ADMISSION EXAM
VERSION B 5

Note:
Each question has only one correct answer
Each correct answer = 3 points
Duration of the exam = 90 minutes

1. **The following data were recorded at a production unit in 2019-2023:**

Year	2019	2020	2021	2022	2023
Production value (thousand euros)	45	55	75	80	95

The adjustment function by the mechanical method based on the average absolute change is:

a) $Y_t = 45*1.148^{(t-1)}$
b) $Y_t = 70+8.5*t$
c) $Y_t = 45+12.5*(t-1)$
d) None of the above

2. **For a time series, the trend was estimated by the analytical method, resulting in the following equation: $\hat{Y}_t = 100 + 0.01 \cdot t$. Specify which of the following statements is correct:**
   a) series is stationary and follows an ARMA process
   b) the cyclic component is equal to 100
   c) the time series shows an upward trend
   d) the seasonal component is 0.01

3. **When modelling the relationship between two variables using the linear regression equation $y = a + bx + \varepsilon$:**
   a) The coefficient ***b*** measures the strength of the link between the variables
   b) The coefficient ***b*** is always positive
   c) the coefficient ***b*** has the same sign as the linear correlation coefficient
   d) The coefficient ***a*** has the same sign as the linear correlation coefficient

4. **The definition of the sampling population involves:**
   a) Parameter estimation
   b) Setting the Starter Sample
   c) Population adjustment based on validity and eligibility proportions
   d) Characterization of the degree of variation in the population according to auxiliary variables

5. **What does the mode represent in a set of numbers?**
   a) the middle number
   b) the average number
   c) the number the appears most often
   d) the highest number

6. **What is the name of the indicator that divides an ordered dataset into two equal parts?**
   a) Median
   b) Quartile 1
   c) Quartile 3
   d) Mode

7. **Over five years, a team of statisticians carried out 335 projects, as follows:**

Year	2019	2020	2021	2022	2023
Number of projects	50	60	65	74	86

The average annual percentage change in the number of projects is:
   a) 114.52%
   b) 9 projects
   c) +14.52%
   d) 67 projects

8. **For two statistical variables, a Pearson correlation coefficient of -0.22 (r=-0.22) was determined. It can be stated that:**
   a) There is a direct relationship between the two variables
   b) There is a strong relationship between the two variables
   c) There is a weak relationship between the two variables
   d) The relationship between the two variables is of medium intensity

9. **The relationship between two statistical variables was modelled by the linear regression equation $\hat{y}_i = 7.13 + 0.14x_i$. Knowing that $\sum_i (x_i - \bar{x})^2 = 800$ and that $\sum_i (y_i - \bar{y})^2 = 32$ specify which variant of the Pearson linear correlation coefficient (r) is true:**
   a) $r = 0.98$
   b) $r = 0.70$
   c) $r = 0.04$
   d) r cannot be estimated

10. Which of the following statements, regarding objective imputation, is true:

a) The advantage of objective imputation is that the value of the imputation harmonizes with the rest of the individual's records.
b) The method is based on a system of 2 equations with 2 unknowns
c) The method is based on the correlation coefficient between the variable for which the non-responses are treated and an auxiliary variable
d) To increase the efficiency of the method it can be combined with random imputation

11. The mortality rate specific to a particular cause of death is calculated as:

a) the ratio of the number of deaths due to that cause to the average total population, multiplied by 100,000
b) the ratio of the number of deaths due to that cause to the female population, multiplied by 1,000
c) the ratio of the number of deaths due to that cause to the male population, multiplied by 100
d) the ratio of the number of deaths due to that cause to the urban population, multiplied by 100,000

12. In a data set [120, 140, 160, 180, 200, 220], the mean and median value are:

a) 170 and 170
b) 160 and 170
c) 170 and 160
d) 175 and 180

13. The following data were recorded at one company, in two consecutive periods:

Product	Quantities sold (pieces)		Unit price (lei/MU)	
	Base period	Current period	Base period	Current period
A	50	60	80	88
B	60	72	90	99

According to the data presented in the table, the group index of quantity is:

a) 110.0%
b) 120.0%
c) 1,880 lei
d) 132.0%

14. In a multiple linear regression model, the coefficients and p-values for the independent variables are:

Variable	Coefficient	p-value
X_1	2.5	0.03
X_2	-1.2	0.15
X_3	0.8	0.001

Which of the independent variables are significant at the 5% significance level?

a) X_1 and X_3
b) X_1 and X_2
c) X_2 and X_3
d) All variables are significant

15. **Which of the following statements regarding the simple random sampling is true:**
    a) To calculate the sample volume, the average of the dispersions of the variables for which we have information from administrative sources is used
    b) It is a survey design recommended for homogeneous populations
    c) The units in the target population can only be persons (not economic agents)
    d) The sample volume shall be directly proportional to the maximum permissible limit error

16. **The owner of a company is interested in the percentage of its consumers that are satisfied with the company's services. In a random sample of 225 consumers 180 said that they are satisfied with the company's services. The confidence interval for all the customers is (use a 95.45% confidence level, z = 2)**
    a) (74%; 86%)
    b) (77%; 83%)
    c) (77%; 86%)
    d) (75%; 85%)

17. **If the female fertility curve peaks within the 20-24 age group, then the type of fertility in that population is:**
    a) early
    b) intermediate
    c) late
    d) mixed

18. **The graph displaying the minimum value, quartile 1, median, quartile 3 and maximum value is:**
    a) Cumulative frequency curve
    b) Histogram
    c) Box-plot
    d) Scatter plot

19. **The percentage change in a company's sale of a given product if the quantity sold increased by 3.4 % and the price increased by 6.5 % in 2023 compared to 2022 is:**
    a) 10.44%
    b) 9.9%
    c) 3.1%
    d) 11.58%

20. **A multiple linear regression model is used to study the influence of temperature X_1 and humidity X_2 on wheat production Y. The estimated model is: $\hat{Y} = 200 + 3X_1 - 2X_2$. Coefficient of determination R^2 had the value 0.78. Based on this information, select the correct statement.**
    a) Each additional degree Celsius is associated with an average increase of 3 units in wheat production
    b) Each additional percentage of moisture is associated with an average decrease of 2 units in wheat production
    c) The model explains 78% of the variation in wheat production
    d) All the above are correct

21. **The estimation of the sampling volume in the case of a stratified survey, under the conditions of a non-repeated selection, is carried out using the relation:**
    a) $$n = \frac{z^2 \cdot \sigma^2}{\Delta_{\bar{x}}^2}$$
    b) $$n = \frac{z^2\sigma^2}{\Delta_{\bar{x}}^2 + \frac{z^2\sigma^2}{N}}$$
    c) $$n = \frac{z^2 N_i \sigma_i^2}{\sum_{i=1}^{k} N_i \sigma_i^2}$$ where k is the number of strata
    d) $$n = \frac{z^2\overline{\sigma}^2}{\Delta_{\bar{x}}^2 + \frac{z^2\overline{\sigma}^2}{N}}$$

22. **Among the 100,000 (infinite population) registered customers of an online shop for pet products, a research is to be carried out on the basis of a statistical survey. The average value of an order is 333 lei with an average square deviation of 80 lei. Determine the sample volume required to be extracted for a 95,45% confidence interval ($z_{(1-\alpha/2)}=2$) and a maximum permissible limit error of ±3% of the average value of an order. The selection procedure used is simple random. The sample volume is:**
    a) 126
    b) 1199
    c) 2038
    d) 256

23. **In the case of the K-Means algorithm, is it possible for the assignment of observations to classes to remain the same between two successive iterations?**
    a) Yes, at any time
    b) Yes, in the case of the final iteration
    c) Never
    d) None of the above

24. **The standard deviation is computed in a dataset as:**
    a) Square root of variance
    b) The sum of the squares of the differences between each value and the arithmetic mean
    c) The difference between the maximum value and the minimum value
    d) The sum of the absolute differences between each value and the arithmetic mean

25. **In a time series, the absolute average increase (the absolute mean change) is determined as:**
    a) simple arithmetic mean of fixed-base changes
    b) simple arithmetic mean of absolute chain-based changes
    c) simple harmonic mean of chain-based changes
    d) weighted arithmetic mean of fixed-base changes

26. **For the simple linear regression model, multicollinearity has the consequence:**
    a) Biased estimators
    b) Autocorrelation of errors
    c) There is no question of multicollinearity in simple linear regression
    d) Increasing the confidence interval for predictions

27. **Which of the following is an advantage of a survey:**
    a) More in-depth characterization of the phenomena studied due to the possibility of including a greater number of features than in the total observation program
    b) Less representativeness errors than in the case of a total research
    c) A statistical survey is carried out in a single phase, that of descriptive statistical analysis
    d) The proportion of response is higher than in the case of exhaustive research

28. **By "process of demographic aging" is understood:**
    a) the change of the population structure by age, in favor of older age groups, as a firm and long-term trend
    b) the change of the population structure by age, in favor of adult age groups, as a firm and long-term trend
    c) a process involving structural changes where the base of the age pyramid broadens
    d) a process involving shifts in the age pyramid resulting in a decrease of the proportion of elderly

29. The number of eigenvalues in a principal component analysis is:

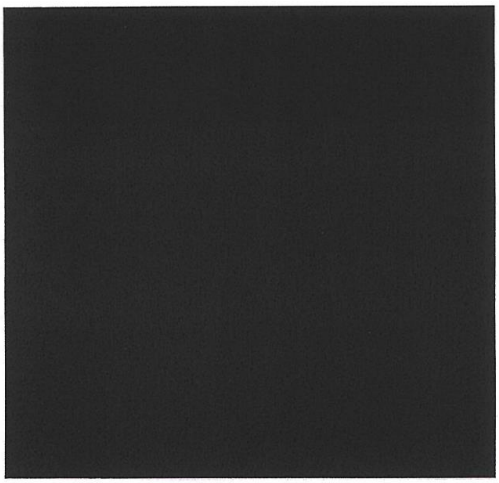
a) 1
b) 100
c) The number of variables in the analysis
d) We cannot say without additional information

30. The number of hours allocated monthly to individual study by 135 students enrolled in the bachelor's program "Statistics" is:

Hours	40 up to 44	44 up to 48	48 up to 52	52 up to 56	56 up to 60
Number of students	10	30	55	30	10

The standard deviation and the coefficient of variation are:

a) 4.07 hours and 8.15%
b) 8.59% and 4.28 hours
c) 18.36 and 4.28 hours
d) 20 and 4.28 hours



Cod Grilă

1	2	3	4	5	6
○	○	○	○	●	○

Disciplină

- ○ AP1
- ○ BT1
- ○ BT2
- ○ CSIE 1
- ○ CSIE 2
- ○ CSIE 3
- ○ CSIE 4
- ● CSIE 5
- ○ CIG1
- ○ CIG2
- ○ DA1
- ○ ETA1
- ○ EAM1
- ○ FIN1
- ○ MAN1
- ○ MK1
- ○ MK2
- ○ REI1
- ○ REI2
- ○ REI5
- ○ MD

30 x 3p = 90 p

Nr.	A	B	C	D
1	○	○	●	○
2	○	○	●	○
3	○	○	●	○
4	○	○	○	●
5	○	○	●	○
6	●	○	○	○
7	○	○	●	○
8	○	○	●	○
9	○	●	○	○
10	●	○	○	○
11	●	○	○	○
12	●	○	○	○
13	○	●	○	○
14	●	○	○	○
15	○	●	○	○
16	●	○	○	○
17	●	○	○	○
18	○	○	●	○
19	●	○	○	○
20	○	○	○	●
21	○	○	○	●
22	○	○	○	●
23	○	●	○	○
24	●	○	○	○
25	○	●	○	○
26	○	○	●	○
27	●	○	○	○
28	●	○	○	○
29	○	○	●	○
30	●	○	○	○
31	○	○	○	○
32	○	○	○	○
33	○	○	○	○
34	○	○	○	○
35	○	○	○	○