MSAI Statistics Home Assignment 2

Problem 1. (3 points) Let $X_1, \dots, X_n \sim Be(p)$ and let $Y_1, \dots, Y_m \sim Be(q)$. Find the plug-in estimator, its bias and estimated standard error:

1. for p
2. for $p - q$

Problem 2. (3 points) Let $X_1, \dots, X_n$ be distinct observations (no ties). Prove that there are exactly $\binom{2n-1}{n}$ possible distinct bootstrap samples.

Problem 3. (4 points, computer experiment) Generate $n = 100$ observations from $\mathcal{N}(0, 1)$. Compute the 95% confidence band for the CDF $F(\cdot)$ using DKW inequality. Repeat this $m = 1000$ times and see how often the confidence band contains:

1. the true CDF
2. the ECDF

Problem 4. (2 bonus points) Find $\mathbb{P}\left(|\widehat{F}(x) - F(x)| > \frac{t}{\sqrt{n}}\right)$. Hint: remember that ECDF is unbiased. Hint 2: you can choose between a lot of statistical instruments for this one. To name a few: CLT, Chebyshev inequality, Chernoof inequality, DKW theorem.

Problem 5. (2 bonus points) In Kolmogorov's theorem, $F(\cdot)$ is required to be continuous. What is the limit of $D_n = \sqrt{n} \sup_x \left|\widehat{F}(x) - F(x)\right|$ if $X_1, \dots, X_n \sim Be(p)$ and $F(x)$ is Bernoulli CDF? Hint: CLT or its special case — de Moivre–Laplace theorem.