Super Enhanced D-Value Method (\(\text{pos}_{\text{sed}}\)):

The Enhanced D-value routine can be improved even more by including a robust estimate for \(P(x)\), the fraction of positives at \(x\). As discussed above, an exact solution to the positive fraction for all \(x\) with \(C_x > 0\) is described by the equation,

\[
\text{POS}(x) := \frac{D_x + P_x}{C_x}
\]

The enhanced D-value routine assumes \(P_x\) is zero and evaluates \(D_x/C_x\) at \(x = x_d\).

One way of possibly improving the enhanced D-value routine would be to try to estimate the \(P(x = x_d)\). Let's look at this question graphically.

If one could estimate \(P(x = x_d)\), then the function,

\[
\frac{D_{x_d} + P_{x_d}}{C_{x_d}} = 0.089 \quad \text{pos} = 0.089
\]

would more precisely estimate the positive fraction. Note, the exact positive fraction value, \(\text{pos}\), equals the above expression.

Recall that the D-value is given by the maximum difference between the normalized cumulative control (C) and test (T) histograms.
As shown above, we can improve upon this estimate by dividing by \( C(x=x_d) \) (enhanced D-value). Thus, the algorithm for estimating \( P(x=x_r) \), where \( r \) is the last channel in the histogram or channel domain, is 1) normalize the cumulative distributions, \( C \) and \( T \), so that they are unity at the end of the \( x \)-axis (\( r \)), 2) find the maximum difference value between these two cumulative distributions and 3) divide by the cumulative control evaluated at this location of maximum difference.

We can potentially estimate \( P(x=x_d) \) by applying the same idea to a subportion of the two cumulative distributions. The example below will make this point clearer.

Suppose we renormalize \( C \) and \( T \) such that they are unity at \( x=x_d \). Our new \( x \) domain will range from 0 to \( x_d \). Mathematically we are just substituting \( x_d \) for \( r \) in the above described algorithm.

\[
x_2 := 0 \ldots x_d
\]

\[
C_{x_2} := \frac{C_{x_2}}{C_{x_d}} \quad T_{x_2} := \frac{T_{x_2}}{T_{x_d}} \quad C_{x_d} = 1 \quad T_{x_d} = 1
\]

Let's find the difference between these two new cumulative distributions over \( x_2 \).

\[
D_{x_2} := C_{x_2} - T_{x_2}
\]

The maximum difference difference is given by

\[
d_{max} := \max(D_2)
\]

and its location is

\[
x_{d_2} := \text{LOC}(D_2, d_{max})
\]

By the same reasoning as shown for the enhanced \( d \)-value, if we divide this maximum difference, \( D_{2_{max}} \), by \( C_2(x=x_{d_2}) \), we should have a reasonable estimate of \( P(x=x_d) \), \( P_{x_d} \).

\[
P_{x_d} := \frac{d_{max}}{C_{x_{d_2}}} \quad P_{x_d} = 8.064 \times 10^{-3} \quad x_d = 0.019
\]

Note that our estimate \( P_{x_d} \) is an estimate of \( P(x=x_d) \). We can now plug this new estimate into the general cumulative distribution formula to better estimate the positive fraction.

\[
\frac{D_{x_d} + P_{x_d}}{C_{x_d}} = 0.077 \quad \text{pos} = 0.089 \quad \text{pos}_{x_d} = 0.068
\]

As we hoped, the estimate is approaching the actual pos value. Let's plug in the above logic into the equation to simplify the equation.

Expanding \( D_{x_d} \), we obtain,

\[
\frac{C_{x_d} - T_{x_d} + P_{x_d}}{C_{x_d}} = 0.077
\]
Substituting,

\[ C_{x_d} - T_{x_d} + \frac{C_{2x_d} - T_{2x_d}}{C_{x_d}} = 0.077 \]

Further substituting,

\[ C_{x_d} - T_{x_d} + \frac{C_{x_d} - T_{x_d}}{C_{x_d}} = 0.077 \]

Simplifying,

\[ C_{x_d} - T_{x_d} + \left( \frac{C_{x_d} - T_{x_d}}{C_{x_d}} \right) = 0.077 \]

Further simplifying,

\[ 1 - \frac{T_{x_d}}{C_{x_d}} + \frac{1}{C_{x_d}} - \frac{T_{x_d}}{C_{x_d}} = 0.077 \]

One could conceptually add additional recursive elements to the above equation to improve the accuracy, but the errors will also likely increase.

The other errors were
\[ \varepsilon_{j} = 66.284 \]
\[ \varepsilon_{d} = -29.629 \]
\[ \varepsilon_{ed} = -23.06 \]
\[ \varepsilon_{ns} = -17.645 \]