Haoying Zhou U48070579

CS542 HW1

1. **Probability**

   (a) Bishop 1.6

   (b) Table 1 represents a hypothetical study of a test for HIV in a population of intravenous drug users, some carrying HIV and others not. "+" denotes a positive outcome from the test and "–" denotes a negative outcome.

Outcome	with HIV	without HIV	Total
+	72	12	84
–	3	71	74
Total	75	83	158

Table 1.1. HIV test summary for intravenous drug users

   i. Calculate the probability of test being positive given the patient has HIV.
   ii. Calculate the probability of test being negative given the patient has HIV.
   iii. If the prevalence of HIV is 12% in intravenous drug users (12 cases per 100 patients), what is the probability of the patient having HIV after a positive test? A negative test?

(a) Bishop 1.6

1.6 (⋆) Show that if two variables x and y are independent, then their covariance is zero.

Proof: Since x and y are two independent variables, then, it has:

$$\mathrm{p}(\mathrm{x},\mathrm{y}) = p_x(x)p_y(y)$$

And according to the definition of expectation:

$$\mathrm{E}[\mathrm{x}] = \int xp_x(x)\,dx, \qquad \mathrm{E}[\mathrm{y}] = \int yp_y(y)\,dy$$

$$E_{x,y}[xy] = \iint xyp(x,y)dxdy$$

Then, it can derive that:

$$\iint xyp(x,y)dxdy = \iint xp_x(x)yp_y(y)dxdy = \left(\int xp_x(x)\,dx\right)\left(\int yp_y(y)\,dy\right)$$

Thus, it can get that:

$$E_{x,y}[xy] = \mathrm{E}[\mathrm{x}]\mathrm{E}[\mathrm{y}]$$

And calculate the covariance:

$$cov[x,y] = E_{x,y}[xy] - E[x]E[y] = 0$$

(b) Solution:

i. According to the table, it can directly obtain the probability of the patient has HIV:

$$p(HIV) = \frac{75}{158}$$

And the probability of the patient has HIV and test being positive is:

$$p(HIV\ and\ +) = \frac{72}{158}$$

it can calculate the probability of test being positive given the patient has HIV:

$$p(+|HIV) = \frac{p(HIV\ and\ +)}{p(HIV)} = \frac{24}{25} = 0.96$$

ii. According to the table, it can directly obtain the probability of the patient has HIV:

$$p(HIV) = \frac{75}{158}$$

And the probability of the patient has HIV and test being negative is:

$$p(HIV\ and\ -) = \frac{3}{158}$$

it can calculate the probability of test being positive given the patient has HIV:

$$p(-|HIV) = \frac{p(HIV\ and\ +)}{p(HIV)} = \frac{1}{25} = 0.04$$

iii. Since the probability of prevalence of HIV in intravenous drug users is:

$$p(pHIV) = 12\% = 0.12$$

Then, it can obtain the probability of non-prevalence of HIV in intravenous drug users is

$$p(\overline{pHIV}) = 1 - p(pHIV)\ = 0.88$$

And according to the table, it can easily obtain that:

$$p(HIV|+) = p(A) = \frac{72}{84} = \frac{6}{7}, p(\overline{HIV}|+) = p(B) = \frac{12}{84} = \frac{1}{7}$$
$$p(HIV|-) = p(C) = \frac{3}{74}, p(\overline{HIV}|+) = p(D) = \frac{71}{74}$$

Then, it can calculate the probabilities of the patient may get prevalence after positive/negative are:

$$p(pos) = p(A) * p(pHIV) + p(B) * p(\overline{pHIV})$$
$$p(neg) = p(C) * p(pHIV) + p(D) * p(\overline{pHIV})$$

Hence, according to Bayes Theorem, it can calculate the probabilities of the patient having HIV after a positive/negative test:

$$p(apos) = \frac{p(A) * p(pHIV)}{p(pos)} = \frac{9}{20}$$

$$p(aneg) = \frac{p(C) * p(pHIV)}{p(neg)} = \frac{9}{1571}$$

2. **Bayes Theorem**

(a) Monty Hall problem: On the game show, Let's make a Deal, you are shown four doors: A, B, C, and D, and behind exactly one of them is a big prize. Michael, the contestant, selects one of them, say door C, because he know from having watched countless number of past shows that door C has twice the probability of being the right door than door A or door B or door D. To make things more interesting, Monty Hall, game show host, opens one of the other doors, say door B, revealing that the big prize is not behind door B. He then offers Michael the opportunity to change the selection to one of the remaining doors (door A or door D). Should Michael change his selection? Justify your answer by calculating the probability that the prize is behind door A, the probability that the prize is behind door C, and the probability that the prize is behind door D, given that Monty Hall opened door B (to show that prize is not there), using Bayes Theorem.

(b) Prof. Chin knows that historically 2 out 75 students in his CS 542 class cheats (!) on his exams. Last semester, he suspected one of the students engaged in cheating during the final, but when he gently confronted the student, the student vehemently denied that he was cheating. If Prof. Chin has 90% accuracy in identifying cheaters (i.e. if a student is cheating, 90% of the time, Prof. Chin will indeed identify that the student as a cheater), but also has 20% false alarm rate (i.e. even though a student is not cheating, Prof. Chin will erroneously identify that student as a cheater). What is the probability that the student Prof. Chin confronted in the last semester's final was indeed cheating?

(a) Solution: according to the description, the game follow these rules:
1. host cannot choose the door which is chosen by player
2. host cannot choose nonempty door
3. host should choose the door randomly

And $\mathrm{p(C) = 2p(A) = 2p(B) = 2p(D)}$, $\mathrm{p(A) + p(B) + p(C) + p(D) = 1}$

Then, it can solve that:

$$\mathrm{p(C) = 0.4, p(A) = p(B) = p(D) = 0.2}$$

After that, it can directly obtain the probability of open door B, B is empty when the prize is in door A/B/C/D:

$$p(B\ open, \bar{B}|A) = \frac{1}{2}$$

$$p(B\ open, \bar{B}|B) = 0$$
$$p(B\ open, \bar{B}|C) = \frac{1}{3}$$
$$p(B\ open, \bar{B}|D) = \frac{1}{2}$$

And the probability of open door B, B is empty can be calculated as:

$$p(B\ open, \bar{B}) = p(B\ open, \bar{B}|A)p(A) + p(B\ open, \bar{B}|B)p(B)$$
$$+p(B\ open, \bar{B}|C)p(C) + p(B\ open, \bar{B}|D)p(D) = \frac{1}{3}$$

According to Bayes theorem, the probability of the big prize is in door A/C/D given the door B is open and empty is:

$$p(A|B\ open, \bar{B}) = \frac{p(B\ open, \bar{B}|A)p(A)}{p(B\ open, \bar{B})} = 0.4$$
$$p(C|B\ open, \bar{B}) = \frac{p(B\ open, \bar{B}|C)p(C)}{p(B\ open, \bar{B})} = 0.3$$
$$p(D|B\ open, \bar{B}) = \frac{p(B\ open, \bar{B}|D)p(D)}{p(B\ open, \bar{B})} = 0.4$$

And he should change his selection.

(b) Solution: According to the description, the probability of cheating/not cheating is:

$$p(C) = \frac{2}{75}, \qquad p(\bar{C}) = \frac{73}{75}$$

And the probability of identifying/not identifying when the student is cheating/not cheating is:

$$p(iC|C) = 90\% = 0.9, p(iC|\bar{C}) = 20\% = 0.2$$

Hence, according to Bayes Theorem, the probability of identifying cheating when indeed cheating is:

$$p(C|iC) = \frac{p(iC|C)p(C)}{p(iC|C)p(C) + p(iC|\bar{C})p(\bar{C})} = \frac{9}{82}$$

3. **Linear Algebra**

(a) Bishop 2.22

(b) Find the eigenvalues and eigenvectors for

$$A = \begin{bmatrix} 3 & 4 & -1 \\ -1 & -2 & 1 \\ 3 & 9 & 0 \end{bmatrix}$$

Is A positive definite?

(a) Bishop 2.22

2.22 ($\star$) **www** Show that the inverse of a symmetric matrix is itself symmetric.

Proof: Suppose A is a symmetric matrix and it is invertible
Since $I = I^T$
Then, it can derive that:

$$AA^{-1} = (AA^{-1})^T$$

According to transpose properties:

$$(AA^{-1})^T = (A^{-1})^T A^T$$

Since A is symmetric, then it has to be square matrix, thus:

$$AA^{-1} = A^{-1}A = I$$

Thus, the equation becomes:

$$A^{-1}A = (A^{-1})^T A^T$$

And $A^T = A$
Then, multiply A^{-1} to both sides and at right of the original element, it can obtain that:

$$A^{-1}AA^{-1} = (A^{-1})^T AA^{-1}$$

Since $AA^{-1} = I$
Hence, it can derive that:

$$A^{-1} = (A^{-1})^T$$

(b) Solution: Since $A = \begin{bmatrix} 3 & 4 & -1 \\ -1 & -2 & 1 \\ 3 & 9 & 0 \end{bmatrix}$

Then, $A - \lambda I = \begin{bmatrix} 3-\lambda & 4 & -1 \\ -1 & -2-\lambda & 1 \\ 3 & 9 & -\lambda \end{bmatrix}$

To solve the eigenvalues, solve following equation:

$$\begin{aligned}\det(\mathrm{A}-\lambda\mathrm{I}) = 0 &= 3(4-2-\lambda) - 9(3-\lambda-1) - \lambda[(\lambda+2)(\lambda-3)+4] \\ &= 6 - 3\lambda - 18 + 9\lambda - \lambda^3 + \lambda^2 + 2\lambda = -\lambda^3 + \lambda^2 + 8\lambda - 12 \\ &= -\lambda^3 - 3\lambda^2 + 4\lambda^2 + 8\lambda - 12 = -\lambda^2(\lambda+3) + 4(\lambda+3)(\lambda-1) \\ &= -(\lambda+3)(\lambda^2 - 4\lambda + 4) = -(\lambda+3)(\lambda-2)^2\end{aligned}$$

Thus, the eigenvalues of matrix A are:

$$\lambda_1 = -3, \qquad \lambda_{2,3} = 2$$

Since not all eigenvalues are positive, thus it is not positive definite.

To find the eigenvectors:

For $\lambda_1 = -3$

$$A - \lambda_1 I = A + 3I = \begin{bmatrix} 6 & 4 & -1 \\ -1 & 1 & 1 \\ 3 & 9 & 3 \end{bmatrix}$$

And $(A - \lambda_1 I)v_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 & 4 & -1 \\ -1 & 1 & 1 \\ 3 & 9 & 3 \end{bmatrix} v_1$

Thus, it can easily solve that:

$$v_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

For $\lambda_2 = \lambda_2 = \lambda = 2$

$$A - \lambda I = A - 2I = \begin{bmatrix} 1 & 4 & -1 \\ -1 & -4 & 1 \\ 3 & 9 & -2 \end{bmatrix}$$

It can easily see that the above matrix is not full-rank. Thus it’s a Jordan canonical form when diagonalizing.

And $(A - \lambda I)v = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 & 4 & -1 \\ -1 & 1 & 1 \\ 3 & 9 & 3 \end{bmatrix} v$

Thus, it can easily solve that:

$$v = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}$$

Then, to get the JCF eigenvector:

$$(A - \lambda I)^2 = (A - 2I)^2 = \begin{bmatrix} 1 & 4 & -1 \\ -1 & -4 & 1 \\ 3 & 9 & -2 \end{bmatrix} \begin{bmatrix} 1 & 4 & -1 \\ -1 & -4 & 1 \\ 3 & 9 & -2 \end{bmatrix} = \begin{bmatrix} -6 & -21 & 5 \\ 6 & 21 & -5 \\ -12 & -42 & 10 \end{bmatrix}$$

Then, find the vectors satisfy following equation:

$$(A - \lambda I)^2 v = 0 = \begin{bmatrix} -6 & -21 & 5 \\ 6 & 21 & -5 \\ -12 & -42 & 10 \end{bmatrix} v$$

Thus, it can solve that:

$$v = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}, v_2 = \begin{bmatrix} 13 \\ -3 \\ 3 \end{bmatrix}$$

Therefore, the last eigenvector is:

$$v_3 = (A - 2I)v_2 = \begin{bmatrix} -2 \\ 2 \\ -6 \end{bmatrix}$$

4. **Probability Distributions**

Bishop 2.2, 2.10*, 2.12, 2.15*

(a) Bishop 2.2

2.2 ($\star\star$) The form of the Bernoulli distribution given by (2.2) is not symmetric between the two values of x. In some situations, it will be more convenient to use an equivalent formulation for which $x \in \{-1, 1\}$, in which case the distribution can be written

$$p(x|\mu) = \left(\frac{1-\mu}{2}\right)^{(1-x)/2} \left(\frac{1+\mu}{2}\right)^{(1+x)/2} \tag{2.261}$$

where $\mu \in [-1, 1]$. Show that the distribution (2.261) is normalized, and evaluate its mean, variance, and entropy.

Solution: Since the distribution is:

$$p(x|\mu) = \left(\frac{1-\mu}{2}\right)^{\frac{1-x}{2}} \left(\frac{1+\mu}{2}\right)^{\frac{1+x}{2}}$$

And x ∈ $[-1.1]$

Thus, it can derive that:

$$\sum_{x_i=-1,1} p(x_i) = \frac{1-\mu}{2} + \frac{1+\mu}{2} = 1$$

Hence this distribution is normalized.

Based on definition, it can calculate its mean:

$$\mathrm{E[x]} = \sum_{x_i=-1,1} x_i p(x_i) = (-1) \cdot \frac{1-\mu}{2} + 1 \cdot \frac{1+\mu}{2} = \mu$$

And it can get its variance:

$$\mathrm{var[x]} = \sum_{x_i=-1,1} (x_i - E[x])^2 p(x_i)$$

$$= (-1-\mu)^2 \cdot \frac{1-\mu}{2} + (1-\mu)^2 \cdot \frac{1+\mu}{2} = (1-\mu)^2$$

Then, according to the formula, it can obtain its entropy:

$$H[x] = -\sum_{x_i=-1,1} p(x_i)\ln p(x_i) = -\frac{1-\mu}{2}\ln\frac{1-\mu}{2} - \frac{1+\mu}{2}\ln\frac{1+\mu}{2}$$

(b) Bishop 2.10

2.10 ($\star\star$) Using the property $\Gamma(x+1) = x\Gamma(x)$ of the gamma function, derive the following results for the mean, variance, and covariance of the Dirichlet distribution given by (2.38)

$$\mathbb{E}[\mu_j] = \frac{\alpha_j}{\alpha_0} \qquad (2.273)$$

$$\mathrm{var}[\mu_j] = \frac{\alpha_j(\alpha_0-\alpha_j)}{\alpha_0^2(\alpha_0+1)} \qquad (2.274)$$

$$\mathrm{cov}[\mu_j\mu_l] = -\frac{\alpha_j\alpha_l}{\alpha_0^2(\alpha_0+1)}, \qquad j \neq l \qquad (2.275)$$

where α_0 is defined by (2.39).

Solution: Since the Dirichlet distribution is:

$$Dir(\mu|\alpha) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)}\prod_{k=1}^{K}\mu_k^{\alpha_k-1}$$

And $\int Dir(\mu|\alpha)\,d\mu = 1,\ \ \Gamma(x+1) = x\Gamma(x)$

Then, based on the definition of mean, it can calculate that:

$$E[\mu_j] = \int \mu_j \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)}\prod_{k=1}^{K}\mu_k^{\alpha_k-1}\,d\mu = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)}\int \mu_j \prod_{k=1}^{K}\mu_k^{\alpha_k-1}\,d\mu$$
$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)}\frac{\Gamma(\alpha_1)\cdots\Gamma(\alpha_j+1)\cdots\Gamma(\alpha_K)}{\Gamma(\alpha_0+1)}\int Dir(\mu|\beta)\,d\mu$$

Where in $Dir(\mu|\beta), \beta_0 = \alpha_0+1, \beta_j = \alpha_j+1$

Thus, it can obtain that:

$$E[\mu_j] = \frac{\Gamma(\alpha_0)\Gamma(\alpha_j+1)}{\Gamma(\alpha_j)\Gamma(\alpha_0+1)}\cdot 1 = \frac{\alpha_j}{\alpha_0}$$

After, it's very similar when calculating $E[\mu_j^2]$

$$E[\mu_j^2] = \int \mu_j^2 \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)}\prod_{k=1}^{K}\mu_k^{\alpha_k-1}\,d\mu = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)}\int \mu_j^2 \prod_{k=1}^{K}\mu_k^{\alpha_k-1}\,d\mu$$
$$= \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)}\frac{\Gamma(\alpha_1)\cdots\Gamma(\alpha_j+2)\cdots\Gamma(\alpha_K)}{\Gamma(\alpha_0+2)}\int Dir(\mu|\gamma)\,d\mu$$

Where in $Dir(\mu|\gamma), \gamma_0 = \alpha_0+2, \gamma_j = \alpha_j+2$

Thus, it can obtain that:

$$E[\mu_j^2] = \frac{\Gamma(\alpha_0)\Gamma(\alpha_j+2)}{\Gamma(\alpha_j)\Gamma(\alpha_0+2)} \cdot 1 = \frac{\alpha_j(\alpha_j+1)}{\alpha_0(\alpha_0+1)}$$

Hence, the variance can be obtained as:

$$var[\mu_j] = E[\mu_j^2] - E[\mu_j]^2 = \frac{\alpha_j(\alpha_j+1)}{\alpha_0(\alpha_0+1)} - \left(\frac{\alpha_j}{\alpha_0}\right)^2 = \frac{\alpha_j(\alpha_0-\alpha_j)}{\alpha_0^2(\alpha_0+1)}$$

And, for getting the covariance, using the formula and j ≠ l:

$$cov[\mu_j\mu_l] = \int (\mu_j - E[\mu_j])(\mu_l - E[\mu_l])Dir(\mu|\alpha)d\mu$$

$$= \int (\mu_j\mu_l - E[\mu_j]\mu_l - E[\mu_l]\mu_j + E[\mu_j]E[\mu_l])Dir(\mu|\alpha)d\mu$$

$$= \int \mu_j\mu_l Dir(\mu|\alpha)d\mu - \int E[\mu_j]\mu_l Dir(\mu|\alpha)d\mu$$

$$- \int E[\mu_l]\mu_j Dir(\mu|\alpha)d\mu + \int E[\mu_j]E[\mu_l]Dir(\mu|\alpha)d\mu$$

$$= \frac{\Gamma(\alpha_0)\Gamma(\alpha_l+1)\Gamma(\alpha_j+1)}{\Gamma(\alpha_j)\Gamma(\alpha_l)\Gamma(\alpha_0+2)} - E[\mu_j]E[\mu_l] - E[\mu_j]E[\mu_l] + E[\mu_j]E[\mu_l]$$

$$= \frac{\alpha_j\alpha_l}{\alpha_0^2(\alpha_0+1)} - \frac{\alpha_j\alpha_l}{\alpha_0^2} = -\frac{\alpha_j\alpha_l}{\alpha_0^2(\alpha_0+1)}$$

(c) Bishop 2.12

2.12 ($\star$) The uniform distribution for a continuous variable x is defined by

$$\mathrm{U}(x|a,b) = \frac{1}{b-a}, \qquad a \leqslant x \leqslant b. \tag{2.278}$$

Verify that this distribution is normalized, and find expressions for its mean and variance.

Solution: Since the uniform distribution is:

$$U(x|a,b) = \frac{1}{b-a}, \quad a \leq x \leq b$$

Hence, it is certain that:

$$\int_a^b \frac{1}{b-a} dx = 1$$

It means the distribution is normalized.

And, to calculate its mean:

$$E[x] = \int_a^b x\frac{1}{b-a} dx = \frac{x^2}{2(b-a)} |_a^b = \frac{a+b}{2}$$

Then, it can calculate that:

$$E[x^2] = \int_a^b x^2 \frac{1}{b-a} \mathrm{dx} = \frac{x^3}{3(b-a)} |_a^b = \frac{(b-a)^2}{3}$$

Thus, it can obtain its variance:

$$var[x] = E[x^2] - E[x] = \frac{(b-a)^2}{12}$$

(d) Bishop 2.15

2.15 $(\star\star)$ Show that the entropy of the multivariate Gaussian $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is given by

$$\mathrm{H}[\mathbf{x}] = \frac{1}{2} \ln |\boldsymbol{\Sigma}| + \frac{D}{2}(1 + \ln(2\pi)) \tag{2.283}$$

where D is the dimensionality of $\mathbf{x}$.

Proof: Since it is multivariate Gaussian, say:

$$\mathrm{p(x)} = \mathcal{N}(\mu|\Sigma)$$

According to the definition of entropy:

$$H[x] = -\int p(x) \ln p(x)\, dx = -\int p(x) \ln \left[\frac{1}{(2\pi)^{\frac{D}{2}}} \frac{1}{|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)} \right] dx$$

$$= -\int p(x) \ln \left[\frac{1}{(2\pi)^{\frac{D}{2}}} \frac{1}{|\Sigma|^{\frac{1}{2}}} \right] dx - \int p(x) \left[-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu) \right] dx$$

$$= -\ln \left[\frac{1}{(2\pi)^{\frac{D}{2}}} \frac{1}{|\Sigma|^{\frac{1}{2}}} \right] + \frac{1}{2} E[(x-\mu)^T \Sigma^{-1}(x-\mu)]$$

And substitute following formula into above equation:

$$\mathrm{E}[(x-a)^T A(x-a)] = \mathrm{tr}(\mathrm{A}\Sigma) + (\mu - a)^T A(\mu - a)$$

Thus, it can derive that:

$$H[x] = -\ln \left[\frac{1}{(2\pi)^{\frac{D}{2}}} \frac{1}{|\Sigma|^{\frac{1}{2}}} \right] + \frac{1}{2} E[(x-\mu)^T \Sigma^{-1}(x-\mu)]$$

$$= -\ln \left[\frac{1}{(2\pi)^{\frac{D}{2}}} \frac{1}{|\Sigma|^{\frac{1}{2}}} \right] + \frac{1}{2} tr(I_D) = \frac{1}{2} \ln|\Sigma| + \frac{D}{2} + \frac{D}{2} \ln(2\pi)$$

(Lecture one) Solve for Beta distribution's covariance

Solution: This problem is almost the same as Bishop 2.10

Since it only has two αs, it just substitutes:

$$\alpha_1 = a, \qquad \alpha_2 = b, \qquad \alpha_0 = \alpha_1 + \alpha_2 = a + b$$

Thus, the covariance of Beta distribution is:

$$cov[ab] = -\frac{ab}{(a+b)^2(a+b+1)}$$

5. **Programming assignment - can you beat the computer in rock-paper-scissors? - a case of counter - Artificial Intelligence**

The strategies of rock(R)-paper(P)-scissor(S) games only have three options: the one beat player's result, the one beat the AI's last result and the one lose player's last result. Although it seems the strategies are chosen randomly, actually it does have some potential rules when choosing the strategy.

Thus I could write some "trap" strategies to confuse the AI. But the AI will find potential repeating strategy. Then, I will generate a random gesture between two cycles which may prevent the program to study it.

After lots of attempts and some calculation, the "trap" strategy is decide to be "SRRSRSRRSRSSRSR", this always works when the first gesture of player is rock for first round. When the number of game is not so large, say 20-30, this strategy will have relatively high possibility to win the game. Even though the number of games is large, say over 100, it still has some possibilities to win and at least it won't be beaten so hard. This means it is improved compared to previous results.

According to input the results manually to the website, if the number of games is 20 or 30, the win rate is 70% (7/10); And if the number of games is 100, it still has win rate as 40% (2/5). This result is reasonable.

The others you can see the programming assignment part as "CS542HW1.py"