Empirical Analysis of Adaptive Kernel Density Estimation in Overlap Estimation

Michelle Chao Chen [1]   Rickard Karlsson [1]   Jesse H. Krijthe [1]

[1]Delft University of Technology

Background

Given data from an observational study or a randomized experiment, the **positivity** assumption must hold in order to draw causal relations between the treatment and outcome [1]. **Overlap estimation** can be used to detect violations of the assumption. We present automatic tools for overlap estimation and empirically analyse different kernel density estimation (KDE) methods.

Main Question: How do adaptive kernel density estimation methods compare to the classical kernel density estimation method in estimating overlap?

Method

Kernel Density Estimation

1. Standard KDE: Selects kernel bandwidth h using the sample variance (Silverman).
2. Adaptive KDE (aKDE): Selects kernel bandwidth h based on local density [2].
3. Variable KDE (vKDE): Selects kernel bandwidth h based on distance to neighbours.

Metrics

1. Intersection-over-Union: Similarity between estimated overlap region and true overlap region.
2. FPR & TPR: Demonstrates the tradeoff between FP and TP for different distributions.
3. MISE: Global error of the density estimate.

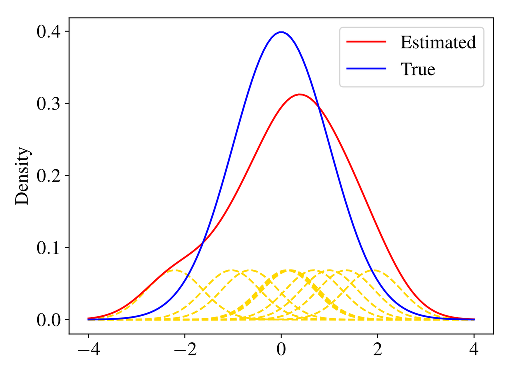


Figure 1. Standard KDE

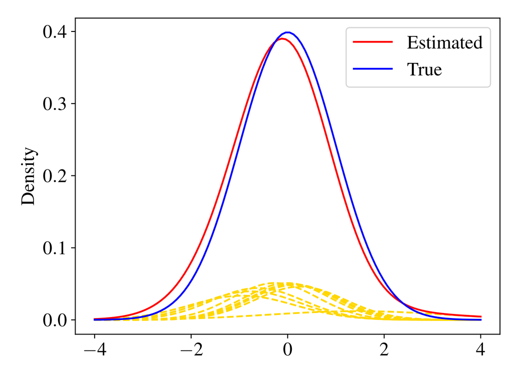


Figure 2. Adaptive KDE

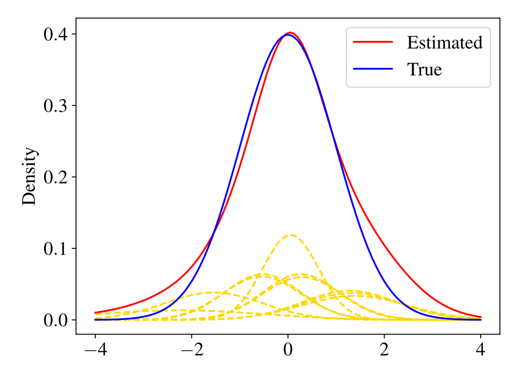
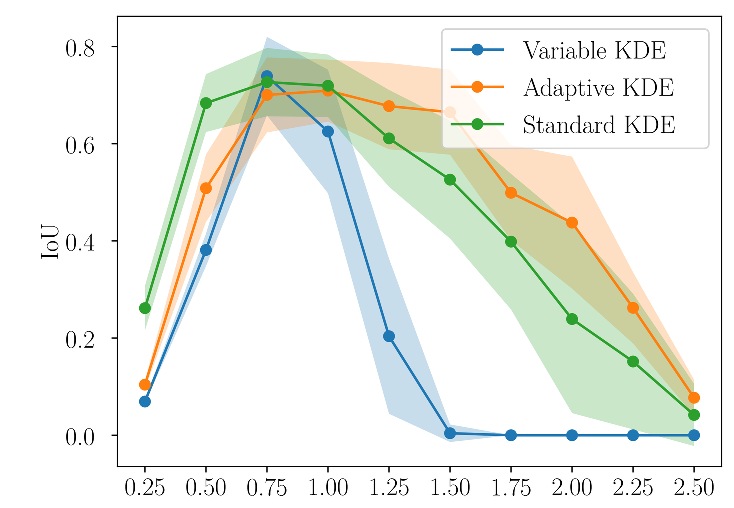


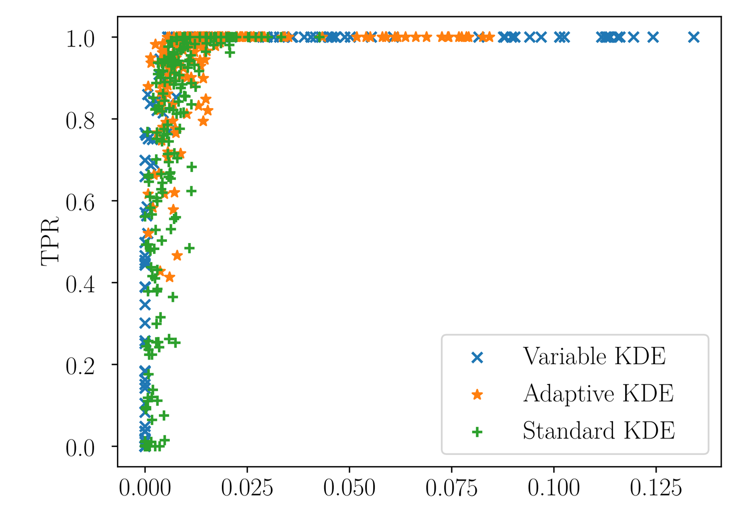
Figure 3. Variable KDE

Results

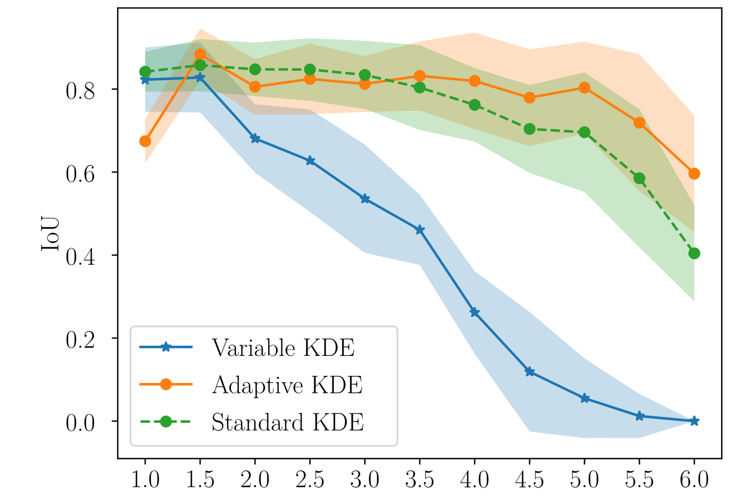
1. Standard KDE, vKDE, and aKDE are compared over different synthetic and natural datasets where the IoU, FPR, TPR, and MISE are measured for each method.
2. IoU of all KDE methods drop as variance in the dataset increases as density estimation in local regions become less accurate. MISE does not capture this as it is a global measure.
3. While aKDE and standard KDE performs similarly in most settings, vKDE performs significantly worse when σ grows.



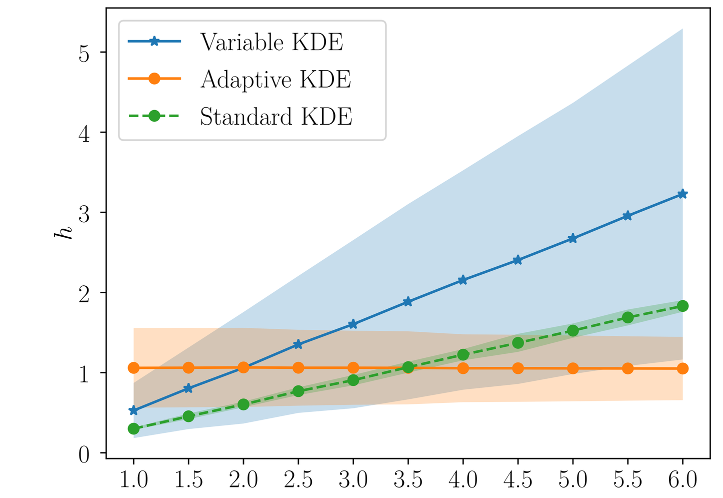
(a) IoU for $\mathcal{N}([0,0],\sigma I)$ and $\mathcal{N}([0,2],\sigma I)$



(b) ROC for $\mathcal{N}([0,0],\sigma I)$ and $\mathcal{N}([0,2],\sigma I)$



(c) IoU for $\mathcal{N}(0,\sigma^2)$ and $\mathcal{N}(3,\sigma^2)$

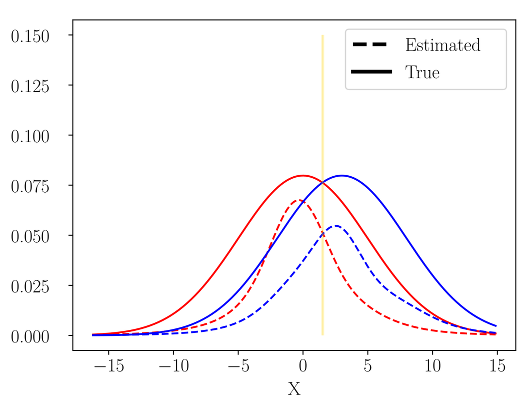


(d) Kernel for $\mathcal{N}(0,\sigma^2)$ and $\mathcal{N}(3,\sigma^2)$

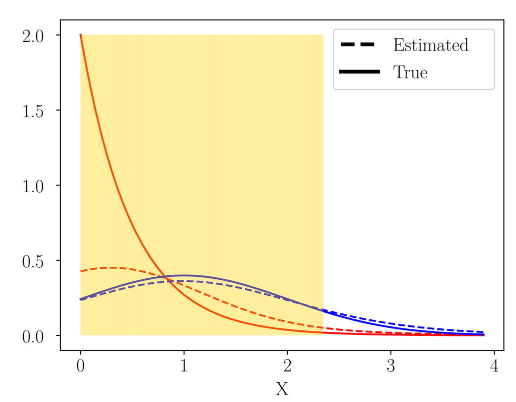
Figure 4. Performance of KDE methods in estimating overlap in 1D and 2D using $n = 200$ samples

Conclusions

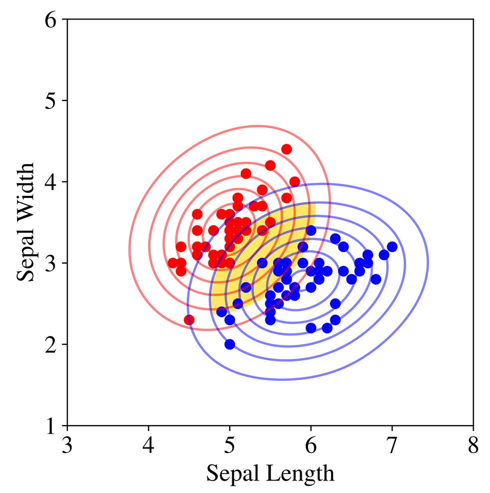
- Standard KDE is recommended in general, and especially when false positives are considered problematic
- aKDE is recommended when the variance of the distribution is high and when sample size is small, as aKDE has the advantage of producing smooth estimates at small sample sizes.
- vKDE is generally not recommended, because without tuning k, which is used to determine the scaling parameter $d_{j,k}$ of a kernel, the density estimate by vKDE can be erroneous.



(a) vKDE (IoU = 0.02)



(b) aKDE (FPR = 0.24)



(c) Overlap (yellow) between Setosa and Versicolor

Figure 5. Cases where each of the methods fail are shown.

Future Work

1. Varying the overlap threshold ϵ.
2. Analysis of KDE methods for higher dimensions and multi-modal distributions.

References

[1] Robins JM Hernán MA. *Causal Inference: What If.* Chapman & Hall/CRC, Boca Raton, 2023.

[2] B. W. Silverman. *Density Estimation for Statistics and Data Analysis.* Chapman & Hall, London, 1986.

C.Chen-18@student.tudelft.nl