CS5340:
Tutorial 1

Asst. Prof. Harold Soh

TA: Eugene Lim

1. Two Numbers Game

- Team 1:
  - Pick 2 **different** numbers between 0 and 10 (inclusive).
  - Write each number on a piece of paper each.
  - Turn the papers face down.
- Team 2:
  - Objective is to pick the **larger number**.
  - Pick one of the pieces of paper.
  - Have a peek at the number.
  - **Decide:** do you keep this number or *switch*?
- **Question:** Can Team 2 win more than 50% of the time?

Yes!

Team 2 (Cheating Strategy)

- Team 2 has a **spy** in Team 1
- Team 1 picks two numbers L and H
- Spy tells you a number between L and H
  - If L and H are next to one another, then spy tells you L
- Example:
  - Team 1 picks 3 and 8
  - Spy says 5
  - Team 2 picks randomly.
  - Team 2: Do you switch?

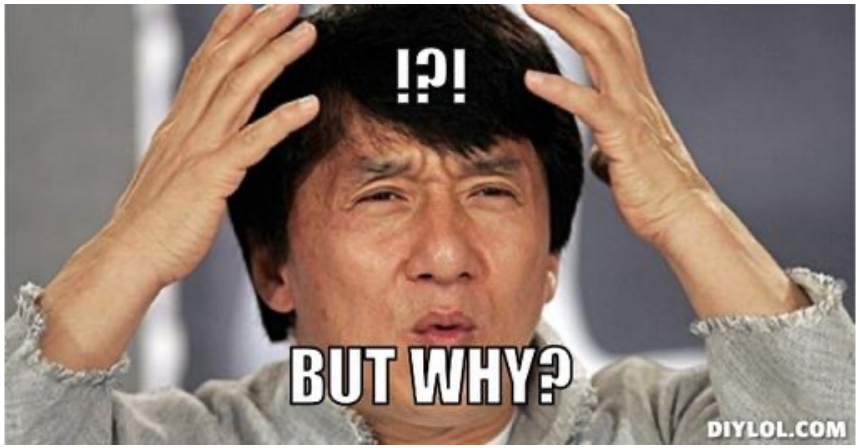
- Team 2 has a 100% of winning!

From Just now:
Academic Honesty

- Please be **academically honest**.
  - *"Give credit where credit is due"*
- Do ***NOT*** cheat
- **Strict** Plagiarism policy:
  - If you cheat, we will report you to the disciplinary board
  - If found guilty, you will get an F (University Policy)

Team 2 Strategy

- Randomly pick $z \in [0, 10)$
- Take a peek at one of the numbers, lets call this x
- If $x \leq z$, switch, else stick with x

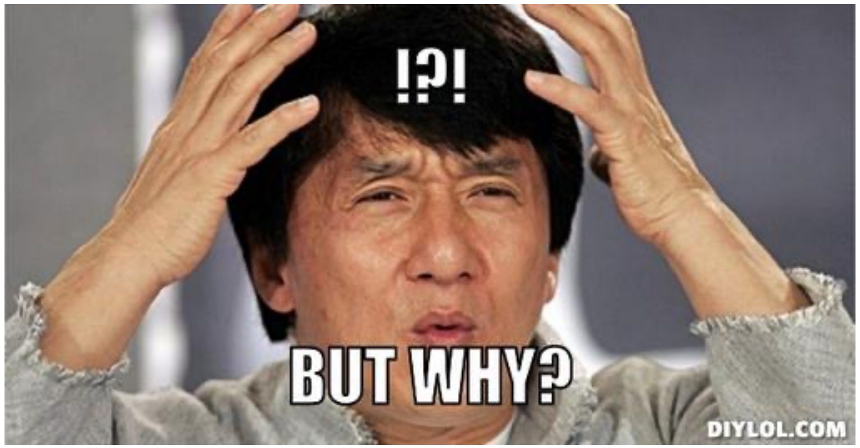


In Groups

- Introduce yourselves.
- Split into Team 1 and Team 2 and play this game.
- Discuss and analyze why the randomized strategy works.
  - Use the techniques learnt in the lecture (tree diagram)
- What is the probability Team 2 wins using this strategy?
- If you are done, move on to the other problems in the tutorial.

Team 2 Strategy

- Randomly pick $z \in [0, 10)$
- Take a peek at one of the numbers, lets call this x
- If $x \leq z$, switch, else stick with x



Strategy Analysis

- Let $L \ < H$ be the numbers chosen by Team 1.
- 3 possible cases:
  - **Case** just-right: $L \leq Z < H$:
    - Team 2 wins always!
    - $p(win|\text{justright}) = 1$ and $p(M) \geq \frac{1}{10}$
  - **Case** too-high: $H \ \leq Z$:
    - Team 2 switch. Only wins if picked L
    - $p(win|\text{toohigh}) = \frac{1}{2}$
  - **Case** too-low: $Z \ < L$:
    - Team 2 stays. Only wins if picked H
    - $p(win|\text{toolow}) = \frac{1}{2}$

$$p(\text{win}) \geq \left(1\times\frac{1}{10}\right) + \frac{1}{2}\left(1 - \frac{1}{10}\right)$$
$$p(\text{win}) \geq \frac{11}{20}$$

More complete derivation in: MIT Math for CS, Chapter 18.3.3
https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-042j-mathematics-for-computer-science-spring-2015/readings/MIT6_042JS15_Session31.pdf

Strategy Analysis (general case)

- Team 1 chooses 2 numbers $[0, \dots, n]$
- Team 2 picks half-integers:

$$\frac{1}{2}, \frac{3}{2}, \dots, \frac{2n-1}{2}$$

- Analysis is similar to before
- 3 cases:
  - $p(z$ just right): $(H - L)/n$
  - $p(z$ too low): L/n
  - $p(z$ too high): $(n - H)/n$

Source: https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-042j-mathematics-for-computer-science-spring-2015/readings/MIT6_042JS15_Session31.pdf

Strategy Analysis (general case)

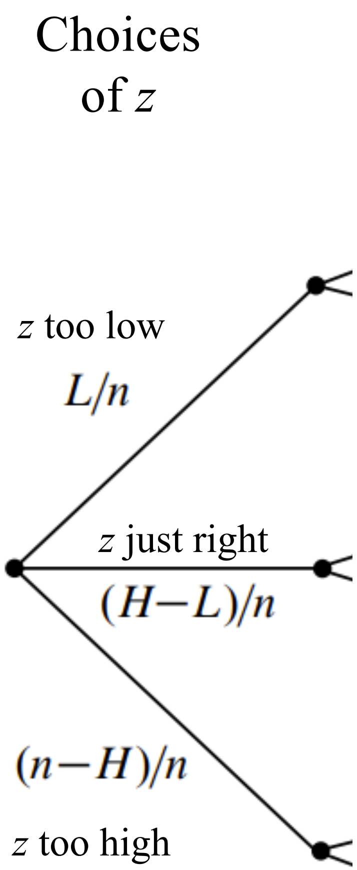


Figure 18.3 The tree diagram for the numbers game.

Strategy Analysis (general case)

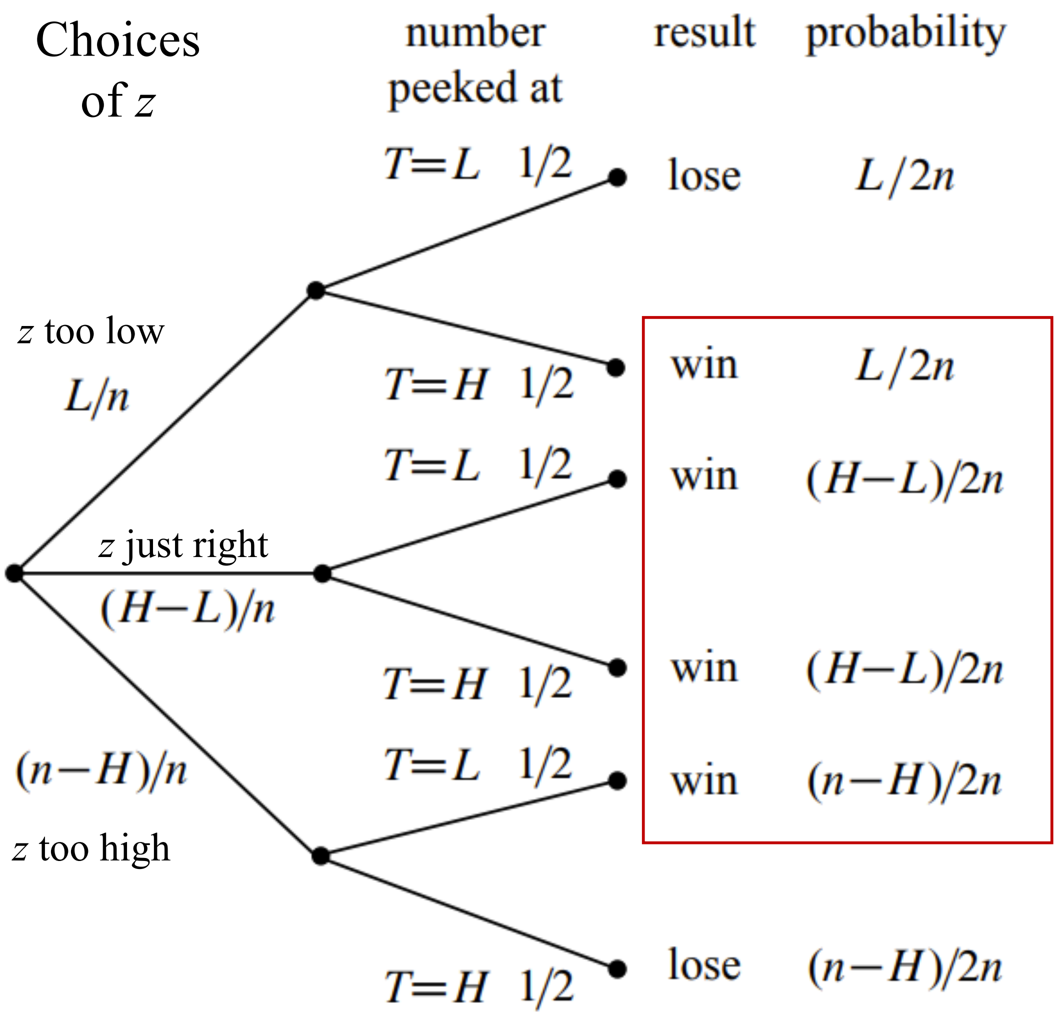


Figure 18.3 The tree diagram for the numbers game.

Probablity of win:

$$
\begin{aligned}
p(\text{win}) &= \frac{L}{2n} + 2\frac{H-L}{2n} + \frac{n-H}{2n} \\
&= \frac{1}{2} + \frac{H-L}{2n} \\
&\geq \frac{1}{2} + \frac{1}{2n}
\end{aligned}
$$

Source: https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-042j-mathematics-for-computer-science-spring-2015/readings/MIT6_042JS15_Session31.pdf

https://www.quantamagazine.org/solution-information-from-randomness-20150722/

More rigorous: https://arxiv.org/pdf/1608.01899.pdf

Questions?

https://pollev.com/haroldsohsoo986

SCAN ME

2. Legal Reasoning

Setup:

- Blood found at a scene.
- Blood type present is type S.
- Type S blood is found in 1% of the population.

Prosecutor: “*there is a 1% chance the defendant would have blood type S if he was innocent. Thus, there is a 99% chance he is guilty!*”

Is the prosecutor correct? Justify your answer.

What is your decision?

- You are the Jury. Do you vote to convict the Defendant?
- Put in your vote onto polleverywhere

SCAN ME

2. Legal Reasoning

The argument is **wrong**. Let:

- A = "*person has blood type S*"
- B = "*person is innocent*"

Prosecutor has quoted $p(A|B)$ (or $1 - p(A|B)$).
However, what is relevant is $p(B|A)$

$$p(\text{innocent}|S) = \frac{p(S|\text{innocent})\, p(\text{innocent})}{p(S)}$$

This is known as the **prosecutor's fallacy**

Questions?

https://pollev.com/haroldsohsoo986

SCAN ME

3. Conjugate Analysis

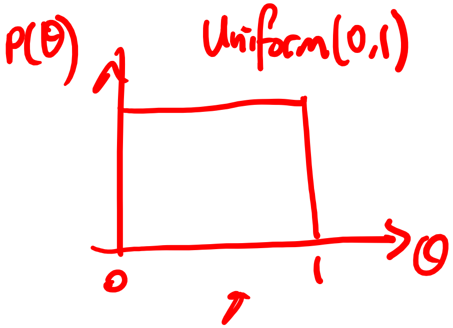
Show that the Beta Distribution is conjugate to the Binomial Distribution.

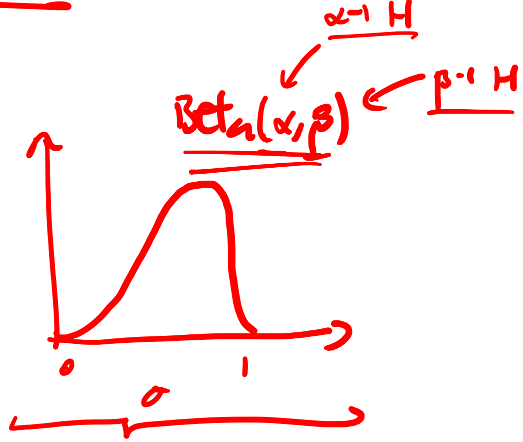
α', β'

$$\underbrace{p(\theta|x)}_{\text{posterior}} = \frac{\overset{\text{Prior}}{p(\theta)}\;\overset{\text{Likelihood.}}{p(x|\theta)}}{\underbrace{p(x)}_{\text{Marginal}}}$$

$|\alpha,\beta$

$$p(x) = \int p(\theta)\,p(x|\theta)\,d\theta$$





$$p(\pi) = \frac{1}{B(\alpha,\beta)}\,\pi^{\alpha-1}(1-\pi)^{\beta-1}$$

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt$$

$$\Gamma(n) = (n-1)!$$

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!} = \frac{\Gamma(n+1)}{\Gamma(x+1)\Gamma(n-x+1)}$$

$$p(x|\pi) = \binom{n}{x}\pi^{x}(1-\pi)^{n-x}$$

https://mathlets.org/mathlets/beta-distribution/ (online demo)

$$p(\pi) = \frac{1}{B(\alpha,\beta)} \pi^{\alpha-1} (1-\pi)^{\beta-1}$$ Prior.

$$p(x|\pi) = \binom{n}{x} \pi^{x} (1-\pi)^{n-x}$$ Likelihood

$$\underline{\underline{p(\pi|x)}} \propto p(x|\pi)\,p(\pi) = \binom{n}{x} \underline{\pi}^{x} (1-\pi)^{n-x} \cdot \frac{1}{B(\alpha,\beta)} \underline{\pi}^{\alpha-1} (1-\pi)^{\beta-1}$$

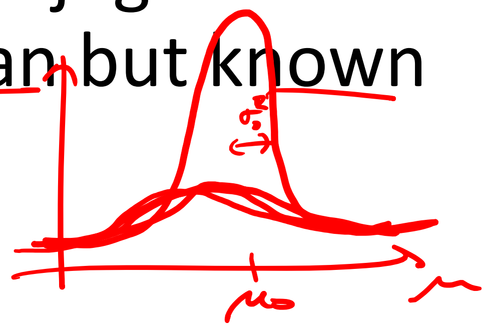
$$= \underbrace{\binom{n}{x} \frac{1}{B(\alpha,\beta)}}_{\text{const w.r.t. } \pi} \pi^{\underline{x+\alpha-1}} \underline{(1-\pi)}^{\underline{n-x+\beta}-1}$$

$$\left[\begin{aligned} \alpha' &= \underline{\underline{x}} + \underline{\alpha} \\ \underline{\beta}' &= \underline{n-x} + \underline{\beta} \end{aligned}\right]$$

3. Conjugate Analysis

$p(\mu \mid X)$ $\qquad$ $p(\mu \mid x_1)$

Show that the Normal distribution is conjugate to the Normal distribution with unknown mean but known variance.



$N(\mu \mid \mu_0, \sigma_0^2)$ $\qquad$ $N(x \mid \mu, \sigma^2)$ known, unknown

$N(\mu \mid \mu_1, \sigma_1^2)$

$$p(\mu \mid x_1) \propto p(\mu)\, p(x \mid \mu)$$

$$= \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}\right) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$= \underbrace{\frac{1}{\sqrt{2\pi\sigma_0^2}} \cdot \frac{1}{\sqrt{2\pi\sigma^2}}}_{\text{const wrt } \mu} \exp\left(-\frac{1}{2} \underbrace{\left[\frac{(\mu-\mu_0)^2}{\sigma_0^2} + \frac{(x-\mu)^2}{\sigma^2}\right]}_{\text{quad in } \mu \text{ (A)}}\right) \approx c \exp\left(-\frac{1}{2} \boxed{\frac{(\mu-\mu_1)^2}{\sigma_1^2}}\right) \text{ (B)}$$

(A) $$\frac{\mu^2 - 2\mu\mu_0 + \mu_0^2}{\sigma_0^2} + \frac{x^2 - 2x\mu + \mu^2}{\sigma^2} = \left[\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}\right]\mu^2 - \left[\frac{2\mu_0}{\sigma_0^2} + \frac{2x}{\sigma^2}\right]\mu + \left[\frac{\mu_0^2}{\sigma_0^2} + \frac{x^2}{\sigma^2}\right]$$

$$= \frac{1}{\sigma_1^2}\mu^2 - \frac{1}{\sigma_1^2} \cdot 2\mu_1\mu + \frac{1}{\sigma_1^2}\mu_1^2$$

$$\underline{\left[\frac{1}{\sigma_1^2}+\frac{1}{\sigma^2}\right]}\mu^2 - \underline{\left[\frac{2\mu_0}{\sigma_0^2}+\frac{2x}{\sigma^2}\right]}\mu + \text{const}$$

$$= \frac{1}{\sigma_1^2}\mu^2 - \underline{\frac{1}{\sigma_1^2}}2\mu_1\mu + \text{const}$$

$$\mathcal{X} = \{x_1, \dots, x_n\}.$$

$$\Downarrow$$

$$\frac{\mu_0}{\sigma_0^2}+\frac{x_1}{\sigma^2} = \frac{\mu_1}{\sigma_1^2}$$

$$\underline{\underline{\sigma_1^2 = \left[\frac{1}{\sigma_0^2}+\frac{1}{\sigma^2}\right]^{-1}}} \Rightarrow \mu_1 = \underline{\underline{\left[\frac{\mu_0}{\sigma_0^2}+\frac{x_1}{\sigma^2}\right]\left[\frac{1}{\sigma_0^2}+\frac{1}{\sigma^2}\right]^{-1}}} \quad \#$$

$$\sigma_2^2 = \left[\frac{1}{\sigma_1^2}+\frac{1}{\sigma^2}\right]^{-1} = \left[\frac{1}{\left[\frac{1}{\sigma_0^2}+\frac{1}{\sigma^2}\right]^{-1}}+\frac{1}{\sigma^2}\right]^{-1} = \left[\frac{1}{\sigma_0^2}+\frac{1}{\sigma^2}+\frac{1}{\sigma^2}\right]^{-1} = \left[\frac{1}{\sigma_0^2}+\frac{2}{\sigma^2}\right]^{-1}$$

$$\mu_2$$

4. Variance of a Sum

Show that

$$V[X+Y] = V[X] + V[Y] + 2\mathrm{Cov}[X,Y]$$ show.

where $\mathrm{Cov}[X,Y] = E[XY] - E[X]E[Y]$ ✓
and $V[X] = E[X^2] - E[X]^2$ ✓

$$V[X+Y] = E[(X+Y)^2] - E[X+Y]^2$$
$$= E[X^2+2XY+Y^2] - E[X+Y]^2$$
$$= E[X^2] + 2E[XY] + E[Y^2] - \underbrace{(E[X]+E[Y])^2}$$
$$= E[X^2] + 2E[XY] + E[Y^2] - E[X]^2 - E[Y]^2 - 2E[X]E[Y]$$
$$= V[X] + V[Y] + 2\mathrm{Cov}[X,Y].$$

Questions?

https://pollev.com/haroldsohsoo986

SCAN ME

Please Prepare for Next Week

- Complete Tutorial 1 (online).
- Watch the next batch of videos and do tutorial 2 (out tomorrow)