# 1 統計の基礎

## 1.1 宝くじの期待値と分散

テキスト

芳賀敏郎（2011）医薬品開発のための統計解析
第１部　基礎　改訂版、サイエンティスト社、p.275

# 第１部　基礎

**1. 統計の基礎**・・・・・**1.1 宝くじの期待値と分散**、1.2 サイコロの目の数の期待値と分散

1.3 分散の加法性・中心極限定理・正規分布、1.4 統計的推測、1.5 モデル

2. １組のデータの解析　2.1 データの特徴の記述、2.2 データのグラフ表示と外れ値

2.3 対数変換と対数正規分布、2.4 平均に関する推測（母標準偏差 $\sigma$ 既知）

2.5 分散に関する推測、2.6 平均に関する推測（母標準偏差 $\sigma$ 未知）

3. ２組のデータの解析　3.1 データのグラフ化、3.2 平均値の差の $t$ 検定、3.3 分散の違いの検定

3.4 分散が異なる場合の平均値の差の比較

3.5 対応のある場合の平均値の差の $t$ 検定、3.6 検出力と $n$ の決め方

3.7 ノンパラメトリック検定

4. 相関・回帰・・・・・4.1 散布図、4.2 相関係数、4.3 回帰モデルとモデルの推定

4.4 誤差を考慮した推定、4.5 回帰分析適用上の諸問題

# 1.3 宝くじの期待値と分散

p.7

テキストの
該当ページ

（1）期待値
（2）分散
（3）標準偏差

使用するファイル
Excel ファイル「基本改1.xls」
サイエンティスト社ホームページからダウンロード

JMP 10.0.2 の出力を表示（本節では未使用）

★プレゼンテーションの
スピーカーノートを、
PDF の注釈に変換してあります

# （1）期待値

# 宝くじ１本の価値



●宝くじの事例

10,000 本の宝くじ


宝くじ　10,000枚
番号　00001 ～10000

表示1.1.1　宝くじの賞金と本数

| 行番号 $i$ | 等級 | 賞金(円) $x_i$ | 本数(本) $n_i$ | 賞金額(円) $x_i n_i$ |
|---|---|---|---|---|
| 1 | 1等 | 10,000 | 10 | 100,000 |
| 2 | 2等 | 1,000 | 200 | 200,000 |
| 3 | 3等 | 100 | 4,000 | 400,000 |
| 4 | はずれ | 0 | 5,790 | 0 |
| | 合計 | | 10,000 $N$ | 700,000 $T$ |

# 宝くじ１本の価値



●宝くじの１本の価値

10,000 本の宝くじ
この中の１本を持っている（番号：00352）
10,000 円、1,000円、100 円、0 円 のいずれかである
抽選前、この１本にはいくらの価値があると考えたらよいか

宝くじ 00352

賞金は 10,000 円？
1,000円？
100 円？
0 円？


宝くじ　10,000枚
番号　00001 ～10000

表示1.1.1　宝くじの賞金と本数

| 行番号 $i$ | 等級 | 賞金(円) $x_i$ | 本数(本) $n_i$ | 賞金額(円) $x_i n_i$ |
|---|---|---|---|---|
| 1 | 1等 | 10,000 | 10 | 100,000 |
| 2 | 2等 | 1,000 | 200 | 200,000 |
| 3 | 3等 | 100 | 4,000 | 400,000 |
| 4 | はずれ | 0 | 5,790 | 0 |
| | 合計 | | 10,000 $N$ | 700,000 $T$ |

# 宝くじ１本の価値

p.7

## ●全部の宝くじを買い占めた場合

賞金額の合計　$T$ = 700,000 円

本数の合計　$N$ = 10,000 本

宝くじ１本あたりの賞金　$T/N$ = 700,000 / 10,000 = 70 円

宝くじ 00352

賞金は 10,000 円？
1,000円？
100 円？
0 円？


宝くじ　10,000枚
番号　00001 ～10000

表示1.1.1　宝くじの賞金と本数

| 行番号 | 等級 | 賞金(円) | 本数(本) | 賞金額(円) |
|---|---|---|---|---|
| $i$ | | $x_i$ | $n_i$ | $x_i n_i$ |
| 1 | 1等 | 10,000 | 10 | 100,000 |
| 2 | 2等 | 1,000 | 200 | 200,000 |
| 3 | 3等 | 100 | 4,000 | 400,000 |
| 4 | はずれ | 0 | 5,790 | 0 |
| | 合計 | | 10,000 | 700,000 |
| | | | $N$ | $T$ |

# 期待値

p.7

●全部の宝くじを買い占めた場合

賞金額の合計　$T$ = 700,000 円
本数の合計　$N$ = 10,000 本
宝くじ 1 本あたりの賞金　$T/N$ = 700,000 / 10,000 = 70 円

宝くじ 00352

賞金は 10,000 円？
1,000円？
100 円？
0 円？

●期待値

抽選前の 1 本の宝くじは、
１等に当たるか、はずれになるか、
分からない
平均的には 70 円の賞金が期待される

「この宝くじ 1 本の
期待値（Expectation）は 70 円」

「この宝くじの賞金 $x$ の期待値は 70 円」

表示1.1.1　宝くじの賞金と本数

| 行番号 $i$ | 等級 | 賞金(円) $x_i$ | 本数(本) $n_i$ | 賞金額(円) $x_i n_i$ |
|---|---|---|---|---|
| 1 | 1等 | 10,000 | 10 | 100,000 |
| 2 | 2等 | 1,000 | 200 | 200,000 |
| 3 | 3等 | 100 | 4,000 | 400,000 |
| 4 | はずれ | 0 | 5,790 | 0 |
| | 合計 | | 10,000 $N$ | 700,000 $T$ |

# 期待値



## ●本数の合計 $N$ と賞金額の合計 $T$ の計算

$i$ 等（1～4 等）の賞金 $x_i$、本数 $n_i$ としたときの、本数の合計 $N$、賞金額の合計 $T$

添え字 $i$ は行番号に対応
1～4 まで変化

$$N = n_1 + n_2 + n_3 + n_4 \qquad (1.1.1)$$
$$= 10 + 200 + 400 + 5790$$
$$= 10000$$

$$T = x_1 n_1 + x_2 n_2 + x_3 n_3 + x_4 n_4 \qquad (1.1.1)$$
$$= 10000 \times 10 + 1000 \times 200$$
$$+100 \times 4000 + 0 \times 5790$$
$$= 700000$$

添え字 $i$ は行番号に対応
1～4 まで変化

表示1.1.1　宝くじの賞金と本数

| 行番号 | 等級 | 賞金(円) | 本数(本) | 賞金額(円) |
|---|---|---|---|---|
| $i$ | | $x_i$ | $n_i$ | $x_i n_i$ |
| 1 | 1等 | 10,000 | 10 | 100,000 |
| 2 | 2等 | 1,000 | 200 | 200,000 |
| 3 | 3等 | 100 | 4,000 | 400,000 |
| 4 | はずれ | 0 | 5,790 | 0 |
| | 合計 | | 10,000 | 700,000 |
| | | | $N$ | $T$ |

# 期待値



## ●本数の合計 $N$ と賞金額の合計 $T$ の計算

$i$ 等（1～4 等）の賞金 $x_i$、本数 $n_i$ としたときの、本数の合計 $N$、賞金額の合計 $T$

$$\begin{aligned} N &= n_1 + n_2 + n_3 + n_4 \\ &= 10 + 200 + 400 + 5790 \\ &= 10000 \end{aligned} \tag{1.1.1}$$

$$\begin{aligned} T &= x_1 n_1 + x_2 n_2 + x_3 n_3 + x_4 n_4 \\ &= 10000 \times 10 + 1000 \times 200 \\ &\qquad + 100 \times 4000 + 0 \times 5790 \\ &= 700000 \end{aligned}$$

$x_1 = 10000$
$x_2 = 1000$
$x_3 = 100$
$x_4 = 0$

$n_1 = 10$
$n_2 = 200$
$n_3 = 4000$
$n_4 = 5790$

$x_1 n_1 = 10000 \times 10$
$x_2 n_2 = 1000 \times 200$
$x_3 n_3 = 100 \times 4000$
$x_4 n_4 = 0 \times 5790$

表示1.1.1　宝くじの賞金と本数

| 行番号 $i$ | 等級 | 賞金(円) $x_i$ | 本数(本) $n_i$ | 賞金額(円) $x_i n_i$ |
|---|---|---|---|---|
| 1 | 1等 | 10,000 | 10 | 100,000 |
| 2 | 2等 | 1,000 | 200 | 200,000 |
| 3 | 3等 | 100 | 4,000 | 400,000 |
| 4 | はずれ | 0 | 5,790 | 0 |
| | 合計 | | 10,000 $N$ | 700,000 $T$ |

# 期待値



●本数の合計 $N$ と賞金額の合計 $T$ の計算

$i$ 等（1～4 等）の賞金 $x_i$、本数 $n_i$ としたときの、本数の合計 $N$、賞金額の合計 $T$

$$N = n_1 + n_2 + n_3 + n_4 = \sum_{i=1}^{4} n_i$$

$$T = x_1 n_1 + x_2 n_2 + x_3 n_3 + x_4 n_4 = \sum_{i=1}^{4} x_i n_i \tag{1.1.2}$$

Σ（シグマ）：総和（summation）
$i$ を 1, 2, 3, 4 と変えて加える

「この宝くじ 1 本の期待値は 70 円」

$$\frac{T}{N} = \sum_{i=1}^{4} x_i n_i \Big/ \sum_{i=1}^{4} n_i = 70$$

表示1.1.1　宝くじの賞金と本数

| 行番号 $i$ | 等級 | 賞金(円) $x_i$ | 本数(本) $n_i$ | 賞金額(円) $x_i n_i$ |
|---|---|---|---|---|
| 1 | 1等 | 10,000 | 10 | 100,000 |
| 2 | 2等 | 1,000 | 200 | 200,000 |
| 3 | 3等 | 100 | 4,000 | 400,000 |
| 4 | はずれ | 0 | 5,790 | 0 |
| | 合計 | | 10,000 $N$ | 700,000 $T$ |

# 期待値



- 期待値の計算

Expectation の頭文字

$i$ 等（1～$m$ 等）の賞金 $x_i$、本数 $n_i$ としたときの、賞金 $x$ の期待値 $\mathrm{E}[x]$

4 を $m$ に拡大

$$N = \sum_{i=1}^{m} n_i$$

$$T = \sum_{i=1}^{m} x_i n_i \tag{1.1.2}$$

$$\mathrm{E}[x] = \frac{T}{N} = \sum_{i=1}^{m} x_i n_i \Big/ \sum_{i=1}^{m} n_i \tag{1.1.3}$$

$$\mathrm{E}[x] = \frac{T}{N} = \frac{1}{N}\sum_{i=1}^{m}(x_i n_i) = \sum_{i=1}^{m}\left(x_i \frac{n_i}{N}\right) \tag{1.1.4}$$

$$= \sum_{i=1}^{m} x_i \pi_i$$

$N$ は $i$ に関わらず一定：定数

$i$ 等の確率
$\frac{n_i}{N} = \pi_i$

積の合計を $N$ (定数)で割る代わりに
個々の等級の本数 $n_i$ を $N$ で割る

# 期待値

## ●期待値の計算

$$\mathrm{E}[x] = \sum_{i=1}^{m} x_i \pi_i \qquad (1.1.4)$$

$$= 10000 \times \frac{10}{10000} + 1000 \times \frac{200}{10000} + 100 \times \frac{4000}{10000} + 0 \times \frac{5790}{10000}$$

$$= 10 + 20 + 40 + 0$$

$$= 70$$

賞金 $x$ の期待値 $\mathrm{E}[x]$ は
各賞金 $x_i$ と
その確率 $\pi_i$ の積の和

$i$ 等の確率
$\frac{n_i}{N} = \pi_i$

1 等の場合

$\pi_1 = 10/10000 = 0.001$

0.1% の割合で 1 等になる

表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i$*$n_i$ | $\pi_i$ | $x_i$*$\pi_i$ |
| 14 | 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |
| | 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| | 18行目 | | =SUM(C14:C17) | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

# 期待値



## ●期待値の計算

賞金 $x$ の期待値　E[$x$] は、各賞金 $x_i$ とその確率 $\pi_i$ の積の和

確率 $\pi_i$ の合計は 1 になる

$$\sum_{i=1}^{m}\pi_i=\sum_{i=1}^{m}\frac{n_i}{N}=\frac{1}{N}\sum_{i=1}^{m}n_i=\frac{N}{N}=1$$

賞金 $x$ の期待値 E[$x$] は
各賞金 $x_i$ と
その確率 $\pi_i$ の積の和

確率 $\pi_i$ の合計は 1

表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i$*$n_i$ | $\pi_i$ | $x_i$*$\pi_i$ |
| 14 | 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |
| | 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| | 18行目 | | =SUM(C14:C17) | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

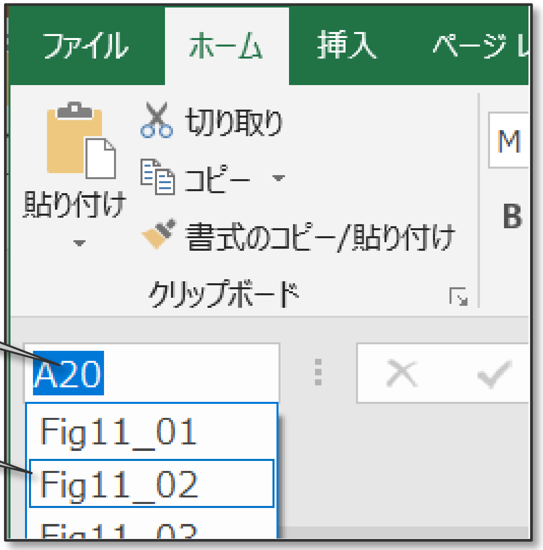
# Excel を使った期待値の計算



●Excel のセル

Excelファイル「基本改1.xls」の読み込み

名前ボックスで「Fig11_2」を選択、表示1.1.2 に移動




名前ボックス

Fig11_02

セルの表示・・・・列の文字＋行の番号

セル範囲の表示・・範囲の左上隅のセル：右下隅のセル

期待値 E[$x$] = 70 が計算される過程を理解する

セル範囲 B14:C17

セル F18

表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i*n_i$ | $\pi_i$ | $x_i*\pi_i$ |
| 14 | 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |
| | 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| | 18行目 | | =SUM(C14:C17) | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

# Excel を使った期待値の計算

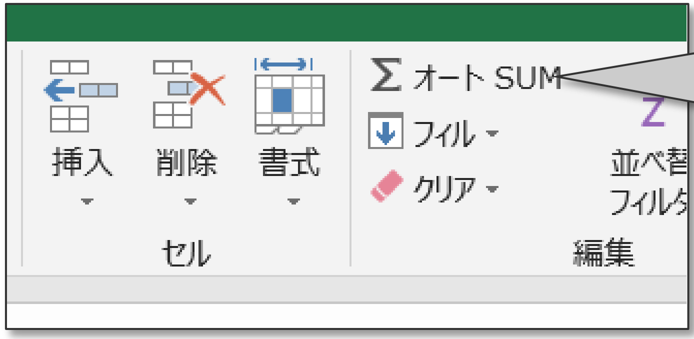
●A 列, B 列, C 列の計算式

A 列：文字列
計算に用いられない

B 列：宝くじの賞金 $x_i$

C 列：宝くじの本数 $n_i$

C18：本数の合計 $T$ 、SUM 関数を入力
「=sum(」を入力後
C14 から C17 まで選択
Enter キーをクリック




セル C18 を選択して
「オートSUMボタン」をクリック
C14 のセルをクリックし、
C17 までドラッグ、Enterキー
p.9 脚注　p.184 参照

**セルC18：** =SUM(C14:C17)
p.69参照

表示1.1.2　期待値の計算(1)

計算表

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | [illegible] | $\pi_i$ | $x_i * \pi_i$ |
| 14 | 1等 | 10,000 | 10 | [illegible],000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |

計算式の説明

| | | | | | | |
|---|---|---|---|---|---|---|
| 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| 18行目 | =SUM(C14:C17) | | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

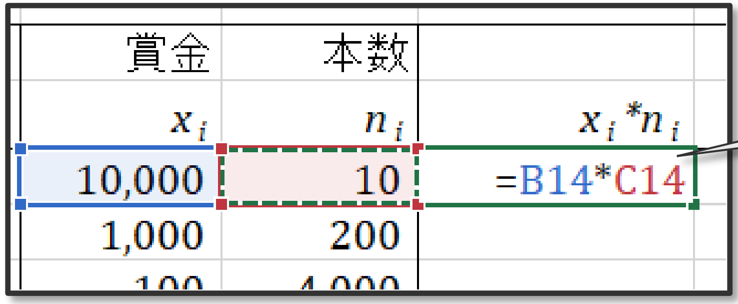
# Excel を使った期待値の計算



- D 列の計算式

  D 列：賞金 $x_i$ と本数 $n_i$ の積
  （賞金額）

  D18：賞金 $x_i$ と本数 $n_i$ の積の和
  SUM 関数を入力




D14 を選択、以下の順に入力
“=“，B14 をクリック，“*”，
C14 をクリック，Enter キー

**セルD14：**
=B14*C14

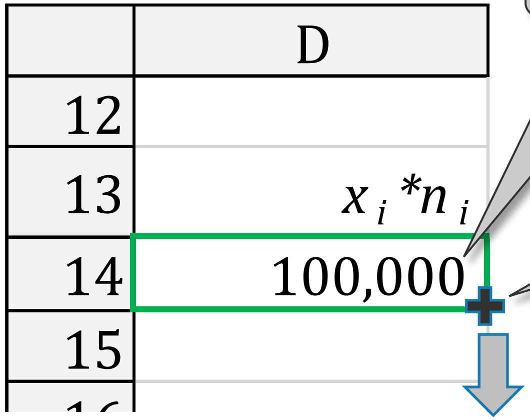
表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i*n_i$ | $\pi_i$ | $x_i*\pi_i$ |
| 14 | 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |
| | 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| | 18行目 | | =SUM(C14:C17) | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

# Excel を使った期待値の計算

p.9

- D 列の計算式
  - セル D14 に式を入力
  - オートフィルでセル D14 を D15:D17 にコピー
  - オートフィル機能を利用




=B14*C14

D14に式を入力後
D14 を選択して「＋マーク」を下にドラッグ
オートフィル機能 p.12参照

=B14*C14　元の数式
=B15*C15
=B16*C16
=B17*C17
自動的に変化

表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i$*$n_i$ | $\pi_i$ | $x_i$*$\pi_i$ |
| 14 | 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |
| | 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| | 18行目 | | =SUM(C14:C17) | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

# Excel を使った期待値の計算

p.9

- E 列の計算式

E列：等級ごとの確率 $\pi_i$

セル E14に　“=E14/C$18” を入力（F4 キーを利用）

E14 に入力した式を、D 列と同様に、E15～E17 にコピー

（オートフィル機能を利用）

C18 の「18」を固定するために C$18 とする（「$マーク」に注目）

=C14/C18
=C15/C19
=C16/C20
=C17/C21
ではない

=C14/C$18
=C15/C$18
=C16/C$18
=C17/C$18

Excel を上手に使いこなすには
「$」 マークを適切に用いる
技術が必須（絶対参照）

表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i * n_i$ | $\pi_i$ | $x_i * \pi_i$ |
| 14 | 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |
| | 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| | 18行目 | | =SUM(C14:C17) | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

# Excel を使った期待値の計算



● F4 キーの利用

絶対参照の指定

| 本数 | | 確率 |
|---|---|---|
| $n_i$ | $x_i{}^*n_i$ | $\pi_i$ |
| 10 | 100,000 | =C14/C18 |
| 200 | 200,000 | 0.0200 |
| 4,000 | 400,000 | 0.4000 |
| 5,790 | 0 | 0.5790 |
| 10,000 | 700,000 | 1.0000 |

**セルE14** は、以下の順序で入力
“=“、セル C14 をクリック、”/” 、セル C18 をクリック、
F4 キーを押すたびに(1), (2), (3), (4), (1)･･･ に変化
今回は (3) の状態で Enter キーを押す　(p.80 参照)

(1) =C14/C18

(2) =C14/$C$18

(3) =C14/C$18

(4) =C14/$C18

# Excel を使った期待値の計算



## ●F 列の入力

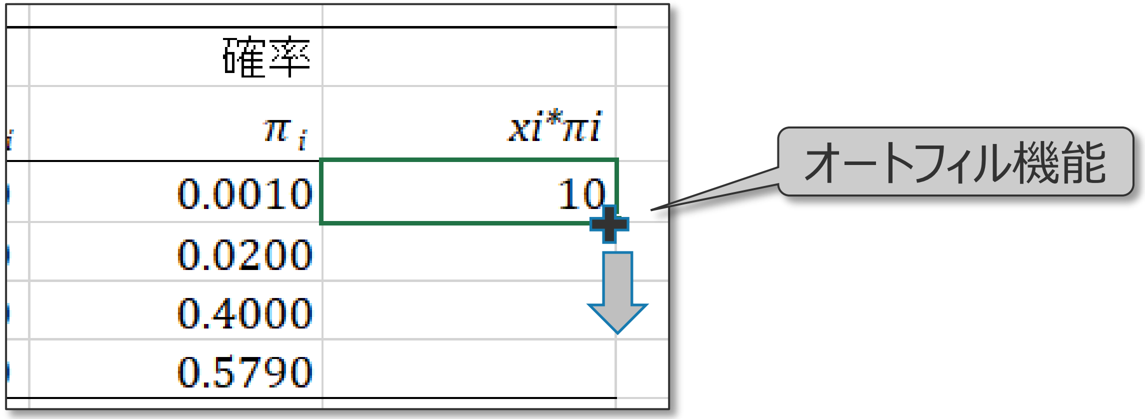
F列：賞金 $x_i$ と確率 $\pi_i$ の積

D 列と同様に入力

F14 に式を入力後、

F14 を選択して＋マークを下にドラッグ

（オートフィル機能を利用）




表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i*n_i$ | $\pi_i$ | $x_i*\pi_i$ |
| 14 | 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |
| | 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| | 18行目 | | =SUM(C14:C17) | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

# Excel を使った期待値の計算

p.10

- ●18 行の入力

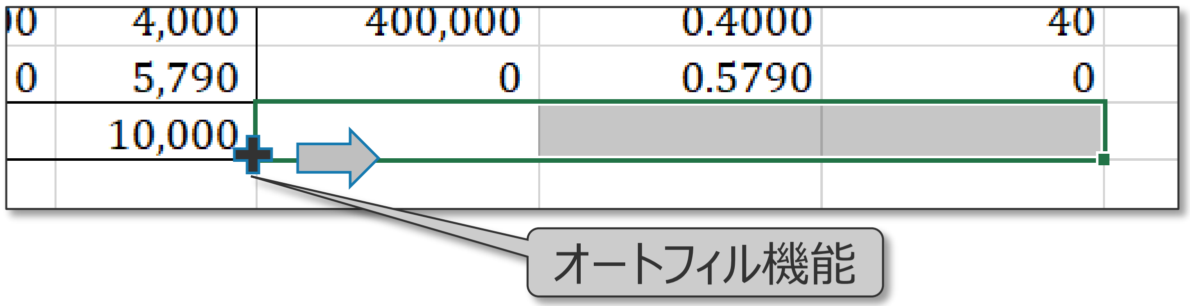
  18行：各列の合計

  D 列と同様に入力

  C18 は、既に式が入力済み。

  C18 を選択して＋マークを右にドラッグ

  （オートフィル機能を利用）




F18 に期待値 E[$x$] が表示

表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i * n_i$ | $\pi_i$ | $x_i * \pi_i$ |
| 14 | 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |
| | 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| | 18行目 | | =SUM(C14:C17) | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

# Excel を使った期待値の計算



## 実習

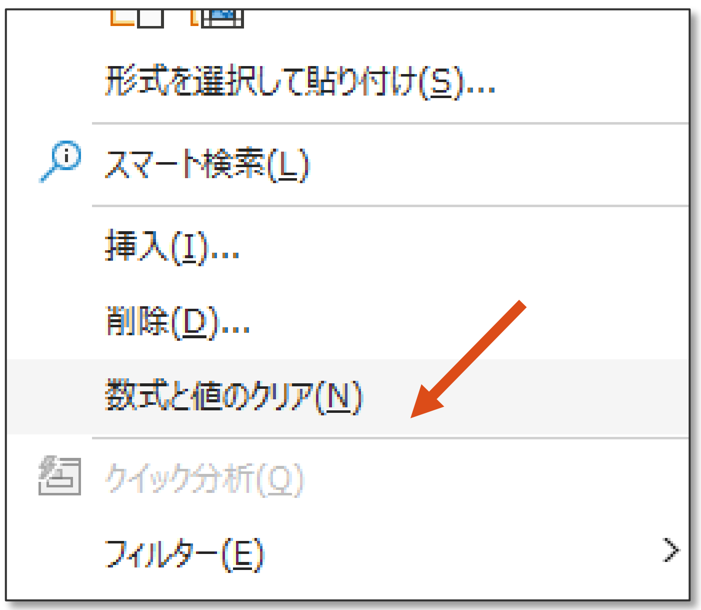
Excelファイルで、表示1.1.2 のセルの内容を一部消去

セル範囲を指定、右クリック、[数式と値のクリア]を選択

or [Del]キー

自分で数式を入力して再現

「数式と値のクリア」で
セルの内容を消去
自分で数式を入力して再現




表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i*n_i$ | $\pi_i$ | $x_i*\pi_i$ |
| 14 | 1等 | 10,000 | 10 | | | |
| 15 | 2等 | 1,000 | 200 | | | |
| 16 | 3等 | 100 | 4,000 | | | |
| 17 | はずれ | 0 | 5,790 | | | |
| 18 | 合計 | | | | | |

# 確率変数と確率分布



## ●確率変数と確率分布

変数 $x$（賞金）の取りうる個々の値 $x_i$ に対して確率 $\pi_i$ が対応づけられている

この対応関係を $x$ の「確率分布」という

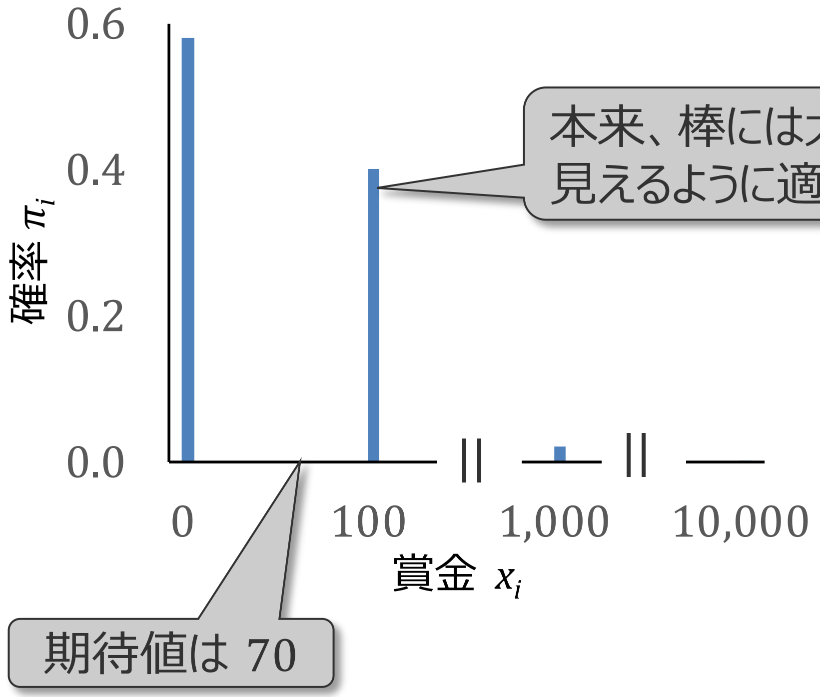
変数 $x$ を「確率変数」という

表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i * n_i$ | $\pi_i$ | $x_i * \pi_i$ |
| 14 | 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |
| | 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| | 18行目 | | =SUM(C14:C17) | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

# 確率変数と確率分布






変数（賞金）$x$ は、0, 100, 1000, 10000 の値を取る確率変数

それぞれの値を取る確率が決まっている

変数（賞金）$x$ の期待値は 70

表示1.1.2　期待値の計算(1)

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 12 | 等級 | 賞金 | 本数 | | 確率 | |
| 13 | $i$ | $x_i$ | $n_i$ | $x_i*n_i$ | $\pi_i$ | $x_i*\pi_i$ |
| 14 | 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 15 | 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 16 | 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| 17 | はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 18 | 合計 | | 10,000 | 700,000 | 1.0000 | 70 |
| | 14行目 | | | =B14*C14 | =C14/C$18 | =B14*E14 |
| | 18行目 | | =SUM(C14:C17) | =SUM(D14:D17) | =SUM(E14:E17) | =SUM(F14:F17) |

# （2）分散

# ばらつき



● 2 種類の宝くじのばらつきを比較

表示1.1.2
期待値の計算
（1）

| 級等 | 賞金 | 本数 | | 確率 | |
|---|---|---|---|---|---|
| $i$ | $x_i$ | $n_i$ | $x_i{}^*n_i$ | $\pi_i$ | $x_i{}^*\pi_i$ |
| 1等 | 10,000 | 10 | 100,000 | 0.0010 | 10 |
| 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| はずれ | 0 | 5,790 | 0 | 0.5790 | 0 |
| 合計 | | 10,000 | 700,000 | 1.0000 | 70 |

↓　1等の賞金が10倍の100,000円、残り9本が外れ

表示1.1.3
期待値の計算
（2）

| 等級 | 賞金 | 本数 | | 確率 | |
|---|---|---|---|---|---|
| $i$ | $x_i$ | $n_i$ | $x_i{}^*n_i$ | $\pi_i$ | $x_i{}^*\pi_i$ |
| 1等 | 100,000 | 1 | 100,000 | 0.0001 | 10 |
| 2等 | 1,000 | 200 | 200,000 | 0.0200 | 20 |
| 3等 | 100 | 4,000 | 400,000 | 0.4000 | 40 |
| はずれ | 0 | 5,799 | 0 | 0.5799 | 0 |
| 合計 | | 10,000 | 700,000 | 1.0000 | 70 |

| 賞金 | 本数 | |
|---|---|---|
| | 表示1.1.2 | 表示1.1.3 |
| 100,000 | | 1 |
| 10,000 | 10 | |
| 1,000 | 200 | 200 |
| 100 | 4,000 | 4,000 |
| 0 | 5,790 | 5,799 |
| 計 | 10,000 | 10,000 |

表示1.1.3の方が

当たり外れの差が大きい

→　ばらつきが大きい

これを定量的に表したい

# 偏差の期待値

p.10

●偏差

偏差：各値と期待値との差（期待値に対してどれだけ偏っているか）

$d_i = x_i - \mathrm{E}[x] \qquad d_1 = x_1 - \mathrm{E}[x] = 10{,}000 - 70 = 9930$

ばらつきが大きいと、当然、
偏差が大きくなることは想像される

●偏差の期待値

偏差の期待値は

それぞれの偏差とその確率の積の和

$$\mathrm{E}[d] = \sum_{i=1}^{m} (d_i \pi_i)$$

$$\mathrm{E}[x] = \sum_{i=1}^{m} (x_i \pi_i) \qquad (1.1.4)$$

表示1.1.4　偏差の期待値の計算

| 等級 $i$ | 賞金 $x_i$ | 本数 $n_i$ | 偏差 $d_i$ | 確率 $\pi_i$ | $d_i * \pi_i$ |
|---|---|---|---|---|---|
| 1等 | 10,000 | 10 | 9,930 | 0.0010 | 9.93 |
| 2等 | 1,000 | 200 | 930 | 0.0200 | 18.60 |
| 3等 | 100 | 4,000 | 30 | 0.4000 | 12.00 |
| はずれ | 0 | 5,790 | -70 | 0.5790 | -40.53 |
| 合計 | 0 | 10,000 | | 1.0000 | 0.00 |

確率 $\pi_i$ の計

偏差の期待値

# 偏差の期待値

p.10

## ●偏差

偏差：各値と期待値との差（期待値に対してどれだけ偏っているか）

$$d_i = x_i - \mathrm{E}[x] \qquad d_1 = x_1 - \mathrm{E}[x] = 10{,}000 - 70 = 9930$$

ばらつきが大きいと、当然、
偏差が大きくなることは想像される

## ●偏差の期待値

$$\begin{aligned}\mathrm{E}[d] &= \sum_{i=1}^{m}(d_i\pi_i) \\ &= \sum_{i=1}^{m}(x_i - \mathrm{E}[x])\pi_i \\ &= \sum_{i=1}^{m}(x_i\pi_i) - \mathrm{E}[x]\sum_{i=1}^{m}\pi_i \\ &= \mathrm{E}[x] - \mathrm{E}[x] \times 1 = 0\end{aligned}$$

表示1.1.4　偏差の期待値の計算

| 等級 $i$ | 賞金 $x_i$ | 本数 $n_i$ | 偏差 $d_i$ | 確率 $\pi_i$ | $d_i * \pi_i$ |
|---|---|---|---|---|---|
| 1等 | 10,000 | 10 | 9,930 | 0.0010 | 9.93 |
| 2等 | 1,000 | 200 | 930 | 0.0200 | 18.60 |
| 3等 | 100 | 4,000 | 30 | 0.4000 | 12.00 |
| はずれ | 0 | 5,790 | -70 | 0.5790 | -40.53 |
| 合計 | 0 | 10,000 | | 1.0000 | 0.00 |

確率 $\pi_i$ の計

偏差の期待値

# 分散

p.12

- 偏差の期待値

偏差の期待値では「ばつらき」の大きさを定量化できない

偏差には正負の符号があって、打消し合ってしまう

- 偏差の２乗の期待値：確率変数の分散

偏差の２乗の期待値　→　分散（Variance）

$$\mathrm{V}[x] = \mathrm{E}[d^2] = \sum_{i=1}^{m}(d_i^2\pi_i) \qquad (1.1.5)$$

表示 1.1.5 分散の計算（一部）

| | 級等 | 賞金 | 本数 | 偏差 | 確率 | | |
|---|---|---|---|---|---|---|---|
| | $i$ | $x_i$ | $n_i$ | $d_i$ | $\pi_i$ | $d_i * \pi_i$ | $d_i^2 * \pi_i$ |
| | 1等 | **10,000** | **10** | 9,930 | **0.0010** | 9.93 | **98,605** |
| (1) | 2等 | 1,000 | 200 | 930 | 0.0200 | 18.60 | 17,298 |
| | 3等 | 100 | 4,000 | 30 | 0.4000 | 12.00 | 360 |
| | はずれ | 0 | **5,790** | -70 | **0.5790** | -40.53 | **2,837** |
| | 合計 | 0 | 10,000 | | 1.0000 | 0.00 | **119,100** |

偏差の期待値

偏差の２乗の期待値

# 分散



## ●分散の比較

宝くじ (2) の分散は宝くじ (1) の分散の約８倍　　1,019,100 / 119,100 = 8.6

分散はばらつきの指標になる

表示 1.1.5 分散の計算

| | 級等 $i$ | 賞金 $x_i$ | 本数 $n_i$ | 偏差 $d_i$ | 確率 $\pi_i$ | $d_i * \pi_i$ | $d_i^2 * \pi_i$ |
|---|---|---|---|---|---|---|---|
| (1) | 1等 | **10,000** | **10** | 9,930 | **0.0010** | 9.93 | **98,605** |
| | 2等 | 1,000 | 200 | 930 | 0.0200 | 18.60 | 17,298 |
| | 3等 | 100 | 4,000 | 30 | 0.4000 | 12.00 | 360 |
| | はずれ | 0 | **5,790** | -70 | **0.5790** | -40.53 | **2,837** |
| | 合計 | 0 | 10,000 | | 1.0000 | 0.00 | **119,100** |
| (2) | 1等 | **100,000** | **1** | **99,930** | **0.0001** | 9.99 | **998,600** |
| | 2等 | 1,000 | 200 | 930 | 0.0200 | 18.60 | 17,298 |
| | 3等 | 100 | 4,000 | 30 | 0.4000 | 12.00 | 360 |
| | はずれ | 0 | **5,799** | -70 | **0.5799** | -40.59 | **2,842** |
| | 合計 | 0 | 10,000 | | 1.0000 | 0.00 | **1,019,100** |

p.11

# （3）標準偏差

# 標準偏差

p.12

## ●分散と標準偏差

分散（偏差の2乗の期待値）の単位は　円$^2$

元の単位に戻すために分散の平方根をとる　→　標準偏差（Standard Deviation）　D[$x$]

$$\mathrm{V}[x] = \mathrm{E}[d^2] = \sum_{i=1}^{m} \left(d_i^2 \pi_i\right) \qquad (1.1.5)$$

$$\mathrm{D}[x] = \sqrt{\mathrm{V}[x]} \qquad (1.1.6)$$

表示 1.1.2 (1)　$\mathrm{D}[x] = \sqrt{119100} =$ 　345 円

表示 1.1.3 (2)　$\mathrm{D}[x] = \sqrt{1019100} = 1010$ 円

約3倍の違いがある

# 標準偏差



● 期待値、分散、標準偏差

期待値（Expectation）

$$\mathrm{E}[x] = \sum_{i=1}^{m} x_i \pi_i \qquad \frac{n_i}{N} = \pi_i \qquad (1.1.4)$$

確率変数の期待値
（確率変数×確率）の和

分散（Variatnce）

$$\mathrm{V}[x] = \mathrm{E}[d^2] = \sum_{i=1}^{m} \left(d_i^2 \pi_i\right) \qquad d_i = x_i - \mathrm{E}[x] \qquad (1.1.5)$$

確率変数の偏差の２乗の期待値
（偏差の２乗×確率）の和

標準偏差（Standerd Deviation）

$$\mathrm{D}[x] = \sqrt{\mathrm{V}[x]} \qquad (1.1.6)$$

分散の平方根

# まとめ

p.12

- 宝くじの例から、期待値、分散、標準偏差を導いた

  これらは、統計の中で、極めて重要

- どのような確率分布であっても、

  確率変数 $x$ の期待値は確率変数 $x$ の分布における位置を示す指標

  確率変数 $x$ の分散、標準偏差は、分布における「ばらつき」の指標

- この節で理解が不十分でも、次節以降で理解が進む

  再度、この節を読み返すことにより理解が深まる

| | |
|---|---|
| ● 作成 | 片瀬雅彦 |
| ● 監修 | 松本一彦、長谷文雄 |
| ● 作成時期 | 2018年8月21日 |
| ● 改訂 | 2019年3月18日、2024年7月14日 |