

統計学I

早稲田大学政治経済学術院

西郷浩

本日の目標

- 多次元の確率分布
 - 条件つき確率分布、同時確率分布、周辺確率分布
- 確率変数の独立性
- 多次元の確率変数の期待値
- 条件つき期待値と条件付き分散
- 確率変数の和の期待値と分散
 - 共分散

条件つき確率分布

- 実験A

- サイコロを1つ投げる。
- その結果に応じて、2回目に投げるサイコロを変える。
 - 結果が奇数 $\rightarrow [1, 1, 3, 3, 5, 5]$
 - 結果が偶数 $\rightarrow [2, 2, 4, 4, 6, 6]$
- 2つの確率変数
 - X = 最初のサイコロの出目
 - Y = 2回目のサイコロの出目
- $X = 1$ のときの Y の条件つき確率分布:
 - $P(Y = 1|X = 1) = \frac{1}{3}; P(Y = 2|X = 1) = 0; P(Y = 3|X = 1) = \frac{1}{3};$
 $P(Y = 4|X = 1) = 0; P(Y = 5|X = 1) = \frac{1}{3}; P(Y = 6|X = 1) = 0.$
- X を条件とした Y の条件つき確率関数:
 - $p_{Y|X}(y|x) = P(Y = y|X = x)$

同時確率分布(1)

- 同時確率

- $P(X = x, Y = y) = P(X = x)P(Y = y|X = x)$

- 例: 実験Aにおいて、 $P(X = 1, Y = 1)$ を求める

- $P(X = 1, Y = 1) = P(X = 1)P(Y = 1|X = 1) = \frac{1}{6} \times \frac{1}{3} = \frac{1}{18}$

- 同時確率関数

- $p_{X,Y}(x, y) = P(X = x, Y = y)$

- $= P(X = x)P(Y = y|X = x) = p_X(x)p_{Y|X}(y|x)$

同時確率分布(2)

表1: 実験Aに関するXとYの同時分布

x(左); y	1	2	3	4	5	6	合計
1	1/18	0	1/18	0	1/18	0	1/6
2	0	1/18	0	1/18	0	1/18	1/6
3	1/18	0	1/18	0	1/18	0	1/6
4	0	1/18	0	1/18	0	1/18	1/6
5	1/18	0	1/18	0	1/18	0	1/6
6	0	1/18	0	1/18	0	1/18	1/6
合計	1/6	1/6	1/6	1/6	1/6	1/6	1

周辺確率分布

- 周辺確率

- $P(Y = y) = P(X = 1, Y = y) + P(X = 2, Y = y) + \cdots + P(X = 6, Y = y).$

- 例: 実験Aにおける $Y = 1$ の周辺確率

- $P(Y = 1) = \frac{1}{18} + 0 + \frac{1}{18} + 0 + \frac{1}{18} + 0 = \frac{1}{6}.$

- $P(X = x)$ を周辺確率とみることもできる。

- $P(X = x) = P(X = x, Y = 1) + P(X = x, Y = 2) + \cdots + P(X = x, Y = 6).$

- 表1の周辺部分の分布が Y (または X) の周辺確率分布

- 周辺確率関数:

- $p_Y(y) = P(Y = y) = \sum_x P(X = x, Y = y) = \sum_x p_{X,Y}(x, y).$

- $p_X(x) = P(X = x) = \sum_y P(X = x, Y = y) = \sum_y p_{X,Y}(x, y).$

同時確率分布と条件つき確率分布(1)

- X を条件としたときの Y の条件つき確率

$$- P(Y = y|X = x) = \frac{P(X=x, Y=y)}{P(X=x)}$$

- X を条件としたときの Y の条件つき確率関数

$$- p_{Y|X}(y|x) = \frac{p_{X,Y}(x,y)}{p_X(x)}$$

- x の値を固定して y の値を変化させると、 Y の条件つき確率をあたえる。

- 例: つぎのスライド

同時確率分布と条件つき確率分布(2)

表2: 実験Aに関する X を条件としたときの Y の条件つき分布

$x(\text{左}); y$	1	2	3	4	5	6	合計
1	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1(1/6)$
2	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	$1(1/6)$
3	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1(1/6)$
4	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	$1(1/6)$
5	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1(1/6)$
6	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	$1(1/6)$
合計	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	1

注: () 内は同時確率を示す。

同時確率分布と条件つき確率分布(3)

- Y を条件としたときの X の条件つき分布

$$- p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

- y の値を固定して x の値を変化させると、 X の条件つき確率をあたえる。

同時確率分布と条件つき確率分布(4)

表3: 実験Aに関する Y を条件としたときの X の条件つき分布

$x(\text{左}); y$	1	2	3	4	5	6	合計
1	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/6$
2	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	$1/6$
3	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/6$
4	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	$1/6$
5	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/6$
6	0	$1/3(1/18)$	0	$1/3(1/18)$	0	$1/3(1/18)$	$1/6$
合計	$1(1/6)$	$1(1/6)$	$1(1/6)$	$1(1/6)$	$1(1/6)$	$1(1/6)$	1

注: () 内は同時確率を示す。

確率変数の独立性(1)

- 実験B
 - サイコロを1つ投げる。
 - その結果が何であれ、2回目に普通のサイコロを投げる。
 - 同時確率分布
 - 次のスライド
 - 条件つき確率分布
 - その次の2つのスライド

確率変数の独立性(2)

表4: 実験Bに関する X と Y の同時分布

$x(\text{左}); y$	1	2	3	4	5	6	合計
1	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
2	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
3	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
4	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
5	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
6	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
合計	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	1

確率変数の独立性(3)

表5: 実験Bに関する X を条件とした Y の条件つき分布

$x(\text{左}); y$	1	2	3	4	5	6	合計
1	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1(1/6)$
2	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1(1/6)$
3	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1(1/6)$
4	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1(1/6)$
5	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1(1/6)$
6	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1(1/6)$
合計	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	1

注: () 内は同時確率を示す。

確率変数の独立性(4)

表6: 実験Bに関する Y を条件とした X の条件つき分布

$x(\text{左}); y$	1	2	3	4	5	6	合計
1	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6$
2	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6$
3	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6$
4	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6$
5	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6$
6	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6(1/36)$	$1/6$
合計	$1(1/6)$	$1(1/6)$	$1(1/6)$	$1(1/6)$	$1(1/6)$	$1(1/6)$	1

注: () 内は同時確率を示す。

確率変数の独立性(5)

- 実験Bの特徴
 - すべての x, y について $p_{X,Y}(x, y) = p_X(x)p_Y(y)$
 - 表4参照
 - 同時確率分布が、周辺確率分布の積で求められる。
 - すべての x, y について $p_{Y|X}(y|x) = p_Y(y)$
 - 表5参照
 - X の実現値は Y の出方(確率分布)に影響を及ぼさない。
 - すべての x, y について $p_{X|Y}(x|y) = p_X(x)$
 - 表6参照
 - Y の実現値は X の出方(確率分布)に影響を及ぼさない。
 - 上の3つの条件のうち1つが成り立てば、他の2つは自動的に成り立つ。
- 3つの条件のうち1つが成り立てば、「 X と Y とが独立である」という。

確率変数の独立性(6)

- 実験Aの特徴
 - 独立性の条件が成り立たない。
 - 例: $p_{Y|X}(1|1) = \frac{1}{3}, p_{Y|X}(1|2) = 0$
 - X の実現値に応じて Y の出方が変わる。
 - 独立性の条件が成り立たないとき、「 X と Y とが独立でない(従属している)」という。
- 独立性の直観的な意味
 - X と Y とが独立である。
 - $\Leftrightarrow X$ の発生と Y の発生とが無関係である。
 - $\Leftrightarrow X$ の値がわかって、 Y についての情報にならない。
 - X と Y とが独立でない。
 - $\Leftrightarrow X$ の発生と Y の発生とが関係ある。
 - $\Leftrightarrow X$ の値がわかると、 Y についての情報になる。

多次元の確率変数の期待値(1)

- 表7: (X, Y) の同時確率分布

x(表側)y	1	2	4
1	1/12	2/12	1/12
2	2/12	1/12	1/12
3	1/12	1/12	2/12

- $Z = XY$ の期待値を計算する。

– Z の確率分布にもとづく計算方法

$$\begin{aligned} E(Z) &= \sum_z z p_Z(z) = 1 \times \frac{1}{12} + 2 \times \frac{4}{12} + 3 \times \frac{1}{12} + 4 \times \frac{2}{12} + 6 \times \frac{1}{12} + 8 \times \frac{1}{12} + 12 \times \frac{2}{12} \\ &= \frac{58}{12} \end{aligned}$$

– (X, Y) の同時分布にもとづく計算方法

$$\begin{aligned} E(XY) &= \sum_x \sum_y x y p_{X,Y}(x, y) = 1 \times 1 \times \frac{1}{12} + 1 \times 2 \times \frac{2}{12} + 1 \times 4 \times \frac{1}{12} + 2 \times 1 \times \frac{2}{12} \\ &\quad + 2 \times 2 \times \frac{1}{12} + 2 \times 4 \times \frac{1}{12} + 3 \times 1 \times \frac{1}{12} + 3 \times 2 \times \frac{1}{12} + 3 \times 4 \times \frac{2}{12} = \frac{58}{12} \end{aligned}$$

多次元の確率変数の期待値(2)

- 多次元の確率変数の関数の期待値
 - $E(g(X, Y)) = \sum_x \sum_y g(x, y) p_{X,Y}(x, y)$

条件つき期待値(1)

- 表8: X を条件としたときの Y の条件つき確率分布

x (表側) y	1	2	4	合計
1	$1/4(1/12)$	$2/4(2/12)$	$1/4(1/12)$	$1(4/12)$
2	$2/4(2/12)$	$1/4(1/12)$	$1/4(1/12)$	$1(4/12)$
3	$1/4(1/12)$	$1/4(1/12)$	$2/4(2/12)$	$1(4/12)$

- 条件つき期待値

- $X=1$ のときの Y の条件つき期待値

- $E_{Y|X}(Y|1) = \frac{1}{4} \times 1 + \frac{2}{4} \times 2 + \frac{1}{4} \times 4 = \frac{9}{4}$

- X の他の値に対応するの Y の条件つき期待値

- $E_{Y|X}(Y|2) = \frac{2}{4} \times 1 + \frac{1}{4} \times 2 + \frac{1}{4} \times 4 = \frac{8}{4}$

- $E_{Y|X}(Y|3) = \frac{1}{4} \times 1 + \frac{1}{4} \times 2 + \frac{2}{4} \times 4 = \frac{11}{4}$

条件つき期待値(2)

- 条件とした確率変数の変化
 - $X = 1$ (確率 $4/12$ で発生) のときの Y の条件つき期待値: $9/4$
 - $X = 2$ (確率 $4/12$ で発生) のときの Y の条件つき期待値: $8/4$
 - $X = 3$ (確率 $4/12$ で発生) のときの Y の条件つき期待値: $11/4$
- 条件つき期待値そのものが、条件とした変数の変化に応じて変化する確率変数とみなせる。
 - $E_{Y|X}(Y|X)$ と記して、「 Y の条件つき期待値が X の値に応じて変化する確率変数である」とみなす。
 - このとき、以下の式が成立する。
 - $E(Y) = E_X\{E_{Y|X}(Y|X)\}$

条件つき分散(1)

- 表9: X を条件としたときの Y の条件つき確率分布と期待値

x (表側) y	1	2	4	合計	$E_{Y X}(Y x)$
1	$1/4(1/12)$	$2/4(2/12)$	$1/4(1/12)$	$1(4/12)$	$9/4$
2	$2/4(2/12)$	$1/4(1/12)$	$1/4(1/12)$	$1(4/12)$	$8/4$
3	$1/4(1/12)$	$1/4(1/12)$	$2/4(2/12)$	$1(4/12)$	$11/4$

- 条件つき分散
 - $X = 1$ のときの Y の条件つき分散

$$\bullet V_{Y|X}(Y|1) = \frac{1}{4} \times \left(1 - \frac{9}{4}\right)^2 + \frac{2}{4} \times \left(2 - \frac{9}{4}\right)^2 + \frac{1}{4} \times \left(4 - \frac{9}{4}\right)^2 = \frac{19}{16}$$

- X の他の値に対応するの Y の条件つき分散

$$\bullet V_{Y|X}(Y|2) = \frac{2}{4} \times \left(1 - \frac{8}{4}\right)^2 + \frac{1}{4} \times \left(2 - \frac{8}{4}\right)^2 + \frac{1}{4} \times \left(4 - \frac{8}{4}\right)^2 = \frac{24}{16}$$

$$\bullet V_{Y|X}(Y|3) = \frac{1}{4} \times \left(1 - \frac{11}{4}\right)^2 + \frac{1}{4} \times \left(2 - \frac{11}{4}\right)^2 + \frac{2}{4} \times \left(4 - \frac{11}{4}\right)^2 = \frac{27}{16}$$

条件つき分散(2)

- 条件とした確率変数の変化
 - $X = 1$ (確率 $4/12$ で発生) のときの Y の条件つき分散: $19/16$
 - $X = 2$ (確率 $4/12$ で発生) のときの Y の条件つき分散: $24/16$
 - $X = 3$ (確率 $4/12$ で発生) のときの Y の条件つき分散: $27/16$
- 条件つき分散そのものが、条件とした変数の変化に応じて変化する確率変数とみなせる。
 - $V_{Y|X}(Y|X)$ と記して、「 Y の条件つき分散が X の値に応じて変化する確率変数である」とみなす。
 - このとき、以下の式が成立する。
 - $V(Y) = E_X\{V_{Y|X}(Y|X)\} + V_X\{E_{Y|X}(Y|X)\}$

条件つき期待値・分散の性質(1)

- 表10: 条件つき期待値・分散と通常の期待値・分散

x(表側)y	1	2	4	合計	$E_{Y X}(Y x)$	$V_{Y X}(Y x)$
1	1/4(1/12)	2/4(2/12)	1/4(1/12)	1(4/12)	9/4	19/16
2	2/4(2/12)	1/4(1/12)	1/4(1/12)	1(4/12)	8/4	6/4
3	1/4(1/12)	1/4(1/12)	2/4(2/12)	1(4/12)	11/4	27/16
合計	(1/3)	(1/3)	(1/3)	(1)	-	-

$$- E(Y) = 1 \times \frac{1}{3} + 2 \times \frac{1}{3} + 4 \times \frac{1}{3} = \frac{7}{3}$$

$$- E_X\{E_{Y|X}(Y|X)\} = \frac{9}{4} \times \frac{4}{12} + \frac{8}{4} \times \frac{4}{12} + \frac{11}{4} \times \frac{4}{12} = \frac{7}{3}$$

条件つき期待値・分散の性質(2)

- 表11: 条件つき期待値・分散と通常の期待値・分散

x(表側)y	1	2	4	合計	$E_{Y X}(Y x)$	$V_{Y X}(Y x)$
1	1/4(1/12)	2/4(2/12)	1/4(1/12)	1(4/12)	9/4	19/16
2	2/4(2/12)	1/4(1/12)	1/4(1/12)	1(4/12)	8/4	6/4
3	1/4(1/12)	1/4(1/12)	2/4(2/12)	1(4/12)	11/4	27/16
合計	(1/3)	(1/3)	(1/3)	(1)	-	-

$$- V(Y) = \frac{1}{3} \times \left(1 - \frac{7}{3}\right)^2 + \frac{1}{3} \times \left(2 - \frac{7}{3}\right)^2 + \frac{1}{3} \times \left(4 - \frac{7}{3}\right)^2 = \frac{14}{9}$$

$$- V_X\{E_{Y|X}(Y|X)\} = \frac{4}{12} \times \left(\frac{9}{4} - \frac{7}{3}\right)^2 + \frac{4}{12} \times \left(\frac{8}{4} - \frac{7}{3}\right)^2 + \frac{4}{12} \times \left(\frac{11}{4} - \frac{7}{3}\right)^2 = \frac{7}{72}$$

$$- E_X\{V_{Y|X}(Y|X)\} = \frac{4}{12} \times \frac{19}{16} + \frac{4}{12} \times \frac{24}{16} + \frac{4}{12} \times \frac{27}{16} = \frac{35}{24}$$

確率変数の和の期待値と分散(1)

- 確率変数の和の期待値

- $E(X + Y) = E_X(X) + E_Y(Y) = \mu_X + \mu_Y$

- 確率変数の和の分散

- $V(X + Y) = V_X(X) + V_Y(Y) + 2Cov(X, Y)$

- 共分散(X と Y との関係の指標)

- $Cov(X, Y) = E\{(X - \mu_X)(Y - \mu_Y)\} = \sum_x \sum_y (x - \mu_X)(y - \mu_Y)p_{X,Y}(x, y)$

- 相関係数(X と Y との関係の指標)

- $\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{V_X(X) \times V_Y(Y)}}$

- X と Y とが独立であるとき:

- $Cov(X, Y) = 0$
 - $V(X + Y) = V_X(X) + V_Y(Y)$
 - $\rho_{X,Y} = 0$

確率変数の和の期待値と分散(2)

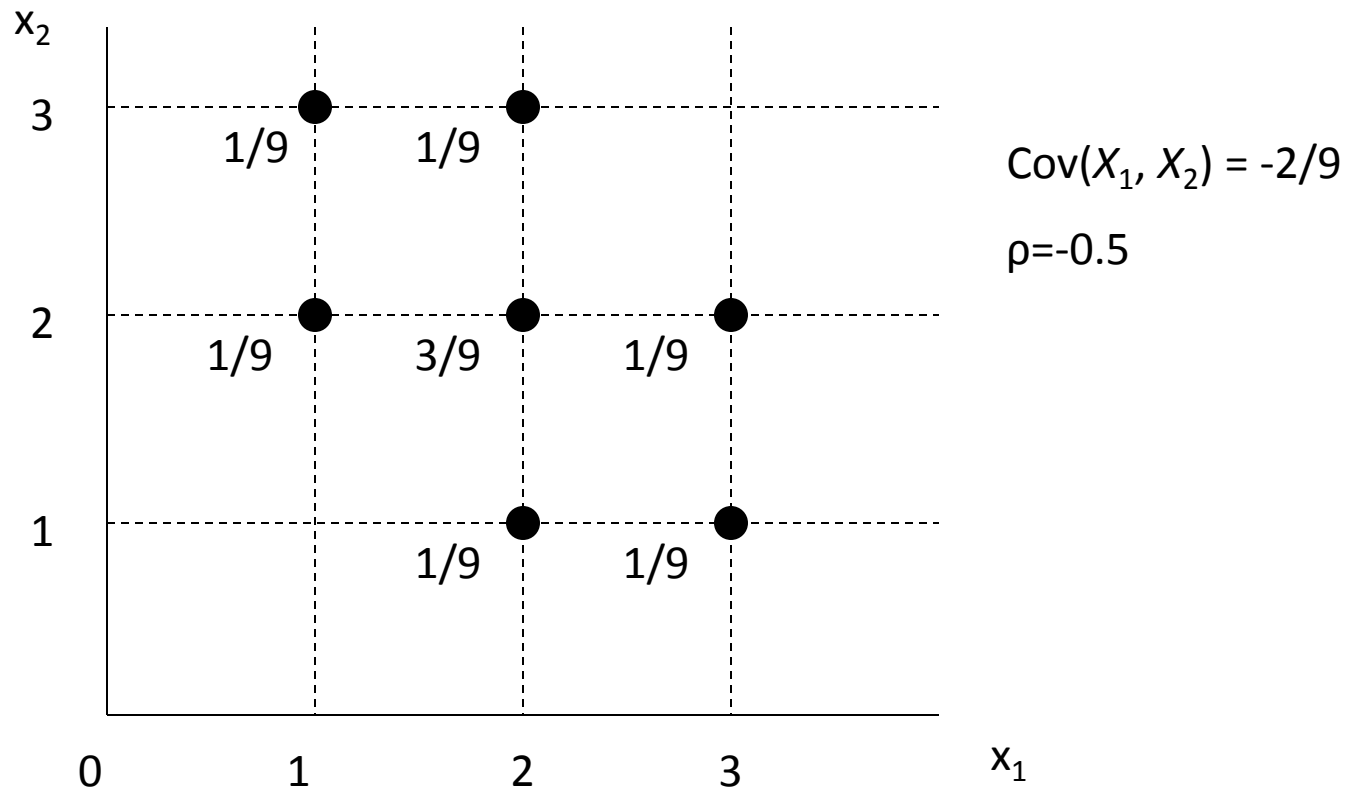
- 応用

- 2項分布の期待値と分散を求める。

- $X \sim \text{Binomial}(n, p)$
 - $X = X_1 + X_2 + \cdots + X_n$, ただし、 $X_i \sim_{iid} \text{Bernoulli}(p)$
 - $E(X) = E(X_1) + E(X_2) + \cdots + E(X_n) = np$
 - $V(X) = V(X_1) + V(X_2) + \cdots + V(X_n) = np(1 - p)$

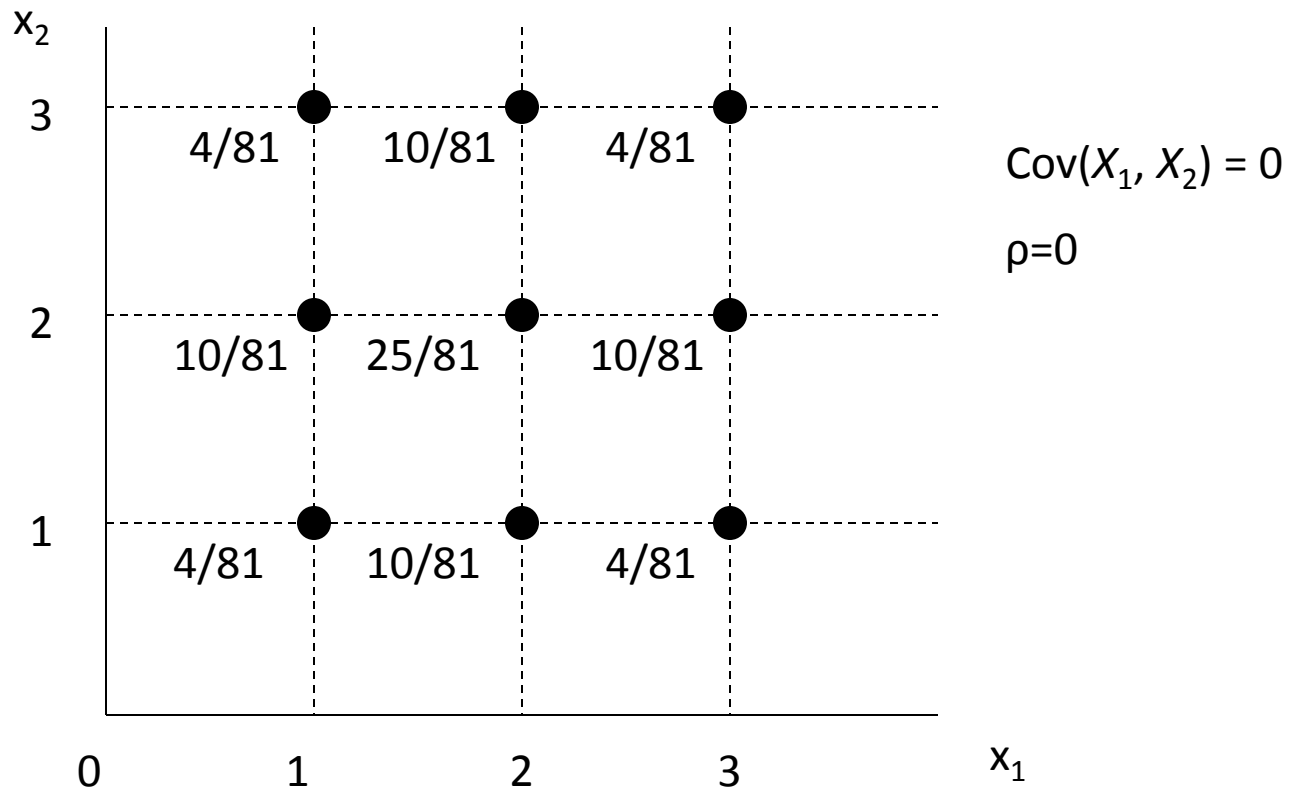
共分散の符号(1)

図1: ケース1(負の相関)



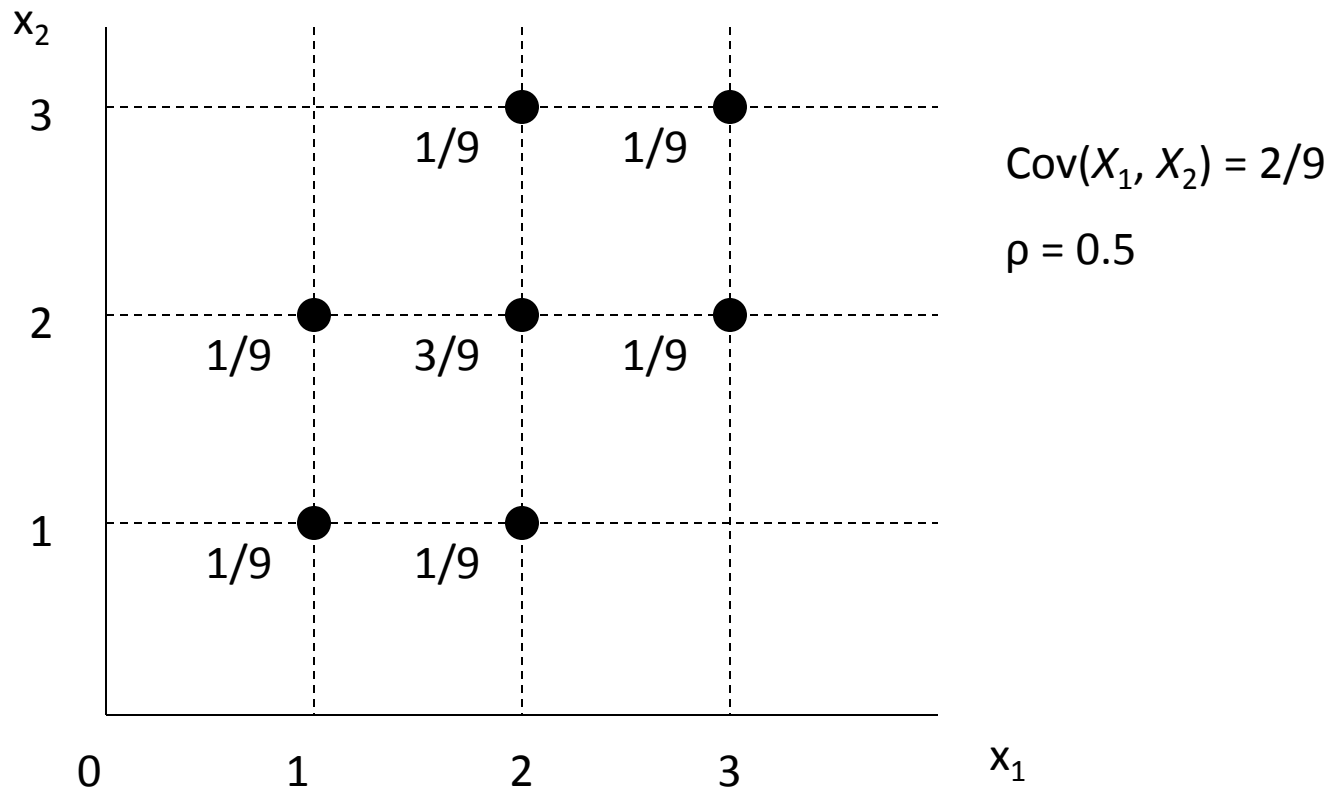
共分散の符号(2)

図2: ケース2 (独立)



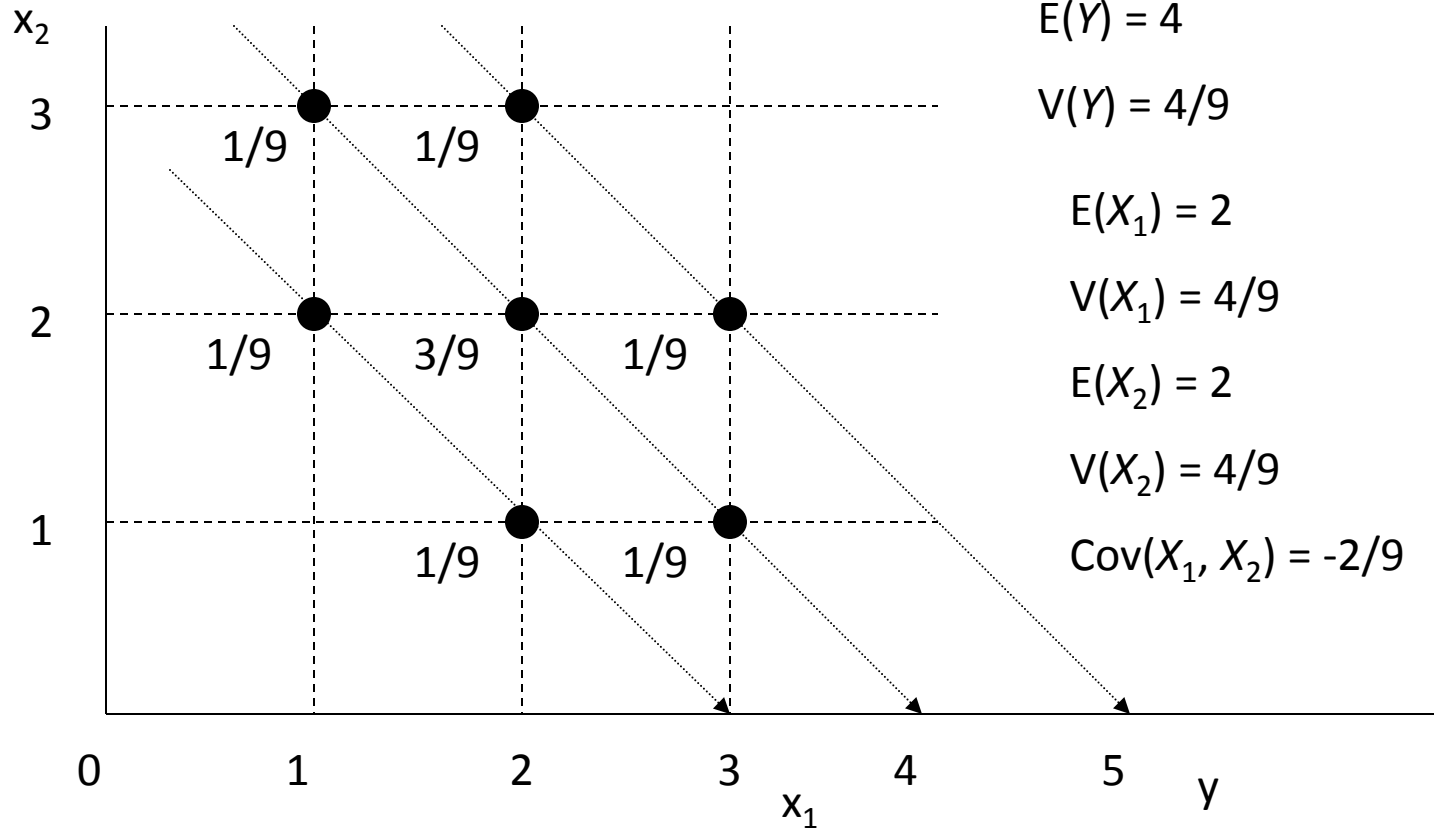
共分散の符号(3)

図3: ケース3(正の相関)



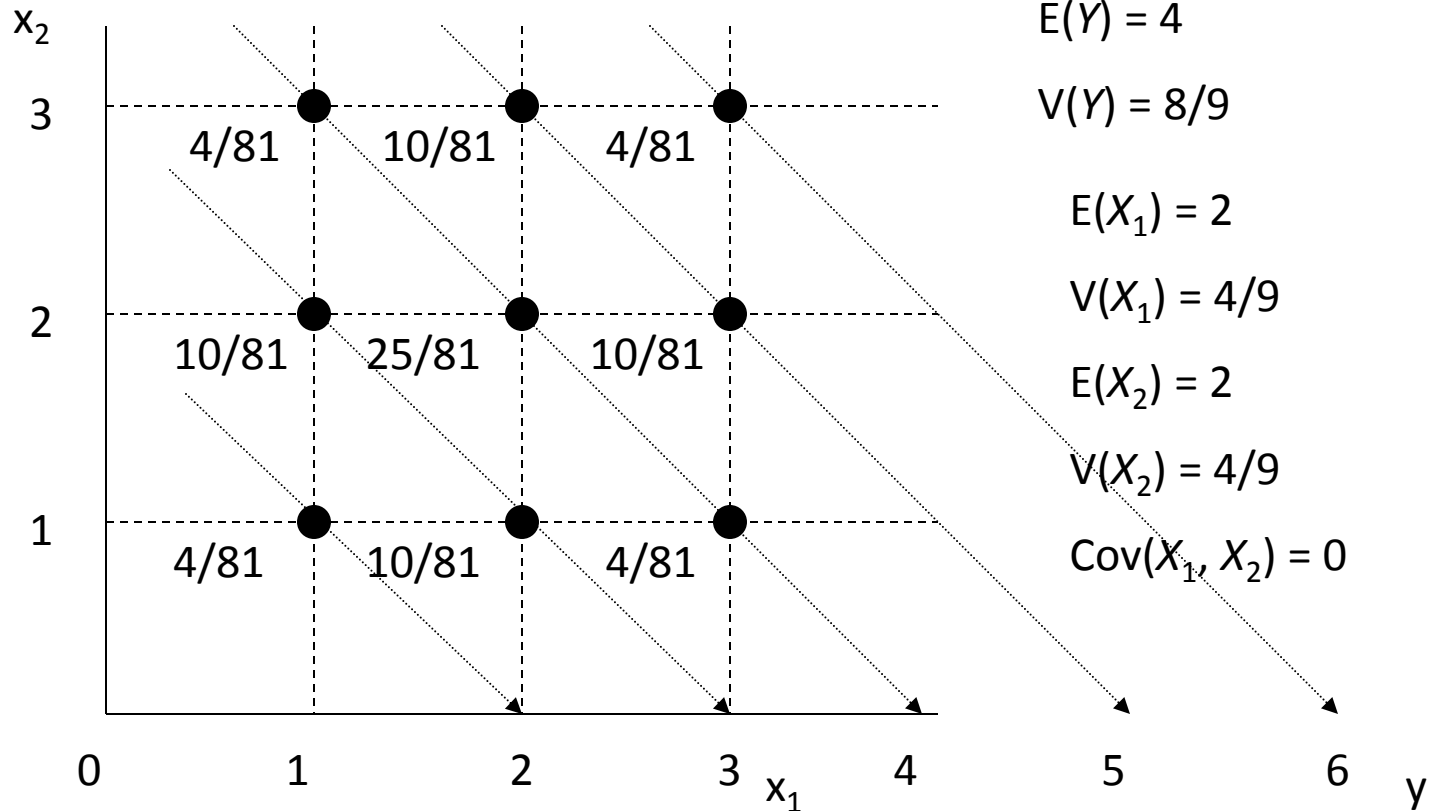
和の分散(1)

図1: ケース1(負の相関)



和の分散(2)

図2: ケース2 (独立)



$$Y = X_1 + X_2$$

$$E(Y) = 4$$

$$V(Y) = 8/9$$

$$E(X_1) = 2$$

$$V(X_1) = 4/9$$

$$E(X_2) = 2$$

$$V(X_2) = 4/9$$

$$\text{Cov}(X_1, X_2) = 0$$

和の分散(3)

図3: ケース3(正の相関)

