

# 不偏分散は、どうして $n - 1$ で割るの？ \*

浜田 宏

## 登場人物

あおば

青葉: 大学生。数学が苦手

かきょういん

花京院: 大学生。数学が好き。スタンドは使えない。

花京院は作業机の隅で、本を読んでいる。

広い机の上には、論文のコピーと計算用紙が散乱している。

青葉は授業課題のレポートを書くために、パソコンのモニタに向かって  
ている。

数理行動科学研究室の分析室には、青葉と花京院の 2 人だけがいた。

「ねえ、花京院くん。いま時間ある？」花京院の読書が一段落したのを見計らって、青葉は声をかけた。

「うん、まあ暇と言えは暇だし、暇でないと言えは暇じゃない」花京院は、壁の時計で時間を確認した。

「どっちなのよ」

「ごめん、暇だった。なに？」

「えっとね、統計学の授業で《不偏分散》って、習ったでしょ……。よくわかんないから教えてほしいなと思って」

---

\* ver1.0. 2022 年 6 月 20 日

「不偏分散？ なにそれ」

「ほら、あの  $n - 1$  で割るやつだよ」

「ああ、この推定量のことかな」

花京院は机のうえの計算用紙に式を書いた。

$$\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

「そうそう。これこれ」

「これ、不偏分散っていう名前なんだ」

「いま自分で式を書いたじゃん。名前覚えてないの？」と青葉が聞いた。

「うーん、式はなんとなく覚えてるけど、名前まで覚えてなかった」

「この式って、すごく変じゃない？」

「30 分」

「え？」

「ちゃんと説明するのに、30 分」

そう言うとき花京院はパソコンの前に座ると、キーボードをたたき始めた。

## 1 青葉の疑問

---

「君は不偏分散のどこが変だと思うの？」

「えーっとね、 $n$  個の項を足してるのに、 $n - 1$  で割るところ」

「そうだね。おかしいね」

「でしょ？ 花京院君だって、おかしいと思うでしょ？」

「そうだね。最初に見たときは、なんでだろうと思ったよ。でも理由を考えたなら納得したよ」

「理由あるんだ」

「ちゃんと理由があるんだよ。実験して確かめてみよう」

「実験？」

花京院は、キーボードでなにやら謎のコードを書くと、ヒストグラム

を一つ画面に描いた。

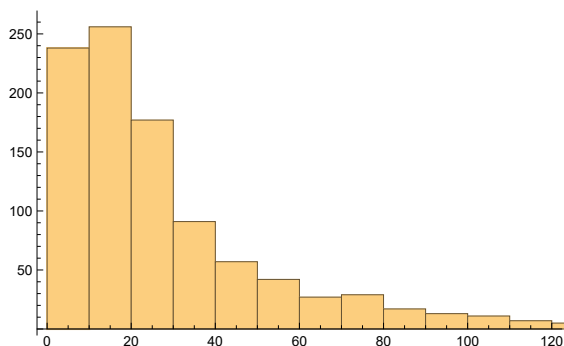


図1 全体の分布

「これをね、観察する対象の《全体》としよう。1000 個の数値をヒストグラムにまとめたものだよ」

「わかった。なんのデータなの？」

「いや、特に意味はないよ。テキトーにコンピュータに作らせたから」

「なにか意味がないと、イメージ湧かないよー」

「しょうがないな。じゃあ柿ピー」

「は？」

「大体 0~120 くらいで分布しているから、1 日に食べる柿ピーのグラム数ってことにしよう。柿ピーの小袋って 30g なんだよ」花京院は机のうえに置いてあったお菓子の袋をひとつ手に取った。

「わかった。じゃあこれは《1 日に食べる柿ピーのグラム数》のデータだね」

「そういうこと。この 1000 人は M 町に住む成人男性全員の数としよう。全員ってところが重要だよ。この《全体》のデータからまず平均値を計算してみよう。1000 人分のデータを

$$x_1, x_2, \dots, x_{1000}$$

で表すと、その平均は

$$\frac{1}{1000} \sum_{i=1}^{1000} x_i$$

で計算できる。実際に計算してみると 31.91 だから、1 人平均 32g ほど柿ピーを食べることが分かる。」

「意外と少ないんだね。私なら 2 袋はいけるよ」

「もちろん、そういう人もいるだろう。そういう食べる量の《バラつき》を表すのが分散だったね。《全体》データから分散を計算してみよう。さきほど計算した平均値を  $\mu$  とおけば分散は

$$\frac{1}{1000} \sum_{i=1}^{1000} (x_i - \mu)^2$$

だね。計算した結果は 1287.11 だよ」

「これは  $n$  で割ったね」

「そう。いま計算したのは《全体》の分散。これを《母集団》の分散ともいう。もし観察対象全てのデータが手元にあるのなら、分散を計算するときただ  $n$  で割ればいい」

「うん。それは分かる。じゃあいつ  $n - 1$  で割るの？」

「標本から《全体》のパラメータを推定するときだよ」

「ちょっとなに言ってるかわからない」

「実例で示そう」

## 2 標本をつかった推定

---

「いま《全体》から 20 個だけデータをランダムに抜き出したとしよう。たとえばこんな感じ」

65.06, 11.44, 13.08, ..., 145.86

「うん」

「省略して書いたけど、いまここに 20 個の数値がある。この 20 個を使って平均値を計算してみよう。全部足して 20 で割るだけだよ。結果は

$$(65.06 + 11.44 + 13.08 + \cdots + 145.86) / 20 = 28.46$$

だよ。ここまではいいかな？」

「うん、まあ平均値の計算くらい分かるよ」

「さらにこの標本平均値を使って、不偏分散を計算する\*1。つまり

全ての値から標本平均値 28.46 を引いて 2 乗した数を全部足し、  
最後に 19 で割る。

という操作をする。こんな感じだ。

$$\{(65.06 - 28.46)^2 + (11.44 - 28.46)^2 + \cdots + (145.86 - 28.46)^2\} / 19 \\ = 1183.32$$

比較のために  $n = 20$  で割った数も計算しておこう。

$$\{(65.06 - 28.46)^2 + (11.44 - 28.46)^2 + \cdots + (145.86 - 28.46)^2\} / 20 \\ = 1124.15$$

となる」

「うーん、計算方法は分かるよ。でもどうして  $n - 1$  で割ったり、 $n$  で割ったりするの？」

「 $n - 1$  で割った推定値と、 $n$  で割った推定値を比べてみよう。どっちが《全体》の分散に近い？」

「えーっと、さっき計算した《全体》の分散は、1287.11 だったね。ということは、

$n - 1$  で割った場合 : 1183.32

$n$  で割った場合 : 1124.15

だから、 $n - 1$  で割った場合のほうが、本当の値に近いよ」

「そのとおり。実験の結果、標本から《全体》の分散を推測するとき、 $n$  で割るより  $n - 1$  で割った《不偏分散》を使うほうが、本当の値に近い値になった」

「でも、いまのは偶然かもしれないじゃん」

---

\*1 不偏分散という確率変数（推定量）の実現値（推定値）を計算する、とも表現できます

青葉は怪しそうに計算結果を見比べた。

「いいねえ。その注意深さ。じゃあ計算を繰り返してみよう。手続きはこうだよ」

1. 《全体》から 20 個の標本を抜き出す
2. 標本から「 $n$  で割る推定値」と「 $n - 1$  で割る推定値」を計算する
3. この操作を 100 回繰り返した後でそれぞれの推定値の平均を求める
4. どちらが本当の値に近いかを比較する

「いいかな？」

「OK」

「それじゃあ計算するよ。100 回ほど標本抽出と計算を繰り返してみよう」花京院は、計算用のコードを書き足した。

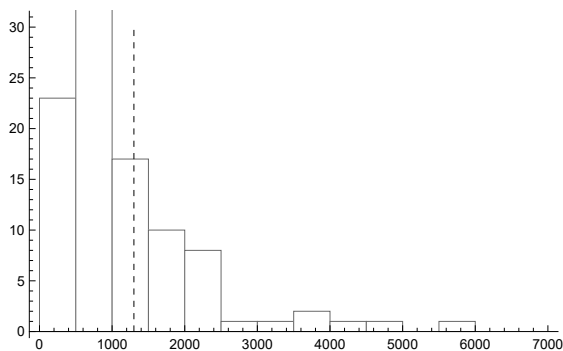


図 2  $n$  で割った場合の結果

「まずこれが  $n$  で割った場合。100 回分の計算結果のヒストグラムだ。点線の位置が本当の値だよ」

「うーん、まあ微妙だね」

「次いこう」

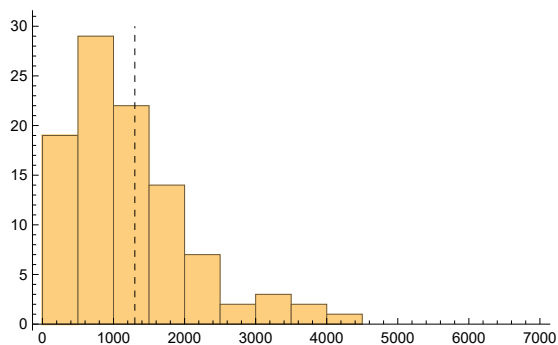


図3  $n-1$  で割った場合の結果

「これが  $n-1$  で割った場合」

「うーん、これも結構ズレてるね」

「二つ並べてみよう」

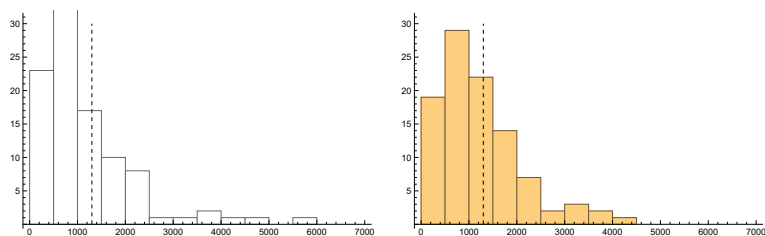


図4 計算結果の比較。左が  $n$ 、右が  $n-1$  で割った場合の結果

「うーん、微妙だけど、右の方が本当の値に近い数値を推測できてるのかなあ」

「ちょっと見た目では分からないね。それぞれの平均を計算して、本当の値と比較してみよう」

$n$  で割った推定値の平均 : 1184.31

$n-1$  で割った推定値の平均 : 1279.28

本当の値 : 1287.11

この結果を見ると、

標本から《全体》の分散を推測するときに、 $n$ で割るより  $n - 1$ で割った《不偏分散》を使うほうが、平均的に本当の値に近い値になる

ことが分かる」

「ちょっと待ってよ。まだなんかしっくりこないな」

青葉は計算結果を見比べながら、つぶやいた。

「どこが？」

「計算の結果がこうなることは分かったよ。でもどういえばいいのかな……。どうしてそうなるのか、まだ分からない」

「ふむ」

「《全体》の分散を計算する式は

$$\frac{1}{1000} \sum_{i=1}^{1000} (x_i - \mu)^2 = 1287.11$$

でしょ。標本の分散を同じ式で計算すると、どうしてズレちゃうの？ 同じ式なら同じ値になるんじゃないの？」

「いい疑問だ。その疑問に答えるために、確率変数という考え方を導入しよう」

「確率変数？」

### 3 標本平均と確率分布

「まず《全体》の平均を計算する式

$$\frac{1}{1000} \sum_{i=1}^{1000} x_i$$

と、標本の平均を計算する式

$$\frac{1}{20} \sum_{i=1}^{20} X_i$$

は、よく似ているけど概念的に少し異なるところがある。」

「データの個数が違うね」



「全体のデータ数は 1000 だから、1 つ 1 つの数値は  $1/1000$  の確率で生じていると見なせる。表で書くとこんな感じだよ。

実現した値	65.06	11.44	...	145.86	計
確率	$1/1000$	$1/1000$	...	$1/1000$	1

この確率分布表を確率変数  $X$  で表す。どういう値がどういう確率で現れるかの対応を示したものだよ。確率を具体的に書くとこうなる」

$$X \text{ が } 65.06 \text{ になる確率} \iff P(X = 65.06) = \frac{1}{1000}$$

$$X \text{ が } 11.44 \text{ になる確率} \iff P(X = 11.44) = \frac{1}{1000}$$

「ふむ」

「確率変数  $X$  の期待値は

$$E[X] = \sum_{i=1}^{1000} P(X = x_i) x_i$$

と定義される。いま

$$P(X = x_i) = \frac{1}{1000}$$

だから

$$E[X] = \sum_{i=1}^{1000} P(X = x_i) x_i = \sum_{i=1}^{1000} \frac{1}{1000} x_i = \frac{1}{1000} \sum_{i=1}^{1000} x_i$$

となり、《全体》の平均値は、《全体》に対応する確率変数  $X$  の期待値と完全に一致する」

「そうだね」

「次に、標本から平均を計算する操作を考えてみよう。サイズ 20 の標本を  $X_1, X_2, X_3, \dots, X_{20}$  という記号で表す。標本から平均を計算する操作を確率変数を使って書けば

$$\frac{X_1 + X_2 + X_3 + \dots + X_{20}}{20}$$

となる。これは期待値ではなく、《複数の確率変数を、しかじかの演算によって1つの確率変数として合成しますよ》という計算方法だ。この操作を記号  $\bar{X}$  で表すことにする。つまり

$$\frac{X_1 + X_2 + X_3 + \cdots + X_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$$

だ。標本を使って計算するので《標本平均》と呼ばれることが多い。この《標本平均》は平均という名前がついているけど、

$$\frac{1}{n} \sum_{i=1}^n X_i$$

という、ひとかたまりで確率変数なんだよ。だから1つの定まった値じゃなくて、確率的にいろいろな値をとる。つまり分布をもっている」

「うーん、ややこしくなってきた。標本平均なのに1つの値じゃないって、どういうこと？」

「《全体》の平均値を考えてみよう。これは1つしか存在しない。1000個のデータを足して1000で割った数だから、必ず一つの定まった値となる。その値を記号  $\mu$  で表すと約束した」

「そうだったね」

「でも  $\bar{X}$  は、 $X_1, X_2, X_3, \dots, X_{20}$  という標本の実現値によって、値が異なる。だから分布を持つ」

「あなるほど。1回目にとりだした20個の値と、2回目にとりだした20個の値は違うから、 $\bar{X}$  の実現値は毎回違う値になるんだ」

「そういうこと。」

$$\frac{1}{n} \sum_{i=1}^n X_i$$

は、これ自体が確率変数だってことがポイントだよ」

## 4 確率変数としての不偏分散

「次に《標本平均（確率変数）》をつかい、《不偏分散》を定義する。ここで  $n - 1$  が登場するよ

$$\text{《不偏分散》} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

さてこの不偏分散は確率変数かな？ 定数かな？」

「うーん、どっちかな。  $\bar{X}$  は確率変数だったけど、それをさらに足し合わせるとどうなるのかな」青葉は首をかしげた。

「じゃあ、ヒントを出そう。かく……」

「かく？」

「り…… つ……」

「りつ？」

「へん…… す……」

「確率変数？ ヒントの出しかた下手すぎるでしょ」

「さて準備は整った。さっき僕らはサイズ 20 の標本を何度もとりだし、不偏分散を計算して、さらにその平均を求めた。あの実験は、じつは不偏分散という確率変数の《期待値》を計算する作業に対応している」

「ちょっとなに言ってるかわからない」

「実際に《不偏分散》という確率変数の期待値を計算してみよう」

---

一般的に、分散の定義を

$$V[X] = E[(X - E[X])^2]$$

とする。

以下では《全体》を表す確率変数  $X$  の平均と分散を具体的に

$$E[X] = \mu$$

$$V[X] = \sigma^2$$

とおく。また標本  $X_1, X_2, \dots, X_n$  は  $X$  と同一であり、独立であるとする。

次にいくつか補題を導入するよ。補題は証明の論理構造を明確にするための命題だよ。

**補題 1.**

$$E[X^2] = V[X] + E[X]^2$$

証明.

$$\begin{aligned} V[X] &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2] && \text{展開する} \\ &= E[X^2] - 2E[X]E[X] + E[E[X]^2] && \text{定数を出す} \\ &= E[X^2] - 2E[X]^2 + E[X]^2 && \text{まとめる} \\ &= E[X^2] - E[X]^2 \end{aligned}$$

言い換えると

$$E[X^2] = V[X] + E[X]^2$$

□

**補題 2.**

$$E[X_i^2] = V[X] + E[X]^2$$

証明. 定義により  $X_1, X_2, \dots, X_n$  は《全体》の分布  $X$  と同じだから、任意の  $i$  について

$$E[X] = E[X_i], \quad V[X] = V[X_i]$$

が成立する。

□

**補題 3.**

$$E[\bar{X}^2] = \frac{1}{n}V[X] + E[X]^2$$

証明. 補題 1 の  $X$  を  $\bar{X}$  に置き換える。

$$E[\bar{X}^2] = V[\bar{X}] + E[\bar{X}]^2$$

次に

$$\begin{aligned}
V[\bar{X}] &= V\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right] \\
&= \frac{1}{n^2} V[X_1 + X_2 + \cdots + X_n] \\
&= \frac{1}{n^2} \{V[X_1] + V[X_2] + \cdots + V[X_n]\} \\
&= \frac{1}{n^2} nV[X] = \frac{1}{n} V[X]
\end{aligned}$$

および

$$\begin{aligned}
E[\bar{X}] &= E\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right] \\
&= \frac{1}{n} E[X_1 + X_2 + \cdots + X_n] \\
&= \frac{1}{n} \{E[X_1] + E[X_2] + \cdots + E[X_n]\} \\
&= \frac{1}{n} nE[X] = E[X]
\end{aligned}$$

より

$$E[\bar{X}^2] = V[\bar{X}] + E[\bar{X}]^2 = \frac{1}{n} V[X] + E[X]^2$$

□

これで準備は整った。あとは不偏分散の期待値を変形していけば証明完了だよ。

$$\begin{aligned}
E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] &= E\left[\sum_{i=1}^n (X_i^2 - 2X_i\bar{X} + \bar{X}^2)\right] \\
&= E\left[\sum_{i=1}^n X_i^2 - 2\bar{X}\sum_{i=1}^n X_i + \sum_{i=1}^n \bar{X}^2\right] && \text{総和をばらす} \\
&= E\left[\sum_{i=1}^n X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right] && \text{総和の計算} \\
&= E\left[\sum_{i=1}^n X_i^2 - n\bar{X}^2\right] && \text{まとめる} \\
&= E\left[\sum_{i=1}^n X_i^2\right] - nE[\bar{X}^2] && \text{期待値をばらす}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n E[X_i^2] - nE[\bar{X}^2] && \text{期待値をばらす} \\
&= nE[X^2] - nE[\bar{X}^2] && \text{補題を適用} \\
&= n(V[X] + E[X]^2) - nE[\bar{X}^2] && \text{補題を適用} \\
&= n(V[X] + E[X]^2) - n\left(\frac{1}{n}V[X] + E[X]^2\right) && \text{補題を適用} \\
&= n(V[X] + E[X]^2) - (V[X] + nE[X]^2) \\
&= nV[X] - V[X] = (n-1)V[X] && \text{まとめる} \\
&= (n-1)\sigma^2 && \text{仮定より}
\end{aligned}$$

まとめると、

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = (n-1)\sigma^2$$

であることが分かった。これを使って

$$\begin{aligned}
E\left[\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2\right] &= \left(\frac{1}{n-1}\right) E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] \\
&= \left(\frac{1}{n-1}\right) (n-1)\sigma^2 = \sigma^2
\end{aligned}$$

つまり不偏分散（確率変数）の期待値は《全体》の分散  $\sigma^2$  と一致する。

「おー、なんだかややこしい計算をしたけど、結果はスッキリしたね」

「もし  $n$  で割るとしたら

$$\begin{aligned}
E\left[\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2\right] &= \frac{1}{n} E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] \\
&= \frac{1}{n} (n-1)\sigma^2 \\
&= \frac{n-1}{n} \sigma^2 < \sigma^2
\end{aligned}$$

だから、推定量の平均は必ず《全体》の分散  $\sigma^2$  よりも小さくなる。さきほどの実験結果は

$n$  で割った推定値の平均 : 1184.31

$n - 1$  で割った推定値の平均 : 1279.28

本当の値 : 1287.11

だから、たしかに  $n$  で割った推定量は  $n - 1$  で割った推定量より小さくなっている」

「なるほどー」

「不偏分散のように、期待値が推定の目的であるパラメータと一致する推定量を不偏推定量っていうんだよ。ちなみに不偏だからといって、よい推定量とは限らない。有効性と言って……」

「それはまた今度でいいよ」

「OK。ちょうど 30 分だし、このへんで説明を終わろう。じゃあ柿ピーでも食べようか」

花京院は柿の種の小袋を一つ開封した。

おわり

## まとめ

Q 不偏分散はどうして  $n - 1$  で割るの？

A  $n - 1$  で割る推定量を使うと、その期待値が《全体》の分散と一致するからです。

- パラメータの推定に用いる確率変数（の関数）を推定量といいます。特に期待値がパラメータと一致する推定量を不偏推定量と言います。
- 標本平均  $\frac{1}{n} \sum_{i=1}^n X_i$  や不偏分散  $\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$  は不偏推定量の一種です。
- 確率変数（不偏推定量）としての不偏分散は分布を持っています。標本のデータから計算した不偏推定値は、不偏推定量の分布の実現値と考えることができます。