Sept. 24, 2015@ELC秋学校(岐阜)

Exploring the
Limits of
Computation

乱択の技法

来嶋秀治

九州大学 大学院システム情報科学研究院

研究の興味

敬称略

確率的アルゴリズム

- 完璧サンプリング法 (MCMC法) (w/松井)
- ランダムウォークの脱乱択化 (w/牧野, 古賀, 梶野, 白髪, 山内, 山下)
- データストリーム中の頻出アイテム検知 (w/緒方, 山内, 山下)
- 置換多面体上の点の乱択丸め (w/畑埜, 瀧本, 安武, 末廣, 竹田, 永野)
- メディアン安定結婚問題 (w/根本)

グラフアルゴリズム

- 区間グラフの部分グラフ同型性判定問題はP? NP-c? (w/斎藤, 大舘, 宇野)
- 区間グラフ上の支配集合数え上げ問題はP? #P-c? (w/岡本, 宇野)
- 最大次数4の剛性サーキットは，いつでも
  2つの互いに疎なハミルトン閉路に分解できるか？ (w/谷川)

分散アルゴリズム

(w/ 溝口, 徐, 小野, 山内, 山下)

- ポピュレーションプロトコルのリーダー選挙問題の領域計算量
- 円盤上の自律ロボットの脱出 (w/ Gasieniec, 山内, 山下)

「MCMCとか決定的酔歩の話をしてほしい．」

by上原先生

確率的技法を使う人に増えてほしい．

確率、苦手

- ✓ 使い所がわからない.
  - ➢ 簡単なの？難しいの？
  - ➢ まるでサーカス.
- ✓ 解析がマニアック.
  - ➢ 式が多い.
- ✓ 難癖つける.
  - ➢ 不偏分散とか，t検定とか，細かい.
  - ➢ そのくせ，「高い確率」とかやたら大雑把.
- ✓ 結局，何をすれば論文になるの？

実は、私の愚痴

今日の趣旨
基礎技術を押さえる

- 確率は定義的には「足し算，線形」の世界
- ここに割り算(条件付き確率)や非線形関数(相関)を持ち込むと，途端にぐちゃぐちゃ(今日は避ける).

確率、使って

- ✓ **確率にはよくlogが出てくる**
  - ➢ 計算量改善のチャンス！
- ✓ ２乗，平方根は確率の歴史
  - ➢ Taylor展開の遺産（原理），
  - ➢ 一番簡単な強凸関数
- ✓ 期待値は線形関数
  - ➢ 線形計画，多面体，線形代数etc. との相性抜群
  （だから皆使ってる．）
- ✓ **全列挙とサンプリング**

- ✓ Computer Scienceの多くの論文では，難しい確率は避けている．
- ✓ 難しい技術を手に入れれば，論文チャンス(しかし、難しい)
  - ➢ Martingale, 測度, 極値理論

乱択の技法: 今日の話題

基礎技術1. $\left(1-\frac{1}{n}\right)^n \simeq \frac{1}{\mathrm{e}}$

基礎技術2. logの出てくる仕組み

基礎技術3. サンプリングのサイズはなぜ $\left(\frac{1}{\varepsilon}\right)^2$?

- レアイベント(起こりにくい事象)の確率

  ～ 確率不等式の世界へようこそ ～

基礎技術4. Median Trick

応用. 高度モンテカルロ法

- サンプリングと積分

今後の課題. BPL=PSPACE?

私のよく使う技法

基礎技術 1. $\left(1-\frac{1}{n}\right)^n \simeq \frac{1}{\mathrm{e}}$

Birthday paradox

問題 1.
40人の人が居ます．誕生日の同じ人の居る確率は？

$$\frac{40}{365} \simeq \frac{1}{9} ?$$

問題 2.(悪用)
ここに300人の人が居ます．各自，1～10000までの数字を一つ書いてください．もしこの中に同じ数字の人が居ると…

$$\frac{300}{1000} = \frac{3}{100} ?$$

例: それは悪魔の仕業です．この場に居る皆さんに不幸が訪れます．お祓いしましょう．１万円です．

ヒント
余事象を考える．

問題

「$\{1, \ldots, n\}$の中から一様ランダムに数字を選ぶ」操作を $k\sqrt{n}$ 回繰り返す．数字が重複しない確率は？

$$\frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k\sqrt{n}}{n} \simeq$$

問題

「$\{1,\ldots,n\}$の中から一様ランダムに数字を選ぶ」操作を $k\sqrt{n}$ 回繰り返す．数字が重複しない確率は？

$$\frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k\sqrt{n}}{n} \leq 1^{\sqrt{n}} \cdot \left(\frac{n-\sqrt{n}}{n}\right)^{(k-1)\sqrt{n}}$$

問題

「$\{1, \ldots, n\}$の中から一様ランダムに数字を選ぶ」操作を $k\sqrt{n}$ 回繰り返す．数字が重複しない確率は？

$$\frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k\sqrt{n}}{n} \le 1^{\sqrt{n}} \cdot \left(\frac{n-\sqrt{n}}{n}\right)^{(k-1)\sqrt{n}}$$

$$= \left(1 - \frac{1}{\sqrt{n}}\right)^{(k-1)\sqrt{n}}$$

問題

「$\{1,\dots,n\}$の中から一様ランダムに数字を選ぶ」操作を $k\sqrt{n}$ 回繰り返す．数字が重複しない確率は？

$$\frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k\sqrt{n}}{n} \le 1^{\sqrt{n}}\cdot\left(\frac{n-\sqrt{n}}{n}\right)^{(k-1)\sqrt{n}}$$

$$=\left(1-\frac{1}{\sqrt{n}}\right)^{(k-1)\sqrt{n}}$$

$$=\left(\left(1-\frac{1}{\sqrt{n}}\right)^{\sqrt{n}}\right)^{k-1}$$

$$\simeq\left(\frac{1}{\mathrm{e}}\right)^{k-1}$$

$$300=3\sqrt{10000}$$

$$\Rightarrow k-1=2$$

$$\Rightarrow\left(\frac{1}{e}\right)^{2}\simeq 0.135$$

基礎技術 2. logのでてくる仕組み

確率には log が良く出てくる．

問題 3.

人気クイズ番組「クイズ超難問！」では，２択の問題が20問出される．問題は超難問ばかりで，たとえば

「強いのはどっち？ A. 赤鬼，B. 青鬼」

といった，勘で答えるより他にない問題ばかりである．20問連続正解すると豪華賞品（100万USD）がもらえるが，現在までに67人が挑戦して，全問正解者は居ない．全問正解が出るのは，いつごろ（何人目ごろ）か？

A. 約千人目ごろ

B. 約百万人目ごろ

9人目までの戦績

1. ○○×
2. ×
3. ○×
4. ×
5. ×
6. ○○○○○○○×
7. ○○×
8. ×
9. ×

9人目までの戦績

1. ○○×○○×○○×××○××○××○○×
2. ×××○○××××○○×○×○○×××○
3. ○×○○××○○×○××××○○×○××
4. ×○○×××○○××○×○×○×○××○
5. ××○××○○××○○×○××○××○○
6. ○○○○○○○×××○×○○○×○○×○
7. ○○×○○○×○×○××○×○×××○×
8. ×××○○×○○○××○○×××○×××
9. ×○×○○×○○××○××○×○×○○×

9人目までの戦績

1. 0.00100100111011011001
2. 0.11100111100101001110
3. 0.01001100101111001011
4. 0.10011100110101010110
5. 0.11011001100101101100
6. 0.00000001110100010010
7. 0.00100010101101011101
8. 0.11100100011001110111
9. 0.10100100110110101001

つまり

0.00000000000000000000から

0.11111111111111111111

の数字が一様ランダム.

実数乱数$r \in [0,1)$が

0.00000000000000000001以下の確率 $= \frac{1}{2^{20}}$

基礎技術 3. サンプリングのサイズはなぜ $\left(\frac{1}{\varepsilon}\right)^2$?

レアイベント(起こりにくい事象)の確率

～確率不等式の世界へようこそ～

全然別の問題: tail boundの世界

問題 4.

２択問題の専門家が101人居て，どの専門家も確率0.6以上で正解する．ある２択問題に対して，101人のうち過半数がAと答えた．Aが正しい確率は？

過半数が不正解する確率

全然別の問題: tail boundの世界

問題 4.

２択問題の専門家が101人居て，どの専門家も確率0.6以上で正解する．ある２択問題に対して，101人のうち過半数がAと答えた．Aが正しい確率は？

過半数が不正解する確率

=正解者の数が50人以下の確率

$$\sum_{k=0}^{50}\binom{101}{k}0.6^k 0.4^{n-k} = ???$$

二項分布の裾

定理[Chvatal]

$$\sum_{j=0}^{(p-t)m} \binom{m}{j} p^j (1-p)^{m-j} \leq \exp(-2t^2 m)$$

$m = 101,\ p = 0.6,\ t = 0.1$とすると,

$$\sum_{k=0}^{50} \binom{101}{k} 0.6^k 0.4^{n-k} \leq \exp(-2 \times 0.1^2 \times 101)$$

$$\leq \frac{1}{\mathrm{e}^2} \simeq 0.135$$

過半数が正解する確率 ≥ 0.865

V. Chvatal, The tail of the hypergeometric distribution, Discrete Mathematics, 25 (1979), 285–287.

一応ベイズも書いておく(省略)

A^*:Aが正解, A:Aが過半数

$$\Pr[\, A \mid A^* \,] \geq 0.865$$

ベイズの定理より

$$\Pr[\, A^* \mid A \,] = \frac{\Pr[A \mid A^*]\Pr[A^*]}{\Pr[A]}$$

$$= \frac{0.865 * 0.5}{0.5 * 0.865 + 0.5 * 0.135}$$

	A (Aが過半数)	B (Bが過半数)
A* (Aが正解)	≥ 0.865	≤ 0.135
B* (Bが正解)	≤ 0.135	≥ 0.865

事前情報が無い（Aが正解の確率が0.5）から.

超幾何分布

超幾何分布 (hypergeometric distribution)

Ωは有限集合とΞはその部分集合とする.

$|\Omega| = M$, $|\Xi| = N$とする.

Ωからm個の非復元抽出を行ったとき,

そのサンプルSに含まれるΞの要素の数がnとなる確率は

$$\frac{\binom{N}{n}\binom{M-N}{m-n}}{\binom{M}{m}} \quad \left(= \frac{\binom{m}{n}\binom{M-m}{N-n}}{\binom{M}{N}}\right)$$

例

トランプのカード52枚から10枚のカードを引いたとき,

ハートが3枚出る確率.

$$\frac{\binom{13}{3}\binom{39}{7}}{\binom{52}{10}}$$

超幾何分布の裾

定理[Chvatal]

$p = \frac{N}{M},\ \ n = (p+t)m$ としたとき,

$$\sum_{j=k}^{m} \frac{\binom{N}{j}\binom{M-N}{m-j}}{\binom{M}{m}} \le \left(\left(\frac{p}{p+t}\right)^{p+t}\left(\frac{1-p}{1-p-t}\right)^{1-p-t}\right)^{n}$$

V. Chvatal, The tail of the hypergeometric distribution, Discrete Mathematics, 25 (1979), 285–287.

確率不等式

確率不等式

- ✓ Markovの不等式
- ✓ Chebyshevの不等式
- ✓ Chernoffの不等式({0,1}確率変数)
- ✓ Hoefdingの不等式(有限値)

参考: 二項分布，幾何分布の裾

- ✓ Chvatal

定理 (マルコフの不等式; Markov's inequality)

定理: (マルコフの不等式; Markov's inequality)

X は非負の確率変数とする. 任意の定数 $a > 0$ に対して,

$$\Pr[X \geq a] \leq \frac{\mathrm{E}[X]}{a}$$

証明

$$\frac{E[X]}{a} = E\left[\frac{X}{a}\right]$$

$$= \int_0^{\infty} \frac{x}{a} f(x)\mathrm{d}x \geq \int_a^{\infty} \frac{x}{a} f(x)\mathrm{d}x \geq \int_a^{\infty} f(x)\,\mathrm{d}x = \Pr[X \geq a]$$

したがって

$$\Pr[X \geq a] \leq \mathrm{E}\left[\frac{X}{a}\right] = \frac{\mathrm{E}[X]}{a}$$

□

Chebyshev’s inequality

定理 [チェビシェフの不等式; Chebyshev’s inequality]

任意の a > 0 に対して,

$$\Pr[|X - E[X]| \geq a] \leq \frac{\mathrm{Var}[X]}{a^2}$$

証明

次の事実に注意する.

$$\Pr[|X - E[X]| \geq a] = \Pr[(X - E[X])^2 \geq a^2]$$

マルコフの不等式より,

$$\Pr[(X - E[X])^2 \geq a^2] \leq \frac{E[(X - E[X])^2]}{a^2} = \frac{\mathrm{Var}[X]}{a^2}$$

Chebyshev’s inequality

系 [チェビシェフの不等式; Chebyshev’s inequality]

任意の a > 0 に対して,

$$\Pr[X \geq (1+t)\cdot \mathrm{E}[X]] \leq \frac{\mathrm{Var}[X]}{(t\cdot \mathrm{E}[X])^2}$$

証明

$$\begin{aligned}\Pr[X \geq (1+t)\cdot \mathrm{E}[X]] &= \Pr[X-\mathrm{E}[X] \geq t\cdot \mathrm{E}[X]] \\ &\leq \Pr\big[|X-\mathrm{E}[X]| \geq t\cdot \mathrm{E}[X]\big] \\ &\leq \frac{\mathrm{Var}[X]}{(t\cdot \mathrm{E}[X])^2}\end{aligned}$$

Chernoff bound

同一分布でなくて良い.

定理 [Chernoff]

$X_1, X_2, \dots, X_n$は相互に独立な {0,1}確率変数とし,
$X := X_1 + \cdots + X_n$の期待値をμとすると,

$$\Pr[X \geq (1+\varepsilon)\mu] < \left(\frac{\mathrm{e}^{\varepsilon}}{(1+\varepsilon)^{(1+\varepsilon)}}\right)^{\mu}$$

$$\Pr[X \leq (1-\varepsilon)\mu] < \left(\frac{\mathrm{e}^{\varepsilon}}{(1-\varepsilon)^{(1-\varepsilon)}}\right)^{\mu}$$

Poisson試行と呼ばれる.

系

$0 < \varepsilon \leq 1$ の場合,

$$\Pr[X \geq (1+\varepsilon)\mu] \leq \exp\left(-\frac{\varepsilon^2\mu}{3}\right)$$

$$\Pr[X \leq (1-\varepsilon)\mu] \leq \exp\left(-\frac{\varepsilon^2\mu}{2}\right)$$

証明のさわり

$$\Pr[X \geq (1+\varepsilon)\mu] = \Pr\left[\mathrm{e}^{tX} \geq \mathrm{e}^{t(1+\varepsilon)\mu}\right] \leq \frac{\mathrm{E}[e^{tX}]}{\mathrm{e}^{t(1+\varepsilon)\mu}} \qquad (1)$$

各X_iは$\{0,1\}$確率変数より，$p_i := \Pr[X_i = 1]$とすると,

$$\begin{aligned}\mathrm{E}[\mathrm{e}^{tX_i}] &= p_i\mathrm{e}^t + (1-p_i)\\ &= 1 + p_i(\mathrm{e}^t - 1)\\ &\leq \exp(p_i(\mathrm{e}^t - 1))\end{aligned}$$

Maclaurin展開

独立性より

$$\begin{aligned}\mathrm{E}[\mathrm{e}^{tX}] &= \prod_{i=1}^{n}\mathrm{E}[\mathrm{e}^{tX_i}] \leq \prod_{i=1}^{n}\exp p_i(\mathrm{e}^t - 1)\\ &= \exp\left((\mathrm{e}^{t-1})\sum_{i=1}^{n}p_i\right) = \exp((e^t - 1)\mu)\end{aligned}$$

$t = \ln(1+\varepsilon)$とすると,

$$(1) \leq \frac{\exp\bigl((e^t - 1)\mu\bigr)}{\mathrm{e}^{t(1+\varepsilon)\mu}} = \frac{\mathrm{e}^{\varepsilon\mu}}{(1+\varepsilon)^{(1+\varepsilon)\mu}} = \left(\frac{\mathrm{e}^{\varepsilon}}{(1+\varepsilon)^{(1+\varepsilon)}}\right)^{\mu}$$

Hoeffding

almost surely bounded
$\exists a, b \in \mathbb{R},\ \Pr\big[X \in [a, b]\big] = 1$
に拡張できる.

定理 [Hoeffding]

$X_1, X_2, \dots, X_n$は相互に独立で$0 \le X_n \le 1$をとる確率変数とする.

$\bar{X} := \frac{X_1 + \cdots + X_n}{n}$ とすると, $0 < t < 1 - \mathrm{E}[\bar{X}]$について,

$$\Pr[\bar{X} - E[\bar{X}] > t] < \exp(-2nt^2)$$

が成り立つ.

確率不等式

確率不等式

Markovの不等式

Chebyshevの不等式

Chernoffの不等式({0,1}確率変数)

Hoefdingの不等式(有限値)

参考: 二項分布，幾何分布の裾

Chvatal

基礎技術 4. Median Trick

全然別の問題: tail boundの世界

問題 5. (成功確率の増幅)

株価予測の専門家が101人居て，どの専門家も確率0.6以上で1年後の株価を誤差1%以内で的中させる．101人の意見からどう株価を予測する？

中央値 (median)を取る

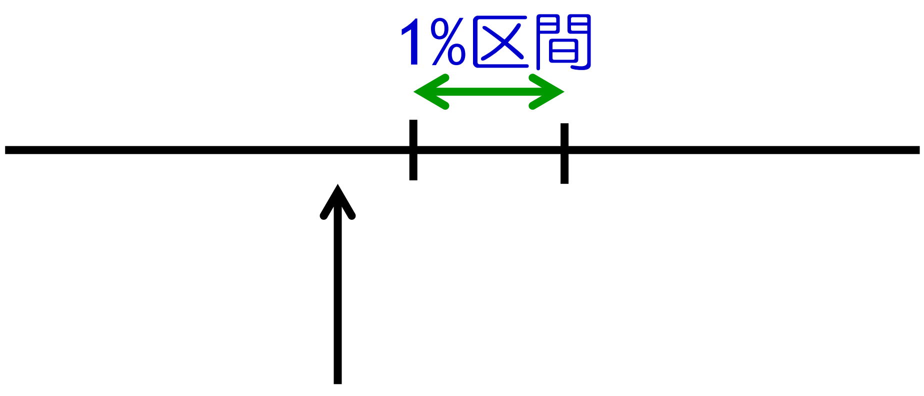
全然別の問題: tail boundの世界

問題 5. (成功確率の増幅)

株価予測の専門家が101人居て，どの専門家も確率0.6以上で1年後の株価を誤差1%以内で的中させる．101人の意見からどう株価を予測する？

中央値 (median)を取る

中央値が誤差1%より小さい⇒51人以上が小さめに外している



全然別の問題: tail boundの世界

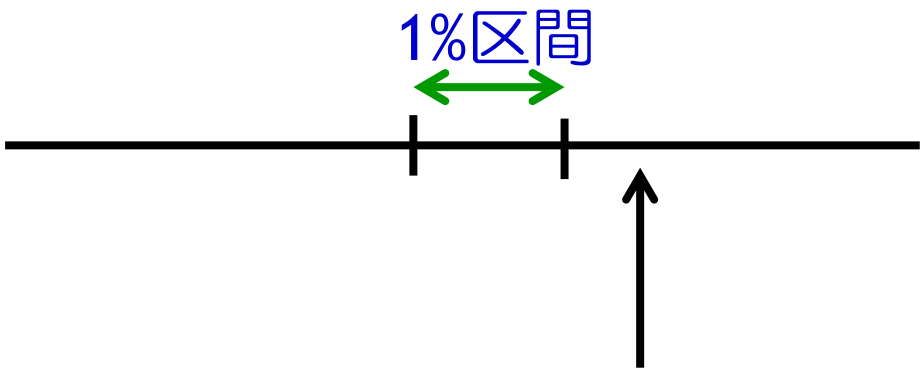
問題 5. (成功確率の増幅)

株価予測の専門家が101人居て，どの専門家も確率0.6以上で1年後の株価を誤差1%以内で的中させる．101人の意見からどう株価を予測する？

中央値 (median)を取る

中央値が誤差1%より小さい⇒51人以上が小さめに外している

中央値が誤差1%より大きい⇒51人以上が大きめに外している



全然別の問題: tail boundの世界

問題 5. (成功確率の増幅)

株価予測の専門家が101人居て，どの専門家も確率0.6以上で1年後の株価を誤差1%以内で的中させる．101人の意見からどう株価を予測する？

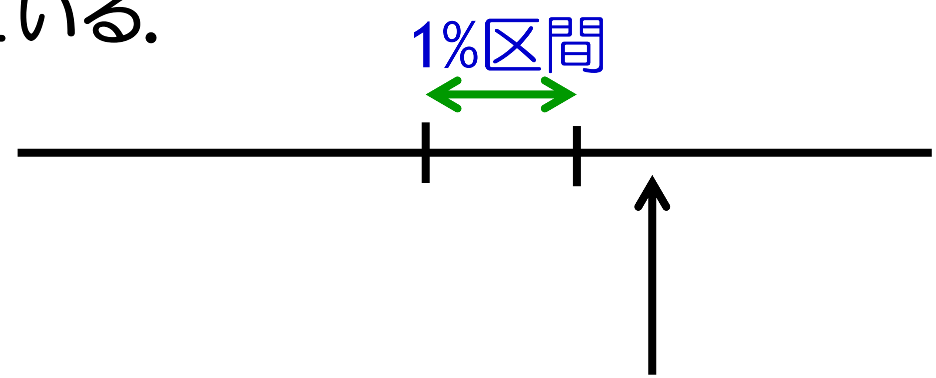
中央値 (median)を取る

中央値が誤差1%より小さい⇒51人以上が小さめに外している

中央値が誤差1%より大きい⇒51人以上が大きめに外している

いずれにせよ，**過半数が外している.**

$$\sum_{k=0}^{50} \binom{101}{k} 0.6^k 0.4^{n-k} = ???$$



Chernoff bound の演習

演習課題

Chernoff boundの系を使って，次の確率の上界を求めよ．

$$\sum_{k=0}^{50} \binom{101}{k} 0.6^k 0.4^{n-k} \leq ???$$

Chernoff boundは
配布資料 p.58 参照

系

$0 < \varepsilon \leq 1$ の場合，

$$\Pr[X \geq (1+\varepsilon)\mu] \leq \exp\left(-\frac{\varepsilon^2 \mu}{3}\right)$$

$$\Pr[X \leq (1-\varepsilon)\mu] \leq \exp\left(-\frac{\varepsilon^2 \mu}{2}\right)$$

乱択の技法: 今日の話題

基礎技術1. $\left(1-\frac{1}{n}\right)^n \simeq \frac{1}{\mathrm{e}}$

基礎技術2. logの出てくる仕組み

基礎技術3. サンプリングのサイズはなぜ $\left(\frac{1}{\varepsilon}\right)^2$?

- ➢ レアイベント(起こりにくい事象)の確率
  ～ 確率不等式の世界へようこそ ～

基礎技術4. Median Trick

応用. 高度モンテカルロ法

- ➢ サンプリングと積分

今後の課題. BPL=PSPACE?

私のよく使う技法

演習問題

問題

人気クイズ番組「クイズ超難問！」では，２択の問題が20問出される．問題は超難問ばかりで，たとえば

「強いのはどっち？A. 赤鬼，B.青鬼」

といった，勘で答えるより他にない問題ばかりである．20問連続正解すると豪華賞品（100万USD）がもらえるが，現在までに67人が挑戦して，全問正解者は居ない．全問正解が出るのは，いつごろ（何人目ごろ）か？

A. 約千人目ごろ

B. 約百万人目ごろ

練習課題 2

$2^{24} \simeq 1.28 \times 10^8$人を超えても達成者のいない確率は？

練習課題 3

達成者が100人になった．およそ何人チャレンジしたか？

Chernoff boundは

配布資料 p.58 参照

演習問題まとめ

演習課題

Chernoff boundの系を使って，次の確率の上界を求めよ．

$$\sum_{k=0}^{50} \binom{101}{k} 0.6^k 0.4^{n-k} \leq ???$$

練習課題 2

$2^{24} \simeq 1.28 \times 10^8$人を超えても達成者のいない確率は？

練習課題 3

達成者が100人になった．およそ何人チャレンジしたか？

何か問題ちょうだい

Thanks

平原さん，渡辺先生，山口さん

~~(たぶん)未解決問題~~

~~BPL=PSPACE?~~

Fact

$\mathrm{BPL} \subseteq \mathrm{P}$

予想

$\mathrm{L} = \mathrm{BPL}$

方針

~~PSPACE完全な問題(例えば遷移問題)は~~

~~乱択 $\log(n)$ spaceで解ける、気がする（解けたらいいな）．~~

The end

Thank you for the attention.

演習問題 1

Chernoff

$$\Pr[X \le (1-\varepsilon)\mu] \le \exp\left(-\frac{\varepsilon^2\mu}{2}\right)$$

$X =$ 当たった人の数

$\mu = 0.6 \times 101,$

$\varepsilon = 0.1749 \qquad (50 < (1-0.1749) \times 60.6)$

$$\sum_{k=0}^{50} \binom{101}{k} 0.6^k 0.4^{n-k} \le \Pr[X \le (1-0.1749) \times 60.6]$$

$$\le \exp\left(-\frac{0.1749^2 \times 60.6}{2}\right)$$

$$\le \exp\left(-\frac{1.8}{2}\right)$$

$$= \frac{1}{\mathrm{e}^{0.9}}$$

演習 3

練習課題 3

達成者が100人になった．およそ何人チャレンジしたか？

- ✓ だいたい $100 \times 2^{20} = 10^8$人 (期待値)
- ✓ 挑戦者の数が3×10^8を超えても達成者が100人に満たない確率

$$\Pr[X < 100] \leq \Pr\left[X < \left(1 - \frac{2}{3}\right) \times 300\right] \leq \exp\left(-\frac{\left(\frac{2}{3}\right)^2 \times 300}{2}\right) \leq \frac{1}{\mathrm{e}^{200}}$$

- ✓ 挑戦者の数が0.6×10^8未満で達成者が100人に超える確率

$$\Pr[X > 100] = \Pr\left[X > \left(1 + \frac{2}{3}\right) \times 60\right] \leq \exp\left(-\frac{\left(\frac{2}{3}\right)^2 \times 60}{3}\right) \leq \frac{1}{\mathrm{e}^{25}}$$

系

$0 < \varepsilon \leq 1$の場合，

$$\Pr[X \geq (1 + \varepsilon)\mu] \leq \exp\left(-\frac{\varepsilon^2 \mu}{3}\right)$$

$$\Pr[X \leq (1 - \varepsilon)\mu] \leq \exp\left(-\frac{\varepsilon^2 \mu}{2}\right)$$