シリーズ・技術調査報告

「統計じかけのオレンジ」
―A Statistics-work Orange―
第4回　推定（2）

一般社団法人 日本下水道施設業協会
技術部長　**堅　田　智　洋**

4.2.2　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

次に、本題の「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」を複数の標本数の場合で解説する。

【例題２】[2)]

お金が入っている箱がたくさん並んでいる。各々の箱の額は不明で、その平均金額を知るために、標本として９箱を無作為に抽出して調べた。その結果は次のとおりである（単位：円）。530、515、470、545、440、530、455、560、455。この標本から、箱の中の平均金額を信頼度95％で推定せよ。ただし、箱の中の金額Xは分散30^2（標準偏差30）の正規分布に従うとする。

先ほどの【例題１】では１だった標本の大きさ（標本数）がこの例題では９になっているが、基本的なアプローチは同じである。ただ、標本数が複数になったことで、ここで、正規母集団からの標本平均が有する以下の性質を利用する。

【正規母集団の標本平均の定理】

平均値μ、分散σ^2の正規分布に従う独立したn個の確率変数X_1、X_2、…、X_nについて、次のように$\overline{X}$を定義する。

$$\overline{X}=\frac{X_1+X_2+\cdots+X_n}{n} \qquad \cdots（式４－８）$$

この確率変数$\overline{X}$（標本平均）は平均μ、分散σ^2／nの正規分布に従う。

Xと$\overline{X}$双方の正規分布を図４－４に示す。$\overline{X}$の分布は、標本数nが大きくなるほど標本の散らばり具合を示す分散σ^2／nの値が小さくなるため、Xの分布よりもピークが鋭くなっていく。

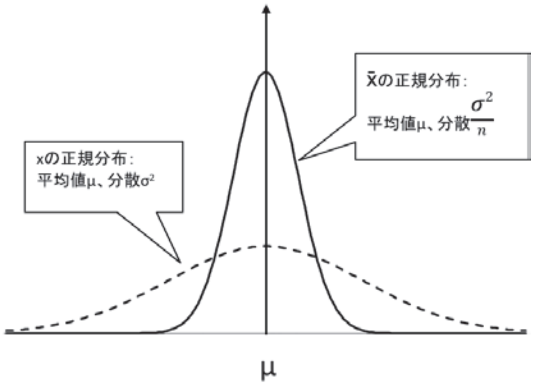

図４－４　Xと$\overline{X}$の正規分布

箱の中の金額Xは「分散30^2の正規分布」に従うので、上記の定理より、

$$\overline{X}=\frac{X_1+X_2+\cdots+X_n}{9} \qquad \cdots（式４－９）$$

で表される標本平均$\overline{X}$は、「平均値μ（未知）、分散30^2／9＝100の正規分布」に従う。そのため、$\overline{X}$の分布においても正規分布の性質が利用できる。今、信頼度は95％と設定されているので、図５－４の$\overline{X}$のグラフにおいて平均値μを中心にして95％の確率の範囲に網を掛けると、その右端（両側５％点）は$\mu+1.96\times\sigma$と与えられる（図４－５）。

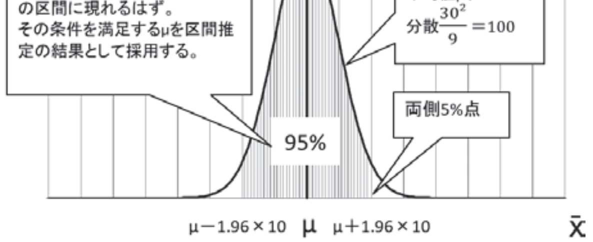

図４－５　$\overline{X}$の正規分布における両側５％点

分散＝100から標準偏差σは10である。したがって、正規分布の対称性から、標本平均$\overline{X}$が95％の確率で生起する範囲は次のように表される。

$\mu-1.96\times10\leqq\overline{X}\leqq\mu+1.96\times10$…（式４－10）

標本平均すなわち無作為に抽出した箱の中の金額の平均値$\overline{X}$は95％の確率でこの区間に現れることになる。

$$標本平均\overline{X}=\frac{530+515+\cdots+455}{9}=500（円）$$
…（式４－11）

を代入して計算すると、

$\mu-1.96\times10\leqq500\leqq\mu+1.96\times10$…（式４－12）

から

$480.4\leqq\mu\leqq519.6$　…（式４－13）

が得られる。これが信頼度95％の平均金額（すなわち母平均）の信頼区間である。先程、標本数１の標本から同一条件で導き出した信頼区間は

$441.2\leqq\mu\leqq558.8$　…（式４－７）（再）

であった。（４）の信頼区間はこれよりも狭まっている。標本の大きさが１から９になり、情報量が多くなったため、推定の精度が高まったことがわかる。

以上の手順を一般化する。

【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】

母分散σ^2の正規母集団から抽出された大きさnの標本の標本平均を$\overline{X}$とするとき、母平均μの信頼度95％の信頼区間は次のようになる。

$$\overline{X}-1.96\times\frac{\sigma}{\sqrt{n}}\leqq\mu\leqq\overline{X}+1.96\times\frac{\sigma}{\sqrt{n}}$$
…（式４－14）

$\frac{\sigma}{\sqrt{n}}$に係っている1.96は信頼度95％の場合に定まる係数であるので、信頼度をαとさらに一般化してみる（ここで、$0<\alpha<1$）。

母分散σ^2の正規母集団から抽出された大きさnの標本の標本平均を$\overline{X}$とするとき、母平均μの信頼度αの信頼区間は次のようになる。

$$\overline{X}-k\frac{\sigma}{\sqrt{n}}\leqq\mu\leqq\overline{X}+k\times\frac{\sigma}{\sqrt{n}}$$
…（式４－15）

ここで、kは標準正規分布の両側100（$1-\alpha$）％点である（例：信頼度95％の場合は両側５％点（＝1.96）、信頼度99％の場合は両側１％点（＝2.58））。

＜次号へ続く＞

【本稿の全体構成】

【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