Assignment 2: Hypothesis Testing

Course: To p or not to p

Problem

You are to test the claim by a mineral water bottle manufacturer that its bottles contain an average of 1000 ml (1 litre). A random sample of $n = 12$ bottles resulted in the measurements (in ml): 992, 1002, 1000, 1001, 998, 999, 1000, 995, 1003, 1001, 997 and 997.

It is assumed that the true variance of water in all bottles is $\sigma^2 = 1.5$, and that the amount of water in bottles is normally distributed.

Test the manufacturer's claim at the 1% significance level (you may use Excel to calculate the pp-value). Also, briefly comment on what the hypothesis test result means about the manufacturer's claim, and if an error might have occurred which type of error it would be.

In summary, the assignment requires:

- the calculation of the sample mean from the raw observations
- the formulation of the hypotheses, H_0 and H_1
- calculation of the test statistic value
- calculation of the p-value
- a decision of whether or not to reject H_0
- an inferential conclusion about what the test result means
- indication of which type of error might have occurred.

Solution

- Sample mean $= \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = 998.75$, here $n = 12$.
- We have $H_0 : \mu = 1000$ vs. $H_1 : \mu \neq 1000$.

- Test statistic $Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{998.75-1000}{\sqrt{1.5}/\sqrt{12}} = -3.535534$, where $\sigma^2 = 1.5$, given.

- p-value $= 2 \times \Phi(Z) = 2 \times \Phi(-3.535534) = 2 \times 0.000203476 = 0.000406952$, where $\Phi(.)$ is the cumulative distribution function (**CDF**) for the normal distribution, the values can be computed from table or using **R** function **pnorm()**.
  As can be seen from the below figure, we need to compute the area of the shaded red region (under the curve), to the left and right of the vertical lines, respectively, to compute the p-value for this **two-tailed hypothesis test** (but the normal distribution being symmetric, the area of these two regions are equal).

- Here p-value $= 0.000406952 < 0.01$, so we shall reject H_0 at $\alpha = 1\%$ level of significance.

- The inferential conclusion is: the mineral water bottles does not contain an average of 1000 ml (1 litre).

- We may have committed a **type I** error with probability $\alpha = 0.01$ (by rejecting a true H_0).