Quiz 6. Jack knife and Bootstrap

Name: Lisa Liubovich Attempt (circle one): BEFORE AFTER

Given 25 samples of water, 16 samples are polluted with arsenic (0) and 9 samples are clean (1). Thus, the proportion of clean samples, p, is estimated by $\hat{p} = 9/25 = 0.36$.

However, the standard approach for water sampling is to estimate $\theta = \sqrt{p}$. An obvious estimator of θ is $\hat{\theta} = \sqrt{p} = 0.6$. But this $\hat{\theta}$ is biased, i.e., $E(\hat{\theta}) = E(\sqrt{p}) \neq \sqrt{E(p)} = \sqrt{p} = \theta$.

Starting with $\hat{\theta} = \sqrt{\hat{p}}$, compute the jackknife estimator $\hat{\theta}_{JK}$.

(a) Delete one at a time and compute $\theta_{(-i)}$, $i = 1, 2, \ldots, 25$, of the obtained data. Note that, given the sample, 16 times the deleted sample is arsenic (0), 9 times the deleted sample is clean (1).

When a clean sample is removed (xi = 1)

$\hat{p}_{(-i)} = 8/24 = 1/3 \rightarrow \hat{\theta}_{(-i)} =$ square root of $1/3 = 0.57735026919$

When a polluted sample is removed (xi = 0)

$\hat{p}_{(-i)} = 9/24 = 3/8 \rightarrow \hat{\theta}_{(-i)} =$ square root of $3/8 = 0.61237243569$

$\hat{\theta}_{(.)} = (9*0.57735026919 + 16 * 0.61237243569)/25 = 14.9941113937/25 = 0.59976445575$

(b) Compute the Jackknife estimate θ_{JK}.

θ_{JK} = n*$\hat{\theta}$ - (n-1) $\hat{\theta}_{(.)} = 25 * 0.60 - 24 * 0.59976445575 = 0.605653062$

(c) Compute the estimated bias of θ.

Bias = $\hat{\theta}$ - $\theta_{JK} = 0.60 - 0.605653062 = -0.005653062$

(d) **Stat-627 only**. Use the following distribution to illustrate that $E(\sqrt{X}) \neq \sqrt{E(X)}$. Recall that: $E(X) = \sum(x \cdot P(X = x))$.

X	P(X = x)	x · P(X = x)	$\sqrt{X}$	$\sqrt{x} \cdot P(X = x)$

0	0.5	0 * 0.5 = 0	0	0
9	0.5	9* 0.5 = 4.5	3	3*0.5 = 1.5
		$E(X) = 0 + 4.5 = 4.5$ $\sqrt{E(X)} =$ 2.12		$E(\sqrt{X}) =$ 0 + 1.5 = 1.5