(10.1a) $X = \log(f|2)$ $\quad x_1, \dots, x_{603}$

normal KDE for X ; Silverman, S-J, Terrell

Silverman :

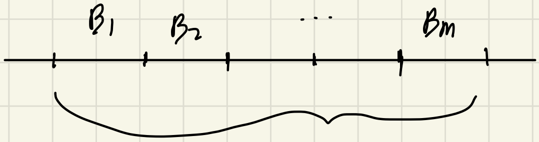
$$\rightarrow h_{opt} = \left(\frac{R(K)}{n \sigma_K^4 R(f'')} \right)^{1/5}$$

$\Rightarrow$ replace $R(f'')$ w/ $\dfrac{R(\phi'')}{\hat{\sigma}^5}$ $\rightarrow$ $h = \left(\frac{4}{3n}\right)^{1/5} \hat{\sigma}$

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

* custom histogram? does # bins relate to data type?
e.g. bin ranges should match data's measurements

$$\hat{f}_n(x) = \sum_{k=1}^{m} \frac{\hat{p}_k}{V_k} I(x \in B_k)$$



$D = \bigcup_{k=1}^{m} B_k$

$$n_k = \sum_{i=1}^{m} I(x_i \in B_k)$$

proportion of obs. in B_k is $\hat{p}_k = \dfrac{n_k}{n}$

V_k : volume (length) of bin B_k

* must determine m through visual inspection

S-J : $\hat{f}''(x) = \frac{d^2}{dx^2}\left(\frac{1}{nh_0}\sum_{i=1}^{n} L\left(\frac{x-x_i}{h_0}\right)\right) = \frac{1}{nh_0^3}\sum_{i=1}^{n} L''\left(\frac{x-x_i}{h_0}\right)$

$h_{opt} = \left(\frac{R(K)}{n\sigma_K^4 R(f'')}\right)^{1/5}$

- initial est. for f, $\hat{f}$
- calculate its $\hat{f}''$
  - $R(\hat{f}'')$
- plug in to h_{opt} or $\hat{h}$
- est. $\hat{f}$ using $\hat{h}$

σ_K^4 ?

$h_0 > h$

$R(g) = \int g(x)^2 dx$

Normal : $K(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}}$; $R(K) = \frac{1}{2\sqrt{\pi}}$

Use Monte-Carlo Integration to est. the roughness of $f''(x)$

$$K'(z) = \frac{\partial}{\partial z}\frac{e^{-z^2/2}}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}}\frac{\partial}{\partial z}e^{-z^2/2} = \frac{-z}{\sqrt{2\pi}}e^{-z^2/2}$$

$$\frac{\partial}{\partial x} = e^u \frac{\partial u}{\partial x} \rightarrow e^{-z^2/2}\frac{\partial}{\partial z}\left(-\frac{z^2}{2}\right) = -ze^{-z^2/2}$$

$$K''(z) = \frac{\partial}{\partial z}\frac{-z}{\sqrt{2\pi}}e^{-\frac{z^2}{2}} = \frac{-1}{\sqrt{2\pi}}\left[e^{-z^2/2} + \left(-z^2 e^{-z^2/2}\right)\right]$$

$$= \frac{-1}{\sqrt{2\pi}}e^{-z^2/2}\left(1-z^2\right)$$

$$R(\hat{f}'') = \int \hat{f}''(x)^2\, dx = \int \left[\frac{1}{nh_0^3} \sum_{i=1}^{n} L''\left(\frac{x-x_i}{h_0}\right)\right]^2 dx$$

$$= \frac{1}{n^2h_0^6} \int \left[\sum_{i=1}^{n} L''\left(\frac{x-x_i}{h_0}\right)\right]^2 dx = \frac{1}{n^2h_0^6} \int \sum_{i=1}^{n} L''\left(\frac{x-x_i}{h_0}\right)^2 + \sum_{\substack{i,j=1\\ i\neq j}}^{n} L''\left(\frac{x-x_i}{h_0}\right) L''\left(\frac{x-x_j}{h_0}\right) dx$$

$$= \frac{1}{n^2h_0^6} \Bigg\{ \int \left(\frac{-1}{\sqrt{2\pi}}\right)^2 \sum_{i=1}^{n} \left(\exp\left[\frac{-1}{2}\left(\frac{x-x_i}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_i}{h_0}\right)\right)\right)^2 +$$

$$\frac{1}{2\pi} \sum_{\substack{i,j=1\\ i\neq j}}^{n} \left(\exp\left[\frac{-1}{2}\left(\frac{x-x_i}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_i}{h_0}\right)\right)\right)\left(\exp\left[\frac{-1}{2}\left(\frac{x-x_j}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_j}{h_0}\right)\right)\right)$$

$$dx \Bigg\}$$

orthogonal? $\qquad \frac{-1}{\sqrt{2\pi}}\, e^{-z^2/2}\,(1-z^2)$ $\quad$ X

$$\langle q_i, q_j \rangle = \int_D q_i(x)\, q_j(x)\, w(x)\, dx = \begin{cases} 0 & i\neq j \\ \lambda_i \geq 0 & i=j \end{cases}$$

$$\Rightarrow \frac{1}{n^2h_0^6} \Bigg\{ \left(\frac{-1}{\sqrt{2\pi}}\right)^2 \int \sum_{i=1}^{n} \left(\exp\left[\frac{-1}{2}\left(\frac{x-x_i}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_i}{h_0}\right)\right)\right)^2 dx +$$

$$\frac{1}{2\pi} \int \sum_{\substack{i,j=1\\ i\neq j}}^{n} \left(\exp\left[\frac{-1}{2}\left(\frac{x-x_i}{h_0}\right)\right]\left(1-\left(\frac{x-x_i}{h_0}\right)\right) \exp\left[\frac{-1}{2}\left(\frac{x-x_j}{h_0}\right)\right]\left(1-\left(\frac{x-x_j}{h_0}\right)\right)\right) dx \Bigg\}$$

$$\Rightarrow \frac{1}{n^2h_0^6} \Bigg\{ \left(\frac{-1}{\sqrt{2\pi}}\right)^2 \sum_{i=1}^{n} \|L''\|^2 +$$

$$\frac{1}{2\pi} \sum_{\substack{i,j=1\\ i\neq j}}^{n} \int \left(\exp\left[\frac{-1}{2}\left(\frac{x-x_i}{h_0}\right)\right]\left(1-\left(\frac{x-x_i}{h_0}\right)\right) \exp\left[\frac{-1}{2}\left(\frac{x-x_j}{h_0}\right)\right]\left(1-\left(\frac{x-x_j}{h_0}\right)\right)\right) dx \Bigg\}$$

$$= \frac{1}{2\pi n^2 h_0^6} \left\{ \underbrace{\int \sum_{i=1}^{n} \left(\exp\left[-\frac{1}{2}\left(\frac{x-x_i}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_i}{h_0}\right)^2\right)\right)^2 dx}_{A} + \right.$$

$$\left. \underbrace{\int \sum_{\substack{i,j=1\\ i\neq j}}^{n} \left(\exp\left[-\frac{1}{2}\left(\frac{x-x_i}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_i}{h_0}\right)^2\right) \exp\left[-\frac{1}{2}\left(\frac{x-x_j}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_j}{h_0}\right)^2\right) dx}_{B} \right\}$$

A:

$$\int \sum_{i=1}^{n} \left(\exp\left[-\frac{1}{2}\left(\frac{x-x_i}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_i}{h_0}\right)\right)\right)^2 dx$$

$$= \sum_{i=1}^{n} \int \left\{ \exp\left[-\frac{1}{2}\left(\frac{x-x_i}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_i}{h_0}\right)^2\right)\right\}^2 dx$$

$$= \sum_{i=1}^{n} \frac{3\sqrt{\pi}}{4} \cdot h_0 = \frac{3}{4} n h_0 \sqrt{\pi}$$

B:

$$\int \sum_{\substack{i,j=1\\ i\neq j}}^{n} \left(\exp\left[-\frac{1}{2}\left(\frac{x-x_i}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_i}{h_0}\right)\right) \exp\left[-\frac{1}{2}\left(\frac{x-x_j}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_j}{h_0}\right)\right) dx$$

$$= \sum_{\substack{i,j=1\\ i\neq j}}^{n} \int \left(\exp\left[-\frac{1}{2}\left(\frac{x-x_i}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_i}{h_0}\right)\right) \exp\left[-\frac{1}{2}\left(\frac{x-x_j}{h_0}\right)^2\right]\left(1-\left(\frac{x-x_j}{h_0}\right)\right) dx$$

$$= \sum_{\substack{i,j=1\\ i\neq j}}^{n} \frac{1}{16h_0^3} \sqrt{\pi} \left(x_j^4 - 4x_i x_j^3 + (6x_i^2 - 12h_0^2)x_j^2 + (24h_0^2 x_i - 4x_i^3)x_j + x_i^4 - 12h_0^2 x_i^2 + 12h_0^4\right) \exp\left[-\frac{1}{4h_0^2}\left(x_j^2 - 2x_i x_j + x_i^2\right)\right]$$

$$\Rightarrow \frac{1}{2\pi n^2 h_0^6} \left\{ \frac{3}{4} n h_0 \sqrt{\pi} + \left[\sum_{\substack{i,j=1\\ i\neq j}}^{n} \frac{1}{16h_0^3} \sqrt{\pi} \left(x_j^4 - 4x_i x_j^3 + (6x_i^2 - 12h_0^2)x_j^2 + (24h_0^2 x_i - 4x_i^3)x_j + x_i^4 - 12h_0^2 x_i^2 + 12h_0^4\right) \exp\left[-\frac{1}{4h_0^2}\left(x_j^2 - 2x_i x_j + x_i^2\right)\right] \right] \right\}$$

σ_K^2 : variance of K $\quad V[\hat{f}(x)] = E[(\hat{f}(x) - E[\hat{f}(x)])^2]$

$$V[K(z)] = E[(K(z) - E[K(z)])^2]$$

$$= E\{K(z)^2 - 2K(z)E(K(z)) + E[K(z)]^2\}$$

$$= E(K(z)^2) - 2E(K(z))E(K(z)) + E(K(z))^2$$

$$K(z)^2 = \left(\frac{e^{-z^2/2}}{\sqrt{2\pi}}\right)^2 = \frac{1}{2\pi}e^{-z^2}$$

$$\Rightarrow E(K(z)^2) = E\left(\frac{1}{2\pi}e^{-z^2}\right) = \frac{1}{2\pi}E(e^{-z^2})$$

$\sigma_K^2 = 1$, $\sigma_K^4 = (1)^2 = 1$

$$h_{opt} = \left(\frac{R(K)}{n\sigma_K^4 R(f'')}\right)^{1/5} = \left(\frac{1}{2\sqrt{\pi}} \cdot \frac{1}{n(1)} \cdot \frac{1}{R(\hat{f}'')}\right)^{1/5}$$

Terrell : $h = 3\left(\frac{R(K)}{35n}\right)^{1/5}\hat{\sigma}$, $R(K) = \frac{1}{2\sqrt{\pi}}$

$$\cdots = 3\left(\frac{1}{70\sqrt{\pi}\,n}\right)^{1/5}\hat{\sigma}$$

(10.1b) uniform, Epanechnikov, triweight

$$R(\hat{f}'') = \int \hat{f}''(x)^2\,dx = \int\left[\frac{1}{nh_0^3}\sum_{i=1}^{n} L''\left(\frac{x-x_i}{h_0}\right)\right]^2 dx$$

$$= \frac{1}{nh_0^6}\int\left(\sum_{i=1}^{n} L''\left(\frac{x-x_i}{h_0}\right)\right)^2 dx$$

$h_{S\text{-}J}$

$$= \frac{1}{nh_0^6}\left\{\int\sum_{i=1}^{n} L''\left(\frac{x-x_i}{h_0}\right)^2 dx + \int\sum_{\substack{i,j=1\\ i\neq j}}^{n} L''\left(\frac{x-x_i}{h_0}\right)L''\left(\frac{x-x_j}{h_0}\right)dx\right\}$$

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

$I\{|z|<1\}$

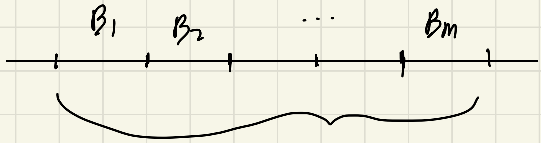
uniform : $K(z) = \frac{1}{2}$

Epanechnikov : $K(z) = \frac{3}{4}(1-z^2)$

triweight : $\frac{35}{32}(1-z^2)^3$

Histogram estimator

$$\hat{f}_n(x) = \sum_{k=1}^{m} \frac{\hat{p}_k}{V_k} I(x \in B_k)$$

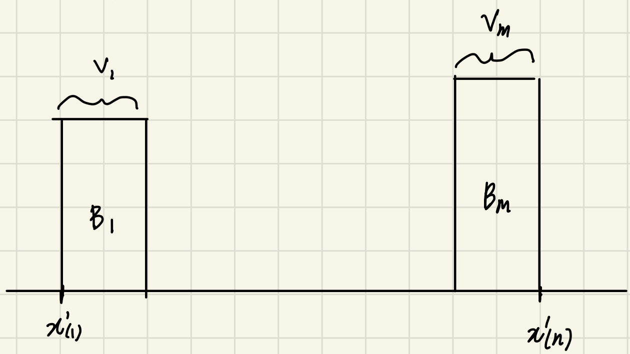


$$D = \cup_{k=1}^{m} B_k$$

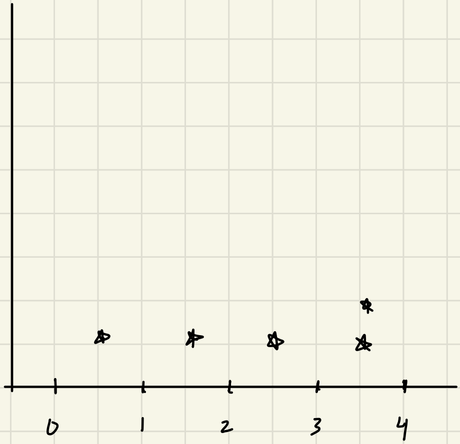
$$n_k = \sum_{i=1}^{m} I(x_i \in B_k)$$

proportion of obs. in B_k is $\hat{p}_k = \frac{n_k}{n}$

V_k : volume (length) of bin B_k



$$\hat{f}_n(x) = \begin{cases} \hat{p}_1 / V_1 & x \in B_1 \\ \hat{p}_2 / V_2 & x \in B_2 \\ \vdots & \\ \hat{p}_m / V_m & x \in B_m \end{cases}$$



$n = 5$

$x_1 = 0.5$
$x_2 = 1.5$
$x_3 = 2.5$
$x_4 = 3.5$
$x_5 = 3.6$

$$\hat{f}_n(x) = \begin{cases} \hat{p}_1/v_1 & x \in B_1 \\ \hat{p}_2/v_2 & x \in B_2 \\ \hat{p}_3/v_3 & x \in B_3 \\ \hat{p}_4/v_4 & x \in B_4 \end{cases}$$

$$\hat{p}_k = \frac{n_k}{n}$$

"$x_3 \in B_3$"

$$n_k = \sum_{i=1}^{m} I(x_i \in B_k)$$

$B_1 = [\overset{1}{x'_{(1)}}, \; x'_{(1)} + \overset{2}{v_1})$ [ ] $\in$ 3, 4

$B_2 = [\overset{2}{x'_{(1)}} + v_1, \; x'_{(1)} + 2\overset{3}{v_1})$ ( ]

$B_3 = [\overset{3}{x'_{(1)}} + 2v_1, \; x'_{(1)} + 3\overset{4}{v_1})$ ( ]

$B_4 = [\overset{4}{x'_{(1)}} + 3v_1, \; x'_{(1)} + 4\overset{5}{v_1}]$ ( ]

$\hat{p}_1 = \hat{p}_2 = \hat{p}_3 = 1$

$\hat{p}_4 = 2$

$$v_1 = \cdots = v_4 = (x'_{(4)} - x'_{(1)})/m$$
$$= (3.6 - 0.5)/4 = 3.1/4$$