

3. Self-Reduction

来嶋 秀治

滋賀大学 データサイエンス学部

計算機科学のMCMC法

FPRAS

“Valiant計画” (since 1979)

#P困難な $A \in \mathbb{R}$ に対して, 近似解 $Z \in \mathbb{R}$ を精度

$$\Pr[(1 - \epsilon)A \leq Z \leq (1 + \epsilon)A] \geq 1 - \delta$$

で多項式時間($\text{poly}(\text{input}, \epsilon^{-1}, \log \delta^{-1})$)で求められるか?

Easy case

単純モンテカルロの場合

準備：単純モンテカルロの精度保証

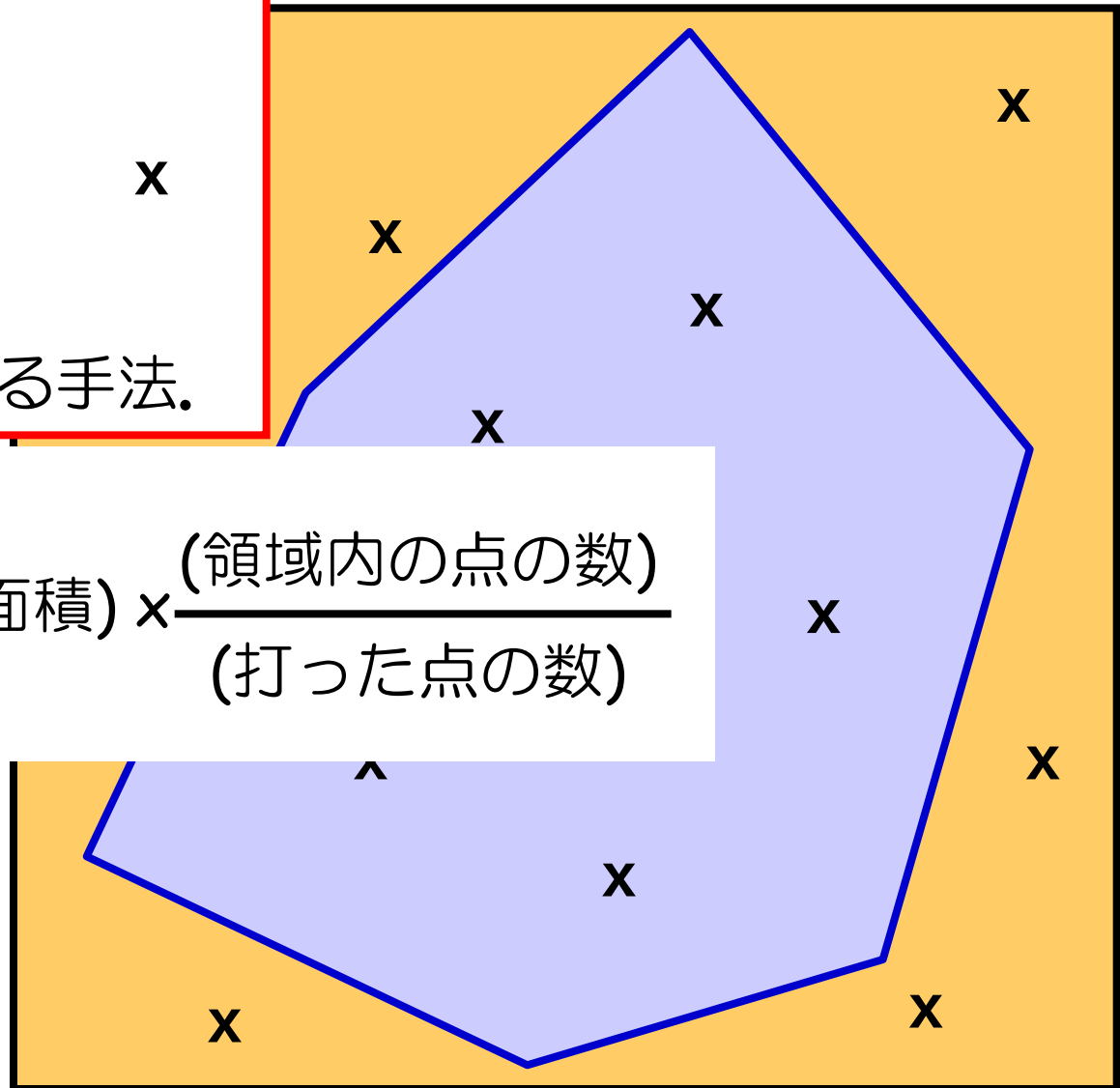
モンテカルロ法

乱数を使用して、

- ・ 比
- ・ 平均

などを近似的に計算する手法。

$$\left(\text{多角形の面積} \right) = \left(\text{正方形の面積} \right) \times \frac{\text{(領域内の点の数)}}{\text{(打った点の数)}}$$



Chernoff bound

$$(\text{多角形の面積}) \doteq (\text{正方形の面積}) \times \frac{(\text{領域内の点の数})}{(\text{打った点の数})}$$

T
 α
 X
 M

定理. (cf. 確率収束)

任意の ϵ ($0 < \epsilon < 1$), δ ($0 < \delta < 1/2$) に対して,
モンテカルロ法の標本数 $M \geq 3p^{-1}\epsilon^{-2} \ln \delta^{-1}$ とすると,

真値 T の近似値 $Z = \alpha \frac{X}{M}$ は

$$\Pr[(1 - \epsilon)T \leq Z_n \leq (1 + \epsilon)T] \geq 1 - \delta$$

を満たす.

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,
 $X := X_1 + \dots + X_n$ の期待値を μ とすると,

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

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

系

$0 < \varepsilon \leq 1$ の場合,

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

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

Hoeffding ineq.

$$\left(\text{多角形の面積} \right) \doteq \left(\text{正方形の面積} \right) \times \frac{\text{(領域内の点の数)}}{\text{(打った点の数)}}$$

T
 α
 X

証明. $p = \frac{T}{\alpha}$ とすると $p \doteq \frac{X}{M}$.

$$\begin{aligned} & \Pr[(1 - \epsilon)T \leq Z_n \leq (1 + \epsilon)T] \\ &= \Pr\left[(1 - \epsilon)p \leq \frac{X}{M} \leq (1 + \epsilon)p\right] \\ &= \sum_{k=\lfloor Mp(1-\epsilon) \rfloor}^{\lfloor Mp(1+\epsilon) \rfloor} \binom{M}{k} p^k (1-p)^{M-k} \quad (*) \end{aligned}$$

Chernoffの不等式より,

$$(*) \geq 1 - 2 \exp\left(-\frac{Mp\epsilon^2}{3}\right) \quad (**)$$

$M \geq 3p^{-1}\epsilon^{-2} \ln \delta^{-1}$ とすると,

$$(**) \geq 1 - 2\delta^2 \geq 1 - \delta \quad \left(0 < \delta < \frac{1}{2} \text{ のとき} \right)$$

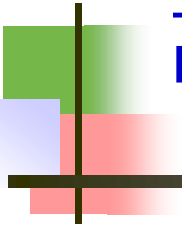
系

$0 < \epsilon \leq 1$ の場合,

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

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





高次元凸体の体積計算

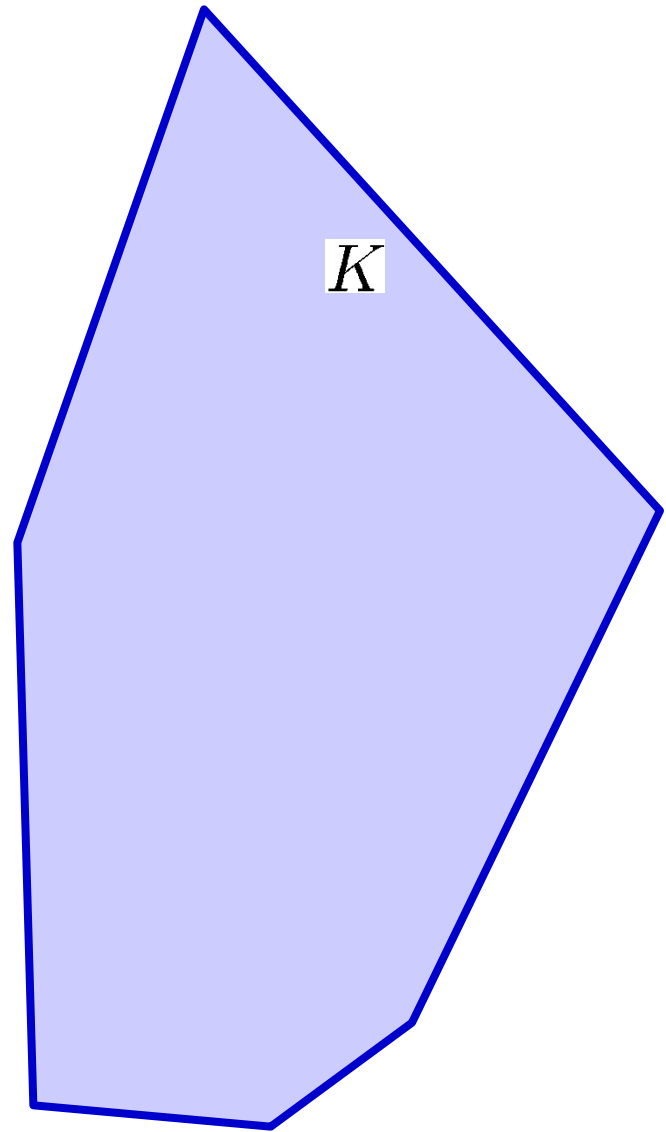
[Dyer, Frieze, Kannan 1991]

再帰的なモンテカルロ法

問題: 高次元凸体の体積計算

K : n 次元凸体

体積 $V(K)$ を計算



仮定: ふっくらとした凸体

K : n 次元凸体

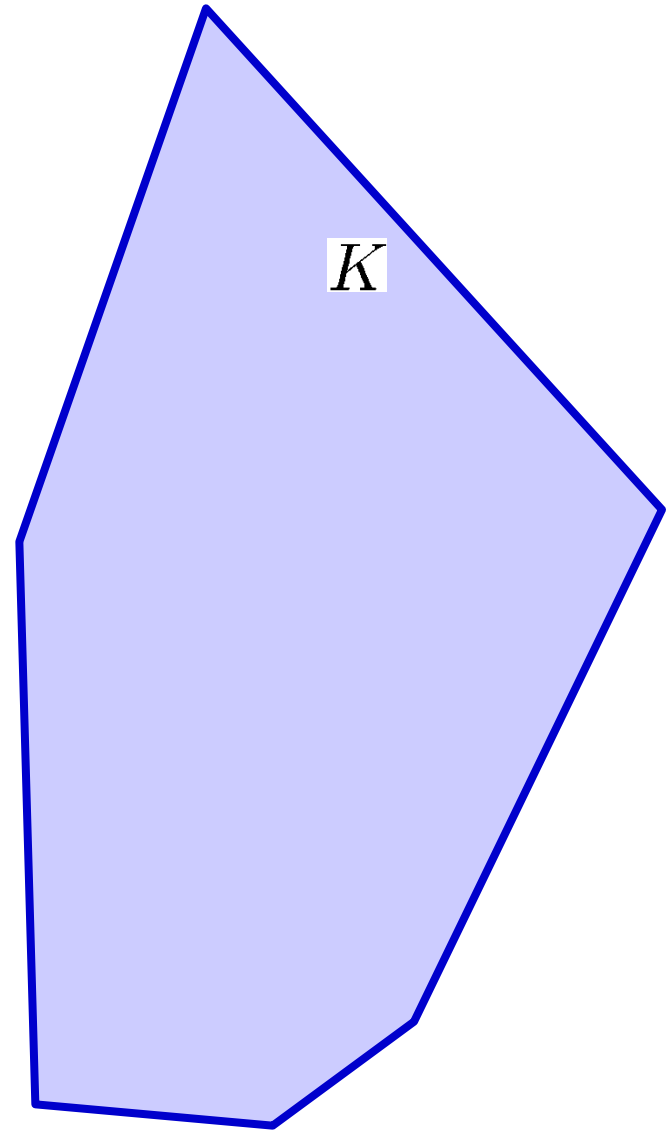
体積 $V(K)$ を計算

仮定

K はふっくらとしている

$$B_n(\mathbf{0}, R) \subseteq K \subseteq B_n(\mathbf{0}, R)$$

$$[R = O(n^c)]$$



仮定: ふっくらとした凸体

K : n 次元凸体

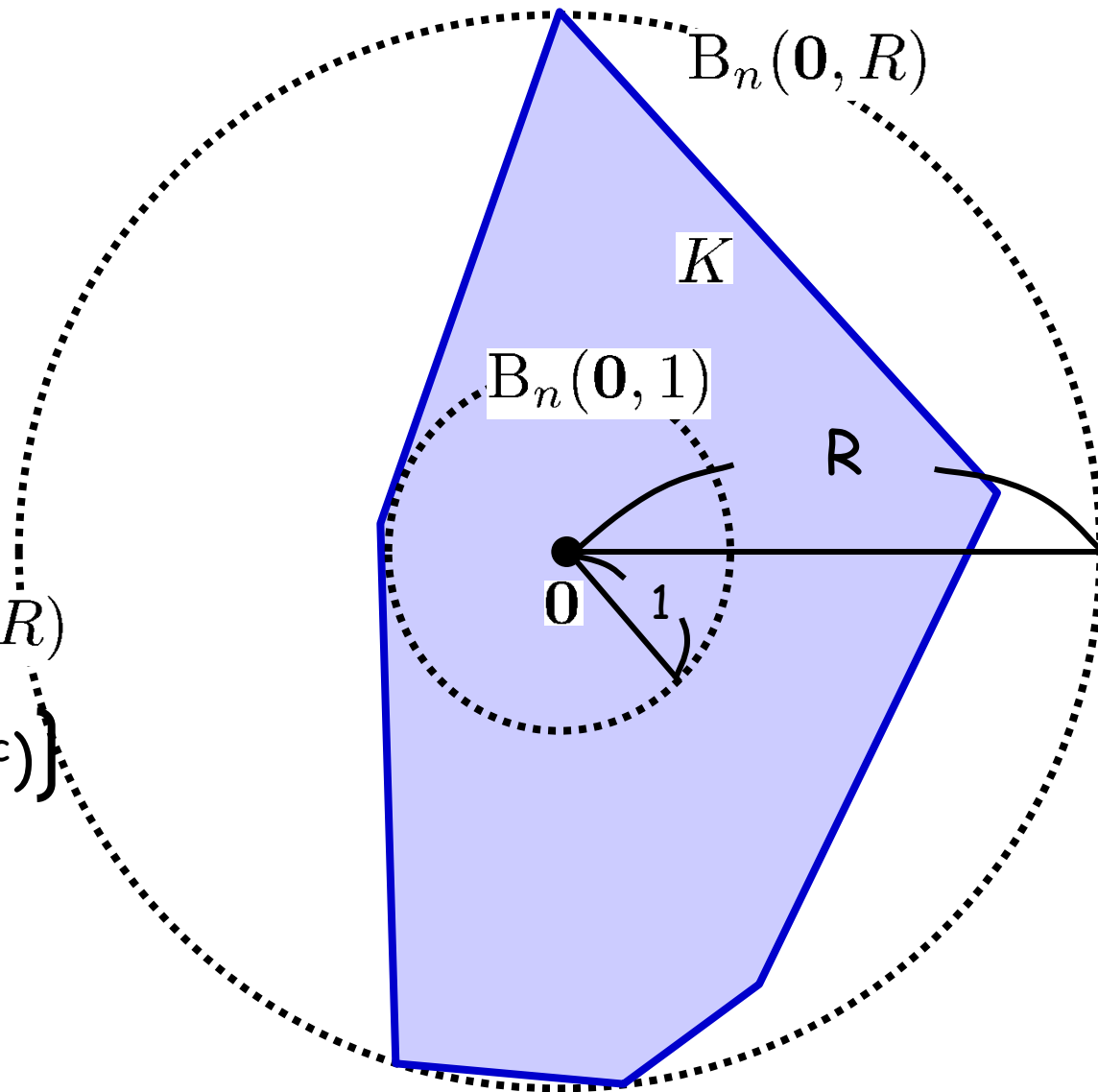
体積 $V(K)$ を計算

仮定

K はふっくらとしている

$$B_n(\mathbf{0}, R) \subseteq K \subseteq B_n(\mathbf{0}, R)$$

$$\left[R = O(n^c) \right]$$



K上の一様ランダム生成 – Ball walk

Bal-walk

X : 現在の点

X' : 次の時刻の点

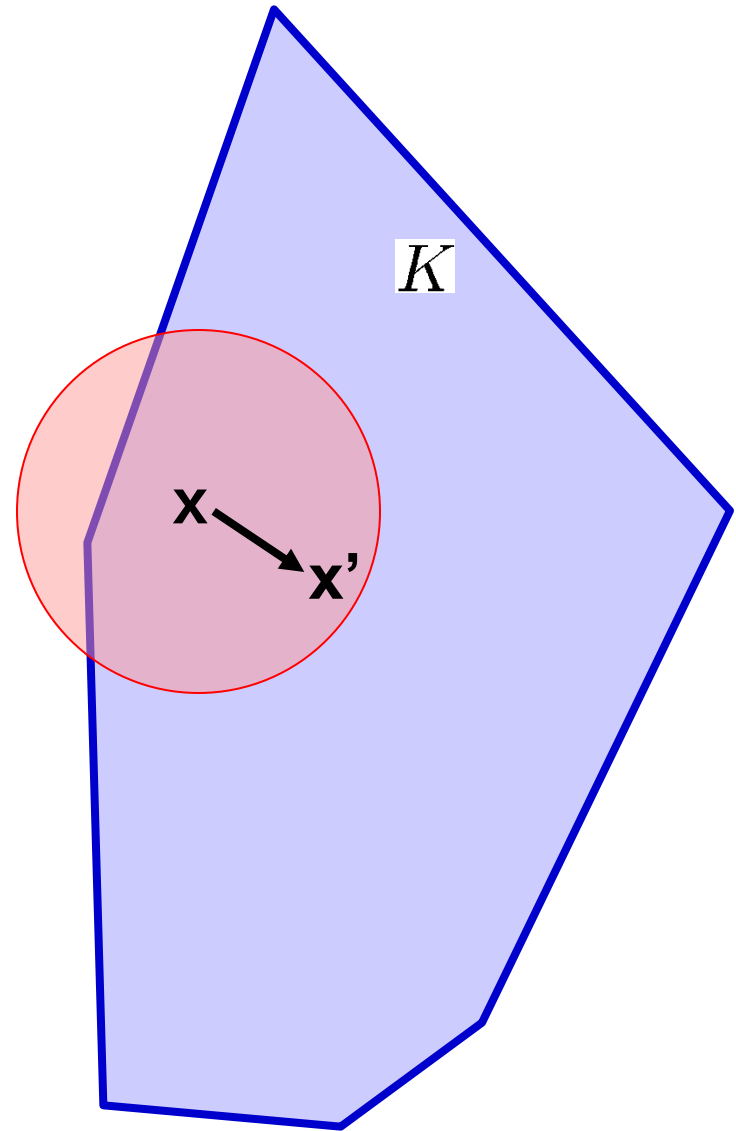
Step 1.

$Y : B_n(X, a)$ 中の
一様ランダムな点を選ぶ.

Step 2.

$$X' = \begin{cases} Y & (Y \in K) \\ X & (Y \notin K) \end{cases}$$

定常分布はK上の一様分布

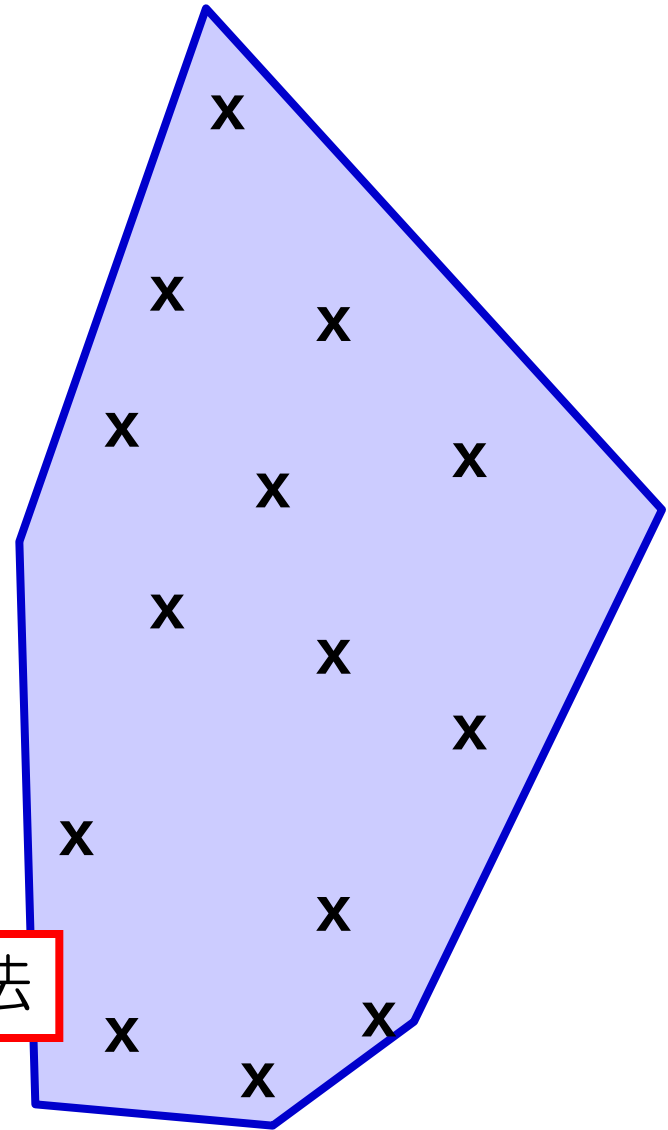


K上の一様ランダム生成

(いっくらとした凸体に対して)
Ball-walkを用いて,
一様ランダムサンプリング可能

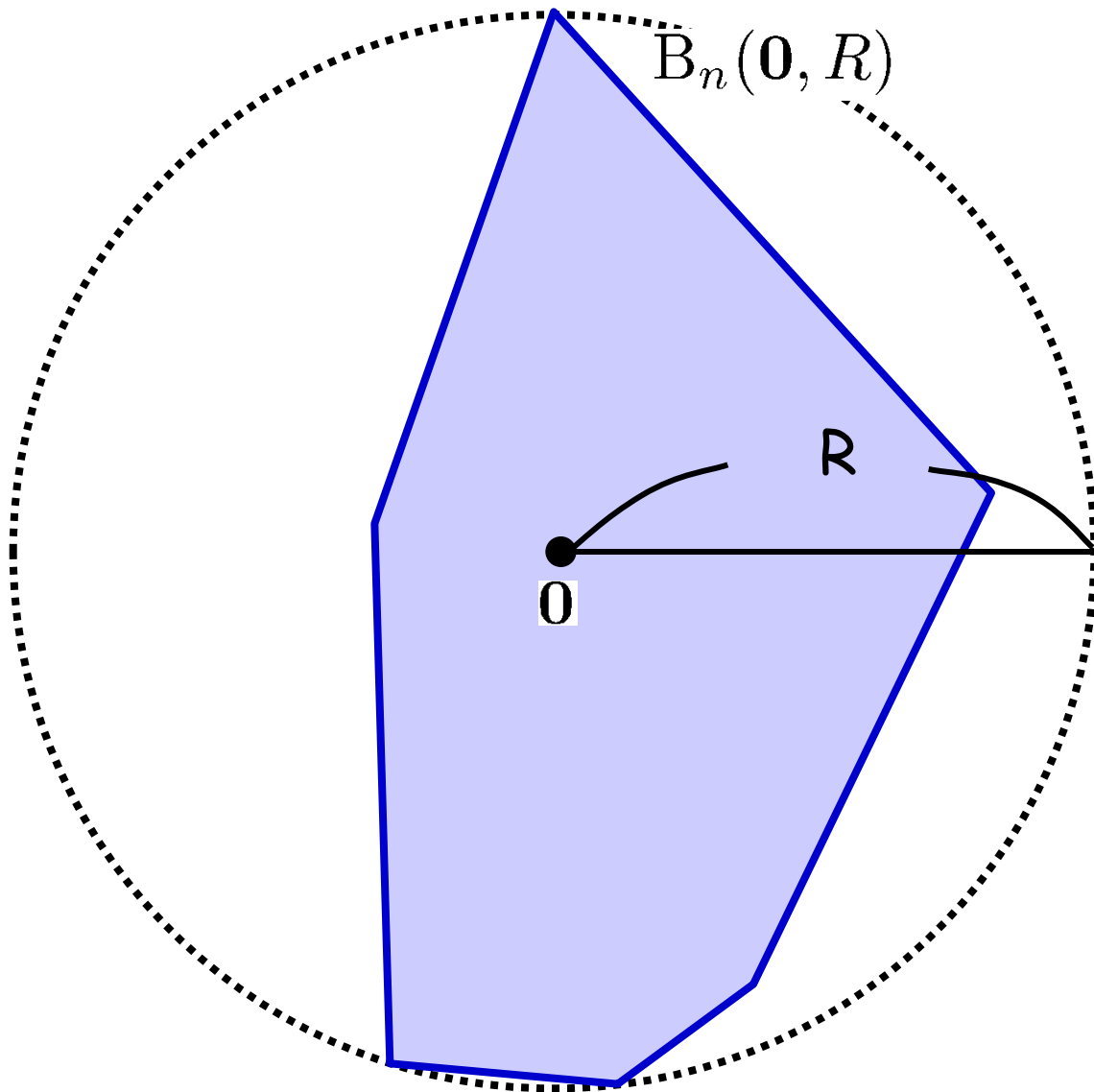
どうやって,
体積を計算するのか?

再帰的アルゴリズム + モンテカルロ法



再帰のアイデア

K : n 次元凸体
体積 $V(K)$ を計算

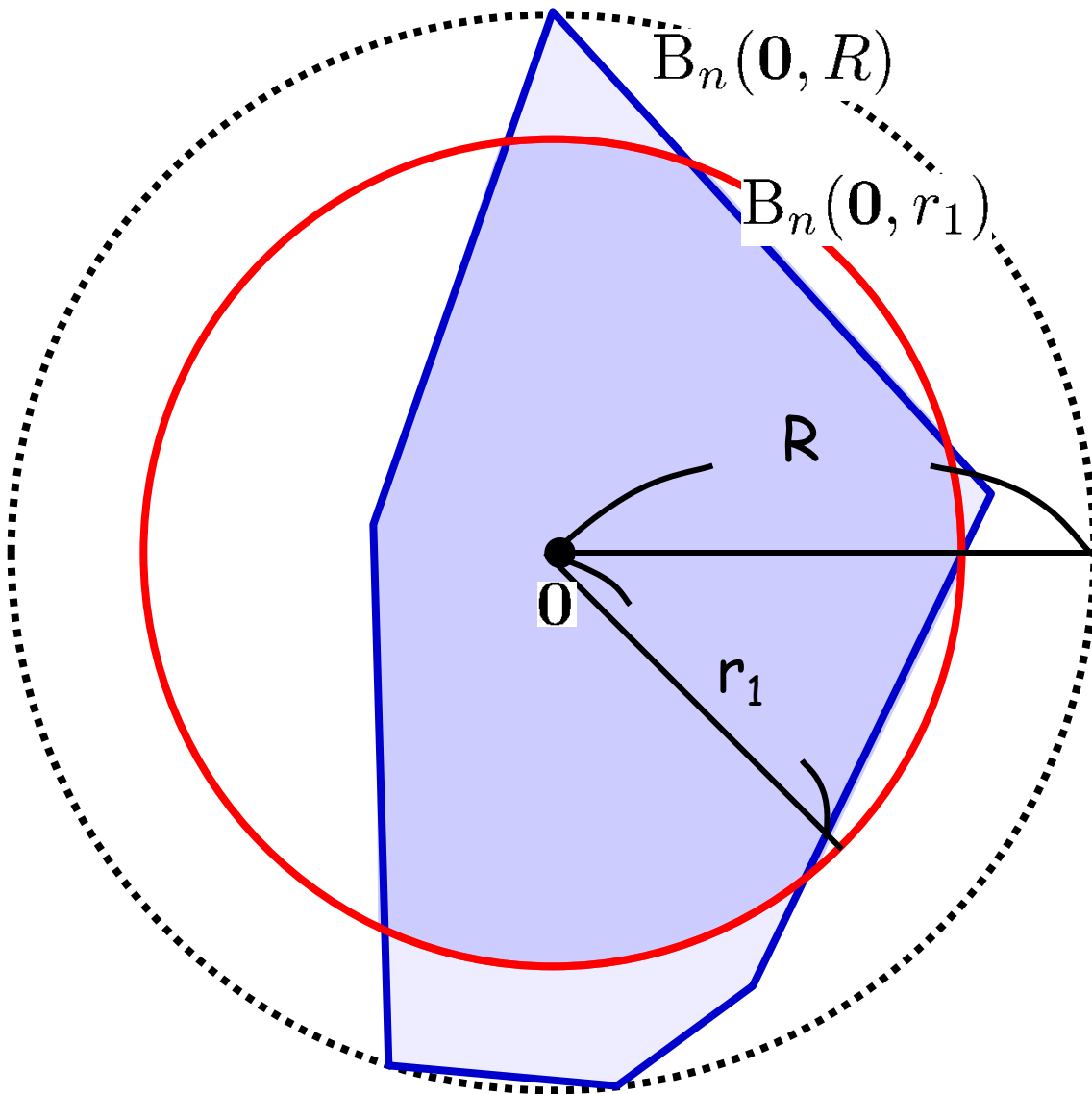


再帰のアイデア

K : n 次元凸体
体積 $V(K)$ を計算

$$K_1 \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_1)$$

$$V(K) = \frac{V(K)}{V(K_1)} \cdot V(K_1)$$



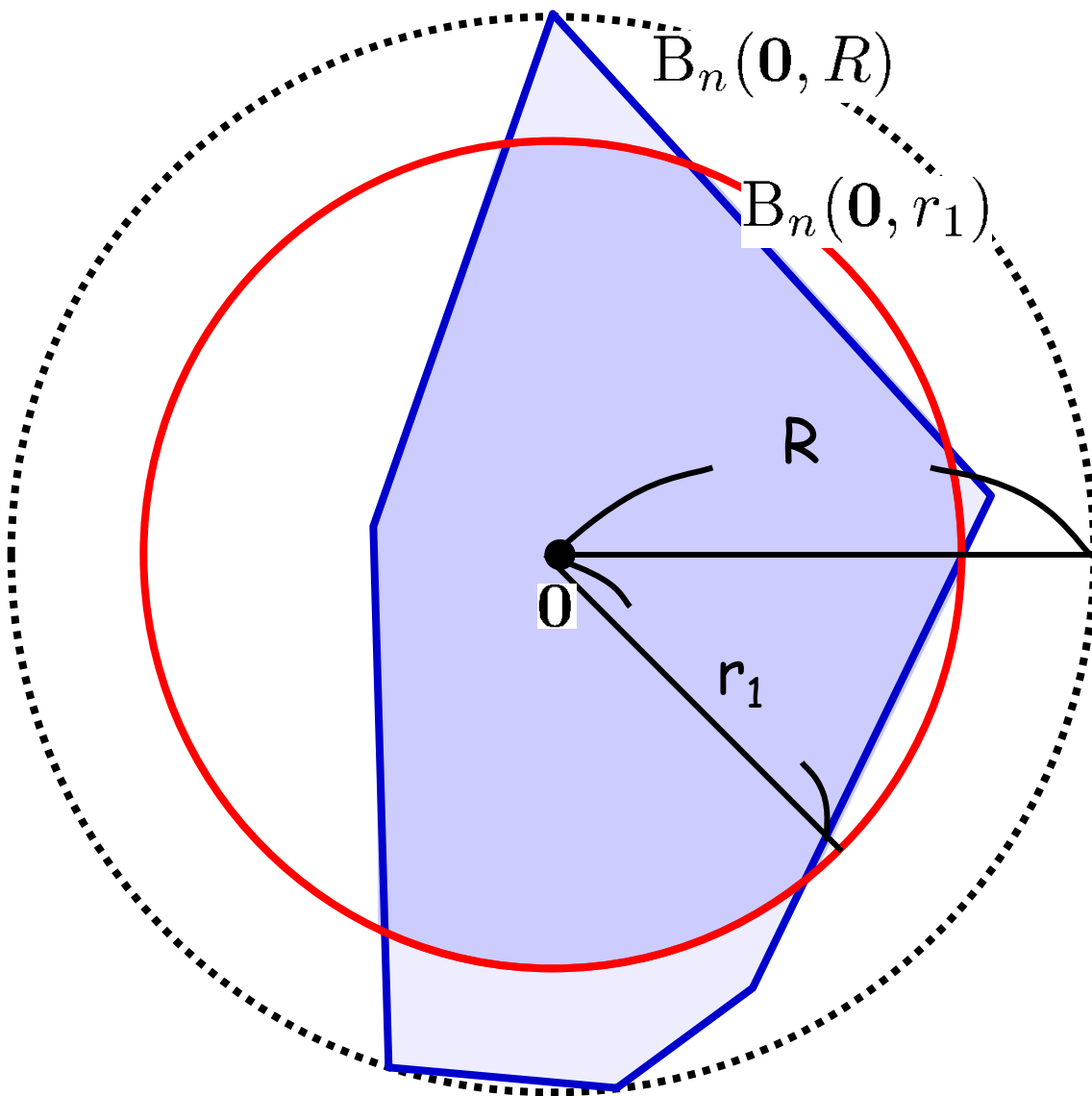
再帰のアイデア

K : n 次元凸体
体積 $V(K)$ を計算

$$K_1 \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_1)$$

$$V(K) = \frac{V(K)}{V(K_1)} \cdot V(K_1)$$

↑
モンテカルロ法で計算



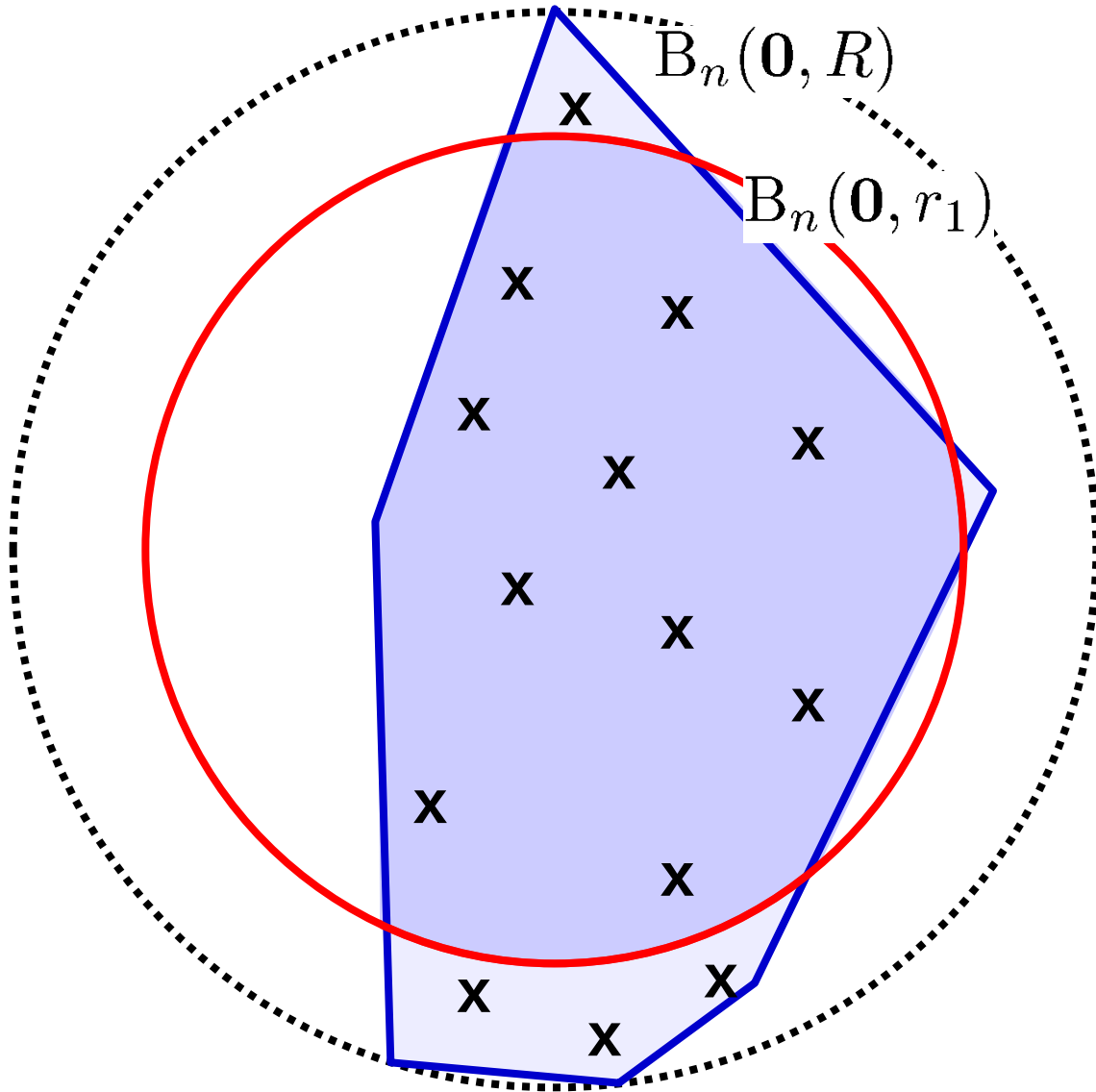
再帰のアイデア

K : n 次元凸体
体積 $V(K)$ を計算

$$K_1 \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_1)$$

$$V(K) = \frac{V(K)}{V(K_1)} \cdot V(K_1)$$

モンテカルロ法で計算



再帰のアイデア

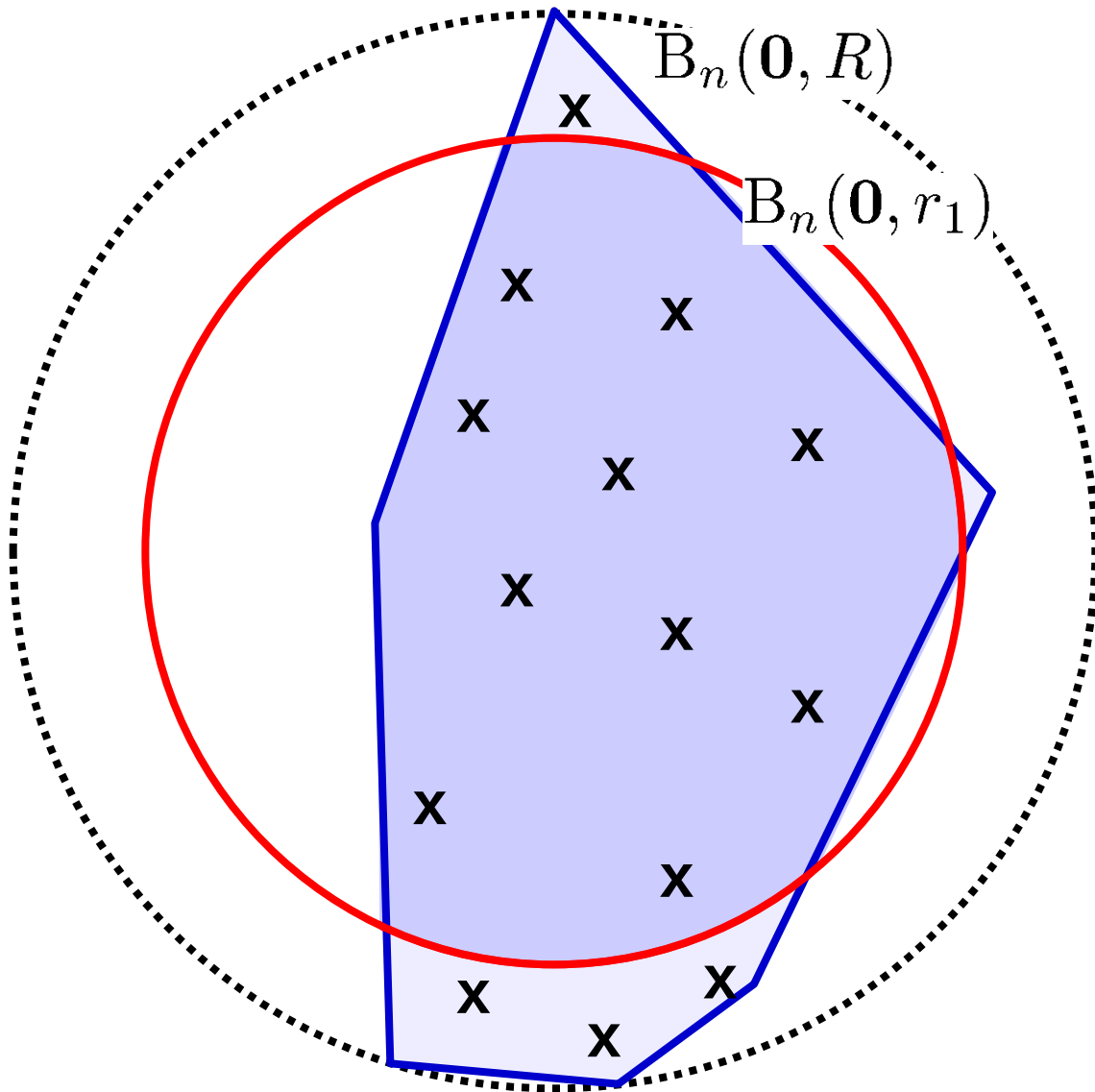
K : n 次元凸体
体積 $V(K)$ を計算

$$K_1 \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_1)$$

$$V(K) = \frac{V(K)}{V(K_1)} \cdot V(K_1)$$

↑
モンテカルロ法で計算

K_1 : n 次元凸体
体積 $V(K_1)$ を計算



再帰のアイデア

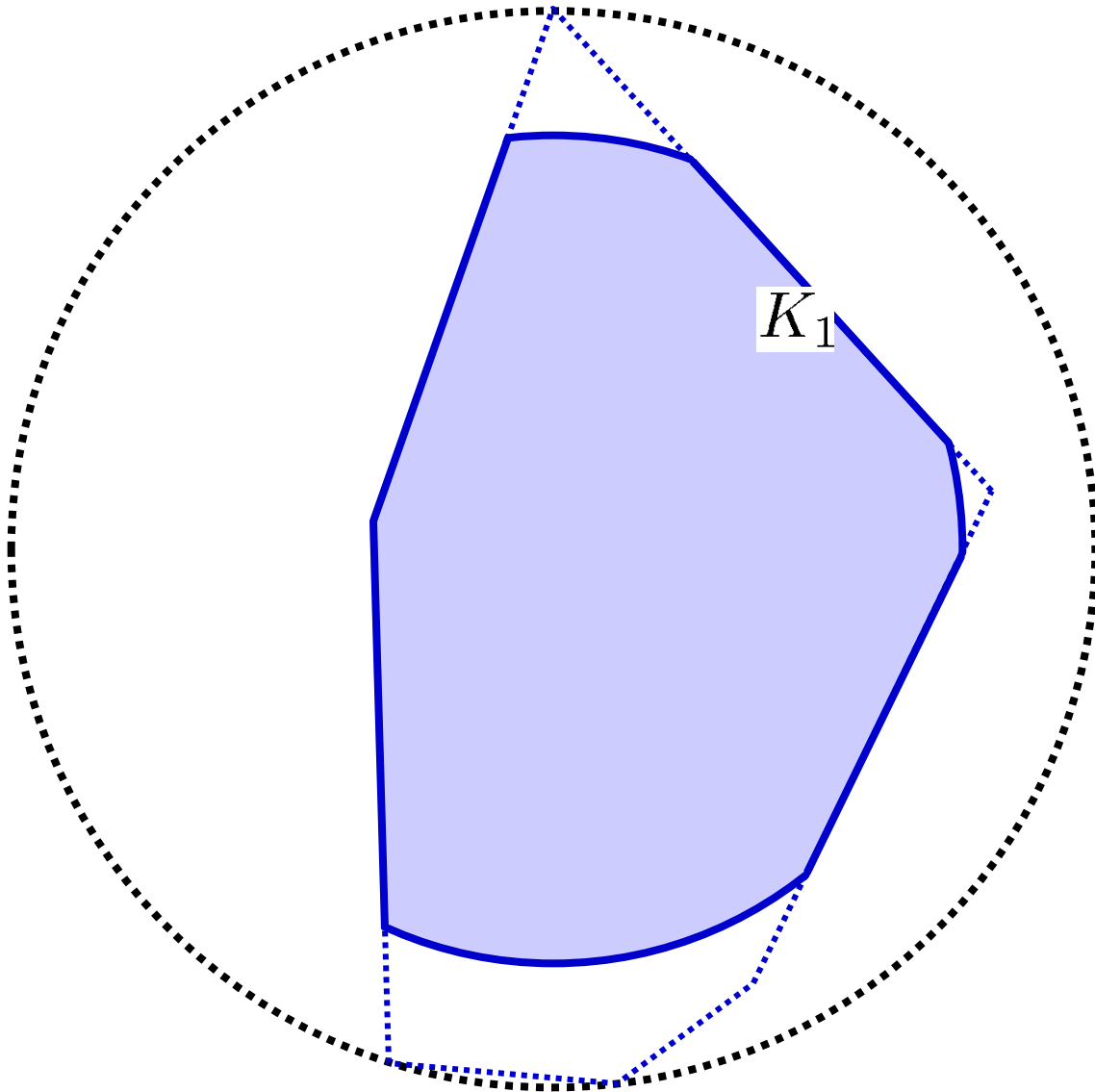
K : n 次元凸体
体積 $V(K)$ を計算

$$K_1 \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_1)$$

$$V(K) = \frac{V(K)}{V(K_1)} \cdot V(K_1)$$

↑
モンテカルロ法で計算

K_1 : n 次元凸体
体積 $V(K_1)$ を計算



再帰のアイデア

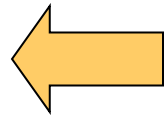
K : n 次元凸体
体積 $V(K)$ を計算

$$K_1 \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_1)$$

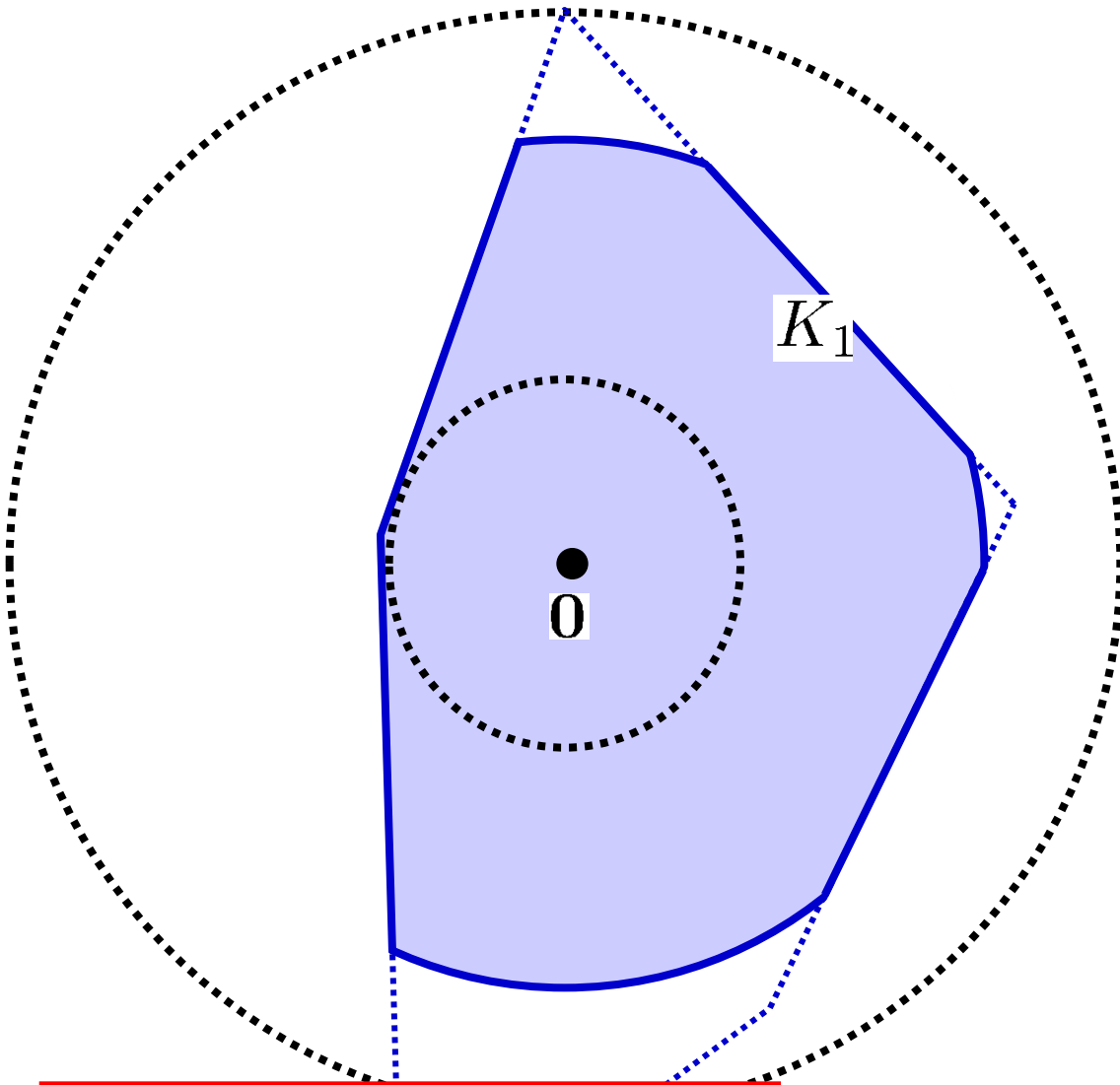
$$V(K) = \frac{V(K)}{V(K_1)} \cdot V(K_1)$$

モンテカルロ法で計算

K_1 : n 次元凸体
体積 $V(K_1)$ を計算

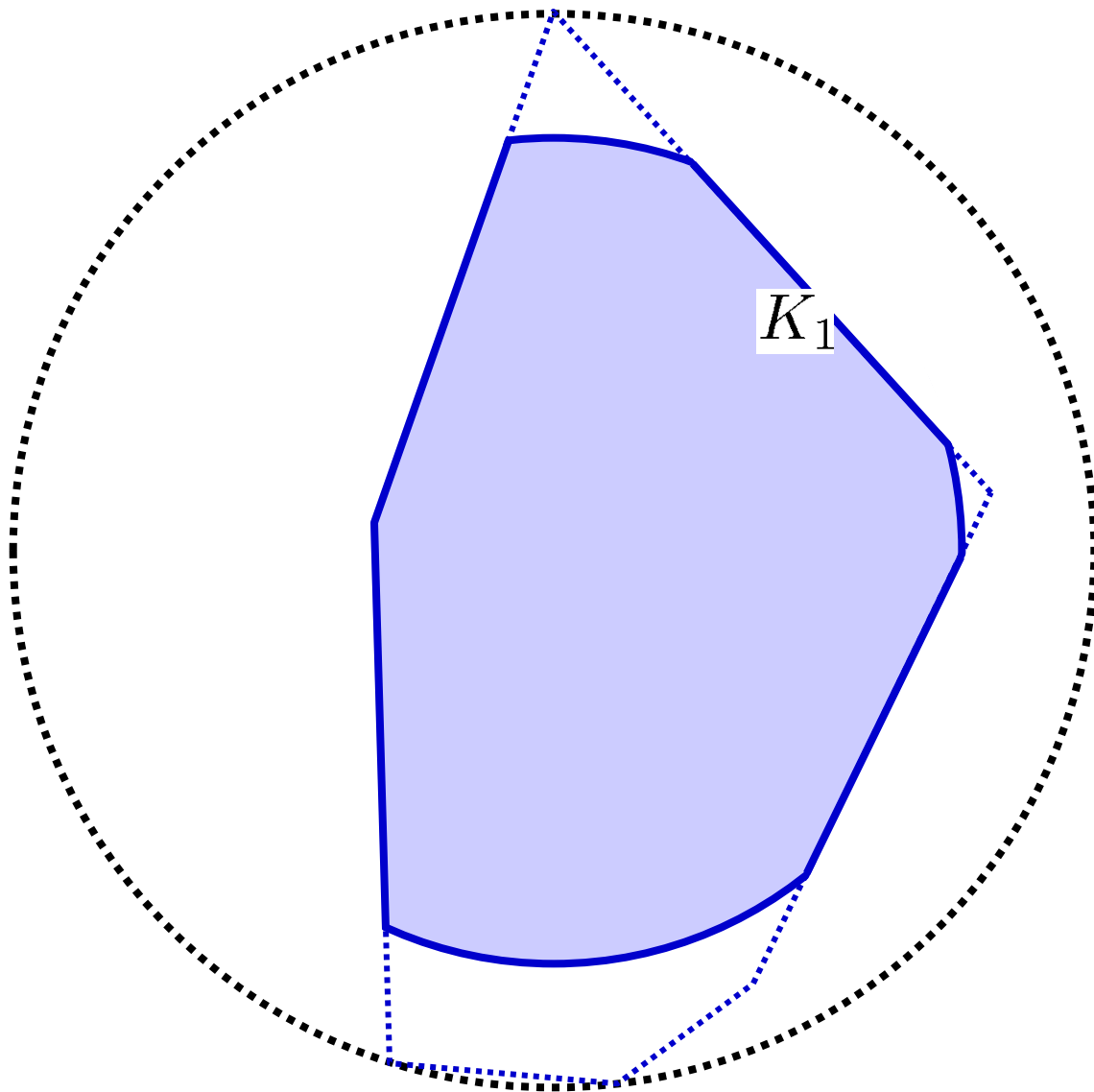


K_1 はふっくらとしている



再帰のアイデア

K_1 : n 次元凸体
体積 $V(K_1)$ を計算

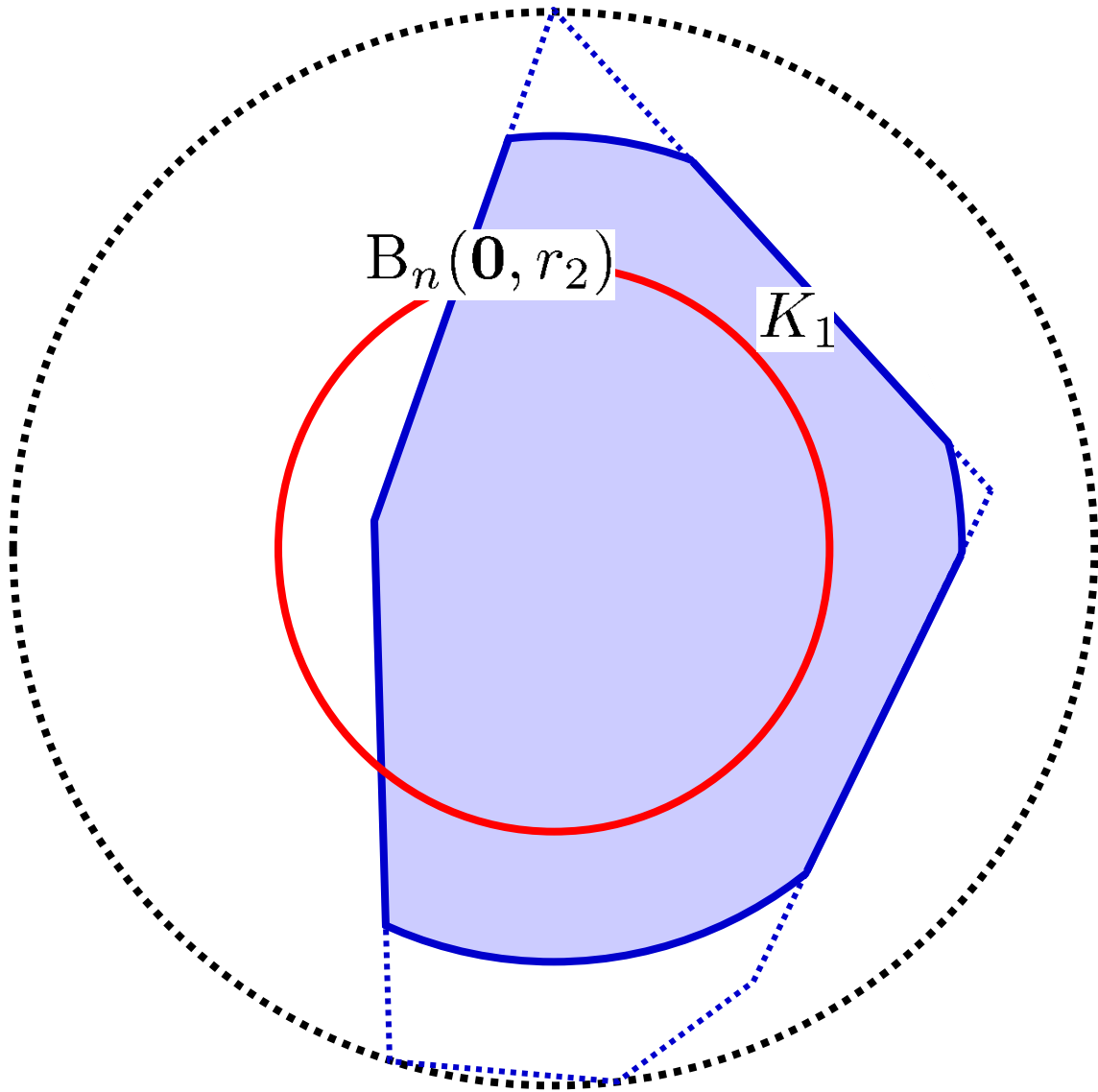


再帰のアイデア

K_1 : n 次元凸体
体積 $V(K_1)$ を計算

$$K_2 \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_2)$$

$$V(K_1) = \frac{V(K_1)}{V(K_2)} \cdot V(K_2)$$



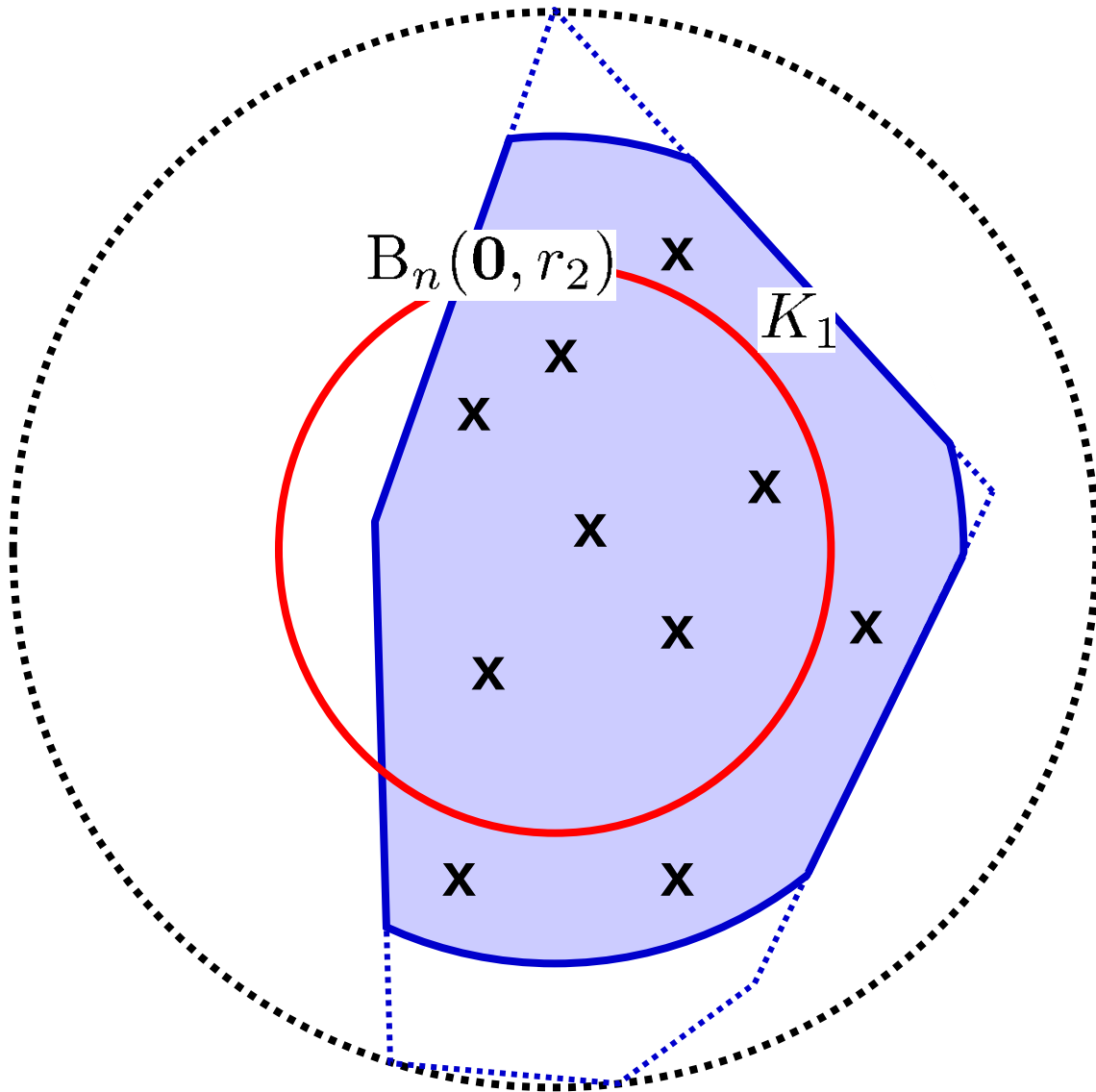
再帰のアイデア

K_1 : n 次元凸体
体積 $V(K_1)$ を計算

$$K_2 \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_2)$$

$$V(K_1) = \frac{V(K_1)}{V(K_2)} \cdot V(K_2)$$

モンテカルロ法で計算



再帰のアイデア

K_1 : n 次元凸体
体積 $V(K_1)$ を計算

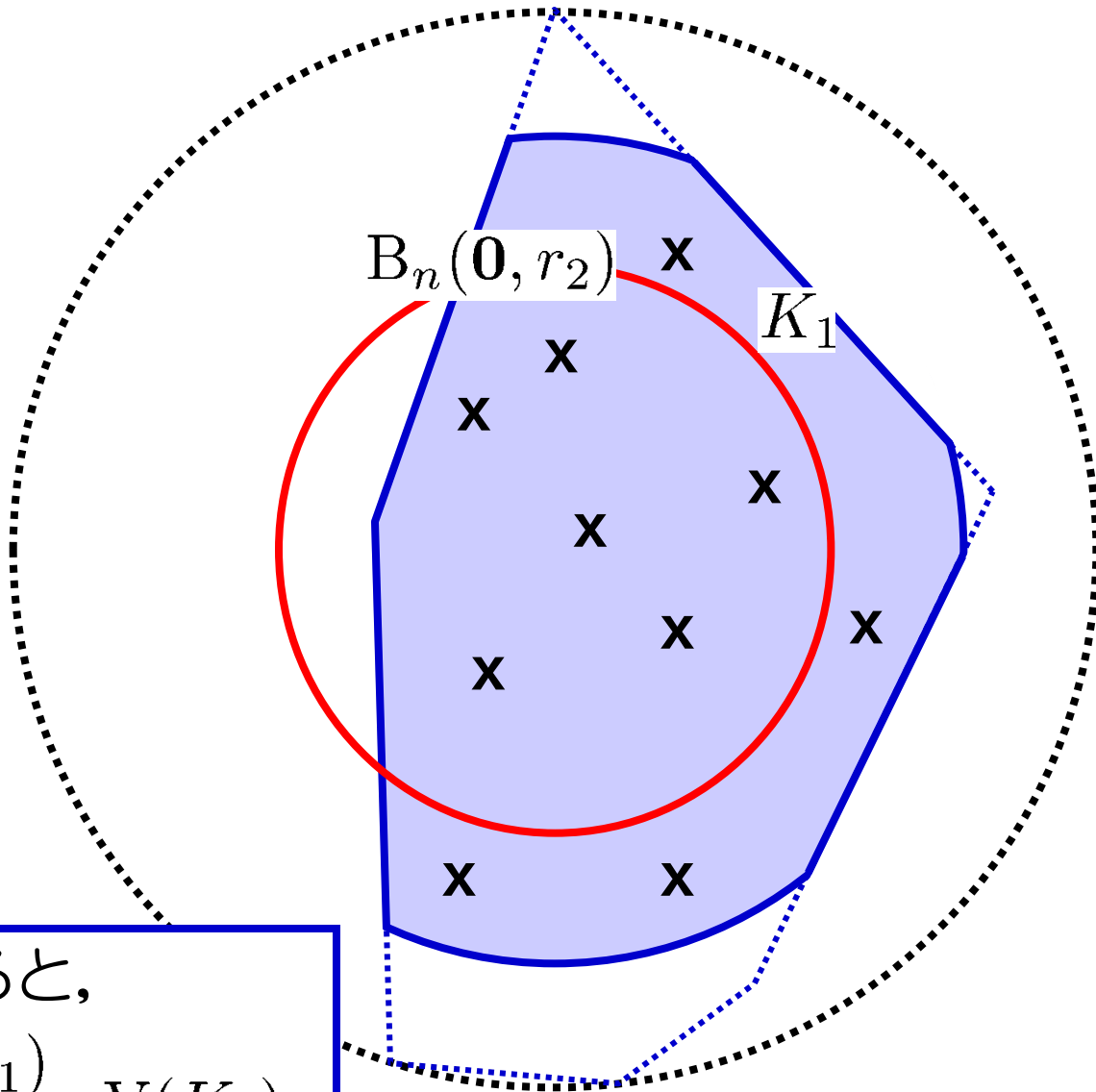
$$K_2 \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_2)$$

$$V(K_1) = \frac{V(K_1)}{V(K_2)} \cdot V(K_2)$$

モンテカルロ法で計算

先程の再帰式に代入すると,

$$V(K) = \frac{V(K)}{V(K_1)} \cdot \frac{V(K_1)}{V(K_2)} \cdot V(K_2)$$

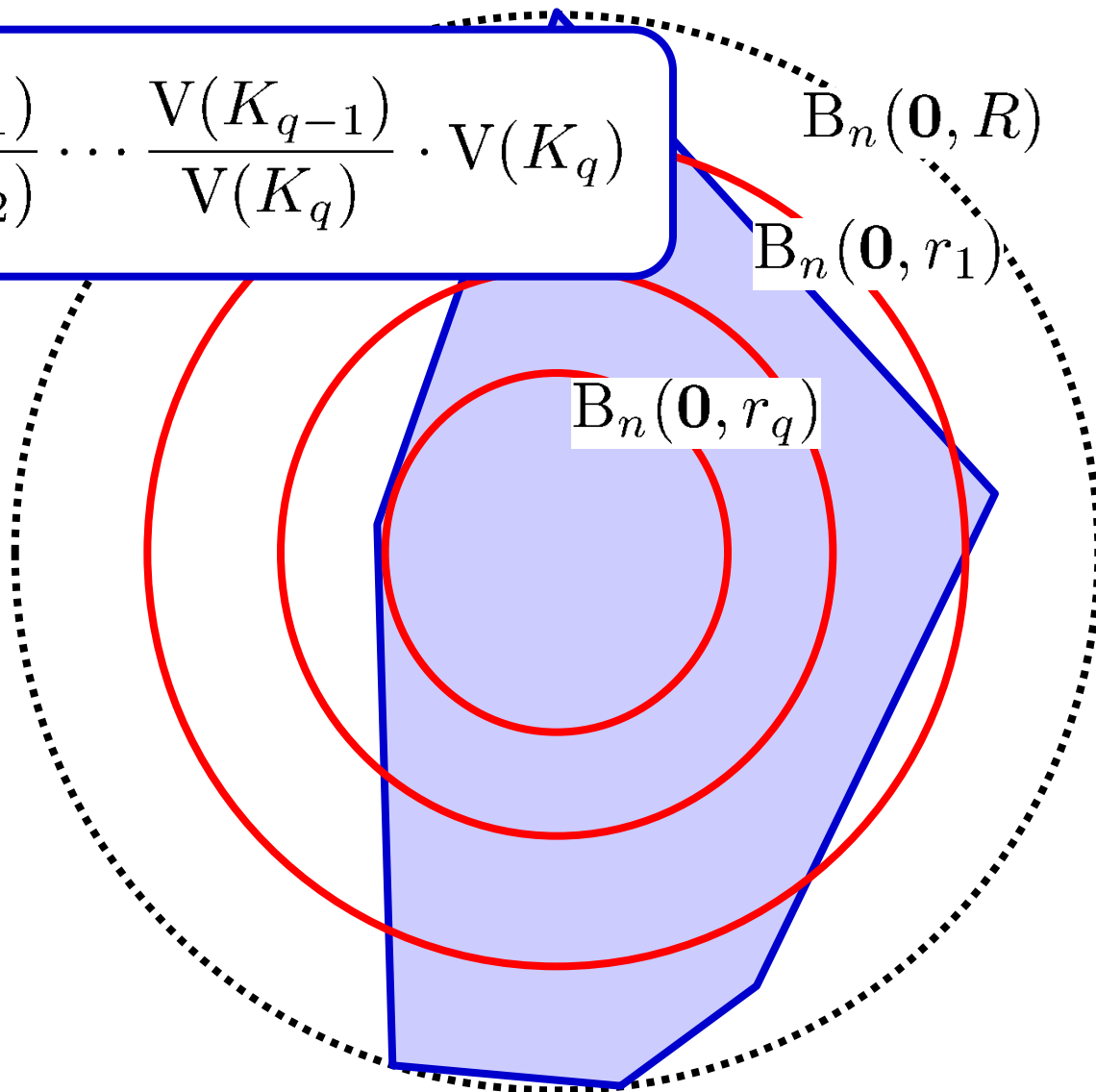


再帰式

$$K_i \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_i)$$

$$V(K) = \frac{V(K)}{V(K_1)} \cdot \frac{V(K_1)}{V(K_2)} \cdots \frac{V(K_{q-1})}{V(K_q)} \cdot V(K_q)$$

(ただし, $r_q = 1$)

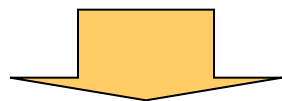


再帰式

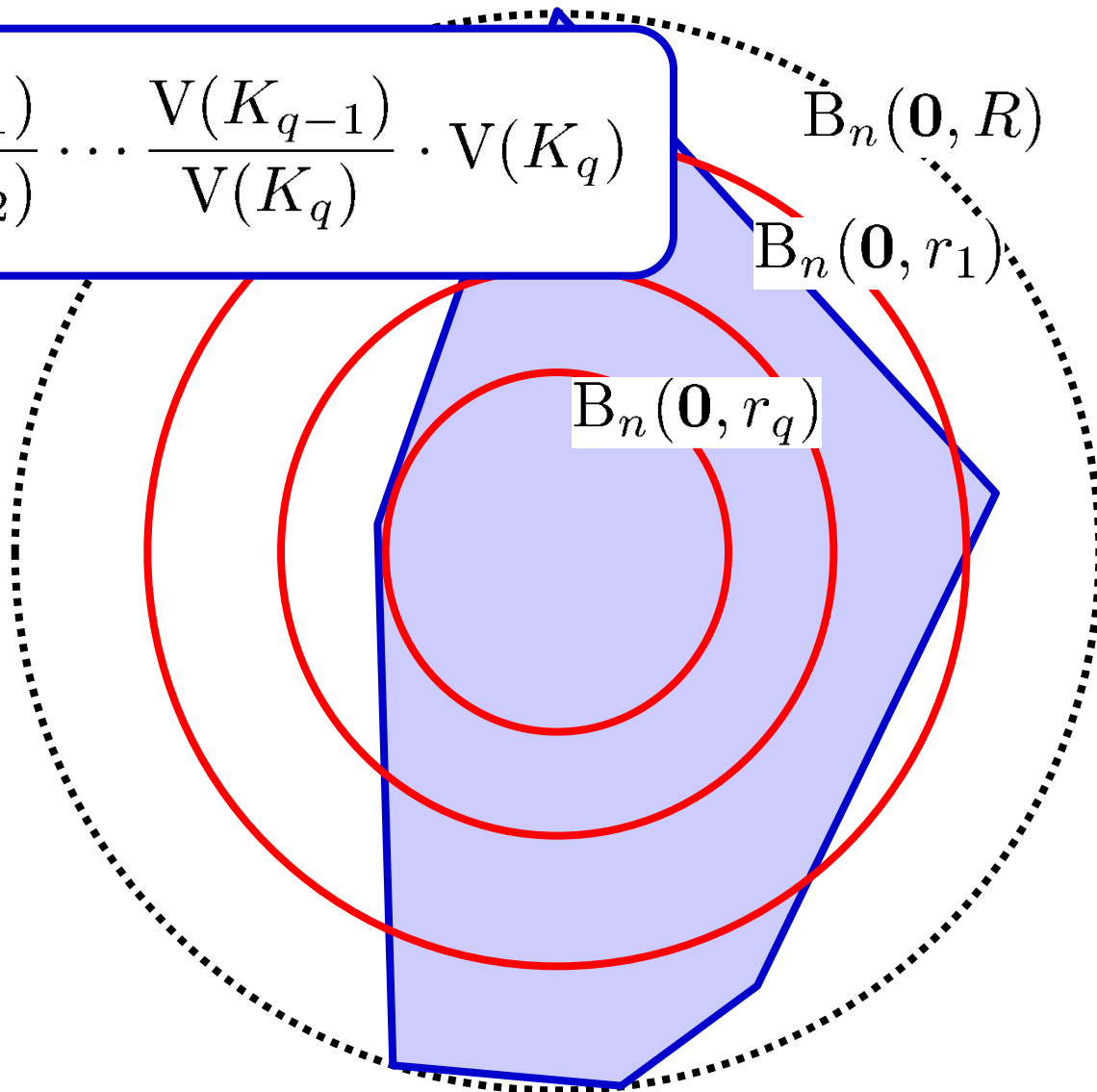
$$K_i \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_i)$$

$$V(K) = \frac{V(K)}{V(K_1)} \cdot \frac{V(K_1)}{V(K_2)} \cdots \frac{V(K_{q-1})}{V(K_q)} \cdot V(K_q)$$

(ただし, $r_q = 1$)



$$\begin{aligned} K_q &= B_n(\mathbf{0}, r_q) \\ &= B_n(\mathbf{0}, 1) \end{aligned}$$



再帰式

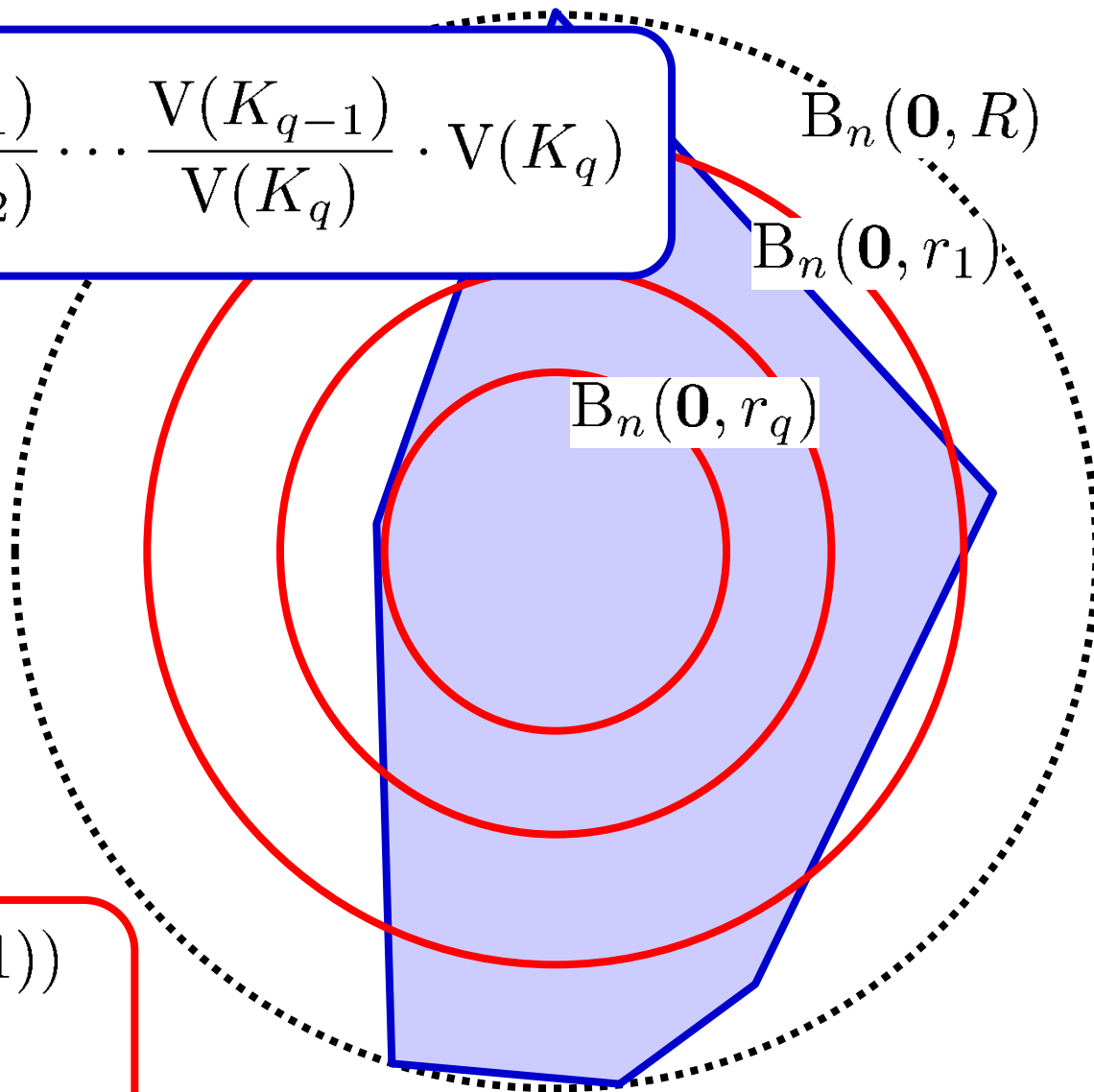
$$K_i \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_i)$$

$$V(K) = \frac{V(K)}{V(K_1)} \cdot \frac{V(K_1)}{V(K_2)} \cdots \frac{V(K_{q-1})}{V(K_q)} \cdot V(K_q)$$

(ただし, $r_q = 1$)

$$\begin{aligned} K_q &= B_n(\mathbf{0}, r_q) \\ &= B_n(\mathbf{0}, 1) \end{aligned}$$

$$\begin{aligned} V(K_q) &= V(B_n(\mathbf{0}, 1)) \\ &= \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)} \end{aligned}$$



再帰式

$$K_i \stackrel{\text{def.}}{=} K \cap B_n(\mathbf{0}, r_i)$$

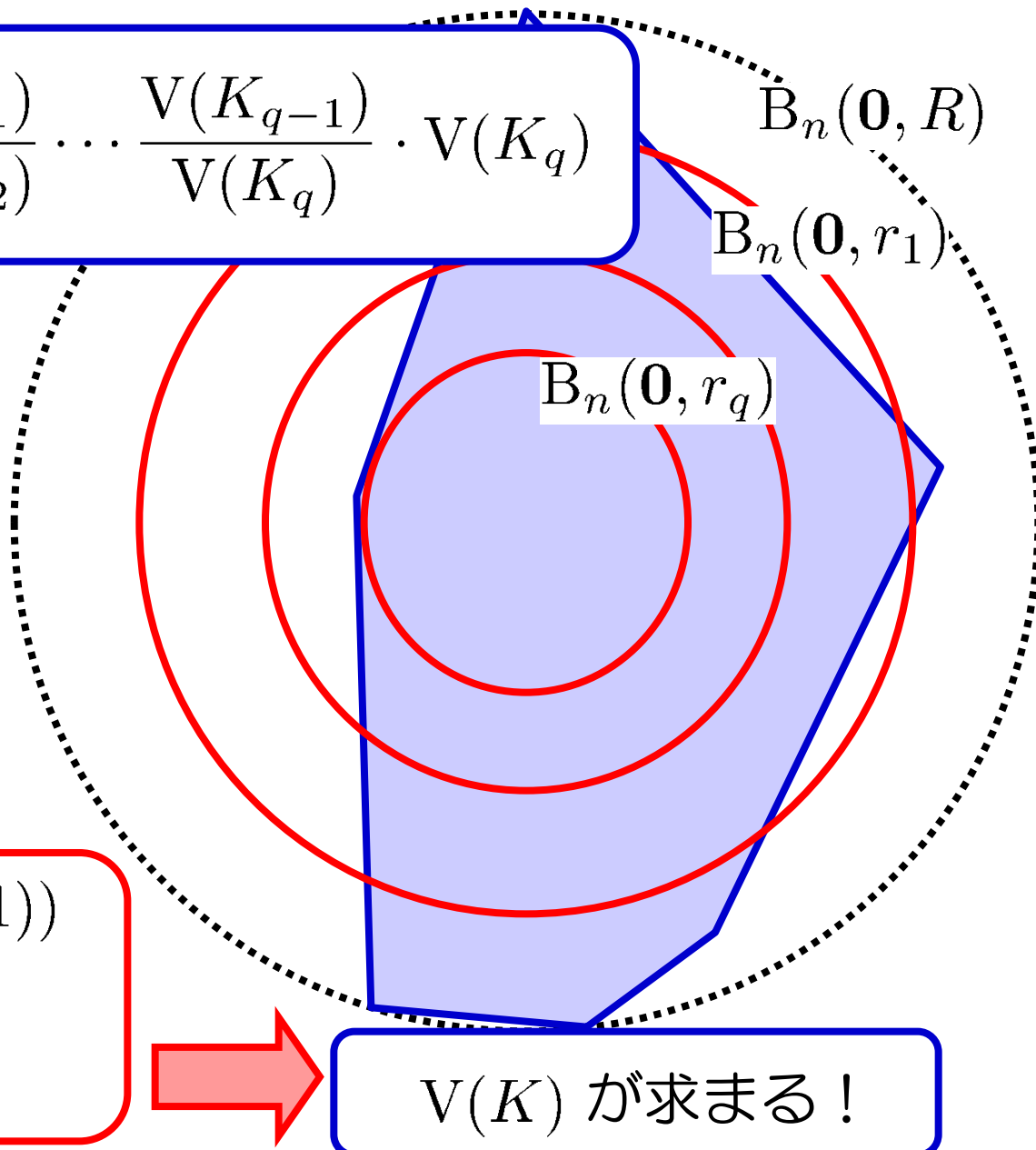
$$V(K) = \frac{V(K)}{V(K_1)} \cdot \frac{V(K_1)}{V(K_2)} \cdots \frac{V(K_{q-1})}{V(K_q)} \cdot V(K_q)$$

(ただし, $r_q = 1$)

$$\begin{aligned} K_q &= B_n(\mathbf{0}, r_q) \\ &= B_n(\mathbf{0}, 1) \end{aligned}$$

$$\begin{aligned} V(K_q) &= V(B_n(\mathbf{0}, 1)) \\ &= \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)} \end{aligned}$$

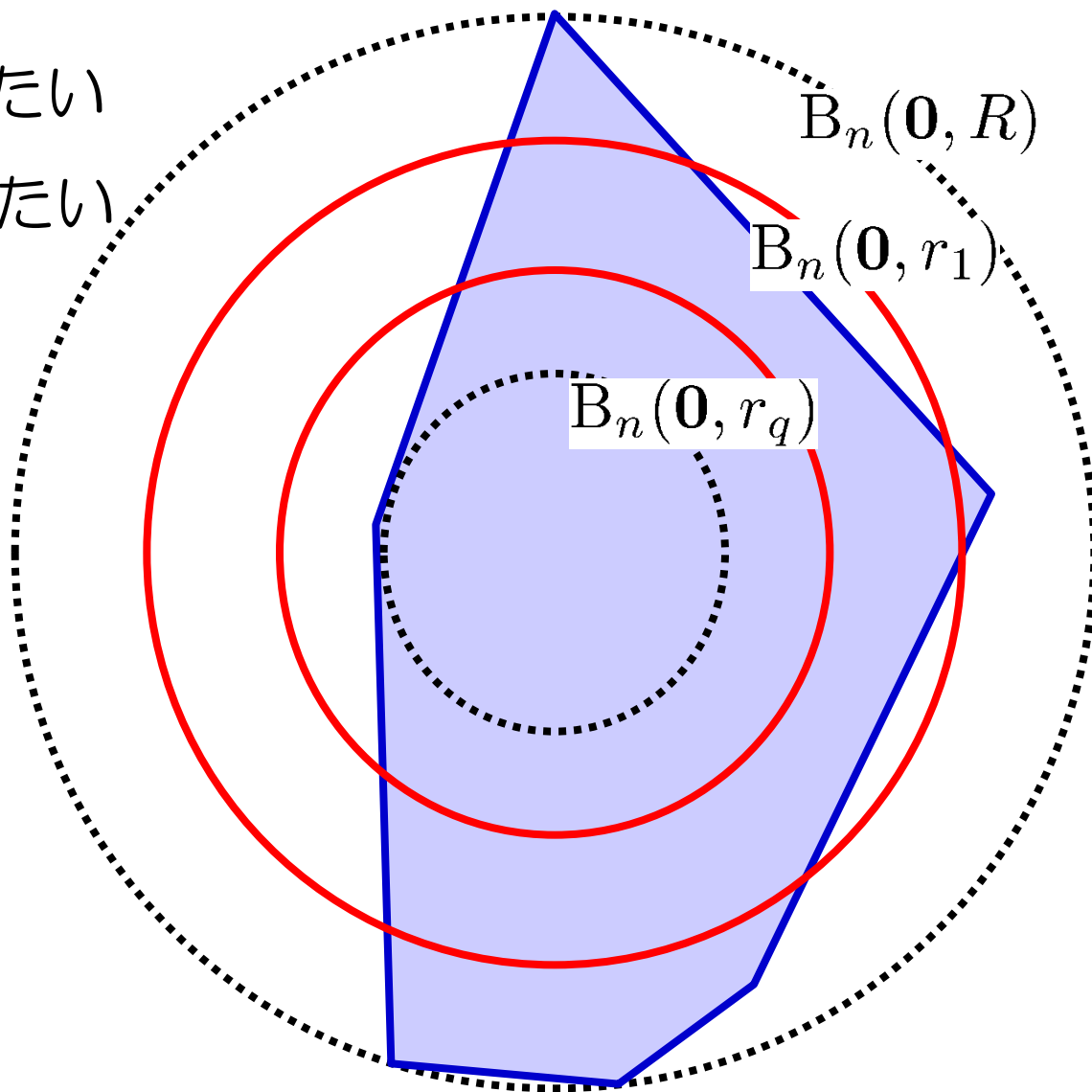
$V(K)$ が求まる!



再帰の回数

再帰の回数 $\Rightarrow$ 少なくしたい

比 $\Rightarrow$ 大きくしたい



再帰の回数

再帰の回数 $\Rightarrow$ 少なくしたい

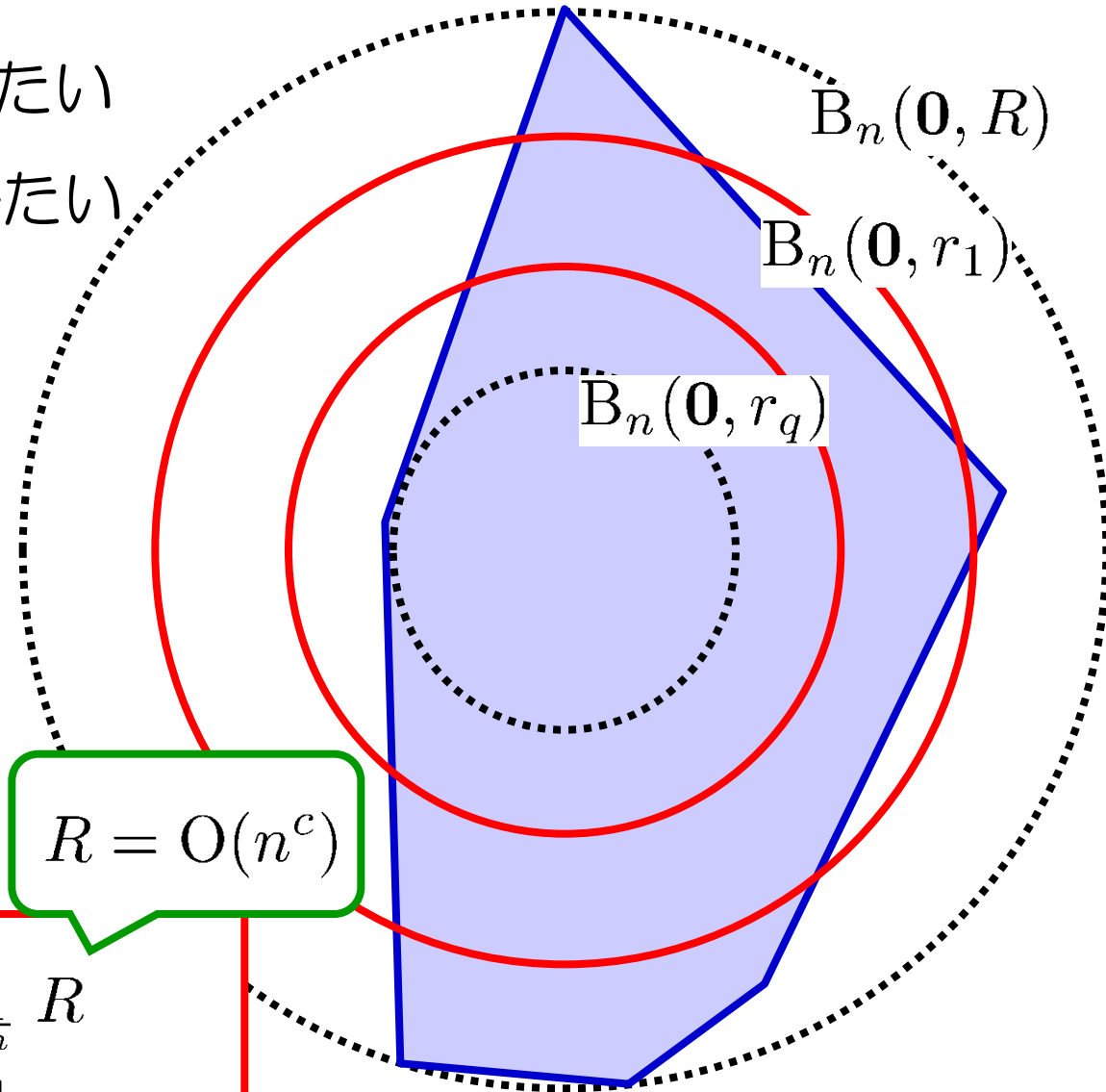
比 $\Rightarrow$ 大きくしたい

$$\frac{r_i}{r_{i-1}} = \frac{1}{\sqrt[n]{2}}$$

比

$$\frac{V(K_i)}{V(K_{i-1})} \geq \frac{1}{2}$$

$$\begin{aligned} (\text{再帰の回数}) &= \log_{2^{\frac{1}{n}}} R \\ &= c \cdot n \log_2 n \end{aligned}$$

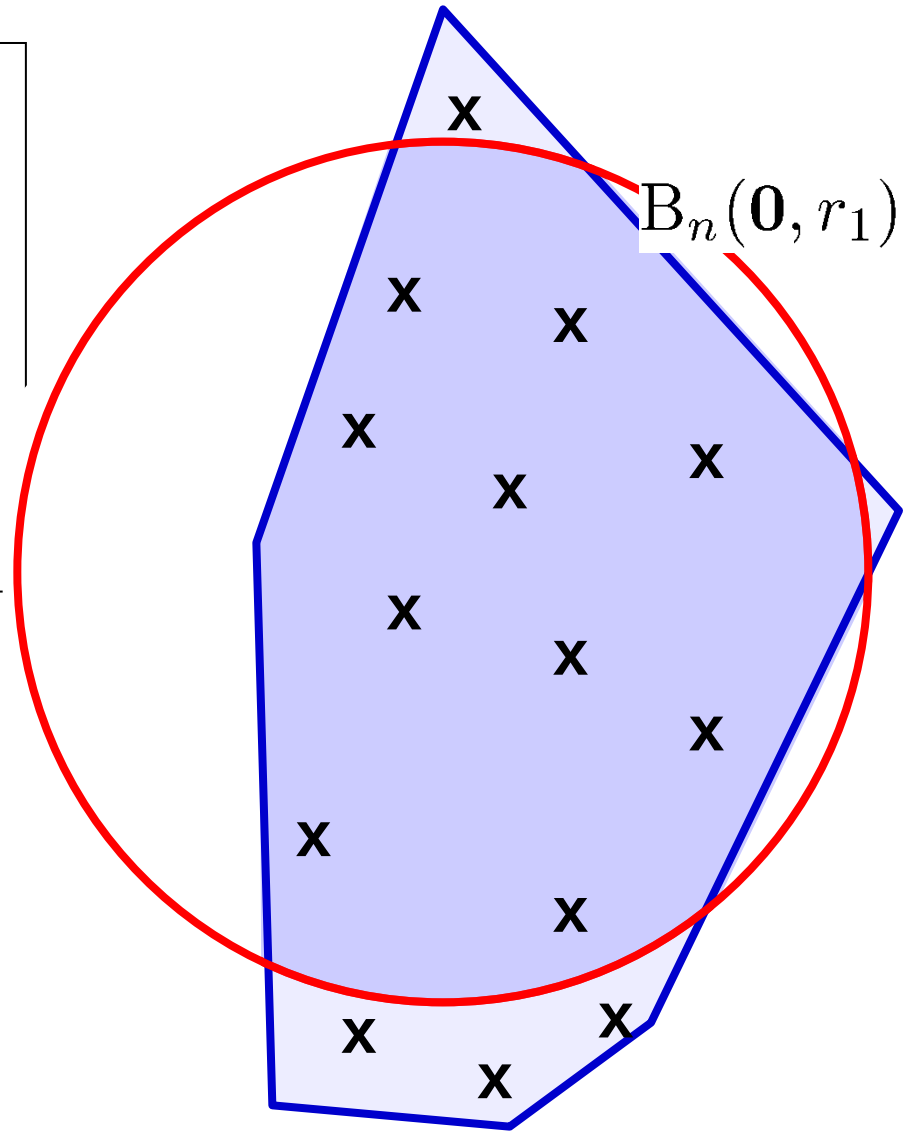


理論的な近似精度保証に向けて

真の比 $\rho_1 = \frac{V(K_1)}{V(K)}$ (unknown).

M 個の一様サンプルのうち,
 X_1 個が $B_n(\mathbf{0}, r_1)$ 内に落ちた.

誤差 $\left| \frac{X_1}{M} - \rho_1 \right|$ はどのくらい?

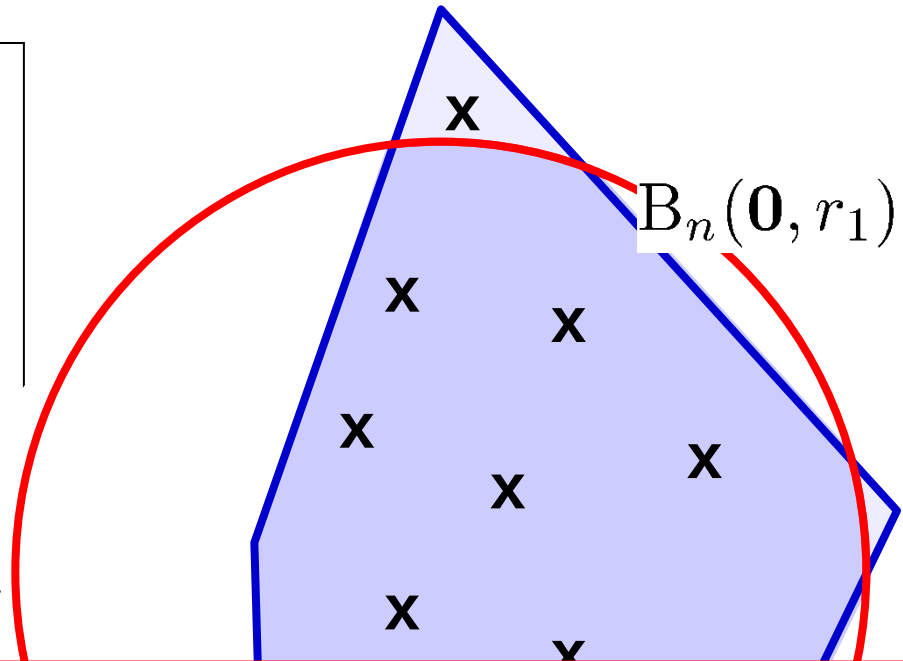


理論的な近似精度保証に向けて

真の比 $\rho_1 = \frac{V(K_1)}{V(K)}$ (unknown).

M 個の一様サンプルのうち,
 X_1 個が $B_n(\mathbf{0}, r_1)$ 内に落ちた.

誤差 $\left| \frac{X_1}{M} - \rho_1 \right|$ はどのくらい?



定理

K_i ($i = 1, 2, \dots, R$) の一様サンプリング・オラクルを用いると,

任意の ε ($0 < \varepsilon \leq 0.5$), δ ($0 < \delta < 1$) に対して, $\text{poly}(n, \varepsilon^{-1}, \log \delta^{-1})$ 時間で

$$\Pr \left[(1 - \varepsilon) \cdot V(K) \leq V(K_q) \prod_{i=1}^R \frac{M}{X_i} \leq (1 + \varepsilon) \cdot V(K) \right] \geq 1 - \delta$$

が成り立つ.

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,
 $X := X_1 + \dots + X_n$ の期待値を μ とすると,

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

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

系

$0 < \varepsilon \leq 1$ の場合,

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

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

証明の方針(パラメータ調整前)

サンプルサイズ M を十分大きな値にしておけば, Chernoff boundより,

$$\Pr[X_1 \geq (1 + \varepsilon)M\rho_1] \leq \exp\left(-\frac{\varepsilon^2 M \rho_1}{3}\right) \leq \delta$$

$$\Pr[X_1 \leq (1 - \varepsilon)M\rho_1] \leq \exp\left(-\frac{\varepsilon^2 M \rho_1}{2}\right) \leq \delta$$

$$\text{つまり } \Pr\left[(1 - \varepsilon)\rho_1 \leq \frac{X_1}{M} \leq (1 + \varepsilon)\rho_1\right] \geq 1 - 2 \exp\left(-\frac{\varepsilon^2 M \rho_1}{3}\right)$$

証明の方針(パラメータ調整前)

サンプルサイズ M を十分大きな値にしておけば, Chernoff boundより,

$$\Pr[X_1 \geq (1 + \varepsilon)M\rho_1] \leq \exp\left(-\frac{\varepsilon^2 M \rho_1}{3}\right) \leq \delta$$

$$\Pr[X_1 \leq (1 - \varepsilon)M\rho_1] \leq \exp\left(-\frac{\varepsilon^2 M \rho_1}{2}\right) \leq \delta$$

[Union bound]

すべての i に対して $(1 - \varepsilon)\rho_i \leq \frac{X_i}{M} \leq (1 + \varepsilon)\rho_i$ となる

確率は $1 - 2\delta$ 以上. つまり

$$\Pr\left[(1 - \varepsilon)^R \prod_{i=1}^R \rho_i \leq \prod_{i=1}^R \frac{X_i}{M} \leq (1 + \varepsilon)^R \prod_{i=1}^R \rho_i\right] \geq 1 - 2R\delta$$

すなわち,

$$\Pr\left[\frac{V(K_q)}{(1 - \varepsilon)^R} \prod_{i=1}^R \frac{1}{\rho_i} \geq V(K_q) \prod_{i=1}^R \frac{X_i}{M} \geq \frac{V(K_q)}{(1 + \varepsilon)^R} \prod_{i=1}^R \frac{1}{\rho_i}\right] \geq 1 - 2R\delta$$

δ は小さくする
必要がある.

Recall

$$V(K_q) \prod_{i=1}^n \frac{1}{\rho_i} = V(K)$$

証明の方針(パラメータ調整前)

(承前)

$$\Pr \left[\frac{V(K_q)}{(1-\varepsilon)^R} \prod_{i=1}^R \frac{1}{\rho_i} \geq V(K_q) \prod_{i=1}^R \frac{M}{X_i} \geq \frac{V(K_q)}{(1+\varepsilon)^R} \prod_{i=1}^R \frac{1}{\rho_i} \right] \geq 1 - 2R\delta$$

$$\Pr \left[\frac{V(K)}{(1-\varepsilon)^R} \geq V(K_q) \prod_{i=1}^R \frac{M}{X_i} \geq \frac{V(K)}{(1+\varepsilon)^R} \right] \geq 1 - 2R\delta$$

$$\frac{1}{(1-\varepsilon)^R} \simeq (1+\varepsilon)^R \simeq 1 + R\varepsilon$$

正確には

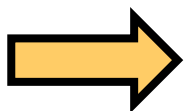
$$\frac{1}{(1-\varepsilon)^R} \leq (1+2\varepsilon)^R \leq 1 + 4R\varepsilon$$

if $\varepsilon < \frac{1}{2}$

$$\frac{1}{(1+\varepsilon)^R} \simeq (1-\varepsilon)^R \simeq 1 - R\varepsilon$$

正確には

$$\frac{1}{(1+\varepsilon)^R} \geq (1-\varepsilon)^R \geq 1 - R\varepsilon$$



パラメータ設定(like ε - δ): $\varepsilon \rightarrow \frac{\varepsilon'}{4R}$, $\delta \rightarrow \frac{\delta'}{2R}$

証明の方針(パラメータ調整後)

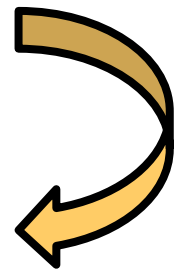
サンプルサイズ M を十分大きな値にしておけば, Chernoff boundより,

$$\Pr \left[X_1 \geq \left(1 + \frac{\varepsilon'}{4R} \right) M \rho_1 \right] \leq \exp \left(- \frac{\left(\frac{\varepsilon'}{4R} \right)^2 M \rho_1}{3} \right) \leq \frac{\delta'}{2R}$$

$$\Pr \left[X_1 \leq \left(1 - \frac{\varepsilon'}{4R} \right) M \rho_1 \right] \leq \exp \left(- \frac{\left(\frac{\varepsilon'}{4R} \right)^2 M \rho_1}{2} \right) \leq \frac{\delta'}{2R}$$

[Union bound]

$$\Pr \left[\left(1 - \frac{\varepsilon'}{4R} \right)^R \prod_{i=1}^R \rho_i \leq \prod_{i=1}^R \frac{X_i}{M} \leq \left(1 + \frac{\varepsilon'}{4R} \right)^R \prod_{i=1}^R \rho_i \right] \geq 1 - 2R \frac{\delta'}{2R} = 1 - \delta'$$



すなわち,

$$\Pr \left[\frac{V(K_q)}{\left(1 - \frac{\varepsilon'}{4R} \right)^R} \prod_{i=1}^R \frac{1}{\rho_i} \geq V(K_q) \prod_{i=1}^R \frac{X_i}{M} \geq \frac{V(K_q)}{\left(1 + \frac{\varepsilon'}{4R} \right)^R} \prod_{i=1}^R \frac{1}{\rho_i} \right] \geq 1 - \delta'$$

証明の方針(パラメータ調整後)

サンプルサイズ $M = 3 \left(\frac{4R}{\varepsilon'} \right)^2 \ln \left(\frac{2R}{\delta'} \right)$ とすると, Chernoff boundより,

$$\Pr \left[X_1 \geq \left(1 + \frac{\varepsilon'}{4R} \right) M \rho_1 \right] \leq \exp \left(- \frac{\left(\frac{\varepsilon'}{4R} \right)^2 M \rho_1}{3} \right) \leq \frac{\delta'}{2R}$$

$$\Pr \left[X_1 \leq \left(1 - \frac{\varepsilon'}{4R} \right) M \rho_1 \right] \leq \exp \left(- \frac{\left(\frac{\varepsilon'}{4R} \right)^2 M \rho_1}{2} \right) \leq \frac{\delta'}{2R}$$

[Union bound]

$$\Pr \left[\left(1 - \frac{\varepsilon'}{4R} \right)^R \prod_{i=1}^R \rho_i \leq \prod_{i=1}^R \frac{X_i}{M} \leq \left(1 + \frac{\varepsilon'}{4R} \right)^R \prod_{i=1}^R \rho_i \right] \geq 1 - 2R \frac{\delta'}{2R} = 1 - \delta'$$



すなわち,

$$\Pr \left[\frac{V(K_q)}{\left(1 - \frac{\varepsilon'}{4R} \right)^R} \prod_{i=1}^R \frac{1}{\rho_i} \geq V(K_q) \prod_{i=1}^R \frac{X_i}{M} \geq \frac{V(K_q)}{\left(1 + \frac{\varepsilon'}{4R} \right)^R} \prod_{i=1}^R \frac{1}{\rho_i} \right] \geq 1 - \delta'$$

証明の方針(パラメータ調整後)

(承前)

$$\Pr \left[\frac{V(K_q)}{\left(1 - \frac{\varepsilon'}{4R}\right)^R} \prod_{i=1}^R \frac{1}{\rho_i} \geq V(K_q) \prod_{i=1}^R \frac{M}{X_i} \geq \frac{V(K_q)}{\left(1 + \frac{\varepsilon'}{4R}\right)^R} \prod_{i=1}^R \frac{1}{\rho_i} \right] \geq 1 - \delta'$$

$$\Pr \left[\frac{V(K)}{\left(1 - \frac{\varepsilon'}{4R}\right)^R} \geq V(K_q) \prod_{i=1}^R \frac{M}{X_i} \geq \frac{V(K)}{\left(1 + \frac{\varepsilon'}{4R}\right)^R} \right] \geq 1 - \delta'$$

$$\frac{1}{\left(1 - \frac{\varepsilon'}{4R}\right)^R} \leq \left(1 + 2\frac{\varepsilon'}{4R}\right)^R \leq 1 + 4R \frac{\varepsilon'}{4R} = 1 + \varepsilon'$$

$$\frac{1}{\left(1 + \frac{\varepsilon'}{4R}\right)^R} \geq \left(1 - \frac{\varepsilon'}{4R}\right)^R \geq 1 - R \frac{\varepsilon'}{4R} \geq 1 - \varepsilon'$$

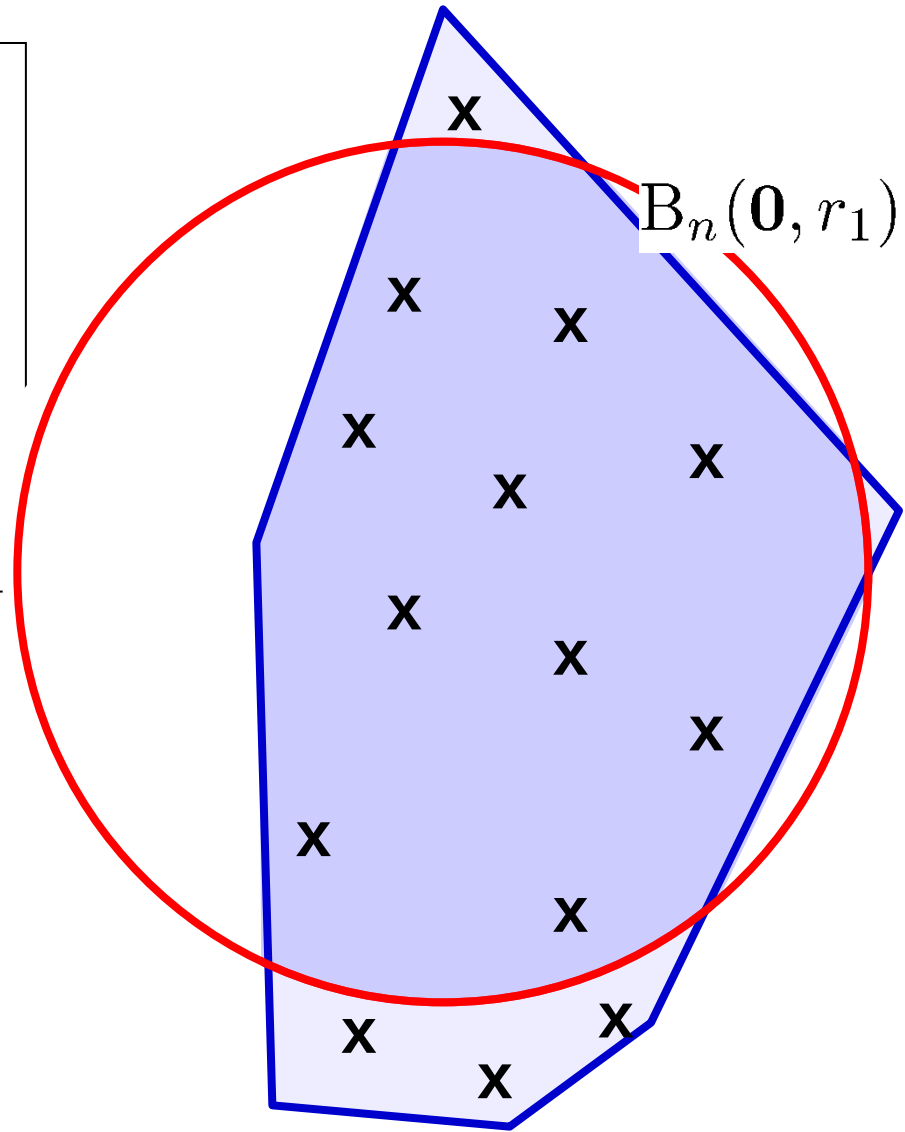
$$\Pr \left[(1 + \varepsilon') \cdot V(K) \geq V(K_q) \prod_{i=1}^R \frac{M}{X_i} \geq (1 - \varepsilon') \cdot V(K) \right] \geq 1 - \delta'$$

理論的な近似精度保証に向けて

真の比 $\rho_1 = \frac{V(K_1)}{V(K)}$ (unknown).

M 個の一様サンプルのうち,
 X_1 個が $B_n(\mathbf{0}, r_1)$ 内に落ちた.

誤差 $\left| \frac{X_1}{M} - \rho_1 \right|$ はどのくらい?



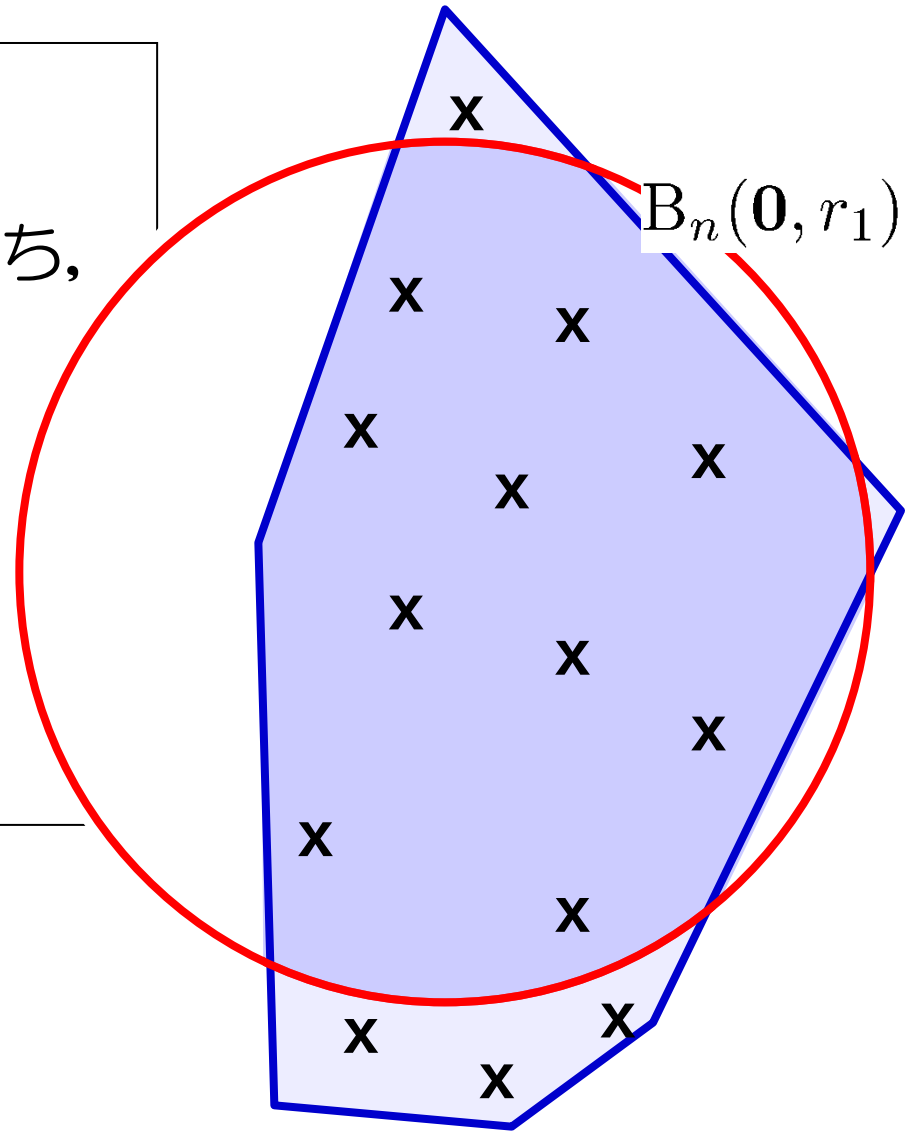
理論的な近似精度保証に向けて

真の比 $\rho_1 = \frac{V(K_1)}{V(K)}$ (unknown).

M 個の近似一様サンプルのうち,
 X_1 個が $B_n(\mathbf{0}, r_1)$ 内に落ちた.

誤差 $\left| \frac{X_1}{M} - \rho_1 \right|$ はどのくらい?

ただし $|\rho_1 - \hat{\rho}_1|$ は
かなり小さくできると仮定.



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,
 $X := X_1 + \dots + X_n$ の期待値を μ とすると,

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

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

系

$0 < \varepsilon \leq 1$ の場合,

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

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

証明の方針(パラメータ調整前)

サンプルサイズ M を十分大きな値にしておけば, Chernoff boundより,

$$\Pr[X_1 \geq (1 + \varepsilon)M\widehat{\rho}_1] \leq \exp\left(-\frac{\varepsilon^2 M \widehat{\rho}_1}{2}\right) \leq \delta$$

$$\Pr[X_1 \leq (1 - \varepsilon)M\widehat{\rho}_1] \leq \exp\left(-\frac{\varepsilon^2 M \widehat{\rho}_1}{2}\right) \leq \delta$$

$$\text{つまり } \Pr\left[(1 - \varepsilon)\widehat{\rho}_1 \leq \frac{X_1}{M} \leq (1 + \varepsilon)\widehat{\rho}_1\right] \geq 1 - 2 \exp\left(-\frac{\varepsilon^2 M \widehat{\rho}_1}{3}\right)$$

Remark

$$|\rho_1 - \widehat{\rho}_1| \leq \gamma \ll 1 \text{ とすると, } \rho_1 \geq \frac{1}{2} \text{ の仮定の下}$$

$$\widehat{\rho}_1 \geq \rho_1 - \gamma \geq (1 - 2\gamma)\rho_1$$

これを代入すると,

$$\Pr[X_1 \geq (1 + \varepsilon)(1 - 2\gamma)M\rho_1] \leq \exp\left(-\frac{\varepsilon^2 M (\rho_1 - \gamma)}{3}\right) \leq \delta$$

$$\Pr[X_1 \leq (1 - \varepsilon)(1 + 2\gamma)M\rho_1] \leq \exp\left(-\frac{\varepsilon^2 M (\rho_1 - \gamma)}{3}\right) \leq \delta$$

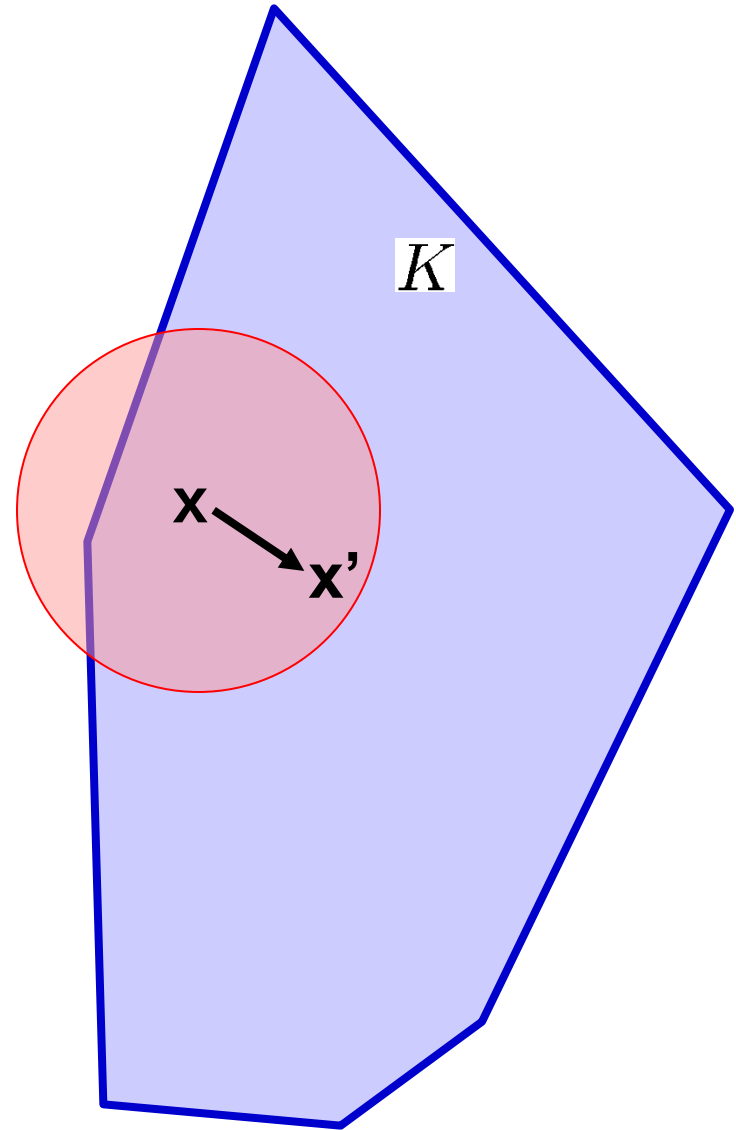
以下、パラメータ調整(略)

理論的な近似精度保証に向けて

Ball-walk

=> 定常分布は一様

十分な回数の推移は何回？



理論的な近似精度保証に向けて

Ball-walk

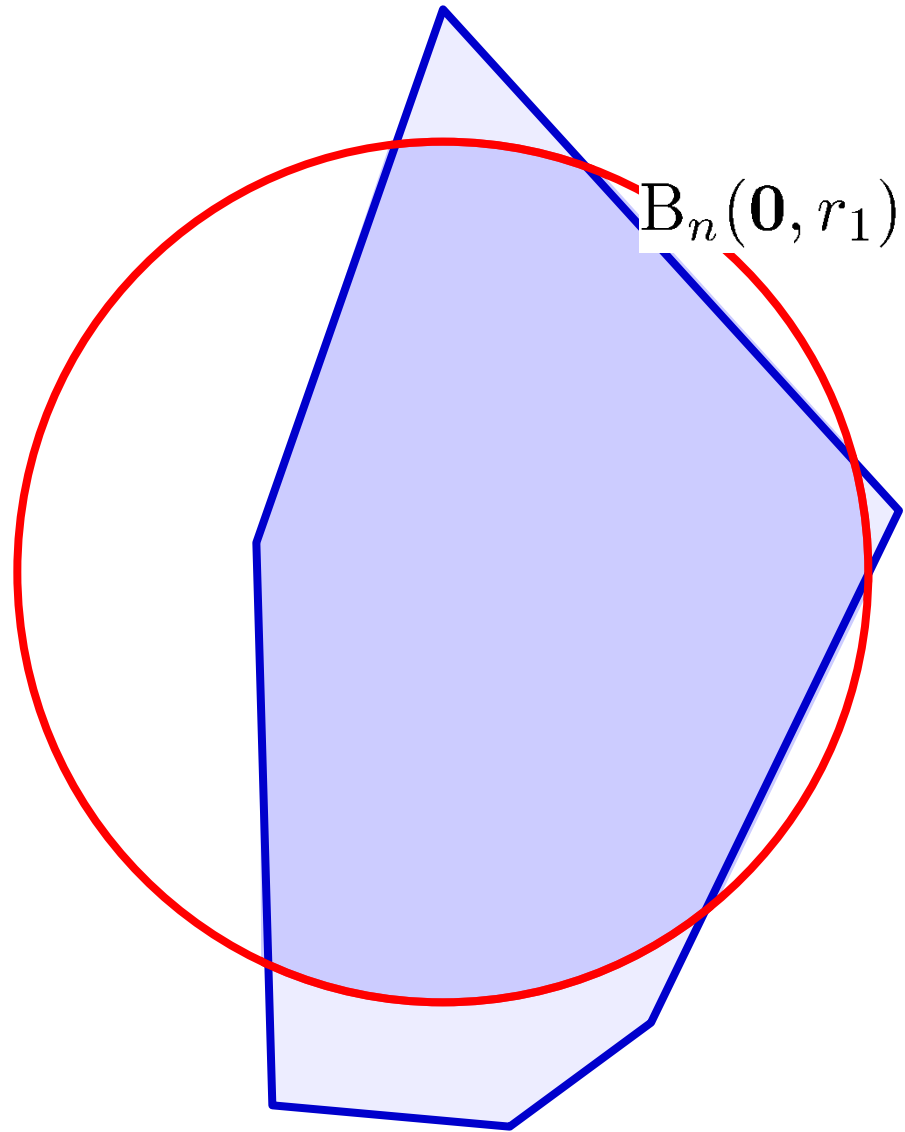
=> 定常分布は一様

十分な回数の推移は何回？

モンテカルロ法

- 真の比 ρ
- 近似比 $\hat{\rho}$ (分布の誤差に起因)

$$|\hat{\rho} - \rho| < \gamma$$



理論的な近似精度保証に向けて

Ball-walk

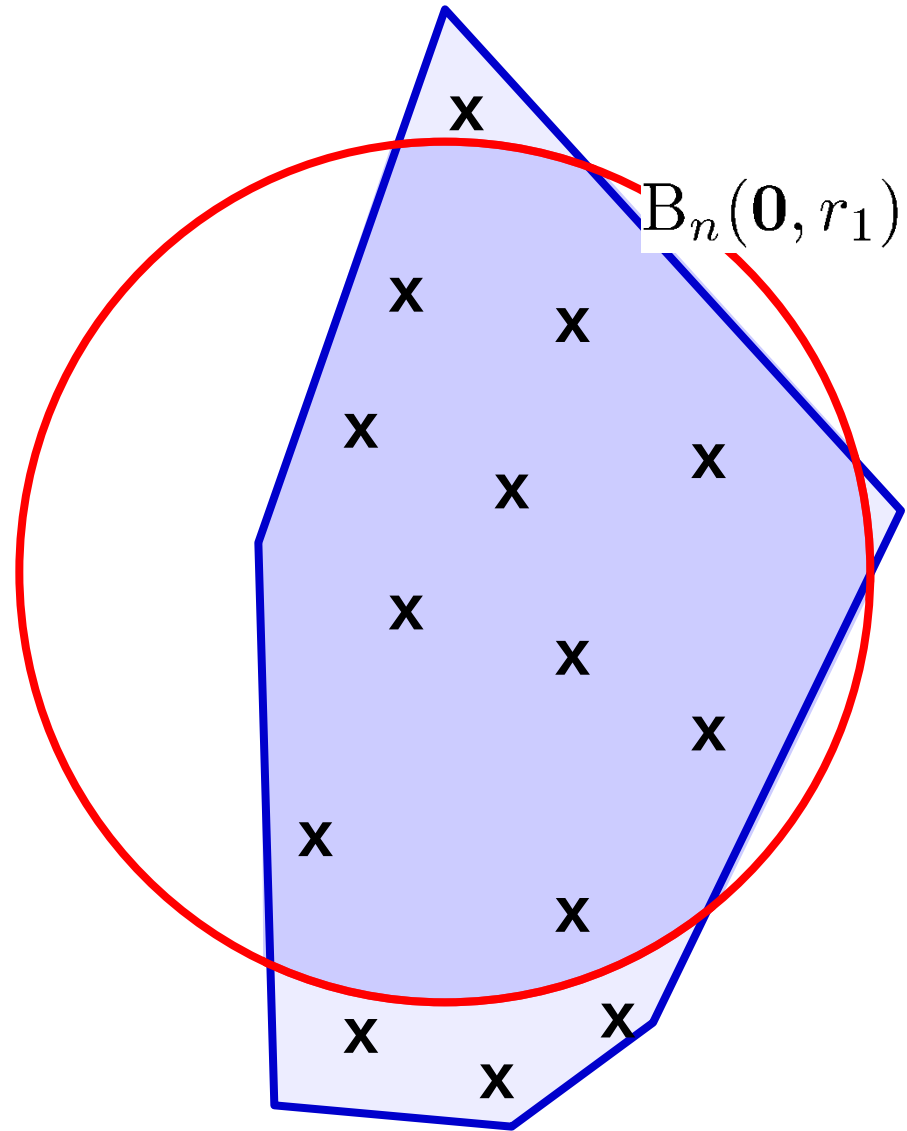
=> 定常分布は一様

十分な回数の推移は何回？

モンテカルロ法

- 真の比 ρ
- 近似比 $\hat{\rho}$ (分布の誤差に起因)

$$|\hat{\rho} - \rho| < \gamma$$



理論的な近似精度保証に向けて

Ball-walk

=> 定常分布は一様

十分な回数の推移は何回？

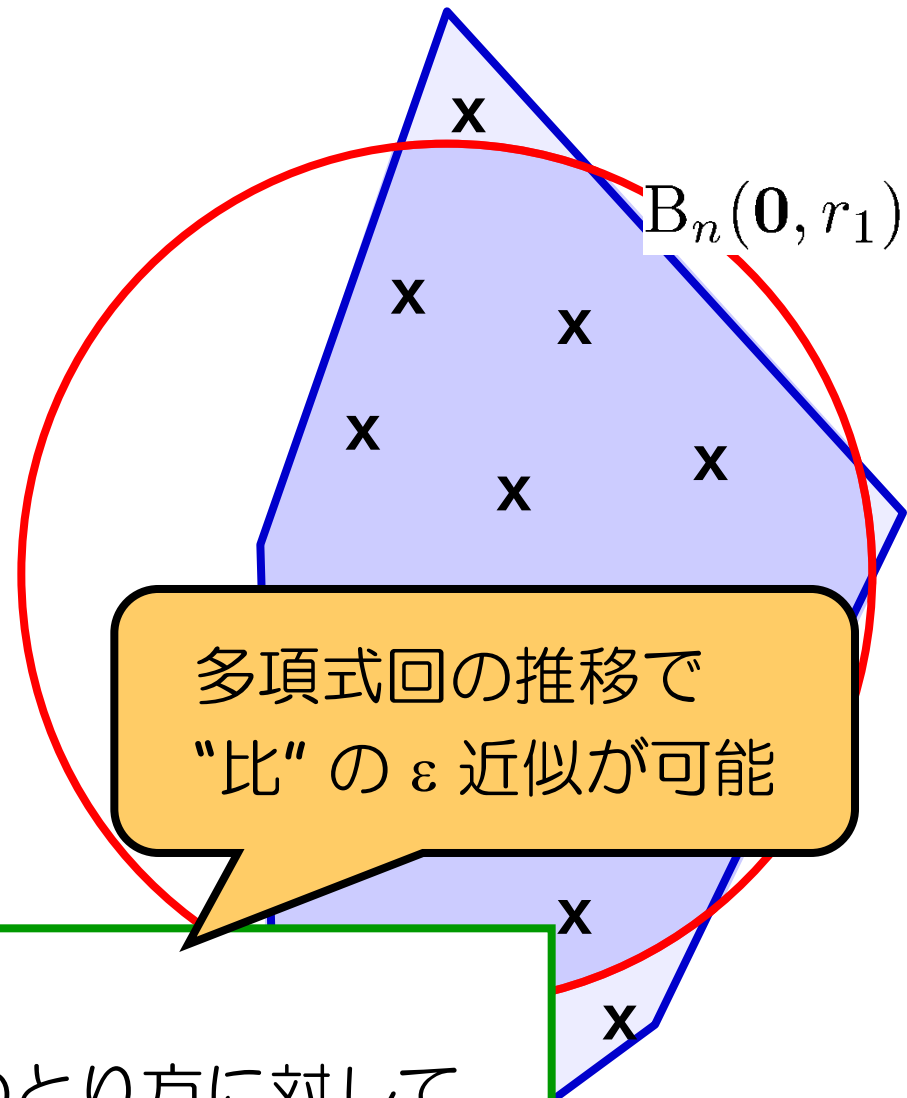
モンテカルロ法

- 真の比 ρ
- 近似比 $\hat{\rho}$ (分布の誤差に起因)

$$|\hat{\rho} - \rho| < \gamma$$

mixing time:

誤差が最大となる部分集合のとり方に対して,
誤差 γ が保証される推移の回数.



FPRAS設計のまとめ

効率的な再帰的モンテカルロ法を設計するには？

1. 帰着先の割合が大きい
2. 再帰の回数が小さい
3. 帰着先のランダムサンプリングが容易

SELF-REDUCTION

計算機科学のMCMC法

計算量クラス #P [Valiant '79]

$\mathcal{L}$ がNPに属するとき, $w \in \mathcal{L}$ に対し,
 $\mathcal{C}(w) \subseteq \Sigma^*$ を w のcertificateの集合とする.
 $|\mathcal{C}(w)|$ を求める問題のクラスを#Pと呼ぶ.

1998年Gödel prize

Cf. 定理 [戸田1991]

$$P^{\#P} = PH$$

“Valiant計画” (since 1979)

#P問題 $|\mathcal{C}(w)| \in \mathbb{R}$ に対して, 近似解 $Z \in \mathbb{R}$ を精度

$$\Pr[(1 - \epsilon)|\mathcal{C}(w)| \leq Z \leq (1 + \epsilon)|\mathcal{C}(w)|] \geq 1 - \delta$$

で多項式時間($\text{poly}(\text{input}, \epsilon^{-1}, \log \delta^{-1})$)で求められるか?

Fully **P**olynomial-time **R**andomized **A**pproximation **S**cheme (**FPRAS**)

$A \in \mathbb{R}$ に対して, 近似解 $Z \in \mathbb{R}$ を精度

$$\Pr[(1 - \epsilon)A \leq Z \leq (1 + \epsilon)A] \geq 1 - \delta$$

で多項式時間($\text{poly}(\text{input}, \epsilon^{-1}, \log \delta^{-1})$)で求めるアルゴリズム.

$(\mathcal{C}(w_i) \cap \mathcal{C}(w_j) \ (i \neq j) \text{が成り立つ})$

Self-reduction

任意の $w \in \mathcal{L}$ に対して, w の certificate によって
 適当な $w_i \in \mathcal{L} \ (i = 1, \dots, k)$ が定まり $\mathcal{C}(w) = \bigcup_{i=1}^k \mathcal{C}(w_i)$ と
 分割できるとき, $\mathcal{L}$ は **self-reducible** (自己帰着可能) という.

定理 [Jerrum Valiant Vazirani 86]

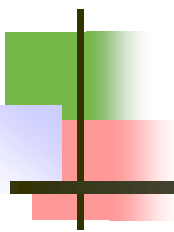
$\mathcal{L}$ が self-reducible なとき, $|\mathcal{C}(w)|$ の FPRAS ができると,
 $\mathcal{C}(w)$ の ϵ -近似一様ランダムサンプリングが $\text{poly}(\text{input}, \epsilon^{-1})$
 時間でできることは等価.

証明の概略

($\Rightarrow$) 確率 $\frac{|\mathcal{C}(w_i)|}{|\mathcal{C}(w)|}$ で self-reduction を繰り返すと $\mathcal{C}(w') = \{a\}$ にたどり着く.

$\mathcal{C}(w) = \bigcup_{i=1}^k \mathcal{C}(w_i)$ が分割なので, $\mathcal{C}(w')$ の唯一の要素 a は
 他の単一集合に現れることがない. 確率の定め方から一様性は明らか.

($\Leftarrow$) 求め方はお姉さんの例と同様. $\text{poly}(\text{input}, \epsilon^{-1}, \log \delta^{-1})$ 時間で
 $\Pr[(1 - \epsilon)A \leq Z \leq (1 + \epsilon)A] \geq 1 - \delta$ は後の証明 (N - ϵ 論法) と同様.



おねえさん問題の乱択近似

柴田友樹¹, 山内由紀子², 来嶋秀治², 山下雅史²

九州大学 工学部電気情報工学科 (1:研究当時)

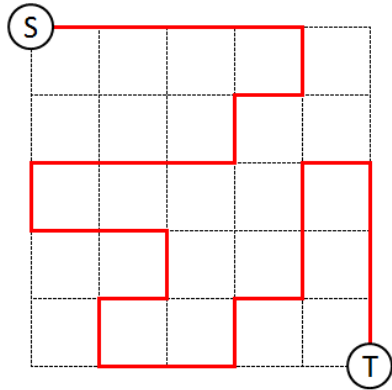
大学院システム情報科学研究所 (2 :研究当時)

[参照]

- ❑ 柴田友樹, 山内由紀子, 来嶋秀治, 山下雅史, 格子上の経路数え上げの乱択近似--24時間で解くおねえさん問題, 2015年度 冬のLAシンポジウム, 京都大学数理解析研究所, 2016年1月28日.
- ❑ 柴田 友樹, 格子上の経路数え上げの乱択近似, 九州大学工学部電気情報工学科卒業論文, 2016年3月.

おねえさん問題

$n \times n$ 格子グラフ上で、左上(s)から右下(t)を結ぶ、 s - t 単純経路の総数を求めよ。



組み合わせの数え方
How to count combinations

画像引用：日本科学未来館、『フカシギの数え方』おねえさんといっしょ！
みんなで数えてみよう！, YouTube, 2012/9/10公開,
<https://www.youtube.com/watch?v=Q4gTV4r0zRs>

おねえさん問題 youtube

search

01

12

212

3184

48512

51262816

6575780564

7789360053252

83266598486981642

941044208702632496804

101568758030464750013214100

11182413291514248049241470885236

1264528039343270018963357185158482118

1369450664761521361664274701548907358996488

14227449714676812739631826459327989863387613323440

152266745568862672746374567396713098934866324885408319028

166874544560914993158763132489232824587945968099457285419306

176344814611237963971310297540795524404944398866480693646369387855336

181782112840842065129893384946652325275167838065704767655931452474605826692782532

1915233449717048799930807428103192969089945425553232329455577602986673735506059287756925844

20396289219982303756020729951713362502106339705739463771515237113377010682364035706704472064940398

2131374751050137102720420538137382214513103312193698723653061351991346433379389385793965576992246021316463868

22755970286667345339661519123315222619353103732072409481167391410479517925792743631234987038883317634987271171404439792

23554354293552374770099143184890614379306903799709643313325569586464840084073348855445663869240208757112420600854085134829333945720

2412371712321207064758338744862673570832373041989012943539678727080484951695515930485641394550792153037191858028212512280926600304581386791094

25840297485788113347100708374543680912729605429377538354982474262393702849789821525692917857708397096012625602506027316549718402106494049978375604247408

26173699315862792793117544042123649890037222958828814060466370372091034241327613476278921819349800610708296223143380491348290026771931129627708738890853908108906396

単連結な彩色のためのマルコフ連鎖

マルコフ連鎖 MC

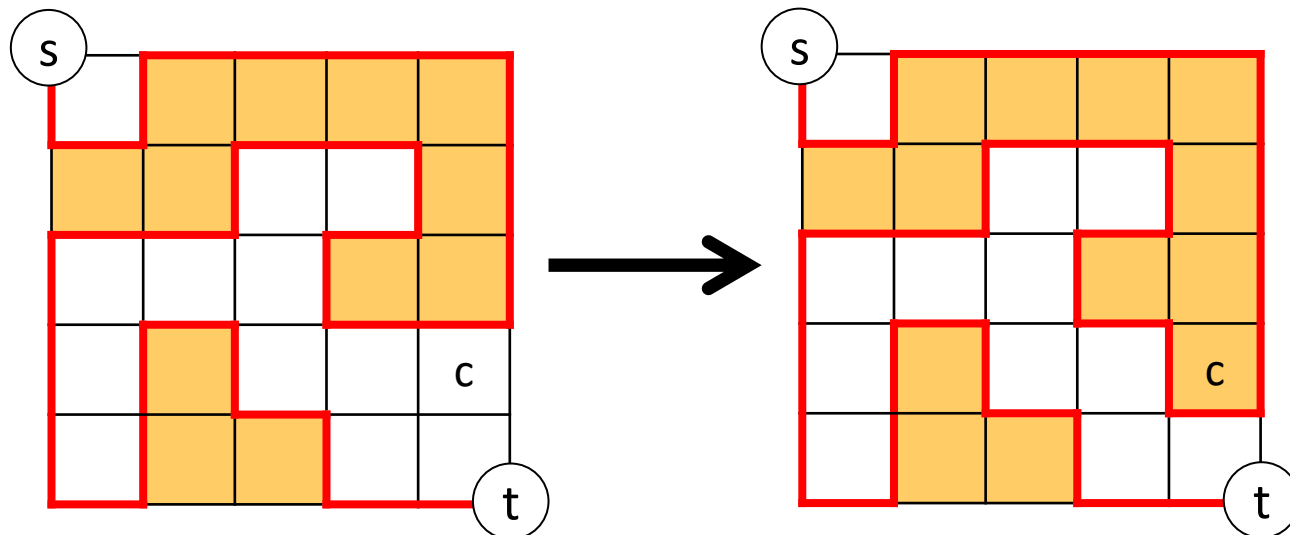
状態空間: Ξ_k ,

状態遷移: $X \rightarrow X'$:

Step 1. セル c を一様ランダムに選ぶ.

Step 2. $Y = X \oplus c$.

Step 3. もし $Y \in \Xi_k$ ならば $X' = Y$, そうでなければ $X' = X$.



アイデア

つまり, 求めたいものは $|\Omega| (= |\mathcal{E}_{L^*}|)$.

記法

□ $\Omega = \{s-t \text{ 単純経路}\}$

□ $\mathcal{E}_k = \{\text{長さが } k \text{ 以下の } s-t \text{ 単純経路}\}$

($k = 2n, 2n + 2, 2n + 4, \dots, L^*$)

ただし L^* は最長路の長さで, 以下で与えられる:

$$L^* = \begin{cases} (n+1)^2 - 1 & (n \text{ 奇数}) \\ (n+1)^2 - 2 & (n \text{ 偶数}) \end{cases}$$

Self-reducibility

$$|\mathcal{E}_{L^*}| = \frac{|\mathcal{E}_{L^*}|}{|\mathcal{E}_{L^*-2}|} \cdot \frac{|\mathcal{E}_{L^*-2}|}{|\mathcal{E}_{L^*-4}|} \cdots \frac{|\mathcal{E}_{2n+4}|}{|\mathcal{E}_{2n+2}|} \cdot \frac{|\mathcal{E}_{2n+2}|}{|\mathcal{E}_{2n}|} \cdot |\mathcal{E}_{2n}|$$

アイデア 2.

逆数の $\frac{|\mathcal{E}_{k-2}|}{|\mathcal{E}_k|}$ はモンテカルロ法で求まる.

アイデア 1.

$$|\mathcal{E}_{2n}| = \# \text{最短経路} \\ = \binom{2n}{n}$$

アルゴリズム

パラメータ:

τ (マルコフ連鎖の遷移回数)

M (モンテカルロ法のサンプル)

入力: n (格子のサイズ)

出力: Z (経路総数の近似値)

Set $Z := 1$;

For($k = 2n + 2$; $k < L^*$; $k := k + 2$) {

Set $X \in \Xi_k$; (X はマルコフ連鎖MCの初期状態)

Set $S := 0$; (S はカウンター)

for($i = 0$; $i < M$; $i++$) {

for($j = 0$; $j < \tau$; $j++$) {

X を更新 (マルコフ連鎖の1回遷移)

}

if($X \in \Xi_{k-2}$) $S++$;

}

Set $Z := Z * \frac{M}{S}$;

}

Output Z ;

Ξ_k から一様サンプリング

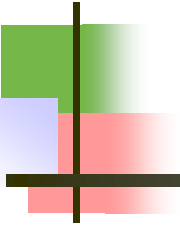
$\Xi_k = \{\text{長さが} k \text{以下} \text{の} s\text{-}t \text{単純経路}\}$

定理

任意の ϵ ($0 < \epsilon < 1$) と任意の δ ($0 < \delta < 1$) に対して,
 Ξ_k ($k = 2n, 2n + 2, \dots, L^*$) からの一様標本の数を
 $M = 12n^3(2n^2\epsilon^{-1})^2 \ln(n^2\delta^{-1})$ とすると, 近似解 Z は
$$\Pr[(1 - \epsilon)|\Omega| \leq Z \leq (1 + \epsilon)|\Omega|] \geq 1 - \delta$$

を満たす.

マルコフ連鎖のmixing timeは未解決



2行分割表

2行分割表のFPRAS

二元分割表

- ✓ 各セルには非負整数が入る
- ✓ (与えられた) 周辺和を満たす

						18
						12
5	4	3	7	5	6	30

5	4	3	6	0	0	18
0	0	0	1	5	6	12
5	4	3	7	5	6	30

状態A

4	3	1	3	1	6	18
1	1	2	4	4	0	12
5	4	3	7	5	6	30

状態B

0	0	0	7	5	6	18
5	4	3	0	0	0	12
5	4	3	7	5	6	30

状態C

問題

入力: 周辺和

出力: 分割表は何通り存在するか?

2行分割表のself-reduction [DG00,KM04]

- $\mathcal{C} = \mathcal{C}(\mathbf{r}, \mathbf{c})$ ($\mathbf{r}$ は行和, $\mathbf{c}$ は列和)
 - 一般性を失うことなく $r_1 \geq r_2$ とする.
- $\mathcal{C}^+ = \left\{ X \mid X_{1n} \geq \frac{c_n}{2} \right\}$
 - $\frac{|\mathcal{C}^+|}{|\mathcal{C}|}$ はモンテカルロで推定可.
- $|\mathcal{C}^+| = |\mathcal{C}(\mathbf{r}', \mathbf{c}')|$
 - ただし $r'_1 = r_1 - \left\lfloor \frac{c_n}{2} \right\rfloor$, $c'_n = c_n - \left\lfloor \frac{c_n}{2} \right\rfloor$

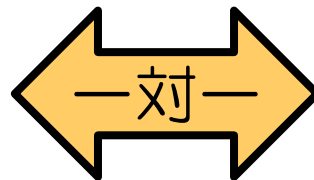
						18
						12
5	4	3	7	5	6	30

定理[KM04]: $\frac{|\mathcal{C}^+|}{|\mathcal{C}|} \geq \frac{1}{2}$

0	0	0	7	5	6	18
5	4	3	0	0	0	12
5	4	3	7	5	6	30

$X \in \mathcal{C}^+$

$$X_{1n} = x \Leftrightarrow X'_{1n} = a - 3$$



0	0	0	7	5	3	15
5	4	3	0	0	0	12
5	4	3	7	5	3	27

$X' \in \mathcal{C}(\mathbf{r}', \mathbf{c}')$

2行分割表のself-reduction [DG00,KM04]

- $\mathcal{C} = \mathcal{C}(\mathbf{r}, \mathbf{c})$ ($\mathbf{r}$ は行和, $\mathbf{c}$ は列和)
 - 一般性を失うことなく $r_1 \geq r_2$ とする.
- $\mathcal{C}^+ = \left\{ X \mid X_{1n} \geq \frac{c_n}{2} \right\}$
 - $\frac{|\mathcal{C}^+|}{|\mathcal{C}|}$ はモンテカルロで推定可.
- $|\mathcal{C}^+| = |\mathcal{C}(\mathbf{r}', \mathbf{c}')|$
 - ただし $r'_1 = r_1 - \left\lfloor \frac{c_n}{2} \right\rfloor$, $c'_n = c_n - \left\lfloor \frac{c_n}{2} \right\rfloor$
- $c_n = 1$ の時は列が減らせる

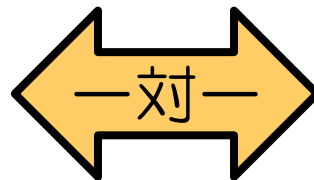
						13
						12
5	4	3	7	5	1	25

定理[KM04]: $\frac{|\mathcal{C}^+|}{|\mathcal{C}|} \geq \frac{1}{2}$

0	0	0	7	5	1	13
5	4	3	0	0	0	12
5	4	3	7	5	1	25

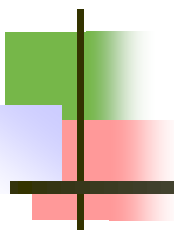
$X \in \mathcal{C}^+$

$$X_{1n} = x \Leftrightarrow X'_{1n} = x - 1$$



0	0	0	7	5	0	12
5	4	3	0	0	0	12
5	4	3	7	5	0	24

$X' \in \mathcal{C}(\mathbf{r}', \mathbf{c}')$



Bipartite matching

- 1996年Gödel prize
- 2004年解決 (Fulkerson Prize)

Bipartite matching

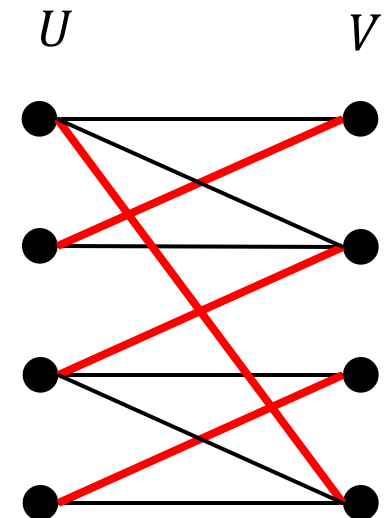
□ 二部グラフ: $G = (U, V; E)$. $|E| = m$ とする.

- 完全マッチングを考えるときは $|U| = |V| = n$ とする.

□ マッチング: $\mathcal{M} = \{M \subseteq E \mid \forall e, f \in M \ (e \neq f), e \cap f = \emptyset\}$

- サイズ k マッチング: $\mathcal{M}_k = \{M \mid |M| = k\}$; $\mathcal{M} = \bigcup_{k=0}^n \mathcal{M}_k$
- 完全マッチング: $\mathcal{P} = \mathcal{M}_n$

**Q. 二部グラフの完全マッチングの個数を求める
FPRASは存在するか? (1979~2004)**

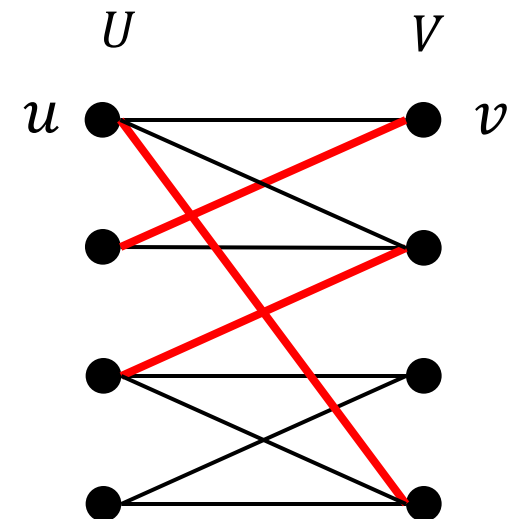


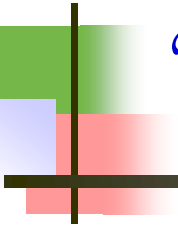
Bipartite matching

枝 $\{u, v\}$ に着目

$\{u, v\}$ を使う $\Leftrightarrow \# \{G[U \setminus \{u\}, V \setminus \{v\}] \text{ の bipartite matching} \}$

$\{u, v\}$ を使わない $\Leftrightarrow \# \{G[E \setminus \{u, v\}] \text{ の bipartite matching} \}$





The end

Thank you for the attention.

TAIL BOUND

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,

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

証明

$$\begin{aligned} \frac{E[X]}{a} &= E\left[\frac{X}{a}\right] \\ &= \int_0^{\infty} \frac{x}{a} f(x) dx \geq \int_a^{\infty} \frac{x}{a} f(x) dx \geq \int_a^{\infty} f(x) dx = \Pr[X \geq a] \end{aligned}$$

したがって

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



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,

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

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,

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

Chebyshev's inequality

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

任意の $t > 0$ に対して,

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

証明

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

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,
 $X := X_1 + \dots + X_n$ の期待値を μ とすると,

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

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

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

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

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

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

Maclaurin展開

独立性より

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

$t = \ln(1 + \varepsilon)$ とすると,

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

Hoeffding

定理 [Hoeffding; 二項分布ver.]

$X \sim B(n, p)$ とすると,

$\Pr[|X - np| > t] < 2 \exp(-2t^2/n)$
が成り立つ.

二項分布のHoeffding証明

$q = p + \epsilon$ とする. マルコフの不等式から

$$\begin{aligned}
 \Pr \left[\frac{\sum_{i=1}^n X_i}{n} > p + \epsilon \right] &= \Pr \left[\sum_{i=1}^n X_i > qn \right] \\
 &= \Pr \left[t \sum_{i=1}^n X_i > tqn \right] \\
 &= \Pr \left[\exp \left(t \sum_{i=1}^n X_i \right) > \exp(tqn) \right] \\
 &< \frac{E[\exp(t \sum_{i=1}^n X_i)]}{e^{tqn}} \quad (\text{by Markov's ineq.}) \\
 &= \frac{\prod_{i=1}^n E[\exp(tX_i)]}{e^{tqn}} \quad (\text{since } X_i \text{ are independent}) \\
 &= \frac{(E[\exp(tX_i)])^n}{e^{tqn}} \quad (\text{since } X_i \text{ are identical}) \\
 &= \frac{(pe^t + (1-p))^n}{e^{tqn}} \\
 &= (pe^{t(1-q)} + (1-p)e^{-tq})^n
 \end{aligned}$$

二項分布のHoeffding

$$t = \log \frac{(1-p)q}{(1-q)p} \text{ を代入}$$

$$\begin{aligned}
 (pe^{t(1-q)} + (1-p)e^{-tq})^n &= \left(p \left(\frac{(1-p)q}{(1-q)p} \right)^{1-q} + (1-p) \left(\frac{(1-p)q}{(1-q)p} \right)^{-q} \right)^n \\
 &= \left(\frac{(1-p)^{1-q} q^{1-q}}{(1-q)^{1-q} p^{-q}} + \frac{(1-p)^{1-q} q^{-q}}{(1-q)^{-q} p^{-q}} \right)^n \\
 &= \left(p^q (1-p)^{1-q} \frac{q^{1-q}}{(1-q)^{1-q}} + \frac{q^{-q}}{(1-q)^{-q}} \right)^n \\
 &= \left(p^q (1-p)^{1-q} \frac{q^{1-q}}{(1-q)^{1-q}} + \frac{(1-q)^q}{q^q} \right)^n \\
 &= \left(p^q (1-p)^{1-q} \frac{q + (1-q)}{(1-q)^{1-q} q^q} \right)^n \\
 &= \left(\left(\frac{p}{q} \right)^q \left(\frac{1-p}{1-q} \right)^{1-q} \right)^n \\
 &= \left(\left(\frac{p}{p+\epsilon} \right)^{p+\epsilon} \left(\frac{1-p}{1-p-\epsilon} \right)^{1-p-\epsilon} \right)^n
 \end{aligned}$$

二項分布のHoeffding証明

雑な変形をすると

$$\begin{aligned}
 \left(\left(\frac{p}{p+\epsilon} \right)^{p+\epsilon} \left(\frac{1-p}{1-p-\epsilon} \right)^{1-p-\epsilon} \right)^n &= \left(\frac{1}{\left(1 + \frac{\epsilon}{p} \right)^{\frac{p}{\epsilon}(\epsilon+\epsilon^2)}} \frac{1}{\left(1 - \frac{\epsilon}{1-p} \right)^{\frac{1-p}{\epsilon}(\epsilon-\epsilon^2)}} \right)^n \\
 &\simeq \left(\frac{1}{e^{(\epsilon+\epsilon^2)}} \frac{1}{e^{-(\epsilon-\epsilon^2)}} \right)^n \\
 &= e^{-2n\epsilon^2}
 \end{aligned}$$

二項分布のHoeffding証明

少し丁寧な変形をすると

$$f(\epsilon) = (p + \epsilon)(\log p - \log(p + \epsilon)) + (1 - p - \epsilon)(\log(1 - p) - \log(1 - p - \epsilon))$$

とおくと

$$\begin{aligned} f'(\epsilon) &= (\log p - \log(p + \epsilon)) - 1 - (\log(1 - p) - \log(1 - p - \epsilon)) + 1 \\ &= \log p - \log(p + \epsilon) - \log(1 - p) + \log(1 - p - \epsilon) \end{aligned}$$

$$f''(\epsilon) = -\frac{1}{p + \epsilon} - \frac{1}{1 - p - \epsilon} = -\frac{1}{(p + \epsilon)(1 - p - \epsilon)}$$

$$f'''(\epsilon) = \frac{1}{(p + \epsilon)^2} - \frac{1}{(1 - p - \epsilon)^2}$$

f のマクローリン展開より

$$\begin{aligned} f(\epsilon) &= f(0) + f'(0)\epsilon + \frac{1}{2}f''(0)\epsilon^2 + \frac{1}{3!}f'''(0)\epsilon^3 + o(\epsilon^3) \\ &= 0 + 0 + \frac{1}{2}\left(-\frac{1}{p(1-p)}\right)\epsilon^2 + \frac{1}{6}\frac{1-2p}{p^2(1-p)^2}\epsilon^3 + o(\epsilon^3) \\ &= -2\epsilon^2\left(\frac{1}{4p(1-p)} + \frac{(1-2p)\epsilon}{12p^2(1-p)^2}\right) + o(\epsilon^3) \leq -2\epsilon^2\left(\frac{1 + \frac{(1-2p)\epsilon}{3p(1-p)}}{4p(1-p)}\right) + o(\epsilon^3) \\ &\leq -2\epsilon^2 \end{aligned}$$

HoeffdingはKL divergenceを計算して証明している。

Hoeffding

定理 [Hoeffding; 二項分布ver.]

$X \sim B(n, p)$ とすると,

$\Pr[|X - np| > t] < 2 \exp(-2t^2/n)$
が成り立つ.

almost surely bounded

定理 [Hoeffding; 一般形]

$X_1, X_2, \dots, X_n$ は相互に独立で, 確率1で $a_i \leq X_i \leq b_i$ とする.

$Y_n := \sum_{i=1}^n X_i$ とすると, 任意の $t > 0$ に対して,

$$\Pr[|Y_n - E[Y_n]| > t] < 2 \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)}\right)$$

が成り立つ.