TASK LIST NO. 11: Introduction to Statistical Decision Theory

Task 1

We consider a batch of goods (e.g., processors) in which we examine defectiveness. We test the hypothesis $H_0 : p = p_0$ (good batch) against $H_1 : p = p_1$ (defective batch). A test was constructed that rejects the batch if the number of defective items in a randomly drawn sample exceeds a certain k.

Calculate:

a) The probability of a Type I error α (rejecting a good batch – producer's risk).
b) The probability of a Type II error β (accepting a bad batch – consumer's risk).

Task 2

For the test from Task 1, determine the power of the test $(1 - \beta)$ for several alternative values of parameter p. Prepare a graph of the test power curve (power function). What does this graph tell us about the "sensitivity" of the decision algorithm to deviations from the norm?

Task 3

For a mean test $H_0 : \mu = 100$ with known $\sigma = 5$ and $n = 25$:

a) Determine the formula for the operating characteristic (OC) function: $L(\mu) = P(\text{acceptance of } H_0|\mu)$.
b) How will changing the sample size to $n = 100$ affect the steepness of this curve (discriminatory ability)?

Task 4

We want to construct a test for the mean that meets strict safety requirements:

- The risk of rejecting the norm when it is met (α) is to be 0.01.
- The risk of accepting the norm when the mean shift is 2 units (β) is not to exceed 0.05.

What is the minimum number of measurements that must be performed?

Task 5

The defectiveness of the production of certain products has been 10% $(p_0 = 0.1)$ so far. A new technology is expected to reduce defectiveness to 5% $(p_1 = 0.05)$. Instead of taking a fixed sample, we take items one by one.

Construct a sequential probability ratio test (Wald test), setting risks $\alpha = 0.05$ and $\beta = 0.10$.

a) Determine the decision lines (acceptance region, rejection region, and continuation region).
b) Present the procedure in the form of an algorithm (pseudocode).

Task 6

For the test from Task 5, suppose the following were drawn sequentially: Good, Good, Bad, Good, Good, Good, Good...

Mark these points on the sequential test chart. In which step (if any) will the algorithm make the decision "The new technology is better"?

Task 7

One of the advantages of sequential methods is that they require less data on average than classical tests. For the test from Task 5, calculate the expected number of steps (samples) needed to make a decision (ASN), assuming that hypothesis H_0 is true.

$$E(n) \approx \frac{(1 - \alpha) \ln A + \alpha \ln B}{E(z)}$$

where A, B are decision thresholds.

Task 8

An automated machine produces parts with a nominal diameter μ_0. We suspect that the machine has become decalibrated and the mean has increased to μ_1. The deviation σ is known.

Construct a sequential test verifying $H_0 : \mu = \mu_0$ against $H_1 : \mu = \mu_1$. Write the "stop" condition for this algorithm.

Task 9

We have two possible decisions d_1 (system deployment) and d_2 (no deployment) and two states of nature θ_1 (system works correctly) and θ_2 (system has errors). The loss (cost) matrix is as follows:

- If d_1 and θ_1: Cost $= 0$
- If d_1 and θ_2: Cost $= 1000$ (failure at client's site)
- If d_2 and θ_1: Cost $= 100$ (lost profit)
- If d_2 and θ_2: Cost $= 0$

What decision should be made using the **Minimax** criterion (minimizing the maximum loss)?

Task 10

For the situation from Task 9, assume that we know from previous tests that the probability of errors occurring is $P(\theta_2) = 0.05$. Calculate the expected loss (Bayes risk) for both decisions. Which decision is optimal in the Bayesian sense?