



入門講座

品質管理のための統計的方法の活用－3

検定と推定 (1)

Statistical Hypothesis Test and Estimation (1)

竹士伊知郎

Ichiro Chikushi

QMビューローちくし 代表
(一財) 日本科学技術連盟 嘱託
関西大学 化学生命工学部
非常勤講師

前号 (Vol.28 No.6 P372) からの続き

5 検定と推定

品質管理を行ううえで重要な目的のひとつは、母集団に関する調査であった。では、その調査の結果をどうあらわすのだろうか？

製造工程における調査なら「Q製品製造工程における製品寸法の平均値は10.00mmである」、「R工場の設備更新後の不適合品率は減少した」などだろうか。

しかし、これらの調査結果は、いずれも対象となる母集団をすべて調べたものではなく、サンプルの調査や測定から得られた情報である。ということは、先に何度も述べたようにサンプルはとるたびに違うものだから、これらの結果は、「たまたま今回そうなっただけでは？」といわれるかもしれない。

だれがみても問題のない、だれもが納得してくれる結果報告をしたいものである。職場の上司やお客様にも、堂々と報告できる調査結果の導き方、それがすなわち検定・推定の極意といえる。

5.1 検定

5.1.1 検定とは

検定は「母集団の平均は従来とは異なる」、「新たな工程では不純物量が減少した」などといった母集団に関することをサンプルから得られたデータで判断するものである。

以下に手順とその基本的な考え方を示す。

1) はじめに主張したい結論を掲げる。もちろんこの段階では、その結論が正しいかどうかわからないので仮説となる。仮説は誤っているかもしれないので、はじめに立てた仮説を否定する仮説も同時に用意しておく。

さきの例でいえば、「仮説A: 母集団の平均は従来と異なる」

とそれを否定する「仮説B: 母集団の平均は従来と等しい」というふたつの仮説になる。

2) 仮説を判定するのだから間違ふことがある。仮説がふたつあるので間違ふ方も2種類考えられる。すなわち、「本当は仮説Bが正しいのに仮説Aが正しいと判定してしまう間違い」と「本当は仮説Aが正しいのに仮説Bが正しいと判定してしまう間違い」である。

間違いがしょっちゅう起こっては信用をなくすので、これらの間違いが起こる確率をあらかじめ決めておく。この確率は、通常5%や1%といった小さい値が使われる。

では、結論としていいたいのは仮説Aであったから、「本当は仮説Bが正しいのに仮説Aが正しいと判定してしまう間違い」の確率を5%としておこう。こうしておけば仮説Aが正しいという判定がでたときに、「その判定が誤っている確率は5%という小さな確率で、めったに起こらない」ということがいえる。逆にいうと、その判定結果は概ね信用してよいというお墨付きが与えられることになる。

3) いよいよサンプルをとって、得られたデータの平均値を求める。その前に、このサンプルは仮説Bの母集団（すなわち従来と同じ平均をもつ集団）からとられたものとすれば（仮説Bが正しいとすれば）、4章で述べたとおり、平均値はどのような分布をするのかを知ることができる。さらにこれを正規分布の標準化をすれば標準正規分布に従うので、これを判定の基準にすればよいのである。正規分布表（付表1 Vol.28 No.5 P311）を使えば、この標準正規分布の値とそのときの確率の関係を知ることができるので、めったに起こらない（すなわち小さな確率）正規分布の値がわかる。この値を「めったに起こらないこと」と判定する境界にすればよい。

4) サンプルから得られたデータを計算した平均値などを使って標準正規分布の値を計算し、さきほどの境界の値と比べる。

境界を越えたとすれば、それは「めったに起こらないことが起こっている」という状況を示すことになる。

しかし、ここはそうは考えずに、「最初の仮説に理由がありそうだ」と考えるほうがよさそうである。

なぜなら母集団から正しくサンプルをとり、そのサンプルから平均値を求め、その分布を決めるという一連の流れは、いつだれがやっても同じようにでき、同じ結果になるはずだからである。

では、「最初の仮説に理由がある」とはどういうことか考えてみよう。今回の場合、仮説Bが正しいということを前提に進めてきたので、「仮説Bを正しいとしたことが間違いだった」とすれば自然である。すなわち、「母集団の平均は従来と等しい」ということが否定されたことになり、もう一つの仮説である仮説Aの「母集団の平均は従来と異なる」が正しかったということになる。

これは最初に掲げた結論と同じであり、思惑通りの結論を導くことができた。

もちろん、いつも、こうはうまくはいかない。境界から外れない場合もある。この場合は、最初に掲げた結論は正しいとはいえないので、「母集団の平均値は従来と異なるとはいえない」という結論になる。

5.1.2 検定の手順

(1) 検定の概要

先に示した検定の基本的な考え方に沿ってもう一度検定の概要を整理する。検定では、普段聞きなれない統計独特の用語を用いるので、その点にも注意してほしい。

検定とは、母集団の分布に関する仮説を統計的に検証するものである。サンプルやそのデータを検証するものではなく、母集団に関する仮説を、データを用いて検証することが目的である。

検定においては、主張したいことを対立仮説 (H_1 と表現する) に置き、この仮説を否定する仮説を帰無仮説 (H_0 と表現する) とする。対立仮説には、両側仮説と片側仮説とがあり、それぞれの場合の検定を両側検定、片側検定という。

表3-1 検定の仮説と誤りの確率 (Online version in color.)

判断 真実 \	H_0 が正しい	H_1 が正しい
H_0 が真	○	× 第1種の誤り 確率: α (有意水準)
H_1 が真	× 第2種の誤り 確率: β	○ 検出力 確率: $(1-\beta)$

帰無仮説が真であるにもかかわらず、対立仮説が真であると判断してしまう誤りを、第1種の誤り (過誤)、またはあわてものの誤りとよび、その確率を有意水準、危険率などといった記号 α であらわす。これに対し、対立仮説が真であるにもかかわらず、帰無仮説が真であると判断してしまう誤りを、第2種の誤り (過誤)、またはぼんやりものの誤りとよび、その確率を記号 β であらわす。一般に α を大きくすると β は小さくなり、 α を小さくすれば β は大きくなる。また、検定では、対立仮説が真のときにそれを正しく検出できることが重要である。この確率は $(1-\beta)$ となり、検出力という。表3-1に検定の仮説と誤りの確率について整理する。

検定における有意水準 (危険率) α とは、帰無仮説が成り立っている場合に、「めったに起こらない」と判断する確率であり、一般的には5%や1%といった小さい値に設定される。したがって、データから求めた検定統計量が、有意水準から求めた棄却域に入った場合は、「めったに起こらないことが起こった」とはせずに、「もとの仮定が間違っていた」と判断し、帰無仮説を棄却するのである。

図3-1に一つの母平均の検定 (対立仮説: $\mu > \mu_0$ 、母分散既知の場合) における棄却域と α 、 β 、検出力 $(1-\beta)$ の関係を示す。

検定においては、データから求めた検定統計量の値が棄却域に入ったとき、帰無仮説が棄却され、対立仮説が成り立っていると判断する。このとき、「検定結果は有意である」などと表現する。

棄却域 R とは、「帰無仮説を棄却すると判断する統計量の範囲」をいう。

- ・両側検定では、棄却域が右、左両側 (上側、下側という) にある。
- ・片側検定では、棄却域が右 (上側) または左 (下側) のいずれかにある。

図3-1は、片側検定で棄却域が右側 (上側) の場合を示している。

- ・棄却域は、有意水準 α によって定まる。

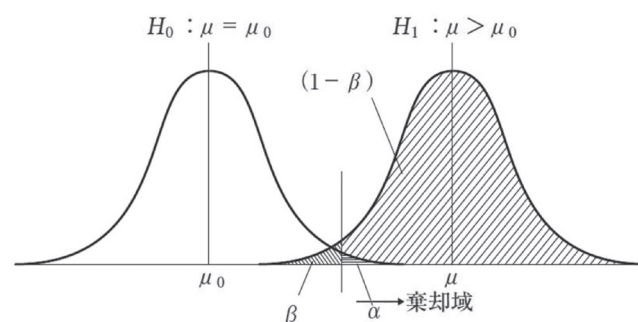


図3-1 母平均の検定における棄却域

有意水準を5%とすると、正規分布は左右対称なので、両側検定の場合、上側に2.5%分、下側に2.5%分の棄却域を設ける必要がある。正規分布の上側2.5%点（これを本稿では、 $u(0.05)$ と表現している）および下側2.5%点（ $-u(0.05)$ ）が帰無仮説を棄却する限界値になる。また、片側検定の場合は、上側または下側に5%分の棄却域を設けるので上側5%点（ $u(0.10)$ ）または下側5%点（ $-u(0.10)$ ）が帰無仮説を棄却する限界値になる。

検定統計量の値が棄却域に入らなかった場合は、「帰無仮説が正しい」とは表現せず、「対立仮説が正しいとはいえない」と表現する。これは先に述べた検出力がからんでおり、棄却域に入らなかったときには、「帰無仮説が正しい」場合の他に、「検出力が十分ではなかった」という可能性があるためである。サンプルの数が多くなれば検出力は大きくなるが、一方で時間やコストがかかるという問題が生じる。したがって、あらかじめ検出したい差と検出力を決めておいて必要なサンプルの数を算定することも行われる。

統計ソフトなどを使って検定を行うと、有意、有意でないに加えて、 p 値というものが表示されることがある。 p 値とは、統計量がサンプルのデータから計算した値よりも分布の中心から離れた側の値をとる確率を示す。例えば、 p 値が2%となっているなら、「帰無仮説が正しいとすると、サンプルのデータから計算した統計量の値は2%という小さな確率でしか起こらない」ということを示している。したがって、有意水準と p 値を比べることで検定の判断ができる。片側検定の場合の棄却域と p 値の関係を図3-2に示す。

(2) 検定の具体的な手順

例として、母分散が既知の場合で1つの母平均 μ が比較する値 μ_0 とは異なるかどうかを検定する場合を考える。

手順1 検定の目的の設定

母分散が既知である1つの母集団の母平均について、母平均が変わったかどうかの検定を行う。

手順2 帰無仮説 H_0 と対立仮説 H_1 の設定

$H_0 : \mu = \mu_0$

対立仮説には、

$H_1 : \mu \neq \mu_0$ (両側仮説)

$H_1 : \mu > \mu_0$ (片側仮説)

$H_1 : \mu < \mu_0$ (片側仮説)

の3つが考えられ、「検定によって何を主張したいか」によっていずれかを選ぶことになる。

- ・特性値の母平均が変化したといたい→ $H_1 : \mu \neq \mu_0$
- ・特性値の母平均が大きくなったといたい→ $H_1 : \mu > \mu_0$
- ・特性値の母平均が小さくなったといたい→ $H_1 : \mu < \mu_0$

手順3 検定統計量の選定

1つの母平均の検定において、母分散が既知の場合の母集団は、正規分布 $N(\mu, \sigma^2)$ に従う。ここからランダムに抜き取られた大きさ n のサンプルの平均値 $\bar{x}$ は、正規分布 $N(\mu, \frac{\sigma^2}{n})$ に従う。これを標準化した $u = \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}}$ は標準正規分布 $N(0, 1^2)$ に従う。

したがって、本検定における検定統計量は $u = \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}}$ である。

手順4 有意水準の設定

有意水準 α （第1種の誤りの確率）を設定する。一般には0.05（5%）、または0.01（1%）を採用する。

注：有意水準は検定に先立って決めておく。検定統計量を計算してから、検定結果を有意になるよう、または有意にならないように変えることはすべきではない。

手順5 棄却域の設定

有意水準と対立仮説に応じた棄却域を設定する。

1つの母平均の検定で、対立仮説が $H_1 : \mu \neq \mu_0$ の両側検定するとき、標準正規分布の棄却域は

$R : |u_0| \geq u(\alpha) = u(0.05) = 1.960$

となる。棄却域は R と表すことが多い。棄却域の値は、正規分布表（付表1 Vol.28 No.5 P311）から両側確率が0.05（上側確率で0.025、下側確率で0.025）になる正規分布の値を1.960と求めている。

両側検定の場合には、上側と下側の両方に棄却域が設定されるので、検定統計量の値が1.960以上または-1.960以下であれば有意と判断するのである。

手順6 検定統計量の計算

検定の対象となる母集団からランダムにサンプルを採取し、測定してデータを得る。データの平均値 $\bar{x} = \frac{1}{n} \sum x_i$ から、

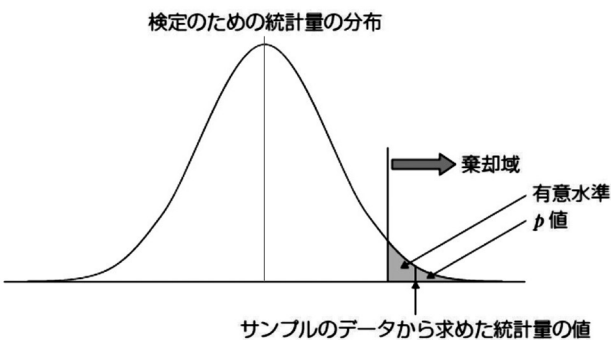


図3-2 棄却域と p 値

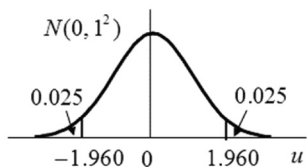


図3-3 正規分布の棄却域

$u = \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}}$ の値を計算する。

手順7 検定結果の判定

計算した検定統計量の値を棄却域の値と比較し検定の結果を判断する。棄却域に入っていれば有意であると判断し、入っていなければ有意ではないと判断する。

手順8 結論

検定の結果、有意であれば帰無仮説が棄却され、対立仮説が採択される。有意でない場合には、帰無仮説は棄却されない。

【参考】

帰無仮説は $H_0: \mu = \mu_0$ などと、常にある値に等しいとされている。その理由を述べる。

検定では、帰無仮説 H_0 のもとで (帰無仮説が正しいとして) データから計算された検定統計量 u が、観測された値 u_0 を超える確率を求め、この確率が小さい値であったときに帰無仮説を棄却する。

(実際の手順は、有意水準 α のもとで仮説が棄却される棄却域 ($u(\alpha)$, $u(2\alpha)$ など) を正規分表などの数値表から読み、 u_0 と比較している)

ここで、帰無仮説 H_0 のもとで検定統計量 u が、観測された値 u_0 を超える確率を求めるためには、 $H_0: \mu = \mu_0$ の場合しか、正しく計算することができない ($\mu < \mu_0$, $\mu > \mu_0$ などでは、分布を特定することができない)。

以上の理由から、帰無仮説は (片側検定の場合でも) 不等号はつけず、 $H_0: \mu = \mu_0$ などとしている。

5.2 推定

5.2.1 推定とは

検定でえられる結論は、「新たな工程の平均は従来と異なる」というようなものであった。これは常に重要な母集団に関する情報を与えてくれたのだが、「じゃあ、変わった平均はどれくらい?」という指摘や問い合わせがやってくることは避けられないだろう。これに答えをだしてくれるのが推定である。

推定については、「検定と推定は2つでひとつ」とか、「検定を行ってから推定をしなくてはならない」とか、「検定で有意でなければ推定には意味がない」などの言説もあるが、検定と推定はまったく別物で、検定だけ、推定だけ行っても問題はない。「視聴率調査」や「世論調査」の結果は、実はこの推定であって、後述する点推定の結果だけが重宝される場合も多い。

推定には点推定と区間推定がある。点推定は、「新たな工程の平均値は100.00kgと推定できる」というように1つの値で推定する。ところがこの推定値もサンプルをとるたびに異なる。サンプルの平均値がばらつくわけである。推定値がどの程度信頼できるかということを、区間を用いて「新たな工程

の平均値は100.00±1.00kg」などと表現する。これが区間推定である。

検定で判定が間違える確率を定めたように、区間推定では信頼区間というものを定める。この信頼区間の幅を決める値としては、95%、90%などが用いられ信頼率とよぶ。

信頼率の意味は、サンプルをとって平均値などを計算し信頼区間を求めることをなんどもなんども行った (だれも、そんなことはしないだろうが) とした場合、得られたたくさんの信頼区間のうち95%のものは真の平均 (母平均) を含んでいるという意味である (5%のものは真の平均を外している)。信頼率は $(1 - \alpha)$ とあらわす。具体的には、検定統計量が $(1 - \alpha)$ の確率で含まれる正規分布などの値の範囲を求め、そこから逆算して真の平均が $(1 - \alpha)$ の確率で含まれる範囲を求めている。

5.2.2 推定の手順

(1) 推定の概要

推定とは、対象とする母集団の分布の母平均や母分散といった母数を推定するものである。1つの推定量により母数を推定する点推定と、区間を用いて推定する区間推定がある。

点推定とは、母平均 μ や母分散 σ^2 などを1つの値で推定することであり、不偏推定量である平均値 $\bar{x}$ 、分散 V などがよく用いられる。

区間推定とは、推定値がどの程度信頼できるかを、区間を用いて推定する方法であり、信頼率 $(1 - \alpha)$ を定めて推定する。信頼率は、一般的には95% (0.95) または90% (0.90) を用いる。「保証された信頼率で母数を含む区間」である信頼区間、すなわち信頼区間の上限値 (信頼上限) と下限値 (信頼下限) である信頼限界を求める。

(2) 推定の具体的な手順

例として、母分散が既知の場合で1つの母平均 μ に関する場合を考える。

手順1 点推定

母平均 μ を点推定すると、

$$\mu = \bar{x}$$

となる。

注：記号 μ はハットとよび、母数の推定値であることをあらわす。この場合はミューハットとよぶ。

手順2 区間推定

母分散が既知の場合の母平均 μ の信頼率95%の区間推定は、

$$\text{統計量 } u = \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}}$$

が標準正規分布 $N(0, 1^2)$ に従うので、 $u = \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}}$ の値が、下側2.5%点 ($-u(0.05)$) と上側2.5%点 ($u(0.05)$) の間にある確率が $(1 - 0.05)$ であることから、

表3-2 計量値データに基づく検定と推定一覧 (Online version in color.)

母集団 の数	検定と推定の目的	母分散の 情報	統計量の 分布	検定統計量	対立仮説と棄却域	推定
1	母平均 μ に関する 検定と推定	母分散 σ^2 が既知	標準正規 分布	$u_0 = \frac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}}$	$H_1: \mu \neq \mu_0 \Rightarrow R: u_0 \geq u(\alpha)$ $H_1: \mu > \mu_0 \Rightarrow R: u_0 \geq u(2\alpha)$ $H_1: \mu < \mu_0 \Rightarrow R: u_0 \leq -u(2\alpha)$	$\bar{x} \pm u(\alpha) \sqrt{\frac{\sigma^2}{n}}$
1	母平均 μ に関する 検定と推定	母分散 σ^2 が未知	t 分布	$t_0 = \frac{\bar{x} - \mu_0}{\sqrt{V/n}}$	$H_1: \mu \neq \mu_0 \Rightarrow R: t_0 \geq t(\phi, \alpha)$ $H_1: \mu > \mu_0 \Rightarrow R: t_0 \geq t(\phi, 2\alpha)$ $H_1: \mu < \mu_0 \Rightarrow R: t_0 \leq -t(\phi, 2\alpha)$	$\bar{x} \pm t(\phi, \alpha) \sqrt{\frac{V}{n}}$
1	母分散 σ^2 に関する 検定と推定	—	χ^2 分布	$\chi_0^2 = \frac{S}{\sigma_0^2}$	$H_1: \sigma^2 \neq \sigma_0^2 \Rightarrow R: \chi_0^2 \geq \chi^2(\phi, \alpha/2)$ または $\chi_0^2 \leq \chi^2(\phi, 1-\alpha/2)$ $H_1: \sigma^2 > \sigma_0^2 \Rightarrow R: \chi_0^2 \geq \chi^2(\phi, \alpha)$ $H_1: \sigma^2 < \sigma_0^2 \Rightarrow R: \chi_0^2 \leq \chi^2(\phi, 1-\alpha)$	$\hat{\sigma}^2 = V = \frac{S}{n-1}$ $\sigma_U^2 = \frac{S}{\chi^2(\phi, 1-\alpha/2)}$ $\sigma_L^2 = \frac{S}{\chi^2(\phi, \alpha/2)}$
2	母平均 μ_A と母平均 μ_B の差に関する検 定と推定	母分散 σ^2 が既知	標準正規 分布	$u_0 = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}$	$H_1: \mu_A \neq \mu_B \Rightarrow R: u_0 \geq u(\alpha)$ $H_1: \mu_A > \mu_B \Rightarrow R: u_0 \geq u(2\alpha)$ $H_1: \mu_A < \mu_B \Rightarrow R: u_0 \leq -u(2\alpha)$	$(\bar{x}_A - \bar{x}_B) \pm u(\alpha) \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}$
2	母平均 μ_A と母平均 μ_B の差に関する検 定と推定 n_A と n_B の比 または V_A と V_B の比が 2倍以内なら、 $\sigma_A^2 = \sigma_B^2$ と見なす	母分散 σ^2 が未知 $\sigma_A^2 = \sigma_B^2$ の場合	t 分布 ただし $V = \frac{S_A + S_B}{n_A + n_B - 2}$	$t_0 = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{V \left(\frac{1}{n_A} + \frac{1}{n_B} \right)}}$ ただし $V = \frac{S_A + S_B}{n_A + n_B - 2}$	$H_1: \mu_A \neq \mu_B \Rightarrow R: t_0 \geq t(\phi, \alpha)$ $H_1: \mu_A > \mu_B \Rightarrow R: t_0 \geq t(\phi, 2\alpha)$ $H_1: \mu_A < \mu_B \Rightarrow R: t_0 \leq -t(\phi, 2\alpha)$ ただし $\phi = n_A + n_B - 2$	$(\bar{x}_A - \bar{x}_B) \pm t(\phi, \alpha) \sqrt{V \left(\frac{1}{n_A} + \frac{1}{n_B} \right)}$
2	母分散の比 $\frac{\sigma_A^2}{\sigma_B^2}$ に関する検定 と推定	—	F 分布	$F_0 = \frac{V_A}{V_B}$	$H_1: \sigma_A^2 \neq \sigma_B^2 \Rightarrow R: F_0 \geq F(\phi_A, \phi_B; \alpha/2)$ または $F_0 \leq F(\phi_A, \phi_B; 1-\alpha/2)$ $H_1: \sigma_A^2 > \sigma_B^2 \Rightarrow R: F_0 \geq F(\phi_A, \phi_B; \alpha)$ $H_1: \sigma_A^2 < \sigma_B^2 \Rightarrow R: F_0 \leq F(\phi_A, \phi_B; 1-\alpha)$	$\frac{\hat{\sigma}_A^2}{\hat{\sigma}_B^2} = \frac{V_A}{V_B}$ $\left(\frac{\sigma_A^2}{\sigma_B^2} \right)_U = \frac{V_A}{V_B} \frac{1}{F(\phi_A, \phi_B; 1-\alpha/2)}$ $= \frac{V_A}{V_B} F(\phi_B, \phi_A; \alpha/2)$ $\left(\frac{\sigma_A^2}{\sigma_B^2} \right)_L = \frac{V_A}{V_B} \frac{1}{F(\phi_A, \phi_B; \alpha/2)}$

注) 検定統計量の値には、下付の0 (帰無仮説 H_0 の0に由来するといわれる) を付けて表記している。

$$\Pr \left(-u(0.05) < \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}} < u(0.05) \right) = 1 - 0.05 = 0.95$$

注: $\Pr (*)$ とは、*の事象が起こる確率を表す記号である。
となり、これを解いて、

$$\text{信頼上限: } \mu_U = \bar{x} + u(0.05) \sqrt{\frac{\sigma^2}{n}} = \bar{x} + 1.960 \sqrt{\frac{\sigma^2}{n}}$$

$$\text{信頼下限: } \mu_L = \bar{x} - u(0.05) \sqrt{\frac{\sigma^2}{n}} = \bar{x} - 1.960 \sqrt{\frac{\sigma^2}{n}}$$

となる。

5.3 計量値の検定・推定

5.3.1 計量値の検定・推定の種類

計量値データに基づく検定と推定には多くの種類がある。
表3-2に1つまたは2つの母集団に関する検定方法および推
定方法についてまとめる。

(次号 Vol.28 No.8に続く)

(2023年3月15日受付)