Asymptotic Mean Estimation

- Given $n$ samples from a distribution, want to estimate mean $\mu$.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Estimator</th>
<th>Converges to</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknown Distribution</td>
<td>Empirical Mean $\mathcal{N}(\mu, \sigma^2/n)$</td>
<td>Central Limit Theorem</td>
<td></td>
</tr>
<tr>
<td>Known Distribution</td>
<td>MLE $\mathcal{N}(\mu, \sigma^2/n)$</td>
<td>$Z$ is the Fisher Information</td>
<td></td>
</tr>
<tr>
<td>Unknown Symmetric Distribution</td>
<td>KDE + MLE $\mathcal{N}(\mu, \sigma^2/n)$</td>
<td>(Stone; 1975)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Classical Asymptotic Results

- In finite-sample setting, when distribution is unknown, [Catoni; 2012], [Lee, Valiant; 2022] show estimator $\hat{\mu}$ such that with probability $1 - \delta$,
  $$|\hat{\mu} - \mu| \leq \sqrt{\frac{2\log(1 + \delta)}{n}} (1 + o(1))$$
- Natural Question: What if distribution is known/symmetric?

Location Estimation (Known Distribution Case)

- Gaussian
- Laplace
- Mixture of Gaussians

Finite-Sample Mean Estimation

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Error</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknown Distribution</td>
<td>$\sqrt{\frac{2\log(1 + \delta)}{n}}$</td>
<td>[Cat12, LV22]</td>
</tr>
<tr>
<td>Known Distribution</td>
<td>$\sqrt{\frac{1}{n}}$</td>
<td>GlPV22, $Z$ is the smoothed Fisher Information</td>
</tr>
<tr>
<td>Unknown Symmetric Distribution</td>
<td>$\sqrt{\frac{2\log(1 + \delta)}{n}}$</td>
<td>This paper</td>
</tr>
</tbody>
</table>

Table 2. Finite-Sample Results

Main Theorem (Informal): For $r \approx \sigma/n^{1/3}$, our estimator $\hat{\mu}$ given $n$ samples from a symmetric distribution satisfies, with probability $1 - \delta$,
  $$|\hat{\mu} - \mu| \leq (1 + o(1)) \sqrt{\frac{2\log(1 + \delta)}{n^2r}}$$

Correcting the KDE – Bias

- Idea: Estimate density using the Kernel Density Estimate (KDE) on the first (say) $n^{1/10}$ samples

- Idea: Estimate density using the Kernel Density Estimate (KDE) on the first (say) $n^{1/10}$ samples

- Naive algorithm: Run (smoothed) MLE on the KDE/Find zero of KDE score using remaining samples. Two issues:
  1. Bias
  2. Variance

- Idea: Estimate density using the Kernel Density Estimate (KDE) on the first (say) $n^{1/10}$ samples

- Naive algorithm: Run (smoothed) MLE on the KDE/Find zero of KDE score using remaining samples. Two issues:
  1. Bias
  2. Variance

- Idea: Estimate density using the Kernel Density Estimate (KDE) on the first (say) $n^{1/10}$ samples

- Naive algorithm: Run (smoothed) MLE on the KDE/Find zero of KDE score using remaining samples. Two issues:
  1. Bias
  2. Variance

Summary

- For a symmetric distribution, MLE with respect to any (possibly different) symmetric distribution is an unbiased estimator
- Idea: Anti-symmetrize the KDE score.

- Idea: Anti-symmetrize the KDE score.

- Idea: Anti-symmetrize the KDE score.

- Idea: Anti-symmetrize the KDE score.

- Idea: Anti-symmetrize the KDE score.

- Idea: Anti-symmetrize the KDE score.

- Idea: Anti-symmetrize the KDE score.

- Idea: Anti-symmetrize the KDE score.