Control questions for DOE1

DOE 2 course 2019 first semester

1. How many parameters characterise a Gauss distribution?
2. How could one tabulate the F(x) distribution function values for a Gauss distributed x random variable (how many dimensions are required?
3. At a filling station the volume of 200 PB bulbs is measured. What are the population and the sample in this case?
4. The volume of one PB bulb is measured 10 times. What is the population and sample in this case?
5. Give examples for variables measured on nominal, ordinal, interval and proportional scales, respectively!
6. From among data measured on nominal, ordinal, interval or proportional scales which ones would deserve ordering?
7. What scale is used to interpret the following types of data: gender, personal identification number, order in the line for gymnastic in the school, caloric value of the gas, temperature of the bath water?
8. What would be the expected value an variance of arithmetic average of a sample taken from any distribution?
9. Give example for a random variable following Student's t distribution!
10. What is the difference between the z and t distribution?
11. Why do we use the average of repeated measurements instead of a single measured value?
12. What does the independence of measurement errors means? Give examples!
13. Why do the sample mean and sample variance differ from the expected value and variance of the population, respectively?
14. When is the difference between means of two samples qualified as significant?
15. What is the difference between one-sample t-test, two-sample t-test and paired t-test?
16. When do we reject the null hypothesis?
17. How do we decide on the allowed probability of error of first kind?
18. What is the power of a test?
19. The effect of cleaning shoes on the life time of the shoes is to be investigated. How would you build up the design with 10 people for two-sample t-test and paired t-test, respectively? Which variant promises more information and why?
20. Can you imagine a situation when the one-at-a-time approach is better than the matrix design?
21. When do we qualify an effect as significant? What does it mean?
22. How can we decide if the linear function is proper for describing the relation and no second order terms are required?
23. What is the meaning of interaction? What does it mean if there is or there is no interaction between two factors?
24. What does the orthogonality of a design means?
25. What do we assume on the form of a function when two-level design is used?
26. What do we assume about the experimental errors?
27. What is the condition for evaluating the effects in a design independently of each other?
28. How can we compute the effect of a factor from the results? What is the meaning of a negative effect?
29. How can we compute the interaction effect? When may we say that there is no interaction between two factors?
30. What is the relation between the effect and coefficient (slope)?
31. If the effect of temperature calculated for levels 20°C and 25°C is given, how can we calculate the effect of 1 ^{0}C rise on y?
32. How can we decide on the significance of an effect if there are repeated experiments?

33. In what sense are the variables in a two-level factorial design orthogonal? What is the advantage of orthogonality?
34. Is it necessary to have equal number of replications in each point of the design?
35. How can we decide on the significance of an effect if there are no repeated experiments? Is there any condition for that?
36. Why do we perform experiments in the centre of the design? How many centre points are required if any?
37. Are there case when we cannot perform experiments in the centre of the design?
38. How can we check if the model (after omitting the not significant effects) is adequate for describing the results?
39. Why and how do we randomise the runs in a design?
40. What is the consequence if the variation interval is too narrow?
41. What should we do if we have suspicion that the variation interval was too narrow?
42. What is the consequence if the variation interval is too wide? What can we do if this occurs?
43. When do we choose quadratic design?
44. How large is the covariance of independent variables?
45. What does it mean if the correlation coefficient between two random variables takes small positive value?
46. Is the correlation coefficient a property of a population or of the sample?
47. Can one be sure about independence of two random variables if the correlation coefficient between them is numerically small?
48. How would you decide if you use correlation analysis or regression analysis?
49. What is the advantage of using the model in the form $\bar{Y}_i = \bar{\alpha} + \bar{\beta}\left(x_i - \bar{x}\right)$ instead of the $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}x_i$ form?
50. Give the meaning of $\bar{\alpha}$.
51. What are the assumptions used for regression analysis and where (at which point of the analysis) do we use them? What should not be done if some assumptions are not hold?
52. Do the true and estimated regression lines necessarily differ from each other?
54. Does it change the estimated line (its parameters) if we do not consider that the number of repetitions at different x_I places is different and the line is fitted to $\bar{y}_i$ values?
55. Does it change the estimated line (its parameters) if we do not consider that the number of repetitions at different x_I places is different and the line is fitted to all y_{ij} values?
56. Why is it important for the user to obtain more efficient estimators for the parameters?
57. Is the run order of experiments to be randomised if we measure a single y_i value at each x_i place?
58. Is the run order of experiments to be randomised if we measure several repeated y values at each x_i place? (if yes, how)
59. When is it sufficient to measure y at only two x values and when is it not?
60. When is it sufficient to measure y at only one x value and when is it not
61. Give the definition of confidence band (confidence curve for the estimated line) and prediction band (confidence curve for the individual values).
62. Which one of them (confidence band or prediction band) is wider?
63. Do you expect experimental points outside of the confidence band?
64. Do you expect experimental points outside of the prediction band?
65. If we accept the null hypothesis that the true line goes through the origin, may we be sure about that? How would you formulate the errors of first and second kind respectively at this decision? How can you reduce the probability of error of second kind?
66. What is the R^2 coefficient of determination and why is it to be adjusted?
67. What does the R^2 depend on?
68. Are you sure that the true regression line exists? What does it mean if not?