Nonparametric Methods

Regression $f(x) = w^T \cdot x + w_0$

Classification $\delta(w^T x + w_0) = \begin{cases} 1 & \text{if } w^T x + w_0 > 0 \\ 0 & \text{otherwise} \end{cases}$

positive class

negative class

Density estimation $N(\mu, \sigma^2)$

$N(\mu, \Sigma)$

similar inputs $\Rightarrow$ similar outputs

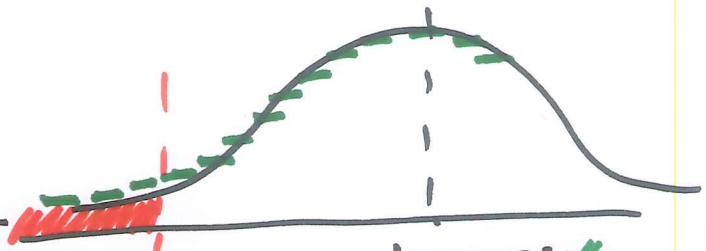
How do we measure similarity?

no parametric form $\Rightarrow$ "data-dependent"

"local models" (neighborhood)

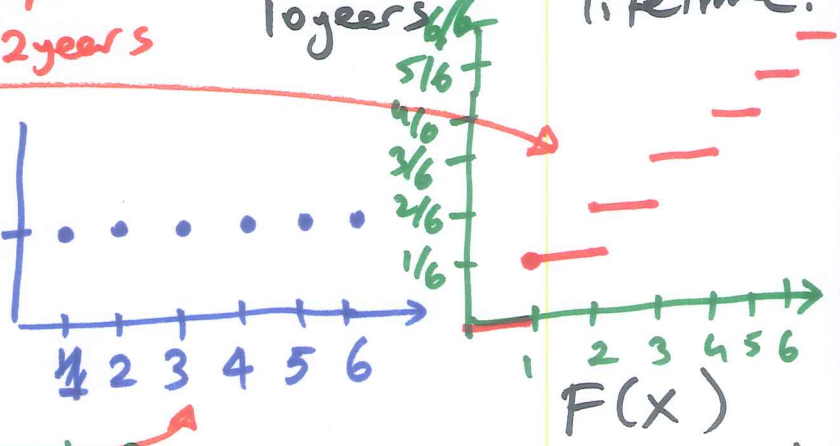
Nonparametric Density Estimation:

$\Pr$(Your product will fail before two years)



cumulative distribution function = ?

density function = ?



$F(x=a) = P(x \le a)$

counting function

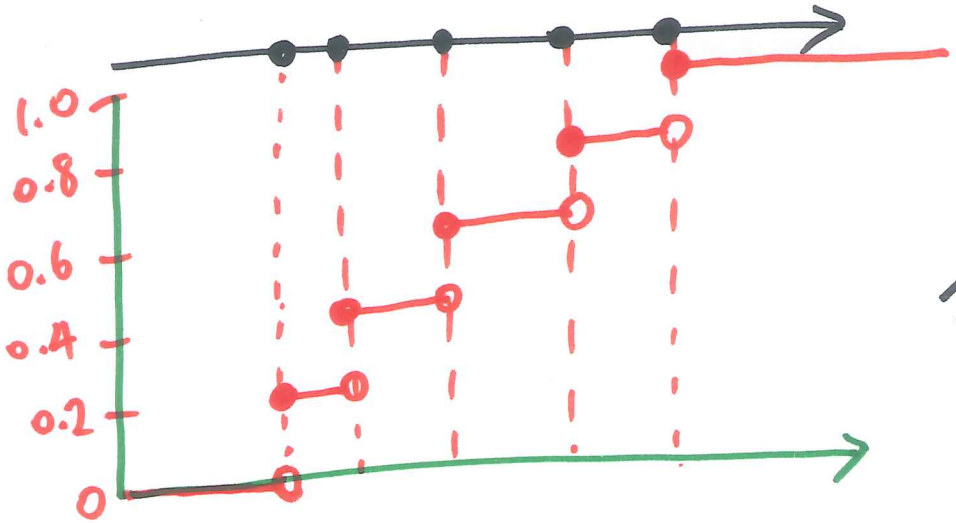
$$\hat{F}(x) = \frac{\#\{x_i \le x\}}{N} = \frac{\sum_{i=1}^{N} 1(x_i \le x)}{N}$$

$N\,\hat{F}(x+h)$ $\qquad N\,\hat{F}(x)$

$$\hat{p}(x) = \frac{1}{h}\left[\frac{\#\{x_i \le x+h\} - \#\{x_i \le x\}}{N}\right]$$

$$= \frac{1}{h}\left[\hat{F}(x+h) - \hat{F}(x)\right]$$

h → bin width



$$\int_{-\infty}^{x} p(x)\,dx \qquad \sum_{-\infty}^{x} p(x)$$

Histogram Estimator

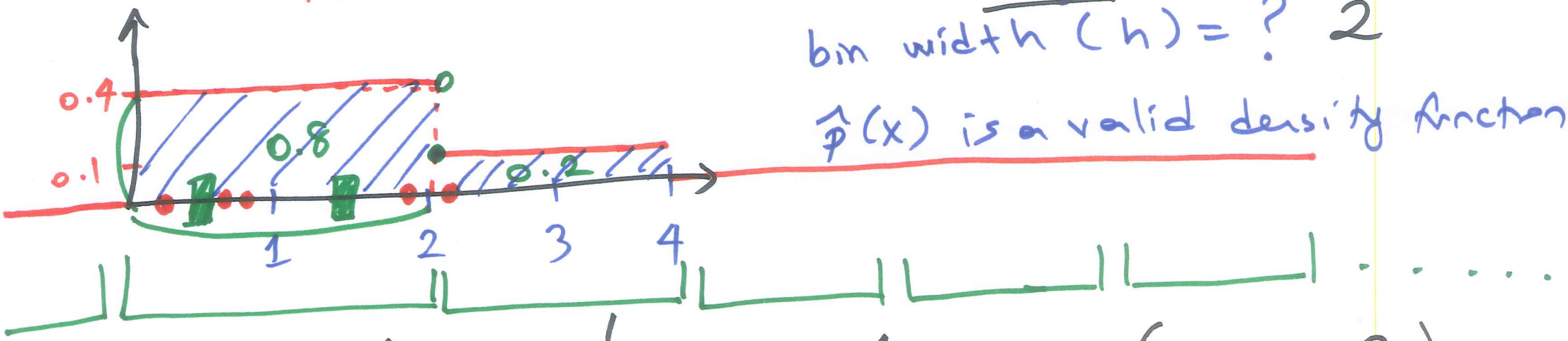
$$\hat{p}(x) = \frac{\#\{x_i \text{ in the same bin as } x\}}{N \cdot h}$$

$$\int_{-\infty}^{+\infty} p(x)\,dx = 1$$

origin = ? 0

bin width (h) = ? 2

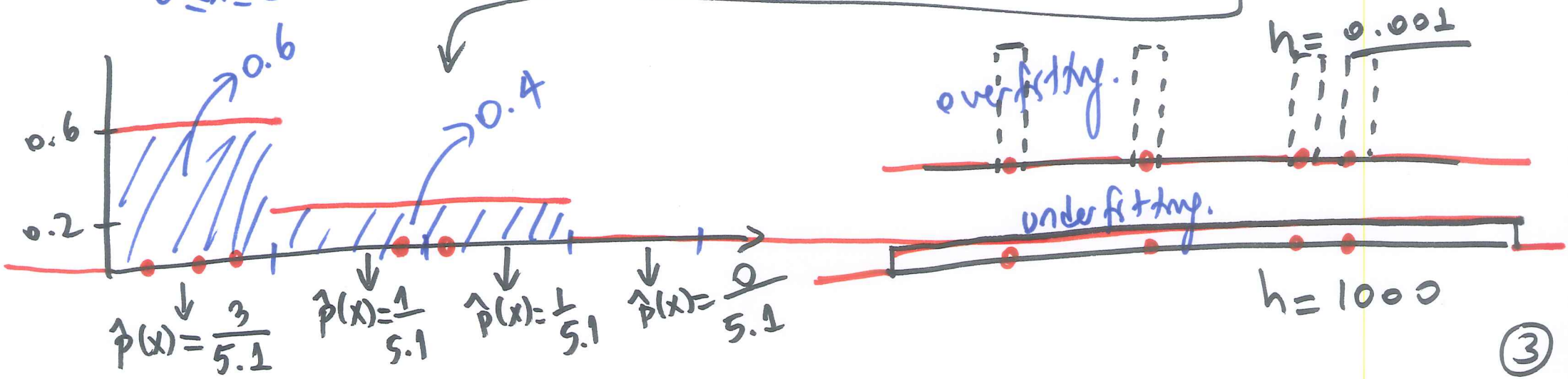
$\hat{p}(x)$ is a valid density function



$$\hat{p}(x) = \frac{4}{5.2} = 0.4 \qquad \hat{p}(x) = \frac{1}{5.2} = 0.1$$

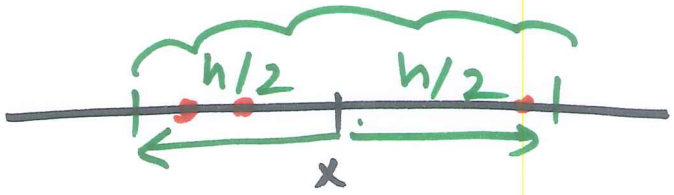
$0 \le x \le 2$

(origin = 0, h = 1)



Naive Estimator

$$\hat{p}(x) = \frac{\#\{x - h/2 < x_i \leq x + h/2\}}{N \cdot h}$$

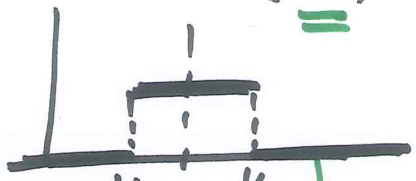


$$\int_{-\infty}^{+\infty} \hat{p}(x)\,dx = 1 \qquad \hat{p}(x) \geq 0 \quad \forall x$$

$$= \frac{1}{Nh} \sum_{i=1}^{N} w\left(\frac{x - x_i}{h}\right)$$

weight

$$w(u) = \begin{cases} 1 & \text{if } |u| \leq 1/2 \\ 0 & \text{otherwise} \end{cases}$$



$$\hat{p}(10) =$$

$$E[x] = \int p(x)\, x\, dx$$

$$\Pr(10 < x < 20)$$

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

$$p(10) =$$

$$p(u) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(u-\mu)^2}{2\sigma^2}\right]$$

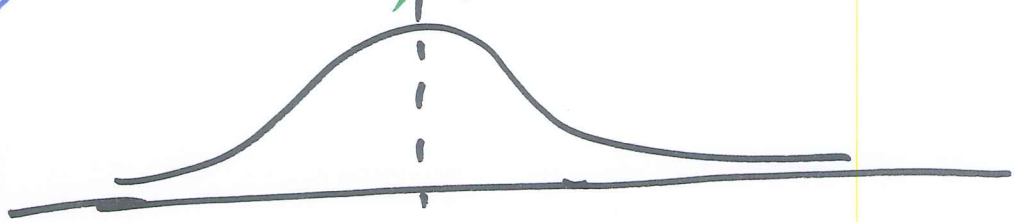
$(u-\mu)^2 \to 0$, $2\sigma^2 \equiv 1$

Kernel Estimator (PARZEN WINDOWS)

$$w(u) = K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$$

$\sigma = 1$, $\mu = 0$

$$\hat{p}(x) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right)$$



"k-nearest neighbor estimator"

$$\hat{p}(x) = \frac{k}{2N d_k(x)} = \frac{k}{N \cdot \underbrace{2 d_k(x)}_{h(x)}}$$

$d_k(x)$: the distance to kth nearest neighbor.