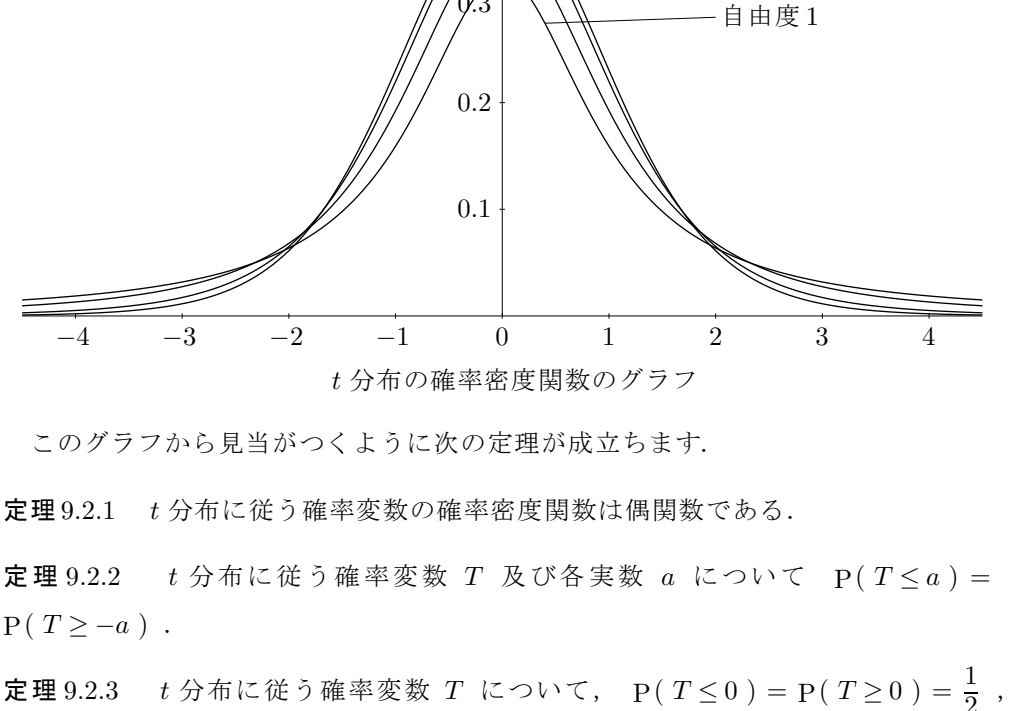


§9.2 t分布

正の自然数 n に対して、標準正規分布 $N(0,1)$ に従う確率変数 Z と自由度 n の χ^2 分布に従う確率変数 X とが互いに独立であるとき、確率変数 $Z\sqrt{\frac{n}{X}}$ が従う確率分布を自由度 n の t 分布といいます。 t 分布の確率密度関数の式は後にしてまずグラフを描きます。



このグラフから見当がつくように次の定理が成立します。

- 定理 9.2.1** t 分布に従う確率変数の確率密度関数は偶関数である。
- 定理 9.2.2** t 分布に従う確率変数 T 及び各実数 a について $P(T \leq a) = P(T \geq -a)$ 。
- 定理 9.2.3** t 分布に従う確率変数 T について、 $P(T \leq 0) = P(T \geq 0) = \frac{1}{2}$ 、 $E[T] = 0$ 、

正の自然数 n 及び実数 μ 及び正の実数 σ に対して、 n 個 ($n \geq 2$) の確率変数 $X_1, X_2, X_3, \dots, X_n$ は総て同じ正規分布 $N(\mu, \sigma^2)$ に従い互いに独立であるとして、確率変数 \bar{X} は $N\left(\mu, \frac{\sigma^2}{n}\right)$ に従うので、確率変数

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}$$

は標準正規分布 $N(0,1)$ に従います。また、定理 9.1.1 より、確率変数 $X_1, X_2, X_3, \dots, X_n$ の偏差平方和を表す確率変数 $S = \sum_{k=1}^n (X_k - \bar{X})^2$ に対して確率変数 $X = \frac{S}{\sigma^2}$ は自由度 $n-1$ の χ^2 分布に従います。故に、 t 分布の定義より、確率変数

$$Z\sqrt{\frac{n-1}{X}} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sqrt{\frac{n-1}{\frac{S}{\sigma^2}}} = \sqrt{\frac{n(n-1)}{S}}(\bar{X} - \mu)$$

は自由度 $n-1$ の t 分布に従います。約分されて σ が消えることが重要です。

定理 9.2.4 n 個 ($n \geq 2$) の確率変数 $X_1, X_2, X_3, \dots, X_n$ は総て同じ正規分布に従い、互いに独立であるとする。それらの平均値を μ とおく。平均を表す確率変数 $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ 及び偏差平方和を表す確率変数 $S = \sum_{k=1}^n (X_k - \bar{X})^2$ に対して、確率変数 $\sqrt{\frac{n(n-1)}{S}}(\bar{X} - \mu)$ は自由度 $n-1$ の t 分布に従う。

この定理を標本分布に適用します。母平均が μ である正規母集団から無作為抽出した大きさ n ($n \geq 2$) の標本の標本確率変数 $X_1, X_2, X_3, \dots, X_n$ は、総て同じ正規分布に従い、互いに独立で、それら平均値は μ です。定理 9.2.4 より、標本確率変数 $X_1, X_2, X_3, \dots, X_n$ の平均を表す確率変数 $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ 及び $X_1, X_2, X_3, \dots, X_n$ の偏差平方和を表す確率変数 $S = \sum_{k=1}^n (X_k - \bar{X})^2$ に対して、確率変数 $\sqrt{\frac{n(n-1)}{S}}(\bar{X} - \mu)$ は自由度 $n-1$ の t 分布に従います。 $X_1, X_2, X_3, \dots, X_n$ の不偏分散を表す確率変数 $U = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X})^2 = \frac{S}{n-1}$ について、 $S = U(n-1)$ なので、

$$\sqrt{\frac{n(n-1)}{S}}(\bar{X} - \mu) = \sqrt{\frac{n(n-1)}{U(n-1)}}(\bar{X} - \mu) = \sqrt{\frac{n}{U}}(\bar{X} - \mu)$$

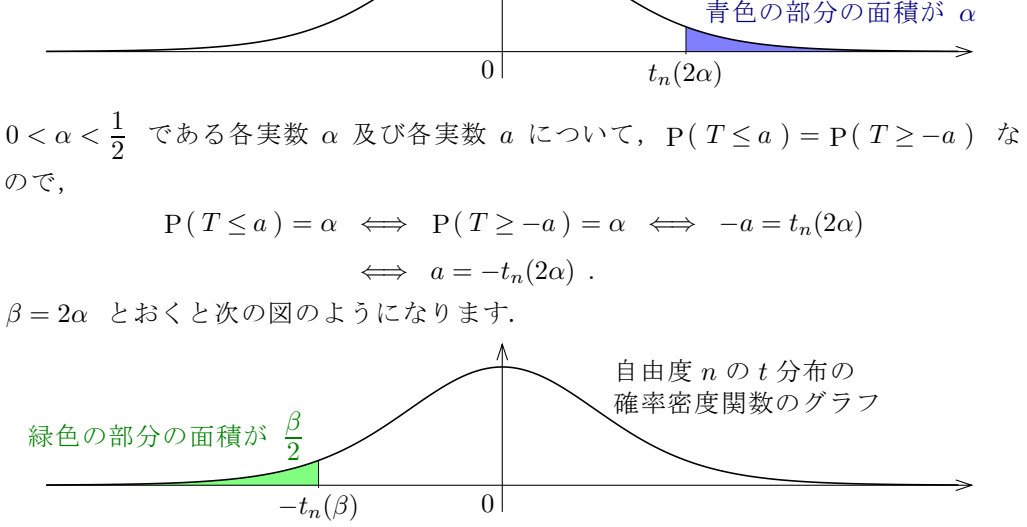
この確率変数が自由度 $n-1$ の t 分布に従います。

定理 9.2.5 母平均が μ である正規母集団から無作為抽出した大きさ n ($n \geq 2$) の正規標本において、平均を表す確率変数を \bar{X} とおき、不偏分散を表す確率変数を U とおくと、確率変数 $\sqrt{\frac{n}{U}}(\bar{X} - \mu)$ は自由度 $n-1$ の t 分布に従う。

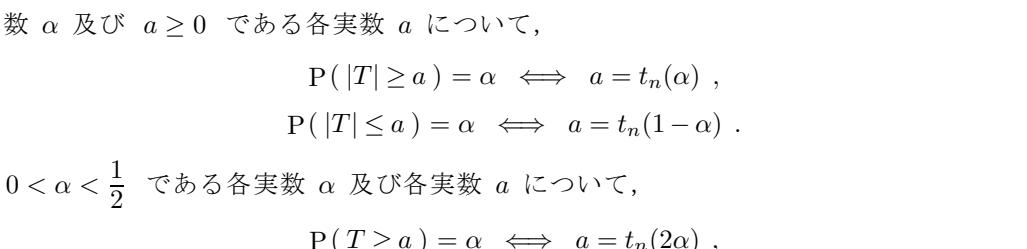
正の自然数 n に対して、確率変数 T は自由度 n の t 分布に従うとします。証明は略しますが、 $0 < \alpha < 1$ である各実数 α に対して、 $P(|T| \geq k) = \alpha$ となる 0 以上の実数 k が唯一つ定まります。この実数 k を、自由度 n の t 分布の両側 α 点といい、 $t_n(\alpha)$ と書き表します。つまり次のことが成り立ちます： $0 < \alpha < 1$ である各実数 α 及び 0 以上の各実数 a について、

$$P(|T| \geq a) = \alpha \iff a = t_n(\alpha)$$

確率密度関数のグラフを考えると次の図のようになります。



$0 < \alpha < 1$ である各実数 α 及び 0 以上の各実数 a について、
 $P(|T| \leq a) = \alpha \iff 1 - P(|T| \geq a) = \alpha \iff P(|T| \geq a) = 1 - \alpha$
 $\iff a = t_n(1 - \alpha)$ 。



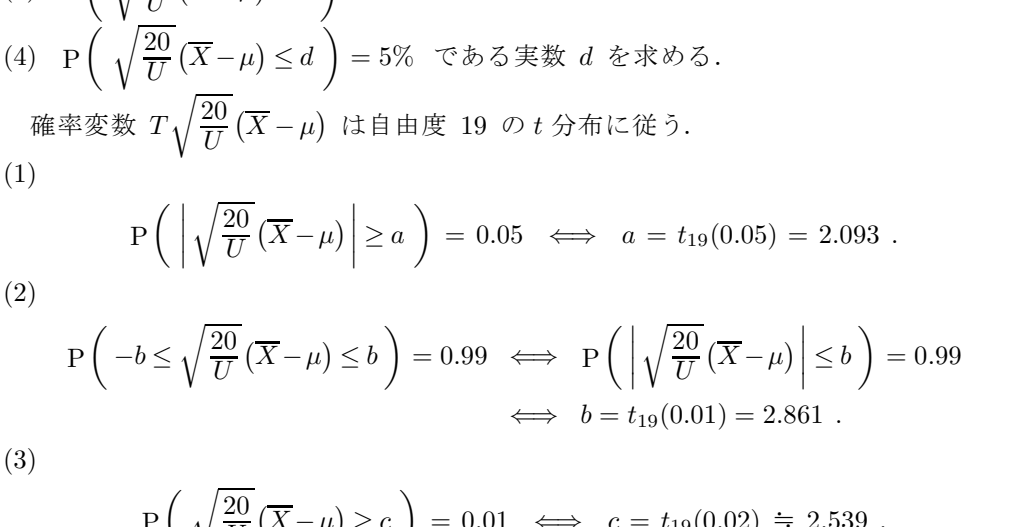
実数 a について、 $a \leq 0$ ならば
 $P(T \geq a) = P(T \geq 0) + P(0 \geq T \geq a) \geq P(T \geq 0) = \frac{1}{2}$ 、
 対偶をとると、 $P(T \geq a) < \frac{1}{2}$ ならば $a > 0$ 。更に、 $a \geq 0$ ならば

$$P(|T| \geq a) = P(T \leq -a) + P(T \geq a) = P(T \geq a) + P(T \geq a) = 2P(T \geq a)$$

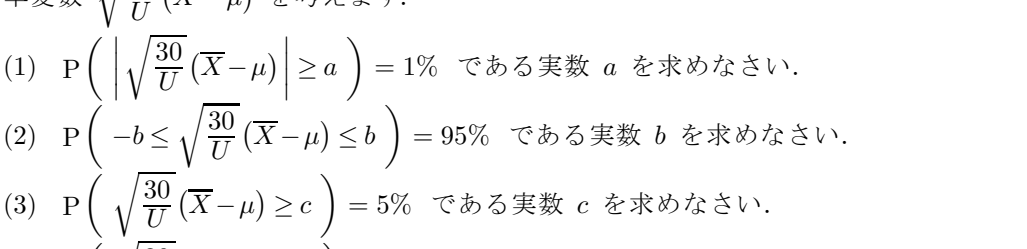
実数 α について $0 < \alpha < \frac{1}{2}$ とします。 $P(T \geq a) = \alpha$ ならば、 $P(T \geq a) < \frac{1}{2}$ 、 $a > 0$ なので $P(|T| \geq a) = 2P(T \geq a) = 2\alpha$ 、よって $a = t_n(2\alpha)$ 。逆に、 $a = t_n(2\alpha)$ ならば、 $P(|T| \geq a) = 2\alpha$ 、 $a \geq 0$ より $P(|T| \geq a) = 2P(T \geq a)$ なので $2P(T \geq a) = 2\alpha$ 、よって $P(T \geq a) = \alpha$ 。このように、 $0 < \alpha < \frac{1}{2}$ である各実数 α について

$$P(T \geq a) = \alpha \iff a = t_n(2\alpha)$$

$\beta = 2\alpha$ とおくと次の図のようになります。



$0 < \alpha < \frac{1}{2}$ である各実数 α 及び各実数 a について、
 $P(T \leq a) = \alpha \iff P(T \geq -a) = \alpha \iff -a = t_n(2\alpha)$
 $\iff a = -t_n(2\alpha)$ 。



このようにして次の定理が導かれます。

定理 9.2.6 確率変数 T は自由度 n の t 分布に従うとする。 $0 < \alpha < 1$ である各実数 α 及び $a \geq 0$ である各実数 a について、
 $P(|T| \geq a) = \alpha \iff a = t_n(\alpha)$ 、
 $P(|T| \leq a) = \alpha \iff a = t_n(1 - \alpha)$ 。

$0 < \alpha < \frac{1}{2}$ である各実数 α 及び各実数 a について、
 $P(T \geq a) = \alpha \iff a = t_n(2\alpha)$ 、
 $P(T \leq a) = \alpha \iff a = -t_n(2\alpha)$ 。

定理 9.2.5 と定理 9.2.6 とを併せませす。

定理 9.2.7 母平均が μ である正規母集団から無作為抽出した大きさ n ($n \geq 2$) の正規標本において、平均を表す確率変数を \bar{X} とおき、不偏分散を表す確率変数を U とおくと、 $0 < \alpha < 1$ である各実数 α 及び $a \geq 0$ である各実数 a について、
 $P\left(\left|\sqrt{\frac{n}{U}}(\bar{X} - \mu)\right| \geq a\right) = \alpha \iff a = t_{n-1}(\alpha)$ 、
 $P\left(\left|\sqrt{\frac{n}{U}}(\bar{X} - \mu)\right| \leq a\right) = \alpha \iff a = t_{n-1}(1 - \alpha)$ 。

$0 < \alpha < \frac{1}{2}$ である各実数 α 及び各実数 a について、
 $P\left(\sqrt{\frac{n}{U}}(\bar{X} - \mu) \geq a\right) = \alpha \iff a = t_{n-1}(2\alpha)$ 、
 $P\left(\sqrt{\frac{n}{U}}(\bar{X} - \mu) \leq a\right) = \alpha \iff a = -t_{n-1}(2\alpha)$ 。

正の自然数 n と各実数 α ($0 \leq \alpha \leq \frac{1}{2}$) とに対する自由度 n の t 分布の両側 α 点 $t_n(\alpha)$ の値を表にしたものを t 分布表といいます。

例題 9.2 母平均が μ である正規母集団から無作為に非復元抽出した大きさ 20 の正規標本の平均を表す確率変数 \bar{X} 及び不偏分散を表す確率変数 U に対して、確率変数 $\sqrt{\frac{20}{U}}(\bar{X} - \mu)$ を考える。

- $P\left(\left|\sqrt{\frac{20}{U}}(\bar{X} - \mu)\right| \geq a\right) = 5\%$ である実数 a を求める。
- $P\left(-b \leq \sqrt{\frac{20}{U}}(\bar{X} - \mu) \leq b\right) = 99\%$ である実数 b を求める。
- $P\left(\sqrt{\frac{20}{U}}(\bar{X} - \mu) \geq c\right) = 1\%$ である実数 c を求める。
- $P\left(\sqrt{\frac{20}{U}}(\bar{X} - \mu) \leq d\right) = 5\%$ である実数 d を求める。

確率変数 $T = \sqrt{\frac{20}{U}}(\bar{X} - \mu)$ は自由度 19 の t 分布に従う。

(1) $P\left(\left|\sqrt{\frac{20}{U}}(\bar{X} - \mu)\right| \geq a\right) = 0.05 \iff a = t_{19}(0.05) = 2.093$ 。

(2) $P\left(-b \leq \sqrt{\frac{20}{U}}(\bar{X} - \mu) \leq b\right) = 0.99 \iff P\left(\left|\sqrt{\frac{20}{U}}(\bar{X} - \mu)\right| \leq b\right) = 0.99$
 $\iff b = t_{19}(0.01) = 2.861$ 。

(3) $P\left(\sqrt{\frac{20}{U}}(\bar{X} - \mu) \geq c\right) = 0.01 \iff c = t_{19}(0.02) \approx 2.539$ 。

(4) $P\left(\sqrt{\frac{20}{U}}(\bar{X} - \mu) \leq d\right) = 0.05 \iff d = -t_{19}(0.1) = -1.729$ 。 □終

問題 9.2 母平均が μ である正規母集団から無作為に非復元抽出した大きさ 30 の正規標本の平均を表す確率変数 \bar{X} 及び不偏分散を表す確率変数を U に対して、確率変数 $\sqrt{\frac{30}{U}}(\bar{X} - \mu)$ を考えます。

- $P\left(\left|\sqrt{\frac{30}{U}}(\bar{X} - \mu)\right| \geq a\right) = 1\%$ である実数 a を求めなさい。
- $P\left(-b \leq \sqrt{\frac{30}{U}}(\bar{X} - \mu) \leq b\right) = 95\%$ である実数 b を求めなさい。
- $P\left(\sqrt{\frac{30}{U}}(\bar{X} - \mu) \geq c\right) = 5\%$ である実数 c を求めなさい。
- $P\left(\sqrt{\frac{30}{U}}(\bar{X} - \mu) \leq d\right) = 1\%$ である実数 d を求めなさい。

次の定理が成り立ちます。

定理 自然数 n が充分大きいとき、自由度 n の t 分布に従う確率変数 T は近似的に標準正規分布 $N(0,1)$ に従う。

t 分布の確率密度関数

正の自然数 n に対して、自由度 n の t 分布の確率密度関数 $f(x)$ は次のようになります：

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)\left(1 + \frac{x^2}{n}\right)^{\frac{n+1}{2}}}$$

ここで関数 $\Gamma(x)$ は前節で述べた Γ 関数です：

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt \quad (x > 0)$$