

「統計じかけのオレンジ」

— A Statistics-work Orange —

第5回 推定 (3)

一般社団法人 日本下水道施設業協会
技術部長 **堅田 智洋**



4.2.3 正規母集団における母平均 μ がわかっているときの母分散 σ^2 の推定

次に、「正規母集団における母平均 μ がわかっているときの母分散 σ^2 の推定」の手順をみていく。

前項の「正規母集団における母分散 σ^2 がわかっているときの母平均 μ の推定」では、「母平均 μ (未知)、母分散 σ^2 (既知) の正規母集団で観測されるデータ x の n 個に対する標本平均 \bar{x} が、母集団と同じ統計値 (μ や σ^2) をもつ『平均値 μ 、分散 σ^2/n の正規分布』に従う」という定理から、標本平均 \bar{x} がある信頼度で出現する区間を母平均 μ (未知)、母分散 σ^2 (既知) との関係式で表すことにより平均値 μ を推定する、というものであった。

今回の「正規母集団における母平均 μ がわかっているときの母分散 σ^2 の推定」においては、「正規母集団で、観測される n 個のデータ x_1, x_2, \dots, x_n と母平均 μ (既知)、分散 σ^2 (未知) から

$$V = \left(\frac{x_1 - \mu}{\sigma}\right)^2 + \left(\frac{x_2 - \mu}{\sigma}\right)^2 + \dots + \left(\frac{x_n - \mu}{\sigma}\right)^2$$

… (式 4-16)

という統計量を作ると、 V は自由度 n のカイ二乗分布に従う。」という定理が基本となる。

これより、カイ二乗分布する統計量 V がある信頼度で出現する区間を設定し、正規母集団の母平均 μ (既知)、母分散 σ^2 (未知)、観測された標本データ x (既知) との関係式で表すことにより母分散 σ^2 を推定するというものである。

まず、カイ二乗分布について説明する。今、標準正規分布 (平均 = 0、標準偏差 = 1) をとる母集団 (標準正規母集団) から 3 個のデータ x_1, x_2, x_3 を観測し、そのデータを 2 乗して足し合わせた

統計量 V を設定する (式 4-17)。

$$V = x_1^2 + x_2^2 + x_3^2 \quad \dots \text{(式 4-17)}$$

例えば、標準正規母集団から +1、+3、-2 の 3 個のデータが観測されたら、

$$V = (+1)^2 + (+3)^2 + (-2)^2 = 14 \quad \dots \text{(式 4-18)}$$

となる。 x_1, x_2, x_3 は標本であるから、観測のたびにまちまちな値を取る。結果、 V もまちまちな値を取るため、 V は統計量の一種となる。その分布をグラフ化したものが図 4-7 である。

この分布は V が 0 以上の数値に限られ、比較的 0 に近いところに多くのデータが密集し、左から右に向かって急激に落ちていく、ジェットコースターのような形になっている。標準正規分布に従う x はその 70% 弱が標準偏差 ± 1 個分の範囲 (すなわち $-1 \leq x \leq 1$) に分布する (No.82、図 3-3 参照)。そのため、そのデータを 3 つ取り出し、それぞれ 2 乗してから足し合わせた V は図 4-7 のような分布となるのである。この分布を「自由度 3 のカイ二乗分布」という。「自由度」という

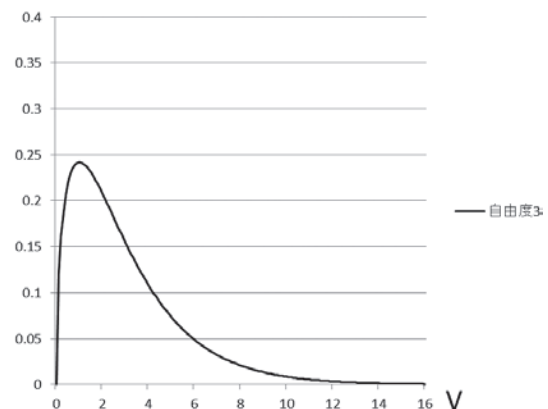


図 4-7 自由度 3 のカイ二乗分布

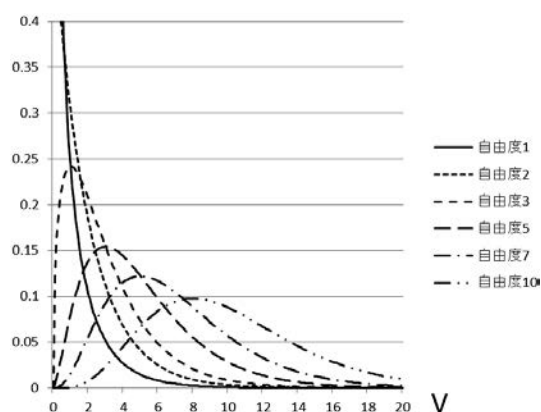


図4-8 自由度nのカイ二乗分布

のは専門的な用語だが、「観測データの数」を意味する。同じように、標準正規母集団から3個ではなくn個のデータを観測し2乗して加えて作った統計量Vは「自由度nのカイ二乗分布」となる。これらの分布（グラフの形）は自由度nによって変わる。それを比較したのが図4-8である。グラフの形は変わるが、それを表す関数は同種なので統一的にカイ二乗分布と呼ばれている。

カイ二乗分布の特徴を以下に挙げる。

- ・0近辺のデータの相対度数が大きく、グラフの形がジェットコースターのような形となる。これは標準正規分布がマイナスの数も含め0近辺の数値の相対度数が大きいためである。

- ・自由度n（観測データ数）が大きくなるにしたがって、山の高さが低くなりながらだんだん右のほうに移動していく。つまりジェットコースターの傾斜が緩やかになる。これは、nが大きくなると0から少し離れたデータが出る相対度数が高まっていくことを表している。

<次号へ続く>

【本稿の全体構成】

1. はじめに
2. さまざまな統計量

2.1 統計量とは

2.2 平均値

2.3 分散、標準偏差

2.3.1 データのバラツキの重要性

2.3.2 分散、標準偏差の算出方法

2.3.3 標準偏差の意味

3. 正規分布

3.1 正規分布の特徴

3.2 標準正規分布

4. 推定

4.1 統計的推定とは

4.2 統計的推定のパターン別アプローチ

4.2.1 統計的推定のファーストアプローチ

4.2.2 正規母集団における母分散 σ^2 がわかっているときの母平均 μ の推定

4.2.3 正規母集団における母平均 μ がわかっているときの母分散 σ^2 の推定

4.2.4 正規母集団における母平均 μ がわからないときの母分散 σ^2 の推定

4.2.5 正規母集団における母分散 σ^2 がわからないときの母平均 μ の推定

5. 検定

5.1 検定とは

5.2 検定のパターン別アプローチ

5.2.1 母平均の検定

5.2.2 t検定

5.2.3 カイ二乗検定

【参考文献】

- 1) 完全独習 統計学入門 小島 寛之 2006年9月ダイヤモンド社
- 2) まずはこの一冊から 意味がわかる統計解析 涌井 貞美 2013年2月 ベレ出版