第 IX 部　母平均の差の検定

目次

第 IX 部　母平均の差の検定 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 1

1　予備的考察 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 4

1.1　はじめに ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 4

1.1.1　ポイント ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・ 4

1.2 総和 · · · · · · · · · · 5
1.2.1 総和の公式 · · · · · · · · · · 9
1.3 平均と分散 · · · · · · · · · · 10
1.3.1 分散 · · · · · · · · · · 13
1.3.2 標準偏差 · · · · · · · · · · 13
1.3.3 標準化 · · · · · · · · · · 16
1.4 標本空間と事象 · · · · · · · · · · 17
1.4.1 事象の演算 · · · · · · · · · · 18
1.4.2 排反な事象 · · · · · · · · · · 19
1.5 確率の定義 · · · · · · · · · · 20
1.5.1 確率の公理主義的定義 · · · · · · · · · · 21

1.6 確率変数とは 24
1.6.1 サイコロを使った例 24
1.6.2 期待値 25
1.6.3 期待値の演算 25
1.7 期待値からの偏差 29
1.8 まとめ 32

1　予備的考察

1.1　はじめに

- 統計学ではいきなり高度な数学の知識を使い始めます。
- 前提がきつすぎるので、これまで説明した内容を確認します。
- これらは一度は説明している内容です。

1.1.1　ポイント

- 総和
- 平均と分散と標準偏差
- 確率
- 確率変数
- 期待値と分散と標準偏差

1.2 総和

定義

n 個の数値があり、i 番目の値を x_i とあらわす。x_i を全て足し合わせた値を『**総和**』といい、$\displaystyle\sum_{i=1}^{n} x_i$ であらわす。

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n \tag{1}$$

全ての x_i が定数 c に等しいとき、添え字 i を使わずに

$$\sum_{i=1}^{n} c \tag{2}$$

とあらわす。

全ての x_i を定数 a 倍した値の総和を

$$\sum_{i=1}^{n} a x_i \tag{3}$$

とあらわす。

2 つの数値の組が n 組あり、i 番目の数値の組を $(x_i,\ y_i)$ とあわらす。$x_i + y_i$ を全て足し合わせた値を

$$\sum_{i=1}^{n} (x_i + y_i) \tag{4}$$

とあらわす。

問題 IX − 1 − 1

$$\sum_{i=1}^{n} c = nc$$

が成り立つことを示しなさい。

問題 IX − 1 − 2

$$\sum_{i=1}^{n} ax_i = a\sum_{i=1}^{n} x_i$$

が成り立つことを示しなさい。

問題 IX − 1 − 3

$$\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$$

が成り立つことを示しなさい。

解例 IX − 1 − 1

$$\sum_{i=1}^{n} c = \overbrace{c + c + \cdots + c}^{n\text{ 個}} = nc$$

解例 IX − 1 − 2

$$\begin{aligned}\sum_{i=1}^{n} ax_i &= ax_1 + ax_2 + \cdots + ax_n \\ &= a\left(x_1 + x_2 + \cdots + x_n\right) \\ &= a\sum_{i=1}^{n} x_i\end{aligned}$$

解例 IX − 1 − 3

$$\sum_{i=1}^{n} (x_i + y_i)$$

$$\begin{aligned}&= (x_1 + y_1) + (x_2 + y_2) + \cdots + (x_n + y_n) \\ &= x_1 + y_1 + x_2 + y_2 + \cdots + x_n + y_n \\ &= x_1 + x_2 + \cdots + x_n + y_1 + y_2 + \cdots + y_n \\ &= (x_1 + x_2 + \cdots + x_n) + (y_1 + y_2 + \cdots + y_n) \\ &= \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i\end{aligned}$$

1.2.1 総和の公式

総和の演算において以下の関係が成り立つ。

$$\sum_{i=1}^{n} c = nc \quad ; c \text{ は定数} \tag{5}$$

$$\sum_{i=1}^{n} a x_i = a \sum_{i=1}^{n} x_i \quad ; a \text{ は定数} \tag{6}$$

$$\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i \tag{7}$$

1.3　平均と分散

定義

$n \geqq 1$ とする。x_i の総和を数値の個数 n で割った値を『**平均**』といい、$\bar{x}$ であらわす。

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{8}$$

(8) 右辺の分母を払うと

$$n\bar{x} = \sum_{i=1}^{n} x_i \tag{9}$$

を得る。

定義

x_i と $\bar{x}$ の差を『**平均からの偏差**』という。

$$x_i - \bar{x} \tag{10}$$

問題 IX − 1 − 4

$$\sum_{i=1}^{n} (x_i - \bar{x})$$

を計算しなさい。

問題 IX − 1 − 5

$$\sum_{i=1}^{n} (x_i - \bar{x})^2$$

を計算しなさい。

解例 IX − 1 − 4

$$\begin{aligned}\sum_{i=1}^{n}(x_i - \bar{x}) &= \sum_{i=1}^{n} x_i - n\bar{x} \\ &= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i \\ &= 0\end{aligned}$$

解例 IX − 1 − 5

$$\begin{aligned}\sum_{i=1}^{n}(x_i - \bar{x})^2 &= \sum_{i=1}^{n}\left({x_i}^2 - 2\bar{x}x_i + \bar{x}^2\right) \\ &= \sum_{i=1}^{n}{x_i}^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2 \\ &= \sum_{i=1}^{n}{x_i}^2 - 2\bar{x}n\bar{x} + n\bar{x}^2 \\ &= \sum_{i=1}^{n}{x_i}^2 - 2n\bar{x}^2 + n\bar{x}^2 \\ &= \sum_{i=1}^{n}{x_i}^2 - n\bar{x}^2\end{aligned}$$

1.3.1　分散

『平均からの偏差』の総和は常に 0 となってしまうため、データのばらつきの指標としては使えない。そこで「『平均からの偏差』の二乗」の平均を『**分散**』といい V_x とあらわす。

$$V_x = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{11}$$

1.3.2　標準偏差

分散 V_x の正の平方根を『**標準偏差**』といい、σ_x または D_x であらわす。

$$\sigma_x = \sqrt{V_x} \tag{12}$$

問題 IX − 1 − 6

$z_i = \dfrac{x_i - \bar{x}}{\sigma_x}$ とする。ただし、$\sigma_x > 0$

$$\bar{z}$$

を計算しなさい。

問題 IX − 1 − 7

$z_i = \dfrac{x_i - \bar{x}}{\sigma_x}$ とする。ただし、$\sigma_x > 0$

$$V_z$$

を計算しなさい。

解例 IX − 1 − 6

$$
\begin{aligned}
\bar{z} &= \frac{1}{n}\sum_{i=1}^{n} z_i \\
&= \frac{1}{n}\sum_{i=1}^{n} \frac{x_i - \bar{x}}{\sigma_x} \\
&= \frac{1}{n}\frac{1}{\sigma_x}\sum_{i=1}^{n} (x_i - \bar{x}) \\
&= \frac{1}{n}\frac{1}{\sigma_x} \times 0 \\
&= 0
\end{aligned}
$$

解例 IX − 1 − 7

$$
\begin{aligned}
V_z &= \frac{1}{n}\sum_{i=1}^{n} (z_i - \bar{z})^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} (z_i)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma_x}\right)^2 \\
&= \frac{1}{{\sigma_x}^2}\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 \\
&= \frac{V_x}{{\sigma_x}^2} = \frac{V_x}{V_x} = 1
\end{aligned}
$$

1.3.3 標準化

『平均からの偏差』を『標準偏差』で割ることを**『標準化』**といい、標準化された値を z_i であらわす。

$$z_i = \frac{x_i - \bar{x}}{\sigma_x} \tag{13}$$

標準化すると z_i の平均は 0、標準偏差は 1 になる。

1.4 標本空間と事象

- 観測されうる結果を『**標本点**』とよび ω_i であらわす。
- 標本点の全体の集合を『**標本空間**』とよび Ω であらわす。
- 標本空間の部分集合を『**事象**』とよぶ。
- 標本点を一つも含まないことも事象とみなし『**空(くう)事象**』とよび $\varnothing$ であらわす。
- 標本点 $\omega_1, \omega_2, \cdots, \omega_n$ からなる事象 A は、$\{\quad\}$ を使って、
  $A = \{\omega_1, \omega_2, \cdots, \omega_n\}$ とあらわす。
- ただ一つの標本点からなる事象を『**根元(こんげん)事象**』とよび、二つ以上の標本点を含む事象を『**複合事象**』とよぶ。

1.4.1　事象の演算

- 事象 A と 事象 B 二つの事象において A と B が共通の標本点を持たない場合 A と B は『**排反事象**（はいはん）』であるという。また、A と B が共通の標本点を持たないことを A と B は排反であるという。
- 事象 A と事象 B 二つの事象のうち少なくとも一つがおこる事象を A と B の『**和事象**（わ）』とよび $A \cup B$ であらわす。
- 事象 A と事象 B 二つの事象が同時に起こる事象を『**積事象**（せき）』とよび $A \cap B$ であらわす。
- 事象 A が起こらない事象を『**補事象**（ほ）』とよび A^C であらわす[1]。

1）　Complement：補完・補足

1.4.2 排反な事象

- X と Y が排反であって、X と Z が排反であって、Y と Z が排反であるとき X と Y と Z は『**互いに排反**』であるという。
- A と A^C の積事象は $\varnothing$ である。
- A と A^C は排反である。
- A と A^C の和事象は Ω である。

1.5　確率の定義

確率とは事象の起こりやすさを定量的に示すもの[2)]であり、事象 A の起こる確率を $\Pr(A)$ であらわす。

2)　可能性は可能か不可能かの二択であり、蓋然性（がいぜんせい）は確実性の度合いを意味し、確率とほぼ同じ意味を持つ。

1.5.1　確率の公理主義的定義

確率の公理は次の 3 つからなる。これに合うならどのような数も確率と認められる。

1. すべての事象 A に対して、
$$0 \leqq \Pr(A) \leqq 1$$
2. 標本空間全体の確率は、
$$\Pr(\Omega) = 1$$
3. 互いに排反な事象 $A_1, A_2, A_3, \cdots$ に対して、
$$\Pr\left(A_1 \cup A_2 \cup A_3 \cup \cdots\right) = \Pr\left(A_1\right) + \Pr\left(A_2\right) + \Pr\left(A_3\right) + \cdots$$

定義

事象 B が起こったとわかっている場合の事象 A の起こる確率を『**B を条件とする A の条件付き確率**』といい $\Pr(A \mid B)$ とあらわし

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} \tag{14}$$

と定義する。

(14) の分母を掃うと

$$\Pr(A \cap B) = \Pr(B)\ \Pr(A \mid B) \tag{15}$$

を得る。

定義

事象 A と 事象 B とそれらの積事象 $A \cap B$ の確率の関係が、

$$\Pr(A \cap B) = \Pr(A)\ \Pr(B) \tag{16}$$

である場合、事象 A と 事象 B は『**独立**』であるという。

積事象の確率を二つの事象の確率の積としてあらわすことができるとき二つの事象は独立である。

　複数の事象が同時に起こる事象つまり積事象の確率を『**同時確率**』という。事象が独立な時、同時確率は事象の確率の積としてあらわすことができる。

1.6　確率変数とは

変数が取りうる各値に対して確率が与えられている変数を『**確率変数**』という。確率変数は大文字を使ってあらわす。

1.6.1　サイコロを使った例

事象を識別する添え字を i とし、それぞれの事象を x_i であらわし x_i の事象がおこる確率を p_i とあらわす。サイコロを振ると x_i のどれかが観測される。この観測値はそれぞれ一定の確率 p_i に従い出現する。サイコロを振って出た目を X とすると、$X = x_i$ には確率 p_i が対応する。従って X は確率変数である。

表 1　サイコロの目と確率

X	1	2	3	4	5	6
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

1.6.2 期待値

確率変数の対応する確率による加重平均を『**期待値**』といい $\mathsf{E}(X)$ であらわす[3)]。期待値 $\mathsf{E}(X)$ は確率変数 $X = x_i$ と、x_i に対応する確率 p_i の積の総和であり、

$$\mathsf{E}(X) = \sum_{i=1}^{n} x_i p_i \tag{17}$$

と定義する。

3） Expected value の頭文字由来。確率変数との違いをあらわすため立体で E() と表記する。

1.6.3 期待値の演算

期待値に関して以下の関係が成り立つ

$$\mathsf{E}(c) = c \qquad ;\ c \text{ は定数} \tag{18}$$

$$\mathsf{E}(X + c) = \mathsf{E}(X) + c \qquad ;\ c \text{ は定数} \tag{19}$$

$$\mathsf{E}(aX) = a\,\mathsf{E}(X) \qquad ;\ a \text{ は定数} \tag{20}$$

問題 IX − 1 − 8

以下の関係が成り立つことを示しなさい。ただし、a, c は定数とする。

(i) $\mathsf{E}(c) = c$

(ii) $\mathsf{E}(X + c) = \mathsf{E}(X) + c$

(iii) $\mathsf{E}(aX) = a\,\mathsf{E}(X)$

問題 IX − 1 − 9

$$\mathsf{E}(X - \mathsf{E}(X))$$

を計算しなさい。

解例 IX − 1 − 8

(i)

$$\begin{aligned}\mathsf{E}(c) &= \sum_{i=1}^{n} c p_i \\ &= c\sum_{i=1}^{n} p_i \\ &= c\end{aligned}$$

(ii)

$$\begin{aligned}\mathsf{E}(X+c) &= \sum_{i=1}^{n} (x_i + c)\, p_i \\ &= \sum_{i=1}^{n} (x_i p_i + c p_i) \\ &= \sum_{i=1}^{n} x_i p_i + \sum_{i=1}^{n} c p_i \\ &= \sum_{i=1}^{n} x_i p_i + c\sum_{i=1}^{n} p_i \\ &= \mathsf{E}(X) + c\end{aligned}$$

(iii)

$$\begin{aligned}\mathsf{E}(aX) &= \sum_{i=1}^{n} a x_i p_i \\ &= a \sum_{i=1}^{n} x_i p_i \\ &= a \; \mathsf{E}(X)\end{aligned}$$

解例 IX − 1 − 9

$$\begin{aligned}\mathsf{E}(X - \mathsf{E}(X)) &= \mathsf{E}(X) - \mathsf{E}(X) \\ &= 0\end{aligned}$$

1.7　期待値からの偏差

確率変数 X の期待値からの偏差 $X - \mathsf{E}(X)$ を考える。$X - \mathsf{E}(X)$ が確率変数なので期待値を求めると

$$\mathsf{E}(X - \mathsf{E}(X)) = 0 \tag{21}$$

となってしまう。

この符号の問題を解決する為に $X - \mathsf{E}(X)$ を二乗して期待値を取ると

$$\begin{aligned}
\mathsf{E}\left(\left(X - \mathsf{E}(X)\right)^2\right) &= \sum_{i=1}^{n} \left(x_i - \mathsf{E}(X)\right)^2 p_i && (22)\\
&= \sum_{i=1}^{n} \left({x_i}^2 - 2\mathsf{E}(X)x_i + \mathsf{E}(X)^2\right) p_i && (23)\\
&= \sum_{i=1}^{n} \left({x_i}^2 p_i - 2\mathsf{E}(X)x_i p_i + \mathsf{E}(X)^2 p_i\right) && (24)\\
&= \sum_{i=1}^{n} {x_i}^2 p_i - 2\mathsf{E}(X)\sum_{i=1}^{n} x_i p_i + \mathsf{E}(X)^2 \sum_{i=1}^{n} p_i && (25)
\end{aligned}$$

第一項において ${x_i}^2$ を確率変数とみると ${x_i}^2$ の期待値なので

$$= \mathsf{E}(X^2) - 2\mathsf{E}(X)\mathsf{E}(X) + \mathsf{E}(X)^2 \tag{26}$$

$$= \mathsf{E}(X^2) - \mathsf{E}(X)^2 \tag{27}$$

これを X の『**分散**』といい、$\mathsf{V}(X)$ であらわす。また分散 $\mathsf{V}(X)$ の正の平方根を X の『**標準偏差**』といい $\mathsf{D}(X)$ であらわす[4)]。

$$\begin{aligned} \mathsf{V}(X) &= \mathsf{E}\left(\left(X - \mathsf{E}(X)\right)^2\right) \\ &= \mathsf{E}(X^2) - \mathsf{E}(X)^2 \end{aligned} \tag{28}$$

4） 分散を ${\sigma_X}^2$、標準偏差を σ_X であらわすことも多い

1.8　まとめ

- 今後利用する予定の数学的な知識を確認した。
- ここで示したものは過去に紹介した内容である。
- 前提は多岐にわたるため、常に参照できるようにしておく必要がある。
- 何度も、見直すことを勧める。

参考文献

1.『経済数学教室 1 巻』小山昭雄著「岩波書店」1994 年 5 月 30 日
2.『新体系 大学数学 入門の教科書 上下』吉沢光雄著「講談社」2022 年 6 月 16 日
3.『技術者のための高等数学 7 確率と統計（原書代 8 版)』E・クライツィィグ著　田栗正章訳「培風館」2004 年 11 月 30 日第 8 版
4.『データ解析のための数理統計入門』久保川達也著「共立出版」2023 年 10 月 15 日初版
5.『統計学入門』東京大学教養学部統計学教室篇「東京大学出版会」1991 年 7 月 9 日初版
6.『行動科学における統計解析法」芝祐順・南風原朝和著　「東京大学出版会」1990 年 3 月 20 日　初版
7.『計量経済学の理論』A.S. ゴールドバーガー著, 福地崇生・森口親司訳「東洋経済新報社」

1970 年 9 月 10 日 第 1 刷

8. 『計量経済学序説』R.J. ウォナコット・T.H. ウォナコット著, 国府田恒夫・田中一盛訳「培風館」 1975 年 6 月 10 日初版
9. 『マーケティング・サイエンス』片平秀貴著 「東京大学出版会」1987 年 4 月 20 日初版
10. 『回帰分析』佐和光男著 「朝倉書店」1979 年 4 月 20 日初版
11. 『完全独習 統計学入門』小島寛之著「ダイヤモンド社」2006 年 9 月 28 日 初版
12. 『分散分析の基礎』高橋敬子著「プレアデス出版」2009 年 10 月 1 日