

t-test & ANOVA (Analysis of Variance)

KUMARI SONA RANI
ASSISTANT PROFESSOR
DEPARTMENT OF BOTANY
S.SINHA COLLEGE, AURANGABAD

FOR M.Sc 4th semester

What are they? The *t*-test is a method that determines whether **two** populations are statistically different from each other, whereas ANOVA determines whether **three or more** populations are statistically different from each other. Both of them look at the difference in means and the spread of the distributions (i.e., variance) across groups; however, the ways that they determine the statistical significance are different.

When are they used? These tests are performed when 1) the samples are independent of each other and 2) have (approximately) normal distributions or when the sample number is high (e.g., > 30 per group). More samples are better, but the tests can be performed with as little as 3 samples per condition.

How do they work?

***t*-test Example**

We want to determine whether the concentration of Proteins 1 – 4 in serum are significantly different between healthy and diseased patients. A *t*-test is performed, which can be visually explained by plotting the protein concentration on the X-axis and the frequency along the Y-axis of the two proteins on the same graph (Figures 1 – 4).

Proteins 1 & 2 have the same difference in protein concentration means but different group variances. Alternatively, Proteins 3 & 4 have similar variances but Protein 4 has a larger difference in protein concentration means between the patient groups.

A *t*-test assigns a “*t*” test statistic value to each biomarker. A good differential biomarker, represented by little to no overlap of the distributions and a large difference in means, would have a high “*t*” value.

Which is a better biomarker of disease: Protein 1 or Protein 2? *Protein 1*

Which is a better biomarker of disease: Protein 3 or Protein 4? *Protein 4*

Same means, different variance

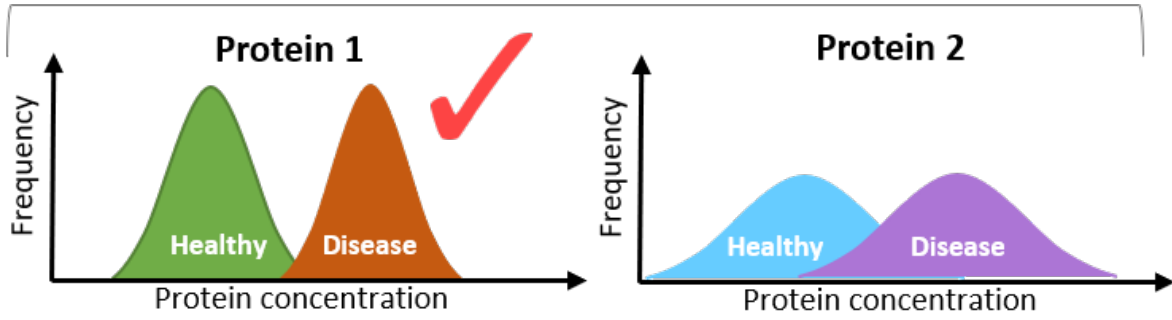


Figure 1. Overlapping histogram plots for concentrations of protein 1 in different populations.

Figure 2. Overlapping histogram plots for concentrations of protein 2 in different populations.

Same variance, different means

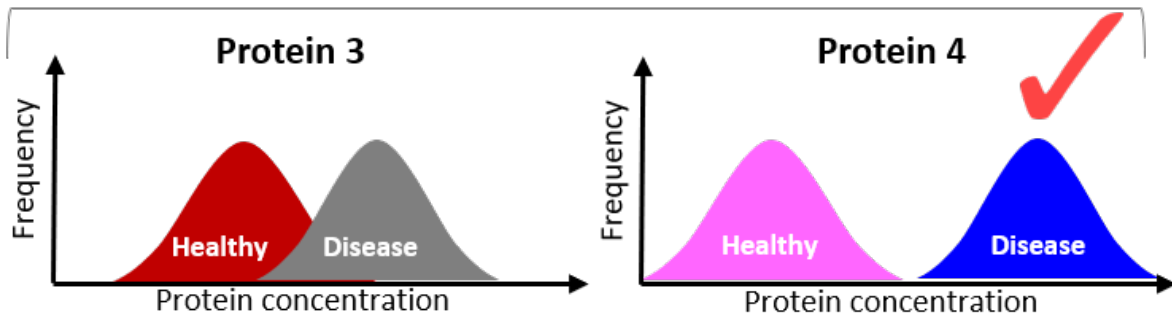


Figure 3. Overlapping histogram plots for concentrations of protein 1 in different populations.

Figure 4. Overlapping histogram plots for concentrations of protein 2 in different populations.