

1) 身長が分布が $N(170.5, 7.5^2)$ 【男性の場合】、 $N(156.5, 6.5^2)$ 【女性の場合】、に従うと仮定できる場合、自分の身長《○○○ cm (小数第1位を四捨五入)》以上になる確率 $P(○○○ \leq X)$ を求めよ

```
> 1-pnorm(○○○,170.5,7.5) または > 1-pnorm(○○○,156.5,6.5)
で確率が求められる。 《> 1-pnorm((○○○-170.5)/7.5) でも可》
```

主な確率の値

身長	150	155	160	165	170	175	180	185	190
確率 (男性)	0.997	0.981	0.919	0.768	0.527	0.274	0.103	0.0266	4.66×10^{-3}
確率 (女性)	0.841	0.591	0.295	0.0955	0.0189	2.21×10^{-3}	—	—	—

2) 母集団が母分散 $(3.25)^2$ の正規分布に従うとき、9つのデータ 32.8, 41.5, 32.5, 38.3, 37.0, 40.6, 36.9, 33.1, 36.7 が与えられたときの、標本平均を計算し、母平均 μ の 95%信頼区間を求めよ。

```
> x <- c(32.8, 41.5, 32.5, 38.3, 37.0, 40.6, 36.9, 33.1, 36.7)
> mean(x)
[1] 36.6
> var(x)
[1] 10.8325
```

ここまでで標本平均 $\bar{x} = 36.6$ 、標本分散 $s^2 = 10.83 \dots$ ($s = 3.291 \dots$) がわかる。95%信頼区間を求めるための ε は

```
> qnorm(1-0.05/2) 《> qnorm((1+0.95)/2) でも可》
[1] 1.959964
```

によって求められるので、信頼区間は

```
> 36.6-1.959964*3.25/sqrt(9)
[1] 34.47671
> 36.6+1.959964*3.25/sqrt(9)
[1] 38.72329
```

のように [34.5, 38.7] であることがわかる。

もし、99%信頼区間を求めるのであれば、

```
> mean(x)-qnorm(1-0.01/2)*3.25/sqrt(9)
[1] 33.80952
> mean(x)+qnorm(1-0.01/2)*3.25/sqrt(9)
[1] 39.39048
```

によって、[33.8, 39.4] である。

- 注) 1. 解答は有効数字3桁以上で答えること。
 2. データの個数が9なのに例題のまま $\sqrt{10}$ とする人がいた。
 3. 区間推定の答えは区間 $[a, b]$ または $a \leq \mu \leq b$ で答えること。
 《 (a, b) や $a < \mu < b$ でも可》
 4. 母分散 $\sigma^2 = (3.25)^2$ なのに、例題のまま3を使う人が多かった。
 5. 母分散既知なのに、 s の値を使う人がいた。
 6. $qnorm$ とまちがえて $pnorm$ を使い $\varepsilon = 0.8352$ としないように注意!!

演習で使うデータ

学籍番号	1 番目のデータ	18 番目のデータ	学籍番号	1 番目のデータ	18 番目のデータ
1822008	39.2	31.3	1822074	33.8	34.3
1822015	36.7	35.3	1822078	35.8	43.6
1822024	35.7	41.8	1822094	40.6	39.6
1822031	42.7	38.2	1822096	37.4	29.2
1822050	28.3	35.1	1822097	37.5	37.1
1822065	33.9	33.5	1822100	31.6	40.4

・平均の区間推定（母分散未知）

基本的には母分散既知と同じ考え方であるが、母分散 σ^2 の値がわからない（未知）のため

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

での変換ができない。そこで Z の変換において、母分散 σ^2 の代わりにその推定量である標本分散 s^2 を使うことを考える。つまり

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

と変換する。このとき、 T は標準正規分布ではなく、自由度 $n - 1$ の t 分布に従うことがわかっている。よって $\Pr\{-z_\alpha \leq Z \leq z_\alpha\} = 1 - \alpha$ ではなく、

$$\Pr\{-t_\alpha \leq T \leq t_\alpha\} = 1 - \alpha$$

になるように t_α を求めてやればよい。確率の括弧の中身は母分散既知と同様に

$$-t_\alpha \leq \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \leq t_\alpha \Rightarrow \bar{X} - t_\alpha \times \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_\alpha \times \frac{S}{\sqrt{n}}$$

と変形できるので、 $[\bar{x} - t_\alpha \times \frac{s}{\sqrt{n}}, \bar{x} + t_\alpha \times \frac{s}{\sqrt{n}}]$ の間に真の平均 μ が入っている確率は、事前に決めた確率にすることができる。

・ $\Pr\{-t_\alpha \leq T \leq t_\alpha\} = 1 - \alpha$ となる t_α の求め方

t 分布は教科書 P.54 にあるように、ほとんど標準正規分布と変わらない形をしている。母分散を推定しているために、誤差が大きくなる分だけ正規分布に比べて少しつぶれた形をしている。自由度 ($= n - 1$) が大きいほど正規分布に近づき、 $n \rightarrow \infty$ のとき t 分布は正規分布となる。コンピュータが無い場合は、 t 分布の両側 $100\alpha\%$ 点の表を使って t_α を求めることになるが、R では正規分布と同様に T が t 分布に従うとき、 $\Pr\{T \leq t_\beta\} = \beta$ となる t_β は $\text{qt}(\beta, \text{自由度})$ で与えられる。データの個数が7の場合、自由度が $6 (= 7 - 1)$ となるので95%信頼区間を求めるには、 $\Pr\{T \leq t_\alpha\} = 1 - 0.05/2 = 0.975$ となる t_α を

```
> qt(1-0.05/2,6)
[1] 2.446912
```

のように求める。データの数が少ないので、標準正規分布の 1.96 に比べるとかなり大きな数字であることがわかる。データの数を増やしていくと

```
> qt(1-0.05/2,50)
[1] 2.008559
> qt(1-0.05/2,100)
[1] 1.983972
```

と正規分布の場合の値である 1.96 に近づいていく。

例. 母平均が未知の場合の区間推定

ある工場で作られた電球7個の寿命を調べたところ、下記のような結果を得た。

76.5, 82.4, 93.0, 78.7, 86.6, 94.2, 82.9 (単位: 日)

このとき、母平均 μ の99%信頼区間を求めよ。

解答例)

標本平均と標本分散を計算すると、

```
> x <- c(76.5, 82.4, 93.0, 78.7, 86.6, 94.2, 82.9)
> mean(x)
[1] 84.9
> var(x)
[1] 45.70667
> sd(x)
[1] 6.76067
```

となるので、標本平均 $\bar{x} = 84.9$ 、標本分散 $s^2 = 45.70667$ 、標本標準偏差 $s = 6.76067$ であることがわかる。母分散が未知なので、

$$T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

と変換すると、 t は自由度 $7 - 1 = 6$ の t 分布に従うことになる。

よって、 $\Pr\{-t_\alpha \leq T \leq t_\alpha\} = 0.99 = 1 - 0.01$ となる t_α は

$$\Pr\{T \leq t_\alpha\} = 1 - \frac{0.01}{2} = 0.995 \Rightarrow \text{qt}(1-0.01/2, 6) = 3.707428$$

となる。よって母平均の99%信頼区間は

$$-3.707428 \leq \frac{84.9 - \mu}{\frac{6.76067}{\sqrt{7}}} \leq 3.707428$$

の不等式を解いて、

$$84.9 - 3.707428 \times \frac{6.76067}{\sqrt{7}} \leq \mu \leq 84.9 + 3.707428 \times \frac{6.76067}{\sqrt{7}}$$

$$75.42643 \leq \mu \leq 94.37357$$

$$75.4 \leq \mu \leq 94.4$$

と求められる。

中間レポートについて

次回11月20日(水)に問題配布、12月4日(水)締め切りで中間レポートを行います。

・母分散が未知の場合の母平均の区間推定

次のように作成した18個のデータに対して、母分散が未知の場合の母平均の99%信頼区間を求めよ。

```
> set.seed(学籍番号)
> x <- round(rnorm(18,35,4.3),digits=1)
```

注) 1番目のデータと18番目のデータは別紙に一覧表で載せておくので確認すること

2019年度神奈川工科大学 数理統計学 演習問題	学科	学年	組	学 籍 番 号	氏 名	

提出先：K3-3309号室前 19番のボックス 提出期限：11月14日（木）17時頃まで