確率と統計　第 5 回

確率変数に親しむ

5 月 18 日

Table of contents

データ行列の中の変数

データ行列の中の変数

ある消費者金融の顧客データが以下のようにデータ行列の形にまとめられているとする.

データ行列の中の変数

ある消費者金融の顧客データが以下のようにデータ行列の形にまとめられているとする.

変数

事例

	借入金額	返済期間	地域
1	200	10	A
2	100	50	C
3	100	20	E
⋮	⋮	⋮	⋮

データ行列の中の変数

データ行列の中の変数

このとき各変数を事例をひとつ選ぶごとに値が決まる変数と考えれば, 各変数は事例の分布に従って値が分布するランダムな変数とみなせる.

データ行列の中の変数

このとき各変数を事例をひとつ選ぶごとに値が決まる変数と考えれば, 各変数は事例の分布に従って値が分布するランダムな変数とみなせる.

このように, 確率的に変動する値をとる変数を確率変数という.

データ行列の中の変数

このとき各変数を事例をひとつ選ぶごとに値が決まる変数と考えれば, 各変数は事例の分布に従って値が分布するランダムな変数とみなせる.

このように, 確率的に変動する値をとる変数を確率変数という.
慣習的に確率変数を大文字 $X, Y, Z, \ldots$ で表し,
確率変数の値を小文字 $x, y, z, \ldots$ で表す.

期待値

期待値

ある授業では教科書 A と問題集 B を指定している．受講生のうち，20%はどちらも購入せず，55%は教科書のみ購入し，25%は両方を購入した．

期待値

ある授業では教科書 A と問題集 B を指定している．受講生のうち, 20%はどちらも購入せず, 55%は教科書のみ購入し, 25%は両方を購入した．

A : 137 ドル
B : 33 ドル.

	A	B	金額
20%	×	×	0
55%	○	×	137
25%	○	○	137 + 33.

期待値

このとき, 一人当たりの購入金額 X (ドル) の平均 (期待値と呼ぶ) はいくつか?

期待値

このとき, 一人当たりの購入金額 X (ドル) の平均 (期待値と呼ぶ) はいくつか?

X の値

$$x_1 = 0,\ x_2 = 137,\ x_3 = 170$$

の確率は

$$P(x_1) = 0.20,\ P(x_2) = 0.55,\ P(x_3) = 0.25$$

であるから,

期待値

このとき，一人当たりの購入金額 X (ドル) の平均 (期待値と呼ぶ) はいくつか?

X の値

$$x_1 = 0,\ \ x_2 = 137,\ \ x_3 = 170$$

の確率は

$$P(x_1) = 0.20,\ \ P(x_2) = 0.55,\ \ P(x_3) = 0.25$$

であるから，

$$(\text{期待値}) = 0 \times 0.20 + 137 \times 0.55 + 170 \times 0.25$$

ドルとなる．

期待値

この例のように確率変数 X のとる値の平均値を期待値と呼ぶ.

期待値

この例のように確率変数 X のとる値の平均値を期待値と呼ぶ.期待値は以下の式で定義される.

$$
\begin{aligned}
E(X) &= x_1 P(x_1) + x_2 P(x_2) + \cdots x_k P(x_k) \\
&= \sum_{i=1}^{k} x_i P(X = x_i).
\end{aligned}
$$

期待値

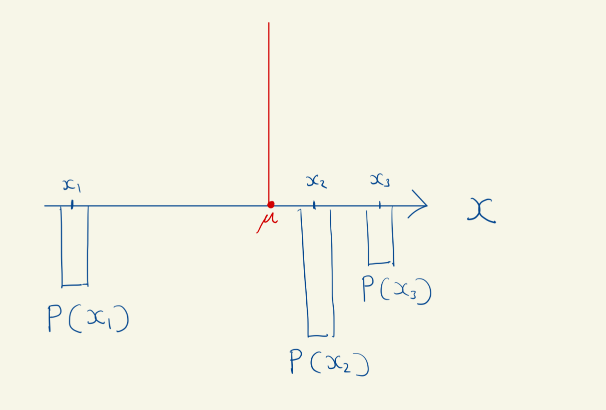
この例のように確率変数 X のとる値の平均値を期待値と呼ぶ.期待値は以下の式で定義される.

$$
\begin{aligned}
E(X) &= x_1 P(x_1) + x_2 P(x_2) + \cdots x_k P(x_k) \\
&= \sum_{i=1}^{k} x_i P(X = x_i).
\end{aligned}
$$

この期待値 $E(X)$ のことをギリシャ文字 μ (ミュー) と書くことも多い.

期待値

確率変数 X の期待値は重心に例えられる.



分散と標準偏差

分散と標準偏差

確率変数 X の値のばらつきを表す分散は以下の式で与えられる.

分散と標準偏差

確率変数 X の値のばらつきを表す分散は以下の式で与えられる.

$$\begin{aligned} V(X) &= (x_1 - \mu)^2 P(x_1) + \cdots + (x_k - \mu)^2 P(x_k) \\ &= \sum_{i=1}^{k} (x_i - \mu)^2 P(x_i). \end{aligned}$$

ここで μ は確率変数 X の期待値 $E(X)$ を表す.

分散と標準偏差

確率変数 X の値のばらつきを表す分散は以下の式で与えられる.

$$\begin{aligned} V(X) &= (x_1-\mu)^2 P(x_1) + \cdots + (x_k-\mu)^2 P(x_k) \\ &= \sum_{i=1}^{k} (x_i-\mu)^2 P(x_i). \end{aligned}$$

ここで μ は確率変数 X の期待値 $E(X)$ を表す.

分散のことをギリシャ文字 σ^2 (シグマの二乗) で表すことも多い. 標準偏差はそのルート $\sigma = \sqrt{\sigma^2}$ と定義される.

分散と標準偏差

先ほどの例での生徒一人当たりの購入金額 X の分散と標準偏差を求めると...

線形結合の期待値

線形結合の期待値

ジョンの通勤にかかる時間は家から駅までの区間, 駅から駅までの区間, 駅から会社までの区間でそれぞれ X_1, X_2, X_3 分かかる.

線形結合の期待値

ジョンの通勤にかかる時間は家から駅までの区間, 駅から駅までの区間, 駅から会社までの区間でそれぞれ X_1, X_2, X_3 分かかる.

各区間でかかる時間は日によって多少変動する確率変数となる.

線形結合の期待値

ジョンの通勤にかかる時間は家から駅までの区間, 駅から駅までの区間, 駅から会社までの区間でそれぞれ X_1, X_2, X_3 分かかる.

各区間でかかる時間は日によって多少変動する確率変数となる.

このとき合計の通勤時間 W の期待値は

$$E(W) = E(X_1) + E(X_2) + E(X_3)$$

と各区間での期待値の和で求まる.

線形結合の期待値

エレナは X ドルで洋服を仕入れて，Y ドルを売り上げる．このとき彼女が得る利益の期待値は

線形結合の期待値

エレナは X ドルで洋服を仕入れて，Y ドルを売り上げる．このとき彼女が得る利益の期待値は

$$E(Y-X)=E(Y)-E(X)$$

ドルとなる．

線形結合の期待値

以上の例のように, 確率変数の線形結合の期待値は各確率変数の期待値の線形結合になる:

線形結合の期待値

以上の例のように, 確率変数の線形結合の期待値は各確率変数の期待値の線形結合になる:

$$E(aX+bY)=aE(X)+bE(Y).$$

線形結合の期待値

以上の例のように, 確率変数の線形結合の期待値は各確率変数の期待値の線形結合になる:

$$E(aX + bY) = aE(X) + bE(Y).$$

この公式は 2 つの確率変数 X と Y が独立でなくても成り立つ.

線形結合の分散

線形結合の分散

レオナルドは A 社と B 社の株にそれぞれ 6000 ドルと 2000 ドルを投資した. A 社と B 社のひと月の株価の上昇率をそれぞれ X%と Y%とする. 確率変数 X と Y の期待値と分散は以下の通りとする:

線形結合の分散

レオナルドは A 社と B 社の株にそれぞれ 6000 ドルと 2000 ドルを投資した. A 社と B 社のひと月の株価の上昇率をそれぞれ X%と Y%とする. 確率変数 X と Y の期待値と分散は以下の通りとする:

	期待値	分散
X	3	4
Y	2	1

線形結合の分散

このとき, ひと月での儲け

$$W = 6000 \times 0.01X + 2000 \times 0.01Y$$

の期待値は

線形結合の分散

このとき, ひと月での儲け

$$W = 6000 \times 0.01X + 2000 \times 0.01Y$$

の期待値は

$$\begin{aligned} E(W) &= 60E(X) + 20E(Y) \\ &= 60 \times 3 + 20 \times 2 \\ &= 220 \end{aligned}$$

ドルとなる.

線形結合の分散

株価の上昇率 X と Y が独立と仮定する.

線形結合の分散

株価の上昇率 X と Y が独立と仮定する.

ひと月での儲け W の分散は以下の式で求めることができる.

線形結合の分散

株価の上昇率 X と Y が独立と仮定する.

ひと月での儲け W の分散は以下の式で求めることができる.

$$
\begin{aligned}
V(W) &= V(60X + 20Y) \\
&= 60^2 V(X) + 20^2 V(Y) \\
&= 60^2 \times 4 + 20^2 \times 1 \\
&= 14800.
\end{aligned}
$$

線形結合の分散

株価の上昇率 X と Y が独立と仮定する.

ひと月での儲け W の分散は以下の式で求めることができる.

$$\begin{aligned} V(W) &= V(60X + 20Y) \\ &= 60^2 V(X) + 20^2 V(Y) \\ &= 60^2 \times 4 + 20^2 \times 1 \\ &= 14800. \end{aligned}$$

よって標準偏差は $\sqrt{V(W)} = 122$ であり, 期待値の計算と合わせてひと月に 220 ± 122 ドルの収益が見込まれる.

線形結合の分散

確率変数 X と Y が独立であると仮定する.

線形結合の分散

確率変数 X と Y が独立であると仮定する.

線形結合 $aX + bY$ の分散について次が成り立つ.

線形結合の分散

確率変数 X と Y が独立であると仮定する.

線形結合 $aX + bY$ の分散について次が成り立つ.

$$V(aX + bY) = a^2 V(X) + b^2 V(Y).$$

線形結合の分散

確率変数 X と Y が独立であると仮定する.

線形結合 $aX + bY$ の分散について次が成り立つ.

$$V(aX + bY) = a^2 V(X) + b^2 V(Y).$$

よって線形結合の標準偏差は $\sqrt{a^2 V(X) + b^2 V(Y)}$ となる.