

Geometric Sampling

Probability (and techniques)

Additive
Chernoff Inequality

Double-sampling

VC dimension

Algorithms

ϵ -samples
and ϵ -nets

Concentration phenomena

let $X_1, \dots, X_n \in [0, 1]$ be independent

$$\mu = E[X_1 + \dots + X_n]$$

Q: probability that

$X_1 + \dots + X_n$ is close to μ ?
(multiplicatively)

A: error probability exponentially small in μ

Multiplicative Chernoff bound

let $X_1, \dots, X_n \in [0, 1]$ be independent

$$\mu = E[X_1 + \dots + X_n]$$

$$\epsilon \in (0, 1)$$

- $P[X_1 + \dots + X_n \geq (1 + \epsilon)\mu] \leq e^{-\epsilon^2 \mu / 3}$
- $P[X_1 + \dots + X_n \leq (1 - \epsilon)\mu] \leq e^{-\epsilon^2 \mu / 2}$

Additive Chernoff bound

let $X_1, \dots, X_n \in [0, 1]$ be independent

$$\mu = E[X_1 + \dots + X_n]/n$$

$$\epsilon \in (0, 1)$$

$$\bullet \quad P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right] \leq e^{-2\epsilon^2 n}$$

$$\bullet \quad P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n}$$

(proof at end if time permits)

$X_1, \dots, X_n \in [0,1]$ independent, $\epsilon \in (0,1)$

Multiplicative Chernoff

$$\mu = E[X_1 + \dots + X_n]$$

$$P[X_1 + \dots + X_n \geq (1+\epsilon)\mu] \leq e^{-\epsilon^2 \mu / 3}$$

$$P[X_1 + \dots + X_n \leq (1-\epsilon)\mu] \leq e^{-\epsilon^2 \mu / 2}$$

Additive Chernoff

$$\mu = E[X_1 + \dots + X_n] / n$$

$$P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right] \leq e^{-2\epsilon^2 n}$$

$$P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n}$$

Lecture 12

Probability (and techniques)

✓ Additive
Chernoff Inequality

Double-sampling

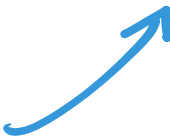
VC dimension

Algorithms

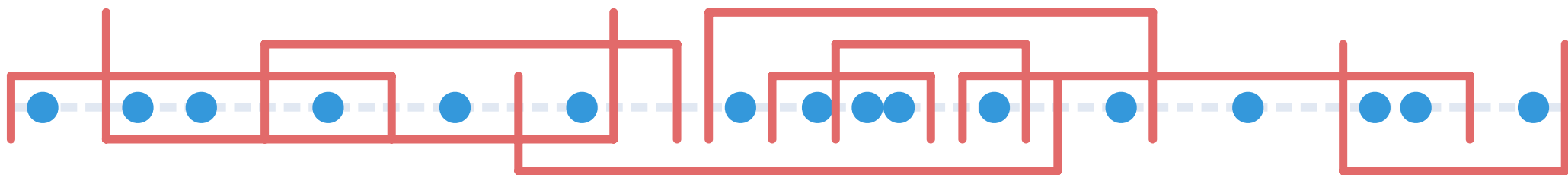
ϵ -samples
and ϵ -nets

Range Space

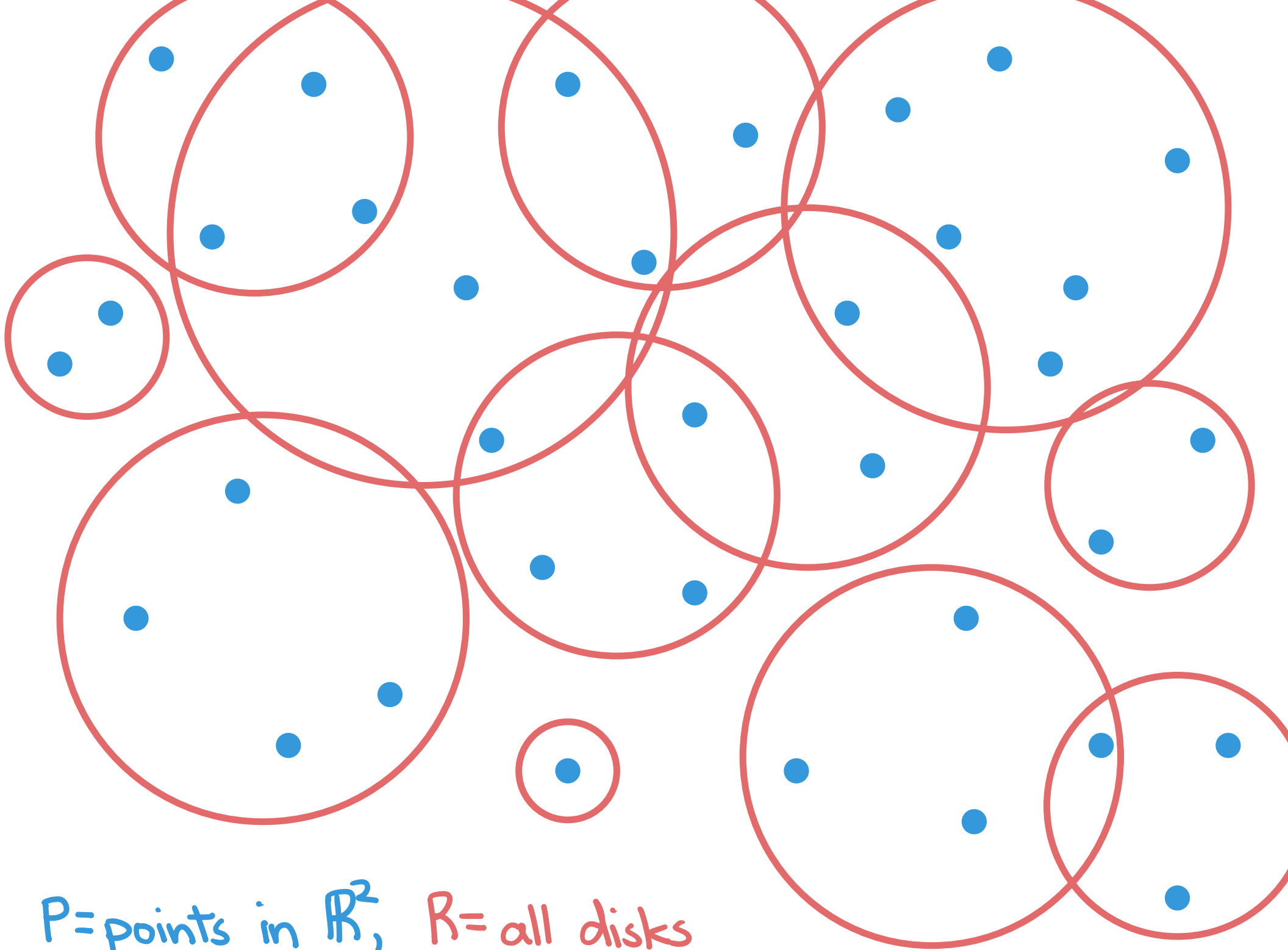
(P, R)

points 

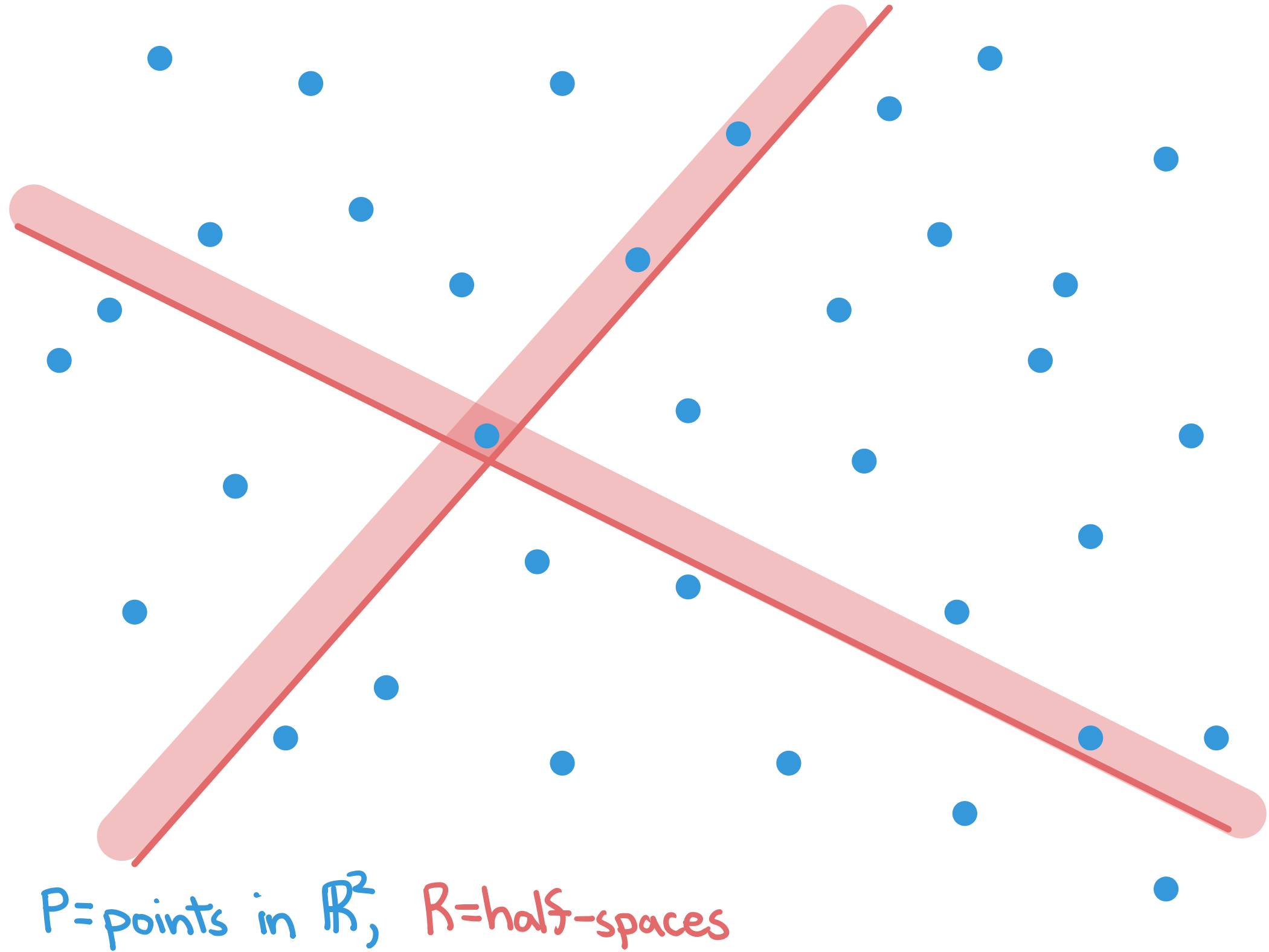
 ranges
(subsets of P)



P = points in $\mathbb{R}$, R = intervals



P =points in $\mathbb{R}^2$, R =all disks

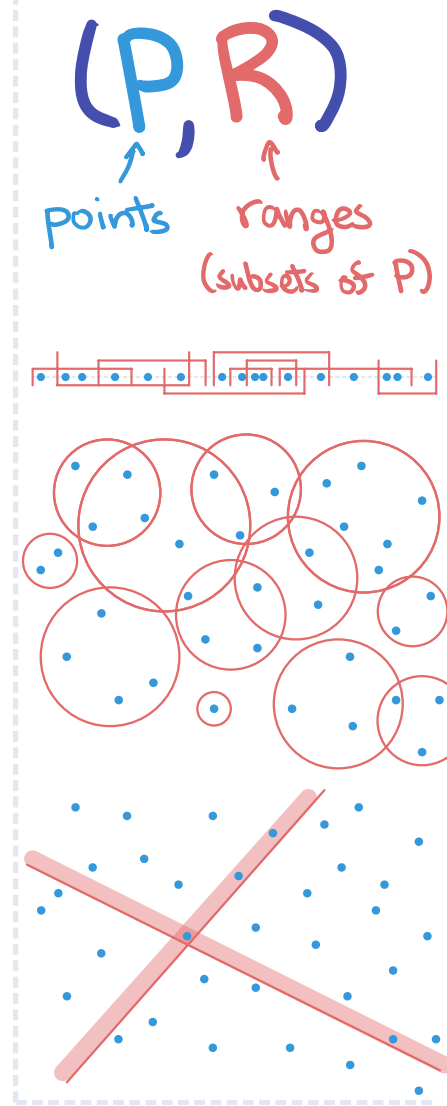


Typical queries: for $r \in R$:

Q: is r empty?

Q: how many points does r contain?

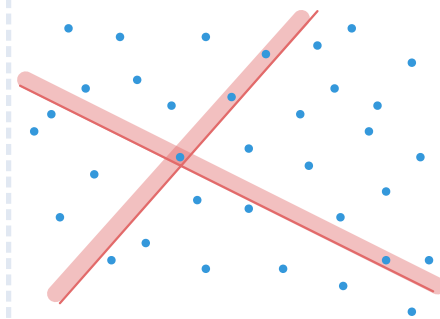
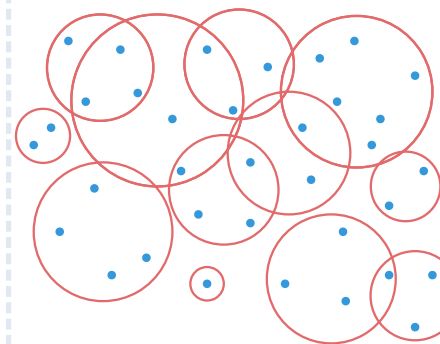
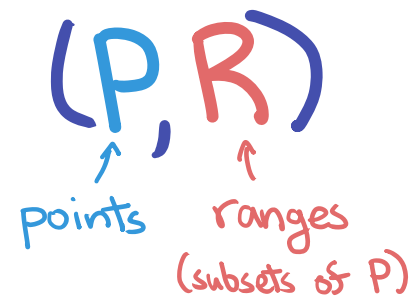
Q: does r contain $\geq \epsilon$ -fraction of P ?
(for $\epsilon \in (0,1)$)



measure $\mu(r) \stackrel{\text{def}}{=} \frac{|P \cap r|}{|P|}$

for $Q \subseteq P$, let $\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$

(high-level)
Question: how big should a random sample $Q \subseteq P$ be so that $\mu_Q \approx \mu$?



Typical queries: for $r \in R$:

Q: is r empty?

Q: how many points does r contain?

Q: does r contain $\geq \epsilon$ -fraction of P ?

(for $\epsilon \in (0,1)$)

ϵ -sample

Q is an ϵ -sample if

$$|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$$

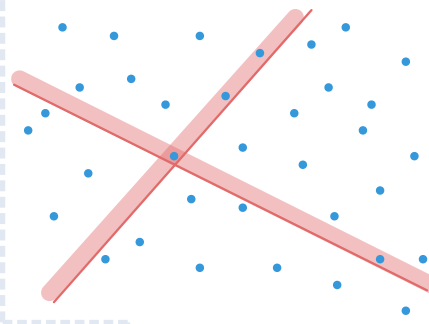
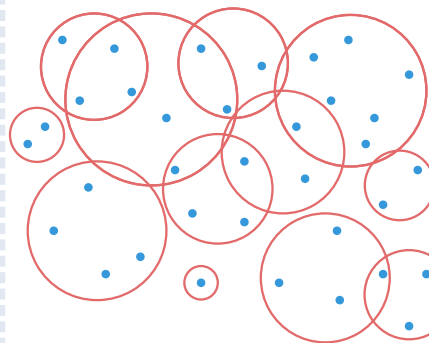
ϵ -net

Q is an ϵ -net if

$$\mu_Q(r) > 0 \text{ whenever } \mu(r) \geq \epsilon$$

(ϵ -sample $\Rightarrow$ ϵ -net)

(P, R)
points ranges
(subsets of P)



$$\mu(r) = \frac{|P \cap r|}{|P|}$$
$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

ϵ -sample

$$|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$$

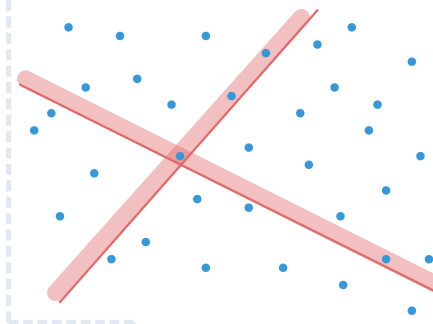
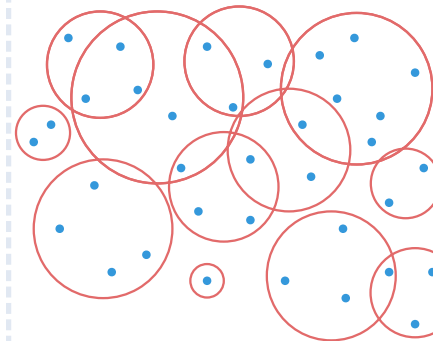
(P, R)
↑ points ↑ ranges
 (subsets of P)

(!) $|Q| < (1-\epsilon)|P|$ is impossible
w/out any assumptions on R

e.g. $R = 2^P$ (=all subsets of P)

if $|Q| < (1-\epsilon)|P|$, then $\bar{Q} = P \setminus Q$ has

$$\mu(\bar{Q}) > \epsilon, \quad \mu_Q(\bar{Q}) = 0$$



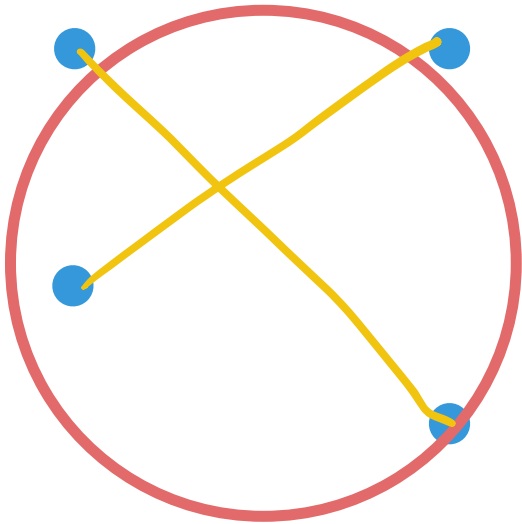
$$\mu(r) = \frac{|P \cap r|}{|P|}$$
$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

(same for ϵ -nets)

(!) $|Q| < (1-\epsilon)|P|$ is impossible w/out any assumptions on R

e.g. $R = 2^P$ (=all subsets of P)

many natural range spaces cannot
induce (nearly) all subsets of P

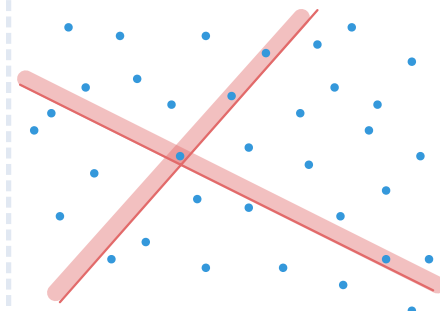
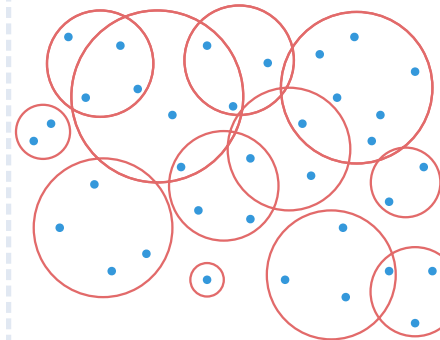


e.g. cannot cover
(only) opposite
corners of a square
w/ a disk

(P, R)

points

ranges
(subsets of P)



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

(!) $|Q| < (1-\epsilon)|P|$ is impossible w/out any assumptions on R

e.g. $R = 2^P$ (=all subsets of P)

many natural range spaces cannot induce (nearly) all subsets of P

let $R \wedge Q = \{r \cap Q : r \in R\}$ for $Q \subseteq P$

growth function $g: \mathbb{N} \rightarrow \mathbb{N}$

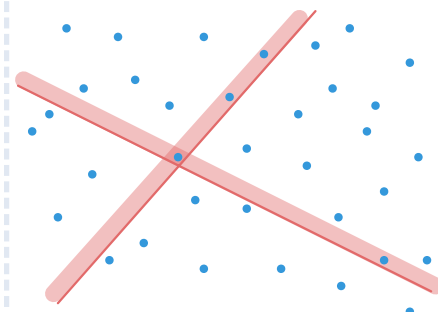
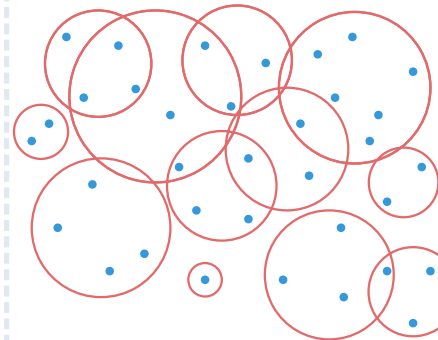
$$g(n) = \max \{ |R \wedge Q| : Q \subseteq P, |Q| \leq n \}$$

e.g. disks have $g(n) \leq O(n^3)$ (which we prove later)

(P, R)

points

ranges
(subsets of P)



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$

let Q sample k points w/ repetition
for range $r \in R$:

$$\mu_Q(r) = \frac{1}{k}(X_1 + \dots + X_k) \text{ w/ } X_i = \begin{cases} 1 & \text{if } i\text{th point in } r \\ 0 & \text{otherwise} \end{cases}$$

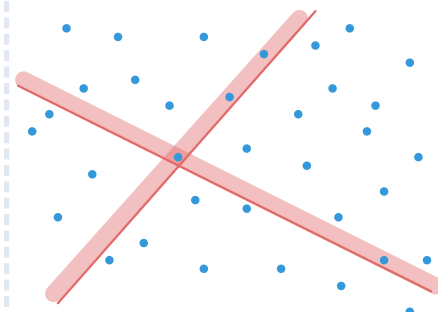
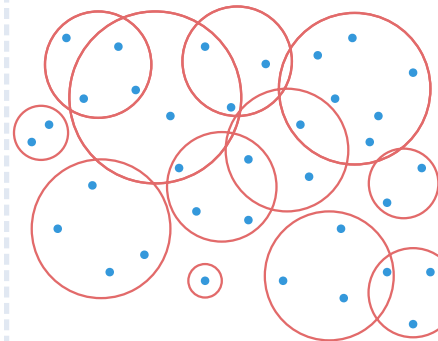
$$P[|\mu_Q(r) - \mu(r)| \geq \epsilon]$$

$$= P\left[\left|\frac{1}{k}(X_1 + \dots + X_k) - \mu(r)\right| \geq \epsilon\right]$$

$$\leq 2e^{-2\epsilon^2 k}$$

(P, R)

points $\uparrow$ ranges
(subsets of P)



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \wedge Q = \{r \cap Q : r \in R\}$$

$$P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right], P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n} \quad g(n) = \max \{|R \wedge Q| : Q \subseteq P, |Q| \leq n\}$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$

let Q sample k points w/ repetition

✓ $P[|\mu_Q(r) - \mu(r)| \geq \epsilon] \leq 2e^{-2\epsilon^2 k}$ for $r \in R$

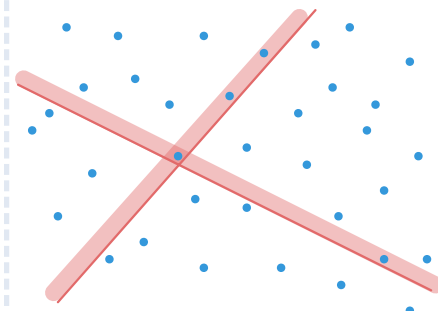
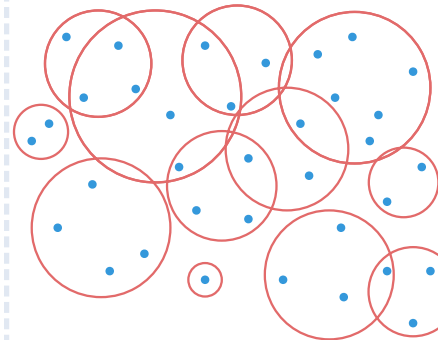
how many r 's?

$$|R \cap P| \leq g(|P|)$$

(P, R)

points

ranges
(subsets of P)



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \cap Q = \{r \cap Q : r \in R\}$$

$$P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right], P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n}$$

$$g(n) = \max \{|R \cap Q| : Q \subseteq P, |Q| \leq n\}$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$

let Q sample k points w/ repetition

✓ $P[|\mu_Q(r) - \mu(r)| \geq \epsilon] \leq 2e^{-2\epsilon^2 k}$ for $r \in R$

✓ $|R \cap P| \leq g(|P|)$

$k = O(\ln(g(|P|))/\epsilon^2) \Rightarrow$

$P_r[Q \text{ not } \epsilon\text{-sample}]$

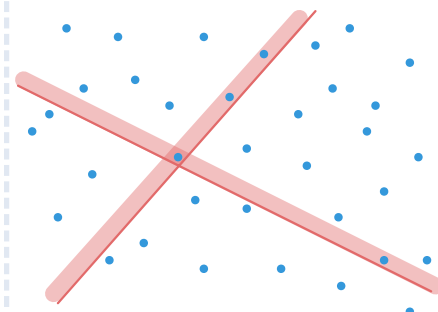
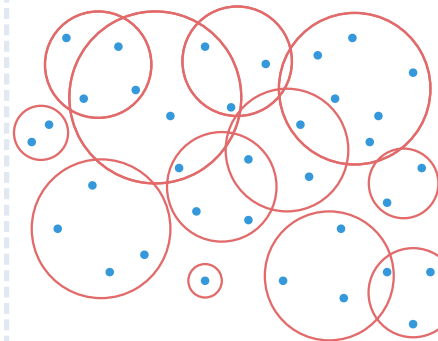
$\leq \sum_{r \in R \cap P} P[|\mu_Q(r) - \mu(r)| \geq \epsilon]$
(union bound)

$\leq g(|P|) e^{-\epsilon^2 k} \leq \frac{1}{2}$

(P, R)

points

ranges
(subsets of P)



$\mu(r) = \frac{|P \cap r|}{|P|}$

$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$

$R \cap Q = \{r \cap Q : r \in R\}$

$P[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon], P[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon] \leq e^{-2\epsilon^2 n}$

$g(n) = \max \{|R \cap Q| : Q \subseteq P, |Q| \leq n\}$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$

let Q sample k points w/ repetition

$$k = O\left(\frac{\ln(g(|P|))}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample}$$

$$(\text{for } g(n) = n^{O(1)})$$

(w/ constant probability)

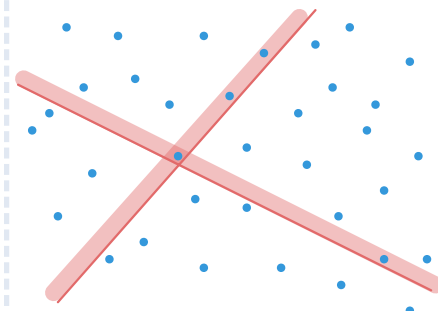
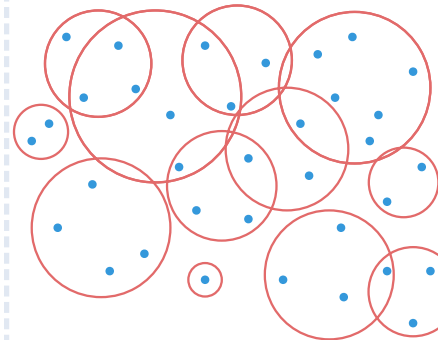
Also (exercise)

$$k = O\left(\frac{\ln(g(|P|))}{\epsilon}\right) \Rightarrow \epsilon\text{-net}$$

(w/ constant probability)

Better?

(P, R)
 $\uparrow$ points $\uparrow$ ranges
 (subsets of P)



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \wedge Q = \{r \cap Q : r \in R\}$$

$$P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right], P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n} \quad g(n) = \max \{|R \wedge Q| : Q \subseteq P, |Q| \leq n\}$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$

let Q sample k points w/ repetition

$$k = O\left(\frac{\ln(g(|P|))}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample} \quad k = O\left(\frac{\ln(g(|P|))}{\epsilon}\right) \Rightarrow \epsilon\text{-net} \quad (\text{w/ constant probability})$$

ϵ -sample theorem

[VC71]

$$k \geq O\left(\frac{\ln(g(2k))}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample} \quad (\text{w/ constant probability})$$

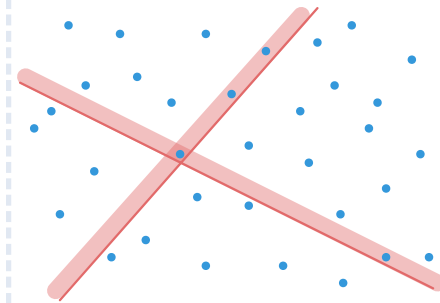
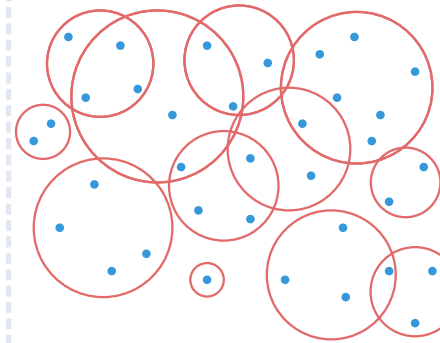
e.g. $g(n) = O(n^d)$

$$k \geq O\left(\frac{d \ln(d/\epsilon)}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample}$$

↑ independent of $|P|$ (!)

$$P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right], P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n}$$

(P, R)
 ↑ points ↑ ranges
 (subsets of P)



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \wedge Q = \{r \cap Q : r \in R\}$$

$$g(n) = \max \{ |R \wedge Q| : Q \subseteq P, |Q| \leq n \}$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$

let Q sample k points w/ repetition

$$k = O\left(\frac{\ln(g(|P|))}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample} \quad k = O\left(\frac{\ln(g(|P|))}{\epsilon}\right) \Rightarrow \epsilon\text{-net} \quad (\text{w/ constant probability})$$

ϵ -net theorem

[HW87]

$$k \geq O\left(\frac{\ln(g(2k))}{\epsilon}\right) \Rightarrow \epsilon\text{-net}$$

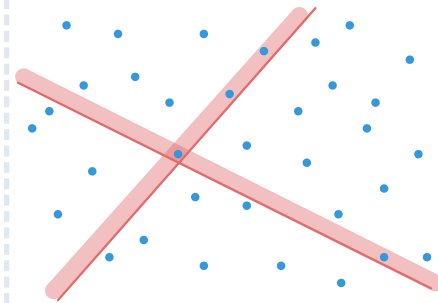
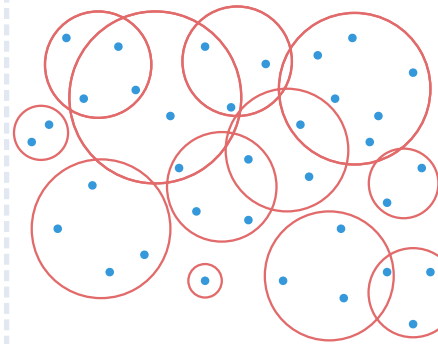
(w/ constant probability)

e.g. $g(n) = O(n^d)$

$$k \geq O\left(\frac{d \ln(1/\epsilon)}{\epsilon}\right) \Rightarrow \epsilon\text{-net}$$

↑ independent of $|P|$ (!)

(P, R)
 ↑ points ↑ ranges
 (subsets of P)



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \wedge Q = \{r \cap Q : r \in R\}$$

$$P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right], P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n} \quad g(n) = \max \{|R \wedge Q| : Q \subseteq P, |Q| \leq n\}$$

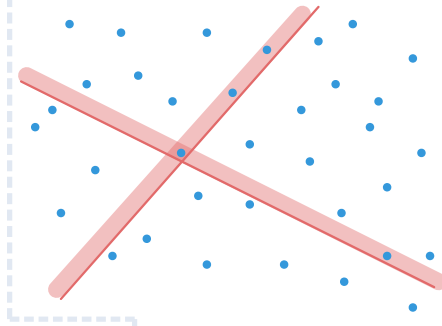
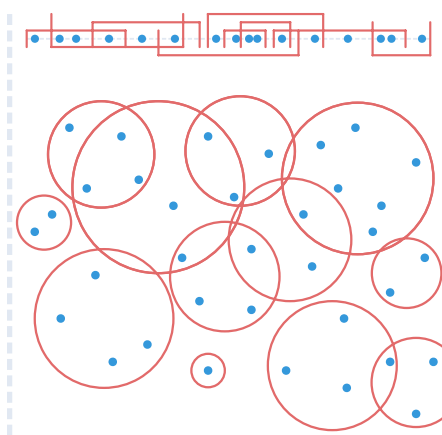
ϵ -sample theorem

[VC71]

$$k \geq O\left(\frac{\ln(g(2k))}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample}$$

(w/ constant probability)

Proof ideas?



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \cap Q = \{r \cap Q : r \in R\}$$

$$g(n) = \max \{|R \cap Q| : |Q| \leq n\}$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r$

ϵ -net $\mu_Q(r) > 0 \quad \forall \mu(r) \geq \epsilon$

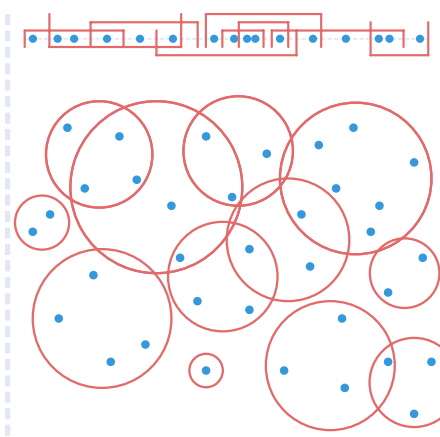
$$\left\{ \begin{array}{l} P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right] \\ P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \end{array} \right\} \leq e^{-2\epsilon^2 n}$$

ϵ -sample theorem

[VC71]

$$k \geq O\left(\frac{\ln(g(2k))}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample}$$

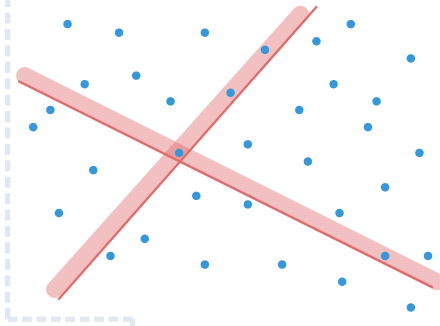
(w/ constant probability)



Proof let Q_1 sample k points (w/ repetition)

define event

$$A = \text{(event that)} |\mu_1(r) - \mu(r)| > \epsilon \text{ for some } r$$



$$\mu(r) = \frac{|P \cap r|}{|P|}$$
$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

goal: prove $P[A] \leq \frac{1}{2}$

$$R \cap Q = \{r \cap Q : r \in R\}$$

$$g(n) = \max \{|R \cap Q| : |Q| \leq n\}$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \forall r$

ϵ -net $\mu_Q(r) > 0 \forall \mu(r) \geq \epsilon$

$$\left\{ \begin{array}{l} P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right] \\ P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \end{array} \right\} \leq e^{-2\epsilon^2 n}$$

let Q_1 sample k points
(w/ repetition)

$$\epsilon\text{-sample theorem}$$

$$k \geq O\left(\frac{\ln(g(2k))}{\epsilon^2}\right)$$

define event

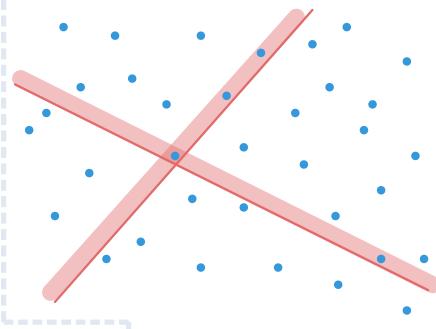
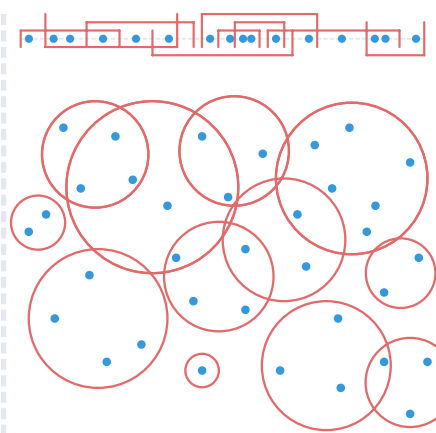
$$A = \text{(event that)} |\mu_1(r) - \mu(r)| > \epsilon \text{ for some } r$$

(!) let Q_2 sample k points

$$B = \text{(event that)} |\mu_2(r) - \mu_1(r)| > \frac{\epsilon}{2} \text{ for some } r$$

$$C = \text{(event that)} |\mu_1(r) - \mu(r)| > \epsilon \text{ and}$$

$$|\mu_2(r) - \mu_1(r)| > \frac{\epsilon}{2} \text{ for some } r$$



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \wedge Q = \{r \cap Q : r \in R\}$$

$$g(n) = \max \{|R \wedge Q| : |Q| \leq n\}$$

$$\epsilon\text{-sample } |\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r$$

$$\epsilon\text{-net } \mu_Q(r) > 0 \quad \forall \mu(r) \geq \epsilon$$

$$\left\{ \begin{array}{l} P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right] \\ P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \end{array} \right\} \leq e^{-2\epsilon^2 n}$$

let Q_1, Q_2 sample k points
(w/ repetition)

$A = (\text{event that}) |\mu_1(r) - \mu(r)| > \varepsilon$ for some r

$B = (\text{event that}) |\mu_2(r) - \mu(r)| > \frac{\varepsilon}{2}$ for some r

$C = (\text{event that}) |\mu_1(r) - \mu(r)| > \varepsilon$ and $|\mu_2(r) - \mu(r)| > \frac{\varepsilon}{2}$ for some r

claim: $P[A] \leq 2P[B]$

$$P[B] \geq P[C] = P[A, C] = P[C|A] P[A]$$

$$P[C|A] \geq 1/2$$

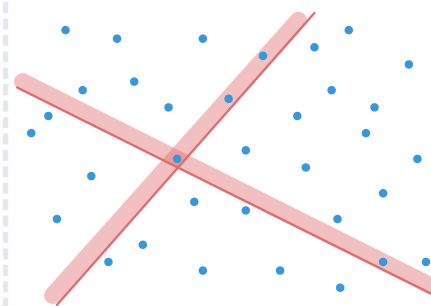
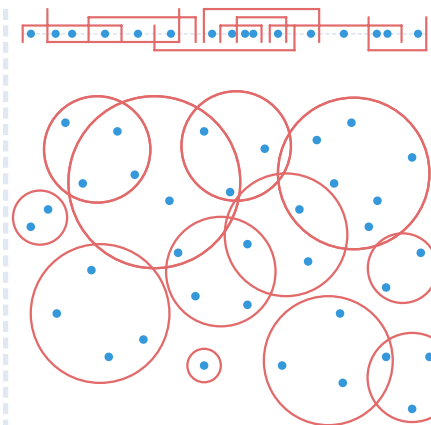
suppose $A \Rightarrow |\mu_1(r) - \mu(r)| > \varepsilon$ for some r

$$|\mu_2(r) - \mu(r)| \leq \frac{\varepsilon}{2} \quad \text{w/ prob} \geq 1/2$$

$$\begin{aligned} \Rightarrow |\mu_1(r) - \mu_2(r)| &\geq |\mu_1(r) - \mu(r)| - |\mu_2(r) - \mu(r)| \\ &\geq \varepsilon - \varepsilon/2 \end{aligned}$$

$$\varepsilon\text{-sample theorem}$$

$$k \geq O\left(\frac{\ln(g(2k))}{\varepsilon^2}\right)$$



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \cap Q = \{r \cap Q : r \in R\}$$

$$g(n) = \max \{|R \cap Q| : |Q| \leq n\}$$

ε -sample $|\mu_Q(r) - \mu(r)| \leq \varepsilon \quad \forall r$

ε -net $\mu_Q(r) > 0 \quad \forall \mu(r) \geq \varepsilon$

$$\left. \begin{aligned} P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \varepsilon\right] \\ P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \varepsilon\right] \end{aligned} \right\} \leq e^{-2\varepsilon^2 n}$$

let Q_1, Q_2 sample k points (w/ repetition)

$A = (\text{event that}) |\mu_1(r) - \mu(r)| > \epsilon$ for some r

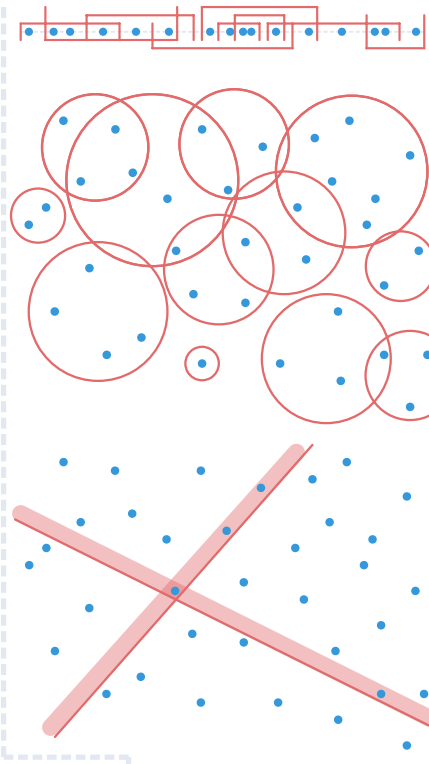
$B = (\text{event that}) |\mu_2(r) - \mu(r)| > \frac{\epsilon}{2}$ for some r

✓ $P[A] \leq 2P[B]$

goal: prove $P[B] \leq \frac{1}{4}$

ϵ -sample theorem

$$k \geq O\left(\frac{\ln(g(2k))}{\epsilon^2}\right)$$



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \cap Q = \{r \cap Q : r \in R\}$$

$$g(n) = \max \{|R \cap Q| : |Q| \leq n\}$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \forall r$

ϵ -net $\mu_Q(r) > 0 \forall \mu(r) \geq \epsilon$

$$\left\{ \begin{array}{l} P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right] \\ P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \end{array} \right\} \leq e^{-2\epsilon^2 n}$$

let Q_1, Q_2 sample k points (w/ repetition)

$A = (\text{event that}) |\mu_1(r) - \mu(r)| > \epsilon$ for some r

$B = (\text{event that}) |\mu_2(r) - \mu(r)| > \frac{\epsilon}{2}$ for some r

$$\checkmark P[A] \leq 2P[B]$$

Suppose we generate Q_1, Q_2 as follows:

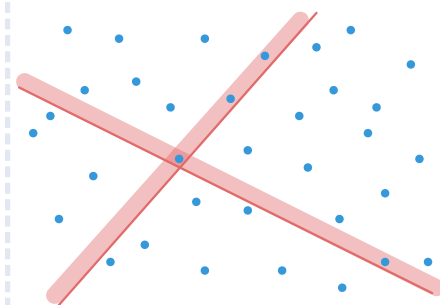
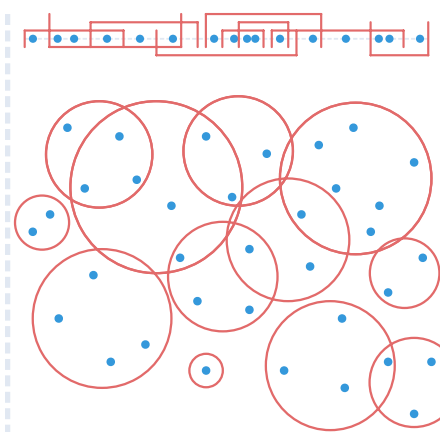
① sample $2k$ points Q_0

② randomly partition Q_0 into halves Q_1, Q_2

$$P[B] = \sum_{Q_0} P[B|Q_0] \cdot P[Q_0] \\ \leq \max_{Q_0} P[B|Q_0]$$

goal: prove $P[B|Q_0] \leq \frac{1}{4} \quad \forall Q_0$

$$\epsilon\text{-sample theorem} \\ k \geq O\left(\frac{\ln(g(2k))}{\epsilon^2}\right)$$



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \wedge Q = \{r \cap Q : r \in R\}$$

$$g(n) = \max \{|R \wedge Q| : |Q| \leq n\}$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r$

ϵ -net $\mu_Q(r) > 0 \quad \forall \mu(r) \geq \epsilon$

$$\left. \begin{aligned} P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right] \\ P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \end{aligned} \right\} \leq e^{-2\epsilon^2 n}$$

$A = (\text{event that}) |\mu_1(r) - \mu(r)| > \varepsilon$ for some r

$B = (\text{event that}) |\mu_2(r) - \mu_1(r)| > \frac{\varepsilon}{2}$ for some r

✓ $P[A] \leq 2P[B]$

① sample $2k$ points Q_0

② randomly partition Q_0 into halves Q_1, Q_2

Fix Q_0 . let $\mu_0 = \text{measure w/r/t } Q_0$

$$|R \cap Q_0| \leq g(2k) \quad (!)$$

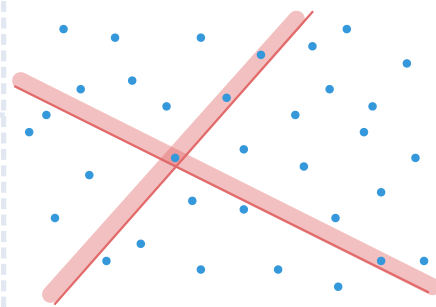
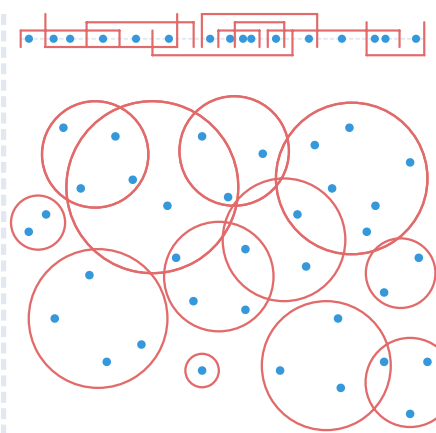
Additive Chernoff:

$$P\left[|\mu_1(r) - \mu_0(r)| \geq \frac{\varepsilon}{4}\right] \leq \frac{1}{8g(2k)}$$

$$P\left[|\mu_2(r) - \mu_0(r)| \geq \frac{\varepsilon}{4}\right] \leq \frac{1}{8g(2k)} \quad \forall r \in R \cap Q_0$$

ε -sample theorem

$$k \geq O\left(\frac{\ln(g(2k))}{\varepsilon^2}\right)$$



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \cap Q = \{r \cap Q : r \in R\}$$

$$g(n) = \max \{|R \cap Q| : |Q| \leq n\}$$

ε -sample $|\mu_Q(r) - \mu(r)| \leq \varepsilon \quad \forall r$

ε -net $\mu_Q(r) > 0 \quad \forall \mu(r) \geq \varepsilon$

$$\left\{ \begin{array}{l} P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \varepsilon\right] \\ P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \varepsilon\right] \end{array} \right\} \leq e^{-2\varepsilon^2 n}$$

Union bound $\Rightarrow$

$$|\mu_1(r) - \mu_0(r)| \leq \varepsilon/4, |\mu_2(r) - \mu_0(r)| \leq \varepsilon/4$$

$$\forall r \in R \cap Q_0 \text{ w/ prob } \geq 3/4$$

$$\forall r \in R \cap Q_0,$$

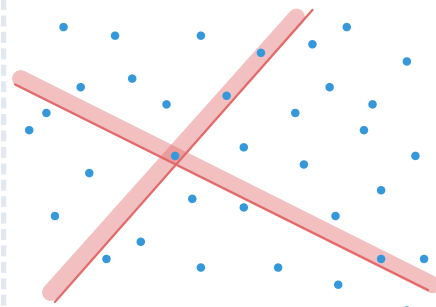
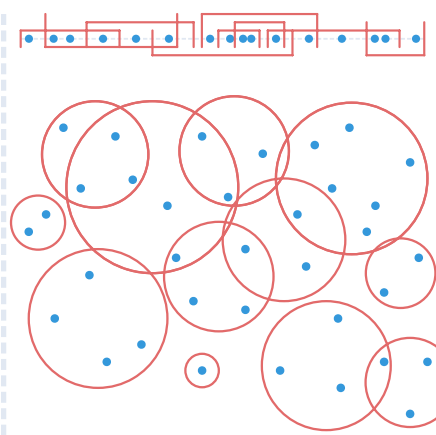
$$|\mu_1(r) - \mu_2(r)| \leq |\mu_1(r) - \mu_0(r)| + |\mu_2(r) - \mu_0(r)|$$

$$\leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}$$

$\Rightarrow$ same $\forall r \in R$

So

$$P[B = \text{(event that)} |\mu_2(r) - \mu_1(r)| > \frac{\varepsilon}{2} \text{ for some } r] \leq \frac{1}{4}$$



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \cap Q = \{r \cap Q : r \in R\}$$

$$g(n) = \max \{|R \cap Q| : |Q| \leq n\}$$

ε -sample $|\mu_Q(r) - \mu(r)| \leq \varepsilon \forall r$

ε -net $\mu_Q(r) > 0 \forall \mu(r) \geq \varepsilon$

$$\left. \begin{aligned} P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \varepsilon\right] \\ P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \varepsilon\right] \end{aligned} \right\} \leq e^{-2\varepsilon^2 n}$$



ϵ -sample theorem

[VC71]

$$k \geq O\left(\frac{\ln(g(2k))}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample}$$

(w/ constant probability)

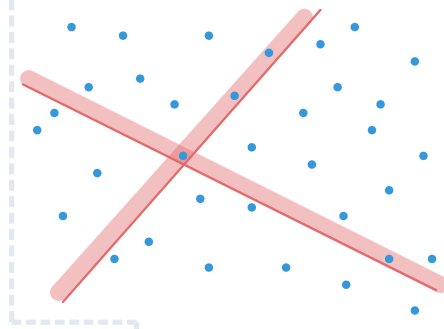
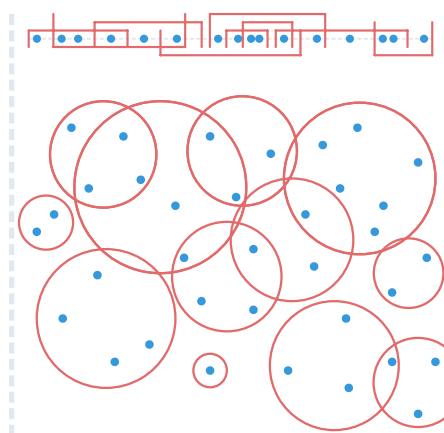
similar,
exercise

ϵ -net theorem

[HW87]

$$k \geq O\left(\frac{\ln(g(2k))}{\epsilon}\right) \Rightarrow \epsilon\text{-net}$$

(w/ constant probability)



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

next question:

how to bound $g(n)$?

$$R \cap Q = \{r \cap Q : r \in R\}$$

$$g(n) = \max \{|R \cap Q| : |Q| \leq n\}$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r$

ϵ -net $\mu_Q(r) > 0 \quad \forall \mu(r) \geq \epsilon$

$$\left\{ \begin{array}{l} P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right] \\ P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \end{array} \right\} \leq e^{-2\epsilon^2 n}$$

Lecture 12

Probability (and techniques)

✓ Additive
Chernoff Inequality

✓ Double-sampling

VC dimension

Algorithms

✓ ϵ -samples
and ϵ -nets

VC dimension

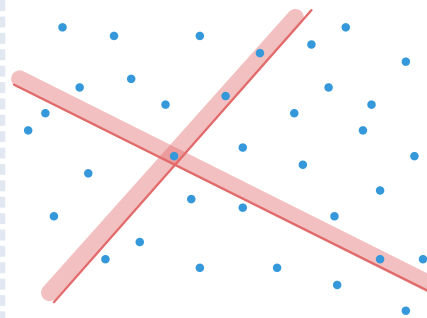
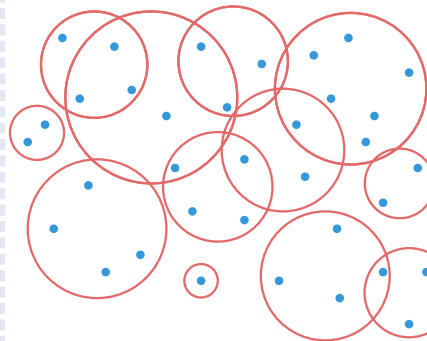
for QCP, R shatters Q if

$R \cap Q = \text{all } 2^{|Q|} \text{ subsets of } Q$

VC dimension of $(P, R) =$

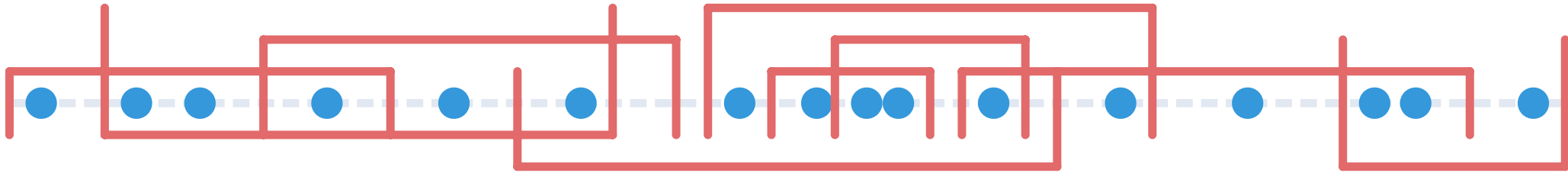
$\max |Q| \text{ s.t. } R \text{ shatters } Q$

(P, R)
points ranges
(subsets of P)



for QCP, R shatters Q if $R \cap Q = \text{all } 2^{|Q|} \text{ subsets of } Q$

VC dimension of $(P, R) = \max |Q| \text{ s.t. } R \text{ shatters } Q$



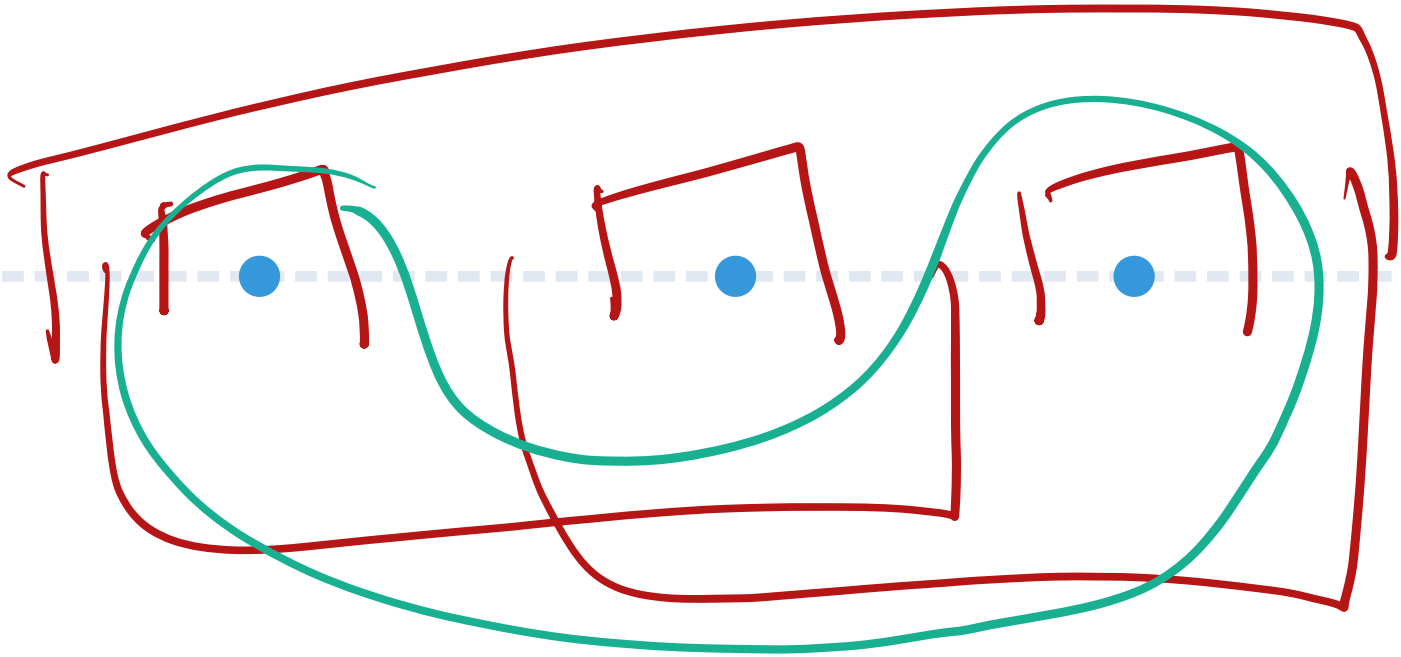
VC-dim =

P = points in $\mathbb{R}$, R = intervals

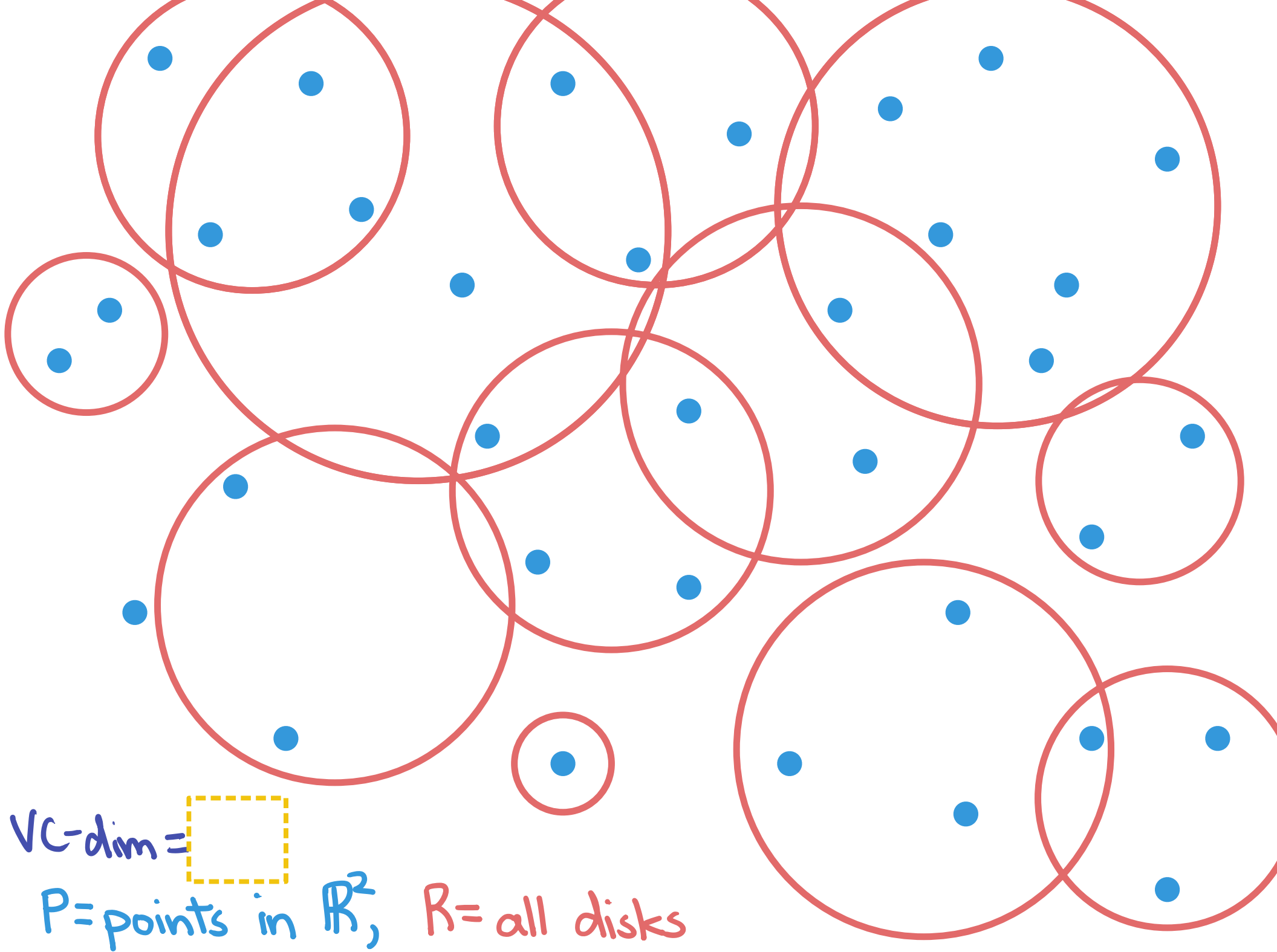
For QCP, R shatters Q if $R \cap Q = \text{all } 2^{|Q|} \text{ subsets of } Q$

VC dimension of $(P, R) = \max |Q| \text{ s.t. } R \text{ shatters } Q$

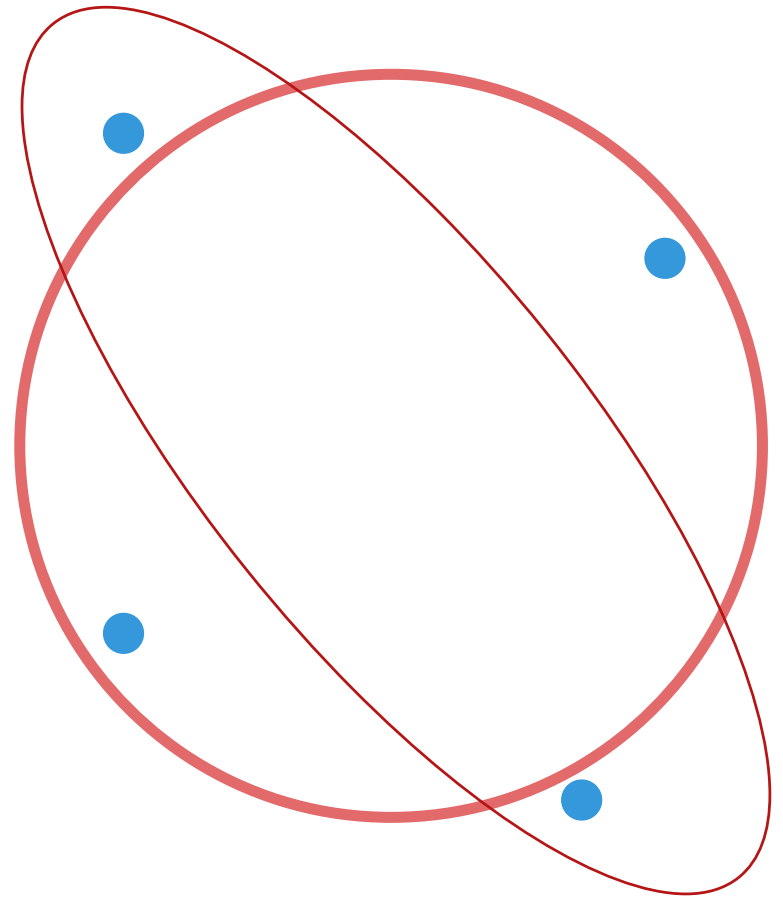
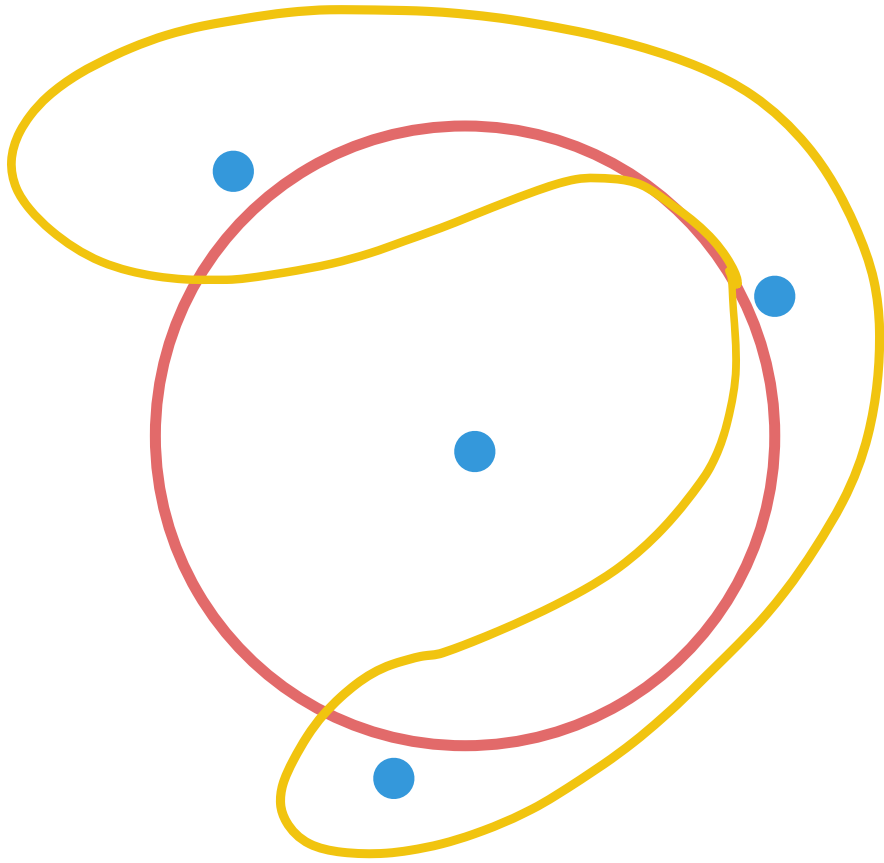
VC-dim = 2



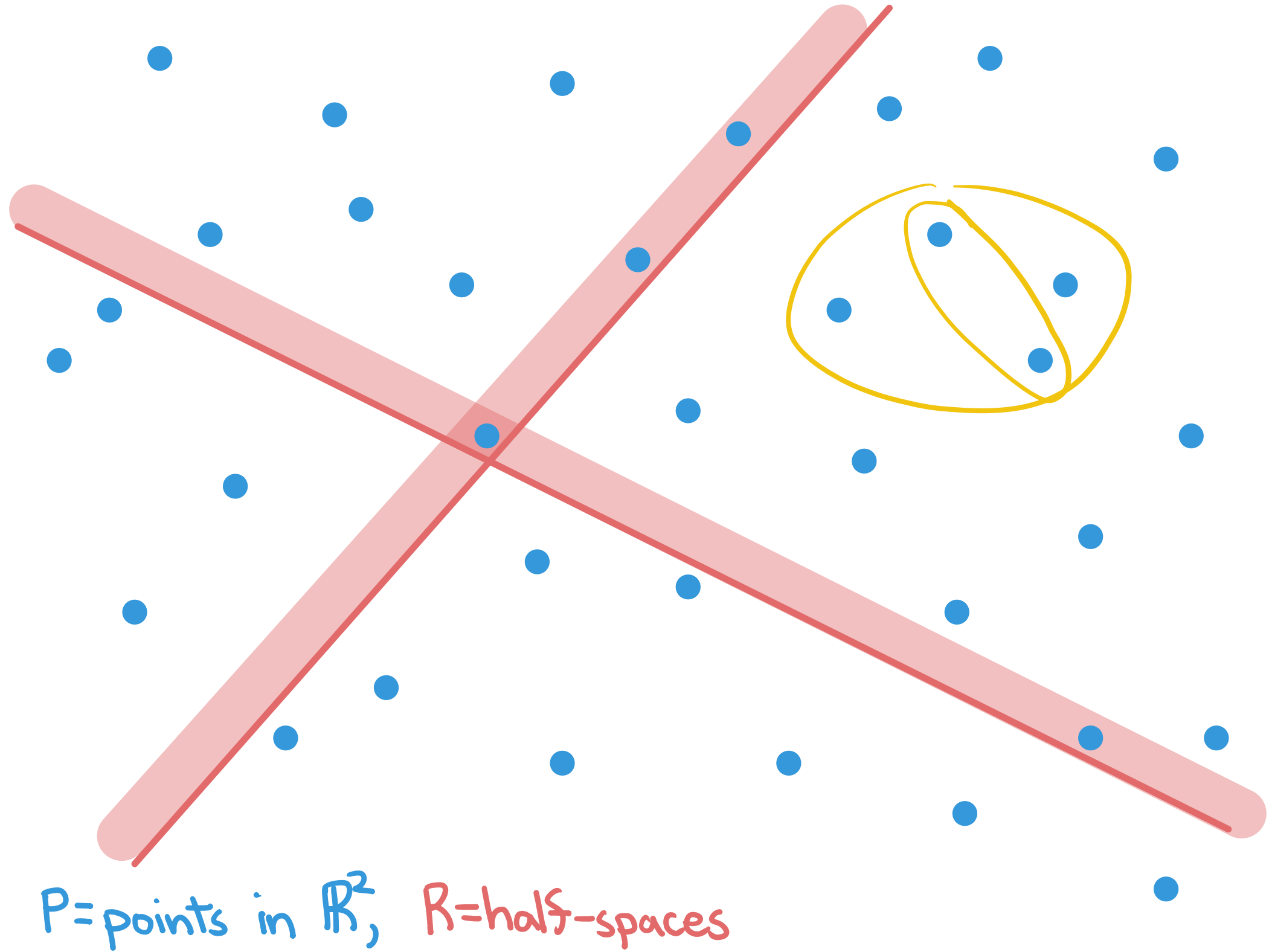
$P = \text{points in } \mathbb{R}$, $R = \text{intervals}$



VC-dim = ~~4~~ 3



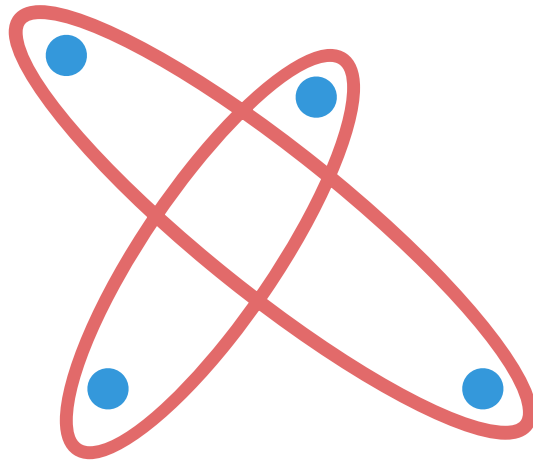
P = points in $\mathbb{R}^2$, R = all disks



Radon's theorem $Q \subset \mathbb{R}^d$, $|Q| = d+2$.

$\exists$ partition $Q = S_1 \cup S_2$ s.t.

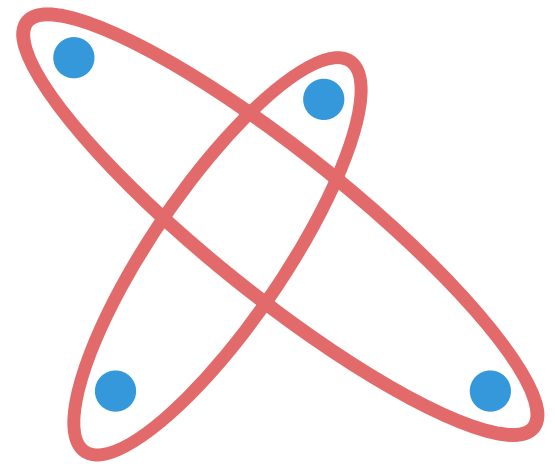
$$\text{convex-hull}(S_1) \cap \text{convex-hull}(S_2) \neq \emptyset$$



Radon's theorem $Q \subset \mathbb{R}^d$, $|Q| = d+2$.

$\exists$ partition $Q = S_1 \cup S_2$ s.t.

$\text{convex-hull}(S_1) \cap \text{convex-hull}(S_2) \neq \emptyset$



$\Rightarrow$ VC-dimension = $d+1$

Sauer's lemma let (P, R) have VC-dim d .
then

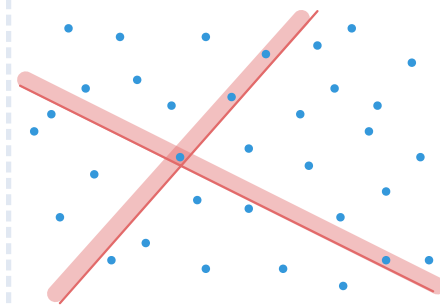
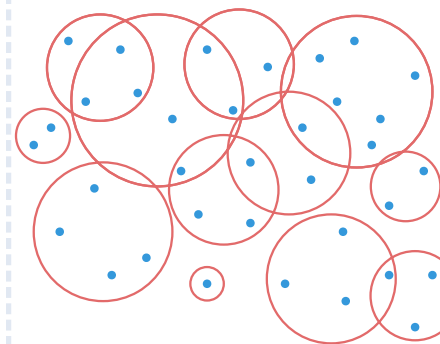
$$g(n) \leq \sum_{i=0}^d \binom{n}{i} \leq n^d$$

e.g. intervals: $g(n) \leq n^2$

disks in $\mathbb{R}^2$: $g(n) \leq n^3$

halfspaces in $\mathbb{R}^d$: $g(n) \leq n^{d+1}$

(P, R)
↑ points ↑ ranges
 (subsets of P)

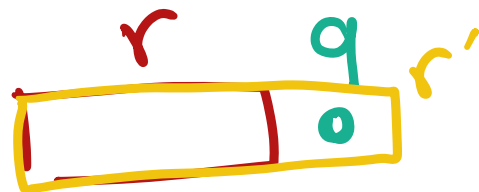


For QCP, R shatters Q if
 $R \cap Q = \text{all subsets of } Q$

VC dimension of $(P, R) =$
 $\max |Q| \text{ s.t. } R \text{ shatters } Q$

Sauer's lemma let (P, R) have VC-dim d .
then

$$g(n) \leq \sum_{i=0}^d \binom{n}{i} \leq n^d$$



Proof let $f(d, n) = \sum_{i=0}^d \binom{n}{i}$.

base cases: $n=0 \Rightarrow g(n)=0 \leq f(d, 0)$

$VC=0 \Rightarrow g(n) \leq 1 \leq f(0, n)$

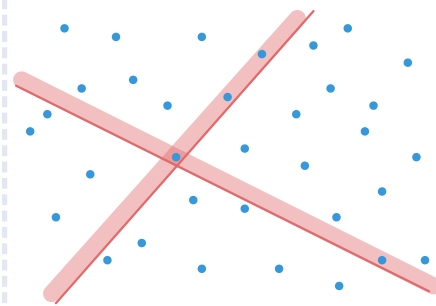
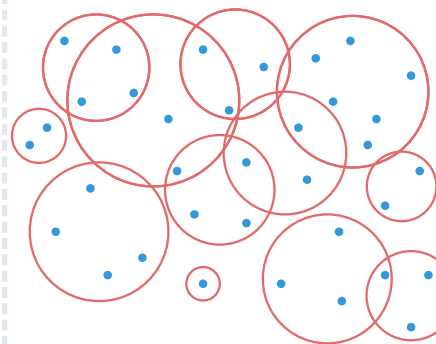
case $n, d > 0$: fix Q w/ $n=|Q|$, $p \in Q$.

let $R_p = \{r-p: r \in R, r \cup \{p\} \in R, r \setminus \{p\} \in R\}$
 $= \{r \in R: p \notin r, r+p \in R\}$

$R^p = \{r-p: r \in R\}$

("r+p" = $r \cup \{p\}$, "r-p" = $r \setminus \{p\}$)

(P, R)
 points $\uparrow$ ranges $\uparrow$
 (subsets of P)



For QCP, R shatters Q if
 $R \upharpoonright Q = \text{all subsets of } Q$

VC dimension of $(P, R) =$
 $\max |Q| \text{ s.t. } R \text{ shatters } Q$

let $f(d,n) = \sum_{i=0}^d \binom{n}{i}$.

case $n,d > 0$: fix Q w/ $n = |Q|$, $p \in Q$. wlog $R = R \cap Q$

$R_p = \{r-p : r \in R, r \cup \{p\} \in R, r \setminus \{p\} \in R\}$ $R^p = \{r-p : r \in R\}$

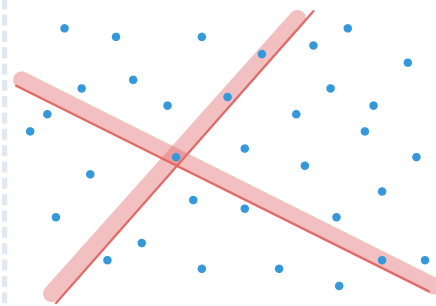
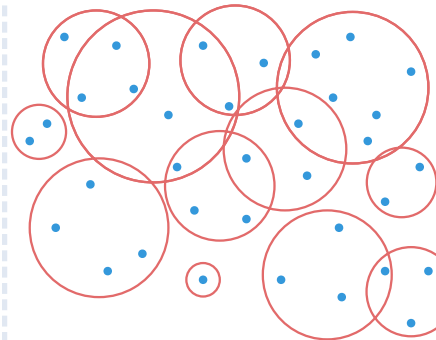
claim: $|R_p| + |R^p| = |R|$

for $r \in R$, 3 cases:

1. $p \notin r$, $r+p \notin R \Rightarrow r \in R^p$
2. $p \notin r$, $r+p \in R \Rightarrow r \in R_p$ and $r \in R^p$
3. $p \in r$, $r-p \in R \Rightarrow$

(P, R)

points ranges
(subsets of P)



for QCP, R shatters Q if
 $R \cap Q = \text{all subsets of } Q$

VC dimension of $(P, R) =$
 $\max |Q| \text{ s.t. } R \text{ shatters } Q$

let $f(d,n) = \sum_{i=0}^d \binom{n}{i}$.

case $n, d > 0$: fix Q w/ $n = |Q|$, $p \in Q$. wlog $R = R \cap Q$

$$R_p = \{r - p : r \in R, r \cup \{p\} \in R, r \setminus \{p\} \in R\} \quad R^P = \{r - p : r \in R\}$$

✓ $|R_p| + |R^P| = |R|$

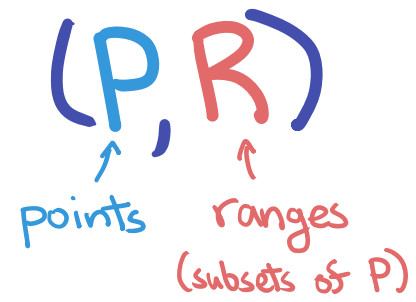
$(Q-p, R_p)$ has VC-dim $d-1$

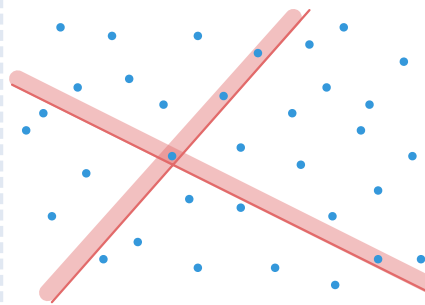
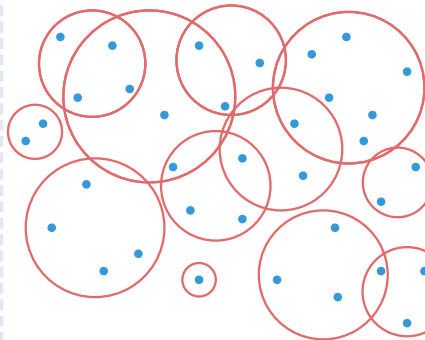
$(Q-p, R^P)$ has VC-dim d

$$|R| = |R_p| + |R^P| \leq f(d-1, n-1) + f(d, n-1)$$

$$= \sum_{i=0}^{d-1} \binom{n-1}{i} + \sum_{i=0}^d \binom{n-1}{i}$$

$$= \sum_{i=1}^d \binom{n-1}{i-1} + \sum_{i=0}^d \binom{n-1}{i}$$

(P, R)

 points ranges
 (subsets of P)



For QCP, R shatters Q if $R \cap Q = \text{all subsets of } Q$

VC dimension of $(P, R) = \max |Q| \text{ s.t. } R \text{ shatters } Q$

$$\leq \sum_{i=0}^d \binom{n}{i} = \mathfrak{F}(d, n)$$

ϵ -sample $|\mu_Q(r) - \mu(r)| \leq \epsilon \quad \forall r \in R$

let Q sample k points w/ repetition

$$k = O\left(\frac{\ln(g(|P|))}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample} \quad k = O\left(\frac{\ln(g(|P|))}{\epsilon}\right) \Rightarrow \epsilon\text{-net} \quad (\text{w/ constant probability})$$

(P, R)

points $\uparrow$ ranges
(subsets of P) $\uparrow$



ϵ -sample theorem

[VC71]

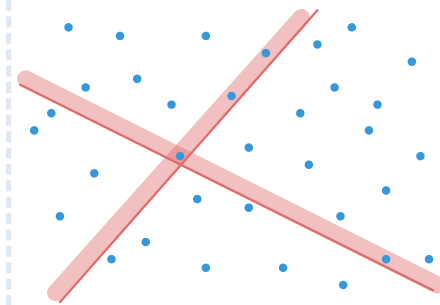
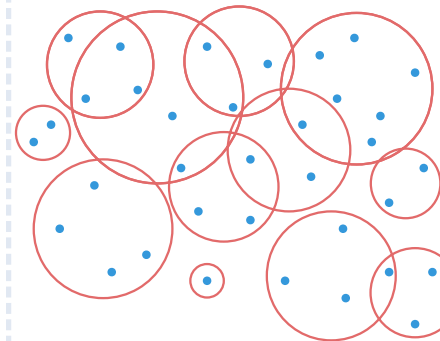
$$k \geq O\left(\frac{\ln(g(2k))}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample} \quad (\text{w/ constant probability})$$

e.g. $g(n) = O(n^d)$

for VC-dimension d

$$k \geq O\left(\frac{d \ln(d/\epsilon)}{\epsilon^2}\right) \Rightarrow \epsilon\text{-sample}$$

$\uparrow$ independent of $|P|$ (!)



$$\mu(r) = \frac{|P \cap r|}{|P|}$$

$$\mu_Q(r) = \frac{|Q \cap r|}{|Q|}$$

$$R \wedge Q = \{r \cap Q : r \in R\}$$

$$P\left[\frac{X_1 + \dots + X_n}{n} \geq \mu + \epsilon\right], P\left[\frac{X_1 + \dots + X_n}{n} \leq \mu - \epsilon\right] \leq e^{-2\epsilon^2 n} \quad g(n) = \max \{|R \wedge Q| : Q \subseteq P, |Q| \leq n\}$$