t値、t分布、t検定の「意味」についてのメモ

ー 改訂版 ー

A note on the meanings of t-value, t-distribution and t-test

2017.2.24
石黒真木夫@統計思考院

概要

t分布は、正規確率変数の実現値から求められる t値の分布と「定義」されている。そして、t値とは　$T=(\hat{\mu}-m)/\hat{\sigma}$　という式で計算される統計量であるというような説明で終わっていることが多い。この言い方は t値が「何を」計算しようとするものなのかを説明していない。これでは t分布なるものが正体不明な量の分布ということになってしまう。

正規分布に従う確率変数が正の値をとる確率を計算する過程で見えてきた、t値と　t分布の「意味」を紹介しておく。

t値の意味を理解することは　t検定がどういう場面でどう役に立つのか考えるときに役に立つ。

1. t 値の意味

1.1 確率変数X > 0 の「認定」という問題

ある正の値、α を固定したとき、

$$Prob.(X > 0) > \alpha$$

なるとき、「確率変数X > 0 」 と認定することにする。

社会的問題でこの形の認定が必要になる場合は多い。α をどうとるか客観的に決める方法があればそれを使えばいい。

客観的に決められない場合にはこの認定を必要とする「当事者」の間で合意された値を使うことになる。

1.2　$X \sim N(\mu, \sigma^2)$の場合　:t値の「意味」

$$Prob.(X>0) = Prob(X-\mu > -\mu) = Prob.\left(\frac{X-\mu}{\sqrt{\sigma^2}} < \frac{\mu}{\sqrt{\sigma^2}}\right)$$
$$= Prob.\left(G < \frac{\mu}{\sqrt{\sigma^2}}\right) = F_G\left(\frac{\mu}{\sqrt{\sigma^2}}\right)$$

であるから、

$$\alpha \ < \ F_G\left(\frac{\mu}{\sqrt{\sigma^2}}\right)$$

なら、「確率変数X > 0 」 と認定されることになる。ただし、G ~$N(0,1)$　である。

X が「 n 個のデータの平均値」という確率変数である場合、その分布は

$$X \sim N\left(\mu, \frac{1}{n}(\text{データの分散})\right)$$

となるから、$\sigma^2 = \frac{1}{n}(\text{データの分散})$ として、計算する必要があることに注意。

1.3　標本に基づく認定

$$X_i \sim N(\mu, \sigma^2) \qquad i = 1,2, \dots, n$$

$x_i = X_i$ の実現値

$$\hat{\mu} = \frac{1}{n}\sum_{i\sim 1}^{n} x_i$$

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu})^2$$

とする。

問題：$\hat{\mu},\ \hat{\sigma}^2$　が与えられた条件のもとで $Prob.(X>0)$　を求めよ。
ただし、$X \sim N(\mu, \sigma^2)$　とする。

1.3.1 解 I

$\hat{\mu}$ と $\hat{\sigma}^2$ を信頼して、

$$\alpha < F_G(\frac{\hat{\mu}}{\sqrt{\hat{\sigma}^2}})$$

なら「確率変数X > 0」と認定する。

1.3.2 解 II

認定にあたって $\hat{\mu}$ と $\hat{\sigma}^2$ の推定誤差を考慮する。「事後分布」、

$$h\left(\frac{\mu}{\sqrt{\sigma^2}} \mid \frac{\hat{\mu}}{\sqrt{\hat{\sigma}^2}}\right)$$

が分かるなら、

$$\alpha < \int F_G(s)\, h\left(s \mid \frac{\hat{\mu}}{\sqrt{\hat{\sigma}^2}}\right) \mathrm{ds}$$

のとき「確率変数X > 0」と認定する。という「手」がある。

$$\varphi(u) = \int F_G(s)\, h(s \mid u)\, \mathrm{ds}$$

で定義される関数を利用して

$$\alpha < \varphi\left(\frac{\hat{\mu}}{\sqrt{\hat{\sigma}^2}}\right)$$

なら「確率変数X > 0」と認定するわけである。

2. t-分布の「意味」

2.1 $\frac{\mu}{\sigma}$ の推定値としての $\frac{\bar{X}}{S_n}$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \mu + \Delta\mu$$
$$V_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$
$$S_n = \sqrt{V_n} = \sigma + \Delta\sigma$$

とすると,

$$\frac{\bar{X}}{S_n} = \frac{\mu + \Delta\mu}{\sigma + \Delta\sigma} = \frac{\mu}{\sigma}\left(1 + \frac{\Delta\mu}{\mu}\right)\left(1 + \frac{\Delta\sigma}{\sigma}\right)^{-1} \approx \frac{\mu}{\sigma}\left(1 + \frac{\Delta\mu}{\mu} - \frac{\Delta\sigma}{\sigma}\right)$$
$$= \frac{\mu}{\sigma} + \frac{\Delta\mu}{\sigma} - \frac{\mu}{\sigma^2}\Delta\sigma$$

2.2 準備: χ^2 分布の平均と分散

$$X_i \sim N(\mu, \sigma^2) \qquad i = 1,2, \dots, n$$

とすると、

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

は自由度 $(n-1)$ の χ^2 分布に従う確率変数であり、

$$E\left\{\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2\right\} = n-1$$

$$E\left\{\frac{1}{\sigma^4}\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^2\right\} - (n-1)^2 = 2(n-1)$$

2.3 分散の推定量 V_n とその分散

$$V_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

とすると、

$$E(V_n) = \sigma^2$$

$$E(V_n^2) - \sigma^4 = \frac{2(n-1)}{(n-1)^2}\sigma^4$$

が得られ、

$$E(V_n^2) = \frac{(n-1)^2 + 2(n-1)}{(n-1)^2}\sigma^4 = \frac{n+1}{n-1}\sigma^4$$

$$E\{(V_n - \sigma^2)^2\} = E(V_n^2) - \sigma^4 = \frac{n+1-n+1}{n-1}\sigma^4 = \frac{2}{n-1}\sigma^4$$

2.4 t 値の推定量の分散$\left\{\left(\frac{\bar{X}}{S_n}-\frac{\mu}{\sigma}\right)^2\right\}$

$$S_n=\sqrt{V_n}$$

$$S_n-\sigma=\sqrt{V_n-\sigma^2+\sigma^2}-\sigma=\sigma\left(1+\frac{V_n-\sigma^2}{\sigma^2}\right)^{\frac{1}{2}}-\sigma\cong\frac{V_n-\sigma^2}{2\sigma}$$

$$E\{(S_n-\sigma)^2\}\cong\frac{1}{4\sigma^2}\,E\{(V_n-\sigma^2)^2\}=\frac{1}{4\sigma^2}\frac{2}{n-1}\sigma^4=\frac{\sigma^2}{2(n-1)}$$

$$E\left((\bar{X}-\mu)^2\right)=\frac{1}{n}\sigma^2$$

$$\frac{\bar{X}}{S_n}=\frac{\mu+\Delta\mu}{\sigma+\Delta\sigma}=\frac{\mu}{\sigma}\left(1+\frac{\Delta\mu}{\mu}\right)\left(1+\frac{\Delta\sigma}{\sigma}\right)^{-1}\approx\frac{\mu}{\sigma}\left(1+\frac{\Delta\mu}{\mu}-\frac{\Delta\sigma}{\sigma}\right)$$

$$=\frac{\mu}{\sigma}+\frac{\Delta\mu}{\sigma}-\frac{\mu}{\sigma^2}\Delta\sigma$$

$$\frac{\bar{X}}{S_n}-\frac{\mu}{\sigma}\cong\frac{\Delta\mu}{\sigma}-\frac{\mu}{\sigma^2}\Delta\sigma$$

$$E\left\{\left(\frac{\bar{X}}{S_n}-\frac{\mu}{\sigma}\right)^2\right\}\cong\frac{1}{\sigma^2}E\left((\bar{X}-\mu)^2\right)+\frac{\mu^2}{\sigma^4}E\{(S_n-\sigma)^2\}=\frac{1}{n}+\frac{1}{2(n-1)}\frac{\mu^2}{\sigma^2}$$

2.5 $\varphi_n(u)$

正規分布の密度関数 g を用いて、

$$h_n\left(\frac{\bar{X}}{S_n}|\frac{\mu}{\sigma}\right)=g\left(\frac{\bar{X}}{S_n}|\frac{\mu}{\sigma},\frac{1}{n}+\frac{1}{2(n-1)}\frac{\mu^2}{\sigma^2}\right)$$

を定義する。この式の $\frac{\bar{X}}{S_n}$ と $\frac{\mu}{\sigma}$ を入れ替えると

$$h_n\left(\frac{\mu}{\sigma}|\frac{\bar{X}}{S_n}\right)=g\left(\frac{\mu}{\sigma}|\frac{\bar{X}}{S_n},\frac{1}{n}+\frac{1}{2(n-1)}\frac{\bar{X}^2}{S_n^2}\right)$$

が得られる。

$$W(u)=\frac{1}{n}+\frac{1}{2(n-1)}u^2$$

とすると、

$$h_n(v|u)=g\left(v|u,W(u)\right)=\frac{1}{\sqrt{2\pi W(u)}}\mathrm{e}^{\left(-\frac{(v-u)^2}{2W(u)}\right)}$$

この $h_n(v|u)$ を用いて

$$\varphi_n(u)=\int_{-\infty}^{\infty}F_G(s)\,h_n(s\mid u)\,\mathrm{ds}$$

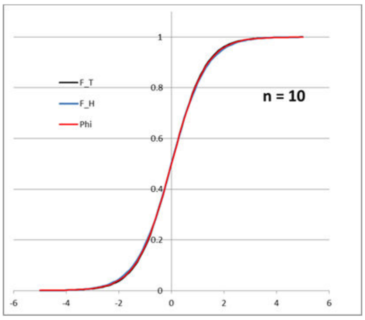
を定義する。

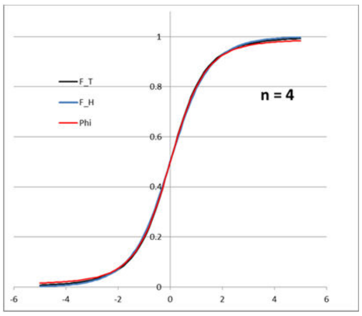
2.6　$\varphi_n(x) \cong \mathrm{F_T}(x)$

t-分布の分布関数 $\mathrm{F_T}(x)$ と $\varphi_n(x)$の関係は？同じもの？それとも「似て非なる」もの？

$\varphi_n(x)$ は $F_G(s)$ が下に凸な領域で $\varphi_n(x) > F_G(s)$ となり、$F_G(s)$ が上に凸な領域で $\varphi_n(x) < F_G(s)$ * < となるのが明らかで、$\mathrm{F_T}(x)$ とかなりよく似ている。

$n = 10$ の場合と $n = 4$ の場合 の両者のグラフを重ねてみると以下の結果が得られる。





この類似を偶然と見なすのは無理だろう。

2.7　$\varphi_n(x) \cong \mathrm{F_T}(x)$　の「証明」

我々は $h_n(y|x)$ を次の式で定義した。

$$h_n(y|x) = g\big(y|x, W(x)\big) = \frac{1}{\sqrt{2\pi W(x)}} \mathrm{e}^{\left(-\frac{(y-x)^2}{2W(x)}\right)}$$

$$b_w(z) = \frac{1}{\sqrt{2\pi W}} \mathrm{e}^{\left(-\frac{z^2}{2W}\right)}$$

とすると、

$$h_n(y|x) = b_{W(x)}(y - x)$$

と表すことができる。b_W を用いると $\varphi_n(x)$ を以下のように変形できる。

$$\varphi_n(x) = \int_{-\infty}^{\infty} F_G(y)\, h_n(y \mid x)\, \mathrm{dy} = \int_{-\infty}^{\infty} F_G(y)\, b_{W(x)}(y - x) dy$$
$$= \int_{-\infty}^{\infty} F_G(x + y)\, b_{W(x)}(y) dy = \int_{-\infty}^{\infty} \int_{-\infty}^{x+y} f_G(u)\, b_{W(x)}(y) du dy$$

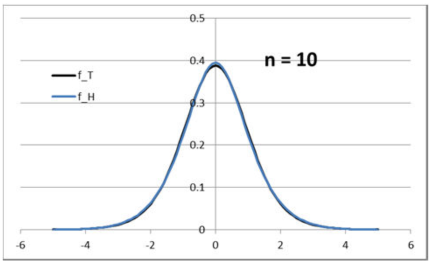
2.7.1 t分布の確率密度関数

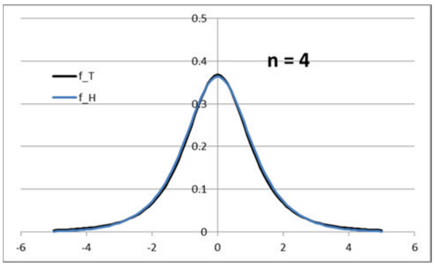
$$A(u) = \frac{2}{n+1} + \frac{1}{(n-1)} u^2$$

として、

$$f_H(u) = \int_{-\infty}^{\infty} f_G(y) b_{A(y)}(y-u)\mathrm{dy}$$

を定義する。$f_H(u)$ と t 分布の確率密度関数 $f_T(u)$ のグラフを重ねてプロットすると、下のような結果が得られる。





この結果は $f_T(u) \cong f_H(u)$ が成り立つことをを暗示している。

2.7.2 t分布の確率分布関数

$$F_H(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_G(y) b_{A(y)}(u-y)dydu = \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_G(y) b_{A(y)}(u-y)dudy$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{x-y} f_G(y) b_{A(y)}(u)dudy = \int_{-\infty}^{\infty}\int_{-\infty}^{x+y} f_G(y) b_{A(y)}(u)dudy$$

を定義する。2.6 節のグラフには、実は、$\varphi_n(x)$ と$\mathrm{F_T}(x)$ だけでなく、$F_H(x)$ も描きこまれている。３本の曲線が重なっていることは、

$$\int_{-\infty}^{\infty}\int_{-\infty}^{x+y} f_G(y) b_{A(y)}(u)dudy \cong \int_{-\infty}^{\infty}\int_{-\infty}^{x+y} f_G(u)\, b_{W(x)}(y)dudy$$

という近似関係と、

$$f_T(u) \cong f_H(u)$$

という2つの近似関係が成り立っていることを暗示している。この2つの近似関係が解析的に証明されれば

$$\varphi_n(x) \cong \mathrm{F_T}(x)$$

が解析的に証明されたことになる。

2.7.3 $\varphi_n(x) \cong F_H(x)$ の「証明」

$\psi_A(y,u) = f_G(y) b_{A(y)}(u)$ と $\psi_W(y,u) = f_G(u) b_{W(x)}(y)$ で定義される2つの2次元確率密度関数が $(y, y+x)$ で定義される直線の下の領域で近似的に同じ確率を持つことが証明できればよい。変数 (y,u) を

$$y = (s+t)/\sqrt{2}$$
$$u = (s-t)/\sqrt{2}$$

で変換すると、$u < y + x$ が $t > -x/\sqrt{2}$ と等価になるので、

$$\int_{-\infty}^{\infty}\int_{-x/\sqrt{2}}^{\infty} f_G\left((s+t)/\sqrt{2}\right) b_{A\left((s+t)/\sqrt{2}\right)}\left((s-t)/\sqrt{2}\right) dtds$$
$$\cong \int_{-\infty}^{\infty}\int_{-x/\sqrt{2}}^{\infty} f_G\left((s-t)/\sqrt{2}\right) b_{W(x)}\left((s+t)/\sqrt{2}\right) dt\, ds$$

がすべての x に対して証明できればよいことになる。

これは、$t = -x/\sqrt{2}$ において

$$\int_{-\infty}^{\infty} f_G\left((s+t)/\sqrt{2}\right) b_{A\left((s+t)/\sqrt{2}\right)}\left((s-t)/\sqrt{2}\right) ds$$
$$\cong \int_{-\infty}^{\infty} f_G\left((s-t)/\sqrt{2}\right) b_{W(x)}\left((s+t)/\sqrt{2}\right) ds$$

が成り立てばOK.

2.7.4 $f_H(x) \cong \frac{d}{dx}\varphi_n(x)$

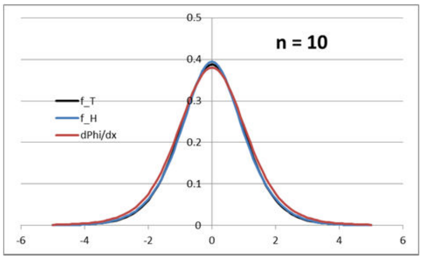
$t(x) = -x/\sqrt{2}$ として、

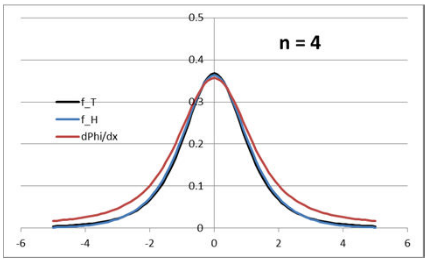
$$f_H(x) = \int_{-\infty}^{\infty} f_G\left(\frac{s+t(x)}{\sqrt{2}}\right) b_{A\left(\frac{(s+t)}{\sqrt{2}}\right)}\left(\frac{s-t(x)}{\sqrt{2}}\right) ds$$

と

$$\int_{-\infty}^{\infty} f_G\left(\frac{s-t(x)}{\sqrt{2}}\right) b_{W(x)}\left(\frac{s+t(x)}{\sqrt{2}}\right) ds = \frac{d}{dx}\varphi_n(x)$$

のグラフを描くと以下のようになる。





2.7.5 解析的証明と図形的証明

2.6 節のグラフに見える $\varphi_n(x)$ と $\mathrm{F_T}(x)$ の「一致」が偶然の結果であると考えるのは無理と思われる。

なんらかの形で、

$$\varphi_n(x) \cong \mathrm{F_T}(x)$$

の解析的証明ができると想像されるが、いまのところ未完である。そのような証明が得られた場合には、おそらく

$$\varphi_n(x) = \mathrm{F_T}(x) + \text{誤差項}$$

という形が「証明」され、誤差項が「小さい」く

$$\varphi_n(x) \cong \mathrm{F_T}(x)$$

が成り立つことが「図形的」に示されることになるのだろうが……

3. t検定の「意味」

$$\varphi_n(x) \cong \mathrm{F_T}(x)$$

という関係を手がかりに t 検定の意味を考えることができる。

$\varepsilon > 0$ とし

$$F_T(t_b) = 1 - \varepsilon$$

$$F_T(t_a) = \varepsilon$$

とする。

$F_T(t)$ の単調性から $(\hat{\mu} - m)/\hat{\sigma}$ が t_b より大きい場合、正規確率変数 X の実現値が m 以上になる確率は $1 - \varepsilon$ より高くなる。

また、 $(\hat{\mu} - m)/\hat{\sigma}$ が t_a より小さい場合、正規確率変数 X の実現値が m 以下になる確率は $1 - \varepsilon$ より高くなる。

3.1 応用

t分布の分布関数を F_T で表すことにすると、
- 正規確率変数の実現値が b 以上になる確率が $F_T((\hat{\mu}-b)/\hat{\sigma})$ となること、
- 正規確率変数の n 個の実現値から求めた算術平均が b 以上になる確率が $F_T(\sqrt{n}(\hat{\mu}-b)/\hat{\sigma})$であること

等々が言える。

確率変数 X について「$X>0$」を認定する方法は、ただちに、2つの確率変数 X と Y について「$X>Y$」を認定する方法として使える。$\hat{\mu}$ と $\hat{\sigma}^2$ を確率変数 $X-Y$ の期待値と分散の推定量と読み替えればいいだけである。

正規確率変数 X の実現値が a 以下になる確率は、正規確率変数 $-$X の実現値が $-a$ 以上になる確率に等しく $F_T((-\hat{\mu}+a)/\hat{\sigma})$ となる。この値は、F_T 関数の対称性から $1-F_T((\hat{\mu}-a)/\hat{\sigma})$に等しい。

$a<b$ の場合、正規確率変数 X の実現値が b 以上あるいは a 以下になる確率は

$$1-F_T((\hat{\mu}-a)/\hat{\sigma})+F_T((\hat{\mu}-b)/\hat{\sigma})$$

となる。これは正規確率変数 X の実現値が a 以上かつ b 以下である確率が

$$F_T((\hat{\mu}-a)/\hat{\sigma})-F_T((\hat{\mu}-b)/\hat{\sigma})$$

であることを意味する。

3.2 個人的問題と社会的問題

「$X>0$ の認定」という問題を「個人的問題」、「$\sum_{i=1}^{n} X_i>0$ の認定」という問題を「社会的問題」と呼ぶことにしよう。

$$E(X)\neq 0$$

である場合、個人的問題に対する答えと、社会的問題に対する答えは必ずしも一致しない。 このメモにおける定義を適用したとき、「$X>0$ 」とは認定できないけれど、「$\sum_{i=1}^{n} X_i>0$ 」と認定されるべき場合があり得る。

たとえば、ある薬剤の個人的レベルでの効果はばらつきが多く目立たないが、社会的にははっきりとした効果を示すようなことである。

平均値の検定を利用するときに留意すべきことであろう。