

問題解決技法入門

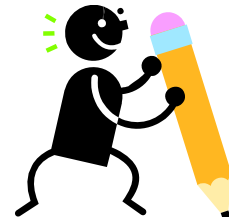
1. 統計・シミュレーションと予測

～サイコロをうまく使おう～

堀田 敬介



問題



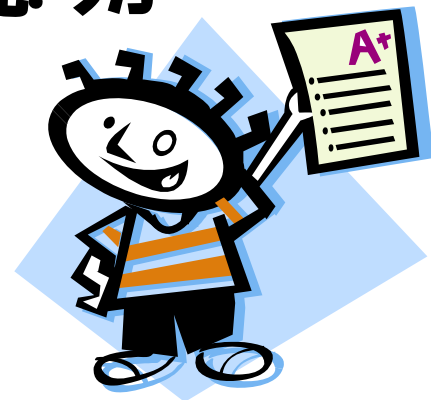
文教太郎君はマークシート形式のテストを受ける予定だ。全部で20問あり、各問題の選択肢は4つ(A,B,C,D)である。

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



わかんなきゃサイコロを振ろう

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>

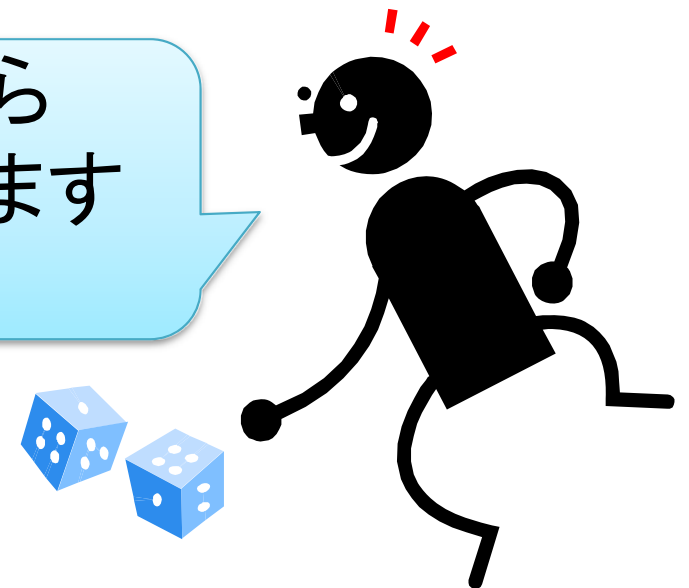
- ✓ 自分で振る場合→ 理論通りの出目 (cf. 正重心サイコロ, 重心がずれた賽の確率計算)
- ✓ コンピュータに振らせる場合→ 疑似乱数生成 (cf. 線形合同法, Mersenne twister, etc.)

※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

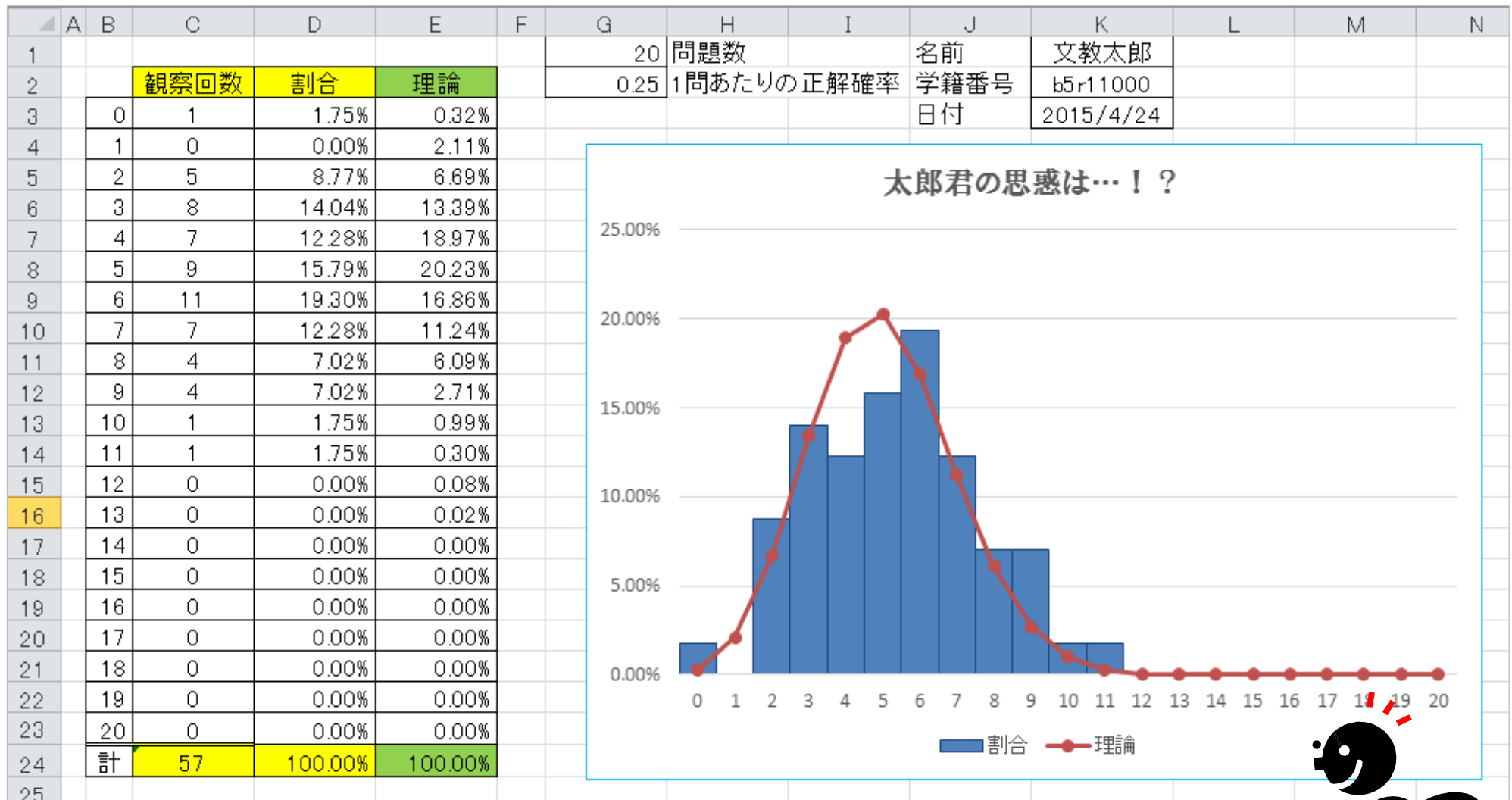
ここでも、本来の発言の趣旨と異なる形で引用しているので注意



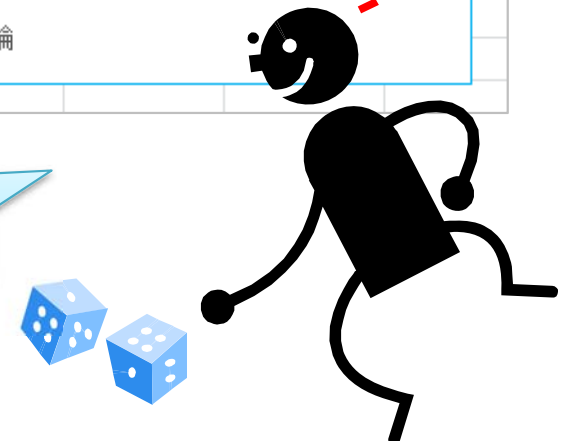
でも、人間ですから
サイコロ使っちゃいます
ゴメンね！



演習



さあ、トライ(試行)して予測しよう



Coffee Break!

ベイズ推定

ベイズの定理

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

確率の乗法定理

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A) \\ &= P(B)P(A|B) \end{aligned}$$

◎Q. データにもとづきコインがイカサマかどうか知りたい(コインの確率分布を知りたい)

◎Step0: (準備)

◎X... 仮定

◎B... 試行から得られた結果(データ)

◎P(X) ... 事前確率: 試行前のXの確率分布

◎P(X|B) ... 事後確率: 試行後のXの確率分布

◎P(B|X) ... 尤度: XのもとでBのおこる度合

◎P(B) ... Bがおきる確率

◎Step1: 試行1回目...10回コインを投げたら表が6回出た!

◎P(X)=1 ... Xの分布は不明なので一様分布(どの値をとる確率も同じ)とする

◎P(B)=6/10 ... 実際に10回中6回表が出た

◎P(B|X)= ${}_{10}C_6 x^6(1-x)^4$... 表の出る確率xのコインで, 10回中表が6回出る確率

$$\rightarrow P(X|B) = \frac{P(X)P(B|X)}{P(B)} = \frac{10}{6} {}_{10}C_6 x^6(1-x)^4 = k_1 x^6(1-x)^4 \rightarrow k_1 = \dots$$

和が1になるよう
係数を調整(積分)

Coffee Break!

ベイズ推定

ベイズの定理

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

確率の乗法定理

$$\begin{aligned} P(A \cap B) &= P(A)P(B|A) \\ &= P(B)P(A|B) \end{aligned}$$

©Step2: 試行2回目...さらに10回コインを投げたら表が9回出た!

© $P(X)=k_1x^6(1-x)^4$

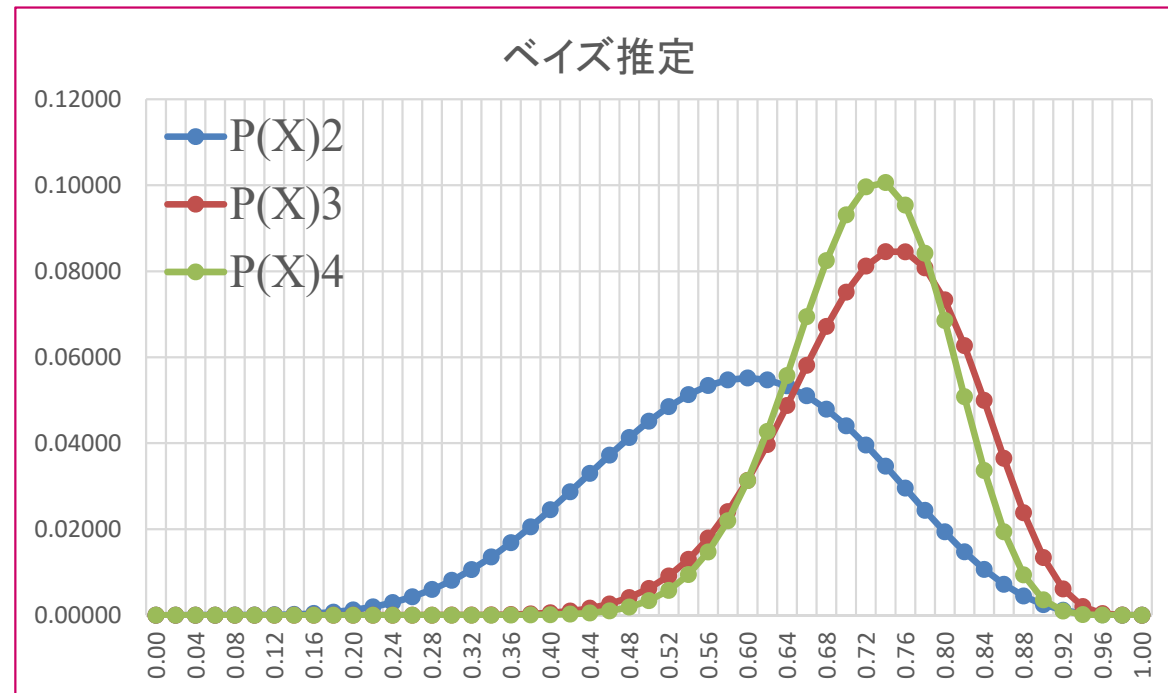
© $P(B)=9/10$... 実際に10回中9回表が出た

© $P(B|X)={}_{10}C_9 x^9(1-x)^1$... 表の出る確率 x のコインで, 10回中表が9回出る確率

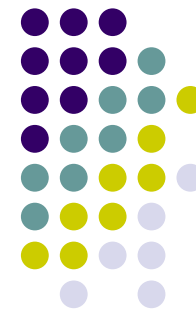
➡
$$P(X|B) = \frac{P(X)P(B|X)}{P(B)} = \frac{10}{9} k_1 x^6 (1-x)^4 {}_{10}C_9 x^9 (1-x) = k_2 x^{15} (1-x)^5$$

➡ $k_2=...$

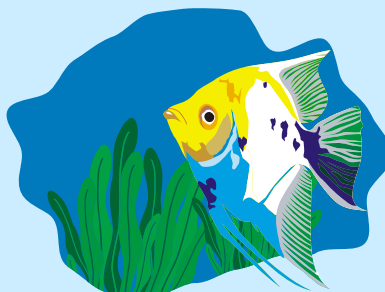
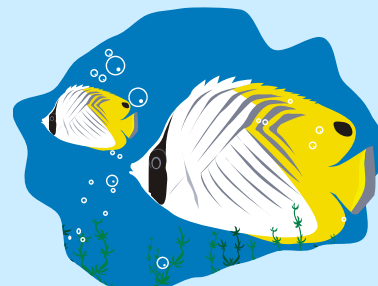
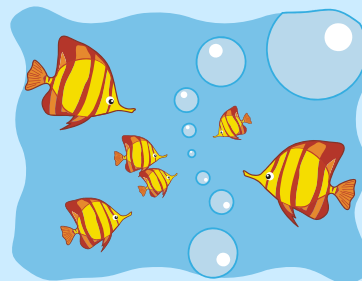
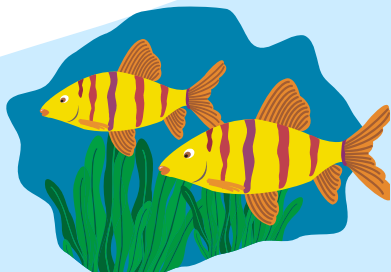
©Step3: 試行3回目...



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(m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(m匹)

未知の推定したい数値

総数N匹: 印●付きM匹

再捕獲n匹: 印●付きm匹

既知の観測数値

→ $N:M = n:m$ より
推定値: $\hat{N} = \frac{Mn}{m}$

例) $M=300, n=500, m=5$ なら

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

cf. 超幾何分布

演習

問 文教太郎君の受けた試験が、1問10点の**全10問**で、
すべて**3択**だった場合、彼が 秘伝・鉛筆転がし で単位をと
れる(60点以上とる)可能性を予測しなさい

シミュレーション回数(試行回数)は、 x 回とする
(つまり、太郎君が同じ試験を x 回受けたこととする)

ただし、「 $x = 30 + \text{あなたの学籍番号下1桁}$ 」とする

<ヒント>

3択なので、6面体サイコロを振った時、1か2の目が出たら正解！とすればよい
試験は全10問なので、1回のシミュレーションで10個サイコロをふり、1と2の目が何回出たかを観察する。それを x 回行う
結果が出たら、60点以上で単位取得なので、該当する部分から確率(予測値)を計算



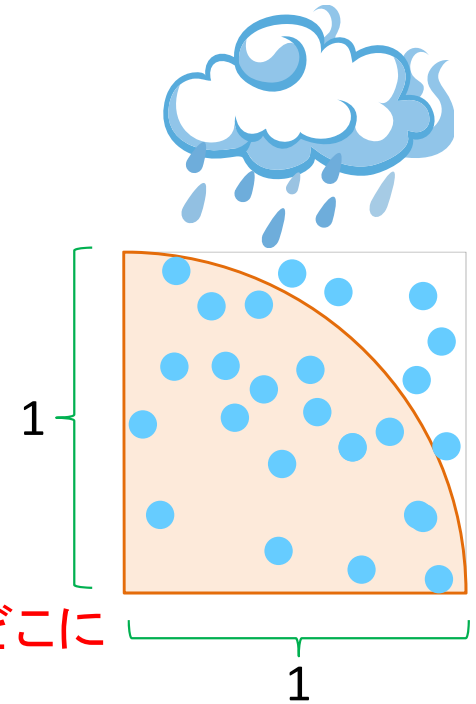
問題

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

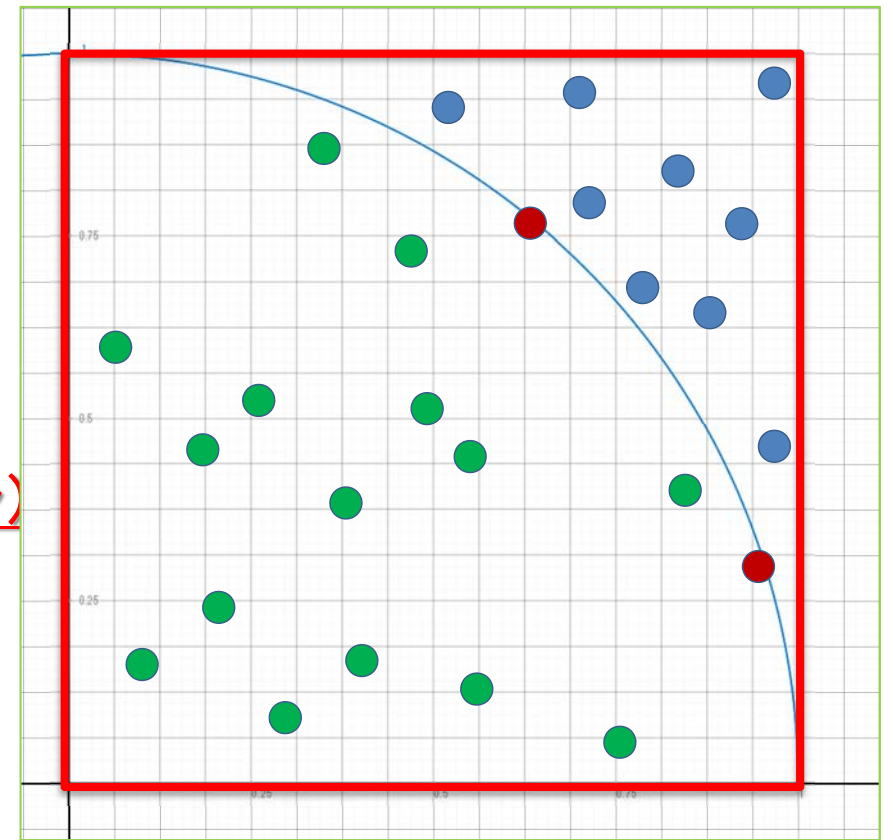
問 円周率 π はいくつか？

モンテカルロ法(モンテカルロ・シミュレーション)

$$\pi = 4 \times R \div T$$

T=降らせる雨粒の数(例:T=10,000粒)

R=円内に降った雨粒の数(数える/計算させる)



原点中心半径 r の円の式は $x^2 + y^2 = r^2$ 半径 1 なので, $x^2 + y^2 = 1$

ある点 (x, y) (=1つの雨粒) が, 円の内部にある(降る)とは,
 $x^2 + y^2 < 1$

を満たすということ

つまり, ある点 (x, y) (=1つの雨粒) が円内に降ったかどうかは,
 $x^2 + y^2$

を計算した結果が1より小さいかどうかを見れば良い

- : 円内にある点(雨粒)
- : 円外にある点(雨粒)
- : 円上にある点(雨粒)

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)

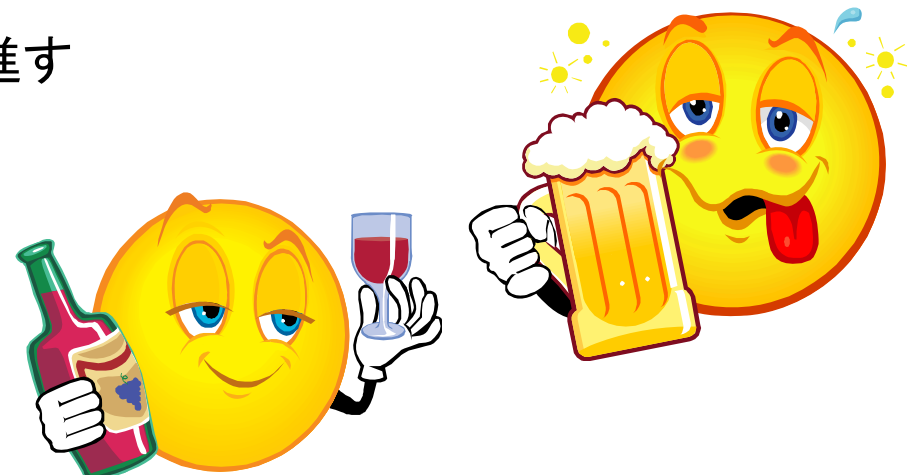
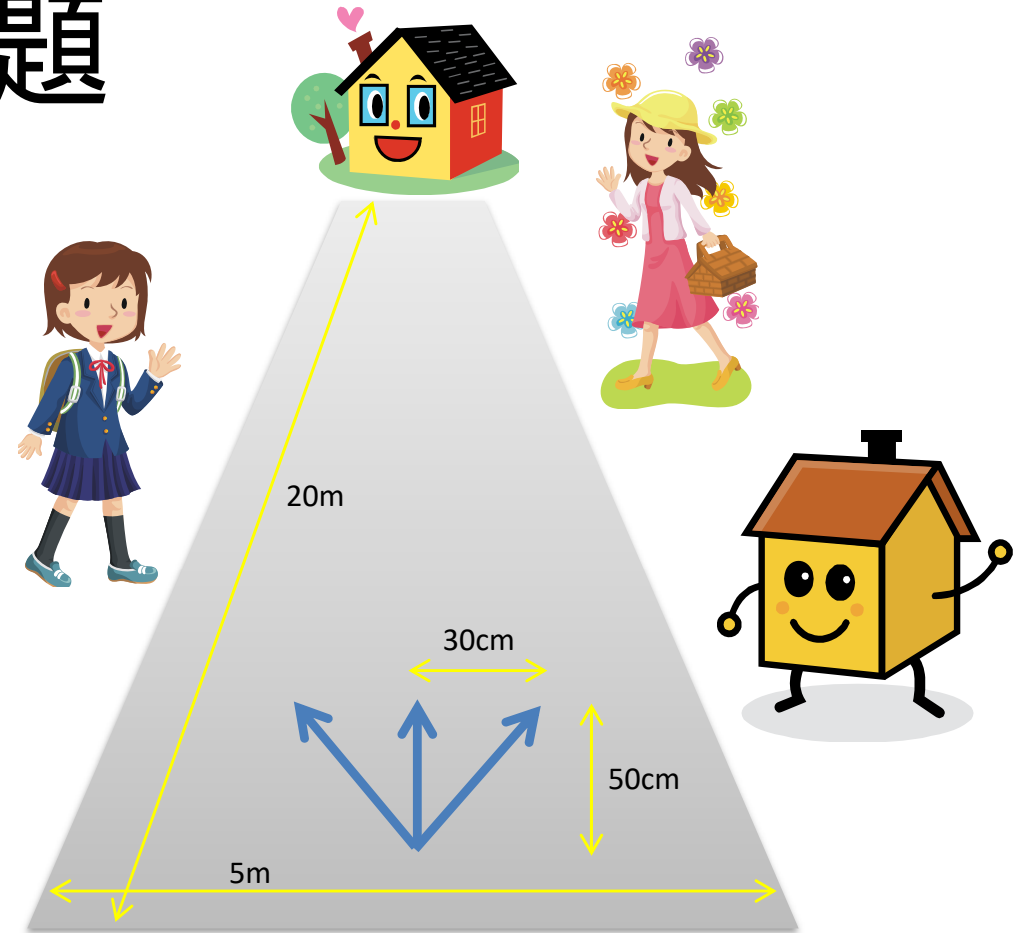
彼の歩みは

- 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

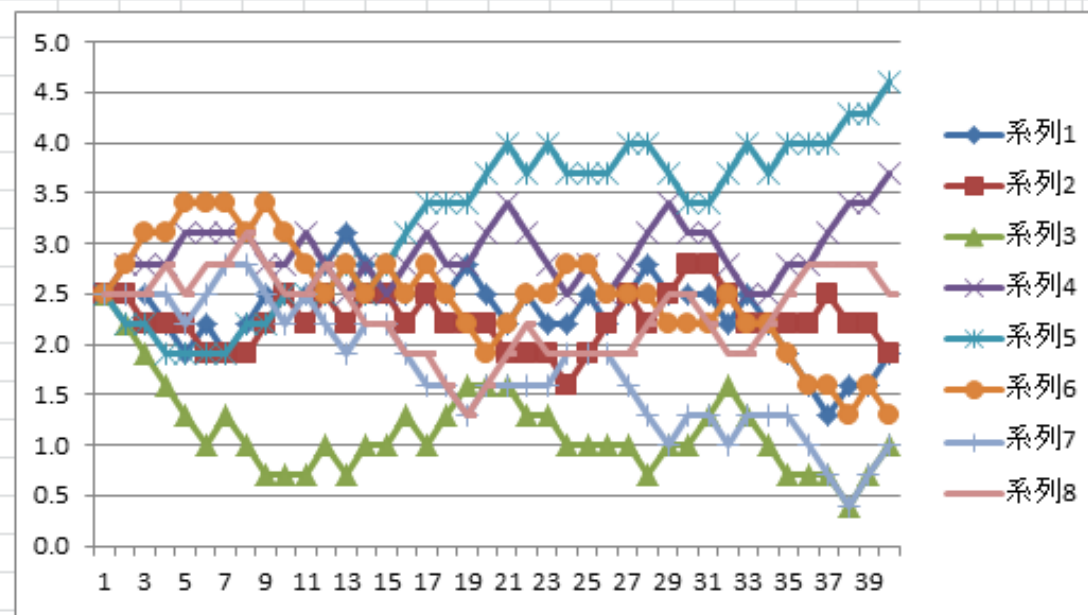
現在位置(初期位置)は道路の真ん中
とし, 家まで20mとする

この設定で**シミュレーション**をするよ



演習

	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AAAAAAAAAAAAAAAAAAAA	AY	AZ	BA	BB														
1		start																																							goal
2	歩数(歩)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21					37	38	39	40											
3	距離(m)	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0	10.5					18.5	19.0	19.5	20.0											
4	位値	2.5	2.5	2.5	2.2	1.9	2.2	1.9	2.2	2.5	2.5	2.5	2.8	3.1	2.8	2.5	2.5	2.5	2.5	2.8	2.5	2.2					1.3	1.6	1.6	1.9											
5		2.5	2.5	2.2	2.2	2.2	1.9	1.9	1.9	2.2	2.5	2.2	2.5	2.2	2.5	2.5	2.2	2.5	2.2	2.2	2.2	1.9					2.5	2.2	2.2	1.9											
6		2.5	2.2	1.9	1.6	1.3	1.0	1.3	1.0	0.7	0.7	0.7	1.0	0.7	1.0	1.0	1.3	1.0	1.3	1.6	1.6	1.6					0.7	0.4	0.7	1.0											
7		2.5	2.8	2.8	2.8	3.1	3.1	3.1	3.1	2.8	2.8	3.1	2.8	2.5	2.8	2.5	2.8	3.1	2.8	2.8	3.1	3.4					3.1	3.4	3.4	3.7											
8		2.5	2.2	2.2	1.9	1.9	1.9	1.9	2.2	2.2	2.5	2.5	2.5	2.8	2.5	2.8	3.1	3.4	3.4	3.4	3.7	4.0					4.0	4.3	4.3	4.6											
9		2.5	2.8	3.1	3.1	3.4	3.4	3.4	3.1	3.4	3.1	2.8	2.5	2.8	2.5	2.8	2.5	2.8	2.5	2.2	1.9	2.2					1.6	1.3	1.6	1.3											
10		2.5	2.5	2.5	2.5	2.2	2.5	2.8	2.8	2.5	2.2	2.5	2.2	1.9	2.2	2.2	1.9	1.6	1.6	1.3	1.6	1.6					0.7	0.4	0.7	1.0											
11		2.5	2.5	2.5	2.8	2.5	2.8	2.8	3.1	2.8	2.5	2.5	2.8	2.5	2.2	2.2	1.9	1.9	1.6	1.3	1.6	1.9					2.8	2.8	2.8	2.5											



無事にたどり着けたかな？



2次元ランダムウォーク

問 酔っ払いが道を歩いている

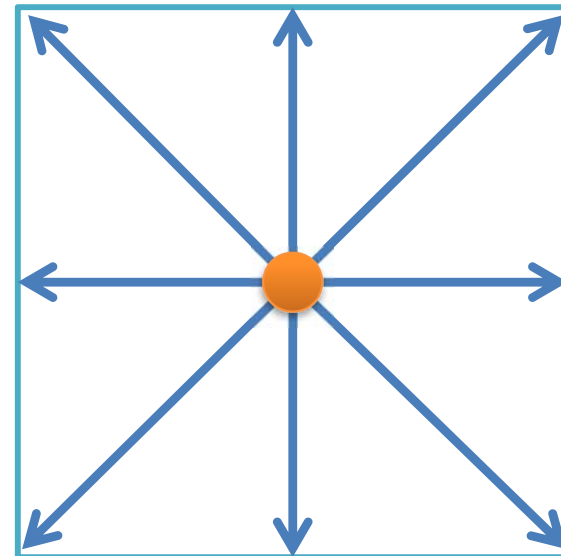
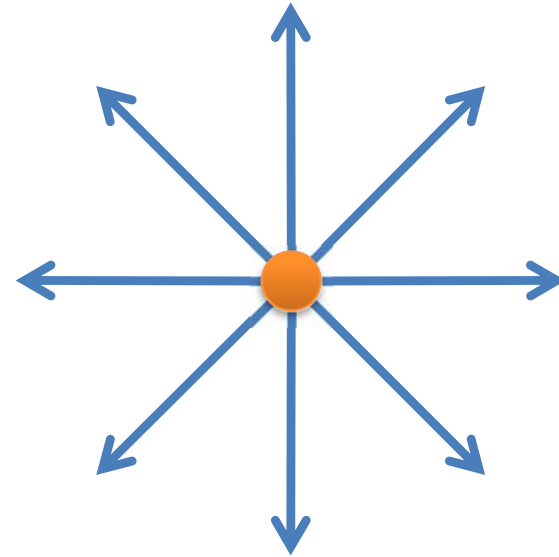
あっちにふらふら, こっちにふら
ふら, …あぶなっかしい

無事に家までたどり着けるだ
ろうか？

2次元ランダム・ウォーク(酔歩)

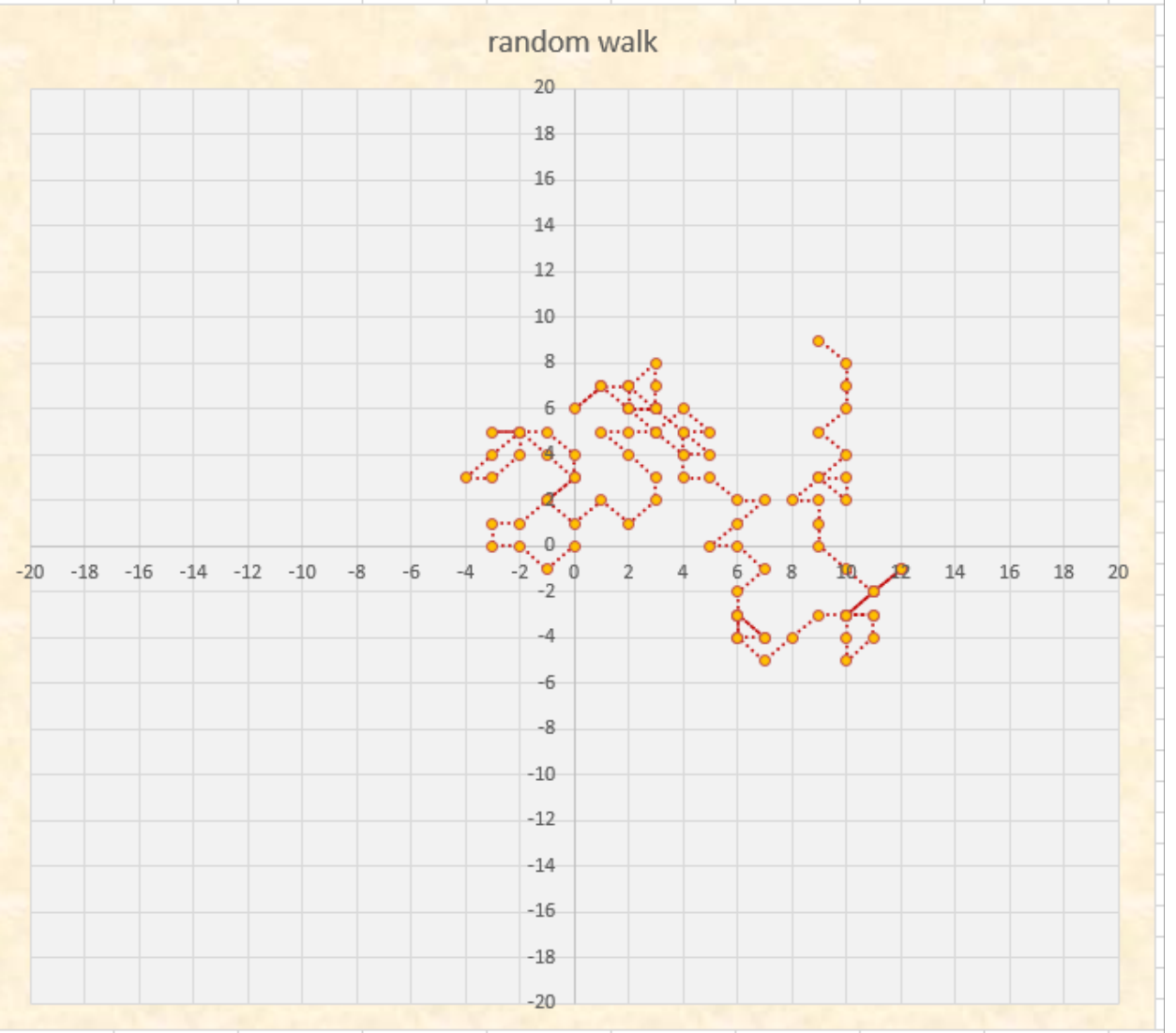
東西南北8方向にそれぞれ1/8の確率で進む

この設定でシミュレーションをするよ



演習

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	2次元ランダムウォーク(酔歩)																					
2									2		3	4										
3		歩幅		1				乱数			x	y										
4		次の一步		x座標	y座標		1				(	0	0)									
5		1	東	1	0		2	4	南西		(	-1	-1)									
6		2	南東	1	-1		3	6	北西		(	-2	0)									
7		3	南	0	-1		4	5	西		(	-3	0)									
8		4	南西	-1	-1		5	7	北		(	-3	1)									
9		5	西	-1	0		6	1	東		(	-2	1)									
10		6	北西	-1	1		7	8	北東		(	-1	2)									
11		7	北	0	1		8	8	北東		(	0	3)									
12		8	北東	1	1		9	7	北		(	0	4)									
13							10	6	北西		(	-1	5)									
14							11	5	西		(	-2	5)									
15							12	3	南		(	-2	4)									
16							13	4	南西		(	-3	3)									
17							14	5	西		(	-4	3)									
18							15	8	北東		(	-3	4)									
19							16	8	北東		(	-2	5)									
20							17	5	西		(	-3	5)									
21							18	1	東		(	-2	5)									
22							19	2	南東		(	-1	4)									
23							20	2	南東		(	0	3)									
24							21	4	南西		(	-1	2)									
25							22	2	南東		(	0	1)									
26							23	8	北東		(	1	2)									
27							24	2	南東		(	2	1)									
28							25	8	北東		(	3	2)									
29							26	7	北		(	3	3)									
30							27	6	北西		(	2	4)									
31							28	6	北西		(	1	5)									
32							29	1	東		(	2	5)									
33							30	1	東		(	3	5)									
34							31	6	北西		(	2	6)									
35							32	1	東		(	3	6)									
36							33	5	西		(	2	6)									



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)

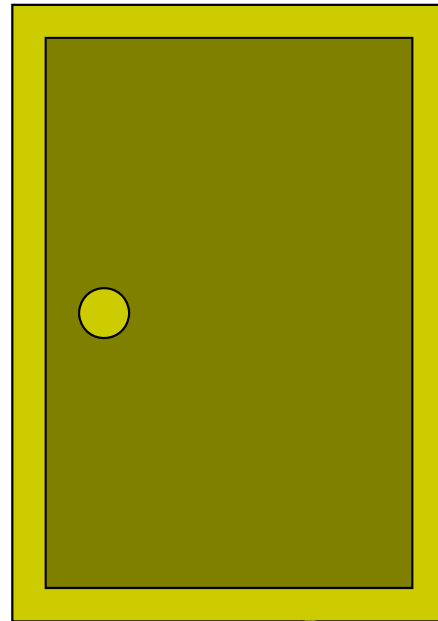
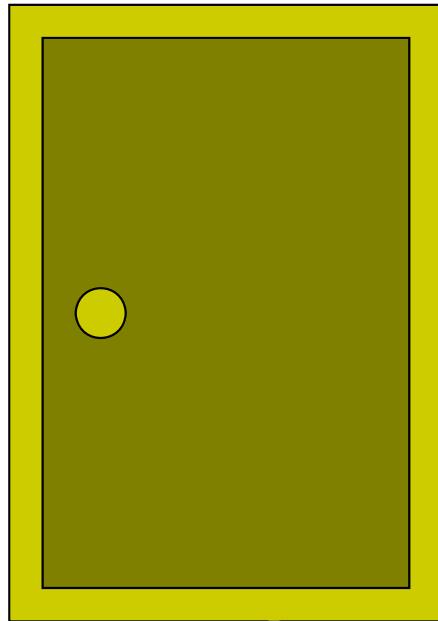
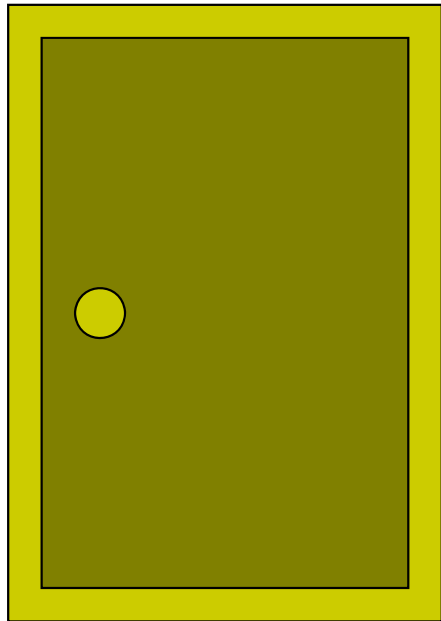
• **山羊** (はずれ)

• **山羊** (はずれ)

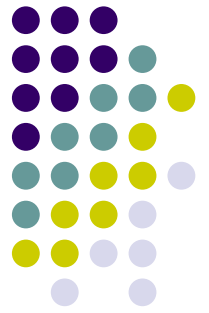
が隠されているよ。
あなたは扉を1つだけ選んでいいのよ。

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。

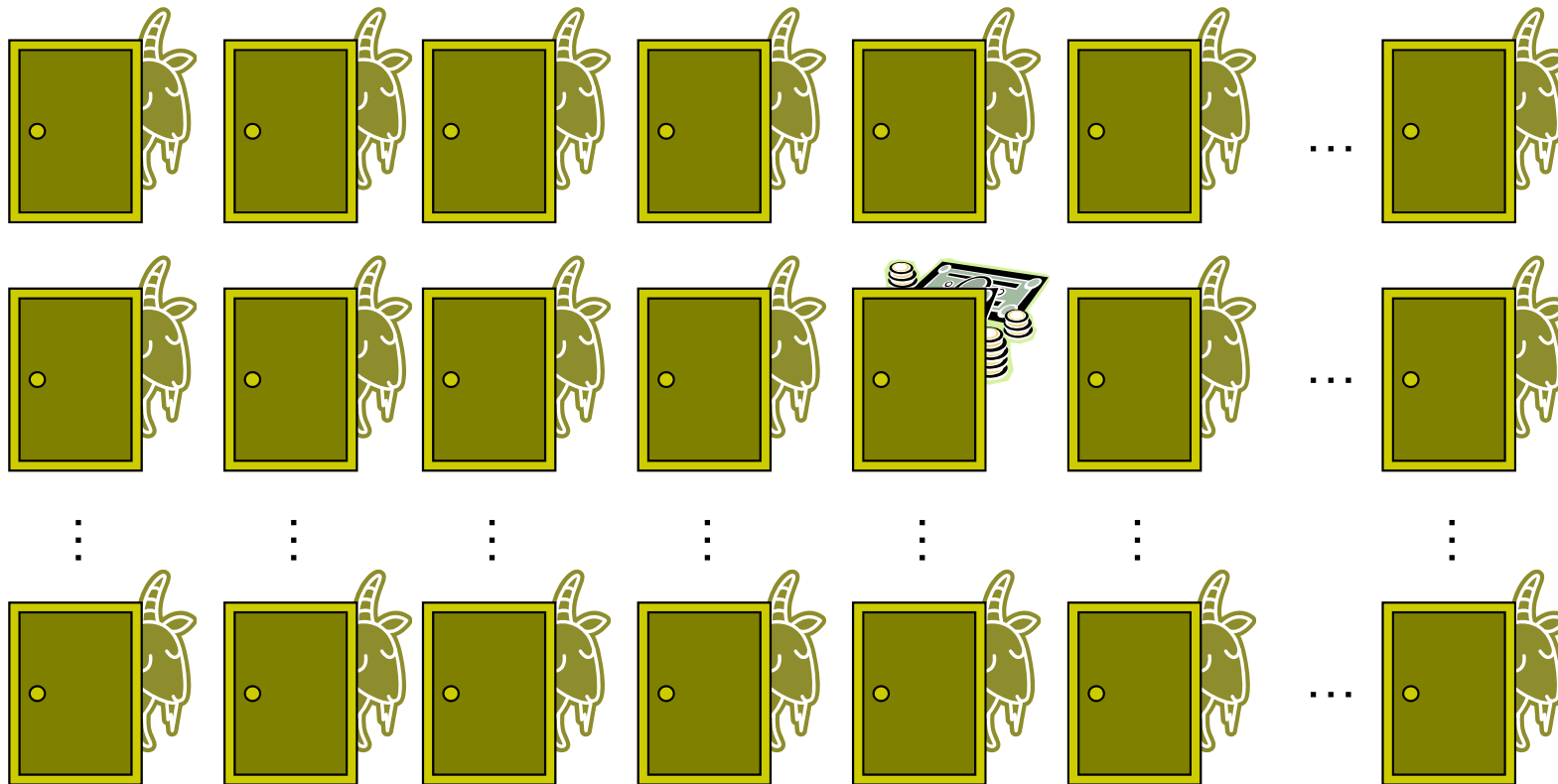
さあ、どうする？



Coffee Break!



Monty-Hole Dilemma



どうしても納得
いかない人の
ため、扉の
数を増やして
みましょう！

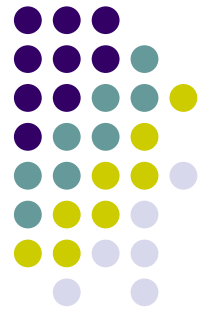
最初に選ぶ
扉が100個
あったらどう
かしら？

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. △△△△

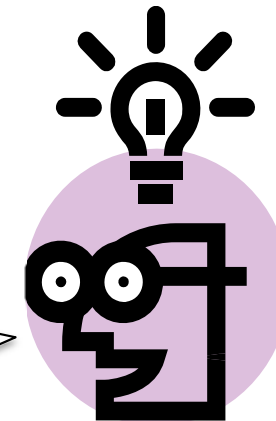
C. □□□□

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. ○○○○

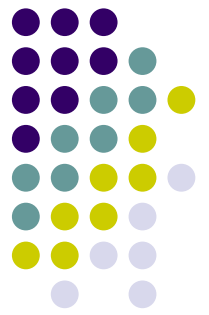
~~B. △△△△~~

~~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...

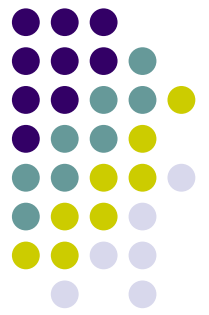
問: 思春期の男女90人に、恋人がいるかどうか調査したい

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...

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

もっと知りたい人へ

- 関連する授業

- 「統計の見方」(1/2セメ)
- 「統計の分析と利用」(2セメ)
- 「データ処理Ⅱ」(2・3セメ)
- 「データ分析」(3セメ)
- 「統計データの扱い方」(3/4セメ)
- 「多変量の統計データ解析」(4セメ)
- 「シミュレーションモデル分析A」(4セメ)
- 「シミュレーションモデル分析B」(5セメ)
- etc...