2020.5.31(日)

情報統計力学 学習7

6月2日(火) 午後1時ー2時半

大偏差原理を解説する。

1 これまでのまとめ

$\{X_i\}$ を互いに独立で、同じ分布を持つ確率変数とする。平均値を μ、標準偏差を $\sigma(<\infty)$ とする。
$S_n = X_1 + X_2 + \cdots + X_n$ とすると、$E(\frac{S_n}{n}) = \mu, V(\frac{S_n}{n}) = \frac{\sigma^2}{n}$ である。

1.1 大数の強法則

$\frac{S_n}{n}$ は、期待値 μ に概収束する。すなわち、ほとんど全ての試行で $\frac{S_n}{n} \to \mu$ となる。
つまり、ほとんど全ての観測において、観測した平均値が真の平均値に収束する。

1.2 リンデベルグ＝レヴィ(Lindeberg-Lévy)の中心極限定理

$n \to \infty$ のとき、$\sqrt{n}\left(\frac{1}{n}S_n - \mu\right)$ は正規分布 $N(0, \sigma^2)$ に法則収束する。書き換えると、

$n \to \infty$ のとき、$\frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}S_n - \mu\right)$ は正規分布 $N(0,1)$ に法則収束する。つまり、
$n \to \infty$ のとき、

$$\frac{\sqrt{n}}{\sigma}\left(\frac{1}{n}S_n - \mu\right) \xrightarrow{d} Z \sim N(0,1). \tag{1}$$

よって、n が十分大きければ、

$$P\left(\frac{1}{n}S_n - \mu \in \frac{\sigma}{\sqrt{n}}[\alpha, \beta]\right) \sim \frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-\frac{1}{2}x^2}dx. \tag{2}$$

つまり、μ からのずれが $\frac{1}{\sqrt{n}}$ のオーダーの小さな値となる場合の確率を与えている。

それでは、μ からのずれが1のオーダーのように大きな場合はどうなるだろうか？つまり、平均値から大きく外れた値をとる確率についての理論が、大偏差原理である。

2 大偏差原理

まず、いくつか準備を行う。

2.1 凸関数

$y = f(x)$ が区間 I 上の凸関数とは、次の関係を満足する場合。
任意の $\lambda \in [0,1]$ と,$x_1, x_2 \in I$ に対して、

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda)f(x_2) \tag{3}$$

が成り立つ。$f''(x) > 0$ ならば、凸関数である。

2.2 イェンゼンの不等式

$f(x)$ が凸関数ならば、

$$E(f(X)) \quad \ge \quad f(E(X)). \tag{4}$$

2.3 ルジャンドル変換

$y = f(x)$ を凸関数とする。$f''(x) > 0$ とする。新しい変数 p を導入する。p を固定して、直線 $y = px$ と曲線 $y = f(x)$ の、同一の x における値の差 $F(p,x) = px - f(x)$ の最大値を与える点を $x(p)$ とし、そのときの差の値を $g(p)$ とする。これを、$f(x)$ の**ルジャンドル変換**という。

$$\frac{\partial}{\partial x}F(p,x) = 0, \tag{5}$$

$$p - f'(x) = 0,\ p = f'(x). \tag{6}$$

$$\text{つまり、} p \text{ が決まれば、} x \text{ は、} p = f'(x) \text{ となる点。 x} = x(p) \text{ とする。} \tag{7}$$

$f''(x) > 0$ であるから、$\frac{\partial^2}{\partial x^2}F(p,x) = -f''(x) < 0$ なので、$x(p)$ は存在すれば一意的である。

$$\begin{aligned}
& p \text{ と } x \text{ は、} p = f'(x) \text{ という関係で 1 対 1 対応する。} && (8)\\
g(p) &= F(p,x(p)) = px(p) - f(x(p)), && (9)\\
g'(p) &= x(p) + px'(p) - f'(x(p))x'(p) = x(p) + px'(p) - px'(p) = x(p), && (10)\\
g''(p) &= x'(p), && (11)\\
p &= f'(x(p)) \text{ より、} 1 = f''(x(p))x'(p), \quad x'(p) = 1/f''(x(p)) > 0, && (12)\\
\text{よって、} g''(p) &= 1/f''(x(p)) > 0. && (13)
\end{aligned}$$

つまり、$g(p)$ も凸関数である。次に、$g(p)$ のルジャンドル変換を $h(y)$ 考える。y を与えたとき、$G(y,p) = yp - g(p)$ の最小値を求める。

$$\frac{\partial}{\partial p}G(y,p) = 0, \tag{14}$$

$$y - g'(p) = 0,\ y = g'(p),\ p = p(y), \tag{15}$$

つまり、y を与えたときに、$G(y,p) = yp - g(p)$ の最小値をあたえる p は、$y = g'(p)$ を満たす。

$$y \text{ と } p \text{ は、} y = g'(p) \text{ という関係で 1 対 1 対応する。} \tag{16}$$

$$\text{ところで、} g'(p) = x(p) \text{ は一意的なので、} y = x(p), \tag{17}$$

$$h(y) = G(y,p(y)) = yp(y) - g(p(y)) = yp(y) - \Bigg(py - f(y)\Bigg) = f(y). \tag{18}$$

つまり、ルジャンドル変換を 2 度行うと、もとの関数に戻る。

独立同分布族 $\{X_1, X_2, \cdots, X_n, \cdots\}$ について、次の積率母関数が有限であるとする。

$$M(t) \quad = \quad E(e^{tX_1}) < \infty,\ t \in \mathbb{R}. \tag{19}$$

さて、

$$c(t) \quad \equiv \quad \lim_{n\to\infty} \frac{1}{n} \ln E(e^{tS_n}) = \ln M(t) \tag{20}$$

と定義する。$E(e^{tS_n}) = M(t)^n$ であるから、$c(t) = \ln M(t) < \infty.$
$c(t)$ は凸関数であることを示す。$E_l = E[(X_1)^l e^{tX_1}]$ とおく。

$$c'(t) \quad = \quad \frac{E[X_1 e^{tX_1}]}{M(t)} = \frac{E_1}{E_0}, \tag{21}$$

$$c''(t) \quad = \quad \frac{E[X_1^2 e^{tX_n}]}{E_0} - \frac{E_1^2}{E_0^2} = \frac{E_2}{E_0} - \frac{E_1^2}{E_0^2}. \tag{22}$$

一方、

$$\begin{aligned} E[\left(X_1 - \frac{E_1}{E_0}\right)^2 \frac{e^{tX_1}}{E_0}] \quad &= \quad E[\left(X_1^2 - 2X_1\frac{E_1}{E_0} + (\frac{E_1}{E_0})^2\right)\frac{e^{tX_1}}{E_0}] \\ &= \quad \frac{E_2}{E_0} - 2\frac{E_1^2}{E_0^2} + (\frac{E_1}{E_0})^2 = \frac{E_2}{E_0} - \frac{E_1^2}{E_0^2}. \end{aligned}$$

左辺は、正の値を持つから、$c''(t) > 0.$ よって、$c(t)$ は凸関数であるから、その**ルジャンドル変換**を考える。

$$I(x) \quad = \quad \sup_{t\in\mathbb{R}} \{tx - c(t)\}. \tag{23}$$

$I(x)$ は**レート関数**と呼ばれる。

$I(x)$ **の性質**。
$I(x) \geq 0, x = \mu$ で最小値 0 を持つ。
証明

$$c(0) \quad = \quad \ln M(0) = \ln 1 = 0.$$

また、$I(x)$ は、$tx - c(t)$ の上限だから、$t = 0$ を入れた値よりも小さくない。つまり、0 より小さくないので、$I(x) \geq 0.$ 一方、$f = -\ln x$ は、$f''(x) = \frac{1}{x^2} > 0$ なので、凸関数だからイェンゼンの不等式より、

$$\begin{aligned} -\ln E[e^{tX_1}] \quad &\leq \quad E[-\ln e^{tX_1}], \\ \ln E[e^{tX_1}] \quad &\geq \quad E[\ln e^{tX_1}] = tE[X_1] = t\mu, \\ c(t) \geq t\mu, & \\ I(\mu) \quad &= \quad \sup_{t\in\mathbb{R}}\{t\mu - c(t)\} \leq 0. \end{aligned}$$

ところが、$I(\mu) \geq 0$ なので、$I(\mu) = 0.$

2.4 クラメール (Cramér) の定理

$$\lim_{n\to\infty}\frac{1}{n}\ln\left(P(\frac{S_n}{n}\in[a,b])\right) \quad = \quad -\inf_{[a,b]} I(x) \tag{24}$$

つまり、試行回数を十分大きくしたとき、試行の平均が区間 $[a,b]$ に入っている確率は、

$$P(\frac{S_n}{n}\in[a,b]) \quad \sim \quad e^{-n\inf_{[a,b]} I(x)} \tag{25}$$

となる。

証明は省略して、具体例を考える。

2.5 例 ランダムフォーク

直線上を運動する質点を考える。但し、移動は ±1 とし、左に移動する確率を p、右に移動する確率を q とする。$0<p,q<1, p+q=1$. 最初、原点にいるとする。i 回目の移動を確率変数 X_i で表す。各回の移動は独立であるとする。

$$P(X_i=-1)=p,\ P(X_i=1)=q, \tag{26}$$
$$E(X_i)=-p+q,\ E(X_i^2)=p+q=1,\ V(X_i)=1-(q-p)^2=4pq=4p(1-p). \tag{27}$$

積率母関数 $M(t), c(t)$ は定義より次のようになる。

$$\begin{aligned} M(t) &= E(e^{tX_1})=pe^{-t}+qe^t, \\ c(t) &= \ln M(t)=\ln(pe^{-t}+qe^t). \end{aligned}$$

レート関数 $I(x)$ を計算する。

$$I(x) \quad = \quad \sup_{t\in\mathbb{R}}\left(xt-\ln(pe^{-t}+qe^t)\right). \tag{28}$$

$h(t)\equiv xt-\ln(pe^{-t}+qe^t)$ とする。

1. $-1<x<1$ のとき。
   $h'(t)=0$ より、$t=\ln\sqrt{\frac{p(1+x)}{q(1-x)}}$ となる。従って、

$$I(x) \quad = \quad x\ln\sqrt{\frac{p(1+x)}{q(1-x)}}-\ln\left(p\sqrt{\frac{q(1-x)}{p(1+x)}}+q\sqrt{\frac{p(1+x)}{q(1-x)}}\right). \tag{29}$$

2. $x=-1$ のとき。

$$\begin{aligned} I(-1) &= \sup_{t\in\mathbb{R}}\left(-t-\ln(pe^{-t}+qe^t)\right)=\sup_{t\in\mathbb{R}}\left(\ln\frac{e^{-t}}{pe^{-t}+qe^t}\right) \\ &= \sup_{t\in\mathbb{R}}\left(\ln\frac{1}{p+qe^{2t}}\right) \end{aligned} \tag{30}$$

   $t=-\infty$ のときが上限値。よって、

$$I(-1) \quad = \quad \ln\frac{1}{p}=-\ln p. \tag{31}$$

3. $x=1$ のとき。

$$
\begin{aligned}
I(-1) &= \sup_{t\in\mathbb{R}}\left(t-\ln(pe^{-t}+qe^{t})\right)=\sup_{t\in\mathbb{R}}\left(\ln\frac{e^{t}}{pe^{-t}+qe^{t}}\right)\\
&= \sup_{t\in\mathbb{R}}\left(\ln\frac{1}{pe^{-2t}+q}\right)
\end{aligned}
\tag{32}
$$

$t=\infty$ のときが上限値。よって、

$$
I(-1) = \ln\frac{1}{q}=-\ln q. \tag{33}
$$

4. $x>1$、または $x<-1$ のとき。
   $x>1$ なら、$t\gg 1$ のとき、$h(t)\sim tx-t$ であるから、$I(x)=\infty$.
   $x<-1$ なら、$-t\gg 1$ のとき、$h(t)\sim tx+t$ であるから、$I(x)=\infty$.

$p=q=0.5$ の場合。つまり、右と左の移動が同確率で完全にランダムな場合。

1. $-1<x<1$ のとき。

$$
I(x) = x\ln\sqrt{\frac{1+x}{1-x}}-\ln\left(\frac{1}{2}\sqrt{\frac{1-x}{1+x}}+\frac{1}{2}\sqrt{\frac{1+x}{1-x}}\right). \tag{34}
$$

2. $x=\pm 1$ のとき。

$$
I(-1) = I(+1)=\ln 2. \tag{35}
$$

3. $x>1$、または $x<-1$ のとき。
   $I(x)=\infty$.

右にのみ移動し続ける確率を求めよう。$S_n=n$ であるから、$\frac{S_n}{n}=1$.

$$
P(\frac{S_n}{n}=1) \sim e^{-nI(1)}=2^{-n} \tag{36}
$$

で、ほとんど有り得ないことがわかる。$n\to\infty$ とすると、確率０。