ASSIGNMENT 2

Antoni Adamczyk

MatNr: 12306508

1. **Z-Test - Crop Yield**
   a. **Hypothesis Definition**

   Let μ be the actual yield level in tons per hectare of fertilized field. Hence, X_1 ~ N(μ, σ^2) and therefore $\dot{X}_n$ ~ N(μ, σ^2/n). The most natural hypotheses are:

   H_0: μ ≤ μ_0, new fertilizer is not increasing the average yield

   H_1: μ > μ_0, new fertilizer is increasing the average yield

   Both hypotheses are composite, however, we can work with H_0: μ = μ_0 which is simple.

   b. **Threshold Determination**

   Formula:

   threshold = μ_0 + $z_{1-\alpha}$*(σ / n^(½))

   Data:

   μ_0 = 40.0

   n = 8

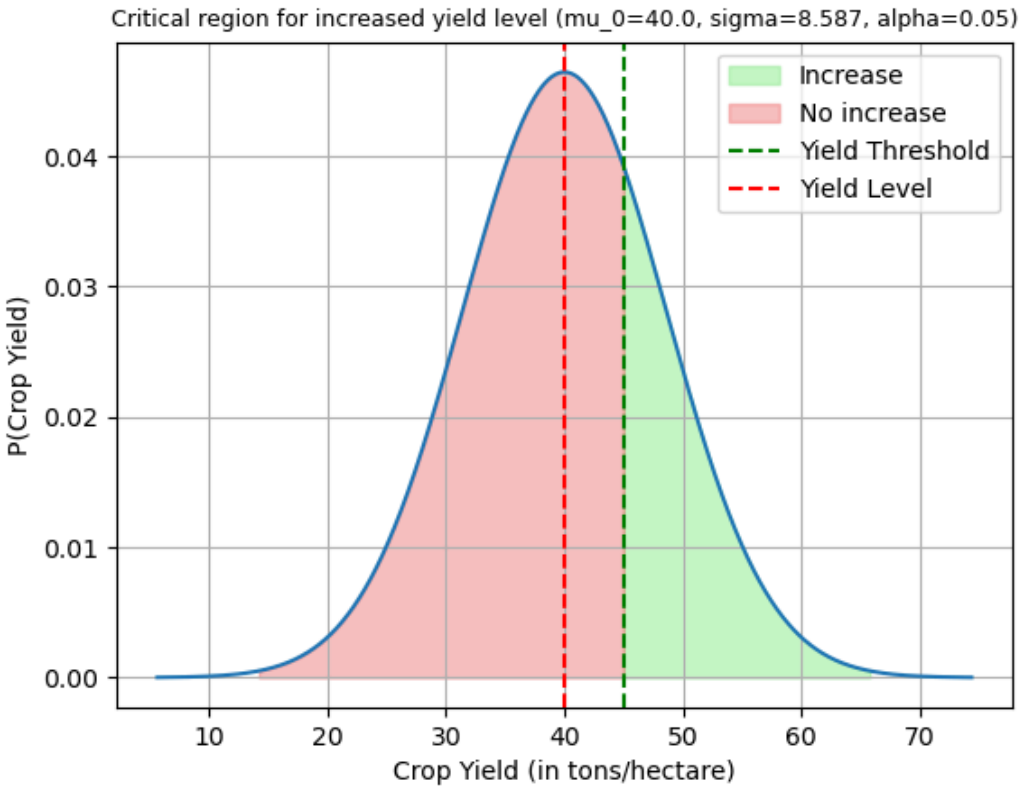
   alpha = 0.05

   σ = 8.587 (draw from N(9,1))

   $z_{0.95}$ = 1.645

   Outcome:

   threshold = 44.994



c. **Power Analysis I**

c1) Power = 0.630

c2) Fields needed for power 0.9 = 18

d. **Power Analysis II**

Yield for power of 0.99 = 52.072

e. **Comparison and Discussion**

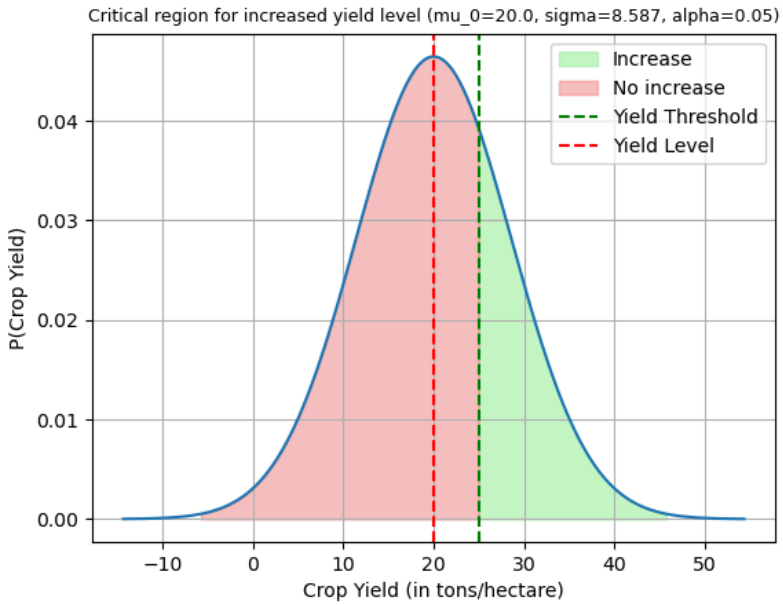
Formula:

threshold = $\mu_0 + z_{1-\alpha}*(\sigma / n^{(1/2)})$

I.

μ_0 = 20.0

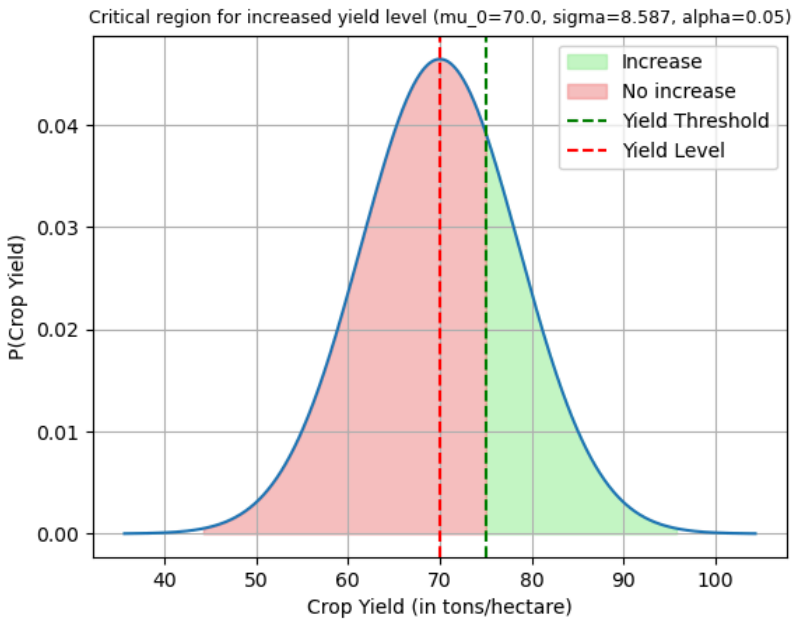
threshold = 24.994



II.

μ_0 = 70.0

threshold = 74.994



III.

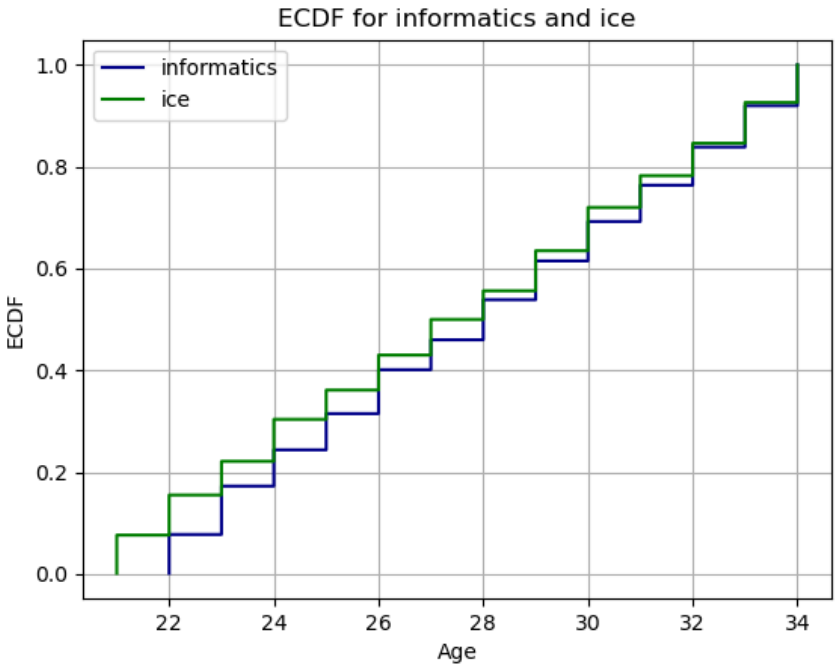
With the different values for μ_0 changes value of the threshold and the range of values of crop yield (in tons/hectare). The difference 'threshold - μ_0' does not change.

If we increase/decrease the number of tested fields (n), the value of the threshold and the difference 'threshold - μ_0' changes. The bigger is n, the smaller becomes the difference threshold - μ_0'.

If we increase/decrease the measurement error (σ), the value of the threshold, the range of values of crop yield (in tons/hectare) and the difference 'threshold - μ_0' changes. The bigger is σ, the bigger becomes the difference threshold - μ_0'.

If we increase/decrease the significance level (α), the value of the threshold and the difference 'threshold - μ_0' changes. The bigger is α, the smaller becomes the difference threshold - μ_0'.

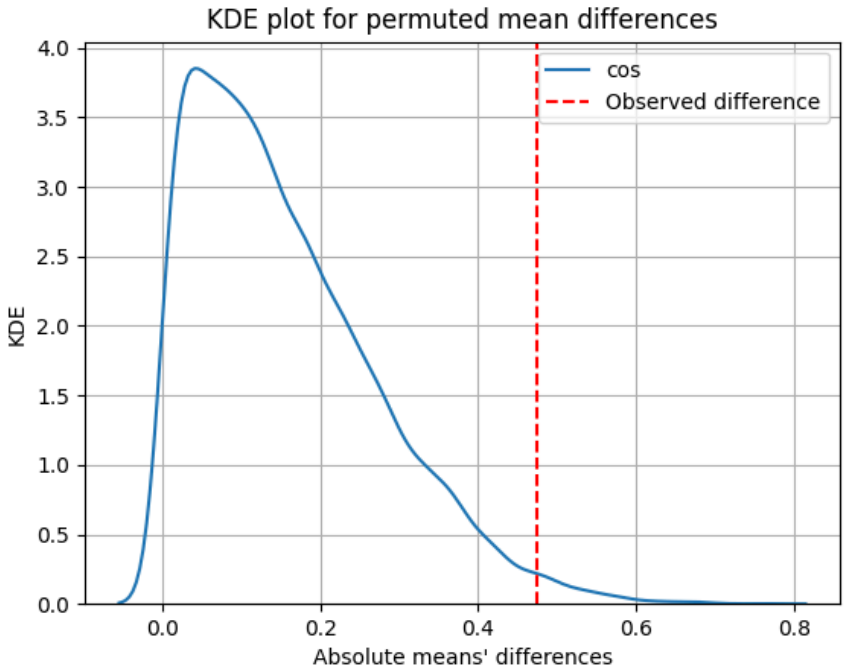
2. **Permutation Test**
   a. **Data Analysis**



   b. **Permutation Test**
      I. p-value = 0.016
      α = 0.04
      ===> p-value < α-significance level

      According to statistical convention, we reject the null hypothesis in favor of the alternative hypothesis. This means that the difference in age between the 'informatics' group and the 'ice' group is statistically significant at the 0.04 significance level.

II.



c. **Bonferroni Correction**

I. H_0 was rejected 24 times out of 100.

II. Bonferroni significance level used for the test is 0.001
This adjustment ensures that the overall Type I error rate does not exceed the chosen significance level α=0.1.

III. We reject the global null hypothesis H_0 if at least one of the hypotheses is rejected. There were 24 rejected hypotheses so we also reject global null hypothesis H_0.

IV. One of the downsides is increased Type II error rate, the Bonferroni correction due to being very conservative has increased likelihood of failing to reject Type II error.
The other downside is also lack of sensitivity, the correction can reduce the power of the statistical test, making it less sensitive to detecting true effects.