Home Assignment 1

1. Consider naive bootstrap for a sample of n observations. We know that y_i are iid uniform on $[0; 10]$.

   (a) What is the probability that original y_5 will enter any particular bootstrap sample exactly 3 times?

   (b) What is the limit of probability you found in (a) when $n \to \infty$?

   (c) What is the limit of probability $\mathbb{P}(\max y_i > \max y_i^*)$ when $n \to \infty$?

   (d) Find $\mathbb{E}(y_1^*)$, $\mathbb{V}\mathrm{ar}(y_1^*)$, $\mathbb{C}\mathrm{ov}(y_1^*, y_2^*)$.

2. Winnie-the-Pooh has collected 15 honey samples from good (G) and bad (B) bees. Winnie has ordered the honey quality from lower to higher on the real line. He has forgot the exact values, but here is the ordering: BBBGB-BGGBG-GGBGG.

   (a) Please, help Winnie and test the null-hypothesis that the honey quality is the same against alternative that good bees produce better honey at 5% significance level.

   (b) Using computer find the exact P-value of this ordering.

3. What should be the minimal total sample size n for the following testing problem?

   The goal is to have first type error $\alpha \le 0.01$, second type error $\beta \le 0.1$, minimal detectable difference in probabilities should be $H_A : p_B - p_A \ge 0.2$, 60% of total sample should go into test group B and 40% $-$ into control group A.

4. Winnie-the-Pooh simulteneously tests four hypothesis. Actually three null hypothesis are true and one is false (the corresponding p-value is uniform on $[0; 0.5]$), but Winnie does not know this.

   (a) Winnie uses Bonferonni correction with target bound 0.05 for $FWER$. What is the actual $FWER$?

   (b) Winnie uses Holm-Bonferonni correction with target bound 0.05 for $FWER$. What is the actual $FWER$?

5. The sample correlation matrix of three variables (a, b, c) is

$$C = \begin{pmatrix} 1 & 0.7 & 0 \\ 0.7 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

   The number of observations is equal to $n = 101$.

   (a) State the weights of a, b and c in the first principal component.

   (b) All you know about observation number 7 that the first principal component is equal to 0.1 and the second $-$ to 0.2. Recover approximate values of a_7, b_7 and c_7.

   (c) How much variance is explained by the first two components?

Deadline: **2022-12-06, 21:00**.