Rules: online test in lms, no proctoring, 20 questions, 60 minutes, only numerical answers are checked, two digits after decimal point are requested, use anything you want (calculators, python/r code, google, ...), don't cheat.

1. (bootstrap) I have a sample X_1, ..., X_{100}.

   I generate one naive bootstrap sample X_1^*, ..., X_{100}^*.

   What is the probability that the first observation will be present in the bootstrap sample 2 times or more?

2. (bootstrap) Nature generates random variables X_1, ..., X_{100} independently and uniformly on $[0; 10]$.

   I generate one naive bootstrap sample X_1^*, ..., X_{100}^*.

   Find the variance $\operatorname{Var}(X_1^*)$.

3. (welch) We have data for an AB-experiment $\bar{X}_a = 10$, $\bar{X}_b = 12$, $n_a = 20$, $n_b = 30$, $\sum(X_i^a - \bar{X}_a)^2 = 100$, $\sum(X_i^b - \bar{X}_b)^2 = 200$.

   Calculate the standard error of $\bar{X}_a - \bar{X}_b$ for the Welch test.

4. (welch) Assume that X_i are independent and identically normally distributed $\mathcal{N}(\mu, \sigma^2)$, sample size is $n = 10$.

   Find $\operatorname{Var}(\sum(X_i - \bar{X})^2/(n-1))$.

5. (mw test) I have five results of two runners A and B for the 5 km race: 25:12 (A), 26:34 (B), 27:43 (A), 28:12 (A), 29:05 (B).

   Calculate Mann-Whitney statistic U_A that tests the null-hypothesis of equal distributions of time.

   (The statistic U_A should positively depend on the ranks of the runner A).

6. (mw test) I have five results of two runners A and B for the 5 km race: three results for A and two results for B. Assume that the running time for both runners are continuously distributed and their distribution are equal.

   What is the probability that the running times of the runner A will get the ranks 1 and 5?

7. (cuped)

8. (cuped)

9. (matching)

10. (matching)

11. (multiple comparison)

12. (multiple comparison)

13. (sample size)

14. (sample size)

15. (contingency table) I eated 10 M&Ms: 2 green, 1 red, 4 yellow, 1 green, 2 red.

    Only these three colors are possible. I assume that yellow and green colors are equally probable.

    Calculate the maximal log likelihood for my model.

16. (contingency table) Consider the following contingency table

	$B = 1$	$B = 2$
$A = 1$	10	20
$A = 2$	30	40

Calculate LR statistic that checks the hypothesis that A and B are independent against dependency alternative.

17. (anova 1+2)

18. (anova 1+2)

19. (partial correlation) The variables X and Y are jointly normal with zero means, unit variances and $\text{Corr}(X, Y) = 0.8$.

    Find α such that $X^* = X - \alpha Y$ is not correlated with Y.

20. (partial correlation) The variables X_1, X_2, ... are independent and identically distributed with mean 5 and variance 7.

    Find $\text{pCorr}(X_1, X_2; S)$ where $S = X_2 + X_3$.