

2019年度春学期 統計学 第12回

分布の平均を推測する — 区間推定

区間推定

ようやく、具体的な推測へ

今日の講義では、母集団分布が正規分布であると仮定できるとき、母平均を標本から推測する方法を考えます。

第10回の講義で、標本平均は母平均の推測結果としてもそうおかしくはない、ただし間違った推測をしてしまう可能性はある、という説明をしました。確かに、母集団全体のデータではなく、ほんのいくつかのデータからなる標本だけで母平均を推測するのですから、母平均の値については不確かなことしか言えません。

そこで、第11回の講義では、間違った推測をしてしまう可能性、つまり推測の不確かさを確率の形で計算するために、確率分布を式であらわす「確率分布モデル」の考えを説明しました。

区間推定の考え方

サイズ1の標本を何度も取り出すと

図1(a)は、標本としてひとつの数値を取り出した時の値を示しています。「標本抽出○回目」の横棒の上に★印がありますが、この横棒は、物差しのように棒の上の位置で数を表す「数直線」だと思ってください。標本は無作為抽出されていますから、標本として取り出される数値は、標本抽出1回目、2回目、... で毎回異なった値になります。

一方、母平均は、標本としてどんな値が取り出されるかには関係なく、すでに決まった値ですから、標本抽出1回目、2回目、... に関係なく、いつも同じ値です。この値を、標本抽出1回目、2回目、... の数直線を縦に貫く線で表しています。

サイズ n の標本を何度も取り出すと

図1(b)は、標本としていくつかの数値のセットを取り出して、仮に何度も標本平均を計算したとするときの、標本平均のばらつきを表したものです。第10回の講義で説明したように、標本平均の期待値は母平均と同じですから、図のように、標本平均は母平均のまわりにばらついていることになります。

標本サイズが大きくなると、(a)の場合に比べてばらつきは小さくなりますが、標本平均が母平均そのものではないことには変わりはありません。

標本平均のまわりに「幅」をもたせる

そこで、図1(c)のように、標本平均のまわりに「幅」をもたせて、例えば「母平均は、標本平均プラスマイナスいくらの範囲に入っているだろう」というように推測することを考えます。ある程度「幅」があれば、何度も標本抽出したとき、母平均は「たいてい」この幅に入っているようにすることができます。しかも、標本平均を用いた(b)の場合、ひとつの数値を用いた(a)の場合に比べてばらつきが小さくなっているので、この「幅」は狭くできます。

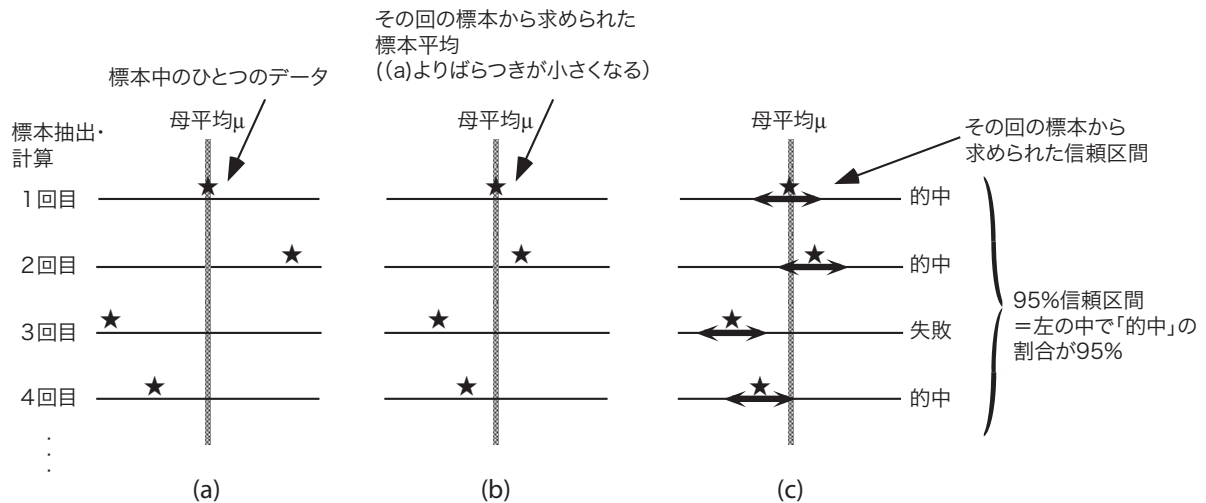


図 1: 区間推定の考え方

信頼区間と信頼係数

この標本平均のまわりにつけた「幅」を、統計学では母平均の**信頼区間**といい、信頼区間が母平均をその中に含んでいる確率を**信頼係数**といいます。信頼係数が95%であるような信頼区間を設定するということは、標本抽出1回目、2回目、...で、それぞれ標本平均のまわりに信頼区間を設定したとき、そのうち95%の信頼区間では母平均が入っていて、5%の信頼区間では入っていないように、信頼区間の幅を設定することです。このときの推測結果は、「標本平均プラスマイナスいくらの信頼区間に、母平均が入っているだろう」という形になります。この「プラスマイナスいくら」という幅が、先に述べた「ほぼ」に相当し、信頼係数95%が「たいてい」の度合いを表しています。

信頼係数の意味

注意すべきなのは、図12.2では母平均を縦の線で表していますが、実際には母平均がいくらなのかはわかっていない、ということです。図では標本抽出3回目で、信頼区間に母平均が入ってなくて「失敗」としてしています。しかし実際には、標本抽出1回目、2回目、...のどの信頼区間が、母平均を含んでいる「的中」で、どの信頼区間が、母平均を含んでいない「失敗」かは、わかりません。ですから、いま1回だけ標本抽出をして、それをもとに標本平均と信頼区間を計算したとしても、その信頼区間が「的中」か「失敗」かはわかりません。ただ、同じやり方を何度も信頼区間を求めれば、確率95%で「的中」する、と言っているだけです。

この考えにもとづき、確率分布モデルを使って、標本平均のまわりにどのくらいの信頼区間をもたせれば、その信頼区間の中に母平均が入っている確率が、ある設定した信頼係数（例えば95%）になるかを計算します。

台風情報と区間推定

台風情報では、区間推定のひとつの例を見ることができます。テレビの画面に出ている予想進路図にある「予報円」は、区間推定によって描かれています。台風情報の「〇〇時に円内の範囲に達すると思われれます」という予報は、「〇〇時に円内の範囲に達する確率が70%である」ことを示しています。

正規分布の場合の、母平均の区間推定

では、母集団分布が正規分布モデルで表されると仮定されるとき、母平均の区間推定の方法を説明します。次の問題を考えてみましょう。

区間推定の例題

ある試験の点数の分布は正規分布であるとし、この試験の受験者から、10人からなる標本を無作為抽出して、この人たちの点数を平均したところ50点でした。この試験の受験者全体の標準偏差が5点であるとわかっているとき、受験者全体の平均点の95%信頼区間を求めてください。

母集団の平均がわからないのに、母集団の標準偏差がわかっているというのはヘンな話ですが、これは説明のために用意した例です。正規分布が仮定でき、母集団の標準偏差が不明な場合については、次の「*t*分布」の項で説明します。

考え方

いまから推定する母平均を μ とし、母分散（こちらはすでにわかっているものとされています）を σ^2 とします。そうすると、母集団分布は平均 μ 、分散 σ^2 の正規分布、すなわち $N(\mu, \sigma^2)$ となります。このとき、標本は無作為抽出されていますから、標本は確率変数で、母集団分布と同じ確率分布にしたがいます。すなわち、 n 人からなる標本のそれぞれの確率分布もまた $N(\mu, \sigma^2)$ です。

このとき、標本平均 \bar{X} は、 n 人からなる標本を X_1, \dots, X_n で表すと、 $(X_1 + \dots + X_n)/n$ で表されます。

「正規分布の性質2」を使う

ここで、第11回の講義で述べた「正規分布の性質2（正規分布の再生性）」、すなわち

X_1, \dots, X_n が独立で、いずれも正規分布 $N(\mu, \sigma^2)$ にしたがうならば、それらの平均 $(X_1 + \dots + X_n)/n$ は $N(\mu, \sigma^2/n)$ にしたがう

を用いると、 n 人からなる標本はこの性質の X_1, \dots, X_n にあてはまっていますから、「標本平均 \bar{X} は $N(\mu, \sigma^2/n)$ にしたがう」ことがわかります。

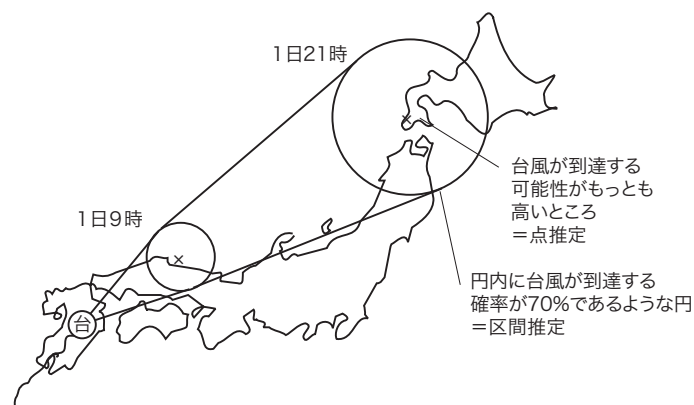


図 2: 台風情報に見る点推定と区間推定

「正規分布の性質 1」を使う

さらに、

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \quad (1)$$

という値を計算すると、第 1 1 回の講義で述べた「正規分布の性質 1」から、 Z は標準正規分布 $N(0,1)$ にしたがうことがわかります。

「 Z が入っている確率が 95%である区間」

そこで、「 Z が入っている確率が 95%である区間」はどういうものか考えてみましょう。第 5 回の講義の「連続型確率分布」のところで説明したように、 Z がある区間に入る確率は、標準正規分布の確率密度関数のグラフの下の、その区間に対応する部分の面積になります。この部分の面積が全体の 95%になるように、左右対称に Z の区間をとることにし、図 3(a)のように表します。このときの Z の区間の両端を $-u$ と u とすると、 Z がこの区間に入る確率すなわち $P(-u \leq Z \leq u) = 0.95$ となります。このとき、図 3(b)のように、 $P(Z \geq u) = 0.025$ となります。 $P(Z \geq u) = 0.025$ となる u は、正規分布の数表から求めることができます。数表によると、 $u = 1.96$ のとき、 $P(Z \geq 1.96) = 0.024998 \approx 0.025$ であることがわかります。すなわち、 $P(-1.96 \leq Z \leq 1.96) = 0.95$ ということがわかります。

信頼区間の式を導く

ところで、(1) 式の関係を $P(-1.96 \leq Z \leq 1.96) = 0.95$ に用いると、

$$P(-1.96 \leq \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \leq 1.96) = 0.95 \quad (2)$$

という関係があることがわかります。ここで、今知りたいのは母集団の平均 μ の範囲ですから、(2) 式を μ の範囲に書き換えると

$$P(\bar{X} - 1.96\sqrt{\sigma^2/n} \leq \mu \leq \bar{X} + 1.96\sqrt{\sigma^2/n}) = 0.95 \quad (3)$$

という関係が得られます。この範囲が、 μ の 95%信頼区間となります。

問題文の数値を入れて計算

この問題では、標本平均 $\bar{X} = 50$ 、母集団の分散 $\sigma^2 = 25$ ですから、これらの数値を (3) 式に入れると、求める 95%信頼区間は「46.9 以上 53.1 以下」となります。「46.9 以上 53.1 以下」という区間を、数学では $[46.9, 53.1]$ と書きます。

「信頼係数 95%」の意味

$P(46.9 \leq \mu \leq 53.1)$ とは書かない

前節で、母平均 μ の区間推定の結果を「求める 95%信頼区間は『46.9 以上 53.1 以下』あるいは $[46.9, 53.1]$ と書きましたが、これを $P(46.9 \leq \mu \leq 53.1)$ とは書きません。

$P(\cdot)$ は、「 (\cdot) の中のことが起きる確率」という意味ですから、 (\cdot) の中にはランダムに決まる数、すなわち確率変数が入っていなければなりません。母平均 μ は、標本を調べている人が知らないだけで、実際には調べる前から 1 つの値に決まっていますから、確率変数ではありません。ですから、 $P(\bar{X} - 1.96\sqrt{\sigma^2/n} \leq \mu \leq \bar{X} + 1.96\sqrt{\sigma^2/n})$ という式では、ランダムなのは μ ではなく \bar{X} であり、不等式の上限と下限がランダムに決まることを示しています。

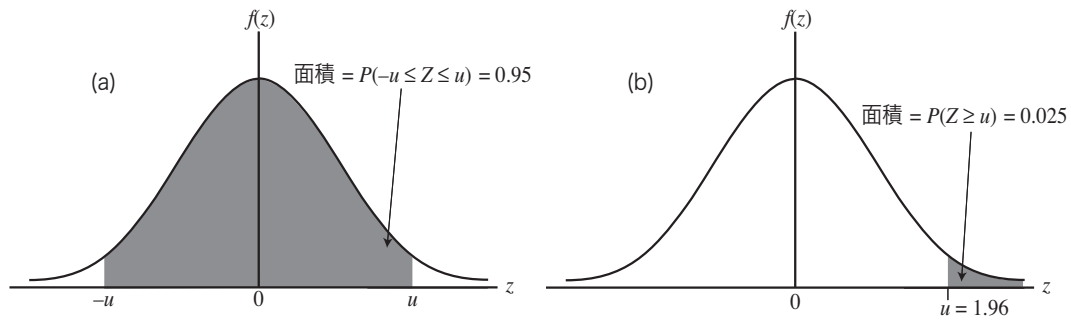


図 3: 95%信頼区間の求め方

ところが、具体的な数値を計算して、 $P(46.9 \leq \mu \leq 53.1) = 0.95$ という式にしてしまうと、この式には確率変数がありません。したがって、この式は間違いです¹。

信頼区間の数値が表す意味は

具体的な数値で表された $[46.9, 53.1]$ という信頼区間は、「いま無作為抽出された標本によって、偶然決まった標本平均 \bar{X} の値を用いて、偶然そうなった値」です。母平均 μ は、 $[46.9, 53.1]$ という区間に「入っているか、入っていないかどちらかに決まっている」のであって、95%の確率で入っているわけではありません。

「母平均が入っている確率が95%であるような区間」とは、今日の冒頭の図1(b)で示したように、「標本を取り出して計算し信頼区間を求める」という操作を何回も行うと、

**100回あたり95回は、求めた信頼区間の中に確かに母平均が入っているが
残り5回は、求めた信頼区間の中に母平均は入っていない**

となるような計算、つまり「95%の確率で当たるような、推測のやりかた」を意味しているのです。

「信頼」とは

現実には、標本を取り出して計算するのは1回だけです。ですから、その時にたまたま取り出された標本から計算された信頼区間、例えば $[46.9, 53.1]$ には、母平均は入っていないかもしれません。

「1回の、信頼係数95%の推測を信じる」ことは、ある人の言っていることについて「この人が今回言っていることは本当かどうか分からないが、この人は95%の確率で本当のことを言うらしいから、今回も信じることにしよう」というのと同じです。

区間推定に関する注意

[1] ここまでの区間推定の説明では、95%信頼区間を求めました。信頼係数としては95%が一番よく使われますが、**信頼係数として95%という値を選ぶ根拠は何もありません**。「95%の確率で当たっている推測」とは、「5%の確率ではずれている推測」でもありますから、信頼係数を95%とすることは、「5%くら

¹正確には、本文で述べている「この式は間違いである」という考え方（ネイマン=ピアソンの解釈）と、「この式は“ $46.9 \leq \mu \leq 53.1$ ”という命題の信頼性が95%であることを表す」という考え方（フィッシャーの解釈）があり、大きな論争を引き起こしました。本文のように前者の立場に立つ統計学は「頻度論統計学」とよばれ、現在の統計学の基本となっています。

いの確率なら、推測がはずれて失敗しても、まあいいか」と考えていることになります。また、信頼係数を例えば99%（この値も95%の次によく用いられます）にすると、図3から明らかなように、95%の場合よりも信頼区間の幅は広がります。信頼区間の幅が広い、とは、推測のあいまいさが大きい、ということですから、場合によっては意味のある推定ができなくなってしまうこともあります。

[2] 区間推定においては、**母集団の大きさは信頼区間の幅には影響しない**ことに注意してください。今回の例題でも、標本サイズが10人という条件が同じであれば、この試験の受験者全体の人数が1000人でも10万人でも、信頼区間の幅は同じです。つまり、「信頼区間の幅は、標本の**サイズそのもの**で決まり、標本サイズの母集団の大きさに対する**割合には無関係**」ということです。

「10人からなる標本」は、「1000人のうちの10人」であっても「10万人のうちの10人」であってもその価値は同じ、というのは一見不思議ですが、これは、「母集団のどの人も同じチャンスで選ばれ、しかも、ある人が選ばれるかどうかは、他の人が選ばれるかどうかには影響をうけない」という理想的な無作為抽出が、第6回の講義で説明した「復元抽出」であることに理由があります。

復元抽出の場合、「ある値のデータが標本として取り出される確率＝その値のデータが母集団中で占める**割合**」という、ここまでの講義で説明した原理が、抽出の順序によらずなりたちます。「割合」は、母集団の大きさには無関係です。したがって、その標本から計算される区間推定の結果も、母集団の大きさには無関係です。

一方、非復元抽出の場合は、標本を抽出するたびに母集団全体の人数が減ってゆきますから、「ある値のデータが標本として取り出される確率＝その値のデータが母集団中で占める割合」が、抽出の途中でだんだん変化してゆきます。この変化のしかたは、母集団のサイズに影響されます。したがって、区間推定の結果も、母集団の大きさに影響されます。この違いは、母集団が大きければさして問題になりませんが、そうでなければ、非復元抽出においては計算で補正をする必要があります。

なお、復元抽出をしている場合は、この理屈でいえば、母集団の大きさが「無限」であっても、本質的違いはありません。これは、「600本中100本の当たりを含むくじを作り、そこから1本ひく」というくじびきと、「さいころを1回ふって、1の目が出たら当たり」というくじびきが、理想的なくじびきであればまったく同じ意味である、というのと同じです。前者は、母集団に含まれるくじの本数は600本です。後者は、さいころはいくらでもふることができますから、母集団の大きさが無限の場合に相当します。

今日の演習

ある製品の長さを10回測定したとき、10個の測定値の平均は10.0(cm)でした。この測定では測定値の標準偏差が0.1(cm)であることがわかっており、測定値が真の長さを期待値とする正規分布にしたがうとすると、真の長さの95%信頼区間を求めてください。また、10.0(cm)というのが20個の測定値の平均であるときは、95%信頼区間はどうなりますか。