DIP HANDOUT II: ADJUSTED DIP QUANTILES FOR LARGE n

If we call the non-adjusted dip quantiles $q(n, \alpha)$, given in the separate table for selected values of n between 4 and 5,000, then the adjusted quantiles are $Q(\alpha, n) := \sqrt{n}q(n, \alpha)$. These are shown in a separate file for the given values of n. Somewhat similarly as for the one-sample Kolmogorov statistic, the adjusted quantiles will converge as $n \to \infty$ to finite, non-zero limits, by a theorem of the Hartigans. So we may be able to predict them by regression on $1/\sqrt{n}$, $\sqrt{n}q(n, \alpha) \doteq \hat{q}(\alpha, n) := a - b/\sqrt{n}$, where a and b depend on α. In this case, we don't know the limiting value a (the exact intercept), so both $\widehat{a}$ and $\widehat{b}$ will be fitted by the regression. For $q = 0.95$, 0.99, and 0.999 such regressions were done based on the adjusted dip quantiles for $n = 500$, 1000, 2000, and 5000.

The results for the predicted adjusted quantile $\hat{q}(\alpha, n)$ where $\alpha = 1 - q$ were as follows:

$$\hat{q}(0.05, n) = 0.5459 - 0.3381/\sqrt{n},$$

with residuals for the n used being -0.0002, 0.0003, 0.0001, -0.0001;

$$\hat{q}(0.01, n) = 0.6335 - 0.432/\sqrt{n},$$

with residuals for the n used being -0.0003, 0.0003, 0.0008, -0.0007; and

$$\hat{q}(0.001, n) = 0.7392 - 0.5176/\sqrt{n},$$

with residuals for the n used being -0.0002, 0.0003, $-.0001$, -0.0001.

It's somewhat disturbing that in each case the residuals show a concave pattern. The regressions are not as successful as for the one-sample Kolmogorov–Smirnov statistic, where quantiles were actually computed, partly because there are actual sampling errors in the quantiles here, but perhaps also for other reasons.

Date: 18.465, Feb. 18, 2015.