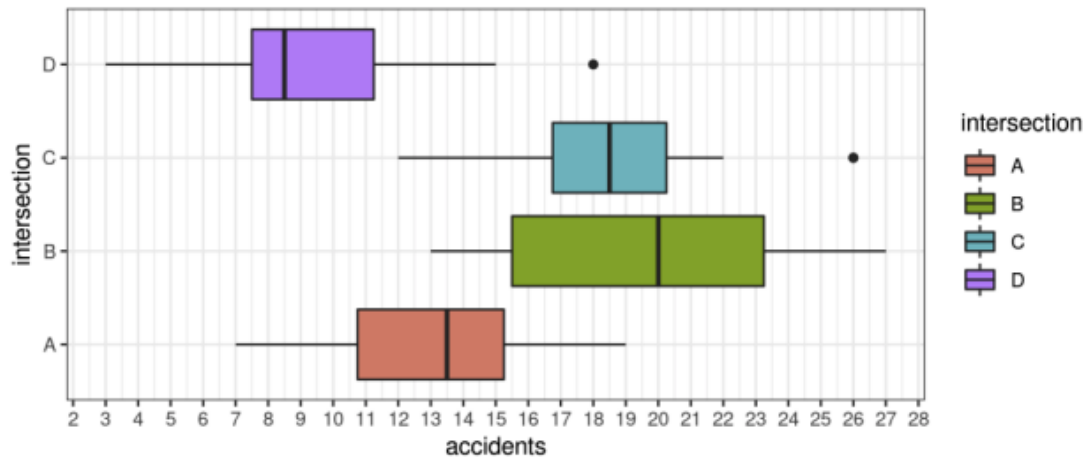


## Exam-30.01.2023

1. Suppose the null hypothesis  $H_0 : p = 0.4$ , and the power of the test for the alternative hypothesis  $H_A: p=0.35$  is 0.75. Which of the following is a valid conclusion?
  - (a) If the null hypothesis is false, the probability of failing to reject it is 0.6
  - (b) If the alternative hypothesis is true, the probability of failing to reject the null is hypothesis 0.25.
  - (c) The probability of committing a Type II error is 0.65.
  - (d) The probability of committing a Type I error is 0.05.
2. Let  $X \text{ Poi}(2)$  and  $Y \text{ Bin}(8,0.5)$  be two random variables with the correlation  $2 \text{ Corr}(X, Y) 0.4$ . Compute  $\text{Cov}(X, Y + 3)$ .
  - (a) -2.2
  - (b) -1.4
  - (c) 0.8
  - (d) 1.6
3. We toss two fair coins simultaneously and independently. If the outcomes of the two coin tosses are the same, we win, otherwise, we lose. Let  $A$  be the event that the first coin comes up heads,  $B$  be the event that the second coin comes up heads and be the event that we win. Which one of the following statements is true?
  - (a) The probability of winning is  $3/4$ .
  - (b) Events  $A$  and  $B$  are not independent.
  - (c) Events  $B$  and  $C$  are not independent.
  - (d) Events  $A$  and  $C$  are independent.
4. Two classes take the same exam. Suppose a certain score is at the 40th percentile for the first class and at the 80th percentile for the second class. Which of the following is the most reasonable conclusion?
  - (a) Students in the second class generally scored higher than students in the first class.
  - (b) A score at the 50th percentile for the first class is at the 90th percentile for the second class.
  - (c) One of the classes has twice the number of students as the other
  - (d) Students in the first class generally scored higher than students in the second class.

5. Consider the two sets  $X = (10, 30, 45, 50, 55, 70, 90)$  and  $Y = (10, 30, 35, 50, 65, 70, 90)$ . Which one of the following answers is false?
- The sets have identical ranges.
  - The sets have identical means.
  - None of the rest are false.
  - The sets have identical medians.
6. Data on the number of yearly accidents were collected from four intersections (A-D) over a 20 year period and are presented below. Which one of the following statements is false?



- During at least 5 years, fewer than 10 accidents occurred at intersection A.
  - During at least 75% of years, intersection C had more accidents than the lowest 75% of years at intersection A.
  - The minimum accident total at intersection C was higher than the number of accidents observed at intersection D in 75% of years.
  - The minimum number accidents that occurred in a single intersection was 3.
7. A study is to be performed to estimate the proportion of voters who believe the economy is "heading in the right direction." Which of the following pairs of sample size  $n$  and population proportion  $p$  will result in the smallest variance for the sampling distribution of  $p$ ?
- $n = 1000$  and  $p = 0.1$
  - $n = 100$  and  $p = 0.99$
  - $n = 1000$  and  $p = 0.5$
  - $n = 100$  and  $p = 0.1$

8. Which one of the following is an incorrect statement?
- (a) The sampling distribution of  $\bar{a}$  has standard deviation  $\sigma/\sqrt{n}$  even if the population is not normally distributed.
  - (b) The sampling distribution of  $\bar{2}$  has mean equal to the population mean even if the population is not normally distributed.
  - (c) When  $n$  is large, the sampling distribution of  $\bar{a}$  is approximately normal even if the population is not normally distributed.
  - (d) The larger the value of the sample size  $n$ , the closer the standard deviation of the sampling distribution of  $\bar{a}$  is to the standard deviation of the population.
9. Let  $X$  be a random variable with a Poisson distribution. If it holds  $P(X = 1) = P(X = 3)$  then the expectation  $EX$  equals.
- (a)  $\sqrt{6}$
  - (b)  $\sqrt{3}$
  - (c) 3
  - (d) 6
10. The income per household in a certain country is assumed to be normally distributed with the mean 9500 Euro and standard deviation of 1750 Euro. The middle 95% of incomes (in Euro) are between what two values?
- (a) 6076 and 12930
  - (b) 8049 and 10951
  - (c) 5422 and 13578
  - (d) 6621 and 12379
11. Which of the following statements about t-distribution are true?
- I Like normal t-distributions are always symmetric.
  - II The smaller the number of degrees of freedom, the closer the curve is to the normal curve.
  - III Twenty degrees of freedom gives the normal curve.
- (a) I only
  - (b) I and III
  - (c) III only
  - (d) I and II

12. In general, how does halving the sample size change the confidence interval size?
- (a) Doubles the interval size
  - (b) Divides the interval size by  $\sqrt{2}$
  - (c) Multiplies the interval size by  $\sqrt{2}$
  - (d) Halves the interval size
13. For statistical test of significance level  $\alpha$  it holds
- (a) the rejection region does not depend on  $\alpha$
  - (b) the rejection area shrinks when  $\alpha$  is increased
  - (c) the rejection area depends on the distribution or the test statistic under the null hypothesis
  - (d) rejection at level  $\alpha$  implies rejection at level  $\alpha/2$
14. Let  $X \approx N(-1, 4)$  Express the probability

$$P(X^2 + 2X \leq 0)$$

in terms of the cumulative distribution function  $\Phi$  of the standard normal random variable.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

- (a)  $2\Phi(0.5) - 1$
  - (b)  $\Phi(2) - 0.5$
  - (c)  $2\Phi(0.25) - 1$
  - (d)  $2 - 2\Phi(0.5)$
15. In a linear regression model ( $y_i$  modeled as a linear function of  $x_i$  plus error) the parameters are estimated via least squares. For the mean and the empirical standard deviation of the  $x$  and  $y$  values we obtain  $\bar{x} = 3$ ,  $s_x = 4$ ,  $\bar{y} = 7$  and  $s_y = 3$ . It holds that
- (a) the slope of the regression line is smaller than  $-3/4$
  - (b) the regression line goes through  $(3, 7)$
  - (c) the slope of the regression line is larger than  $3/4$
  - (d) the regression line goes through  $(7, 3)$

16. Let  $X_1, \dots, X_{64}$  be a random sample from a distribution with the expectation  $-1.2$  and variance  $4$ . Let

$$\bar{X} = \frac{1}{64} \sum_{i=1}^{64} X_i$$

be the sample mean. Determine the approximate value of  $P(\bar{X} > -0.9)$  using the Central limit theorem and express it in terms of a suitable R-function.

- (a)  $pnorm(-0.9, 1.2, 0, 5)$
  - (b)  $pnorm(-0.9, -1.2, 0, 25)$
  - (c)  $pnorm(1.2)$
  - (d)  $pnorm(-1.2)$
17. Out of the students in a class, 60% are playing chess, 70% live ice skating, and 40% fall into both categories. Compute the probability that a randomly selected student is neither a chess player nor an ice skating lover.
- (a) 0.9
  - (b) 0.1
  - (c) 0.6
  - (d) 0.4
18. A fast food chain advertises that their large bag of french fries has a weight of 150 grams. Some high school students, who enjoy french fries at every lunch, suspect that they are getting less than the advertised amount. With a scale from their physics teacher, they weigh a random sample of 16 bags. Assuming the level of significance  $\alpha = 10\%$ , what would be the conclusion if the sample mean is 144 g and standard deviation is 15 g? Assume that all conditions for inference are met.
- (a) There is sufficient evidence to prove the fast food chain advertisement is false.
  - (b) There is sufficient evidence to prove the fast food chain advertisement is true.
  - (c) The students do not have sufficient evidence to reject the fast food chain's claim.
  - (d) The students have sufficient evidence to reject the fast food chain's claim.
19. Two features of a novel operating system are compared using a two-sample t-test. The statistics for the first feature are  $\bar{x} = 21$ ,  $s_x = 10$  and  $n_x = 4$  and those for the second feature are  $\bar{y} = 29$ ,  $s_y^2 = 55$  and  $n_y = 5$ . The rejection Region is given though  $R = (-\infty, -q] \cup [q, \infty)$ . Then it holds
- (a) we do neither reject for  $q=0.4$  nor for  $q=1.2$
  - (b) we reject for  $q = 0.4$  but not for  $q = 1.2$
  - (c) we reject for both  $q=0.4$  and  $q=1.2$
  - (d) we do not reject for  $q=0.4$  but for  $q=1.2$

20. For a project, a high school student randomly picks 100 fellow Statistics students to survey on whether each has either a PC or Apple at home (all students in the school have a home computer) and what score (1. 2. 3. 4. 5) each expects to receive on the Statistics exam. A chi square test of independence results in a test statistic of 8. How many degrees of freedom are there?
- (a) 4
  - (b) 7
  - (c) 9
  - (d) 1