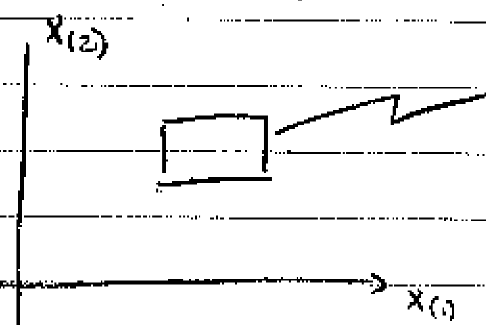
# ECE 544
## 9/10/2015

**TODAY**

1. HISTOGRAM, PARZEN, KNN
2. BIAS, VARIANCE, KLD

## NON-PARAMETRIC ESTIMATION IN GENERAL




① Choose some region $R$ w/ VOLUME $V$

$$P = \int_R p(\vec{x}')\, d\vec{x}' = \text{PROBABILITY ANY } \vec{x} \text{ FALLS INSIDE } R$$

② $N$ = # DATA OBSERVED

$K$ = # DATA THAT FALL INTO $R$

$$K \sim \text{binom}(N, P)$$

$$p_K(k) = \begin{cases} \binom{N}{k} P^k (1-P)^{N-k} & 0 \le k \le N \\ 0 & \text{ELSE} \end{cases}$$

$$E[K] = NP, \quad Var(K) = NP(1-P)$$

$$\boxed{\hat{P} = \frac{K}{N}} \quad E[\hat{P}] = P, \quad Var(\hat{P}) = \frac{P(1-P)}{N}$$

(3) ASSUME $p(\vec{x})$ UNIFORM INSIDE $R$,

SO $$\boxed{\tilde{p}(\vec{x}) = \frac{\hat{P}}{V} = \frac{K}{NV}}$$

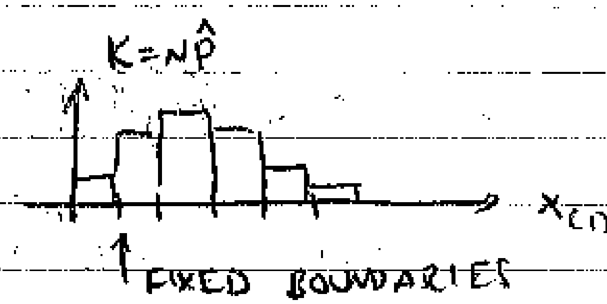
$$E\left[\tilde{p}(\vec{x})\right] = \frac{P}{V} = p(x) \quad \text{IF UNIFORM IN } R$$

$$Var\left(\tilde{p}(\vec{x})\right) = \frac{P(1-P)}{NV} \approx p(x)\left(\frac{1-P}{N}\right) \xrightarrow[N\to\infty]{} 0$$

NON-PARAMETRIC ESTIMATORS

| | | $R$: $V$ | CENTERED AT | $K$ |
|---|---|---|---|---|
| I. | HISTOGRAM | FIXED | FIXED LOCATION | RANDOM |
| II. | PARZEN WINDOWS | FIXED $h^D$ | TRAINING DATUM | RANDOM |
| III. | K-NEAREST NEIGHBORS | RANDOM | TEST DATUM | FIXED |

I HISTOGRAM




ADVANTAGE: EASY TO ANALYZE; $K \sim \text{binom}(N, P)$

DISADVANTAGE: INFLEXIBLE

## II PARZEN WINDOWS

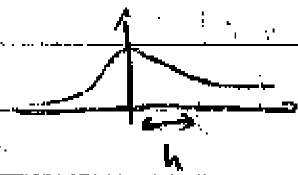
ⓐ CHOOSE $H(\vec{x})$ s.t. $\begin{cases} H(\vec{x}) \geq 0 & \forall \vec{x} \\ \int H(\vec{x})\,dx = 1 \end{cases}$

ⓑ SCALE $\frac{1}{h^D} H\left(\frac{\vec{x}}{h}\right)$ BY FIXED VOLUME $V = h^D$

COMMON CHOICES

① HYPERCUBE $\frac{1}{h^D} H\left(\frac{\vec{x}}{h}\right) = \begin{cases} \frac{1}{h^D} & \max_d |x_{(d)}| \leq \frac{h}{2} \\ 0 & \text{ELSE} \end{cases}$


② GAUSSIAN $\frac{1}{h^D} H\left(\frac{\vec{x}}{h}\right) = \frac{1}{(2\pi h^2)^{D/2}} e^{-\frac{1}{2}\left\|\frac{\vec{x}}{h}\right\|_2^2}$




ⓒ PUT A KERNEL AT EACH TRAINING POINT

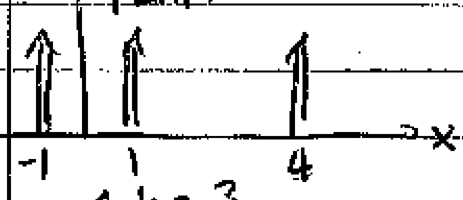
$$\hat{p}(\vec{x}) = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{h^D} H\left(\frac{\vec{x} - \vec{x}^n}{h}\right) = p_{EMP}(\vec{x}) * \frac{1}{h^D} H\left(\frac{\vec{x}}{h}\right)$$

$$p_{EMP}(\vec{x}) = \frac{1}{N} \sum_{n=1}^{N} \delta(\vec{x} - \vec{x}^n)$$

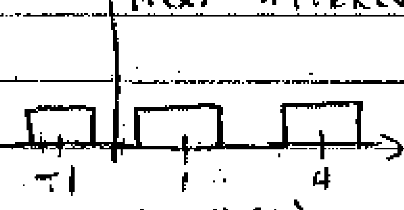
BISHOP'S NOTATION

EXAMPLE $x^1 = 1,\ x^2 = 4,\ x^3 = -1$

$p_{EMP}(x)$

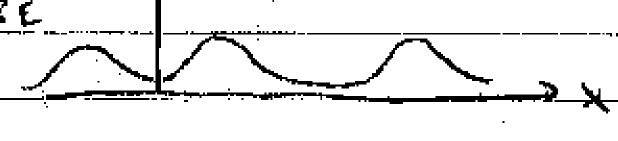



$H(\vec{x})$ = HYPERCUBE, $h = 1$




$h = 3$, $H(\vec{x})$ = CUBE


$H(\vec{x})$ = GAUSSIAN



## BIAS, VARIANCE, & KLD

$$\underline{\text{BIAS}}\left(\tilde{p}(\vec{x})\right) = \left| \mathcal{E}[\tilde{p}(\vec{x})] - p(\vec{x}) \right|$$

EXPECTATION

TRUE VALUE OF THING BEING ESTIMATED

A FUNCTION OF $\vec{x}$ !!!

OVER $p(\underline{X}) = \prod_{n=1}^{N} p(\vec{x}^n)$

$$\underline{\text{VARIANCE}}\left(\tilde{p}(\vec{x})\right) = \mathcal{E}\left[\left(\tilde{p}(x) - \mathcal{E}[\tilde{p}(x)]\right)^2\right]$$

MEAN SQUARED ERROR

$$\text{MSE}\left(\tilde{p}(\vec{x})\right) = \mathcal{E}\left[\left(\tilde{p}(x) - p(\vec{x})\right)^2\right]$$

$$= \mathcal{E}\left[\left(\tilde{p}(x) - \mathcal{E}[\tilde{p}(\vec{x})] + \mathcal{E}[\tilde{p}(\vec{x})] - p(\vec{x})\right)^2\right]$$

$$= \mathcal{E}\left[\left(\tilde{p}(x) - \mathcal{E}[\tilde{p}(x)]\right)^2\right] + \left(\mathcal{E}[\tilde{p}(x)] - p(\vec{x})\right)^2$$

$$+ 2\left(\mathcal{E}[\tilde{p}(x)] - p(\vec{x})\right) \underbrace{\mathcal{E}\left[\left(\tilde{p}(\vec{x}) - \mathcal{E}\tilde{p}(\vec{x})\right)\right]}_{0}$$

$$\boxed{\text{MSE} = \text{VARIANCE} + (\text{BIAS})^2}$$

KARHUNEN-LOEVE DIVERGENCE

$$\text{KLD}\left(\tilde{p}(\vec{x}); p(\vec{x})\right) = \mathcal{E}\left[-\ln \tilde{p}(\vec{x}) + \ln p(\vec{x})\right]$$

$$= \int p(\vec{x}) \ln\left(\frac{p(\vec{x})}{\tilde{p}(\vec{x})}\right) d\vec{x}$$

## BIAS & VARIANCE OF PARZEN WINDOWS

**BIAS**

$$\mathcal{E}\left[\tilde{p}(\vec{x})\right] = \mathcal{E}\left[\frac{1}{N}\sum_{n=1}^{N}\frac{1}{h^D} H\left(\frac{\vec{x}-\vec{x}^n}{h}\right)\right]$$

OVER $p(\mathbf{X}) = \prod_{n=1}^{N} p(\vec{x}^n)$

$$= \frac{1}{N}\sum_{n=1}^{N}\frac{1}{h^D}\,\mathcal{E}\left[H\left(\frac{\vec{x}-\vec{x}^n}{h}\right)\right]$$

$$= \frac{1}{N}\sum_{n=1}^{N}\frac{1}{h^D}\underbrace{\int H\left(\frac{\vec{x}-\vec{x}^n}{h}\right) p(\vec{x}^n)\, d\vec{x}^n}_{\text{EXPECTATION OVER } \vec{x}^{(n)}}$$

$$= \frac{1}{h^D}\int H\left(\frac{\vec{x}-\vec{x}'}{h}\right) p(\vec{x}')\, d\vec{x}' = \frac{1}{h^D}\left(H\left(\frac{\vec{x}}{h}\right) * p(\vec{x})\right)$$

$\longrightarrow p(\vec{x})$ AS $h \to 0$

$\frac{1}{h^D} H\left(\frac{\vec{x}}{h}\right) \to \delta(\vec{x})$

**VARIANCE**

$$\mathcal{E}\left[\left(\tilde{p}(\vec{x}) - \mathcal{E}[\tilde{p}(\vec{x})]\right)^2\right]$$

$$= \left(\frac{1}{h}\right)^{2D}\mathcal{E}\left[\left(\frac{1}{N}\sum_{n=1}^{N} H\left(\frac{\vec{x}-\vec{x}^n}{h}\right) - \mathcal{E}[\tilde{p}(\vec{x})]\right)^2\right]$$

$$= \frac{1}{V^2}\,\mathcal{E}\left[\left(\frac{1}{N}\sum_{n=1}^{N}\left(H\left(\frac{\vec{x}-\vec{x}_n}{h}\right) - \mathcal{E}\left[H\left(\frac{\vec{x}-\vec{x}^n}{h}\right)\right]\right)\right)^2\right]$$

$$= \frac{1}{V^2}\frac{1}{N^2}\sum_{n=1}^{N}\mathcal{E}\left[\left(H\left(\frac{\vec{x}-\vec{x}_n}{h}\right) - \mathcal{E}\left[H\left(\frac{\vec{x}-\vec{x}^n}{h}\right)\right]\right)^2\right]$$

$$= \frac{1}{N^2 V^2}\, N \cdot \text{CONSTANT}$$

$\longrightarrow 0$ IFF $NV^2 \longrightarrow \infty \Leftrightarrow V \gg \frac{1}{\sqrt{N}}$

$V \longrightarrow 0$ AS $N \to \infty$

## METHOD III : K NEAREST NEIGHBORS

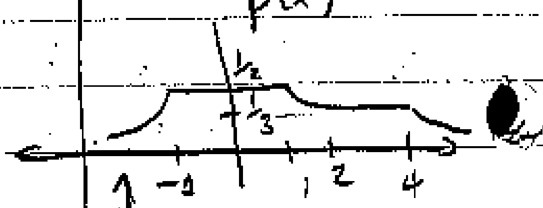
(a) FIX $K$

(b) SET $V(\vec{x})$ = VOLUME (SMALLEST HYPERSPHERE ENCLOSING $K$ DATA POINTS)

(c) $\tilde{p}(\vec{x}) = \dfrac{K}{V(\vec{x})}$

EXAMPLE $\quad p_{emp}(\vec{x}) = \frac{1}{N} \sum \delta(\vec{x} - \vec{x}^n)$


KNN, $K=2$




TAIL DECAYS AS $\frac{1}{|x|} \Rightarrow \int \tilde{p}(x)\,dx = \infty$

MOST PRACTICAL USE

$N_k$ = # OBSERVED FROM CLASS $C_k$

$K_k$ = " " IN $R$

$$\tilde{p}(x) = \frac{K}{N} \qquad \tilde{p}(x|C_k) = \frac{K_k}{N_k}$$

$$\tilde{p}(C_k|x) = \frac{\tilde{p}(x|C_k)\,\tilde{p}(C_k)}{\tilde{p}(x)} = \frac{\left(\frac{K_k}{N_k}\right)\left(\frac{N_k}{N}\right)}{\left(\frac{K}{N}\right)}$$

$$= \frac{K_k}{K}$$