検出力解析について（2）

高際　睦*

1　はじめに

高際 [6] を書いた後に，特に，研究で統計を良く使う人たちから，正規検定だけでなく，もっと頻繁に使うことのある t 検定やカイ 2 乗検定の検出力に関しても解説して欲しいと言われた．その中でも述べたことであるが，正規検定の検出力や必要な標本数の求め方は正規分布の知識だけあれば十分理解できるものであるし，しかも，Excel などの表計算ソフトウェアさえあれば，誰でもそれらの計算を行うことが可能である．それに対して，t 検定の検出力を説明しようとした場合，非心 t 分布などのかなり複雑な分布に関する知識が必要となるだけでなく，実際に検出力を計算しようとしても，高際 [6] を書いていた当時には，これらの分布の確率計算を行うための道具がほとんど存在していなかった．そのため，もし，解説を書いたとしても，数学的な記述に終始したものにしかならなかったはずである．現在でも，そのような状況が大幅に改善されているわけではないが，最近，様々な分野で使われるようになってきた R と呼ばれるフリーの統計ソフトウェアには，t 分布だけでなく，非心 t 分布の確率計算を行うための関数が用意されている．そのため，t 検定の検出力に関する最低限の理論がわかれば，このソフトウェアを使うことによって，その計算が誰にでも自由に行える．したがって，これらの理論をまとめておくこともまったくの無駄ではないので，ここでは，特に，一標本の t 検定の検出力と必要な標本数の計算に関しての解説を行う．高際 [6] 同様，かなり数学的な記述になっているが，出来る限り難しい議論は避け，数学が不得手な人でも頑張れば理解できる内容にしたつもりでいる．

本稿では，まず，その後の議論で必要となる確率分布に関して簡単に触れ，次に，

* 東京歯科大学　数学研究室

一標本 t 検定，および，その検出力についての解説を行う．さらに，t 検定における標本数の求め方を説明し，最後に，R を使ってこれらを実際に計算する方法を示す．

2 いくつかの確率分布

t 検定では t 分布と呼ばれる確率分布に基づいて検定が行なわれる．したがって，t 検定を理解するうえでは，t 分布の基礎的な知識が不可欠であるが，その t 分布はカイ 2 乗分布と呼ばれるまた別の確率分布を用いて定義される．カイ 2 乗分布それ自体も重要な分布ではあるが，本稿ではカイ 2 乗分布を単独で用いることはないので，この分布についてはごく簡単に述べ，その後，t 分布，さらに t 検定の検出力を求めるために必要な非心 t 分布について説明する．これらの分布に関する詳細な解説は，例えば，Johnson[2]，蓑谷 [4] などを参照してもらいたい．

2.1 カイ 2 乗分布

平均 0，分散 1 の正規分布（標準正規分布）に従う確率変数を Z とするとき，その 2 乗 Z^2 が従う確率分布を自由度 1 のカイ 2 乗分布と言い，さらに，n 個の独立な標準正規分布に従う確率変数 $Z_1, Z_2, \cdots, Z_n$ に対して，その二乗和を

$$Y \;=\; Z_1^2 + Z_2^2 + \cdots + Z_n^2 \;=\; \sum_{k=1}^{n} Z_k^2$$

としたとき，Y を自由度 n のカイ 2 乗分布に従う確率変数と言う．カイ 2 乗分布の確率密度関数の式や平均（期待値），分散などは参考文献を見てもらいたい．カイ 2 乗分布に従う確率変数として，特に重要なものは以下の統計量である．

平均 μ，分散 σ^2（標準偏差 σ）の正規母集団からの n 個の独立な標本 $X_1, X_2, \cdots, X_n$ に対して，標本分散（もしくは不偏分散）は

$$S^2 \;=\; \frac{1}{n-1} \sum_{k=1}^{n} (X_k - \overline{X})^2 \tag{1}$$

で定義される．ただし，$\overline{X}$ は標本平均，つまり，$\overline{X} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$ である．このとき，統計量 $\dfrac{(n-1)\,S^2}{\sigma^2}$ は自由度 $n-1$ のカイ 2 乗分布に従うことが知られている．そ

れは，この統計量が

$$\frac{(n-1)\,S^2}{\sigma^2} \;=\; \frac{1}{\sigma^2}\sum_{k=1}^{n}(X_k-\overline{X})^2 \;=\; \sum_{k=1}^{n}\left(\frac{X_k-\mu}{\sigma}\right)^2-\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right)^2$$

と変形できること，$\overline{X}$ が平均 μ，分散 $\dfrac{\sigma^2}{n}$ の正規分布に従うこと，また，一般に X が平均 μ，分散 σ^2 の正規分布に従う確率変数であるとき $\dfrac{X-\mu}{\sigma}$ が標準正規分布に従うことなどより導くことができる．

2.2　t 分布

Z を標準正規分布に従う確率変数，Y を Z と独立な自由度 n のカイ 2 乗分布に従う確率変数としたとき，

$$T \;=\; \frac{Z}{\sqrt{Y/n}}$$

で定義された T を自由度 n の t 分布（もしくは，スチューデントの t 分布）に従う確率変数と言う．自由度 n の t 分布に従う確率変数 T の平均（期待値），分散はそれぞれ，0 $(n>1)$，$\dfrac{n}{n-2}$ $(n>2)$ である．t 分布の確率密度関数の式は省略するが，いくつかの自由度の確率密度関数のグラフを図 1 に示す．

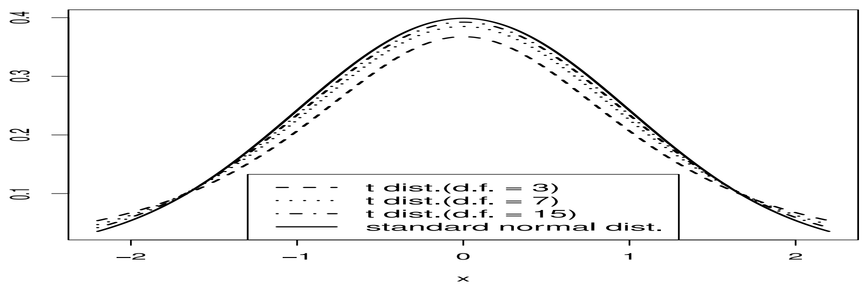

図 1　下から，自由度 3，7，15 の t 分布の確率密度関数のグラフで，一番上のグラフが標準正規分布の確率密度関数のグラフである．

図 1 からもわかるように，t 分布の確率密度関数は，自由度によらず，平均 $x=0$ に関し対称であり，また，自由度が小さい場合，標準正規分布の確率密度関数に較べ

両裾が長くなっている（分散が大きい）が，自由度が大きくなるに従い，標準正規分布の確率密度関数のグラフに近づいていく[*1].

平均 μ，分散 σ^2 の正規分布に従う n 個の独立な確率変数 $X_1, X_2, \cdots, X_n$ に対して，その標本平均 $\overline{X}$ の確率分布が平均 μ，分散 $\dfrac{\sigma^2}{n}$ の正規分布であることより，

$$Z = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} = \frac{\sqrt{n}\,(\overline{X} - \mu)}{\sigma} \tag{2}$$

が標準正規分布に従う確率変数となることはすでに述べた通りであるが，もし，ここで，分散 σ^2 の代わりに (1) 式で定義した標本分散 S^2 を用いて $\overline{X}$ を標準化した

$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} = \frac{\sqrt{n}\,(\overline{X} - \mu)}{S} \tag{3}$$

はもはや正規分布に従う確率変数ではなく，以下のように変形すればわかるように，

$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} = \frac{\dfrac{\overline{X} - \mu}{\sqrt{\sigma^2/n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}\Big/(n-1)}}$$

は自由度 $n-1$ の t 分布に従う確率変数になる[*2].

2.3　非心 t 分布

t 分布のときと同様に，Z を標準正規分布に従う確率変数，Y を Z と独立な自由度 n のカイ 2 乗分布に従う確率変数とし，さらに，δ を定数としたとき

$$T' = \frac{Z + \delta}{\sqrt{Y/n}}$$

で定義した T' を自由度 n，非心度（もしくは，非心パラメータ）δ の非心 t 分布に従う確率変数と言う．定義から明らかなように，$\delta = 0$ のときは自由度 n の t 分布と一致する．非心 t 分布に関してもその確率密度関数の式は書かないが，図 2 に，自

[*1] t 分布の分散の式からわかるように，自由度 n が大きくなれば，その値は 1（標準正規分布の分散）に近づく．実際，n が 30 以上であれば標準正規分布とほとんど変わらない．

[*2] t 分布に従うことを示すには，さらに，$\overline{X}$ と S^2 が独立であることも証明しなくてはならないが，ここでは省略する．

由度が同じ（$n = 10$）で，非心度を変化させた確率密度関数のグラフを，図３には，非心度を一定（$\delta = 3$）にし，自由度を変化させた確率密度関数のグラフを示す．定義からも予想はつくが，非心 t 分布の確率密度関数は，非心度 δ 近辺でその確率密度は高くなるが，t 分布のそれと異なり対称ではない．ただし，自由度が大きくなるに従い，分布の裾が短くなっていくことは t 分布の確率密度関数と同じ性質である．

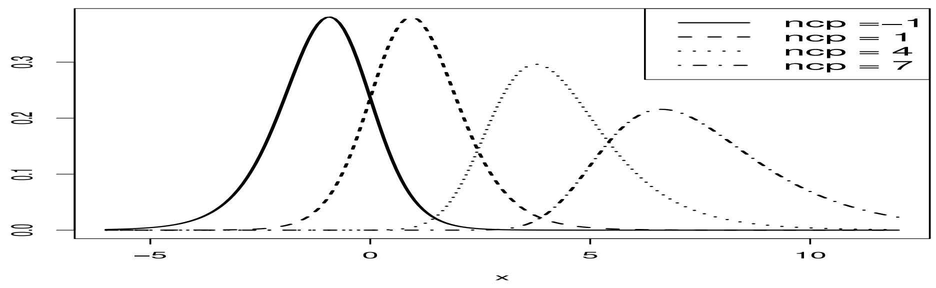

図２　自由度 10 で，左から非心度 $-1, 1, 4, 7$ の非心 t 分布の確率密度関数のグラフ．

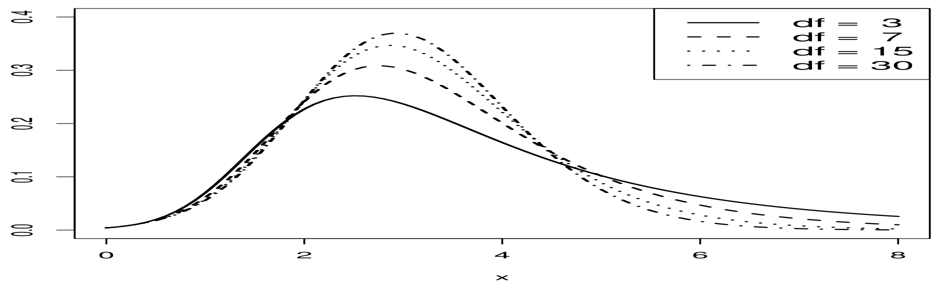

図３　非心度 3 で，下から自由度 3, 7, 15, 30 の非心 t 分布の確率密度関数のグラフ．

平均 μ，分散 σ^2 の正規母集団に従う n 個の独立な確率変数 $X_1, X_2, \cdots, X_n$ に対して，その標本平均，標本分散をそれぞれ $\overline{X}$，S^2，また，μ_1 をある定数としたとき，

$$T' = \frac{\overline{X} - \mu_1}{\sqrt{S^2/n}}$$

で定義された T' は，$\mu_1 = \mu$ であれば，(3) 式と同じであるので t 分布に従う確率

変数であるが，$\mu_1 \neq \mu$ ならば，以下の最後の式から

$$T' \;=\; \frac{\overline{X}-\mu+\mu-\mu_1}{\sqrt{S^2/n}} \;=\; \frac{\dfrac{\overline{X}-\mu}{\sqrt{\sigma^2/n}} \;+\; \dfrac{\mu-\mu_1}{\sqrt{\sigma^2/n}}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}\Big/(n-1)}} \tag{4}$$

は自由度 $n-1$，非心度 $\sqrt{n}\,(\mu-\mu_1)/\sigma$ の非心 t 分布に従う確率変数であることがわかるであろう．

3 t 検定とその検出力

この節では，まず，一標本の t 検定について説明し，その後，この t 検定における検出力の求め方を示す．一般的な検定の方法や検定における過誤，および，検出力などについては Mendenhall[3] や [7] を参照してもらいたい．

3.1 一標本 t 検定

分散 σ_0^2 が既知な正規母集団の平均 μ に関する検定は，帰無仮説 $H_0 : \mu = \mu_0$ が正しいときに, 母集団からの n 個の独立な標本 $X_1, \cdots, X_n$ の標本平均 $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$ を標準化した

$$Z_0 \;=\; \frac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}} \;=\; \frac{\sqrt{n}\,(\,\overline{X}-\mu_0\,)}{\sigma_0} \tag{5}$$

が標準正規分布に従うことを使って検定を行えばよいことは高際 [6] で説明した通りである．この Z_0 を検定統計量と言い，Z_0 の分布が正規分布であることよりこの検定を正規検定と呼ぶ.

ここで，少し上の状況を考えてもらいたい．母平均 μ に関する検定を行うということは，μ の真の値がわからないからであるが，そのときには，母分散の値もわかっていないことが多い．なぜならば，通常，分散は母平均の値を使って求めるからである．したがって，実際に正規検定が適用される場面はそう多くはない．では，母集団分布が正規分布であることはわかっているが，その分散が未知である場合の検定はどのように行えばよいであろうか．分散の値はわからないが，やはり, 正規検定と同じ

様な方法を行うのが自然であろう．そこで，正規検定の検定統計量である (5) 式において，母分散の代わりにその最も良い推定量である標本分散 S^2 を用いて，標本平均 $\overline{X}$ を標準化すれば,

$$T_0 \;=\; \frac{\overline{X}-\mu_0}{S/\sqrt{n}} \;=\; \frac{\sqrt{n}\,(\,\overline{X}-\mu_0\,)}{S} \tag{6}$$

は，帰無仮説 $H_0 : \mu = \mu_0$ が正しければ，2.2 で説明した通り，自由度 $n-1$ の t 分布に従う確率変数（検定統計量）となる[*3]．あとは，正規検定をはじめとする他の検定同様，あらかじめ定めてあった有意水準 α から棄却域を求め，検定統計量 T_0 がこの棄却域に入るかで，帰無仮説を棄却するかどうかを判断すればよい．t 検定の棄却域，棄却点は t 分布が 平均 0 に関して対称な分布であるので，正規検定のとき同様に求めればよい．つまり，対立仮説が $H_1 : \mu \neq \mu_0$ である両側検定の場合は，その棄却点，棄却域をそれぞれ $\pm t_{\alpha/2}(n-1)$ ，$|T_0| > t_{\alpha/2}(n-1)$ とし，対立仮説が $H_1 : \mu > \mu_0$ の片側 t 検定の場合は $t_\alpha(n-1)$，$|T_0| > t_\alpha(n-1)$ とする．ただし，$t_\alpha(n)$ は自由度 n の t 分布の上側 100α パーセント点である．例えば，有意水準 0.05 で，標本数 16 の両側検定の棄却域は $|T_0| > t_{0.025}(15) = 2.131$，片側検定の棄却域は $T_0 > t_{0.05}(15) = 1.753$ である[*4]．この棄却域は，同じ条件のもとでの正規検定の棄却域より狭くなっている．

3.2　t 検定の検出力

検定の検出力とは第 2 種の過誤（帰無仮説が間違っているのに帰無仮説を棄却しない）が起こらない確率のことを言い，第 2 種の過誤が起こる確率を β とすれば，検出力は $1-\beta$ で表わされる．まずは，帰無仮説，対立仮説がそれぞれ，$H_0 : \mu = \mu_0$，$H_1 : \mu \neq \mu_0$，有意水準 α である一標本の両側 t 検定における検出力を考える．もし，母平均の真の値が $\mu_1 (\neq \mu_0)$ であるならば，それは帰無仮説が間違っていることであるので，このときに帰無仮説を棄却する確率が検出力である．つまり，このときに検定統計量 T_0 が $|T_0| > t_{\alpha/2}(n-1)$ となる確率のことであるが，ただし，このときの T_0 は t 分布でなく，(4) 式で示した通り，自由度 $n-1$，非心度 $\sqrt{n}\,(\mu_1-\mu_0)/\sigma$

[*3] したがって，この検定を（一標本） t 検定と呼ぶ.

[*4] t 分布の棄却点，つまり，上側 100α パーセント点の計算は易しくはないが，検定などで用いる主だった棄却点の値はすでに計算されており，たいていの統計学の教科書には載っている．また，後で例示するように R を使って求めることもできる.

の非心 t 分布に従う確率変数である．したがって，$\mathcal{F}(x\,;n,d)$ を自由度 n，非心度 d の非心 t 分布の分布関数とすれば，検出力は

$$\begin{aligned}1-\beta &= \mathrm{P}\left(|T_0| > t_{\alpha/2}\right)\\ &= 1-\mathrm{P}\left(-t_{\alpha/2} < T_0 < t_{\alpha/2}\right)\\ &= 1-\mathcal{F}\left(t_{\alpha/2}\,;n-1,\sqrt{n}\,(\mu_1-\mu_0)/\sigma\right)+\\ &\qquad\qquad \mathcal{F}\left(-t_{\alpha/2}\,;n-1,\sqrt{n}\,(\mu_1-\mu_0)/\sigma\right) \qquad (7)\end{aligned}$$

である．ここで，$t_{\alpha/2}=t_{\alpha/2}(n-1)$ とする．この式より一標本 t 検定の検出力が求められそうだが，(7) 式の非心度 $\sqrt{n}\,(\mu_1-\mu_0)/\sigma$ の分母である標準偏差 σ が未知（そのために，t 検定を行っている）であるので，実際には検出力の正確な値は求められない．通常は，σ の代わりに標本の標準偏差（標本分散の平方根）を使って検出力を求めるが，それで求めた値はあくまでも近似値であることに注意してもらいたい[*5]．帰無仮説，対立仮説がそれぞれ，$H_0:\mu=\mu_0$，$H_1:\mu>\mu_0$ である有意水準 α の片側 t 検定の，真の母平均が μ_1 であるときの検出力も，理論的には

$$\begin{aligned}1-\beta &= \mathrm{P}\left(T_0 > t_{\alpha}\right)\\ &= 1-\mathcal{F}\left(t_{\alpha}\,;n-1,\sqrt{n}\,(\mu_1-\mu_0)/\sigma\right) \qquad (8)\end{aligned}$$

となる．

例 1　分散が未知な正規母集団の平均に関して, 帰無仮説, 対立仮説を $H_0:\mu=100$, $H_1:\mu\neq 100$ とした，有意水準 $\alpha=0.05$ の両側検定を $n=16$ 個の標本を用いて行う．ただし，標本分散は 100 とする．このとき，母平均 μ の真の値が 104 であるときの検出力は，棄却域が $|T_0|>t_{0.025}(15)=2.13145$ であることより，(7) 式を使えば

$$\begin{aligned}1-\beta = 1-\mathcal{F}&\left(2.13145\,;15,\,\sqrt{16}\,(104-100)/10\right)+\\ &\mathcal{F}\left(-2.13145\,;15,\,\sqrt{16}\,(104-100)/10\right) \;=\; 0.32237\end{aligned}$$

となる．同様に，対立仮説が $H_1:\mu>100$ である片側 t 検定の母平均が 104 であるときの検出力も棄却点が $t_{0.05}(15)=1.75305$ であることと (8) 式より

$$1-\beta \;=\; 1-\mathcal{F}\left(1.75305\,;15,\,\sqrt{16}\,(104-100)/10\right) \;=\; 0.45326$$

[*5] ただし，実用上はこの近似値で十分である．

と求められる．

第１種の過誤（帰無仮説が正しいのに帰無仮説を棄却する）が起こる確率は有意水準 α によってコントロールすることができる．もし，この確率を小さくしたいのであれば，有意水準の値を小さくすればよいし，それほど小さくする必要がなければ，比較的大きな有意水準の値を使えばよい．ただし，有意水準を小さくして，第１種の過誤が起こる確率を小さくすれば，棄却域は狭くなり，結果として，第２種の過誤が起こる確率が大きくなる（検出力は小さくなる）．実際，

例２　（**例１**つづき）もし，有意水準 α が 0.01 であれば，母平均の真の値が 104 であるときの両側検定の検出力は，棄却点が $t_{0.005}(15) = 2.946713$ であることより，

$$1 - \beta = 1 - \mathcal{F}\left(2.946713\,;\,15,\,\sqrt{16}\,(104-100)/10\right) + \mathcal{F}\left(-2.946713\,;\,15,\,\sqrt{16}\,(104-100)/10\right) \;=\; 0.12585\ .$$

対立仮説が $H_1 : \mu > 100$ である片側検定の検出力も，同様に，

$$1 - \beta \;=\; 1 - \mathcal{F}\left(2.60248\,;\,15,\,\sqrt{16}\,(104-100)/10\right) \;=\; 0.19286$$

となる．どちらの場合も有意水準 0.05 のときの検出力よりはるかに小さくなっている．

どの検定でもそうであるが，第１種の過誤，第２種の過誤の両方の起こる確率を同時に小さくするためには，標本数を大きくするしかない．図４に **例１** と同じ両側検定の３つの標本数 $n = 10$，$n = 20$，$n = 30$ に対する検出力関数（母平均の真の値を変数として，その値に対する検出力を求める関数）のグラフを示す．母平均の値に関わらず，標本数が大きくなるほど，検出力も大きくなっていることがわかるであろう．

4　検出力を用いた標本数の計算方法

3.2 でも述べたが，有意水準を小さくして，第１種の過誤が起こる確率を小さくすれば，第２種の過誤が起こりやすくなり，逆に，第２種の過誤が起こる確率を小さくするためには，有意水準を比較的大きな値にしなければならないが，そうすると今度は第１種の過誤が起こりやすくなる．実際に検定を行う場合には，どちらの過誤が起こる確率を小さくすることがより重要であるかを考え，有意水準を決めなければなら

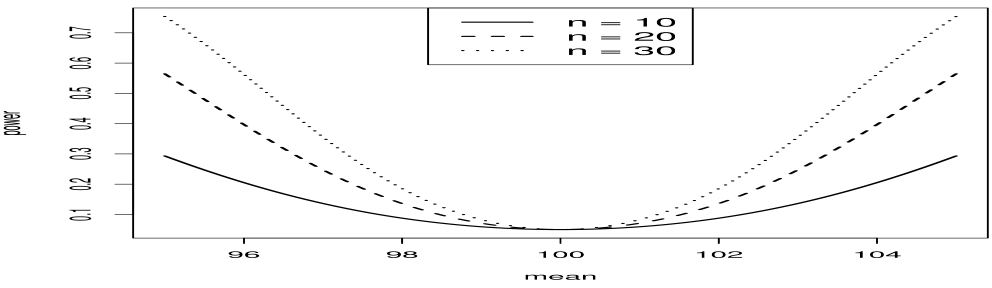

図 4　**例 1** と同じ両側検定で，下から標本数 $n = 10$，$n = 20$，$n = 30$ の母平均 μ を 95 から 105 まで変化させたときの検出力関数のグラフ．

ない．ただし，定められた有意水準の値に対して，検出力もある程度高い値が要求されることがある．この要求に対する解決策は，やはり標本数を大きくすることしかない．では，要求された検出力に対して必要とされる標本数はどのように求めればよいだろうか．

まず，正規母集団で分散が既知である，帰無仮説，対立仮説が $H_0 : \mu = \mu_0$，$H_1 : \mu \neq \mu_0$ である有意水準 α の両側正規検定に関してその標本数の求め方を思い返してみる．この場合，母平均の真の値が μ_1 であるときの検出力は，$\Phi(x)$ を標準正規分布の分布関数，z_α を標準正規分布の上側 100α パーセント点とすると

$$1 - \beta \ = \ 1 - \Phi\left(z_{\alpha/2} - \frac{\sqrt{n}(\mu_1 - \mu_0)}{\sigma_0}\right) + \Phi\left(-z_{\alpha/2} - \frac{\sqrt{n}(\mu_1 - \mu_0)}{\sigma_0}\right)$$

であることより，母平均が μ_1 のときの検出力を γ とする標本数 n は，近似的に

$$n \ = \ \left(\frac{(z_{\alpha/2} - z_\gamma)\,\sigma_0}{\mu_1 - \mu_0}\right)^2 \tag{9}$$

と求めることができた（高際 [6] を参照）．一標本の t 検定においても，母平均の真の値が μ_1 であるときの検出力が (7) 式で表わされるので，検出力が γ となる標本数も

$$\gamma = 1 - \mathcal{F}\left(t_{\alpha/2}\,;\, n-1, \sqrt{n}\,(\mu_1 - \mu_0)/\sigma\right) + \mathcal{F}\left(-t_{\alpha/2}\,;\, n-1, \sqrt{n}\,(\mu_1 - \mu_0)/\sigma\right)$$

の式から同じように求められそうに思えるが，実は，そう簡単な話ではない．正規検定の場合は，母平均の値がいくつであっても，検定統計量 Z_0 の分布は正規分布であり，検出力が γ となる確率点 z_γ を容易に計算することができた．それに引き換え，

t 検定の場合では，母平均の値が $\mu_1(\neq \mu_0)$ であれば，検定統計量 T_0 は非心 t 分布に従い，この分布における検出力 γ の確率点がそう簡単に求められないだけでなく，求められたとしても，その値は標本数 n に依存するので，n を簡単な式で表すことはできない．では，t 検定における標本数はどのようにして求めるかというと，通常は，非心 t 分布を正規分布で近似して求めるのだが，その方法はかなり専門的なので，ここでは，結果を述べるだけにする．有意水準 α の一標本両側 t 検定における真の母平均が μ_1 であるときの検出力を γ とするために必要な標本数 n は，近似的に

$$n \ = \ \left(\frac{(z_{\alpha/2} - z_{\gamma})\,\sigma}{\mu_1 - \mu_0}\right)^2 + \frac{z_{\alpha/2}^2}{2} \tag{10}$$

であり，対立仮説が $H_1 : \mu > \mu_0$ である片側検定の場合も，母平均の真の値が μ_1 であるとき，その検出力を γ とするのに必要な標本数は

$$n \ = \ \left(\frac{(z_{\alpha} - z_{\gamma})\,\sigma}{\mu_1 - \mu_0}\right)^2 \ + \ \frac{z_{\alpha}^2}{2} \tag{11}$$

である．両方とも，正規検定の標本数を求める式に最後の項が加わった式となるので，同じ状況では t 検定の方が標本数が数個多く必要になることがわかる．もし，非心 t 分布の正規近似や t 検定の標本数の導き方について詳しく知りたいときには，例えば，永田 [5] が参考になる．

例３　（**例１** つづき）　もし，母平均の真の値が 105 であるときの検出力を 0.5 としたいならば，それに必要な標本数は，(10) 式より

$$n \ = \ \left(\frac{(1.96 - 0) \times 10}{105 - 100}\right)^2 + \frac{1.96^2}{2} \ = \ 17.2872$$

となる．また，片側検定で，母平均の真の値が 105 であるときの検出力を 0.5 とするために必要な標本数は

$$n \ = \ \left(\frac{(1.645 - 0) \times 10}{105 - 100}\right)^2 \ + \ \frac{1.645^2}{2} \ = \ 12.17711$$

である．

5 R を用いた検出力の計算

せっかく理論的なことがわかっても，実際にそれらの計算ができなければあまり意味がないかもしれない．そこで，この節では，R という統計ソフトウェアを使って，実際に t 検定の検出力や必要な標本数を計算する方法を述べる．R の詳しい使い方は，舟尾 [1] をはじめ多くの書籍や参考ホームページがあるので，それらを参照してもらいたい．

R は AT&T ベル研究所で開発された対話型環境（言語）である S 言語を参考にして作られたソフトウェアである．対話型環境と言うのは，ユーザが式を入力すれば，その式に対するアクションをすぐに行ってくれる環境のことであり，R も基本的にはコンソール・ウィンドウと呼ばれるウィンドウで対話的に操作するものである[*6]．実際の R の操作は簡単で，R を起動すると，`R Console` ウィンドウにバージョンなどの様々な情報が表示された後に `>` 記号（プロンプト）が現れるが，あとはこのプロンプトの後に行いたい式を入力するだけでよい．例えば，$1 + 2 \times 3 \div 4$ という演算を行いたいならば，`>` の後に以下のような式を入力し，エンターキーを押せば，

```
> 1+2*3/4
[1]  2.5
```

`2.5` という答えを出力してくれる．R には様々な関数が用意されており，その関数の引数に適当な数値（もしくは式）を与えれば，その関数の値を求めたり，様々なグラフィックスを描画したりしてくれる[*7]．もし，関数にどのような引数があるかとかその関数の詳しい使い方を知りたいときには，以下のように，

```
> ?pt
```

`?` の後に関数名を入力すれば，その関数のオンラインドキュメントが表示される．ここに出てくるものをはじめ，使い方のよくわからない関数があるときには，このオンラインドキュメントを利用してもらいたい．

それでは，t 検定の検出力を計算してみよう．検出力を求めるためには，まずは，棄却域（点）が必要になる．これを求めるためには，`qt` という関数を使えばよい．`qt`

[*6] 他の商用統計ソフトウェア同様に，GUI を使って行うためのパッケージも用意されている．

[*7] 本稿の図はすべて R で作成した．

は引数で与えられた p と df の値に対して，自由度 df の t 分布の下側確率が p である確率点を求めてくれる．したがって，**例 1** の自由度 15，有意水準 0.05 の両側検定，片側検定の棄却点はそれぞれ

```
> qt(1 - 0.05/2, df = 15)
[1]  2.13145
> qt(1 - 0.05, df = 15)
[1]  1.75305
```

で求められる[*8]．次に，検出力を求めるために必要なのは，t 分布（非心 t 分布）の下側確率を求める関数であるが，これを行う関数は pt である．この関数の主な引数には q，df，ncp などがあって，ncp 引数を省略したときには，自由度 df の t 分布の，ncp を与えたときには，自由度 df，非心度 ncp の非心 t 分布の q 以下の累積確率（下側確率）を求めてくれる．この関数を使えば，**例 1** の検出力は (7), (8) 式より

```
> 1 - pt(2.13145, df=15, ncp=1.6) - pt(-2.13145, df=15, ncp=1.6)
[1]  0.3223701
> 1 - pt(1.75305, df=15, ncp=1.6)
[1]  0.4532642
```

と求められる．さらに，t 検定の必要な標本数を求めたいならば，z_α，z_γ などの値を関数 dnorm から求め，あとは (10)，(11) 式を使えばよいが，ここでは省略する．実は便利なことに，R には t 検定の検出力や必要な標本数を求めてくれる関数 power.t.test が最初から用意されている．この関数の引数には n，delta，sd，power，sig.level，alternative などがあり，それぞれ標本数，平均の差（非心度ではない），標準偏差，検出力，有意水準，対立仮説（両側か片側か）を与える．例えば，先ほどと同じ検出力を求めたいときには，power 以外の引数に適切な数値を与えれば（引数は他の引数と区別できるところまで入力すればよい），

```
> power.t.test(n = 16, delta = 4, sd=10, type="one",
+  alt="two.sided", sig=0.05, strict=T)

     One-sample t test power calculation
```

[*8] 本稿の $t_\alpha(n)$ は上側 $100\,\alpha$ パーセント点であるが，qt は下側確率点を求める関数である．

```
            n = 16
        delta = 4
           sd = 10
    sig.level = 0.05
        power = 0.3223702
  alternative = two.sided
```

この検定の検出力が，結果の power のところに表示される．同様に，必要な標本数を求めたいときには，引数 n の代わりに，要求する検出力を引数 power に与えれば，必要な標本数が結果の n のところに表示される．

```
> power.t.test(delta = 2, sd=10, power = .5, type="one",
+  alt="two.sided", sig=0.05, strict=T)

     One-sample t test power calculation
            n = 17.35044
        delta = 5
           sd = 10
    sig.level = 0.05
        power = 0.5
  alternative = two.sided
```

6　終わりに

何人かの人たちから t 検定の検出力の解説もと頼まれていたこともあるが，筆者自身も高際 [6] の続きをいつか書きたいと思い続けていた．ただし，どうせ書くならば，高際 [6] を書いたあとに受けた “検出力についてはわかったけど，では，実際に検出力や標本数を計算するにはどうすればいいの？” という要望に応えられるものにしたいと思っていたのだが，t 検定の検出力を計算するための手段がなかなか見つからず，実際の計算方法はという要望に応えることは半ばあきらめかけていた．そんなとき，たまたま別の研究で使っていた R 言語に非心 t 分布の確率を計算する関数があることを知り，それならばと思って本稿を書き始めたのだが，ただでさえ t 検定や

その検出力に関する説明には様々な確率分布の話が必要なうえに，実際の計算例を示すには，その道具である R 言語の簡単な使い方の紹介も必要になりと，予想外に内容が増えてしまった．そのため，最初は書くつもりでいた二標本の t 検定や Welch 検定，また，それらの検出力に関する記述を省いてしまったが，そのことが非常に残念であり，また心残りでもある．もし，さらにまたこの続きを書く機会があれば，それらのことについても，是非説明したいと思っている．

参考文献

[1] 舟尾暢男,「The R tips ー　データ解析環境 R の基本技・グラフィックス活用集」，オーム社，2009.

[2] N. L. Johnson, S. Kotz, N. Balakrishnan, *Continuous Univariate Distributions*, Wiley, 1995.

[3] W. Mendenhall, R. I. Beaver，B. M. Beaver，*Introduction to probability and statistics*, Brooks/Cole, 2009.

[4] 蓑谷千凰彦,「正規分布ハンドブック」，朝倉書店，2005.

[5] 永田　靖,「サンプルサイズの決め方」，朝倉書店，2003.

[6] 高際　睦,「検出力解析について」，東京歯科大学　教養系研究紀要，26, 15-26, 2011.

[7] 東京大学教養学部統計学教室 (編),「自然科学の統計学」，東京大学出版会，1992.