確率分布　Q2500 t 分布

組　　　　番　名前

標本から母平均を推定する際、これまで学んだ知識では、

標本平均の分布

母集団が正規分布 $N(m, \sigma^2)$ に従うとき、
標本平均 $\bar{X}$ の分布：$\bar{X} \sim N\left(m, \frac{\sigma^2}{n}\right)$
標準化すると：

$$Z = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

実際のデータ分析では、母標準偏差 σ は**未知**であるため、σ の代用として、標本標準偏差 s を用いる。

$$t = \frac{\bar{X} - m}{\frac{s}{\sqrt{n}}}$$

この σ の代わりに s で置き換えた統計量を t とすると、これは Z とは違うので t は厳密には標準正規分布 $N(0, 1)$ には**従わない**。では、どのような分布に従うのだろうか？

t 分布

標本平均とその標本標準偏差から母平均を推定するために正規分布を修正した確率分布として t 分布が作られた。

t 分布の定義

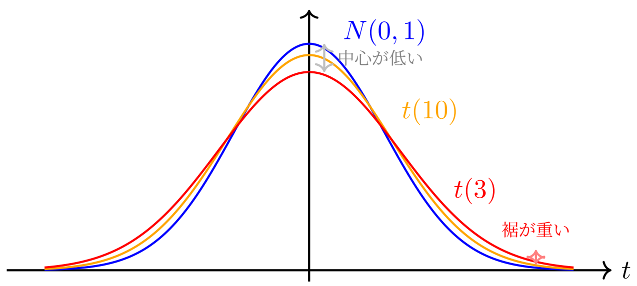
母集団が正規分布 $N(m, \sigma^2)$ に従い、
標本サイズが n のとき、

$$t = \frac{\bar{X} - m}{s/\sqrt{n}}$$

と定義される統計量 t は、**自由度 $(n-1)$ の t 分布**に従う。　　記号：$t \sim t(n-1)$
(n：標本サイズ、$(n-1)$：自由度)

t 分布の性質：

- 平均 0、左右対称（正規分布と似ている）
- 正規分布 $N(0, 1)$ より**裾が重い**
- 自由度 $(n-1)$ が大きくなると、正規分布 $N(0, 1)$ に近づく
- 自由度が小さいほど、裾が重い



【なぜ裾が重いのか】

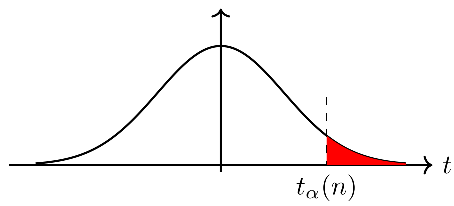
σ が既知なら、標本平均の標準誤差 $\frac{\sigma}{\sqrt{n}}$ は確定。
σ が未知で s を使うと、s 自体がランダムに変動する。
→ 不確実性が増す → 分布の裾が重くなる

t 分布表の使い方

t 分布は自由度ごとに異なる分布なので、表から値を読み取る。

【t 分布表（抜粋）】

自由度	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
10	1.372	1.812	2.228	2.764	3.169
20	1.325	1.725	2.086	2.528	2.845
30	1.310	1.697	2.042	2.457	2.750
∞	1.282	1.645	1.960	2.326	2.576

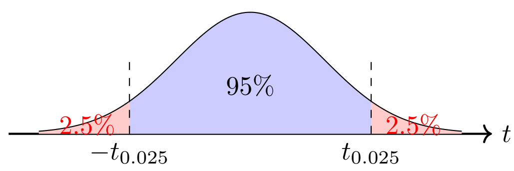


読み方：

$t_\alpha(k)$：自由度 k の t 分布で、右側の面積が α となる値
例：$t_{0.025}(4) = 2.776$
→ 自由度 4 の t 分布で、$P(t > 2.776) = 0.025$

95% 信頼区間の場合：

両側 5%（片側 2.5%）なので、$t_{0.025}(n-1)$ を使う。



例 1　自由度が 10 のとき、次の値を t 分布表から読みとりなさい。
(1) $t_{0.025}(10) =$
(2) $t_{0.005}(10) =$

問 1　自由度が 30 のとき、次の値を t 分布表から読みとりなさい。
(1) $t_{0.025}(30) =$
(2) $t_{0.005}(30) =$

++*+*+*+*+*+*+ **【解答】** *+*+*+*+*+*+*+*+

例 1　自由度が 10 のとき、次の値を t 分布表から読みとりなさい。

(1) $t_{0.025}(10) = \boxed{2.228}$

(2) $t_{0.005}(10) = \boxed{3.169}$

問 1　自由度が 30 のとき、次の値を t 分布表から読みとりなさい。

(1) $t_{0.025}(30) = \boxed{2.042}$

(2) $t_{0.005}(30) = \boxed{2.750}$