SML _ Übung 1

Task 1:

· 1a): $\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \ddots & & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & & a_{nn} \end{bmatrix}$ $\quad \underline{\underline{B}} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & \ddots & & b_{2n} \\ \vdots & & & \vdots \\ b_{n1} & \cdots & & b_{nn} \end{bmatrix}$ $\quad \underline{\underline{C}} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & \ddots & & c_{2n} \\ \vdots & & & \vdots \\ c_{n1} & \cdots & & c_{nn} \end{bmatrix}$

① commutative: $(\underline{\underline{A}}\,\underline{\underline{B}}) \cdot \underline{\underline{C}} = \begin{bmatrix} \sum_{i=1}^{n} a_{1i} \cdot b_{i1} & \cdots & \sum_{i=1}^{n} a_{1i} \cdot b_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} a_{ni} \cdot b_{i1} & \cdots & \sum_{i=1}^{n} a_{ni} \cdot b_{in} \end{bmatrix} \cdot \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix}$

$$= \begin{bmatrix} \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{1i} \cdot b_{ij} \right) \cdot c_{j1} & \cdots & \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{1i} \cdot b_{ij} \right) \cdot c_{jn} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ni} \cdot b_{ij} \right) \cdot c_{j1} & \cdots & \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ni} \cdot b_{ij} \right) \cdot c_{jn} \end{bmatrix}$$

$$\underline{\underline{A}} \cdot (\underline{\underline{B}} \cdot \underline{\underline{C}}) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} \sum_{i=1}^{n} b_{1i} \cdot c_{i1} & \cdots & \sum_{i=1}^{n} b_{1i} \cdot c_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} b_{ni} \cdot c_{i1} & \cdots & \sum_{i=1}^{n} b_{ni} \cdot c_{in} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{j=1}^{n} a_{1j} \cdot \left(\sum_{i=1}^{n} b_{ji} \cdot c_{i1} \right) & \cdots & \sum_{j=1}^{n} a_{1j} \cdot \left(\sum_{i=1}^{n} b_{ji} \cdot c_{in} \right) \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{n} a_{nj} \cdot \left(\sum_{i=1}^{n} b_{ji} \cdot c_{in} \right) & \cdots & \sum_{j=1}^{n} a_{nj} \cdot \left(\sum_{i=1}^{n} b_{ji} \cdot c_{in} \right) \end{bmatrix}$$

element analys: $\because \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{1i} \cdot b_{ij} \right) \cdot c_{j1} \neq \sum_{j=1}^{n} a_{1j} \cdot \left(\sum_{i=1}^{n} b_{ji} \cdot c_{i1} \right)$

$\therefore (\underline{\underline{A}} \cdot \underline{\underline{B}}) \cdot \underline{\underline{C}} \neq \underline{\underline{A}} \cdot (\underline{\underline{B}} \cdot \underline{\underline{C}})$

$\Rightarrow$ commutative disproved

② distribute: $(\underline{\underline{A}} + \underline{\underline{B}}) \cdot \underline{\underline{C}} = \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{nn}+b_{nn} \end{bmatrix} \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix}$

$$= \begin{bmatrix} \sum_{i=1}^{n} (a_{1i} + b_{1i}) \cdot c_{i1} & \cdots & \sum_{i=1}^{n} (a_{1i} + b_{1i}) \cdot c_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} (a_{ni} + b_{ni}) \cdot c_{i1} & \cdots & \sum_{i=1}^{n} (a_{ni} + b_{ni}) \cdot c_{in} \end{bmatrix}$$

associative

$$= \begin{bmatrix} \sum_{i=1}^{n} a_{1i} \cdot c_{i1} & \cdots & \sum_{i=1}^{n} a_{1i} \cdot c_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} a_{ni} \cdot c_{i1} & \cdots & \sum_{i=1}^{n} a_{ni} \cdot c_{in} \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^{n} b_{1i} \cdot c_{i1} & \cdots & \sum_{i=1}^{n} b_{1i} \cdot c_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} b_{ni} \cdot c_{i1} & \cdots & \sum_{i=1}^{n} b_{ni} \cdot c_{in} \end{bmatrix}$$

$= \underline{\underline{A}} \cdot \underline{\underline{C}} + \underline{\underline{B}} \cdot \underline{\underline{C}}$ $\Rightarrow$ distribute proved

③ associative: $(\underline{\underline{A}} + \underline{\underline{B}}) + \underline{\underline{C}} = \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \cdots & a_{nn}+b_{nn} \end{bmatrix} + \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{bmatrix}$

$$= \begin{bmatrix} a_{11}+b_{11}+c_{11} & \cdots & a_{1n}+b_{1n}+c_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1}+c_{n1} & \cdots & a_{nn}+b_{nn}+c_{nn} \end{bmatrix}$$

$$\underline{\underline{A}} + (\underline{\underline{B}} + \underline{\underline{C}}) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} + \begin{bmatrix} b_{11}+c_{11} & \cdots & b_{1n}+c_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1}+c_{n1} & \cdots & b_{nn}+c_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}+b_{11}+c_{11} & \cdots & a_{1n}+b_{1n}+c_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1}+c_{n1} & \cdots & a_{nn}+b_{nn}+c_{nn} \end{bmatrix}$$

$\Rightarrow (\underline{\underline{A}} + \underline{\underline{B}}) + \underline{\underline{C}} = \underline{\underline{A}} + (\underline{\underline{B}} + \underline{\underline{C}})$

$\Rightarrow$ associative proved

1b) ① method 1: elementary transformation:

$$[\underline{\underline{A}} \ \ \underline{\underline{I}}] \xrightarrow{\text{elementary transformation:}} [\underline{\underline{I}} \ \ \underline{\underline{A}}^{-1}]$$

$$= \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 4 & 6 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{3}{2} \\ 1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

② methode 2: use adjoint matrix

$$\underline{\underline{A}}^{-1} = \frac{1}{|A|} \underbrace{\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix}}_{\text{adjoint}}$$

wherein A_{ij} is algebraic cofactor

$$= \frac{1}{\det(\underline{\underline{A}})} \cdot \begin{bmatrix} 4 & -2 & 0 \\ -1 & 1 & -3 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

with $\det(\underline{\underline{A}}) = |A| = 2$

If A is changed to $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 6 \\ 1 & 0 & 0 \end{bmatrix}$

$\det(\underline{\underline{A}}) = 0 \Rightarrow$ A is a singular matrix

$\Rightarrow$ irreversible

1c) for $A \in \mathbb{R}^{n \times m}$

① left Moore - Penrose pseudoinverse: $\underline{\underline{A}}^{+}_{left} = (\underline{\underline{A}}^{T} \cdot \underline{\underline{A}})^{-1} \cdot \underline{\underline{A}}^{T}$

right Moore - Penrose pseudoinverse: $\underline{\underline{A}}^{+}_{right} = \underline{\underline{A}}^{T} \cdot (\underline{\underline{A}} \cdot \underline{\underline{A}}^{T})^{-1}$

② dimensional analysis

if $A \in \mathbb{R}^{2\times 3}$

$\underline{\underline{A}}^{+}_{left} = (\underline{\underline{A}}^{T} \cdot \underline{\underline{A}})^{-1} \cdot \underline{\underline{A}}^{T}$

$\rightarrow (3\times2)\cdot(2\times3)\cdot(3\times2) \rightarrow (3\times2)$

$A^{+}_{right} = \underline{\underline{A}}^{T} \cdot (\underline{\underline{A}} \cdot \underline{\underline{A}}^{T})^{-1}$

$\rightarrow (3\times2)\cdot[(2\times3)\cdot(3\times2)] \rightarrow (3\times2)$

$\Rightarrow A^{+} \in \mathbb{R}^{3\times2}$

1d): eigenvector : represent the orientation of each principal components

eigen value : represent the weight of each principal components

both of them represent the components distribution of all features

(PCA - perspective)

Task 2

2a):

i): for Ω is a finite set

① Expectation: $E[f(w)] = \sum_{w \in \Omega} f(w) \cdot P(w)$

Variance : $V[f(w)] = \sum_{w \in \Omega} P(w) \cdot (f(w) - E(w))^2$

$= E[(f(w) - E[f(w)])^2]$

② Expectation and Variance are linear operators

Definition and first consequences [edit]

Let V and W be vector spaces over the same field K. A function $f : V \to W$ is said to be a *linear map* if for any two vectors $\mathbf{u}, \mathbf{v} \in V$ and any scalar $c \in K$ the following two conditions are satisfied:

$f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$ additivity / operation of addition

$f(c\mathbf{u}) = cf(\mathbf{u})$ homogeneity of degree 1 / operation of scalar multiplication

For Expectation : the rules of expectation prove it :

$E[a\underline{X}] = aE[\underline{X}]$

$E[\underline{X} + \underline{y}] = E[\underline{X}] + E(\underline{y})$

$E[\sum_i a_i X_i] = \sum_i a_i E[X_i]$

Variance is essentially an expectation operation :

$V[f(w)] = E[(f(w) - E[f(w)])^2]$

$= E[(f(w))^2] - 2E[f(w)] \cdot E[f(w)] + (E[f(w)])^2$

$= E[(f(w))^2] - E[f(w)]^2$

$= \sum_{w \in \Omega} P(w) \cdot (f(w))^2 - E[f(w)]^2$

	1	2	3	4	5	6
A	4	1	6	2	1	4
B	5	6	1	1	4	1
C	3	3	4	2	3	3

2)

	1	2	3	4	5	6
A	4	1	6	2	1	4
B	5	6	1	1	4	1
C	3	3	4	2	3	3

X	1	2	3	4	5	6
P(X|A)	2/9	1/18	1/3	1/9	1/18	2/9
P(X|B)	5/18	1/3	1/18	1/18	2/9	1/18
P(X|C)	1/6	1/6	2/9	1/9	1/6	1/6

For dice A, set $P(A)=1 \Rightarrow P(x,A) = P(x|A)\cdot P(A) = P(x|A) = P_A(x)$

$$\bar{X}_A = \sum_{i=1}^{6} X_i \cdot P(X_i, A) = 1\cdot 2/9 + 2\cdot 1/18 + 3\cdot 1/3 + 4\cdot 1/9 + 5\cdot 1/18 + 6\cdot 2/9$$
$$= \frac{61}{18} \approx 3,3889$$

unbiased $V[(X,A)] = \frac{1}{N-1} \sum_{i=1}^{N=18} (X_i - \bar{X}_A)^2$ (Bessel's Correction)

$$= \frac{1}{17}\cdot\left(4\cdot\left(1-\frac{61}{18}\right)^2 + 1\cdot\left(2-\frac{61}{18}\right)^2 + 6\cdot\left(3-\frac{61}{18}\right)^2 + 2\cdot\left(4-\frac{61}{18}\right)^2 + 1\cdot\left(5-\frac{61}{18}\right)^2 + 4\cdot\left(6-\frac{61}{18}\right)^2\right)$$

$$= \frac{1013}{306} \approx 3.3105$$

For dice B, set $P(B)=1 \Rightarrow P(X,B) = P(X|B)\cdot P(B) = P(X|B) = P_B(X)$

$$\bar{X}_B = \sum_{i=1}^{6} X_i \cdot P(X_i, B) = 1\cdot 5/18 + 2\cdot 1/3 + 3\cdot 1/18 + 4\cdot 1/18 + 5\cdot 2/9 + 6\cdot 1/18$$
$$= 25/9 \approx 2.7778$$

unbiased $V[(X,B)] = \frac{1}{N-1} \sum_{i=1}^{N=18} (X_i - \bar{X}_B)^2$

$$= \frac{1}{17}\left(5\cdot\left(1-\frac{25}{9}\right)^2 + 6\cdot\left(2-\frac{25}{9}\right)^2 + 1\cdot\left(3-\frac{25}{9}\right)^2 + 1\cdot\left(4-\frac{25}{9}\right)^2 + 4\cdot\left(5-\frac{25}{9}\right)^2 + 1\cdot\left(6-\frac{25}{9}\right)^2\right)$$

$$= \frac{460}{153} \approx 3,0065$$

For dice C, set $P(C)=1 \Rightarrow P(X,C) = P(X|C)\cdot P(C) = P(X|C) = P_C(X)$

$$\bar{X}_C = \sum_{i=1}^{6} X_i \cdot P(X_i, C) = 1\cdot 1/6 + 2\cdot 1/6 + 3\cdot 2/9 + 4\cdot 1/9 + 5\cdot 1/6 + 6\cdot 1/6$$
$$= \frac{31}{9} \approx 3,4444$$

unbiased $V[(X,C)] = \frac{1}{N-1} \sum_{i=1}^{N=18} (X_i - \bar{X}_C)^2$

$$= \frac{1}{17}\cdot\left(3\cdot\left(1-\frac{31}{9}\right)^2 + 3\cdot\left(2-\frac{31}{9}\right)^2 + 4\cdot\left(3-\frac{31}{9}\right)^2 + 2\cdot\left(4-\frac{31}{9}\right)^2 + 3\cdot\left(5-\frac{31}{9}\right)^2 + 3\cdot\left(6-\frac{31}{9}\right)^2\right)$$

$$= \frac{472}{153} \approx 3,0850$$

Dice	Expectation	Variance
A	3,3889	3.3105
B	2,7778	3,0065
C	3,4444	3,0850

3) KL-divergence $= -\sum_x P_i(x) \cdot \ln \frac{Q(x)}{P_i(x)} = \sum_x P_i(x) \cdot \ln \frac{P_i(x)}{Q(x)}$ $i =$ A, B or C

for unbaised distribution $Q(x)$, $\mu = 3.5$ $\sigma = \frac{35}{12}$

$\Rightarrow Q(x) = \frac{1}{6}$ for $x = 1, 2, \cdots, 6$

Dice A: $KL = \sum_{x=1}^{6} P_A(x) \cdot \ln \frac{P_A(x)}{Q(x)}$

$= \frac{2}{9} \cdot \ln(\frac{2}{9}/\frac{1}{6}) + \frac{1}{18} \cdot \ln(\frac{1}{18}/\frac{1}{6}) + \frac{1}{3} \cdot \ln(\frac{1}{3}/\frac{1}{6}) + \frac{1}{9} \ln(\frac{1}{9}/\frac{1}{6}) + \frac{1}{18} \ln(\frac{1}{18}/\frac{1}{6}) + \frac{2}{9} \cdot \ln(\frac{2}{9}/\frac{1}{6})$

$= 0.1918$

Dice B: $KL = \sum_{x=1}^{6} P_B(x) \cdot \ln \frac{P_B(x)}{Q(x)}$

$= \frac{5}{18} \cdot \ln(\frac{5}{18}/\frac{1}{6}) + \frac{1}{3} \cdot \ln(\frac{1}{3}/\frac{1}{6}) + \frac{1}{18} \cdot \ln(\frac{1}{18}/\frac{1}{6}) + \frac{1}{18} \ln(\frac{1}{18}/\frac{1}{6}) + \frac{2}{9} \cdot \ln(\frac{2}{9}/\frac{1}{6}) + \frac{1}{18} \ln(\frac{1}{18}/\frac{1}{6})$

$= 0.2538$

Dice C: $KL = \sum_{x=1}^{6} P_C(x) \cdot \ln \frac{P_C(x)}{Q(x)}$

$= \frac{1}{6} \cdot \ln(\frac{1}{6}/\frac{1}{6}) + \frac{1}{6} \cdot \ln(\frac{1}{6}/\frac{1}{6}) + \frac{2}{9} \cdot \ln(\frac{2}{9}/\frac{1}{6}) + \frac{1}{9} \cdot \ln(\frac{1}{9}/\frac{1}{6}) + \frac{1}{6} \cdot \ln(\frac{1}{6}/\frac{1}{6}) + \frac{1}{6} \ln(\frac{1}{6}/\frac{1}{6})$

$= 0.0189$

2b)

1) $A:\{$ a Person has a cold $\}$ $\quad$ $B:\{$ a person has back-pain. $\}$

2) $C:\{$ Word population $\}$ $\quad$ $D:\{$ heathy people, who do not have cold $)$

3) ① $\{$ A person with a cold has back-pain 30% of the time $\}$

$$\Rightarrow \quad P(B|A) = 30\%$$

② $\{$ 3% of the world population has a cold $\}$

$$\Rightarrow P(A|C) = 3\% \quad \text{mit} \quad P(C) = 1 \Rightarrow P(A) = 3\%$$

③ $\{$ 10% of those who do not have a cold still have back-pain $\}$

$$\Rightarrow P(B|D) = 10\% \quad \text{mit} \quad P(D) = 1 - P(A|C) = 97\%$$

4) $$P(B) = P(B,A) + P(B,D)$$
$$= P(B|A)\cdot P(A) + P(B|D)\cdot P(D)$$
$$= 30\%\cdot 3\% + 10\%\cdot 97\% = 10.6\%$$

Bayes' theorem

$\Downarrow$

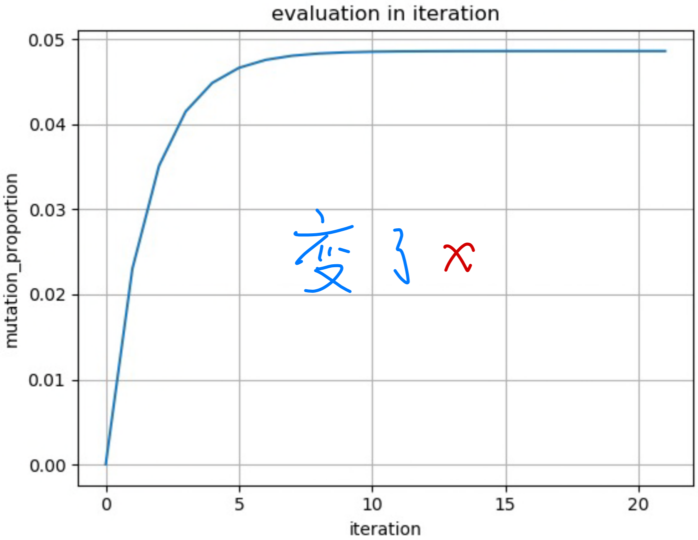
$$P(A|B) = \frac{P(B|A)\cdot P(A)}{P(B)}$$
$$= \frac{30\%\cdot 3\%}{10.6\%} = 8.49\%$$

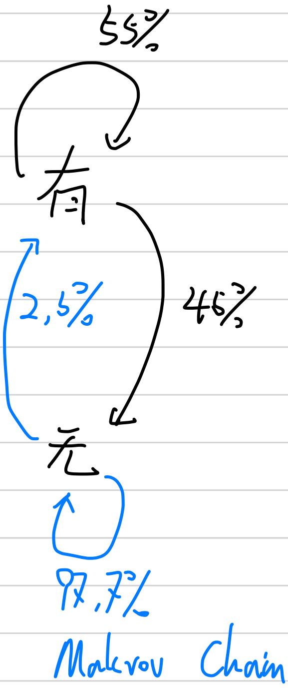
2c) Markov chain

1): $\vec{S}_t = \begin{pmatrix} m \\ \hat{m} \end{pmatrix}$ with probabilty matrix: $I = \begin{bmatrix} 55\% & 2,3\% \\ 45\% & 97,7\% \end{bmatrix}$

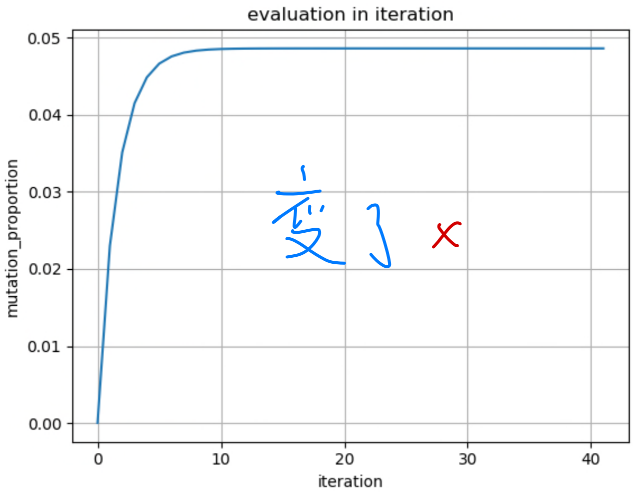
$$\Rightarrow \vec{S}_{t+1} = \underline{I} \cdot \underline{S_t} = \begin{bmatrix} 45\% & 2,3\% \\ 55\% & 97,7\% \end{bmatrix} \begin{bmatrix} m \\ \hat{m} \end{bmatrix} = \begin{pmatrix} 45\% m + 2,3\% \hat{m} \\ 55\% m + 97,7\% \hat{m} \end{pmatrix}$$

2)





3)



approximate at ~~5~~ 20 step will culture contain stop to change mutation

if $45\% m + 2.3\% \hat{m} = m$ mit $\hat{m} = (1-m)$ $\Rightarrow m \approx 4{,}01\%$

$\Rightarrow$ Evaluation process got stabel probability

Task 3

3a) $h(p_i) = -\log_2 p_i$

	S_1	S_2	S_3	S_4
p_i	0.02	0.67	0.23	0.08
$h(p_i)$	5,6439	0,5778	2,1203	3,6439

① average information: $H(p) = E[h(p)] = \sum_{i=1}^{4} p_i h(p_i) = 1,2792$ bit

② In general: max: $\log_2(4) = 2$ bit

③ distribution? ~~高斯分布~~? 平均分布? ⇐ 有序程度最高

	S_1	S_2	S_3	S_4
p_i	0,25	0,25	0,25	0,25
$h(p_i)$	2	2	2	2

$\Rightarrow H(p) = 4 \cdot 0.25 \cdot 2 = 2$ bit

3b) 1) additional condition : $1 = \sum_{i=1}^{4} p_i$ with $1 \geq p_i \geq 0$, $i = 1, 2, 3$ or 4

2) Cost function : $\max_{p} H(p) = -\sum_{i=1}^{4} p_i \ln(p_i)$ (here Entrope with unit "nat" not unit "bit")

$$\text{s.t. } f_1(p) = \sum_{i=1}^{4} 2p_i \cdot i - 6 = 0$$

$$f_2(p) = \sum_{i=1}^{4} p_i - 1 = 0$$

Lagrangian $L(\underline{p}, \lambda) = -\sum_{i=1}^{4} p_i \ln(p_i) + \lambda_1 \left(\sum_{i=1}^{4} 2p_i \cdot i - 6 \right) + \lambda_2 \left(\sum_{i=1}^{4} p_i - 1 \right)$

3) $$\frac{\partial L}{\partial \lambda_1} = \sum_{i=1}^{4} 2p_i \cdot i - 6$$

$$\frac{\partial L}{\partial \lambda_2} = \sum_{i=1}^{4} p_i = 1 \Rightarrow \text{monotone increasing}$$

(Vector) $\underline{P}$ is correlative with $\frac{\partial L}{\partial \lambda_1} \Rightarrow$ very hard to analytically solve

4) Dual Formulation

$$\max_{p} H(p) = -\sum_{i=1}^{4} p_i \ln(p_i)$$

$$\text{s.t. } f_1(p) = \sum_{i=1}^{4} 2p_i \cdot i - 6 = 0$$

$$f_2(p) = \sum_{i=1}^{4} p_i - 1 = 0$$

① $$L(\underline{p}, \lambda) = -\sum_{i=1}^{4} p_i \ln(p_i) + \lambda_1 \left(\sum_{i=1}^{4} 2p_i \cdot i - 6 \right) + \lambda_2 \left(\sum_{i=1}^{4} p_i - 1 \right)$$

② $$\frac{\partial L}{\partial p_i} = -\ln(p_i) + 1 + \lambda_1 2i + \lambda_2 = 0$$

$$\Rightarrow \ln(p_i) = 1 + 2i \cdot \lambda_1 + \lambda_2$$

$$p_i^* = e^{(1 + 2i \cdot \lambda_1 + \lambda_2)}$$

③ replace p_i^* in $L(\underline{p}, \underline{\lambda})$

$$\Rightarrow G(\underline{\lambda}) = -\sum_{i=1}^{4} (1 + \cancel{2i\lambda_1} + \cancel{\lambda_2}) \cdot e^{(1+2i \cdot \lambda_1 + \lambda_2)} + \left(\cancel{\sum_{i=1}^{4} 2i\lambda_1 \cdot e^{(1+2i\lambda_1 + \lambda_2)}} - 6\lambda_1 \right) + \left(\cancel{\sum_{i=1}^{4} \lambda_2 \cdot e^{(1+2i\lambda_1 + \lambda_2)}} - \lambda_2 \right)$$

$$= -\sum_{i=1}^{4} e^{(1+2i\lambda_1 + \lambda_2)} - 6\lambda_1 - \lambda_2$$

assume that $G(\underline{x})$ is a convex function

$$\begin{cases} \dfrac{\partial G}{\partial \lambda_1} = -\sum_{i=1}^{4} 2i \cdot e^{(1+2i\lambda_1+\lambda_2)} - 6 = 0 \\ \dfrac{\partial G}{\partial \lambda_2} = -\sum_{i=1}^{4} e^{(1+2i\lambda_1+\lambda_2)} - 1 = 0 \end{cases}$$

$\Rightarrow$ $\lambda_1 =$ solve_1 $\qquad \lambda_2 =$ solve_2

```
solve_1 =

log(6*exp(1)*root(32*z^3*exp(3) + 32*z^2*exp(2) + 10*z*exp(1) - 1, z, 1) + 8*exp(2)*root(32*z^3*exp(3) + 32*z^2*exp(2) + 10*z*exp(1) - 1, z, 1)^2 + 1)/2


solve_2 =

log(root(32*z^3*exp(3) + 32*z^2*exp(2) + 10*z*exp(1) - 1, z, 1))

fx >>
```

5) 最速下降

Quasi - Newton

或者 CG-Verfahren 随意描述一个

3c)

3d) Python