Provided we have 40 batteries, for how many days do we have enough energy with a 0.99 probability?
    (b) How many batteries must we get provided we want the probability of having enough energy for 80 days being 0.98?
110. In a factory the daily average production is 146 units, with a standard deviation of 10 units, following an unknown distribution.
    (a) How much production can we guarantee for 30 days with a 0.99 probability?
    (b) Which term in days should we give provided we must produce an order of 5000 units with a 0.94 probability?
111. The number of failure in a machine may be 0, 1, 2 with 0.2, 0.5, 0.3 probabilities respectively. Following a maintenace plan, in 100 machines there have been 90 failures. May we conclude that the maintenance plan has been successful? Significance level: 2%.
112. Daily production in a factory follows a U(100,200) distribution, in kilograms. In 50 days, the average production has been 140 kg.

(a) Should we conclude that production has decresed? Significance level: 2%.

(b) Which is the production level in order to claim that average production has decreased? Significance level: 2%.

(c) Which is the production level in order to claim that average production has just changed? Significance level: 2%.

8 Statistical inference: validation

8.1 Goodness of fit: chi-square test

113. 60 customers tasted a recent marketed yoghourt and were asked about their favourite one. A, B, C and D labeled yoghourts were chosen 20, 14, 12 and 14 times respectively. Should we conclude that yoghourts are equally probable to be chosen? Significance level: 10%.

114. We collected throughout 100 days the daily number of failures in a machine:

Number of failures	0	1	2	3	> 3
Number of days	21	19	15	20	25

Can we conclude that the number of failures follows a Poisson distribution? Significance level: 10%.

115. Scores obtained by a group of students are given below (original data: 21, 46, ...):

Score	0-20	20-40	40-60	60-80	80-100
Number of students	2	14	34	38	12

Test whether data follow a normal distribution. Significance level: 10%.

116. A restaurant serves menus from Monday to Friday, at noon and evenings. The total number of menus served ari geven below:

Menus	Monday	Tuesday	Wednesday	Thursday	Friday	Total
Midday	38	45	38	58	40	219
Evening	18	31	26	30	36	141
Total	56	76	64	88	76	360

(a) Conclude whether number of menus served from Monday to Friday follows an uniform distribution, by means of the chi-square test. Significance level: 5%.

(b) Test whether menus served at noon are twice menus served at evening. Significance level: 5%.

117. Times in days till a failure in a machine are given below:

$26.2, 22.3, 33.5, 19.0, 24.7, 25.6, 26.2, 28.9, 27.6, 26.5, 27.1$

$32.4, 36.2, 34.1, 28.7, 26.5, 25.4, 23.4, 21.6, 22.0, 20.6, 30.2$

Setting < 19, $19-21$, $21-23$, ... intervals, test whether data follow a normal distribution, by means of the chi-square test, provided that we must previously estimate the mean and the standard deviation. Significance level: 10%.

118. Times in days till a failure in a machine are given below:

$26.2, 22.3, 33.5, 19.0, 24.7, 25.6, 26.2, 28.9, 27.6, 26.5, 27.1$

$32.4, 36.2, 34.1, 28.7, 26.5, 25.4, 23.4, 21.6, 22.0, 20.6, 30.2$

Setting $0-10$, $10-20$, $20-30$, $30-40$ eta > 40 intervals, test whether data follow a exponential distribution, by means of the chi-square test, provided that we must previously estimate the mean of the distribution. Significance level: 10%.

8.2 Testing randomness and independence: the runs test

119. Allegedly we have collected some random data about the production in a factory. The data are given below:

$34.6 - 26.2 - 31.0 - 28.7 - 29.5 - 33.4 - 35.6 - 27.2 - 30.8 - 28.9$

$36.1 - 27.5 - 34.5 - 29.7 - 35.6 - 32.8 - 28.8 - 32.3 - 27.2$

Apply the runs test to conclude if data were collected independently. Significance level: 5%.

120. We have compiled the share price's change in the last days:

$$\downarrow\uparrow\uparrow\downarrow\downarrow\downarrow\uparrow\downarrow\downarrow\uparrow\uparrow\downarrow\uparrow\uparrow$$

Should we conclude they happen randomly and independently? Significance level: 5%.

121. We have compiled producted pieces have been faulty or faultless in a production process:

$$OOOOOXXOXXXXOOOOXOOOOOOOOXXXOOXOXOOOOOOOXOOOOXXX$$

Test whether the process develops independently and drawn pieces provide a random sample, by means of the runs test. Significance level: 10%.

8.3 Testing homogeneity: Wilcoxon rank-sum test

122. We hace compiled califications laid down by two proofreaders in a exam:

Selektibitateko azterketa batean bi zuzentzailek jarritako kalifikazio zenbait jaso dira:

A proofreader	B proofreader
5.4	6.5
6.2	7.4
8.7	4.7
9.5	5.6
7.6	5.4
8.2	8.6
3.5	7.2
7.8	

Sould we conclude that both proofreaders set the califications in the same way? Significance level: 10%.

123. We have compiled some invoice amounts paid in a shop by sex:

$Men : 3, 3, 5, 6, 8, 10, 10, 11, 11, 12, 12, 12, 16, 19, 20$

$Women : 2, 7, 9, 11, 13, 13, 15, 17, 17, 18, 20, 21, 23, 24, 25, 25, 27, 32, 36, 39$

Test with those data whether men and women have the same buying behavior. Significance level: 10%.

124. Children with math comprehension problems were given a special training program last year. They carried a test before and after the program. The results are given below:

$Before : 22, 32, 43, 28, 27, 36$

$After : 25, 42, 50, 35, 35, 42$

Test whether the program has been successful. Significance level: 5%.

125. Data about number of movie tickets sold on Saturday and Sunday are given below:

Saturday : $126 - 91 - 68 - 122 - 113 - 137 - 111 - 86 - 100 - 82 - 96 - 121 - 97 - 95 - 89$

Sunday : $81 - 98 - 129 - 101 - 121 - 124 - 133 - 108 - 84 - 89 - 86 - 131$

Test if more tickets are sold on Sunday, by means of Wilcoxon rank-sum test. Significance level: 1%.

9 Sampling distributions

126. A population is composed by 1-2-3 elements. We draw a sample of size 3, with devolution. Give the sampling distribution of the arithmetic mean and interpret it.

10 Parametric tests

10.1 Tests for the mean

127. Sales in a shop follow a normal distribution. We suppose the standard deviation is 100. In a normal day sales are 1000 on average. Throughout 10 days, average sales have been 1100. Should we conclude that averages sales have increased? Solve both by the p-value and the critical region. Significance level: 5%.

128. We have drawn these data from a normal population: 22-26-24-24-25-23.

    (a) Test whether the population mean is 22. Significance level: 10%.

    (b) Without any other calculation, which other hypothesis should we accept or reject from the previous results?

    (c) Given 8 data, sample mean has been 28 and sample variance 2. Should we conclude that population mean has increased? Significance level: 1%.

129. It has been set that a null hypothesis establishes that a given flight is 300 minutes on average. We know that standard deviation is 40 minutes, baina we cannot establish the model for the data. Provided that sample mean is 310 minutes, test the given hypothesis by both the p-value and the critical region. Significance level: 5%.

130. We have drawn 50 data aobut time following an alleged exponential distribution. Normally mean time is 100 minutes. Set the test in order to conclude whether mean time has decreased. Significance level: 5%.

131. We have compiled water comsumption data from 80 families: $\overline{x} = 7$; $\sum x_i^2 = 4300$. The model for data is supposed to be a gamma named distribution, with unknown mean and variance.

    (a) Decide by means of the p-value if population mean may be 7.5 or bigger. Significance level: 1%.

    (b) Now decide if population mean may be 7.5 or less. Significance level: 1%.

132. A component's duration follows a normal distribution, with a standard deviation of 4 units. Quality specifications require mean duration to be al least 10 days. Sample mean from 9 data has been 8 units. Test whether the specification is met. Significance level: 1%.

133. A machine's output is 40 units on average, and allegedly follows a normal distribution. In order to have independent data, we have drawn 4 output data from different days: 38-39-35-36. What should we decide with those data about the average output? Significance level: 5%.

10.2 Tests for the proportion

134. In a market, 22% of the consumers buys our product. This year 44 consumers of 200 have declared they will follow to buy our product. With those data, draw a conclusion about the change in the proportion of consumers, by means of both the p-value and the critical region. Significance level: 5%.

135. Our supplier claims that at most 4% of pieces are faulty. We suspect that it's not true. In order to tess our suspicions, we draw randomly and inspect 200 pieces. With a 10% significance level, how many faulty items must we have in order to reject the suppier's claim? (Namely, calculate the critical value) Idem, with 5% and 1% significance levels. Build a table with the results and draw a conclusion.

10.3 Tests for the population variance.

136. We have drawn 10 data from an alleged normal distribution. Sample variance (without correction) is 36. Test whether the standard deviation in the population is 5 or less. Significance level: 0.01.

137. We have compiled the number of microorganisms in two pills, in hundred of millions: 2.2, 4.2. Quality specifications require that the variance of that number must be 2 or less. Should we decide that the given specification is fulfilled? Significance level: 0.10.

11 Confidence intervals

11.1 Confidence intervals about the mean

138. We have compiled the duration of a bus tour (in minutes)

$$12 - 11 - 13 - 16 - 18 - 20 - 10 - 14$$

Duration follows an alleged normal distribution, that should be validated by a goodness-of-fit test.

(a) Estimate the mean and the standard deviation. Are they good estimates?

(b) Give a 90% confidence interval about the population mean.

(c) Idem, with a 99% confidence level.

139. We have compiled sales number in a shop throughout 10 Mondays. Sample mean is 282 currency units and corrected standard deviation 32. A normal population is assumed.

(a) Why do we take only sales on Mondays?

(b) Give a 95% confidence interval about the population mean.

(c) If we had drawn 20 data, instead of 10, which would be the confidence interval? Interpret the new result.

(d) If standard deviation had been 44 with 10 data, which would be the new interval? Interpret the result.

140. We have compiled some outputs per hour in a machine:

$$56 - 44 - 48 - 52 - 60$$

(a) Assumed a normal distribution, give the 90% confidence interval about the population mean.

(b) Which 90% confidence interval should give the machine dealer?

(c) Which 95% confidence interval should give the production manager?

141. We a drawn a random sample of 86 data from a normal population. Sample mean is 144 units, and sample variance 121 units. Give the symmetrical 98% confidence interval.

142. We have collected daily outputs during 10 days, with these results:

$$\sum_{i=1}^{10} x_i = 108 \; ; \sum_{i=1}^{10} x_i^2 = 1234$$

(a) Give a point estimation and a 99% symmetrical confidence interval about the population mean.

(b) May we conclude that the average output is bigger than 10 units with a 90% confidence?

11.2 Confidence intervals about the proportion

143. 120 faulty items have been found in 1000 items.

(a) Give 90% ad 99% confidence intervals about the proportion.

(b) Which would be the confidence interval if 1200 faulty items had been fonud in 10.000 pieces? Interpret the new result in connection with the latter.

144. 16 faulty items have been found in 256 items.

(a) Give the error in the estimate of the proportion with a 80% confidence.

(b) Which interval should give the machine dealer with the same confidence?

(c) Which interval should give the customer in order to make a complaint with the same confidence?

145. We have compiled the amount of a given substance in some batchs of a product (mg/l):

22.0 – 17.7 – 23.0 – 22.7 – 21.5 – 20.6 – 16.2 – 28.7 – 27.4 – 15.6 – 28.4 – 20.1 – 17.1 – 19.5 – 18.6
24.0 – 22.6 – 22.9 – 24.5 – 25.4 – 24.4 – 20.7 – 14.6 – 18.3 – 23.6 – 20.5 – 24.8 – 22.9 – 23.4 – 25.2
24.6 – 24.4 – 26.0 – 26.8 – 21.2 – 13.4 – 20.2 – 17.4 – 25.4 – 23.6

When the amount ofsubstance is bigger than 25.5 mg/l, we reject the batch.

(a) Give an estimate of the proportion of rejected batchs and calculate the corresponding error with a 96% confidence.

(b) Give the corresponding symmetrical confidence interval.

(c) We want to claim the product has a good quality. Which confidence interval should give give for that with the same confidence.

146. In a election, we want to estimate the proportion of the voters for X party with a 90% confidence, with an error of $\pm 2\%$.

  (a) How many voters should we polled for that? Be cautious.

  (b) Give the new sampling size if the error switches to $\pm 1\%$? Be cautious.

  (c) With a $\pm 1\%$ error, if the number of voters for X party are eventually 2.000, give the corresponding confidence interval.

147. Give the sampling size for a 99% interval confidence interval for a proportion of faulty items with a 4% error, provided that in a pilot sample we got 22 faulty items in 100 pieces. Idem, with a 90% interval.