2. How confident are we that h2 and h3 are worse than h1?

We need to compare the error rates of h2 and h3 with h1. All hypotheses were tested on the same N = 45 instances.

- E^1 = 6.67% = 0.0667
- E^2 = 8.89% = 0.0889
- E^3 = 13.3% = 0.1330

We are interested in the confidence that hypothesis h(A) performs worse than hypothesis h(B), which is the confidence that E(A) >E(B), or E(A) − E(B) > 0. We use the approximation for the difference between two hypotheses tested on the same data set.

The number of instances where h(A) is wrong and h(B) is right (k(AB)) and where h(B) is wrong and h(A) is right (k(BA)) are not provided, so we will approximate the standard deviation of the difference in error **STD(**d^**)** using the general form, where d^ = E^(A) − E^(B).

Approximation for Standard Deviation **:**

STD(d^) = Root((E^(A) * (1 - E^(A)) / N) + E^(B) * (1 - E^(B)) / N **)**

This is an approximation and generally an overestimate when using the same test set.

Comparison of h2 and h1

- Difference d^2, 1 = E^2 − E^1 = 0.0889 − 0.0667 = 0.0222
- E^1(1−E^1) = 0.0667 × 0.9333 = 0.0622
- E^2(1−E^2) = 0.0889 × 0.9111 = 0.0809

STD(d^(2,1)**)** = **Root((**0.0622/45) + (0.009/45)**)**

= **Root(**0.001382 + 0.001798**)**

= **Root(**0.00318**)**

= 0.0564

The number of standard deviations (Z) that the difference d^(2,1) is above zero:

Z = (d^2,1 − 0) / **STD(**d^(2,1)**)**

= 0.0222 / 0.0564

= 0.3936

Using the Z(N) table (or a standard Normal Distribution table, which is not fully provided), a Z score of 0.3936 corresponds to a one-tailed confidence level (the probability of being above zero). The table provides Z(N) for two-sided N% confidence intervals. We can use the closest values in Table 3.

- Z = 0.41 is for N = 30% two-sided CI, which means a one-tailed probability of (1 − 0.3) / 2 + 0.3 = 0.85 (or 0.65 if we're looking at the right tail, i.e., 1 − (1 − 0.3) / 2).

A Z value of 0.3936 is slightly less than 0.41. The one-tailed confidence is approximately **65%**.

Comparison of h3 and h1

- Difference d^(3,1) = E(^(3) −E(^(1) = 0.1330 − 0.0667 = 0.0663
- E(^3(1 − E(^3) = 0.1330 × 0.8670 = 0.1153

STD(d^(3, 1)**) = Root(**(0.0622 / 45) + (0.1153 / 45)**)**

= Root(0.001382 + 0.002562**)**

= Root(0.003944**)**

= 0.0628

Z = (d^(3, 1) - 0) / **STD(**d^(3, 1)**)**

= 0.0663 / 0.06288

= 1.056

Using the Z(N) table (Table 3),

Z = 1.00 corresponds to a two-sided 68% CI. The one-tailed probability (confidence that h3 is worse than h1) is 1 − (1 − 0.68) / 2 = 0.84.

A Z value of 1.056 is slightly greater than 1.00. The one-tailed confidence is approximately **85%**.

- Confidence for h2 worse than h1: **65%**.
- Confidence for h3 worse than h1: **85%**.