1. **KS Test practice**

Recall from class that the one-sided and two-sided KS statistics for two samples are:

\[ D^+_{nm} = \sup_x \{ F_n(x) - G_m(x) \} \]

\[ D_{nm} = \sup_x |F_n(x) - G_m(x)| \]

and they have asymptotic null distribution

\[ P \left( \sqrt{\frac{mn}{m+n}} D^+_{mm} \leq k \right) = 1 - e^{-2k^2} \]

\[ P \left( \sqrt{\frac{mn}{m+n}} D_{mn} \leq k \right) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2k^2} \]

(As an aside, for \( m = n \), in the one-sided case, we find that \( m(D^+_{mm})^2 \) is exponential with mean 1.)

To convert this into a one-sample situation, imagine taking \( G \) to be a *theoretical* CDF, rather than an EDF (imagine \( m \) going to infinity, then \( \frac{m}{m+n} \) is about 1). Then the following holds for \( D^+ \), using the theoretical CDF \( G \):

\[ D^+_n = \sup_x \{ F_n(x) - G(x) \} \]

\[ P \left( \sqrt{n} D^+_n \leq k \right) = 1 - e^{-2k^2} \]

The two-sided test actually really hard to use by hand, so you can just use the one-sided test, or use R, to test the following hypothesis:

- Could \( \{0.52, 0.65, 0.13, 0.71, 0.55\} \) been a sample from \( \text{Unif}[0,1] \)?

It will help to plot the EDF. Note that if you use a theoretical CDF for \( G \), rather than an EDF, that corresponds to taking \( m \to \infty \), by the Glivenko-Cantelli theorem, which turns it into a one-sample test against a theoretical distribution. (This theorem was mentioned in class, but only in passing.)