Johns Hopkins Engineering

625.464 Computational Statistics

Kernel Estimators : Choice of Bandwidth part 2

Module 11 Lecture 11D

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Choice of Bandwidth h

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right)$$

- h is too small $\Rightarrow$ $\hat{f}$ wiggly / false modes and high variance.
- h is too big $\Rightarrow$ lose features / higher bias.

Method 2 : Plug-in methods

Recall

$$\text{Opt } h = \left(\frac{R(K)}{n \sigma_K^k R(f'')} \right)^{1/5}$$

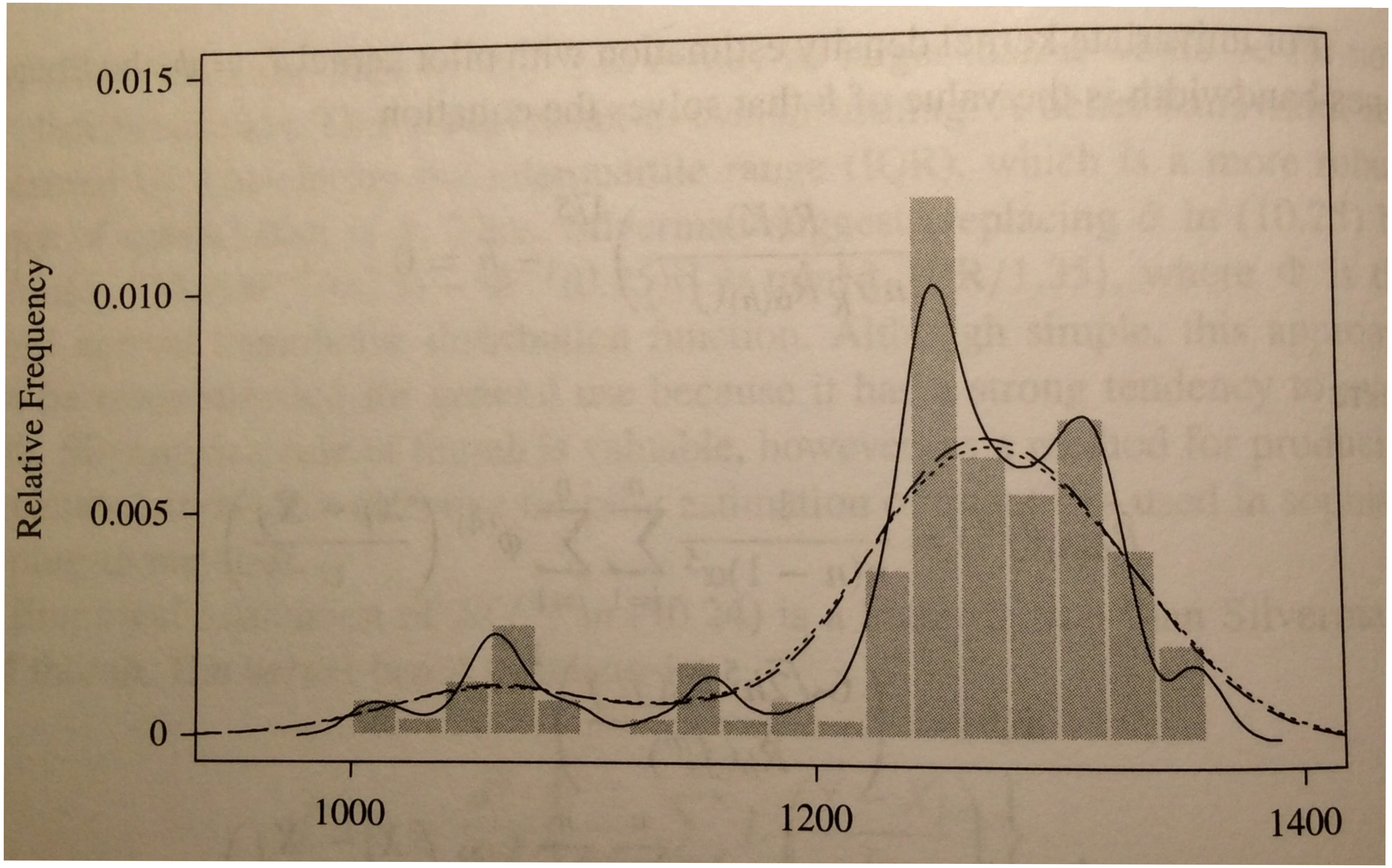
Option 1 : Silverman's rule Replace $R(f'')$ w/ $R(\phi'')$ $\frac{}{\sigma^5}$

- assume f is normal w/ var = sample variance
- Then solve for $h = \left(\frac{4}{3n}\right)^{1/5} \hat{\sigma}$ and find $\hat{f}$.

Option 2 Sheather-Jones Method Replace $R(f'')$ w/ $R(\hat{f}'')$

- choose pilot h_0 and Kernel L
- estimate $\hat{f}''(x) = \frac{d^2}{dx^2}\left(\frac{1}{nh_0}\sum_{i=1}^{n} L\left(\frac{x-x_i}{h_0}\right)\right) = \frac{1}{nh_0^3}\sum_{i=1}^{n} L''\left(\frac{x-x_i}{x_0}\right)$ $h_0 \to h$
- compute opt $\hat{h}$ based on $\hat{f}''$ - estimate $\hat{f}$ using $\hat{h}$

Plug-in Method Example



S–J h = 10.22 Silverman's h = 32.96 Terrel h = 35.60

Method 3: (Terrell's) Maximal Smoothing Principle

Replace $R(f'')$ w/ most cons. value

- Terrell looked at a collection of h that would minimize opt B.W. for various f.

$$h = 3\left(\frac{R(K)}{35n}\right)^{1/5}\hat{\sigma}$$

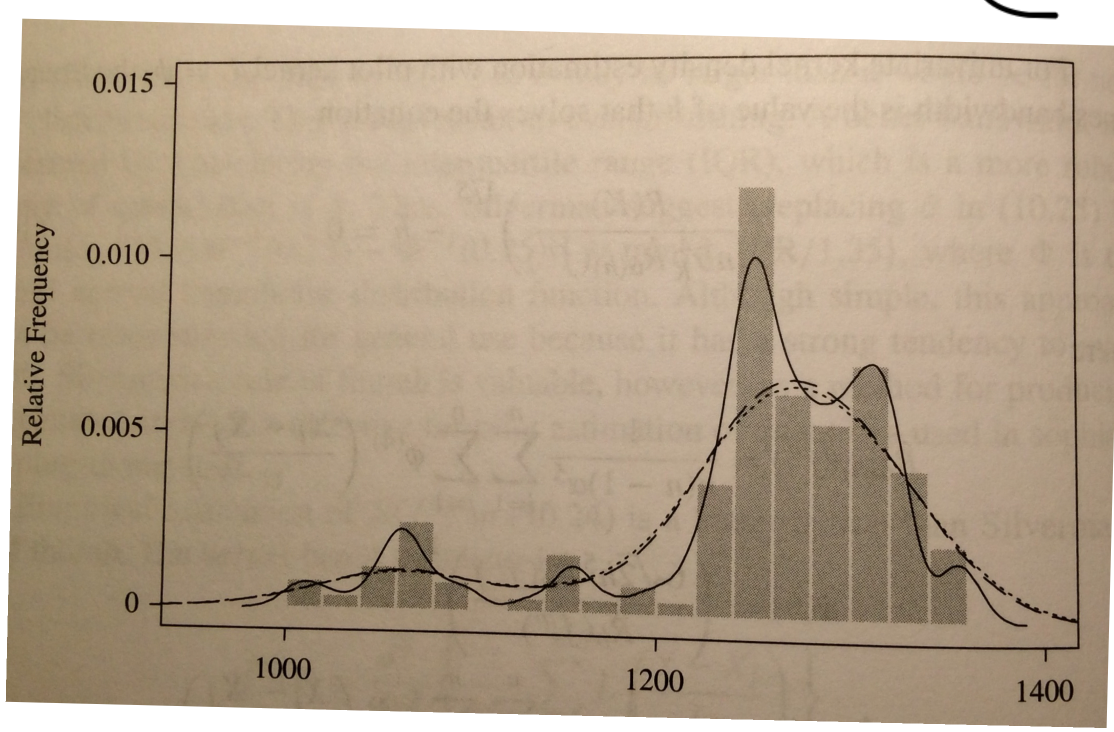


Table 10.1 pg 339.