CS 6100

Lecture - 6 + 7 + 8

# Basics of probability Theory

Ref :- Heiko Röglin's notes on Prob. anal. of algorithms. (available in the google shared folder).

Recall

Analysis of 2-opt with manhatten-metric (instead of Euclidean)

→ random points distributed on $[0,1]^2$, $[0,1]^d$ $d \geq 2$

$$\Delta(S) = |x_1 - x_2| + \ldots$$

$$\Pr[\Delta(S) \leq \Delta_{min}]$$

## Probability Spaces

### Discrete Probability Space $(\Omega, p)$.

- $\Omega$ : samplespace. finite or countable set.
  - set of all possible outcomes.

- $p : \Omega \to [0, 1]$.

* $\sum_{x \in \Omega} p(x) = 1$

An event $S \subseteq \Omega$.

$p(S) = \sum_{x \in S} p(x)$.

die with six faces.
$\{1, \dots 6\}$

$\Omega = \{1, \dots 6\}$.

$p(1) = p(2) = \dots p(6) = \frac{1}{6}$

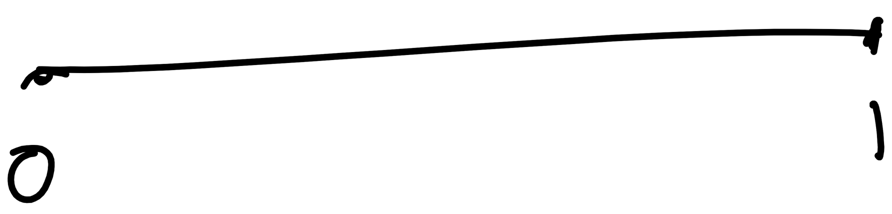
event - outcome is even no

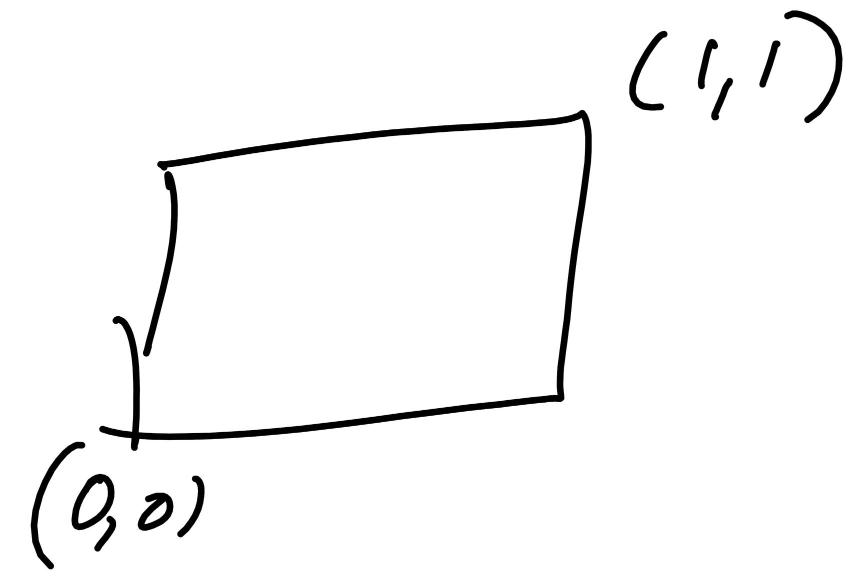
$S = \{2, 4, 6\}$.

$p(S) = \frac{1}{2}$

---

$[0, 1]$




$[-\epsilon, \epsilon]$.

---

## Continuous case

A probability space is $(\Omega, \underline{F}, P)$ s.t

1 $\Omega$ : sample space

2 $F \subseteq 2^{\Omega}$ with the following properties.

(i) $\Omega \in F$

(ii) $F$ should be closed under complementation

ie $X \in F \Rightarrow \Omega \setminus X \in F$.

(iii) $\mathcal{F}$ is closed under countable unions

ie if $X_1, X_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_i X_i \in \mathcal{F}$.

(ie $\mathcal{F}$ should be a $\sigma$ - algebra).

3. $P$ is a probability measure

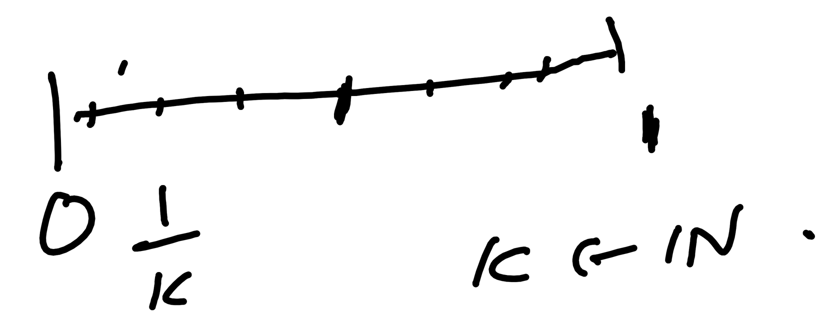
ie $P : \mathcal{F} \to [0,1]$

(i) $P(\Omega) = 1$

(ii) if $X_1, X_2, \ldots \in \mathcal{F}$ s.t $X_i \cap X_j = \phi$ (disjoint)

then $P(\bigcup_i X_i) = \sum_i P(X_i)$

$[0,1]$




$\frac{1}{1}, \frac{1}{2}, \ldots \ldots,$

example of $\mathcal{F}$ : Borel $\sigma$-algebra.

set of all intervals in $\mathbb{R}$.

& their unions.

---

Lemma : $X_1, X_2 \ldots X_n$ are events in $(\Omega, \mathcal{F}, P)$.

then $P[X_1 \cup X_2 \cup \ldots \cup X_n] \leq \sum_{i=1}^{n} P(X_i)$

$\to$ Union bound

# Random variables

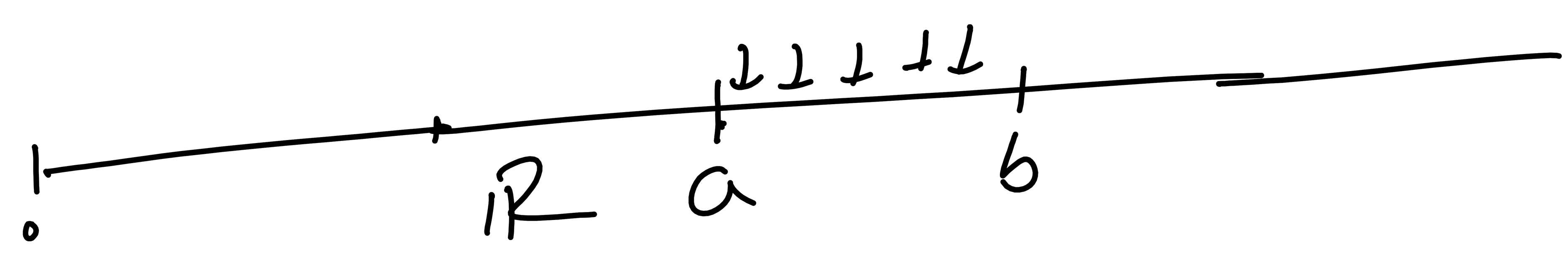
Let $(\Omega, \mathcal{F}, P)$ - prob. space.

A random variable is a fn $X : \Omega \to \mathbb{R}$.
(real valued). $X$: continuous fn

s.t. for an $a \le b$.

$$X^{-1}([a,b]) = \{z \in \Omega \mid X(z) \in [a,b]\} \in \mathcal{F}.$$

$X$




$\Omega = [0,1]$

$X : \Omega \to \mathbb{R}$.
$\alpha \mapsto \alpha^2$

$P_1 \cdots P_n$

eg:- cost of a TSP tour
value of $\Delta(S)$

$\Delta(S) \geq \cdots$
$\Delta(S) = -$

$X$

$Pr[X = x]$: probability that r.v $X$ takes the value $x$.

$$Pr[X = x] \triangleq P[X^{-1}(\{x\})] \qquad Pr[X \le t] \triangleq P[X^{-1}([-\infty, t])]$$

Cumulative Distribution Function (CDF) & Probability density function (PDF).

Let $X$ be a continuous random variable. The CDF of $X$ is the fn $F_X : \mathbb{R} \to \mathbb{R}_{\ge 0}$ given by

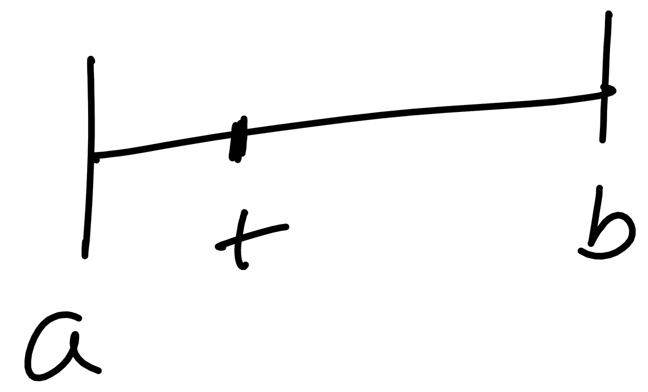
$$\forall t \quad F_X(t) = Pr[X \le t]$$

A function $f_X : \mathbb{R} \to \mathbb{R}_{\ge 0}$ is called PDF of $X$ if.

$$\boxed{\Pr[X \in [a,b]] = \int_a^b f_X(t)\,dt.}$$

## Examples :

① $X$ be a random variable that is uniformly distributed in the interval $[a,b]$




$$\Pr[X \le t] = \frac{t-a}{b-a}.$$

$$F_X(t) = \Pr[X \le t] = \begin{cases} 0 & t < a \\ \frac{t-a}{b-a}. & \\ 0 & t > b \end{cases}$$

$$f_X(t) = \begin{cases} 0 & t > b \text{ or } t < a \\ \frac{1}{b-a}. & \end{cases}$$

---

## Gaussian / Normal distribution

Parameter: $\mu$ : mean.

$\sigma$ : standard deviation.

The PDF of a normal distribution with mean $\mu$ & std. dev $\sigma$ is given by

$$f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{(t-\mu)^2}{2\sigma^2}\right)} \longrightarrow (*)$$

# Lecture 7

Exercise.

→ More examples of random variables

→ Expectation and variance.

→ Conditional probability.

→ Interval lemmas

Obs:- $f_X(t) = \frac{d(F_X(t))}{dt}$

## Expected value (mean)

Suppose $X$ is a discrete random variable taking values from $S$.

Then the expectation of $X$

$$E[X] \triangleq \sum_{x \in S} x \cdot \Pr[X = x]$$

For $X$ continuous

$$\boxed{E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx.}$$

Example:- $X$ be uniformly distributed in $[0,1]$

$f_X(t) = 1$

$$E[X] = \int_0^1 x f(x) dx = \int_0^1 x dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2} - 0 = \frac{1}{2}.$$

$X$ uniform in $[a,b]$

$$E[X] = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2}$$

Ref: Sheldon M Ross.

② $X$ be s.t. the PDF is given by

$$f_X(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1 \\ 0 & \text{o.w.} \end{cases}$$

$$E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx = \int_0^1 x \cdot 2x\,dx = \left[\frac{2x^3}{3}\right]_0^1$$

$$= 2/3$$

③ Let $X$ be given by

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{o.w.} \end{cases}$$

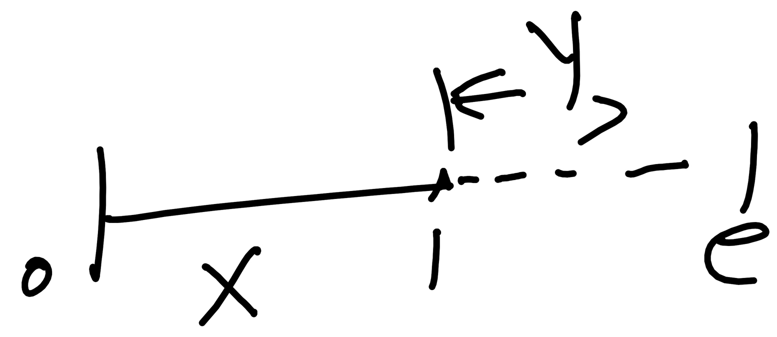
$X$

$e^X, x^2,$

$x^3 - x^2 + X$

To compute $E[e^X]$

$Y = e^X$. $Y$ is a random variable.




$$F_Y(x) = Pr[Y \le x]$$
$$= Pr[e^X \le x] = Pr[X \le \log x]$$
$$= \int_0^{\log x} f_X(t)\,dt = \int_0^{\log x} 1\,dt$$
$$= \log x$$

~~Now~~ ie $F_Y(x) = \log x$

$$f_Y(x) = \frac{d F_Y(x)}{dx} = \frac{d \log x}{dx} = \frac{1}{x}.$$

for $\phi \le x \le e$

Then $E[Y] = \int_1^e x \cdot f_Y(x)\,dx = \int_1^e x \cdot \frac{1}{x}\,dx$

$= \underline{\underline{e-1}}$

<u>Exercies</u> : Compute $E[X^3 - X^2]$

$E[(X+1)^2]$

---

<u>Lemma</u>:- Let $X$ be a non-negative random variable.

Then $E[X] = \int_0^\infty \Pr[X \geq x]\,dx.$

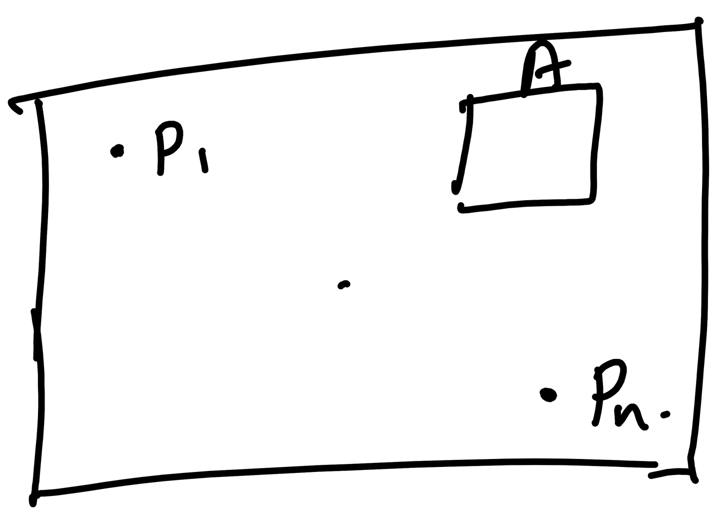
<u>Ex</u>:- Prove this

<u>Lemma</u>: Let $g$ be any function. Then

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\,dx.$$

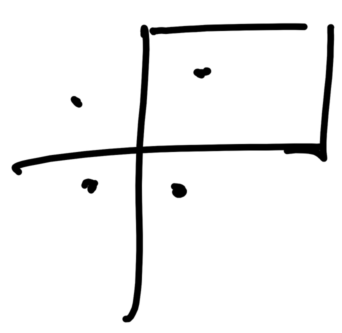
---

<u>Ex</u>:- ① $A$ be a square of area $a$ in $[0,1]^2$




Suppose $n$ points are sampled uniformly and independently at random from $[0,1]^2$.

What is the expected number of points inside $A$?



② Suppose $P_1, \dots, P_n$ are $n$ points in $[0,1]^2$. $\xi_1, \dots, \xi_n$ be $n$- points distributed independ. according to Gaussian with $\mu = 0$ & s.dev $\sigma$. Consider $P_1+\xi_1, P_2+\xi_2, \dots, P_n+\xi_n$.

Compute: the expected # of points that lie outside $[0,1]^2$.

---

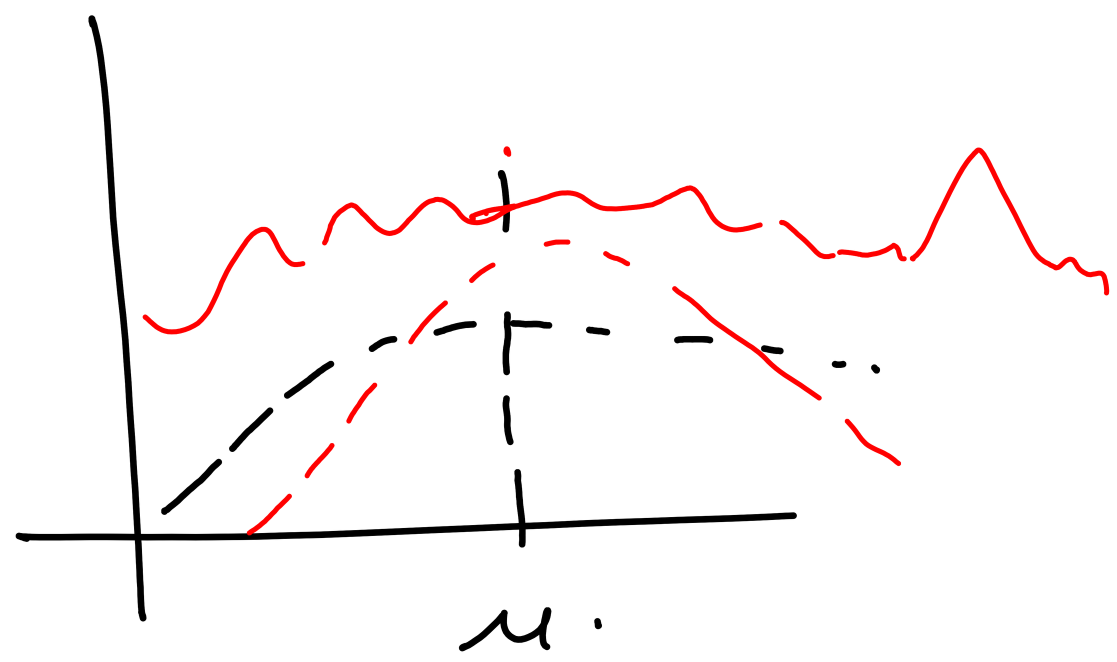
<u>Lemma</u> [Linearity of Expectation] Let $X$ & $Y$ be two random variables

then $E[X+Y] = E[X] + E[Y]$

---

<u>Defn</u>:- [variance].

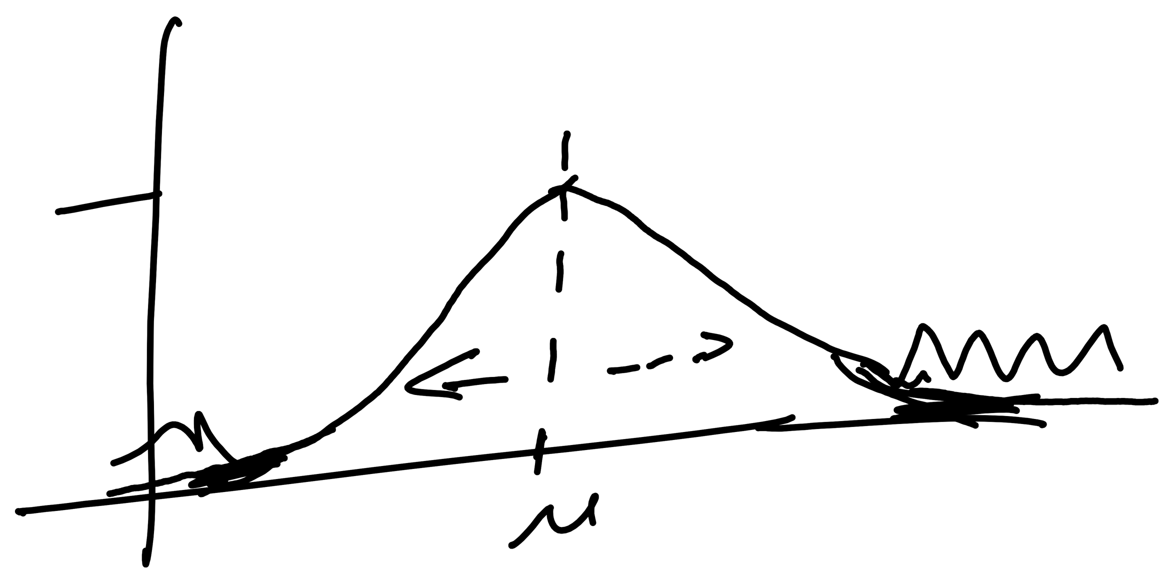
$$\mathrm{var}(X) = E\left[(X - E[X])^2\right]$$




$$Y = X - E[X]$$

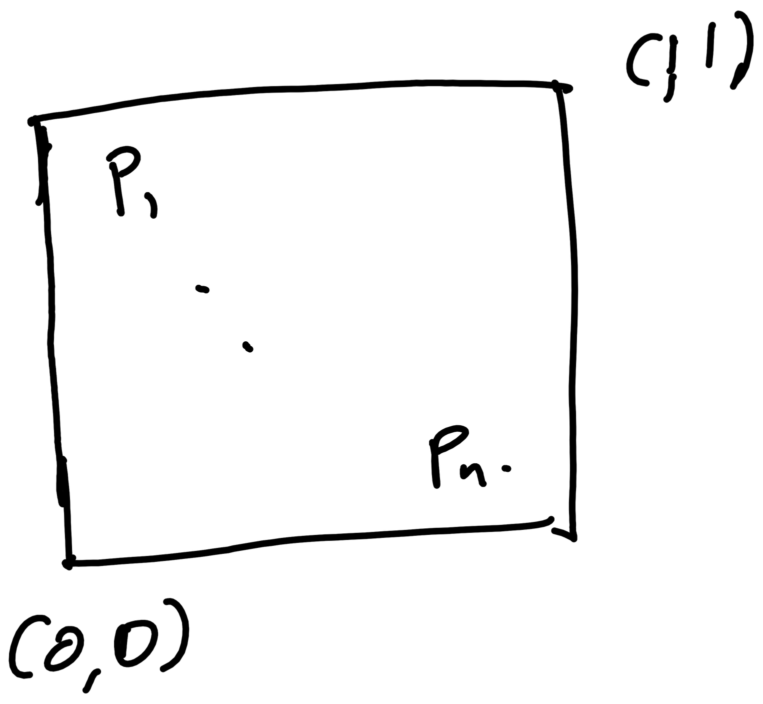
<u>Obs</u>:- $\underline{\mathrm{var}(X) = E[X^2] - (E[X])^2}$

---



Ex:- ① Let $p_1, \dots p_n \in [0,1]^2$.

$T = \{p_1, \dots p_n\}$.




Show that
$$TSP(T) = O(\sqrt{n}).$$

② Show that the above bound is tight.

# Lecture-8

# Smoothed Analysis of 2-OPT

ref: Notes by Manthey.

## Interval Lemmata

**Lemma 1:** Let $\phi > 0$ & $X$ be a r.v. with PDF $f: \mathbb{R} \to [0, \phi]$ i.e. $\forall x\ f(x) \le \phi$

$\forall \epsilon > 0$ & $t$

$$\Pr[X \in (t, t+\epsilon)] \le \epsilon\phi.$$

$$\hookrightarrow \int_t^{t+\epsilon} f(x)\,dx \le \int \phi\, dx = \phi\epsilon.$$

**Lemma 2:-** $X$ - Gaussian with mean $\mu$ & std. dev $\sigma$

$\forall \epsilon > 0, t$

$$\Pr[X \in (t, t+\epsilon)] \le \frac{\epsilon}{2\sigma}$$

---

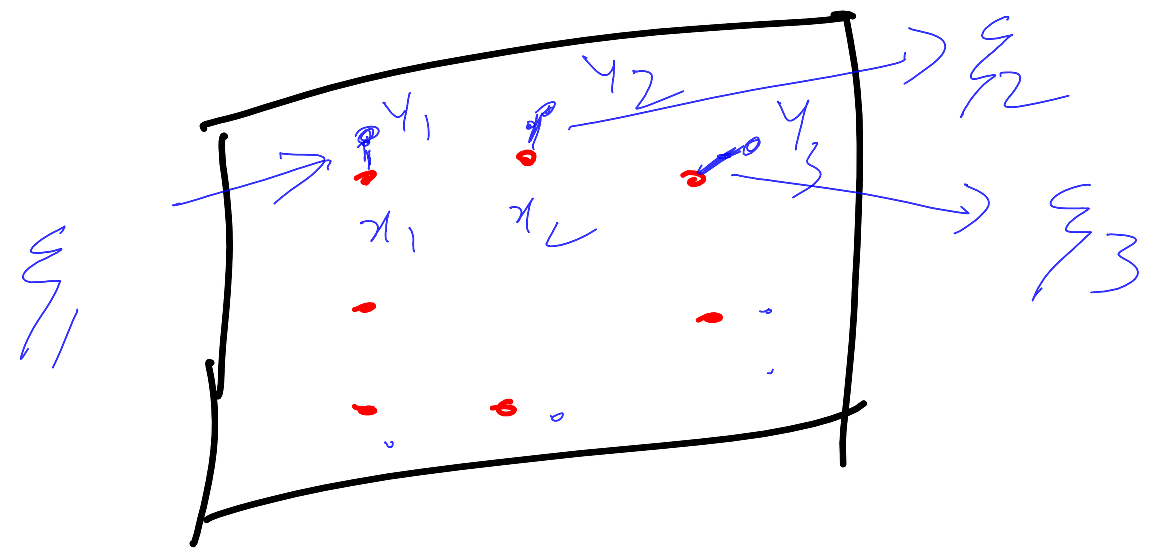
## 2-OPT

initial tour $T$.

→ sequence of 2-opt exchanges.

Let $x_1, \ldots x_n$ be $n$-points in $[0,1]^2$ - fixed by adversary.

$\xi_1, \ldots \xi_n$ - be independent Gaussian r.v. with $\mu = 0$ & std. dev $\sigma$.




$$y_i = x_i + \xi_i \qquad Y = \{y_1, \ldots y_n\}.$$

## Assumptions

→ We consider Squared Euclidean distances

i.e. $d((a_1, b_1)(a_2, b_2))$
$= (a_1 - a_2)^2 + (b_1 - b_2)^2.$

→ $\sigma \le \frac{1}{2\sqrt{n\log n}}$ — in fact the arg works for general $\sigma$, but a bit more involved.

Main idea

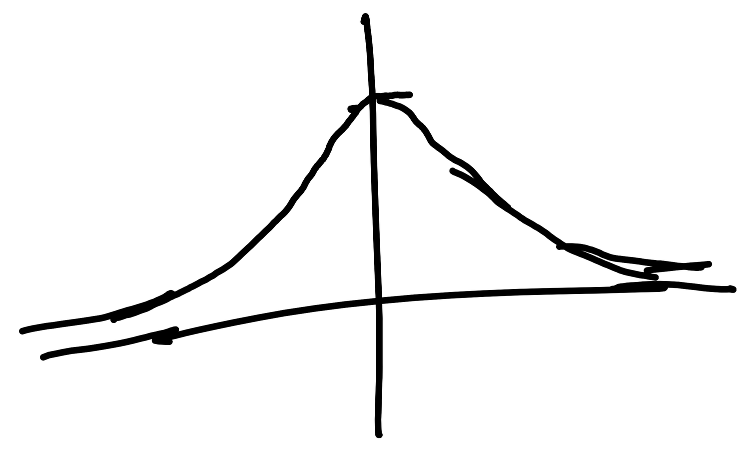
Let $L_{init}$ be the cost of initial tour $T$.

$\Delta_{min}$ be the minimum gain among all 2-opt exchanges. $(n^4)$

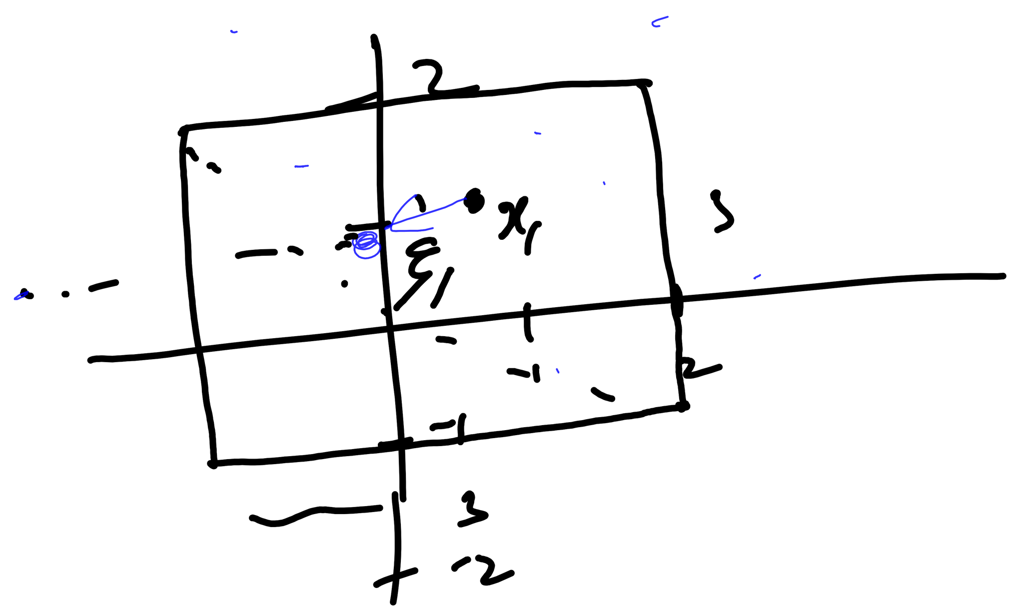
Then bound on the running time $\le \dfrac{L_{init}}{\Delta_{min}}$.

Goal:- Obtain a good upper bound on $L_{init}$

" " Lowerbound for $\Delta_{min}$.



Lemma :- $\Pr[L_{init} > 18n] \le \frac{1}{n!}$




Pf:- If $Y \subseteq [-1,2]^2$

then $L_{init} \le 18n$.

Then enough to estimate $\Pr[Y \notin [-1,2]^2]$.

ie $\Pr[L_{init} > 18n] \le \Pr[Y \notin [-1,2]^2]$.

$Y \notin [-1,2]^2 \Rightarrow \exists\ i$ s.t. $y_i = x_i + \xi_i \notin [-1,2]^2$.

ie $\max\{x_i^1 + \xi_i^1,\ x_i^2 + \xi_i^2\} \ge 1$.

ie $\max\{|\xi_i^1|, |\xi_i^2|\} \ge 1$.

**Lemma** [Tail bound for Gaussian]. $X \sim N(\mu, \sigma^2)$

then $\forall t$ $\Pr[X \geq \mu + \sigma t] = \Pr[X \leq \mu - \sigma t]$
$$\leq \frac{1}{t\cdot\sqrt{2\pi}} \cdot e^{-t^2/2}.$$

set $t = \frac{1}{\sigma}$.

$$\Pr[\xi_i^j \geq 0 + \sigma \cdot \tfrac{1}{\sigma}] \leq \frac{\sigma}{\sqrt{2\pi}} \cdot e^{-\sigma^2/2}.$$

$$\sigma \leq \frac{1}{2\sqrt{n\log n}} \Rightarrow \Pr[\xi_i^1 \geq 1] \leq \left(\frac{1}{n!}\right)^2.$$

Applying union bound over $i \in \{1..n\}$ & $j \in \{1,2\}$.

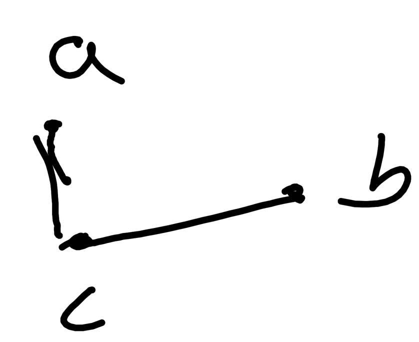
we get $\Pr[Y \notin [-1,2]^2] \leq \frac{1}{n!}$

**To Lower bound $\Delta_{min}$.**

$a, b, c \in \mathbb{R}^2$. $a = (a_1, a_2)$ . .

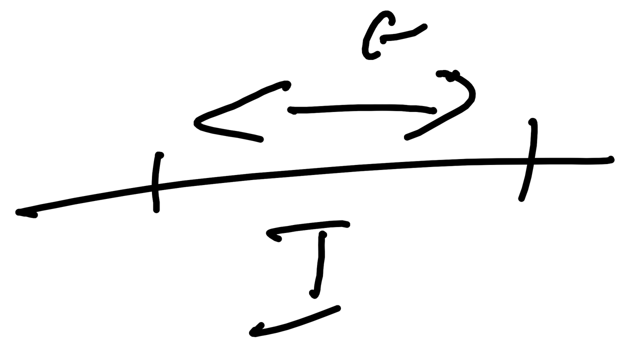
$$d(a,b) = (a_1 - b_1)^2 + (a_2 - b_2)^2$$

Define: $\Delta_{a,b}(c) \stackrel{def}{=} d(c,a) - d(c,b)$




**Lemma**:- $a, b \in \mathbb{R}^2$ & $c \in \mathbb{R}^2$ distributed as Gaussian with $\mu$ & $\sigma$.

$I \subseteq \mathbb{R}$ be any interval of length $\epsilon$



$$\Pr[\Delta_{a,b}(c) \in I] \leq \frac{\epsilon}{4\sigma\sqrt{d(a,b)}}$$

$\hookrightarrow$ Euclidean dist b/n a & b.

Pf Hint: set $a = (0, \dots 0)$ & $b = (\delta, 0 \dots 0)$

**Lemma** $\forall \epsilon > 0 \quad \Pr[\Delta_{min} \leq \epsilon] = O\left(\frac{n^4 \epsilon}{\sigma^2 \sqrt{2}}\right)$.

Assume the Lemma.

**Thm:-** Let $Y \subseteq \mathbb{R}^2$ obtained via a Gaussian perturbation $\mu = 0$ & $\sigma$. Then the maximum # of iterations of the 2-opt algorithm on $Y$ with respect to the squared Euclidean distance is bounded by $O((n^6 \log n)/\sigma)$ in expectation.

**Pf:-** Suppose 2-opt runs for at least $t$ steps.

Then either $\underbrace{\Delta_{min} \leq \frac{18n}{t}}_{A}$ or $\underbrace{L_{init} \geq 18n}_{B}$

$$\Pr[A \text{ or } B] \leq \Pr[A] + \Pr[B]$$

$$\leq \frac{n^4 \cdot 18n}{\sigma^2 t} + \frac{1}{n!}$$

$$= \frac{18n^5}{\sigma^2 t} + \frac{1}{n!}.$$

T : # of iterations of 2-opt on Y

ie $\Pr[T \geq t] \leq \Pr[A \text{ or } B] \leq \frac{18 n^5}{\sigma^2 t} + \frac{1}{n!}$

$$\therefore E[T] = \sum_{t=1}^{n!} \left( \frac{1}{n!} + \frac{18 n^5}{\sigma t} \right)$$

$$= 1 + \frac{18 n^5}{\sigma} \sum_{t=1}^{n!} \frac{1}{t} \quad \leq \log n! \leq n \log n$$

$$\leq 1 + \frac{18 n^6 \log n}{\sigma}$$

$$= O\left( \frac{n^6 \log n}{\sigma} \right)$$

$$E[X] = \int_x \Pr[X \geq x]$$

$$E[X] = \sum_{t=0}^{\infty} \Pr[X \geq t]$$

<u>Lemma [re-stated]</u>

$$\Pr[\Delta_{min} \in (0, \varepsilon]] \leq O\left( \frac{n^4 \varepsilon}{\sigma^2 \cdot \sqrt{2}} \right)$$

<u>Proof:-</u> Considuan arbitrary 2-opt exchange involving vertices $Y_1\, Y_2\, Y_3\, Y_4$ : S

ie edges $(Y_1, Y_2)$ & $(Y_3, Y_4)$ are replaced by $(Y_1, Y_3)$ & $(Y_2, Y_4)$

then $\Delta(S) = d(Y_1, Y_2) + d(Y_3, Y_4) - d(Y_1, Y_3) - d(Y_2, Y_4)$.

$$= \Delta_{y_2 y_3}(y_1) - \Delta_{y_2 y_3}(y_4) \quad \text{(prove!)}$$

notation

$$= \Delta_{23}(1) - \Delta_{23}(4)$$

Suppose adversary chooses $y_3$ and then we draw $y_2$.

this fixes $\delta \stackrel{\Delta}{=} \sqrt{d(y_2, y_3)}$. (Euclidean dist b/n $y_2$ & $y_3$)

Now we draw $y_4$

then this fixes $\Delta_{23}(4)$ (to some value say $\alpha$)

Now, $\Delta(S) < \epsilon$ only if $\Delta_{23}(1) - \Delta_{23}(4) < \epsilon$

only if $\Delta_{23}(1) - \alpha < \epsilon$.

ie ~~the~~ the quantity $\Delta_{23}(1)$ lies in some interval of length at most $\epsilon$

$|\Delta_{23}(1) - \boxed{\Delta_{23}(4)}| \in [0, \epsilon]$

$\hookrightarrow \Delta_{23}(1) \in I$

$|I| = \epsilon$

Then ~~$\Pr[\Delta(S) \le \epsilon]$~~ $\le \dfrac{\epsilon}{4\sigma \sqrt{d(y_2, y_3)}}$

$$\Pr[\Delta(S) \le \epsilon \mid d(y_2, y_3) = \delta^2] \le \frac{\epsilon}{4\sigma\delta} \rightarrow (**)$$

Then $\Pr[\Delta(S) \le \epsilon] = \displaystyle\int_0^{\infty} \Pr[\Delta(S) \le \epsilon \mid d(y_2, y_3) = \delta^2] \, g(\delta) \cdot d\delta$

$A_1 \cup \ldots \cup A_n$

$\Pr[B \mid A_1] \ldots$

$\Pr[B \mid A_n]$

$\Pr[B] = \sum_{i=1}^{n} \Pr[B \mid A_i] \cdot \Pr[A_i]$

where $g(\delta)$ is the density function of $\delta$.

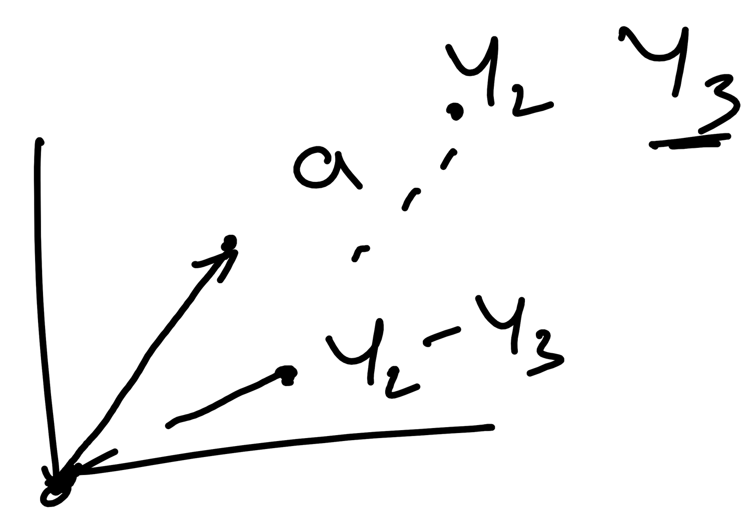
from (**)

$$\le \int_0^{\infty} \frac{\epsilon}{4\sigma\delta} \cdot g(\delta) \cdot d\delta$$

Chi-distribution: $\chi_d$:

is distribution of $\|a\|_2$ where $a$ is distributed according to $N_d(\overset{\mu}{0}, \sigma)$

$a \in \mathbb{R}^d$ as per $N_d(\bar{\mu} = 0, \sigma)$




length of a $= \|a\|_2 = \sqrt{a_1^2 + a_2^2 + \dots a_d^2}$

Obs [without proof]

if $a \sim N_d(0, \sigma)$

then $\|a\|_2 \le \delta$.

($\delta$ is distance b/n $y_2$ & $y_3$)

ie $\frac{1}{\delta} \le \frac{1}{\|a\|_2}$. (ie $\chi_2$).

ie $\frac{1}{\delta}$ stochastically dominated by $\frac{1}{\|a\|_2}$.

---

$\therefore \Pr[\Delta(S) \le \epsilon] \le \frac{\epsilon}{4\sigma} \int^{\infty} \frac{1}{\delta} g(\delta) \cdot d\delta$.

density fn of Chi-distrb.

Chi-d of

$$\le \frac{\epsilon}{4\sigma} \cdot \underbrace{\int_0^{\infty} \frac{1}{\chi_d} f_d(\delta) \cdot d\delta}_{\le O\left(\frac{1}{\sigma}\right) \text{ (Known)}}$$

$$\le O\left(\frac{\epsilon}{4\sigma^2}\right)$$

$$\Pr[\Delta_{min} \le \epsilon] \le \Pr[\exists S \ \Delta(S) \le \epsilon]$$

$$\le \sum_S \Pr[\Delta(S) \le \epsilon]$$

$$\le \sum_S \cdot O\left(\frac{\epsilon}{4\sigma^2}\right)$$

$$\le \underline{O\left(\frac{n^4 \epsilon}{4\sigma^2}\right)}.$$