Chapter2:
確率的バンディット問題の基礎知識

機械学習プロフェッショナルシリーズ輪読会

バンディット問題の理論とアルゴリズム

@takeru0911

構成

- 2.1 中心極限定理による確率近似
- 2.2 裾確率の評価
- 2.3 大偏差原理

構成

- 2.1 中心極限定理による確率近似
- 2.2 裾確率の評価
- 2.3 大偏差原理

中心極限定理による確率近似

- バンディット問題でよくある問題
  - 広告のクリック率が$\hat{\mu}$が5%以下であるとき，
    その真のクリック率μが10%である可能性は？

中心極限定理による確率近似

- バンディット問題でよくある問題
  - 広告のクリック率が$\hat{\mu}$が5%以下であるとき，
    その真のクリック率μが10%である可能性は？

一般化

  - 真のクリック率がμであるとき，$x \in [0,1]$に対して
    標本平均が $\hat{\mu} \le x$となる確率は？

$\hat{\mu} \le x$となる確率が最も低い広告を表示したい！

中心極限定理による確率近似

- 「標本平均が $\hat{\mu} \leq x$ となる確率は？」をどうやって求める？

中心極限定理による確率近似

- 「標本平均が $\hat{\mu} \leq x$ となる確率は？」をどうやって求める？
  - →中心極限定理による近似だ！

中心極限定理による確率近似

- 中心極限定理

標準化された標本平均$\frac{\sqrt{n}(\hat{\mu}-\mu)}{\sigma}$ の分布は標準正規分布に弱収束する.
すなわち任意の$x \in \mathbb{R}$で次が成り立つ

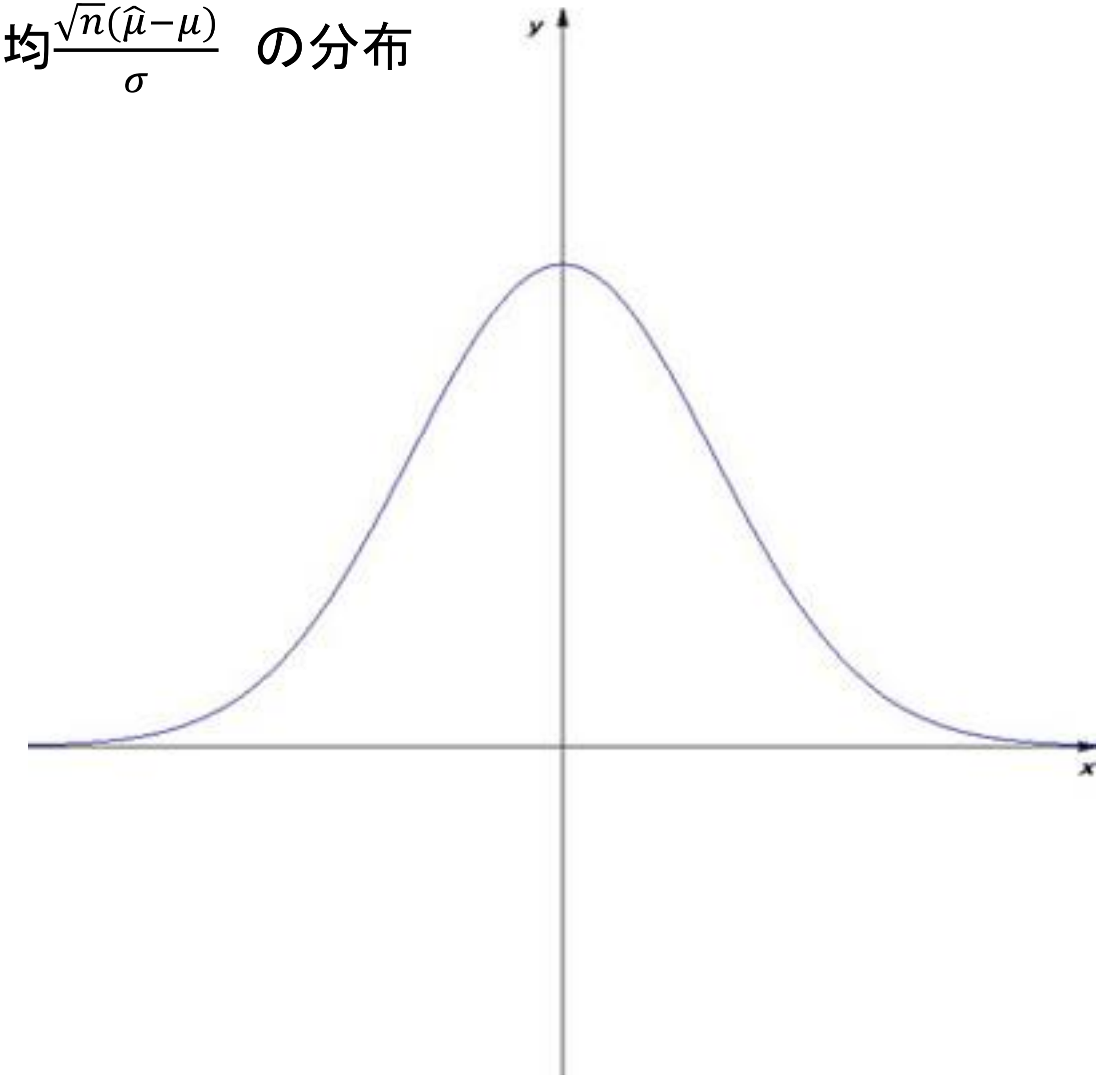
$$\lim_{n\to\infty} \mathbb{P}\left[\frac{\sqrt{n}(\hat{\mu}-\mu)}{\sigma} \leq x\right] = \Phi(x)$$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

ただし, $\Phi(x)$は標準正規分布の累積分布関数を表す

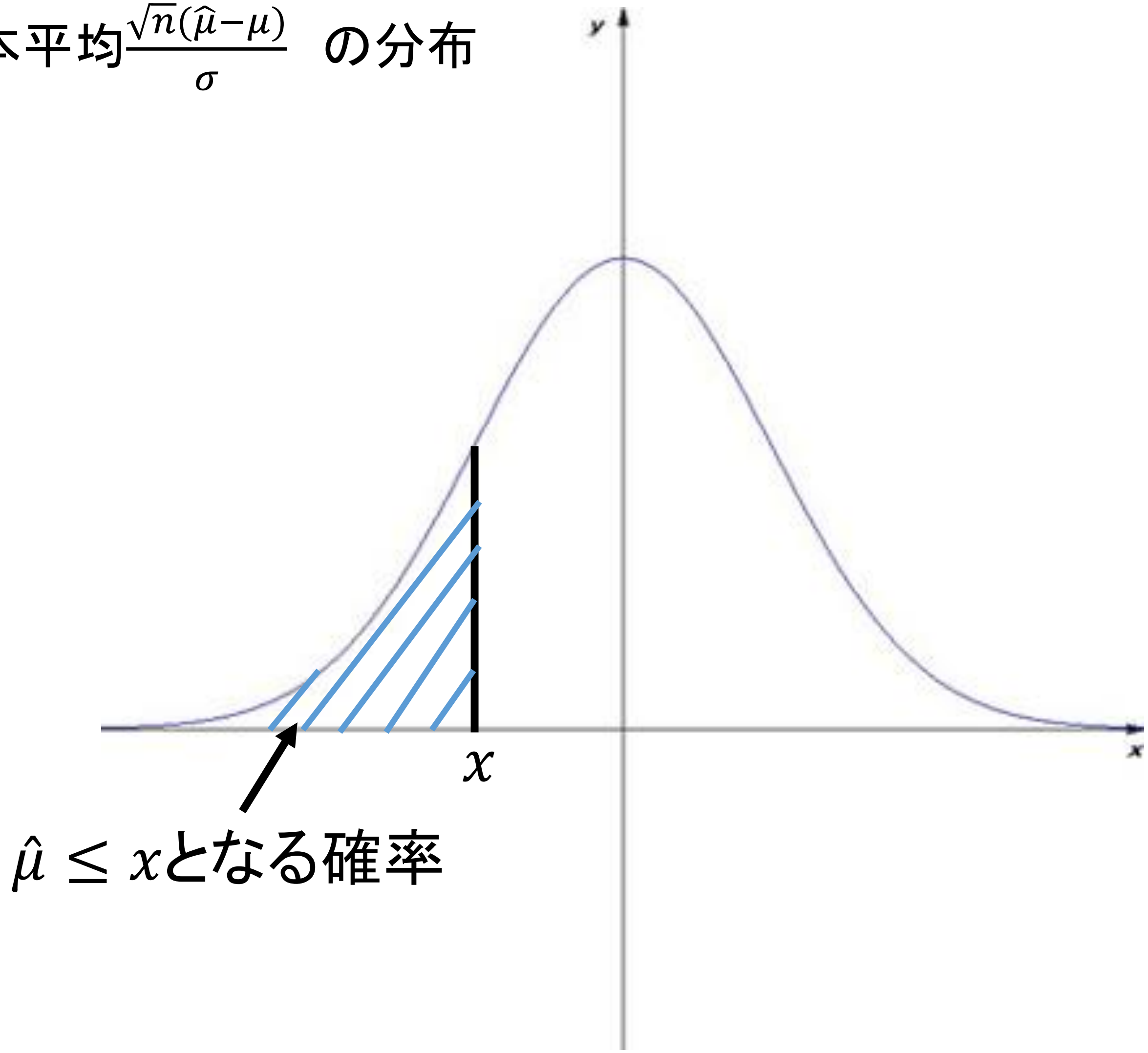
中心極限定理による確率近似

標本平均$\frac{\sqrt{n}(\hat{\mu}-\mu)}{\sigma}$ の分布



中心極限定理による確率近似

標本平均$\frac{\sqrt{n}(\hat{\mu}-\mu)}{\sigma}$ の分布



中心極限定理による確率近似

- しかし中心極限定理による近似では
  ベリー・エッセンの定理より
  $$\varepsilon = O(\frac{1}{\sqrt{n}})$$
  程度の絶対誤差があることが示されている。
  つまり誤差をε以下で近似するには
  $$n = O(\frac{1}{\varepsilon^2})$$
  程度のサンプル数が必要となる

中心極限定理による確率近似

- バンディット問題で扱う確率
  - 生起確率が小さな確率

  →小さな誤差で近似したい

  しかし，中心極限定理では誤差を
  小さくするには膨大なサンプルが必要

ちょっと中心極限定理による近似は適してなさそう･･･

構成

- 2.1 中心極限定理による確率近似
- 2.2 裾確率の評価
- 2.3 大偏差原理

裾確率の評価

- 中心極限定理では裾確率で発生するような事象の確率の評価は難しい

裾確率の評価

- 中心極限定理では裾確率で発生するような事象の確率の評価は難しい

  → ヘフディングの不等式

裾確率の評価
- ヘフディングの不等式 -

- 中心極限定理では裾確率で発生するような事象の確率の評価は難しい

  → ヘフディングの不等式

i.i.d. 確率変数$X_i \in [0,1]$と任意の$\Delta > 0$に対して,

$$\mathbb{P}[\hat{\mu} \leq \mu - \Delta] \leq e^{-2n\Delta^2}$$
$$\mathbb{P}[\hat{\mu} \geq \mu + \Delta] \leq e^{-2n\Delta^2}$$

が成り立つ

裾確率の評価
- ヘフディングの不等式 -

- 中心極限定理では裾確率で発生するような事象の確率の評価は難しい

　→ ヘフディングの不等式

真の平均と標本平均とのずれ

i.i.d. 確率変数$X_i \in [0,1]$と任意の$\Delta > 0$に対して,

$$\mathbb{P}[\hat{\mu} \leq \mu - \Delta] \leq e^{-2n\Delta^2}$$
$$\mathbb{P}[\hat{\mu} \geq \mu + \Delta] \leq e^{-2n\Delta^2}$$

が成り立つ

裾確率の評価
- ヘフディングの不等式 -

- 中心極限定理では裾確率で発生するような事象の確率の評価は難しい

→ ヘフディングの不等式

i.i.d. 確率変数$X_i \in [0,1]$と任意の$\Delta > 0$に対して,

$$\mathbb{P}[\hat{\mu} \le \mu - \Delta] \le e^{-2n\Delta^2}$$
$$\mathbb{P}[\hat{\mu} \ge \mu + \Delta] \le e^{-2n\Delta^2}$$

が成り立つ

X_iやμに依存せずに確率の評価が可能！

裾確率の評価

- ヘフディングの不等式により裾確率の評価ができた
  - とは言え精度は良くない・・・

裾確率の評価

- ヘフディングの不等式により裾確率の評価ができた
  - とは言え精度は良くない･･･

X_iやμに対する依存を許す

- チェルノ･ヘフディングの不等式（Chernoff-Hoeffding's inequality）
  - カルバック･ライブラー･ダイバージェンス（KLダイバージェンス）およびピンスカーの不等式を導入して精度を向上

裾確率の評価
-チェルノ･ヘフディングの不等式-

- KLダイバージェンス
  - 2つの確率分布の差異を測る指標
（厳密ではないと思います）
  - 以下の式で定義される

$$D(P||Q) = \sum_{x} P(x) \log \frac{P(x)}{Q(x)}$$

裾確率の評価
-チェルノ・ヘフディングの不等式-

- KLダイバージェンス
  - 2つの確率分布の差異を測る指標（厳密ではないと思います）
  - 以下の式で定義される

$$D(P||Q) = \sum_{x} P(x) \log \frac{P(x)}{Q(x)}$$

  - 分布間の距離の2乗に対応し，ピンスカーの不等式より

$$D(P||Q) \geq 2||P - Q||_1^2$$

が成り立つ

ここで

$$||P - Q||_1 = \frac{1}{2} \sum_{x} |P(x) - Q(x)|$$

裾確率の評価
-チェルノ・ヘフディングの不等式-

期待値pのベルヌーイ分布（Ber(p)）間のKLダイバージェンスを

$$d(p,q)=D\big(\mathrm{Ber}(p),\mathrm{Ber}(q)\big)=p\log\frac{p}{q}+(1-p)\log\frac{1-p}{1-q}$$

とすると,

裾確率の評価
-チェルノ・ヘフディングの不等式-

期待値pのベルヌーイ分布（Ber(p)）間のKLダイバージェンスを

$$d(p,q)=D\big(\mathrm{Ber}(p),\mathrm{Ber}(q)\big)=p\log\frac{p}{q}+(1-p)\log\frac{1-p}{1-q}$$

とすると，

チェルノ・ヘフディングの不等式

i.i.d. 確率変数$X_i\in[0,1]$と任意の$0\le x\le\mu$に対して，

$$\mathbb{P}[\hat{\mu}\le x]\le e^{-nd(x,\mu)}$$

が成り立ち，また任意の$\mu\le x\le 1$に対して，

$$\mathbb{P}[\hat{\mu}\ge x]\le e^{-nd(x,\mu)}$$

が成り立つ

裾確率の評価
-チェルノ・ヘフディングの不等式-

チェルノ・ヘフディングの不等式

i.i.d. 確率変数$X_i \in [0,1]$と任意の$0 \leq x \leq \mu$に対して,

$$\mathbb{P}[\hat{\mu} \leq x] \leq e^{-nd(x,\mu)}$$

が成り立ち，また任意の$\mu \leq x \leq 1$に対して,

$$\mathbb{P}[\hat{\mu} \geq x] \leq e^{-nd(x,\mu)}$$

が成り立つ

ここでピンスカーの不等式より

$$d(x,\mu) \geq 2(x-\mu)^2$$

となり，ヘフディングの不等式より精度が高いことがわかる

裾確率の評価
-チェルノ・ヘフディングの不等式-

チェルノ・ヘフディングの不等式

i.i.d. 確率変数$X_i \in [0,1]$と任意の$0 \leq x \leq \mu$に対して,

$$\mathbb{P}[\hat{\mu} \leq x] \leq e^{-nd(x,\mu)}$$

が成り立ち, また任意の$\mu \leq x \leq 1$に対して,

$$\mathbb{P}[\hat{\mu} \geq x] \leq e^{-nd(x,\mu)}$$

が成り立つ

ここでピンスカーの不等式より

$$d(x,\mu) \geq 2(x-\mu)^2$$

ヘフディングの不等式における∆

となり, ヘフディングの不等式より精度が高いことがわかる

構成

- 2.1 中心極限定理による確率近似
- 2.2 裾確率の評価
- 2.3 大偏差原理

大偏差原理

- ここまでは標本平均に関する裾確率の評価だった
- しかし標本分布に関しても確率の評価を行いたい

大偏差原理

- ここまでは標本平均に関する裾確率の評価だった
- しかし標本分布に関しても確率の評価を行いたい
  →サノフの定理（Sanov’s theorem）

大偏差原理
- サノフの定理 -

サノフの定理

$\mathbb{R}$上の確率分布全体の集合を$\mathcal{P}$とする。このとき任意の分布$P \in \mathcal{P}$および開集合$A \subset \mathcal{P}$，開集合$B \subset \mathcal{P}$に対して

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\hat{P} \in A\right] \geq -\inf_{Q\in A} D(Q||P)$$

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\hat{P} \in B\right] \geq -\inf_{Q\in B} D(Q||P)$$

が成り立つ

※ここで$\widehat{P_n}$は確率分布Pからのサンプルn個の標本分布

大偏差原理
- サノフの定理 -

サノフの定理

$\mathbb{R}$上の確率分布全体の集合を$\mathcal{P}$とする。このとき任意の分布$P \in \mathcal{P}$および開集合$A \subset \mathcal{P}$，閉集合$B \subset \mathcal{P}$に対して

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\hat{P} \in A\right] \geq -\inf_{Q\in A} D(Q||P)$$

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\hat{P} \in B\right] \geq -\inf_{Q\in B} D(Q||P)$$

が成り立つ

「分布Pからサンプリングされた集合$\hat{P}$が
分布Qからのものであるように振る舞う」確率を以下のように表す

$$\mathbb{P}[\hat{P}_n \approx Q] \approx e^{-nD(Q||P)}$$

大偏差原理

- ここまで話したような低確率で起こる事象を指数関数形で評価する理論体系
  - → 大偏差原理（large deviation principle）
- バンディット問題では，この原理を用いて確率の評価を行う

大偏差原理
-厳密漸近論-

- サンプル数nの多項式倍の誤差については考えなかった

大偏差原理
-厳密漸近論-

- サンプル数nの多項式倍の誤差については考えなかった
- 1次元確率変数の標本平均については以下の式が成り立つ

$$\lim_{n\to\infty} \frac{\mathbb{P}[\hat{\mu}_n \leq x]}{\frac{C}{\sqrt{n}} e^{-n \sup_{\lambda\leq 0}\{\lambda_x - \Lambda(\lambda)\}}} = 1$$

大偏差原理
-厳密漸近論-

- サンプル数nの多項式倍の誤差については考えなかった
- 1次元確率変数の標本平均については以下の式が成り立つ

標本平均がx以下になる確率

$$\lim_{n\to\infty} \frac{\mathbb{P}[\hat{\mu}_n \leq x]}{\frac{C}{\sqrt{n}} e^{-n \sup_{\lambda \leq 0}\{\lambda_x - \Lambda(\lambda)\}}} = 1$$

大偏差原理
-厳密漸近論-

- サンプル数nの多項式倍の誤差については考えなかった
- 1次元確率変数の標本平均については以下の式が成り立つ

標本平均がx以下になる真の確率

$$\lim_{n\to\infty} \frac{\mathbb{P}[\hat{\mu}_n \leq x]}{\frac{C}{\sqrt{n}} e^{-n\sup_{\lambda\leq 0}\{\lambda_x - \Lambda(\lambda)\}}} = 1$$

大偏差原理に基づき，近似された確率

大偏差原理
-厳密漸近論-

- サンプル数nの多項式倍の誤差については考えなかった
- 1次元確率変数の標本平均については以下の式が成り立つ

標本平均がx以下になる真の確率

$$\lim_{n\to\infty}\frac{\mathbb{P}[\hat{\mu}_n \leq x]}{\frac{C}{\sqrt{n}}e^{-n\sup_{\lambda\leq 0}\{\lambda x-\Lambda(\lambda)\}}}=1$$

大偏差原理に基づき，近似された確率

大偏差原理
-厳密漸近論-

- サンプル数nの多項式倍の誤差については考えなかった
- 1次元確率変数の標本平均については以下の式が成り立つ

$$\lim_{n\to\infty} \frac{\mathbb{P}[\hat{\mu}_n \leq x]}{\frac{C}{\sqrt{n}} e^{-n \sup_{\lambda \leq 0}\{\lambda_x - \Lambda(\lambda)\}}} = 1$$

定数Cを変化させることで真の確率に対して任意の精度で評価することができる
→ 厳密漸近論（exact asymptotics）

大偏差原理
-厳密漸近論-

- 例えば元の分布がベルヌーイ分布の場合・・・

$$\lim_{n\to\infty} \frac{\mathbb{P}[\hat{\mu}_n \leq x]}{\sqrt{\frac{1-x}{2\pi x n}\frac{\mu}{\mu - x}}\, e^{-nd(x,\mu)}} = 1$$

が成り立つ

※C= $\sqrt{\frac{1-x}{2\pi x}\frac{\mu}{\mu-x}}$, $\sup_{\lambda\leq 0}\{\lambda_x - \Lambda(\lambda)\} = d(x,\mu)$