1. (Question 3 of Option A, Selectividad Exam 2011/2012) In a court of access test to University were examined 80 students from school A, 70 students from school B and 50 students from school C. The test has been passed by 80% of students from school A, 90% of students from school B and 82% of students from school C.

   a. What is the probability that a randomly chosen student has passed the test?
   b. A randomly selected student has not passed the test, what is the probability that this student belongs to school B?

2. (Question 4 of Option A, Selectividad Exam 2011/2012) It is assumed that students’ weight in kilograms of a Primary School the first day of the course can be approximated by a normal distribution random variable with standard deviation equal to 2.8 kg. A simple random sample of 8 students of that school provides the following results (in kg):

   26 27.5 31 28 25.5 30.5 32 31.5

   a. Determine a confidence interval at 90% level for the students’ average weight of that school the first day of the course.
   b. Determine the minimum sample size required for the difference’s absolute value between the sample mean and the population average is less than or equal to 0.9 kg at 97% confidence level.

3. (Question 3 of Option B, Selectividad Exam 2011/2012) Let A and B be two events in a random experiment such that

   \[ P(A \cap B) = 0.1 \quad P(\bar{A} \cap \bar{B}) = 0.6 \quad P(A | B) = 0.5 \]

   Compute:

   a. \( P(B) \)
   b. \( P(A \cup B) \)
   c. \( P(A) \)
   d. \( P(B | \bar{A}) \)

   Note: \( S \) denotes the complementary event of an event \( S \). \( P(S | T) \) denotes the event’s conditional probability \( S \) to the event \( T \).

4. Question 4 of Option B, Selectividad Exam 2011/2012) It is assumed that the expenditure made by individuals in a given population on Christmas gifts can be approximated by a random variable with normal distribution with mean \( \mu \) and standard deviation \( \sigma = 45 \) euros.

   a. It is taken a simple random sample, and it is obtained the confidence interval \((251.6, 271.2)\) for \( \mu \), with a confidence level of 95%. Compute the sample mean and the chosen sample size
   b. It is taken a simple random sample of size 64 to estimate \( \mu \). Compute the maximum error for this estimate with a confidence level of 90%.

5. The following table give the joint probability distribution between employment status and college graduation among those either employed or looking for work (unemployed) in the working age in U.S on 2008.

<table>
<thead>
<tr>
<th>Joint Distribution of Employment Status and College of population older than 25 years old in U.S, 2008</th>
<th>Unemployed ((Y = 0))</th>
<th>Employed ((Y = 1))</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non college grads ((X = 0))</td>
<td>0.037</td>
<td>0.622</td>
<td>0.659</td>
</tr>
<tr>
<td>College grads ((X = 1))</td>
<td>0.009</td>
<td>0.332</td>
<td>0.341</td>
</tr>
<tr>
<td>Total</td>
<td>0.046</td>
<td>0.954</td>
<td>1.000</td>
</tr>
</tbody>
</table>
9. A survey of 1055 registered voters is conducted, and the voters are asked to choose between candidate A and candidate B. Let \( \hat{p} \) denote the fraction of voters who prefer candidate A in the population and in the sample, respectively.

a. You are interested in the competing hypotheses \( H_0 : p = 0.5 \) and \( H_1 : p \neq 0.5 \). Suppose that you decide to reject \( H_0 \) if \( |\hat{p} - 0.5| > 0.02 \).

i. What is the size of this test?

ii. Compute the power of this test if \( p = 0.53 \).

b. In the survey \( \hat{p} = 0.54 \).

i. Test \( H_0 : p = 0.5 \) versus \( H_1 : p \neq 0.5 \), using a 5% significance level.

ii. Test \( H_0 : p = 0.5 \) versus \( H_1 : p > 0.5 \), using a 5% significance level.

6. Compute the following probabilities:

a. If \( Y \) is distributed \( N(50, 25) \), find \( \Pr(40 \leq Y \leq 52) \).

b. If \( Y \) is distributed \( t_{15} \), find \( \Pr(Y \geq 1.75) \).

c. If \( Y \) is distributed \( t_{90} \), find \( \Pr(-1.99 \leq Y \leq 1.99) \).

d. If \( Y \) is distributed \( N(0, 1) \), find \( \Pr(-1.99 \leq Y \leq 1.99) \).

e. Are educational achievement and employment status independent?

f. Compute \( E(Y) \).

g. The unemployment rate is the fraction of the labor force that is unemployed. Calculate the unemployment rate.

h. Compute \( E(Y \mid X = 1) \) and \( E(Y \mid X = 0) \).

i. Calculate the unemployment rate for college graduates and non-college graduates.

j. A randomly selected member of this population reports being unemployed. What is the probability that the worker is a college graduate? and a non-college graduate?

7. Let \( \{Y_i\}_{i=1}^n \) independent and identically Bernoulli distributed random variables with mean 0.4.

a. Use the central limit theorem to compute approximations for

i. \( \Pr(\bar{Y} > 0.43) \) with \( n = 100 \).

ii. \( \Pr(\bar{Y} \leq 0.37) \) with \( n = 400 \).

b. How large would the sample size need to be to ensure that \( \Pr(0.39 \leq \bar{Y} \leq 0.41) \geq 0.95 \)? (Use the central limit theorem to compute an approximate answer).

8. Suppose that \( \{Y_i\}_{i=1}^n \) are random variables with identical mean \( \mu_Y \), common variance \( \sigma_Y^2 \) and the same correlation \( \rho \) for each pair, that is, the correlation between \( Y_i \) and \( Y_j \) is \( \rho \) for all \( i \neq j \).

a. Show that \( \text{Cov}(Y_i, Y_j) = \rho \sigma_Y^2 \) for \( i \neq j \).

b. Suppose that \( n = 2 \). Show that \( E(\bar{Y}) = \mu_Y \) and \( \text{Var}(\bar{Y}) = \sigma_Y^2 \left( 1 + \rho \right) / 2 \).

c. For every \( n \geq 2 \), prove that \( E(\bar{Y}) = \mu_Y \) and \( \text{Var}(\bar{Y}) = \sigma_Y^2 \left( 1 + (n - 1) \rho \right) / n \).

d. For \( n \) really large, prove that \( \text{Var}(\bar{Y}) \approx \rho \sigma_Y^2 \).

9. A survey of 1055 registered voters is conducted, and the voters are asked to choose between candidate A and candidate B. Let \( p \) and \( \hat{p} \) denote the fraction of voters who prefer candidate A in the population and in the sample, respectively.

a. You are interested in the competing hypotheses \( H_0 : p = 0.5 \) and \( H_1 : p \neq 0.5 \). Suppose that you decide to reject \( H_0 \) if \( |\hat{p} - 0.5| > 0.02 \).

i. What is the size of this test?

ii. Compute the power of this test if \( p = 0.53 \).
iii. Construct a 95% confidence interval for \( p \).
iv. Construct a 99% confidence interval for \( p \).
v. Construct a 50% confidence interval for \( p \).

c. Suppose that the survey is carried out 20 times using independently selected voters in each survey. For each of these 20 surveys a 95% confidence interval for \( p \) is constructed
i. What is the probability that the true value of \( p \) is contained in all 20 of these confidence intervals?
ii. How many of these confidence intervals do you expect to contain the true value of \( p \)?

d. In survey jargon, the "margin of error" is \( 1.96 \times \text{Standard Error}(\hat{p}) \); that is, the half of the length of 95% confidence interval. Suppose you wanted to design a survey that had a margin of error of at most 1%. That is, you wanted \( \Pr (|\hat{p} - p| > 0.01) \leq 0.05 \). How large should the sample size \( n \) be if the survey uses simple random sampling?

10. To investigate possible gender discrimination in a firm, a sample of 100 men and 64 women with similar job descriptions were selected at random. A summary of the resulting monthly salaries follows:

<table>
<thead>
<tr>
<th></th>
<th>Average Salary (( \bar{Y} ))</th>
<th>Standard Deviation (( s_Y ))</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>3.100$</td>
<td>200$</td>
<td>100</td>
</tr>
<tr>
<td>Women</td>
<td>2.900$</td>
<td>320$</td>
<td>64</td>
</tr>
</tbody>
</table>

a. What do these data suggest about wage differences in the firm? Do they represent statistically significant evidence that wages of men and women are different? (To answer this question, first state the null and alternative hypothesis; second, compute the relevant \( t \) statistic; third, compute the \( p \)-value associated with the \( t \) statistic; and finally use the \( p \)-value to answer the question.)

b. Do these data suggest that the firm is guilty of gender discrimination in its compensation policies? Explain.

ANSWERS: 1. a) 84%; b) 21.87%.
2. a) (27'4,30'6); b) \( n \geq 46 \).
3. a) 0.2; b) 0.4; c) 0.3; d) 0.8571.
4. a) \( \bar{X} = 261.4\€, n = 81 \); b) \( \epsilon_{\text{max}} = 9.3\€ \).
5. a) 0.954; b) 0.046; c) \( E(Y|X = 1) = 0.974, E(Y|X = 0) = 0.944 \); d) 0.026 y 0.056; e) 0.196 y 0.804; f) No, ex. \( \Pr(X = 0|Y = 0) = 0.804 \neq 0.659 = \Pr(X = 0) \).
6. a) 0.6326; b) 0.05; c) 0.950; d) 0.950; e) \( t_{\infty} \stackrel{d}{=} N(0,1) \) and 20 g.l. is a big number; f) 0.90; g) 0.10; h) 0.05; i) 0.05; j) \( F_{10,\infty} \stackrel{d}{=} \chi^2_{10} \).
7. a) i) 0.27, ii) 0.11; b) \( n \geq 9220 \).
8. a) i) 0.19; ii) 0.74; b) i) \( t = 2.61 \), with \( \Pr(|N(0,1)| > 2.61) = 0.0094 \) and reject to 5%; ii) \( \Pr(N(0,1) > 2.61) = 0.0047 \) and reject to 5%; iii) (0.51,0.57), iv) (0.5,0.58), v) (0.53,0.55); c) i) 0.36, ii) The 95% of the 20 confidence intervals d) If \( n > 9604 \), the margin of error is less than 0.01 for any \( p \).