1. **Problem**
   I have a sample $X_1, \ldots, X_{90}$. I generate one naive bootstrap sample $X_1^*, \ldots, X_{90}^*$.

   What is the probability that the first observation will be present in the bootstrap sample exactly 3 times?

2. **Problem**
   We have data of an AB experiment: $\bar{X}_a = 5.4$, $\bar{X}_b = 6$, $n_a = 18$, $n_b = 15$, $\sum(X_i^a - \bar{X}_a)^2 = 890$, $\sum(X_i^b - \bar{X}_b)^2 = 800$.

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

3. **Problem**
   I have five results of two runners A and B for the 5 km race:

   16:49 (B), 21:17 (A), 18:30 (B), 6:18 (B), 20:16 (A), 15:39 (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).

4. **Problem**
   I have 30 hypothesis with independent statistics. The null hypothesis for all 30 cases is actually true, but I don't know this.

   I calculate all p-values. If the 4 lowest p-value are simultaneously lower than 0.01 I wrongly conclude that not all H_0 are true. Otherwise I correctly conclude that all H_0 are true.

   What is the probability that I will get the correct conclusion?

5. **Problem**
   My target variable is binary and I wish minimal detectable effect equal to 0.04, probability of I-error not greater than 0.01, probability of II-error not greater than `beta`. The control and experimental group are of the same size equal to n.

   Which minimal value of n is sufficient in the worst case?

6. **Problem**
   Vasiliy loves to eat shaurma. He has 5 local shaurma dealers. Vasiliy bought 4 shaurmas from each dealer and measured their weight. He would like to test the hypothesis that mean weight is the same for all dealers.

   Total sum of squares is 600, between sum of squares is 100 Calculate the F -statistic to test the hypothesis.

1. Let random variables $Y_1, \ldots, Y_n$ be iid uniform $U[0;1]$. Consider the naive bootstrap sample $Y_1^*, \ldots, Y_n^*$.

   Find $\mathbb{P}(Y_1^* = Y_1)$, $\mathrm{Cov}(Y_1^*, Y_2)$, $\mathbb{P}(\max\{Y_1, \ldots, Y_n\} = \max\{Y_1^*, \ldots, Y_n^*\})$.

2. The eigenvalues of sample correlation matrix are 2.5, 0.3 and 0.2. The eigenvalue $\lambda = 2.5$ corresponds to eigenvector $v = (3, 4, -2)$.

   (a) James Bond predicts every original variable using multivariate regression on the first two components. What is the average value of R^2 he will get?

   (b) Express the first principal component in terms of scaled original variables a, b and c.

3. Winnie-the-Pooh simulteneously tests h null hypothesis using independent samples. All the null hypothesis are true but Winnie does not know it.

   (a) What is the expected value of the lowest P-value?

   (b) What is the expected number of wrongly regected hypothesis if Winnie rejects all the hypothesis with P-value less 0.1?

4. There are three continuosly distributed samples of the same size n, $X_1, \ldots, X_n$, $Y_1, \ldots, Y_n$, $Z_1, \ldots, Z_n$. Imagine that the null hypothesis that all samples have the same distribution is true.

   Consider the random variable R_X — the sum of ranks of the X sample in the pooled sample.

   (a) Find the expected value $\mathrm{E}(R_X)$.

   (b) What is the probability that R_X will be equal to $n(n+1)/2$?