確率統計 前回の解答

2019.11. 8 分

・母分散が既知の場合の区間推定

母集団が母分散 $\sigma^2 = 2.75^2$ の正規分布に従うとき、９つのデータ

34.7, 39.2, 42.9, 40.1, 36.6, 41.5, 35.2, 39.2, 38.0

が与えられたとき。

(1) 関数電卓を使い標本平均 $\bar{x}$ 及び標本分散 s^2 を求めよ。

標本平均 $\bar{x}$ と標本分散 s^2 を求めると、

$$\bar{x} = \frac{34.7 + 39.2 + 42.9 + 40.1 + 36.6 + 41.5 + 35.2 + 39.2 + 38.0}{9} = \frac{347.4}{9} = 38.6$$

$$s^2 = \frac{1}{9-1}\left(13470.64 - 9 \times 38.6^2\right) = \frac{61}{8} = 7.625\ ,\quad s = 2.7613\cdots$$

参考）標本分散の通常電卓を使った計算方法

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2) = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)$$

であるから、データの２乗和 $\displaystyle\sum_{i=1}^{n} x_i^2$ を計算できれば良い。

(2) 母分散が既知の場合の 99%信頼区間を求めよ。　母分散が $\sigma^2 = (2.75)^2$ と既知なので Z を使って変換すると

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

は標準正規分布 $N(0,1)$ に従うので、$P(-\varepsilon \leqq Z \leqq \varepsilon) = 0.99 = 1 - 0.01$ となる ε は 2.576 である。つまり信頼区間は

$$-2.576 \leqq \frac{38.6 - \mu}{\frac{2.75}{\sqrt{9}}} \leqq 2.576$$

を満たせばよいので、この不等式を μ について解くと、

$$-2.576 \times \frac{2.75}{\sqrt{9}} \leqq 38.6 - \mu \leqq 2.576 \times \frac{2.75}{\sqrt{9}}$$

$$38.6 - 2.576 \times \frac{2.75}{\sqrt{9}} \leqq \mu \leqq 38.6 + 2.576 \times \frac{2.75}{\sqrt{9}}$$

$$38.6 - 2.3613\cdots \leqq \mu \leqq 38.6 + 2.3613\cdots$$

$$36.2386\cdots \leqq \mu \leqq 40.9613\cdots$$

となる。したがって、有効数字３桁で答えると $36.2 \leqq \mu \leqq 41.0$ または $[36.2, 41.0]$ となる。

注）途中計算は有効数字より１〜２桁多めに計算した方が良い。

確率統計　区間推定

・母平均 μ に関する区間推定

・母分散が未知の場合

基本的には母分散既知と同じ考え方であるが、母分散が未知のため σ の値を使う Z での変換ができない。そこで母分散 σ^2 の代わりに標本分散 s^2 を使って

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

と変換を行う。この t は標準正規分布ではなく、自由度 $n-1$ の t 分布に従うことが知られている。そこで $100(1-\alpha)\%$ の信頼区間を求める場合、

$$P(-t_\alpha \leqq T \leqq t_\alpha) = 1 - \alpha$$

になるように t_α を決める。この t_α は表をつかって求めることになる。

確率の括弧の中身は母分散既知のときと同様に

$$-t_\alpha \leqq \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \leqq t_\alpha \quad \Rightarrow \quad \bar{x} - t_\alpha \times \frac{s}{\sqrt{n}} \leqq \mu \leqq \bar{x} + t_\alpha \times \frac{s}{\sqrt{n}}$$

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、事前に決めた確率となる。

例. 母集団が母分散未知の正規分布に従うとする。そこから無作為に 10 個のデータが得られ、標本平均が 17.44、標本分散が $(2.95)^2$ だったとする。このとき、母平均 μ の 99%信頼区間を求めよ。

母分散が未知なので、t をつかって変換すると、

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

が自由度 $10 - 1 = 9$ の t 分布に従う。よって、$P(-t_\alpha \leqq T \leqq t_\alpha) = 0.99 = 1 - 0.01$ となる t_α を両側 $100\alpha\%$ の t 分布の表から求めると、自由度 9 で確率 $1 - 0.99 = 0.01$ の部分の値を調べればよいので、$t_\alpha = 3.2498$ であることがわかる。よって母平均の 99%信頼区間は

$$-3.2498 \leqq \frac{17.44 - \mu}{\frac{2.95}{\sqrt{10}}} \leqq 3.2498$$

を満たせばよいので、この不等式を μ について解くと、

$$17.44 - 3.2498 \times \frac{2.95}{\sqrt{10}} \leqq \mu \leqq 17.44 + 3.2498 \times \frac{2.95}{\sqrt{10}}$$

$$14.4083 \cdots \leqq \mu \leqq 20.4716 \cdots$$

である。つまり、母平均の 99%信頼区間を有効数字 3 桁で答えると $[14.4, 20.5]$ となる。

自由度 m の t 分布の両側 $100\alpha\%$ 点

m	0.10	0.05	0.02	0.01
1	6.3137	12.706	31.821	63.656
2	2.9200	4.3027	6.9645	9.9250
3	2.3534	3.1824	4.5407	5.8408
4	2.1318	2.7765	3.7469	4.6041
5	2.0150	2.5706	3.3649	4.0321
6	1.9432	2.4469	3.1427	3.7074
7	1.8946	2.3646	2.9979	3.4995
8	1.8595	2.3060	2.8965	3.3554
9	1.8331	2.2622	2.8214	3.2498
10	1.8125	2.2281	2.7638	3.1693

・平均の区間推定（二標本の差の区間推定）

いままでは、一つの母集団に対して標本平均や標本分散の推定を行ってきたが、男女の平均の差などのように、二つの母集団の平均の差を調べたいこともある。ここでは、二つの母集団がそれぞれ母分散が等しい正規分布 $N(\mu_1, \sigma^2), N(\mu_2, \sigma^2)$ に従うとき、母平均の差 $\mu_1 - \mu_2$ の区間推定を行う。

まず、２つの母集団からそれぞれ

$$(x_1, x_2, \cdots, x_n)\ ,\ (y_1, y_2,\ \cdots\ , y_m)$$

と１つ目の母集団から n 個、２つ目の母集団から m 個の標本が無作為に得られたとする。このとき母平均の差 $\mu_1 - \mu_2$ の点推定は

$$\bar{x} - \bar{y} = \frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{m}\sum_{j=1}^{m} y_j$$

のように、それぞれの標本平均のを計算した差で求めることができる。次にこの点推定の値を使って区間推定を求める。正規分布の和や差が正規分布に従うことから、$\bar{x} - \bar{y}$ も正規分布に従うことになる。実際に $\bar{x} - \bar{y}$ は母平均 $\mu_1 - \mu_2$、母分散 $\frac{\sigma^2}{n} + \frac{\sigma^2}{m}$ の正規分布 $N(\mu_1 - \mu_2, \frac{\sigma^2}{n} + \frac{\sigma^2}{m})$ に従うことが理論的にわかっている。よって、母分散が既知の場合と未知の場合で次のように区間推定を行うことができる。

・母分散が既知のとき

$\bar{x} - \bar{y}$ が正規分布 $N(\mu_1 - \mu_2, \frac{\sigma^2}{n} + \frac{\sigma^2}{m})$ に従うので、

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma^2}{n} + \frac{\sigma^2}{m}}} = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

と変換すれば Z が標準正規分布に従うので、前回同様に次のような z_α を決め、

$$\Pr\{-z_\alpha \leqq Z \leqq z_\alpha\} = 1 - \alpha$$

括弧の中の不等式に実際の値 $\bar{x}, \bar{y}$ を代入し、$\mu_1 - \mu_2$ について解くことで

$$-z_\alpha \leqq \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}} \leqq z_\alpha$$

$$-z_\alpha \times \sigma\sqrt{\frac{1}{n} + \frac{1}{m}} \leqq (\bar{x} - \bar{y}) - (\mu_x - \mu_y) \leqq z_\alpha \times \sigma\sqrt{\frac{1}{n} + \frac{1}{m}}$$

$$(\bar{x} - \bar{y}) - z_\alpha \times \sigma\sqrt{\frac{1}{n} + \frac{1}{m}} \leqq \mu_x - \mu_y \leqq (\bar{x} - \bar{y}) + z_\alpha \times \sigma\sqrt{\frac{1}{n} + \frac{1}{m}}$$

のような区間推定を得る。

資料置場

https://www.gen.kanagawa-it.ac.jp/takeda/class/

中間試験は１１月２２日（金）に行います。

2019年度 確率統計S　小テスト解答用紙　2019.11.15

・母分散が未知の場合の区間推定

母集団が母分散未知の正規分布に従うとき、７つのデータ

52.3, 55.9, 65.3, 53.9, 61.2, 50.1, 59.6

が与えられたときの母平均 μ の 95%信頼区間を次の通り求めよ。

(1) 関数電卓を使い標本平均 $\bar{x}$, 標本分散 s^2, 標本標準偏差 s を求めよ。

(2) t 分布の表を用いて母分散が未知の場合の 95%信頼区間を求めよ。

<table>
<tr><td rowspan="2">2019 年度神奈川工科大学
確率統計 S
演習問題</td><td>学科</td><td>学年</td><td>組</td><td>学　籍　番　号</td><td>氏　　　名</td><td rowspan="2"></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td></tr>
</table>

提出先：K3-3309 号室前　18 番のボックス　提出期限：１１月１８日（月）１７時頃まで