漸近展開へのいざない

秋田 智之[*]

平成 21 年 3 月 24 日

概 要

　このセミナーおよび本稿では統計量の分布の漸近展開について紹介する．漸近展開の必要性がわかりやすいように仮説検定問題を導入に使い，項別積分などの数学的な厳密性はひとまず置いて大雑把な流れを解説することに留めている．なお，本稿は統計サマーセミナー 2008（統計数理研究所 共同研究集会 課題番号 20-共研-5008）の予稿から一部抜粋・加筆したものである．

Key Word：漸近展開，キュムラント，エッジワース展開，エルミート多項式，コーニッシュ・フィッシャー展開

導入 –仮説検定問題–

仮説検定では必ずといってよいほど扱われる次の問題を考える．

例 1.

$X_1, X_2, \ldots, X_n$ を母平均 μ，母分散 σ^2 の **正規分布** からの独立標本とする．このとき帰無仮説 $H_0 : \mu = \mu_0$ v.s. 対立仮説 $H_1 : \mu > \mu_0$ の検定問題を考えよう．但し，σ^2 は既知であるとする．

この仮説検定問題では正規分布の再生性により標本平均 $\bar{X}$ の分布がわかるので，次のような方法で検定を行うことが出来る：

検定法：$X_1, X_2, \ldots, X_n$ は $N(\mu, \sigma^2)$ に従うので，$S_n = \dfrac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$（ここで $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$）は 標準正規分布 $N(0,1)$ に従う．今，$N \sim N(0,1)$ として z_α を $P(N > z_\alpha) = \alpha$ を満たす上側 α 点とする．実際の標本平均 $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$ に対して，$\dfrac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}} > z_\alpha$ が成り立っていれば H_0 を棄却し，成り立っていなければ H_0 を採択する．

ここで S_n の分布が 正確に 標準正規分布 $N(0,1)$ に従っている点に注意する．次に以下の仮説検定問題を考える．

---

[*]広島大学大学院理学研究科数学専攻 D1

例 2.

$X_1, X_2, \dots, X_n$ を母平均 μ，母分散 σ^2 の **正規分布とは限らない母集団分布** からの独立標本とする．このとき帰無仮説 $H_0 : \mu = \mu_0$ v.s. 対立仮説 $H_1 : \mu > \mu_0$ の検定問題を考えよう．但し，σ^2 は既知であるとする．

この場合は先程と状況が変っている．統計量 S_n の分布は未知であるかも知れないし，分かっていても扱いにくいものであるかもしれない．しかし，中心極限定理により，S_n の漸近分布が $N(0,1)$ であることは分かる．従って n が大きいときは先と同様に次の検定方式が得られる．

検定法：$X_1, X_2, \dots, X_n$ は母平均 μ，母分散 σ^2 の確率分布に従うので，$S_n = \dfrac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$（ここで $\bar{X} = n^{-1}\sum_{i=1}^n X_i$）は漸近的に標準正規分布 $N(0,1)$ に従う．実際の標本平均 $\bar{x} = n^{-1}\sum_{i=1}^n x_i$ に対して，$\dfrac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}} > z_\alpha$ が成り立っていれば H_0 を棄却し，成り立っていなければ H_0 を採択する．

では n が大きくない場合はどうだろうか．S_n の正確な分布を求めにくい場合は漸近展開により分布を近似できる場合がある．

キュムラントとモーメントの関係

漸近展開の方法は Z_n の特性関数を特別な形で展開し，反転公式を用いて分布関数 $P(Z_n \le x)$ の展開式を求めるものである．これについて以下で具体的に見ていく．まずキュムラントを定義する．

キュムラント.

X を確率変数，$\varphi(t) = E[e^{itX}]$ を X の特性関数とする．$E(|X|^n) < \infty$ のとき $\log\varphi(t)$ は次の展開が出来る：

$$\log\varphi(t) = \sum_{j=1}^n \frac{\kappa_j}{j!}(it)^j + o(t^n) \quad (t \to 0)$$

上を満たす κ_j を X の j **次キュムラント**という．

一方，$\varphi(t)$ はモーメントを用いても表すことが出来る：

$$\varphi(t) = \sum_{j=0}^n \frac{E(X^j)}{j!}(it)^j + o(t^n) \quad (t \to 0)$$

従って $\log\varphi(t)$ は次のように表すことも出来る：

$$\log\varphi(t) = \log\left\{1 + \sum_{j=1}^n \frac{E(X^j)}{j!}(it)^j + o(t^n)\right\} = \sum_{k=1}^\infty \frac{(-1)^{k+1}}{k}\left(\sum_{j=1}^n \frac{(it)^j}{j!}E(X^j) + o(t^n)\right)^k$$

この係数を比較することにより

$$
\begin{aligned}
\kappa_1 &= E(X) \\
\kappa_2 &= E(X^2) - E(X)^2 = \mathrm{Var}(X) \\
\kappa_3 &= E(X^3) - 3E(X^2)E(X) + 2(EX)^3 \\
&= E(X - E(X))^3 \\
\kappa_4 &= E(X^4) - 4E(X^3)E(X) - 3(EX^2)^2 + 12E(X^2)(EX)^2 - 6(EX^2)^2 \\
&= E(X - EX)^4 - 3(\mathrm{Var}(X))^2
\end{aligned}
$$

が分かる．

エッジワース展開とエルミート多項式

次に以下の関係に注意する．

補題 1. $\varphi_n(t)$ を S_n の特性関数，$\varphi(t)$ を基準化した $Y = (X-\mu)/\sigma$ の特性関数とするとき，$\varphi_n(t) = \{\varphi(t/\sqrt{n})\}^n$ が成り立つ．

証明. 次の変形

$$
S_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}}{\sigma} \cdot \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma}
$$

に注意する．

$$
\varphi_n(t) = E[\exp(itS_n)] = E\left[\exp\left(it\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i - \mu}{\sigma}\right)\right] = \prod_{i=1}^{n} E\left[\exp\left(i\frac{t}{\sqrt{n}}Y_i\right)\right]
$$

この右辺の積の各項が $\varphi(t/\sqrt{n})$ であることから結論を得る．

□

従って Y の特性関数が分かれば S_n の特性関数も分かる．ところで Y は基準化されているので $E(Y) = 0$, $\mathrm{Var}(Y) = 1$ である．これより

$$
\log\varphi(t) = -\frac{1}{2}t^2 + \sum_{j=3}^{n} \frac{\kappa_j}{j!}(it)^j + o(t^n)
$$

が分かる．従って

$$
\begin{aligned}
\log\varphi_n(t) = n\log\varphi(t/\sqrt{n}) &= n\left\{-\frac{1}{2}\left(\frac{t}{\sqrt{n}}\right)^2 + \sum_{j=3}^{n}\frac{\kappa_j}{j!}\left(\frac{it}{\sqrt{n}}\right)^j + o(t^n)\right\} \\
&= -\frac{1}{2}t^2 + \sum_{j=3}^{n}\frac{\kappa_j}{j!}(it)^j n^{-j/2+2} + o(t^n)
\end{aligned}
$$

である．これより

$$
\begin{aligned}
\varphi_n(t) &= \exp\left\{-\frac{1}{2}t^2 + \sum_{j=3}^{n} \frac{\kappa_j}{j!}(it)^j n^{-j/2+2} + o(t^n)\right\} \\
&= e^{-t^2/2} \exp\left\{\sum_{j=3}^{n} \frac{\kappa_j}{j!}(it)^j n^{-j/2+2} + o(t^n)\right\} \\
&= e^{-t^2/2} \left\{1 + \sum_{k=1}^{\infty} \left(\sum_{j=3}^{n} \frac{\kappa_j}{j!}(it)^j n^{-j/2+2} + o(t^n)\right)^k\right\} \\
&= e^{-t^2/2} \left\{1 + n^{-1/2}\frac{1}{6}\kappa_3(it)^3 + n^{-1}\left(\frac{1}{24}\kappa_4(it)^4 + \frac{1}{72}\kappa_3^2(it)^6\right) + \cdots\right\} \\
&= e^{-t^2/2} + n^{-1/2}e^{-t^2/2}r_1(it) + n^{-1}e^{-t^2/2}r_2(it) + \cdots
\end{aligned}
$$

が得られる．$\phi(x)$ の特性関数は $e^{-t^2/2}$ であるから，反転公式より

$$
\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}e^{-t^2/2}dt = \phi(x)
$$

であり，更に項別微分をすることにより

$$
\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}(-it)^j e^{-t^2/2}dt = \phi^{(j)}(x)
$$

も分かる．

従って，反転公式から S_n の密度関数を $f_n(x)$ としたとき，

$$
\begin{aligned}
f_n(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\varphi_n(t)dt \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}e^{-t^2/2}\left\{1 + n^{-1/2}\frac{1}{6}\kappa_3(it)^3 + n^{-1}\left(\frac{1}{24}\kappa_4(it)^4 + \frac{1}{72}\kappa_3^2(it)^6\right) + \cdots\right\}dt \\
&= \phi(x) - n^{-1/2}\frac{\kappa_3}{6}\phi^{(3)}(x) + n^{-1}\left(\frac{1}{24}\kappa_4\phi^{(4)}(x) + \frac{1}{72}\kappa_3^2\phi^{(6)}(x)\right) + \cdots \\
&= \phi(x) + n^{-1/2}\phi(x)\tilde{r}_1(x) + n^{-1}\phi(x)\tilde{r}_2(x) + \cdots
\end{aligned}
$$

となり，S_n の分布関数も

$$
\begin{aligned}
P(S_n \le x) &= \int_{-\infty}^{x} f_n(s)ds \\
&= \Phi(x) - n^{-1/2}\frac{\kappa_3}{6}\phi^{(2)}(x) + n^{-1}\left(\frac{1}{24}\kappa_4\phi^{(3)}(x) + \frac{1}{72}\kappa_3^2\phi^{(5)}(x)\right) + \cdots \\
&= \Phi(x) + n^{-1/2}\phi(x)R_1(x) + n^{-1}\phi(x)R_2(x) + \cdots
\end{aligned}
$$

と展開できる．これを**エッジワース展開**という．$r, \tilde{r}$ および R はいずれも多項式で，次に挙げるエルミート多項式で表すことが出来る．

エルミート多項式.

次の関係式：

$$\phi(x)h_j(x) = (-1)^j \frac{d^j}{dx^j}\phi(x)$$

を満たす多項式 $h_j(x)$ を**エルミート多項式**という.

コーニッシュ・フィッシャー展開

冒頭の仮説検定問題では%点を求める必要がある．最後にこの求め方に触れる．w_α を S_n の上側 α%点，z_α を標準正規分布の上側 α%点とする．このとき，適当な正則条件のもとで以下の展開が出来る：

$$w_\alpha = z_\alpha + n^{-1/2}p_1(z_\alpha) + n^{-1}p_2(z_\alpha) + \cdots$$

ここで p_j はある多項式である．これを**コーニッシュ・フィッシャー展開**と言う．以下ではこの多項式を具体的に求める．w_α をエッジワース展開の式に代入すると

$$\begin{aligned}
&\alpha = \Phi(z_\alpha + n^{-1/2}p_1(z_\alpha) + n^{-1}p_2(z_\alpha) + \cdots) \\
&\quad + n^{-1/2}\phi(z_\alpha + n^{-1/2}p_1(z_\alpha) + n^{-1}p_2(z_\alpha) + \cdots)\cdot R_1(z_\alpha + n^{-1/2}p_1(z_\alpha) + n^{-1}p_2(z_\alpha) + \cdots) \\
&\quad + n^{-1}\phi(z_\alpha + n^{-1/2}p_1(z_\alpha) + n^{-1}p_2(z_\alpha) + \cdots)\cdot R_2(z_\alpha + n^{-1/2}p_1(z_\alpha) + n^{-1}p_2(z_\alpha) + \cdots) + \cdots
\end{aligned}$$

ここで Φ, R_j, ϕ を z_α の周りでテーラー展開し，n の冪で整理すると

$$\begin{aligned}
\alpha = \alpha &+ n^{-1/2}\{p_1(z_\alpha) + R_1(z_\alpha)\}\phi(z_\alpha) \\
&+ n^{-1}[p_2(z_\alpha) - \frac{1}{2}z_\alpha p_1(z_\alpha)^2 p_1(z_\alpha)\{R_1'(z_\alpha) - z_\alpha R_1(z_\alpha)\} + R_2(z_\alpha)]\phi(z_\alpha) + \cdots
\end{aligned}$$

を得る．これより

$$p_1(x) = -R_1(x), \quad p_2(x) = R_1(x)R_1'(x) - \frac{1}{2}xR_1(x)^2 - R_2(x)$$

が導かれる.

まとめ

中心極限定理により，統計量の漸近分布が正規分布や χ^2-分布になることが多い．漸近展開では，特性関数を先にあげた標準正規分布の反転公式または χ^2-分布の反転公式

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}(1-2it)^{-f/2}dt = g_f(x), \quad (x > 0)$$

が使える形に展開することがキーである．また，母集団のモーメントやキュムラントを求める必要がありその導出が面倒だったり，剰余項が収束するかなどの問題がある.