

問題解決技法入門

1. 統計・シミュレーションと予測

～サイコロをうまく使おう～

堀田 敬介



問題



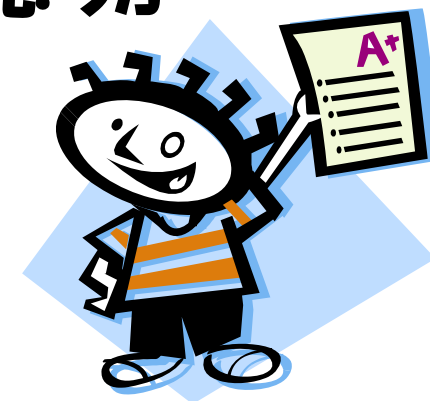
文教太郎君はマークシート形式のテストを受ける予定だ。全部で20問あり、各問題の選択肢は4つ(A,B,C,D)である。

ところが、いざ受けてみたら、まずいことにさっぱり分からない。どうしよう？ 困ったな…。 そうだ！こんな時は秘伝の「鉛筆転がし」を使おう！ 幸いにも、太郎君の鉛筆の断面は正方形だったので、各面にA, B, C, Dと書き、転がした…。

さて、60点以上で合格だが、彼は単位を貰えるだろうか？

問1. 彼が5問以上誤答する確率はどれくらいあると思うか
予測しなさい

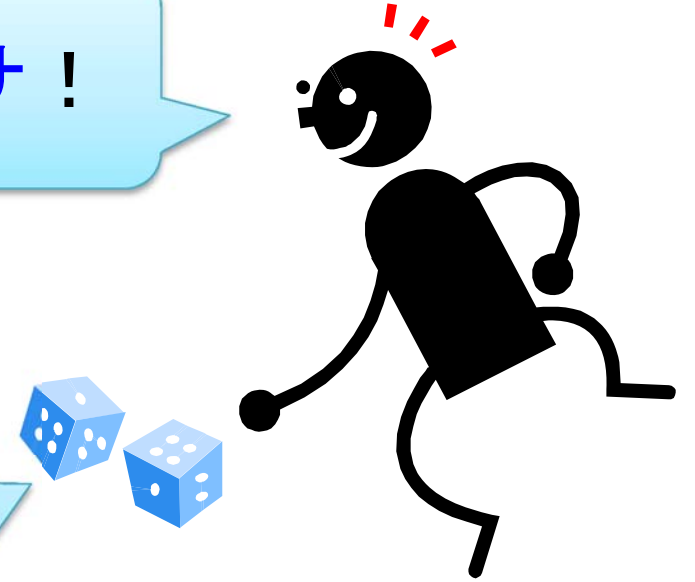
問2. 彼が単位をとれる可能性はどれくらいか
予測しなさい



わかんないやサイコロを振ろう

理屈(確率統計)はいろいろあるサ！

でも、確率は示されているのだから、
四の五の言わずに何千回何万回と
トライ(試行)してみればいいさ～



人類の偉大な発明の1つはサイコロである
サイコロによって人は神に一步近づく(エッ？)

<注意事項>

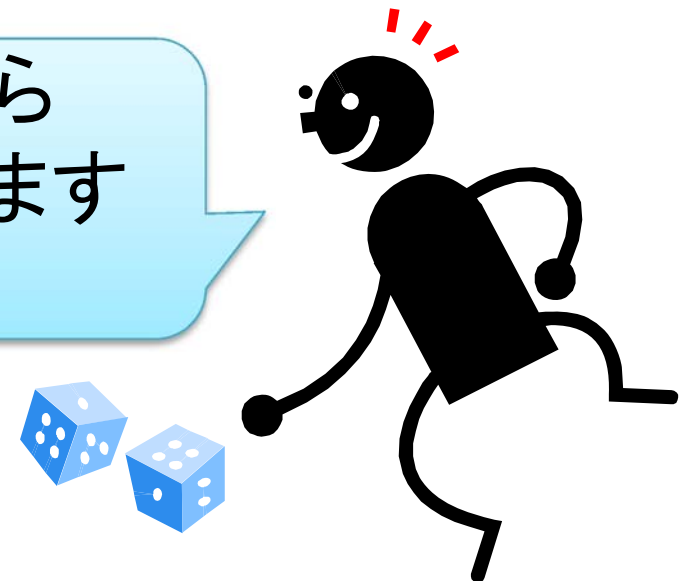
- ✓ 自分で振る場合→ 理論通りの出目 (cf. 正重心サイコロ, 重心がずれた賽の確率計算)
- ✓ コンピュータに振らせる場合→ 疑似乱数生成 (cf. 線形合同法, Mersenne twister, etc.)

※アインシュタインの言葉として有名なため、言葉が一人歩きすることが多いようです

ここでも、本来の発言の趣旨と異なる形で引用しているので注意

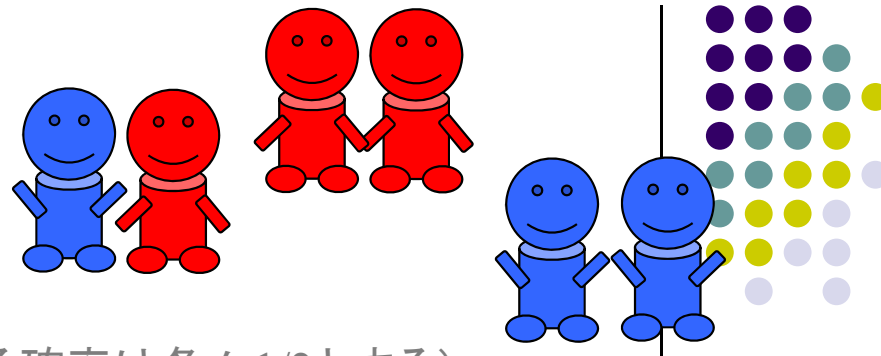


でも、人間ですから
サイコロ使っちゃいます
ゴメンね！



Coffee Break!

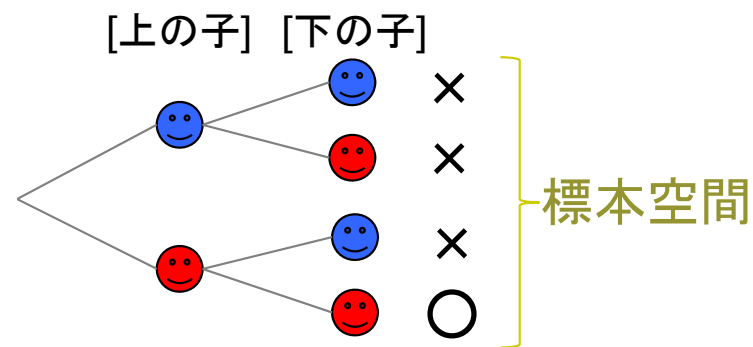
ベイズ推定初歩



© 2人の子供を持つ家庭がある (注: 男女の生まれる確率は各々1/2とする)

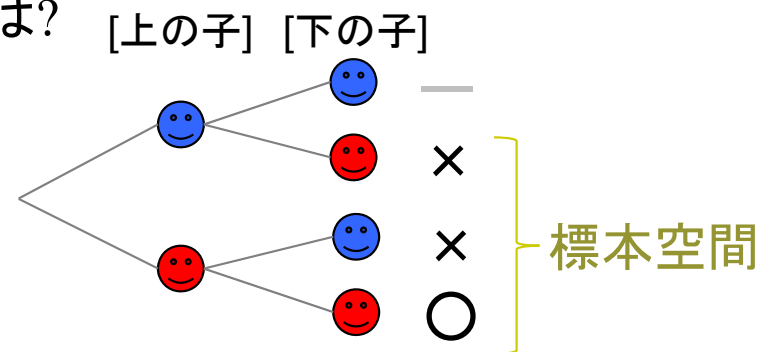
© **Q1**: 2人とも女の子である確率は?

$$\frac{1}{4}$$



© **Q2**: 1人が女の子の時, 2人とも女の子である確率は?

$$\frac{1}{3}$$



© **Q3**: 1人が女の子で名前がAliceの時, 2人とも女の子である確率は?

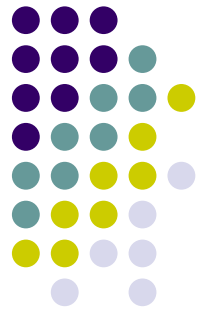
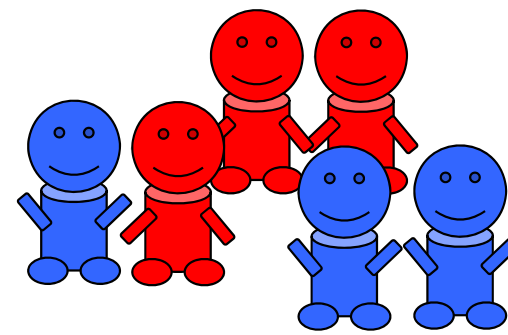
$$\frac{3}{5}$$

or

$$\frac{1}{2}$$

Coffee Break!

ベイズ推定初歩



Q3: 1人が女の子で名前がAliceの時, 2人とも女の子である確率は?

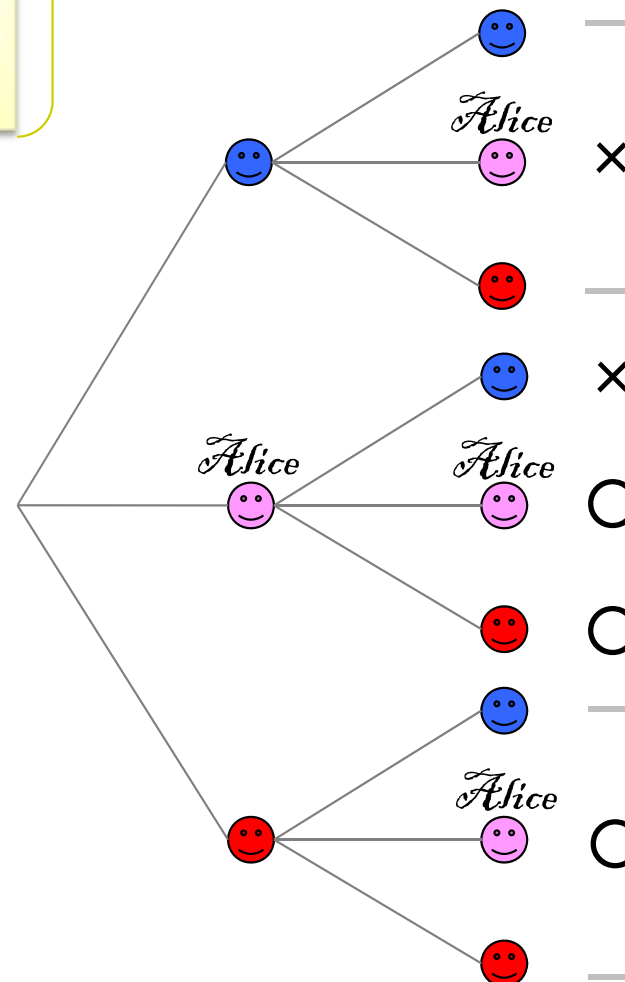
$$\frac{3}{5}$$

or

$$\frac{1}{2}$$

[上の子]

[下の子]



標本空間

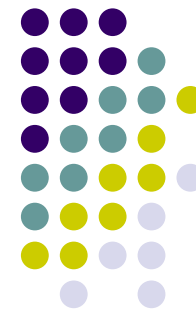
注) 女の子がAliceと名付けられる場合とそうでない場合の確率を同じとみなしている

実際にはAliceと名付けられる確率は、そうでない場合よりずっと低いだろう

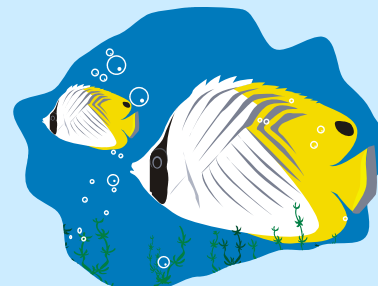
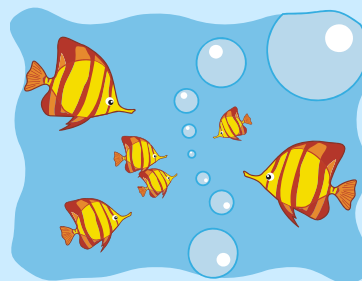
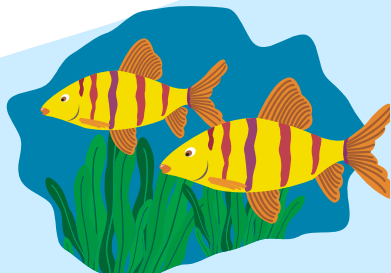
また, ここでは, どちらかの名前がAliceであると判明した(新しい情報が付加された)場合の話であることに注意

注) 2人ともにAliceという名前であることを許さないなら, この場合はない

Coffee Break!



湖の中にいる，特定の魚の**数**を**推定**したい
どうしたらよいか？



Coffee Break!



標識再捕獲法
(mark-recapture method)
ともいう

「捕獲再捕獲法 capture-recapture method」

- 湖の中の魚の個体数(N匹)を推定したい
 - Step1 (ランダムに) 魚を捕獲(M匹), 印●をつけて放す
 - Step2: しばらくおいて, (ランダムに) 魚を再捕獲(n匹)し, 印の付いている魚を数える(x匹)



Step1: 捕獲後,
印●をつけた魚(M匹)

Step2: 再捕獲(n匹)

再捕獲魚のうち
印の付いた魚(x匹)

未知の推定したい数値

個体総数 = N匹
全● M匹, 全● N-M匹

再捕獲 = n匹
● x匹, ● n-x匹

既知の観測数値

➡ 推定値: $\hat{N} = \frac{Mn}{x}$

例) $M=300, n=500, x=5$ なら

$$\hat{N} = \frac{300 \cdot 500}{5} = 30000$$

演習

あなたと友達はボーリングをしている。が、二人とも財布を忘れたことに気がついた。困っていたあなた達に店主がこう持ちかけてきた

「これからそれぞれ1ゲーム(10フレーム)行い、2人あわせて8回以上ストライクを出せたら、ゲーム代はいらないよ。駄目だったら、これから3時間働いてもらう」

あなたは平均5回に2回はストライクがだせる

友達は平均3回に1回はストライクがだせる

問 あなたと友達がゲーム代を支払わずに帰れる可能性はどれくらいあるか？ 予測しなさい



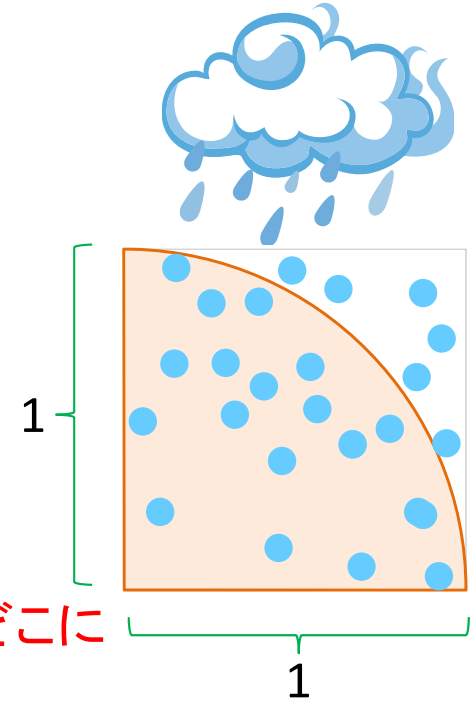
問題

問 円周率 π はいくつか？

モンテカルロ法(モンテカルロ・シミュレーション)

一辺1の正方形と内接する半径1の $\frac{1}{4}$ 円を描く(右図)

右の正方形の領域に「雨が降る」としよう. 雨は, 正方形領域のどこに降るだろう? もちろん, $\frac{1}{4}$ 円の内部か外部のどちらかだよ



そして, 全雨粒の内, 円の内部に降る雨粒の数は, 正方形の面積(1)に対する $\frac{1}{4}$ 円の面積($\pi/4$)の比に近い値になりそうだよね

だから, 全雨粒の数を T , 円内に降る雨粒数を R とすると

$$\text{正方形面積} : \frac{1}{4}\text{円面積} = 1 : \pi/4 = T : R$$

となる. 即ち, $\pi/4 = R / T \Leftrightarrow \pi = 4 \times R \div T$

この式は何を意味するのか? R と T が分かれば π を計算できるってこと!

さあ, 雨粒をたくさん降らせ($T=10,000$ 粒としよう), R の数を数えよう!

これが, モンテカルロ法による, 円周率 π の近似法(の1つ)だよ

問題

問 酔っ払いが道を歩いている
あっちにふらふら, こっちにふら
ふら, …あぶなっかしい

無事に家までたどり着けるだ
ろうか？

ランダム・ウォーク(酔歩)

彼の歩みは

1/3の確率で右にふらふら

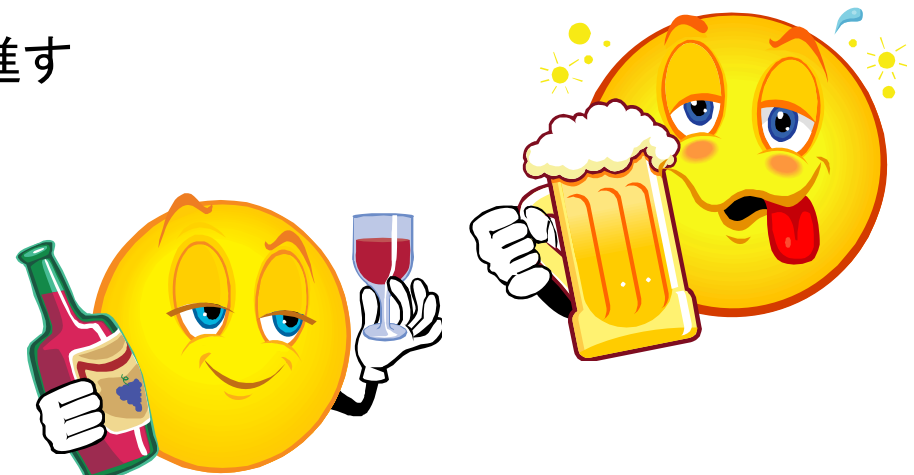
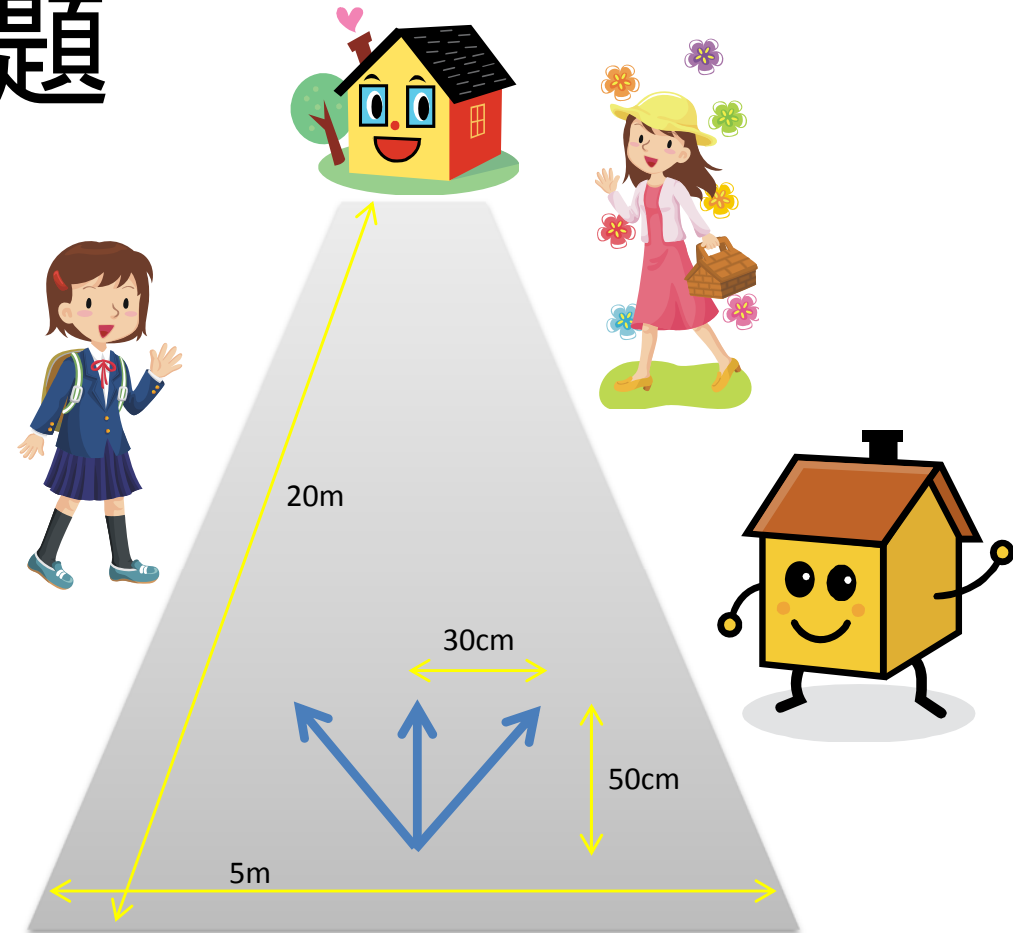
1/3の確率で左にふらふら

1/3の確率でまっすぐ

としよう. ただし, いつも1歩(50cm)は前進す
るとし, 左右へふらつく距離は30cmとする
道幅を5m

現在位置(初期位置)は道路の真ん中
とし, 家まで20mとする

この設定でシミュレーションをするよ



Coffee Break!



Monty-Hole Dilemma

確率的直感

3枚の扉の向こうに

- **百万ドル** (当たり)

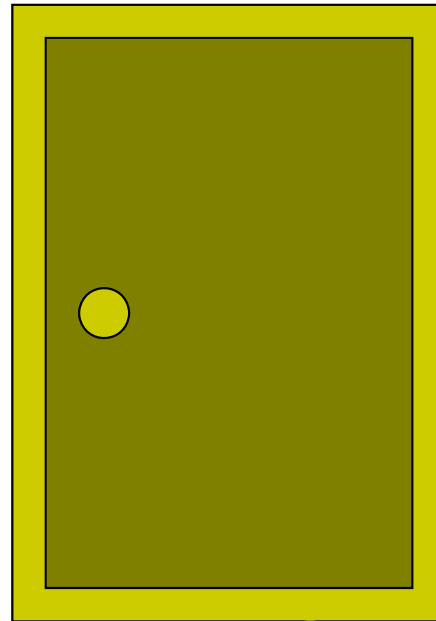
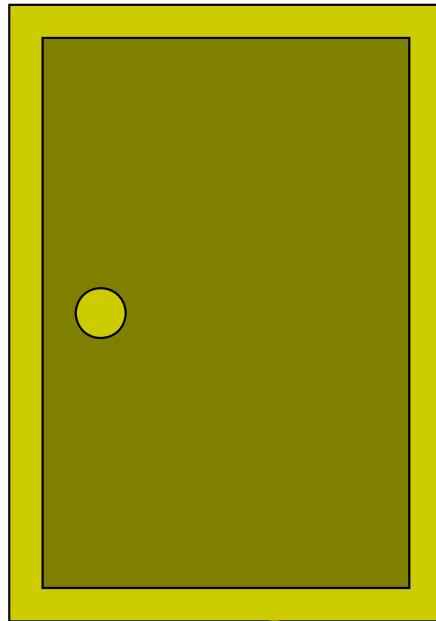
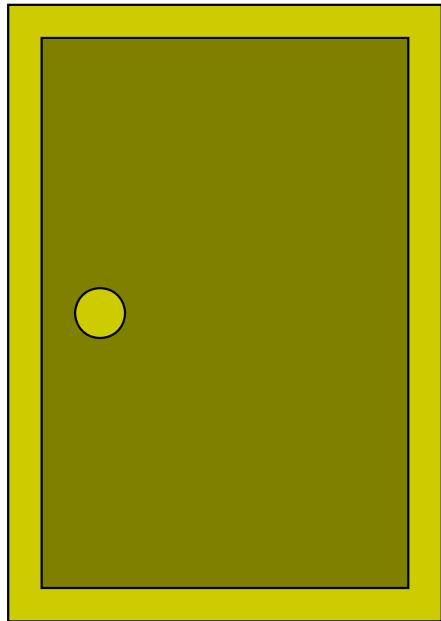
- **山羊** (はずれ)

- **山羊** (はずれ)

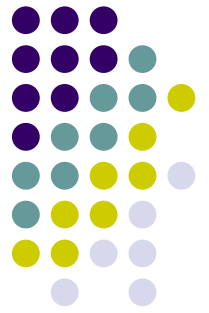
が隠されているよ。
あなたは扉を1つだけ選んでいいのよ。

ところで、あなたが選ばなかった2つの扉のうち、山羊の扉を開くから、それを見た後で、開く扉を変えてもいいよ。

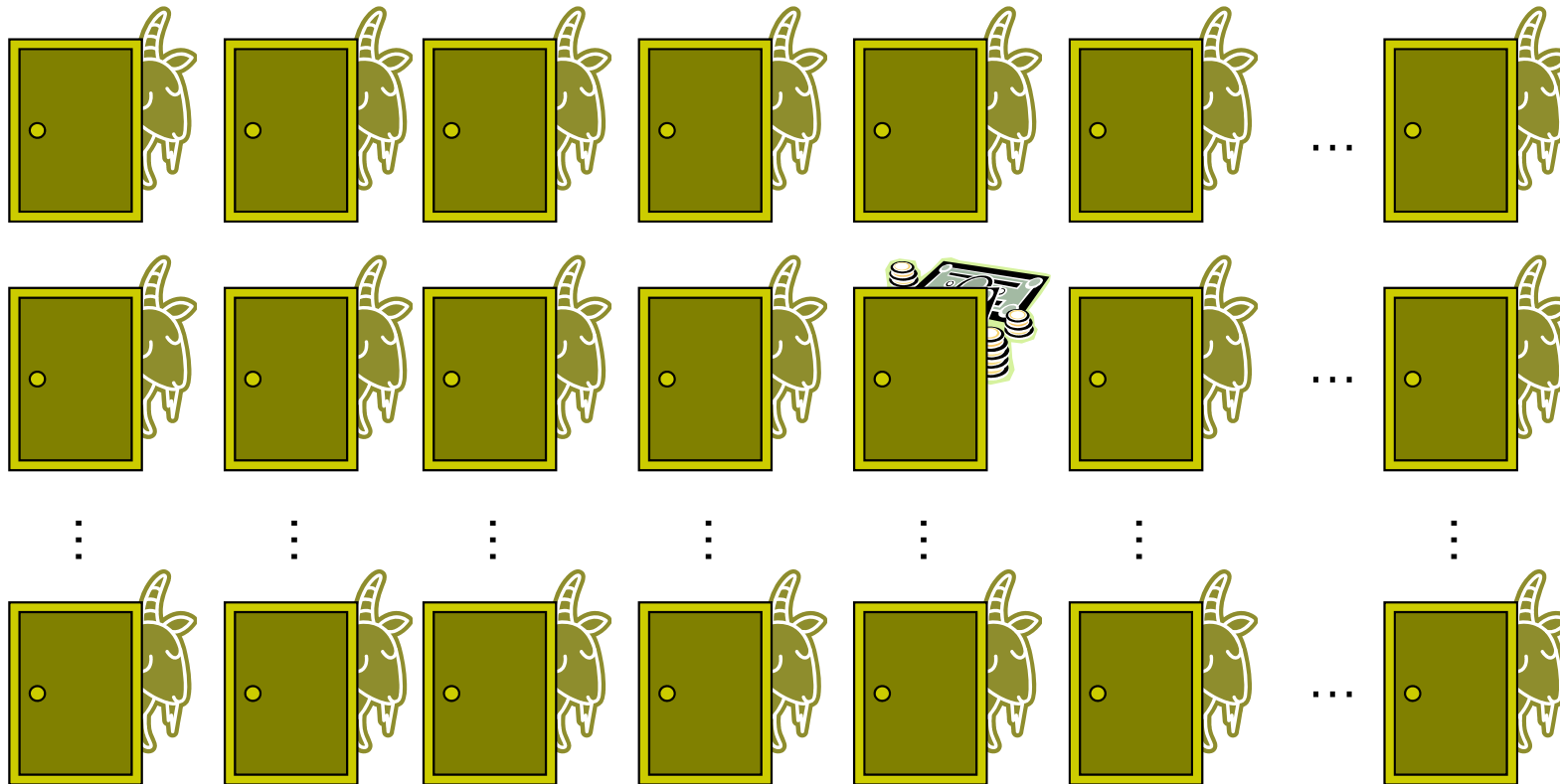
さあ、どうする？



Coffee Break!



Monty-Hole Dilemma



どうしても納
得いかなない人
のため、扉の
数を増やして
みましょう！

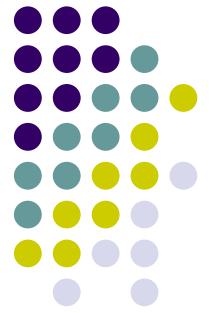
最初を選ぶ
扉が100個
あったらどう
かしら？

100個の扉からあなたが1つを選んだ後で、残り99の扉のうち山羊(はずれ)の98の扉を開いて見せます。さあ、開く扉を変えてもいいよ。それともやっぱり、あなたは開く扉を変えない？

あなたの最初の選択は神懸かり的な幸運に恵まれているのかしら？



Coffee Break!



クイズ・ミリオネアとお助けルール50-50

世の中で役に立つ確率統計3

問題: × × × × ...

A. ○○○○

B. △△△△

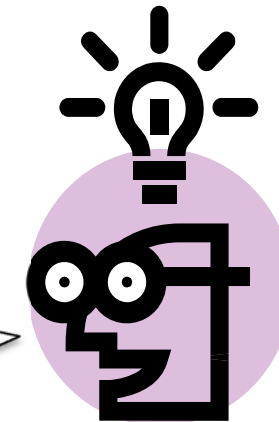
C. □□□□

D. ◇◇◇◇



さっぱり分からないよ(泣)

ならば、考えずにサイコロを
振って選びなさい



さいころ
の教え

A. ○○○○

B. △△△△

C. □□□□

D. ◇◇◇◇

A. ○○○○

~~B. △△△△~~

~~C. □□□□~~

D. ◇◇◇◇

50-50で2つ消してもらおう

A. ○○○○

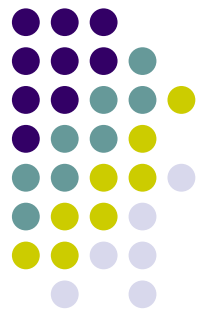
~~B. △△△△~~

~~C. □□□□~~

D. ◇◇◇◇

選択を変更して
ファイナルアンサー!
(当たる確率3/4)

Coffee Break!



答えづらい質問に答えて欲しいけど...

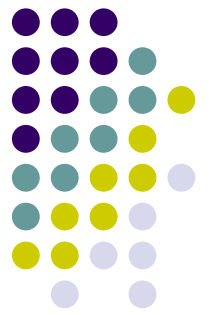
問: 思春期の男女90人に、恋人がいるかどうか調査したい

時間が無いので、「恋人はいるか」という質問に、『Yes』ならその場で手をあげてもらうことにしよう。手を上げなかった人は『No』ということだよね。

自分の答えが**みんなにばれてしまう**ので、
いないのに見栄を張って『YES』と答える人がいるかも...
いるのに、隠しておきたくて『NO』と答える人もいるよね...

正確な調査(正直な答え)を期待できるかな？ どうしたらいい？

Coffee Break!



実現例：ランダム回答法

90人全員にサイコロを振って貰う。出た目は誰にも知らせないこと

- ◆ 1,2が出た人『恋人はいるよね?』に回答 ... 「いる=Yes」「いない=No」
- ◆ 3-6が出た人『恋人はいないよね?』に回答...「いる=No」「いない=Yes」

Yesの人全員に手を上げて貰う(数えたら47人だった)。おわり。

本人以外は、どちらの質問に答えているのかわからない
= プライバシーの保護が確保される
= 正直な答えを期待できる(わざわざ嘘をつく理由がない)
そして、恋人がいる人数をある程度正確に推定できる!

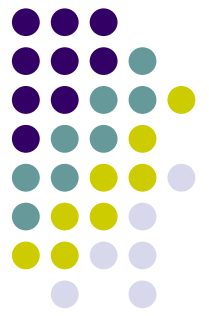
集計後の計算: 真の値(恋人がいる人の数)を x 人とする. よって, いない人の数は $90-x$ 人

- 恋人がいる x 人のうち, 上の質問に答える人は全体の $1/3$ いて, Yesと答える
- 恋人がいない $90-x$ 人のうち, 下の質問に答える人は全体の $2/3$ いて, Yesと答える

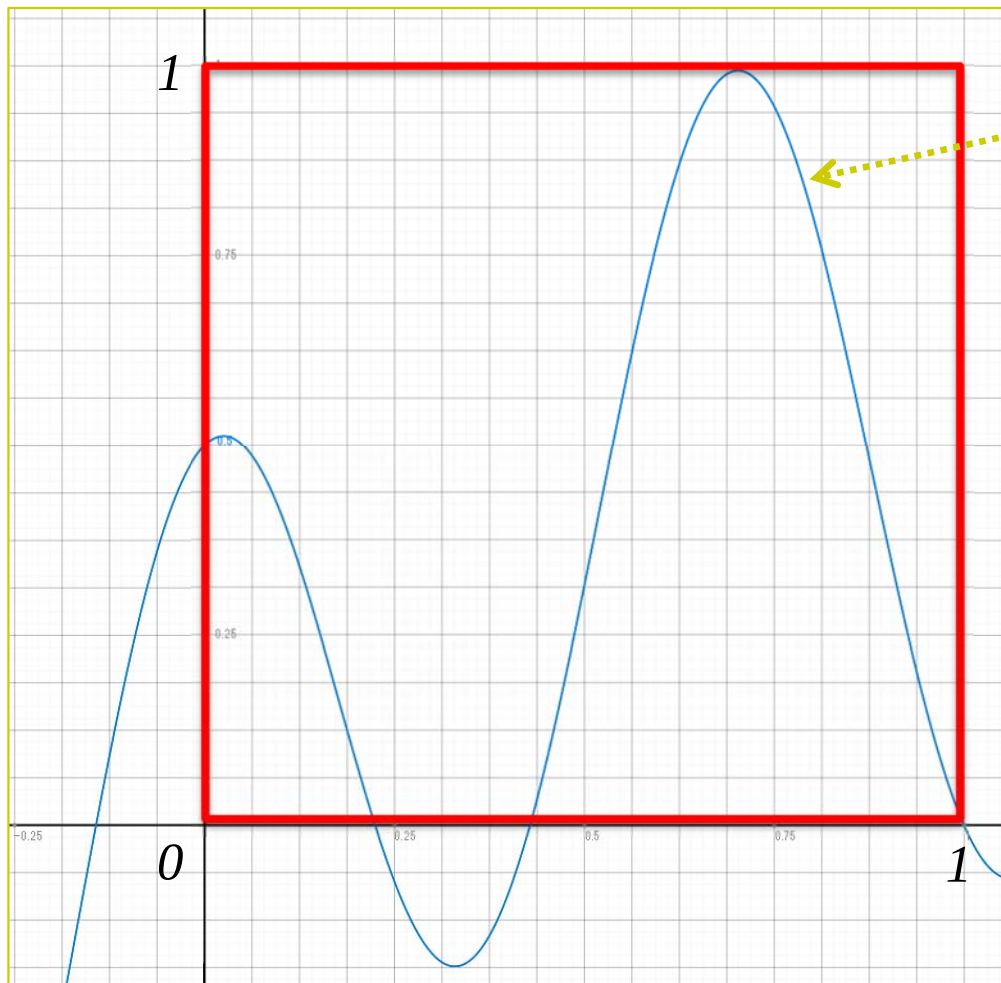
Yesと答えたのは全部で47人だったから...

$$\frac{1}{3}\hat{x} + \frac{2}{3}(90 - \hat{x}) = 47 \quad \therefore \hat{x} = 39$$

演習



問 原点を左下とする一辺1の正方形領域において、与えられた関数の下側にくる面積を求めなさい



$$f(x) = \frac{1}{2}(\sin(2x) + \cos(9x))$$

もっと知りたい人へ

- 関連する授業

- 「統計の分析と利用」(2セメ)
- 「データ分析」(3セメ)
- 「シミュレーションモデル分析」(4セメ)
- 「データ処理Ⅱ」(2・3セメ)
- etc...