Part 1 : Voting

Question 1

In the American presidential elections, while the popular vote is used up to the state level, the electoral college decides the winner at the **national level**. Assuming there are only two candidates, is this system strategy-proof? Does it elect a Condorcet winner? Justify your answers. (Note: You can assume for simplicity that each state gets a single "vote" in a national election, and the state then runs an election by popular vote with however many people are in that state to determine which vote to cast at the national level. If you give a counter-example, you may choose freely the number of states and the number of voters in each state.)

Answer

This system is strategy-proofed

.

Proof

Let's assume by **contradiction** that the system is **not strategy-proofed**.

Let Social-choice context be $\Gamma = (n, \Omega, T, v)$:

- n: number of voters
- Ω: set of candidates: A, B
- T_i: ranking of candidates $\mu = x_1, \dots, x_{|\Omega|}$
  - T_i indicates the winner in state i, denoted by $(A, B)_i$ if A won in state i or $(B, A)_i$ if B won in state i **by majority of votes in the state**.
- v_i: valuation based on ranking $v_i(\mu, x) \geq v_i(\mu, x')$ if $x >_\mu x'$
  - Strict: $v_i(\mu, x) > v_i(\mu, x')$ if $x >_\mu x'$.

if Γ is not strategy proofed than it is not DST , therefore becuase of

Definition 10.1: A mechanism M is *dominant-strategy truthful (DST)* for the context $\Gamma = (n, \Omega, T, v)$ if, for every $\tilde{t} \in T$, t_i is a dominant strategy for player i in $G_{\Gamma,\tilde{t},M}$.

Then : there is a $\tilde{t} \in T$, t_i is **NOT** a dominant strategy for player i in $G_{\Gamma,\tilde{t},M}$.

However, $\tilde{t} \in T$, t_i is the strategy of **who won by majority of votes in the state when there are 2 candidates.** and this is a
DST and therefor strategy proof according to what we learned in class. (**Lecture 7 : slides 49, 53**) $\rightarrow$
CONTRADICTION , therefore This system is strategy-proofed.

This system does'nt elect always Condorcet winner.

Definition 11.2 of Condorcet Winner from the book

Given a set of voter preferences $\mu_1, \ldots, \mu_n$ over a set of candidates $X = \{x_1, \ldots, x_m\}$, we say that a candidate $x \in X$ is a Condorcet winner if for every other candidate $x' \in X$, at least $\frac{n}{2}$ voters prefer x to x' (according to the preferences $\tilde{\mu}$).

Example Scenario

Consider a simplified model of the U.S. Electoral College with three states, each awarding one electoral vote based on a majority vote within the state:

- **State A:** 60 voters
  - Candidate X: 35 votes
  - Candidate Y: 25 votes
- **State B:** 30 voters
  - Candidate X: 10 votes
  - Candidate Y: 20 votes
- **State C:** 10 voters
  - Candidate X: 1 vote
  - Candidate Y: 9 votes

Results

- **State A:** Elects Candidate X
- **State B:** Elects Candidate Y
- **State C:** Elects Candidate Y

Electoral College Votes

- Candidate X: 1 vote (from State A)
- Candidate Y: 2 votes (from States B and C)

National Popular Vote Total

NOTE: it is a group group of n voters , ($n > \frac{n}{2}$).

- Candidate X: 46 votes (35 from A, 10 from B, 1 from C)
- Candidate Y: 54 votes (25 from A, 20 from B, 9 from C)

Therefore Candidate Y would win a national popular vote by a majority, indicating that Y is the Condorcet winner since more voters prefer Y to X in a head-to-head comparison. However, Candidate Y does not win the Electoral College, demonstrating that this system does not necessarily elect the **Condorcet winner.**
In conclusion , we found a group (the whole group of voters) which is **bigger than** $\frac{n}{2}$ such that the Condorcet feature doesn't exists.
Therefore , this system does'nt elect always Condorcet winner.

Question 2

Suppose we have an election between n candidates $c_1, \ldots, c_n$, where each candidate is assigned a ranking (x_i, y_i) representing its stance on two political issues. Each voter has an ideal point (x, y), and it ranks candidates based on the distance $\sqrt{(x - x_i)^2 + (y - y_i)^2}$, ranking candidates with a smaller distance higher. Is it always possible to pick a Condorcet winner? Prove or give a counterexample.

Answer

Counter example

Suppose all the points are on a circle in $\mathbb{R}^2$ and their ideal point (x, y) is the center of the circle. Therefore, all the candidates are ranked the same (it will be R), and thus there is no strict Condorcet winner because no one prefers any candidate over the other (also for every group over $\frac{n}{2}$) (all distances are equal).

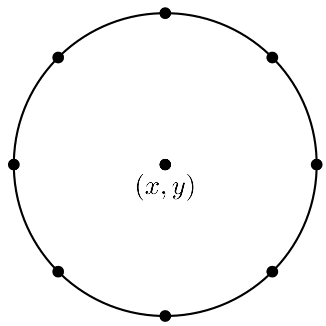


Figure 1: Circle with candidates on the circumference and the ideal point at the center.

Question 3

(a)

Compute a value of m so that the result of the poll is incorrect with probability at most 1%? Use the Hoeffding/Chernoff bounds and show your work.

Answer

We use Theorem 16.2 to determine the required number of samples, n, to achieve the desired confidence in the polling result.
Theorem 16.2 states: Let $W \in \{0, 1\}$, and let $X_1, \ldots, X_n \in \{0, 1\}$ be independent random variables such that $\Pr[X_i = W] \geq \frac{1}{2} + \varepsilon$. Then:

$$\Pr[\text{Majority}(X_1, \ldots, X_n) = W] \geq 1 - 2e^{-2\varepsilon^2 n}$$

Given that we want the probability that the poll result is incorrect to be at most 1%, i.e.,

$$\Pr[\text{Majority}(X_1, \ldots, X_n) \neq W] \leq 0.01,$$

by Theorem 16.2, we require:

$$1 - 2e^{-2\varepsilon^2 n} \geq 0.99.$$

Rearranging, we find:

$$2e^{-2\varepsilon^2 n} \leq 0.01.$$

Taking natural logarithms on both sides, we obtain:

$$\ln(2) + \ln(e^{-2\varepsilon^2 n}) \leq \ln(0.01),$$
$$\ln(2) - 2\varepsilon^2 n \leq \ln(0.01).$$

Solving for n, we get:

$$-2\varepsilon^2 n \leq \ln(0.01) - \ln(2),$$

$$n \geq \frac{\ln(0.01) - \ln(2)}{-2\varepsilon^2}.$$

Hence, the required sample size n can be calculated by substituting a specific value for ε. For example, assuming $\varepsilon = 0.05$, we have:

$$n \geq \frac{\ln(0.01) - \ln(2)}{-2 \times 0.05^2}.$$

The precise calculation yields $n \approx 1059.663$. Therefore, to ensure that the poll's result is incorrect with a probability of at most 1%, the number of samples m **needs to be greater than** 1060**.**

(b)

Let n be the number of people in the population, ε be defined such that $\left(\frac{1}{2} + \varepsilon\right) \cdot n$ prefer A to B, and let δ be the desired accuracy (so the probability the result is incorrect is at most δ). Write your bound m as a function of n, ε, and δ.

- If the number of people in the population increased by a factor of 10, how would that affect m?
- If ε decrease by a factor of 2, how would that affect m?
- If we want to increase our confidence by a factor of 10 ($\delta' = \delta/10$), how would that change m?
- If $\varepsilon = 1/n$ (so 1 person would be the deciding vote), what would this imply about m given your bound from above?

Answer

Derivation of Minimum Sample Size m

Given:

- n is the total number of people in the population.
- ε such that $\left(\frac{1}{2} + \varepsilon\right) \cdot n$ people prefer A over B.
- δ is the desired accuracy, such that the probability that the result is incorrect is at most δ.

Applying to Polling

In polling, the X_i are Bernoulli trials where $X_i = 1$ if the i-th respondent prefers A over B, and $\mu = \frac{1}{2} + \varepsilon$ represents the proportion of the population that prefers A over B. The inequality becomes from **Thoerem 12.6**:

$$\Pr[\text{Majority incorrectly predicted}] \leq 2\exp(-2m\varepsilon^2)$$

To meet the requirement that the probability of an incorrect prediction is at most δ, we set:

$$2\exp(-2m\varepsilon^2) \leq \delta$$

Solving for m, we get:

$$\exp(-2m\varepsilon^2) \leq \frac{\delta}{2},$$

$$-2m\varepsilon^2 \leq \ln\left(\frac{\delta}{2}\right),$$

$$m \geq \frac{\ln\left(\frac{2}{\delta}\right)}{2\varepsilon^2}.$$

Conclusion

Thus, the minimum sample size m necessary to ensure that the polling result is incorrect with a probability of at most δ is given by:

$$m \geq \frac{\ln\left(\frac{2}{\delta}\right)}{2\varepsilon^2}.$$

Effects of Changes in Parameters on the Minimum Required Sample Size m

Given the formula for the minimum required sample size:

$$m \geq \frac{\ln\left(\frac{2}{\delta}\right)}{2\varepsilon^2}$$

Effect of Increasing the Population Size by a Factor of 10

The formula for m does not directly depend on the total population size n. Thus, increasing n by any factor does not affect m, as m is solely a function of ε and δ.

Effect of Decreasing ε by a Factor of 2

Decreasing ε impacts m significantly. If ε is halved ($\varepsilon' = \varepsilon/2$):

$$m' \geq \frac{\ln\left(\frac{2}{\delta}\right)}{2(\varepsilon/2)^2} = 4 \times \frac{\ln\left(\frac{2}{\delta}\right)}{2\varepsilon^2}$$

This implies that m increases by a factor of 4, illustrating the inverse square relationship between ε and m.

Effect of Increasing Confidence by a Factor of 10 ($\delta' = \delta/10$)

To achieve a tenfold increase in confidence ($\delta' = \delta/10$), we modify m:

$$m' \geq \frac{\ln\left(\frac{2}{\delta/10}\right)}{2\varepsilon^2} = \frac{\ln(20/\delta)}{2\varepsilon^2}$$

Considering $\ln(20/\delta) = \ln(2/\delta) + \ln(10)$, the required m increases due to the added logarithmic term $\ln(10) \approx 2.302$.

Setting $\varepsilon = 1/n$

When $\varepsilon = 1/n$, implying one person can swing the preference:

$$m \geq \frac{\ln\left(\frac{2}{\delta}\right)}{2(1/n)^2} = \frac{n^2 \ln\left(\frac{2}{\delta}\right)}{2}$$

Here, m increases quadratically with n, suggesting that for large populations, the required sample size becomes impractically large, reflecting the sensitivity of m to small changes in ε.

(c)

In practice, what might be wrong with the above assumptions (i.e. why might we not we use polls to run our elections)?

Answer

Uniform Independent Distribution: The formula assumes that the preferences of the voters (represented by the random variables) are independent and **identically distributed.** In reality, voters' decisions can be correlated due to **shared information, social influences, demographic factors** such like Locations, neighborhoods and cities that correspond to culutre and political views, and other regional variables.

All of these factors can make the Distribution not a **Uniform Independent Distribution**.

Part 2: Stable Matchings

Question 4

For the following setting, find the female-optimal stable matching. Describe in detail how the Gale-Shapley algorithm arrives at the matching you find.

Females Preferences

$$A : X > W > Y > Z$$
$$B : X > W > Y > Z$$
$$C : X > W > Z > Y$$
$$D : Y > W > Z > X$$

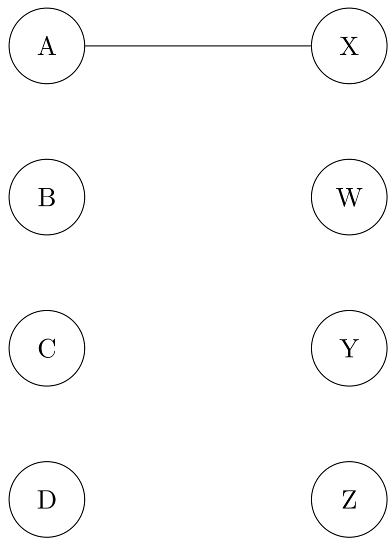
Males Preferences

$$W : D > B > C > A$$
$$X : D > B > A > C$$
$$Y : C > B > D > A$$
$$Z : D > B > C > A$$

Answer

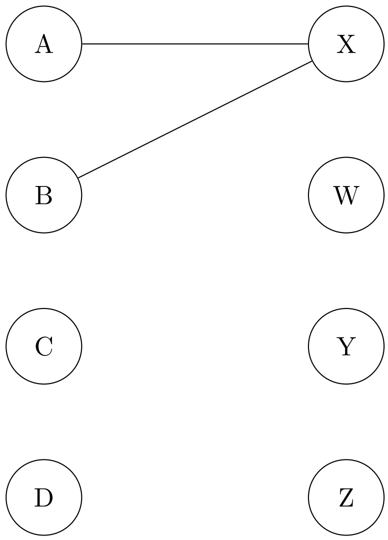
Let's describe step by step how the Gale-Shapley algorithm arrives at the matching we found

Step 1: A proposes to X



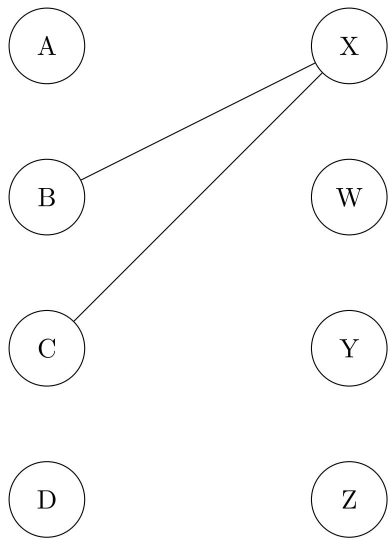
X accepts A

Step 2: B proposes to X



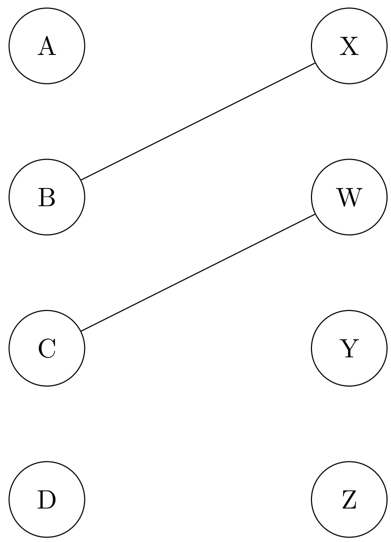
X is currently engaged to A
X prefers B over A
X is now engaged to B

Step 3: C proposes to X



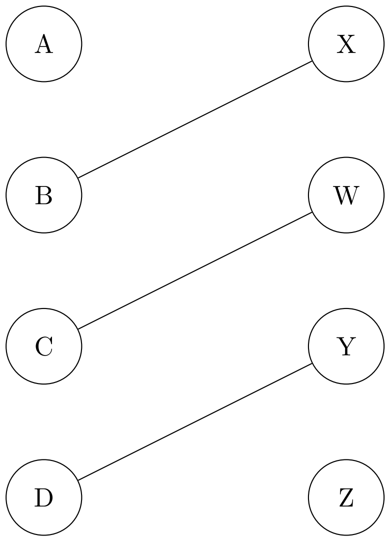
X is currently engaged to B
X decides to stay with B

Step 4: C proposes to W



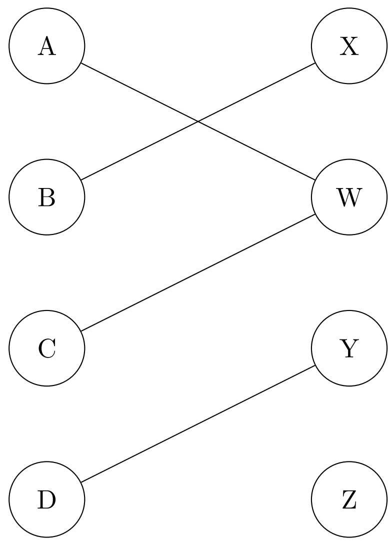
W accepts C

Step 5: D proposes to Y



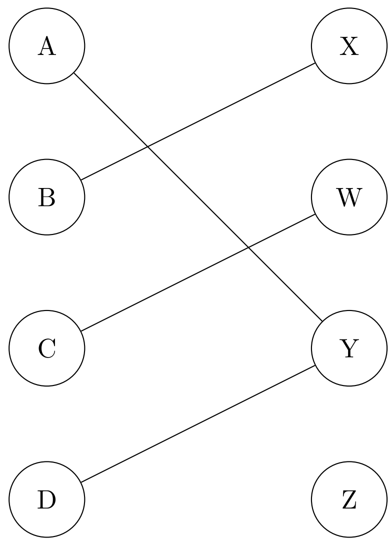
Y accepts D

Step 6: A proposes to W



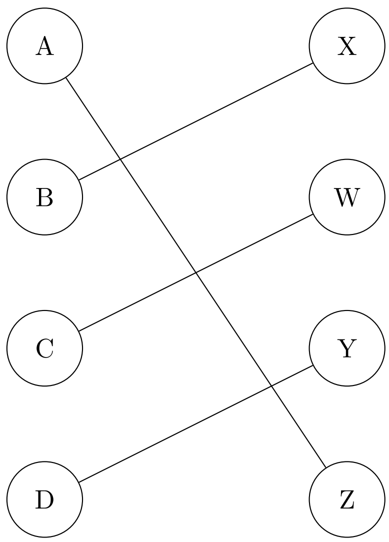
W is currently engaged to C
W decides to stay with C

Step 7: A proposes to Y



Y is currently engaged to D
Y decides to stay with D

Step 8: A proposes to Z



Z accepts A

We have reached to a stable match!

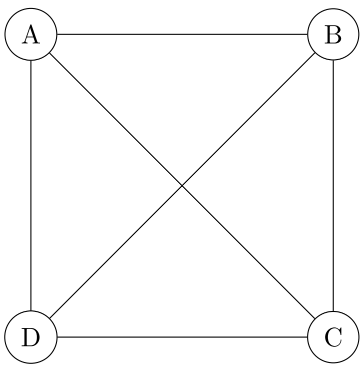
Question 5

Prove that there exists a non-bipartite matching setting (where every individual has preferences over all other individuals) for which no stable matching exists

Answer

Let's prove it:

Graph



Preferences Table

Vertex	Preferences
A	$B > C > D$
B	$C > A > D$
C	$A > B > D$
D	$A > B > C$

First, of course this is not a bipartite graph because there is an odd cycle (for example: **ABC**) .

Second, reminder:

Definition 14.2. An outcome M of a two-sided matching problem $\Gamma = (X, Y, v)$ is ***stable*** if there does not exist $x \in X$, $y \in Y$ such that $v_x(y) > v_x(M(x))$ and $v_y(x) > v_y(M(y))$.

A, B, and C is the most preferable for B, C and A respectively. And D is the least preferable by all of them.
Since we want to have a Match, someone needs to be matched with D.

However, whoever is matched with D, there always exists someone who makes the match M not **stable**.

Case 1: If A is mapped to D , (then C mapped to B)

$v_A(C) > v_A(M(A)) = v_A(D)$ **and**
$v_C(A) > v_C(M(C)) = v_C(B)$.
M is not stable.

Case 2: If B is mapped to D , (then A mapped to C)

$v_B(A) > v_B(M(B)) = v_B(D)$ **and**
$v_A(B) > v_A(M(A)) = v_A(C)$.
M is not stable.

Case 3: If C is mapped to D , (then A mapped to B)

$v_C(B) > v_C(M(C)) = v_B(D)$ **and**
$v_B(C) > v_B(M(C)) = v_B(D)$.
M is not stable.

Eventually we can see, that no matter what, there is no **Stable Match** as desired.

Part 3: Wisdom of Crowds

Question 6

In the setting described in Recitation 9 Slide 8, what happens if players choose sequentially, where Player 2 known Player 1's choice, and every player i where i > 2 knows the choices of players i1 and i2? Are we in the "Wisdom" case where the majority is correct with high probability, or in the "Foolishness" case where there is substantial probability all choices are wrong? Assume players always choose their own signal when all else is equal. Explain your answer.

Answer

In this case , we are in **"Foolishness of crowds case"**.
We will explain.
Player 1: Chooses its signal $g_1 = X_1$.
Player 2: Chooses its signal $g_2 = X_2$.
What about Player 3 ? Player 3 will see what Player 2 and Player 1 chose.

> If $g_1 = g_2 = x$, Then because of rationality of the players, Player 3 will see choose x no matter what.

Also, for every player $i > 2$, If $g_1 = g_2 = x$ then Player i will choose also x.

It is proved by induction.

Base case i = 3 : we already proved.

Assumption for i = n, and proof for n+1 by the assumption of the induction $g_1 = g_2 = g_3 =, , , , , g_n = x$ Player n+1 sees only g_n and g_{n-1}, because of rationality of the players, Player n+1 will choose $g_1 = g_2 = g_{n-1} = g_n = x$ as desired. Therefore $P[All \neq W] = P[X_1, X_2 \neq W] = (\frac{1}{2} - \varepsilon)^2$.
Therefore, we are at **Foolishness of crowds case**