

§ 6.1 点推定

標本調査では母集団から抽出した一部の標本の属性値だけを調べるので、母平均や母分散などの母数は推測するしかない。それ故、標本調査をするとき、得られた標本のデータから母数をどのように推測するかが問題になる。

大きさ n の標本に属す 1 番めの事物の属性値 x_1 、2 番めの事物の属性値 x_2 、3 番めの事物の属性値 x_3 、 $\dots$ 、 n 番めの事物の属性値 x_n に対して、それらの関数 $f(x_1, x_2, x_3, \dots, x_n)$ の値が母数 θ に近いであろうと推測するとする；このとき、標本確率変数 $X_1, X_2, X_3, \dots, X_n$ の関数 $f(X_1, X_2, X_3, \dots, X_n)$ が標本統計量であるならば、つまり標本調査によって標本確率変数 $X_1, X_2, X_3, \dots, X_n$ の実現値が分かれば $f(X_1, X_2, X_3, \dots, X_n)$ の実現値が計算できるならば、この標本統計量 $f(X_1, X_2, X_3, \dots, X_n)$ を θ の**推定量** (estimator) といい、推定量の実現値を**推定値** (estimate) という。

例 大きさ 4 の標本に属す 1 番めの事物の属性値 x_1 、2 番めの事物の属性値 x_2 、3 番めの事物の属性値 x_3 、4 番めの事物の属性値 x_4 が分かるとき、通常は母平均は $\frac{x_1 + x_2 + x_3 + x_4}{4}$ に近いであろうと見当を付けることが多い。このとき、標本確率変数 X_1, X_2, X_3, X_4 の関数 $\bar{X} = \frac{X_1 + X_2 + X_3 + X_4}{4}$ は統計量であり、母平均の推定量である。この推定量の実現値が推定値である。

推定量を定めて未知の母数の推定値を求めることを**点推定** (point estimation) という。点推定では推定量をどのように定めればよいかを問題にする。どのような標本統計量が推定量であるかについてははっきりとした規準がある訳ではない。しかし、未知の母数を推定量の実現値で推定するので、推定量として望ましい性質がいくつかある。ここでは、推定量が持つべき性質として、不偏性、有効性、一致性、の 3 つを説明する。

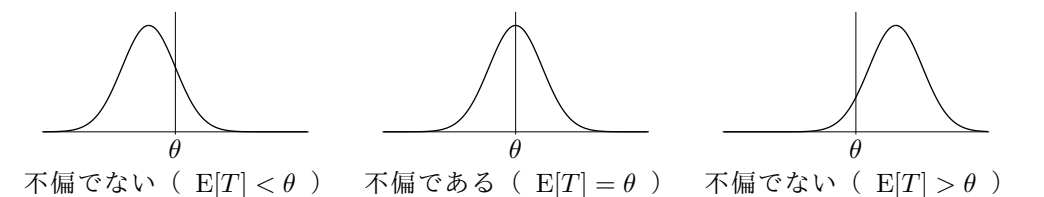
まず不偏性について述べる。

定義 標本統計量 T が母数 θ の**不偏推定量** (unbiased estimator) であるとは、

$$E[T] = \theta$$

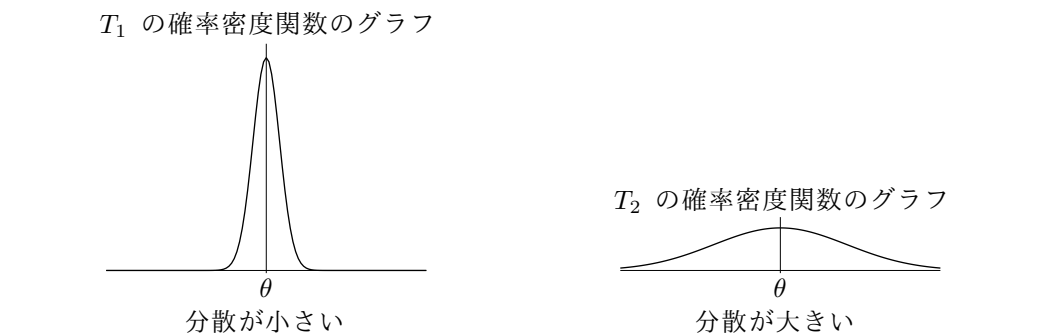
となることである。

統計量 T が母数 θ の不偏推定量であるとは次のことを意味する：推定量 T による母数 θ の推定値は、 θ より大きかったり小さかったりするが、平均すると θ になる。 T の確率密度関数のグラフで図解すると例えば次のようになる。



次に有効性について述べる。

母数 θ の不偏推定量 T_1 と T_2 について、確率密度関数のグラフが例えば下図のようになるとする（各々の座標系の座標軸の尺度は同じであるとする）。



このようなとき、 T_1 の実現値は θ に近い値になる確率が大きいのにに対して、 T_2 の実現値は θ から離れた値になる確率が小さくない。なので θ の値を推定するためには T_2 より T_1 の方が望ましい。

母数 θ の不偏推定量 T_1 と T_2 について、 $V[T_1] < V[T_2]$ のとき、 T_2 より T_1 の方が母数 θ に近い実現値になりやすい、つまり、 T_1 の方が T_2 より母数 θ に近い推定値になることが多い。

定義 母数 θ の不偏推定量 T_1 と T_2 について、 T_1 が T_2 より**有効** (efficient) であるとは

$$V[T_1] < V[T_2]$$

となることである。

最後に一致性について述べる。大きさ n の標本の標本確率変数 $X_1, X_2, X_3, \dots, X_n$ の関数 $T_n = T_n(X_1, X_2, X_3, \dots, X_n)$ が母数 θ の**一致推定量** (consistent estimator) であるとは、標本の大きさ n を大きくしていくと統計量 T_n の実現値が母数 θ の値に近い値になる確率が大きくなることである（正確な定義は省く）；つまり標本の大きさ n が大きいほど統計量 T_n の実現値が母数 θ の推定値として信頼性が高いということである。