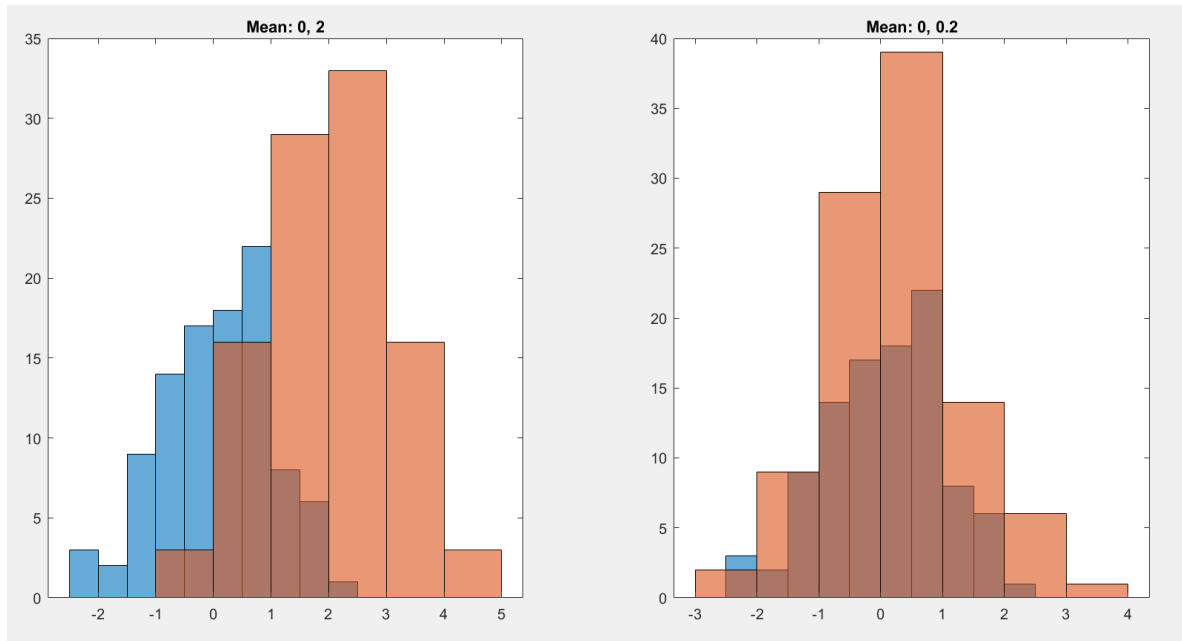


Computer Experiment 5.1



- (a) H_1
- (b) H_0
- (c) H_1
- (d) H_0

In question (a) and (c), the mean difference between two distributions is equal to 2 which is big, and the t test result is 1 meaning they're significantly different as so-called alternative hypothesis. In question (b) and (d), the mean difference between two distributions is equal to 0.2 which is small, and the t test result is 0 meaning they're almost equal as so-called null hypothesis.

Table 5.1 Computation

The table 5.1 from the text is very similar to the z table. The main algorithm behind it is by calculating the pdf of normal distribution by using the equation below in image 1. Then calculate the integral of the pdf from negative infinite to value z. That is CDF (Image 2) and shows the percentage of other numbers lie all the values before x. We can then calculate acceptance intervals by subtracting the entire percentage area (values smaller than positive x) by the small percentage area (values smaller than negative x).

$$f(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

(Image 1: Equation of pdf from Normal Distribution)

$$\int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

(Image 2: Equation of CDF from Standard Normal Distribution - $N(0,1)$)

“Spectral Learning”

1. The kernel enters the approach during data clustering.
2. Affinity matrix is calculated by finding out the similarities between two objects/points. It's very a decimated version of the similarity matrix.
3. Classification algorithm is processed like this. For affinity matrix, if a pair of points (i, j) are in the same class, the matrix A is equal to 1. If a pair of points (i, j) are not in the same class, the matrix A is equal to 0. Then normalize the affinity matrix A. So the data used for classification algorithm is in pairs.