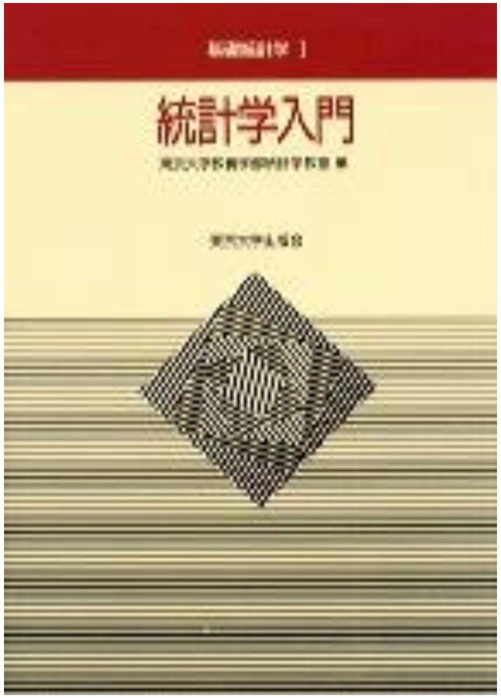
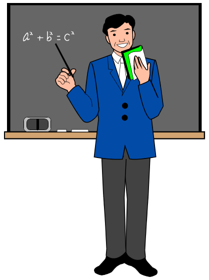
確率と統計（1）：
基礎概念

杉山将・本多淳也
sugi@k.u-tokyo.ac.jp, jhonda@k.u-tokyo.ac.jp
http://www.ms.k.u-tokyo.ac.jp

参考書

- 東京大学教養学部統計学教室（編），統計学入門，東京大学出版会，1991年
- 杉山将，機械学習のための確率と統計，講談社，2015年





講義の流れ

1. 確率変数と確率分布
2. 確率分布の性質を表す指標
3. 同時確率
4. 条件付き確率

確率変数と確率分布

- 確率(probability):事象の起こりやすさを定量的に示す0から1の値
- 事象Aの起こる確率を $P(A)$ で表す
- 確率変数(random variable): その変数に関する任意の事象に対して確率が定義されている変数
- 確率分布(probability distribution):確率変数に関する各事象に対してその確率を与える関数
- 確率変数は大文字で, 実現値は小文字で表わすことが多い

  $X = x$:「確率変数 X が x という実現値をとった」

離散型の確率変数と確率関数

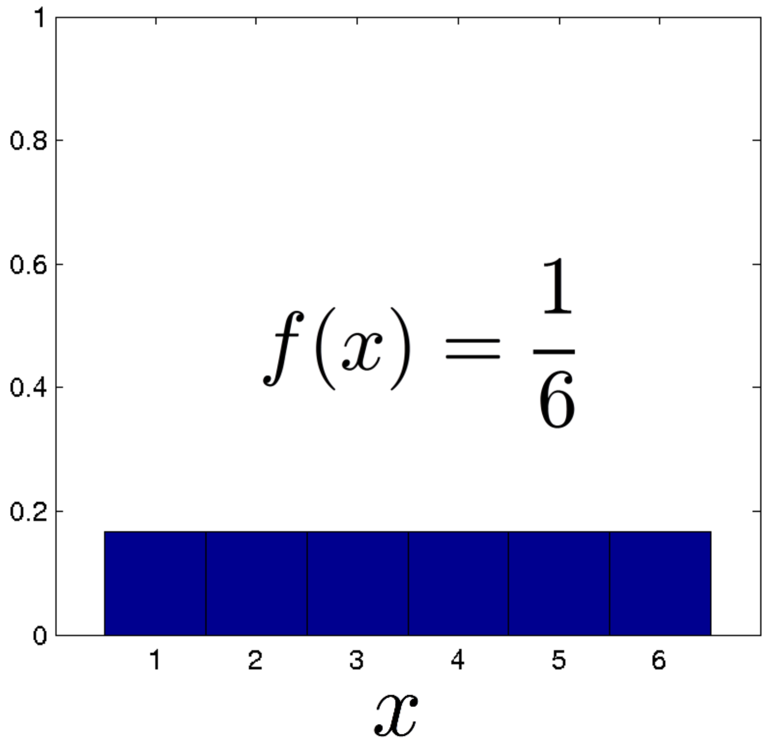
- 離散型(discrete type)確率変数: 可算集合の中の値をとる確率変数
- 離散型の確率変数の確率分布は確率質量関数によって一意に定まる

$$P(X = x) = f(x)$$

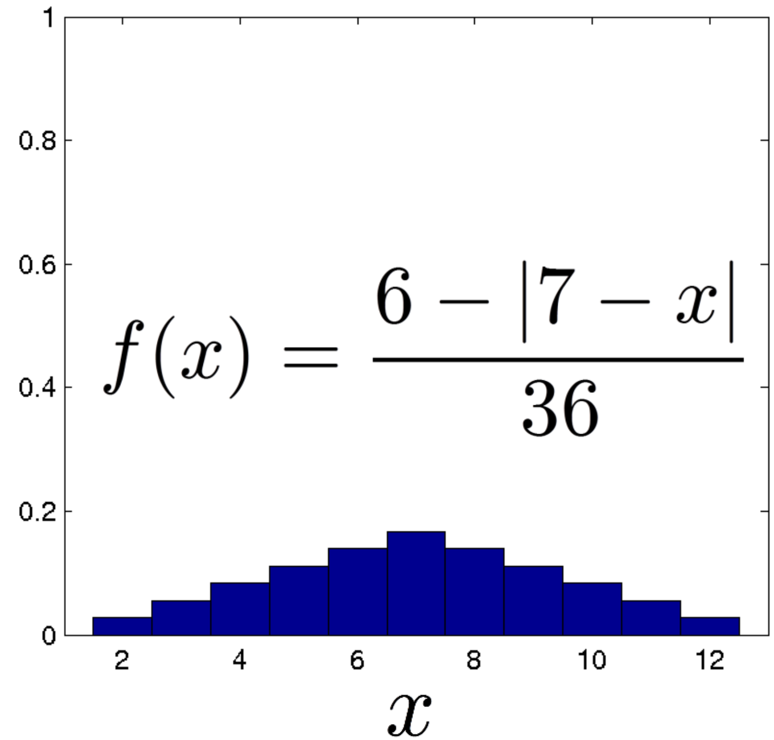
$f(x)$:確率質量関数(probability mass function)

$$f(x) \geq 0, \qquad \sum_{x} f(x) = 1$$

離散型の確率分布の例



さいころの出る目



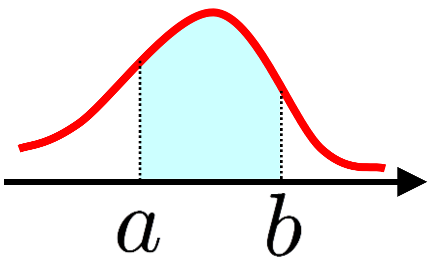
二つのさいころの出る目の和

- 名前がよく知られている離散型確率分布：
  - (離散)一様分布，二項分布，超幾何分布，ポアソン分布，負の二項分布，幾何分布など

連続型の確率変数と確率密度関数

- 連続型(continuous type)確率変数: 連続値をとる確率変数
- 連続型の確率変数の確率分布は確率密度関数で一意に定まる

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$



$f(x)$:確率密度関数(probability density function)

$$f(x) \geq 0, \quad \int f(x)dx = 1$$

- 注意: 連続型の確率変数がある値 a をとる確率はゼロ！

$$P(X = a) = \int_a^a f(x)dx = 0$$

累積分布関数

- 確率変数が x 以下の値をとる確率

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du \ \text{ or } \ \sum_{i:i\leq x} f(i)$$

$F(x)$:累積分布関数(cumulative distribution function)

- 連続型の場合は $F'(x) = f(x)$
- 広義単調増加:

$$x_1 < x_2 \quad \Longrightarrow \quad F(x_1) \leq F(x_2)$$

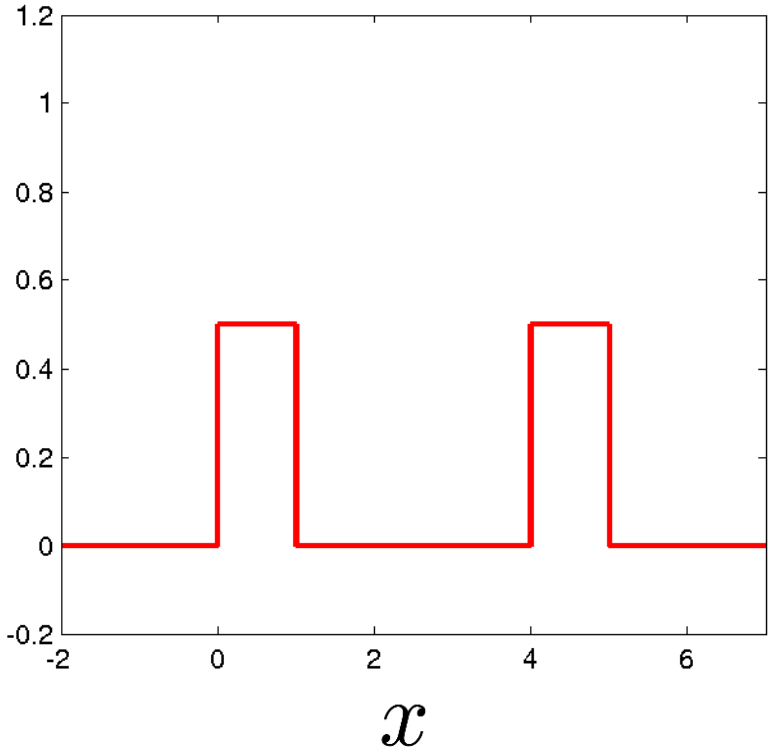
- 範囲:

$$x \to -\infty \quad \Longrightarrow \quad F(x) \to 0$$

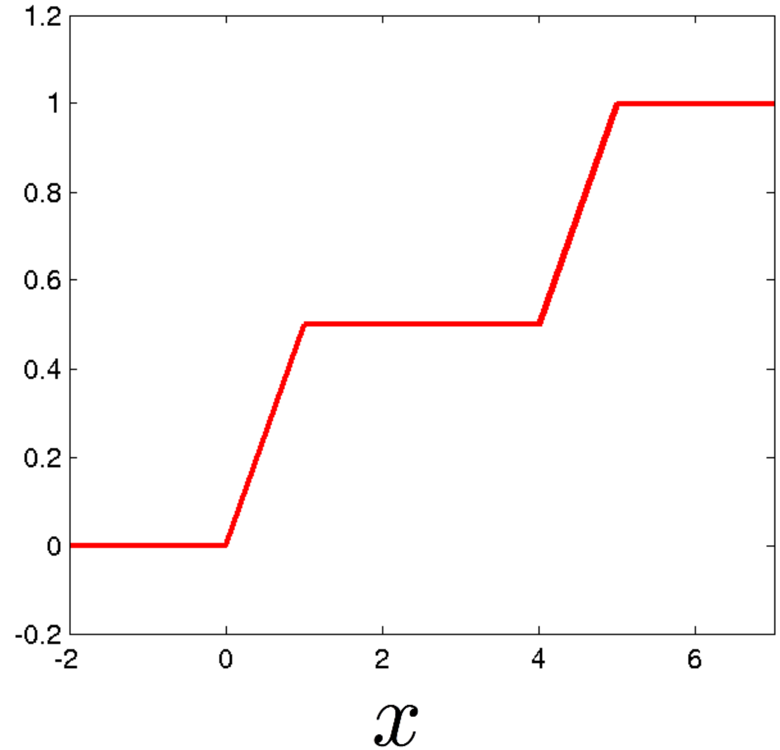
$$x \to \infty \quad \Longrightarrow \quad F(x) \to 1$$

- 任意の(1次元)確率分布は累積分布関数で表現可能

主な連続型の確率分布



確率密度関数

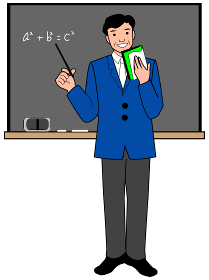


累積分布関数

- 名前がよく知られている連続型確率分布:
  - (連続)一様分布，正規分布，指数分布，カイ二乗分布，ガンマ分布，ベータ分布，コーシー分布，t分布など

確率変数と確率分布：まとめ

- 確率変数と確率分布
- 離散型の確率変数
  - 確率質量関数
- 連続型の確率変数
  - 確率密度関数
- 累積分布関数



講義の流れ

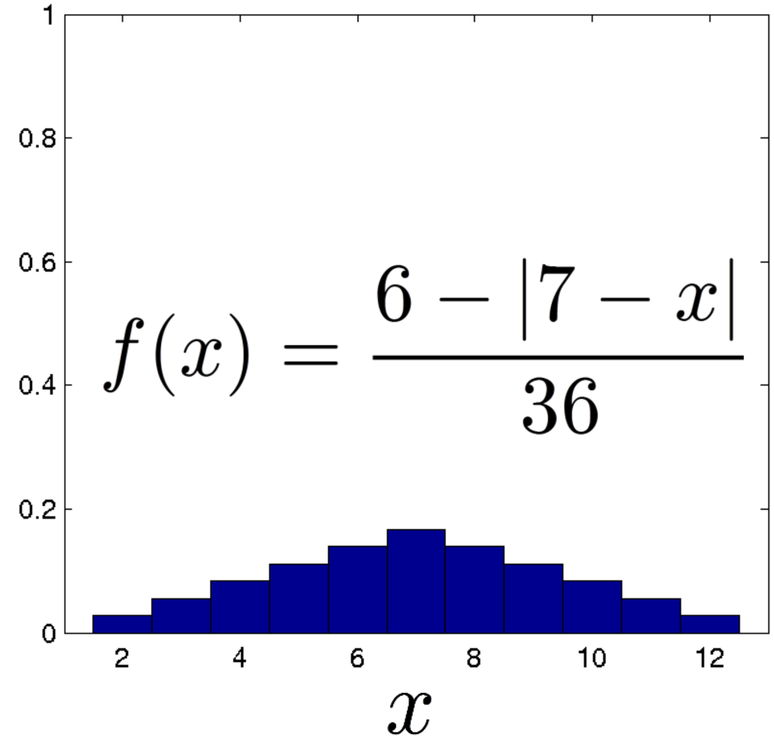
1. 確率変数と確率分布
2. 確率分布の性質を表す指標
3. 同時確率
4. 条件付き確率

確率変数の性質を表わす指標

- 期待値(expectation):確率変数の値の平均（正確には確率による重み付きの平均）
- 確率変数 X の期待値を $E(X)$ で表す
  - 離散型: $$E(X) = \sum_x x f(x)$$
  - 連続型: $$E(X) = \int x f(x) dx$$
- 何の確率変数に関する期待値なのかを正確にするために$E_X(\cdot)$ のように表記する場合もある

例

- 二つのさいころの目の和の期待値を求めよ



$$f(x) = \frac{6 - |7 - x|}{36}$$

二つのさいころの出る目の和

$$E(X) = \sum_{x=2}^{12} x f(x) = 7$$

演習：期待値演算の性質

$$E(X) = \sum_x x f(x)$$

- 定数は期待値をとっても値は変わらない：

$$E(c) = c$$

- 定数を足した期待値は，期待値に定数を足したものと等しい：

$$E(X + c) = E(X) + c$$

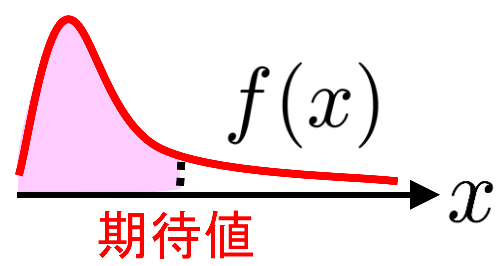
- 定数倍の期待値は，期待値の定数倍と等しい：

$$E(cX) = cE(X) \Longrightarrow \text{期待値演算は線形}$$

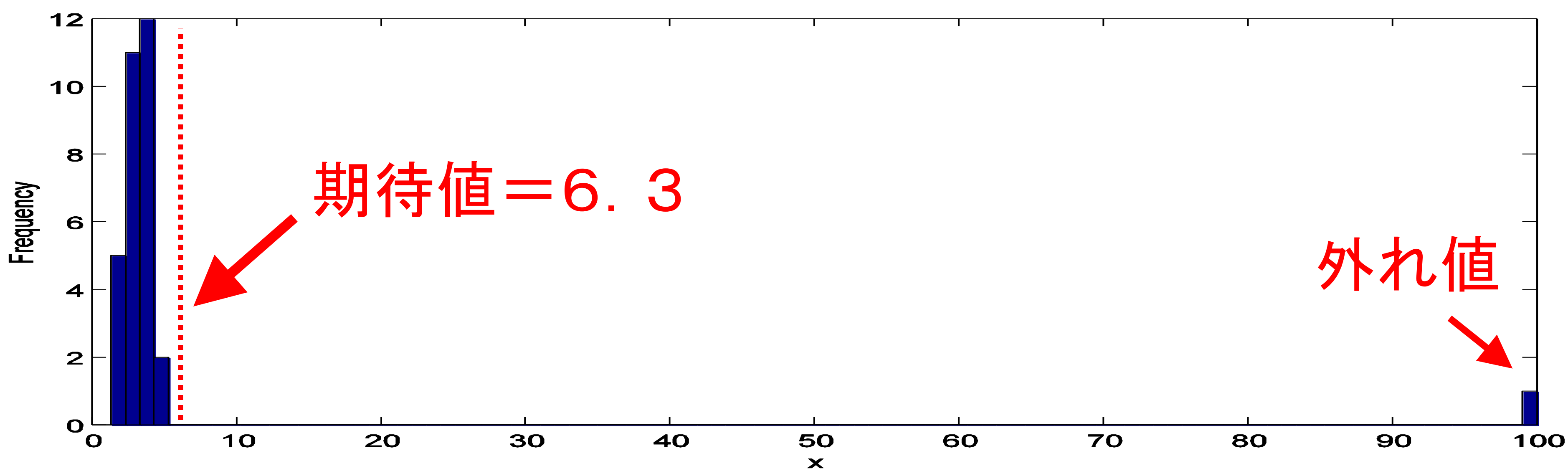
期待値の問題点

- 期待値は，外れ値(outlier)があるときに直感と合わない値になることがある．

$$E(X) = \int x f(x) dx$$



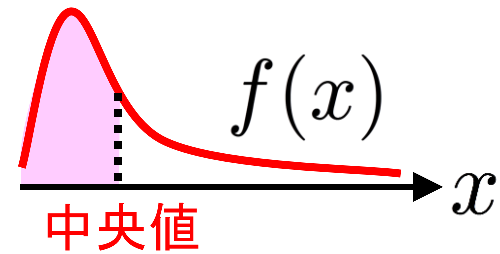
- 例：収入分布．一人大金持ち(外れ値)がいると，その人以外全員が期待値以下になってしまう．



その他のよく用いる指標

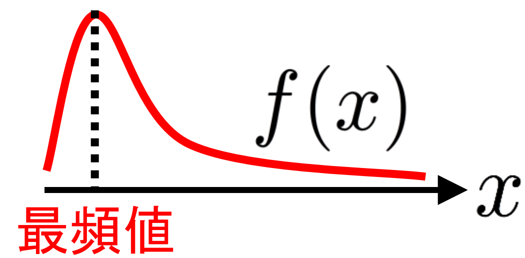
■ 中央値(median):

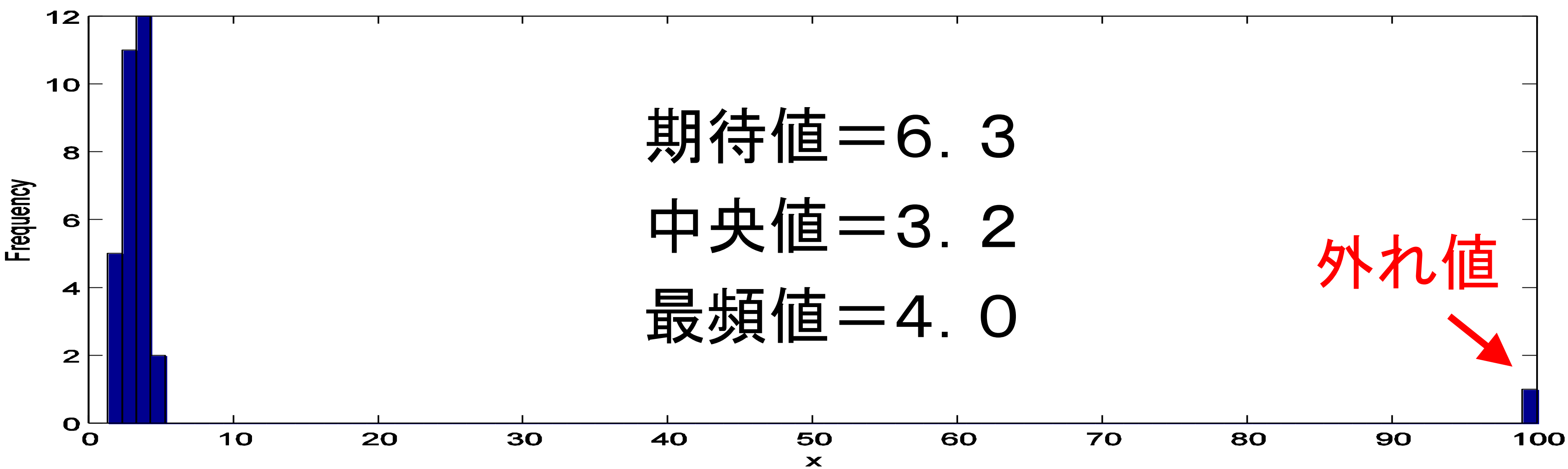
$$P(X \leq x) = \frac{1}{2} \text{となる } x$$



■ 最頻値(mode):

$$f(x) \text{を最大にする } x$$





確率変数のばらつきの指標

- 期待値は確率変数の代表する値を表す指標
- 分散(variance):確率変数の散らばり具合を表す指標

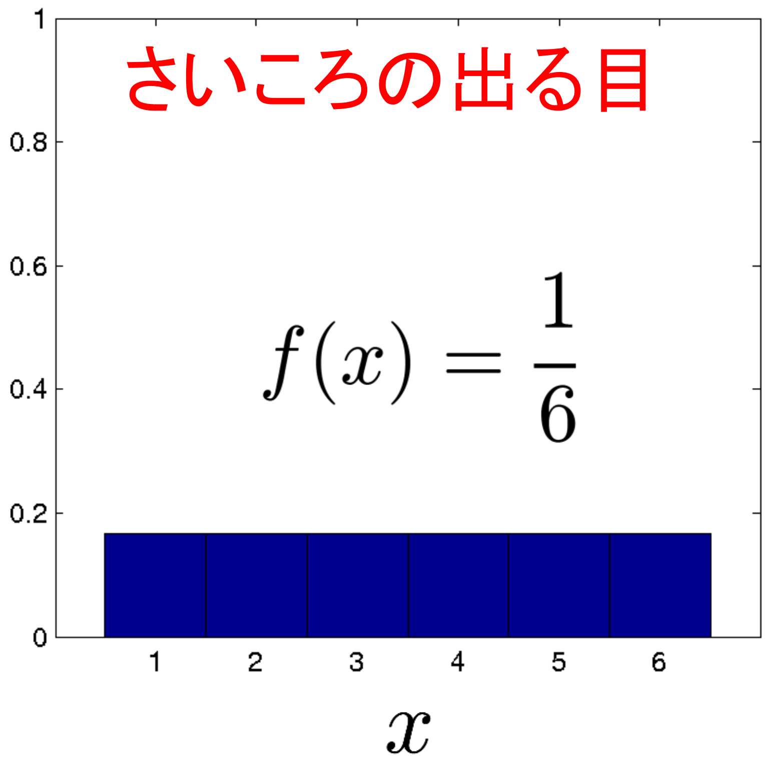
$$V(X) = E\{(X - E[X])^2\}$$

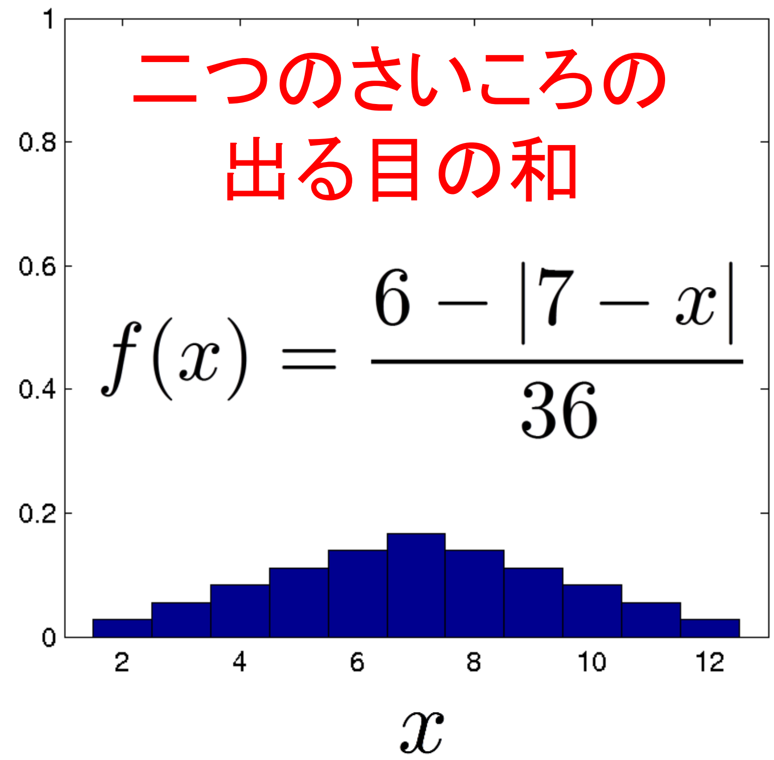
- 次式の方が計算しやすいこともある

$$\begin{aligned} V(X) &= E\{X^2 - 2E[X]X + (E[X])^2\} \\ &= E[X^2] - E[2E[X]X] + (E[X])^2 \\ &= E[X^2] - 2E[X]E[X] + (E[X])^2 \\ &= E[X^2] - (E[X])^2 \end{aligned}$$

  - ただし数値的には不安定になる場合もある

例





期待値：7／2
分散：35／12

期待値：7
分散：35／6

演習：分散演算の性質

1. 定数の分散はゼロ

$$V(c) = 0$$

2. 定数を足したものの分散は，もとの分散と等しい

$$V(X + c) = V(X)$$

3. 定数倍の分散は，もとの分散に定数の2乗をかけたものと等しい

$$V(cX) = c^2 V(X)$$

標準偏差と標準化

- 標準偏差(standard deviation): 分散の平方根 $\sqrt{V(X)}$
- 分散の値を σ^2 で，標準偏差の値を σ で表わすことが多い
- 標準化(standardization): 任意の確率変数 X に対して

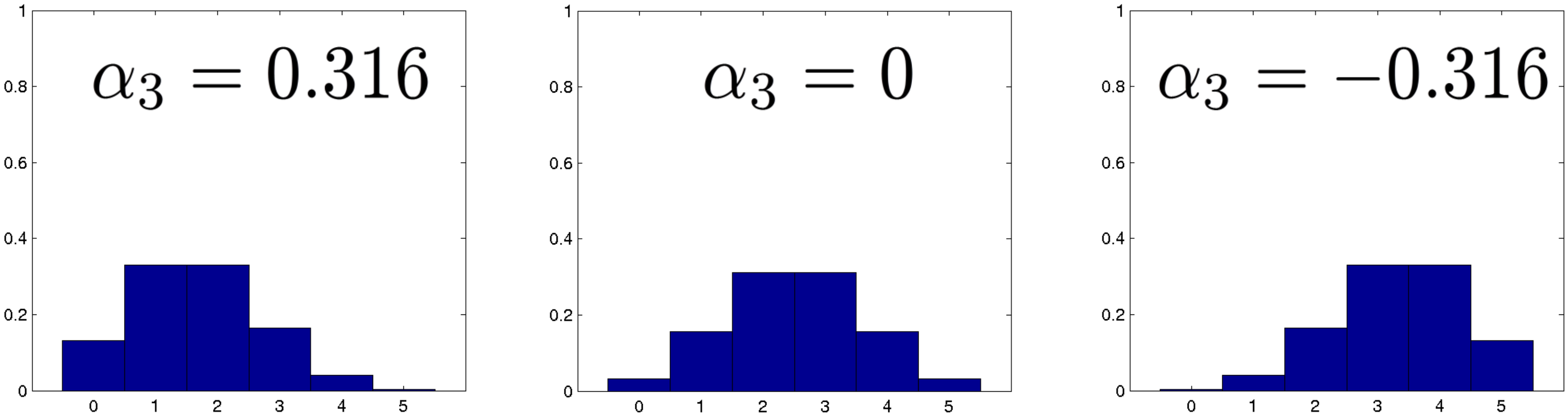
$$Z = \frac{X - E(X)}{\sqrt{V(X)}}$$

と定義すれば，Z は期待値0，分散1になる

確率分布の形の指標(1)

- 歪度(skewness) α_3 :確率分布の非対称性を表わす

$$\alpha_3 = \frac{E\{[X - E(X)]^3\}}{\{V(X)\}^{3/2}}$$

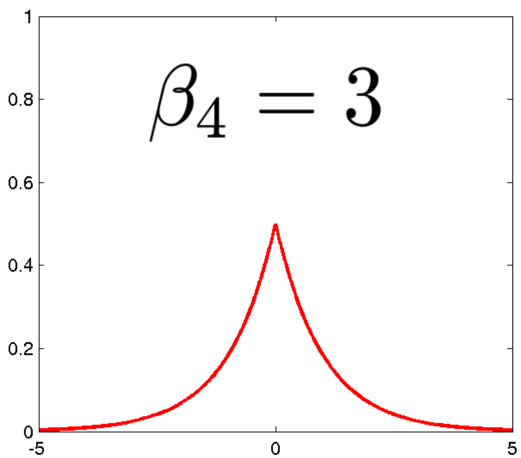


右すそが長い　　左右対称　　左すそが長い

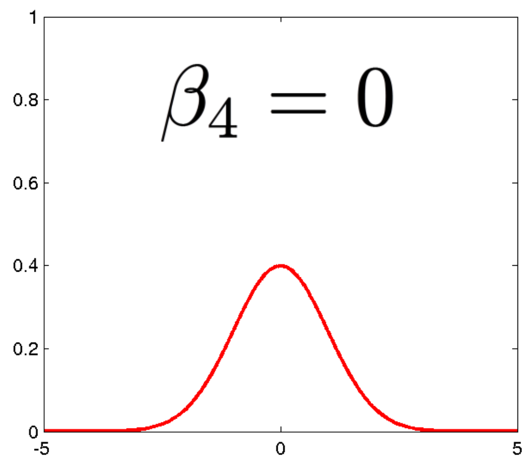
確率分布の形の指標(2)

- 尖度(kurtosis) β_4 :確率分布の尖り具合を表わす

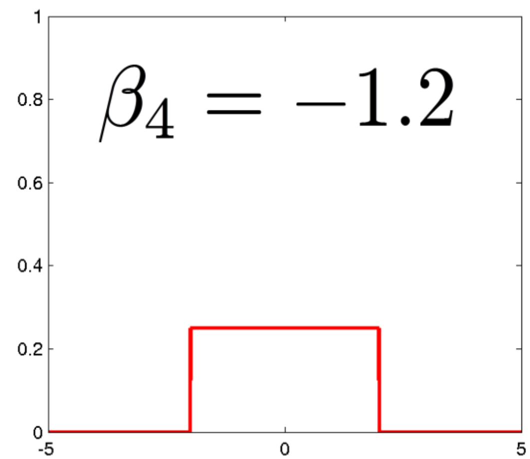
$$\beta_4 = \frac{E\{[X - E(X)]^4\}}{\{V(X)\}^2} - 3$$



尖っている
(“すそが重い”)



標準的な尖り具合
(正規分布)



尖っていない
(“すそが軽い”)

積率

- 確率分布は，期待値，分散，歪度，尖度を指定していくと，形が限定されていく
- r次の積率(moment)：

$$\mu_r = E[X^r]$$

- 期待値まわりのr次の積率：

$$\nu_r = E\left[(X - E[X])^r\right]$$

- 全ての次数の積率を指定すれば，確率分布を一意に決定することができる

演習

$$\alpha_3 = \frac{E\{[X - E(X)]^3\}}{\{V(X)\}^{3/2}}$$

$$\beta_4 = \frac{E\{[X - E(X)]^4\}}{\{V(X)\}^2} - 3$$

$$\mu_r = E[X^r]$$

- 歪度，尖度を（原点周りの）積率を用いて表せ．
- ヒント：分散は，原点周りの積率を用いて

$$V(X) = E[X^2] - (E[X])^2 = \mu_2 - \mu_1^2$$

と表せる

積率母関数

- 積率母関数(moment generating function)：全ての次数の積率を生成する関数（詳細は次ページ）

$$
M_X(t) = E[e^{tX}] = \begin{cases} \displaystyle\sum_x e^{tx} f(x) \\ \displaystyle\int e^{tx} f(x) dx \end{cases}
$$

- ただし，積率母関数は存在しない（無限大に発散する）こともある

積率母関数と積率

- **定理**: 積率母関数の導関数にゼロを代入すれば積率が得られる

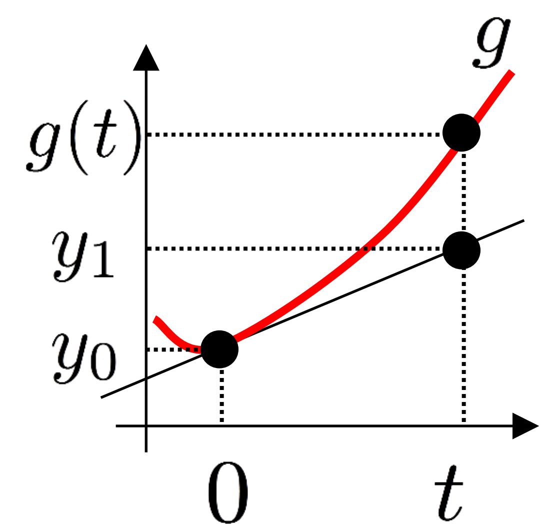
$$M_X^{(r)}(0) = \mu_r$$

- **証明の前に復習**:

(原点周りの)テーラー展開:

$g(t)$ が無限回微分可能のとき

$$g(t) = g(0) + t\frac{g'(0)}{1!} + t^2\frac{g''(0)}{2!} + \cdots$$



$$y_0 = g(0)$$

$$y_1 = g(0) + t\frac{g'(0)}{1!}$$

積率母関数と積率(続き)

- 証明:まず, e^{tX} を原点周りでテーラー展開する

$$e^{tX} = 1 + (tX) + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots$$

$$g(t) = e^{tX}$$

$$g^{(r)}(t) = X^r e^{tX}$$

$$g^{(r)}(0) = X^r$$

両辺の期待値を取れば,

$$M_X(t) = E[e^{tX}]$$

$$= E[1] + tE[X] + t^2\frac{E[X^2]}{2!} + t^3\frac{E[X^3]}{3!} + \cdots$$

$$= 1 + t\mu_1 + t^2\frac{\mu_2}{2!} + t^3\frac{\mu_3}{3!} + \cdots$$

$$\mu_r = E[X^r]$$

積率母関数と積率(続き)

- 両辺を微分すれば

$$M_X(t) = 1 + \mu_1 t + \frac{\mu_2}{2!}t^2 + \frac{\mu_3}{3!}t^3 + \cdots$$

$$M'_X(t) = \mu_1 + \mu_2 t + \frac{\mu_3}{2!}t^2 + \frac{\mu_4}{3!}t^3 + \cdots$$

$$M''_X(t) = \mu_2 + \mu_3 t + \frac{\mu_4}{2!}t^2 + \frac{\mu_5}{3!}t^3 + \cdots$$

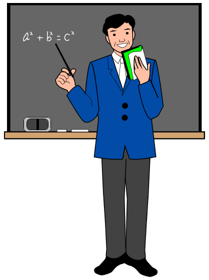
$$\vdots$$

$$M_X^{(r)}(t) = \mu_r + \mu_{r+1} t + \frac{\mu_{r+2}}{2!}t^2 + \frac{\mu_{r+3}}{3!}t^3 + \cdots$$

$$\Rightarrow \quad M_X^{(r)}(0) = \mu_r$$

確率分布の性質を表す指標：まとめ

- 確率変数の性質を表わす指標
  - 期待値，中央値，最頻値
  - 分散，標準偏差
  - 歪度
  - 尖度
  - 積率
  - 積率母関数：全ての次数の積率を生成する関数



講義の流れ

1. 確率変数と確率分布
2. 確率分布の性質を表す指標
3. 同時確率
4. 条件付き確率

多次元の確率分布

- 複数の確率変数が与えられる場合，確率変数間の関係を調べることは重要である
- 簡単のため，2つの確率変数 X と Y の性質を考える

同時確率

- 確率変数 X と Y の同時確率分布(joint probability distribution) :

$$P(a \leq X \leq b,\ c \leq Y \leq d) = \int_c^d \int_a^b f(x, y) dx dy$$

$f(x, y)$: 確率変数 X と Y の同時確率密度関数 (joint probability density function) :

$$f(x, y) \geq 0, \qquad \int \int f(x, y) dx dy = 1$$

周辺確率

- 周辺確率分布(marginal probability distribution):
  確率変数 X および Y 単独の確率分布．同時確率密度関数から，次式で求められる．

$$P(a \leq X \leq b) = \int_a^b \int f(x, y) dy dx = \int_a^b g(x) dx$$

$$P(c \leq Y \leq d) = \int_c^d \int f(x, y) dx dy = \int_c^d h(y) dy$$

$g(x), h(y)$:周辺確率密度関数(marginal probability density function)

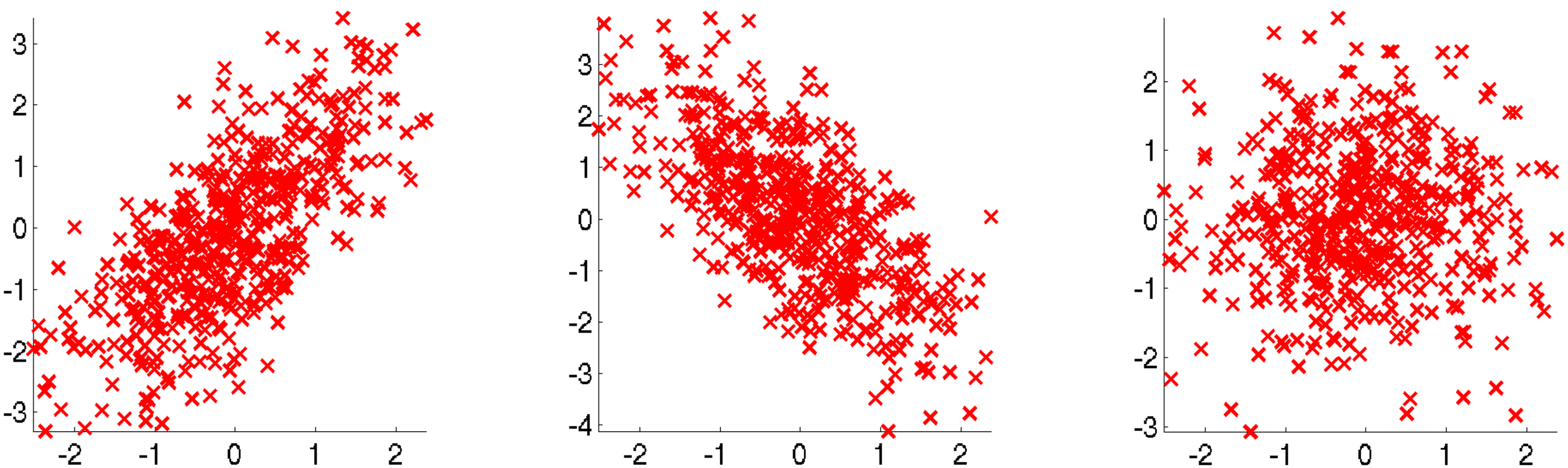
$$g(x) = \int f(x, y) dy \qquad h(y) = \int f(x, y) dx$$

共分散

- XとX'の共分散(covariance):

$$\mathrm{Cov}(X,X') = E[(X-E[X])(X'-E[X'])]$$

  - $\mathrm{Cov}(X,X') > 0$ のとき，X とX'の増減は同傾向
  - $\mathrm{Cov}(X,X') < 0$ のとき，X とX'の増減は逆傾向
  - $\mathrm{Cov}(X,X') = 0$ のとき，X とX'の増減は無関係



$\mathrm{Cov}(X,X') > 0$ $\mathrm{Cov}(X,X') < 0$ $\mathrm{Cov}(X,X') \approx 0$

分散と共分散の性質

1. 二つの確率変数 X と X' の和の期待値は，それぞれの期待値の和と等しい：

$$E(X + X') = E(X) + E(X')$$

2. しかし X と X' の和の分散は，一般にはそれぞれの分散の和とは等しくない．

$$V(X + X') = V(X) + V(X') + 2\mathrm{Cov}(X, X')$$

$$\mathrm{Cov}(X, X') = E[(X - E[X])(X' - E[X'])]$$

証明は演習

共分散の使用法の例

$$V(X+X') = V(X) + V(X') + 2\mathrm{Cov}(X, X')$$

- A社の株価を X, B社の株価を X' とする
- $\mathrm{Cov}(X, X') > 0$ のとき,

$$V(X+X') > V(X) + V(X')$$

  - A, B両社の株を買うと, 分散が拡大
  - 変動リスクが増大し, 資産価値は不安定

- $\mathrm{Cov}(X, X') < 0$ のとき,

$$V(X+X') < V(X) + V(X')$$

  - A, B両社の株を買うと, 分散が縮小
  - 変動リスクが抑制され、資産価値は安定

分散共分散行列

- 分散共分散行列(variance-covariance matrix):

$$\boldsymbol{\Sigma} = E\left[\left\{\begin{pmatrix} X \\ X' \end{pmatrix} - E\begin{pmatrix} X \\ X' \end{pmatrix}\right\}\left\{\begin{pmatrix} X \\ X' \end{pmatrix} - E\begin{pmatrix} X \\ X' \end{pmatrix}\right\}^{\top}\right]$$

$$= \begin{pmatrix} V(X) & \mathrm{Cov}(X, X') \\ \mathrm{Cov}(X, X') & V(X') \end{pmatrix}$$

- 対角成分は分散, 非対角成分は共分散.
- 単に共分散行列ともよぶ

相関

- 相関係数(correlation coefficient):共分散を標準偏差で割った値

$$\rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$$

- 相関係数は $-1 \le \rho_{XY} \le 1$ を満たす（証明は宿題）

- $\rho_{XY} > 0$ のとき，正の相関があるという.
- $\rho_{XY} < 0$ のとき，負の相関があるという.
- $\rho_{XY} = 0$ のとき，無相関(uncorrelated)であるという.
- $|\rho_{XY}| \approx 1$ のとき，X と Y の増減関係はより確定的になる.

独立性

- X と Y は互いに独立(independent)：全ての(x, y)で

$$f(x, y) = g(x)h(y)$$

- 2つの確率変数が独立のとき,
  a. 積の期待値は各々の期待値の積と一致:

$$E[XY] = E[X]E[Y]$$

  b. 和の積率母関数は各々の積率母関数の積と一致:

$$M_{X+Y}(t) = M_X(t)M_Y(t)$$

  c. 2つの確率変数は無相関:

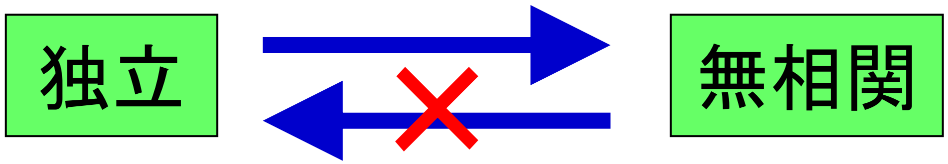
$$\mathrm{Cov}(X, Y) = 0$$

証明は宿題

$$M_X(t) = E[e^{tX}]$$

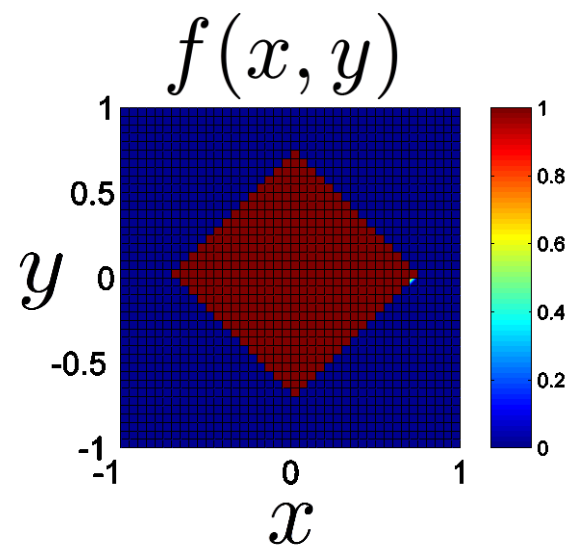
独立性と無相関性

- ２つの確率変数が独立ならば無相関である.
- しかし，逆は一般には正しくない.



- 反例：

$$f(x,y) = \begin{cases} 1 & \text{if } |x| + |y| \leq \frac{1}{\sqrt{2}} \\ 0 & \text{otherwise} \end{cases}$$



無相関性の証明

$$
\begin{aligned}
\mathrm{Cov}(X,Y) &= E[(X-E[X])(Y-E[Y])] \\
&= E[XY] \qquad\qquad E[X]=0,\ E[Y]=0 \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} y f(x,y)\,dy\,dx \\
&= \int_{-\frac{1}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}} x \left(\int_{-\frac{1}{\sqrt{2}}+|x|}^{\frac{1}{\sqrt{2}}-|x|} \underbrace{y}_{\text{奇関数}}\,dy \right) dx \\
&= \int_{-\frac{1}{\sqrt{2}}}^{\frac{1}{\sqrt{2}}} x \cdot 0\,dx = 0
\end{aligned}
$$

独立でないことの確認

- 周辺確率密度関数 $g(x)$ は,
  - $-\frac{1}{\sqrt{2}} \le x \le \frac{1}{\sqrt{2}}$ に対して

$$g(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{-\frac{1}{\sqrt{2}}+|x|}^{\frac{1}{\sqrt{2}}-|x|} dy = \sqrt{2} - 2|x|$$

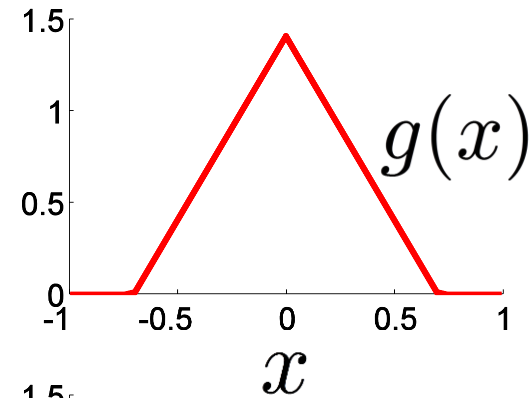
  - それ以外の場合は

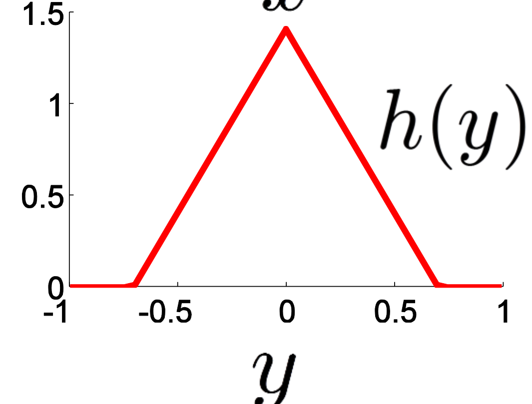
$$g(x) = 0$$

よって, $g(x) = \max(0, \sqrt{2} - 2|x|)$

同様に,

$$h(y) = \max(0, \sqrt{2} - 2|y|)$$





独立でないことの確認

■ 従属性の証明:

- 同時確率密度関数は

$$f(x,y) = \begin{cases} 1 & \text{if } |x|+|y| \le \frac{1}{\sqrt{2}} \\ 0 & \text{otherwise} \end{cases}$$

- 周辺確率密度関数の積は

$$g(x)h(y) = \begin{cases} (\sqrt{2}-2|x|)(\sqrt{2}-2|y|) & \text{if } -\frac{1}{\sqrt{2}} \le x, y \le \frac{1}{\sqrt{2}} \\ 0 & \text{otherwise} \end{cases}$$

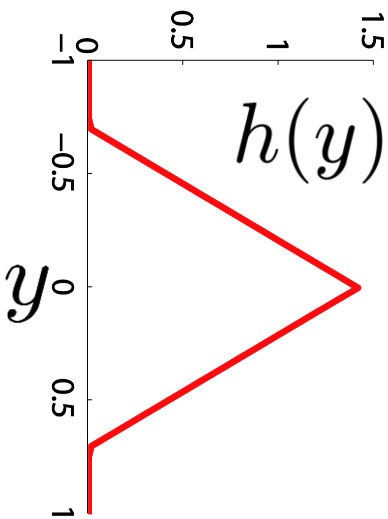
$$g(x) = \max(0, \sqrt{2}-2|x|)$$

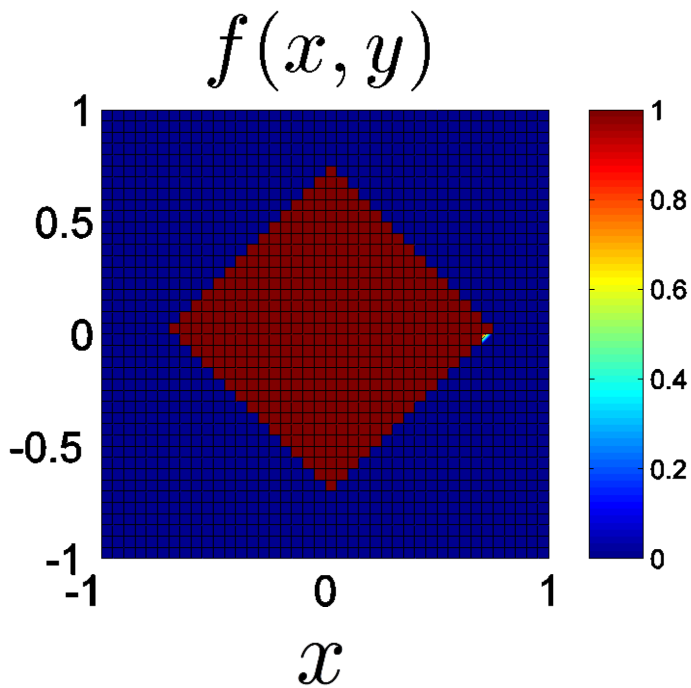
$$h(y) = \max(0, \sqrt{2}-2|y|)$$

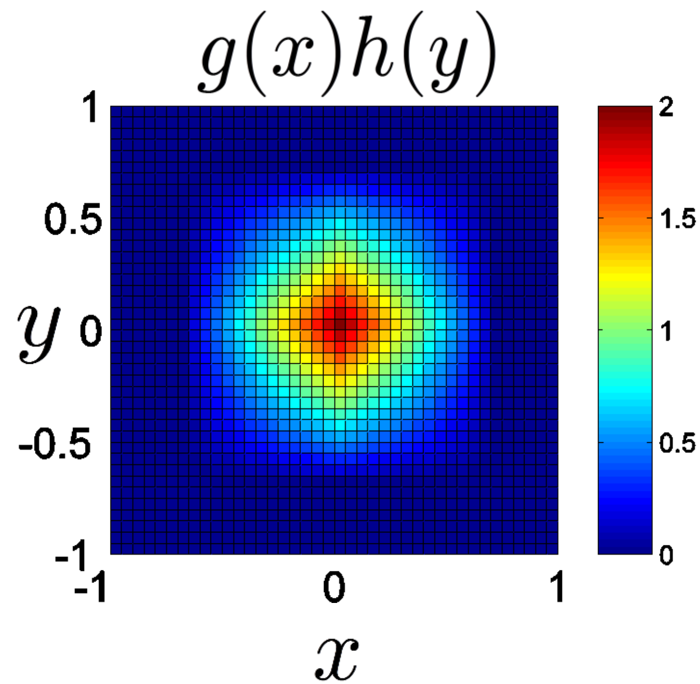
- 従って,

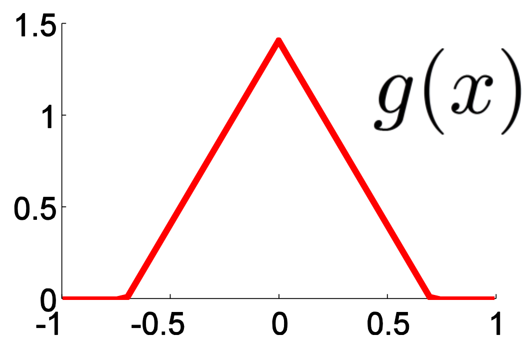
$$g(x)h(y) \ne f(x,y)$$

独立でないことの確認









$$g(x) = \max(0, \sqrt{2} - 2|x|)$$

$$h(y) = \max(0, \sqrt{2} - 2|y|)$$

$$f(x,y) = \begin{cases} 1 & \text{if } |x| + |y| \leq \frac{1}{\sqrt{2}} \\ 0 & \text{otherwise} \end{cases}$$

確率変数の和

- X, Y : 独立な確率変数
- $g(x), h(y)$: それぞれの周辺確率密度関数
- $Z = X + Y$ の確率密度関数 $k(z)$は

$$k(z) = \int g(x)h(z-x)dx \qquad y = z - x$$

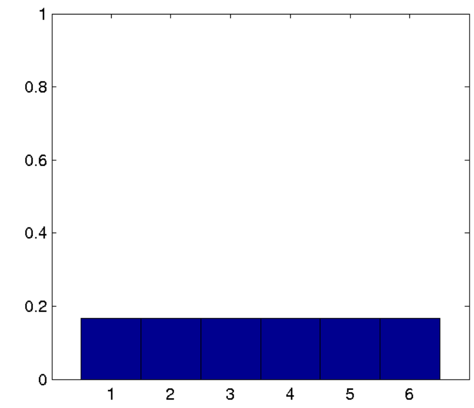
- 6面体のサイコロ2個の和が7のとき，それぞれの値は(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)なので，それらの確率を足しあわせれば良い

- これを畳み込み(convolution)とよび，$k = g * h$で表す．
- 再生的(reproductive) : 同じ種類の確率分布の畳み込みの結果が再び同じ種類の確率分布になる（例 : 正規分布）

畳み込みの例

- さいころ1の出る目 X:

$$g(x) = \frac{1}{6} \text{ for } x = 1, 2, \ldots, 6$$



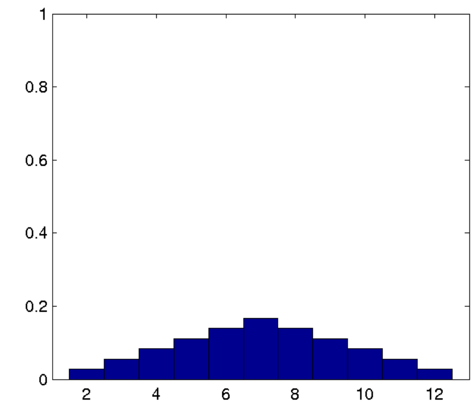
- さいころ2の出る目 Y:

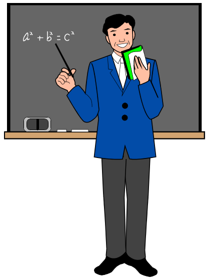
$$h(y) = \frac{1}{6} \text{ for } y = 1, 2, \ldots, 6$$

- 二つのさいころの出る目の和 $Z = X + Y$:

$$k(z) = \frac{6 - |7 - z|}{36}$$

$$\text{for } z = 2, 3, \ldots, 12$$





講義の流れ

1. 確率変数と確率分布
2. 確率分布の性質を表す指標
3. 同時確率
4. 条件付き確率

条件付き確率(離散型の場合)

- 離散型の確率変数 X, Y の条件付き確率分布 (conditional probability distribution):
  $Y = y$ が起こったもとで $X = x$ が起こる確率

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

$$\text{for } P(Y = y) \neq 0$$

- 離散型の確率変数に対する条件付き確率質量関数 (conditional probability mass function):

$$p(x|y) = \frac{f(x, y)}{h(y)} \text{ for } h(y) \neq 0$$

$$h(y) = \sum_x f(x, y)$$

条件付き確率の例

■ 分割表(contingency table)

X　　Y	授業に出席	授業に欠席	合計
晴れ	0.6	0.1	0.7
雨	0.1	0.2	0.3
合計	0.7	0.3	1

P(X=晴)=0.7, P(X=雨)=0.3

P(X=晴 | Y=出)=6/7, P(X=雨| Y=出)=1/7

P(X=晴 | Y=欠)=1/3, P(X=雨| Y=欠)=2/3

条件付き確率(連続型の場合)

- 条件付き確率関数を連続型に変えることにする.
- 条件付き確率密度関数(conditional probability density function):

$$p(x|y) = \frac{f(x,y)}{h(y)} \ \text{for} \ h(y) \neq 0$$

$$h(y) = \int f(x,y)dx$$

- $p(x|y)$ が確率密度関数になっていることの証明:
  - 非負性: $f(x,y) \geq 0, \ h(y) > 0$ より $p(x|y) \geq 0$.
  - 正規性: $\displaystyle \int p(x|y)dx = \int \frac{f(x,y)}{h(y)}dx = \frac{h(y)}{h(y)} = 1$

条件付きの期待値と分散

- 条件付き確率も確率なので，期待値や分散が定義できる
- 条件付き期待値(conditional expectation)：

$$E(X|y) = \int xp(x|y)dx$$

- 条件付き分散(conditional variance)：

$$V(X|y) = \int (x - E(X|y))^2 p(x|y)dx$$

ベイズの定理

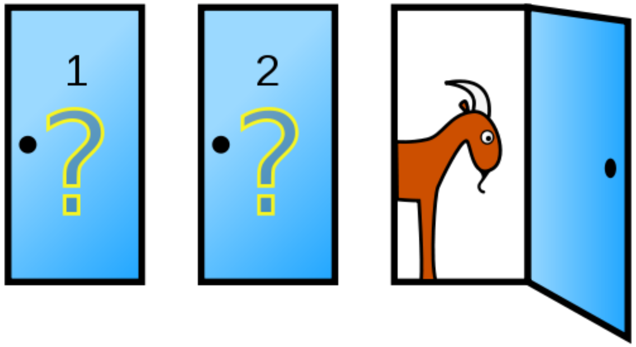
- ベイズの定理(Bayes' theorem):

$$p(x|y) = \frac{q(y|x)g(x)}{\int q(y|x)g(x)dx}$$

- $g(x)$: x の事前確率(prior probability)
  y を知る前の x の確率
- $p(x|y)$: x の事後確率(posterior probability)
  y を知った後の x の確率
- 事前確率 $g(x)$ と条件付き確率 $q(y|x)$ から,
  事後確率 $p(x|y)$ が計算できる

モンティ・ホール問題（米クイズ番組）

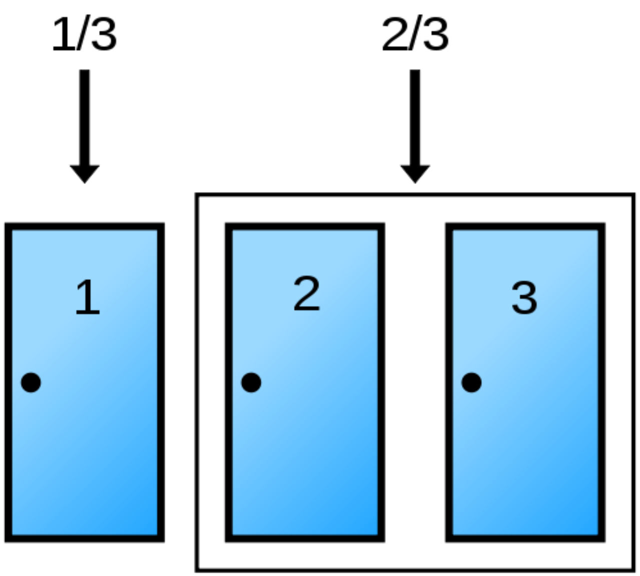
- プレーヤーの前に3つのドアがある.
  - 1つのドアの後ろには景品の車がある
  - 2つのドアの後ろには，はずれを意味するヤギがいる.
- プレーヤーは新車のドアを当てると新車がもらえる.
- プレーヤーが1つのドアを選択した後，モンティ（司会者）が残りのドアのうちヤギがいるドアを（ランダムに）開けてヤギを見せる.
- ここでプレーヤーは，最初に選んだドアを，残っている開けられていないドアに変更してもよいと言われる.
- プレーヤーはドアを変更すべきか？

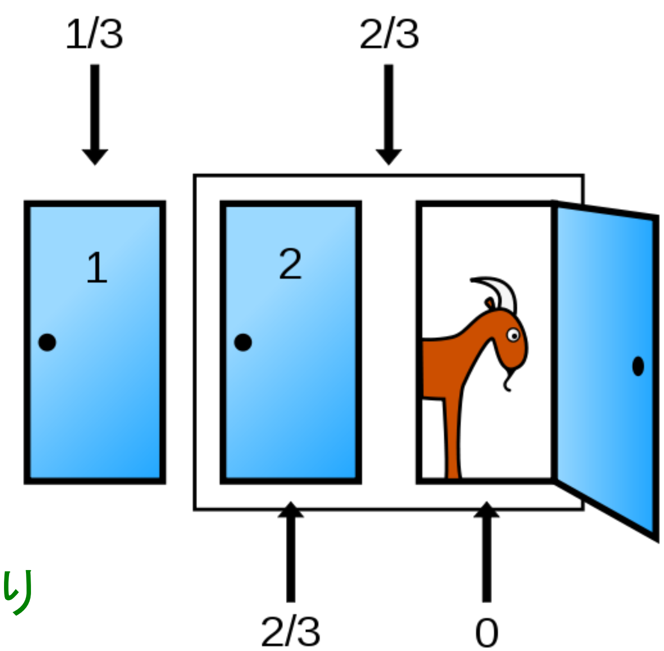


Wikipediaより

解（直感的な説明）

- ドアを変更すると当たる確率が2倍になる
- 最初に選んだドア１が当たりの確率は１／３
  - ドア２かドア３が当たりの確率は２／３
- ドア３にヤギがいたとすると，ドア２が当たりの確率は２／３





Wikipediaより

解(ベイズの定理)

- X: 当たりのドアの番号
  - $P(X=1) = P(X=2) = P(X=3) = 1/3$

等確率とする

- プレーヤーはドア1を選んだとする
- Y: モンティが開けるドアの番号
  - $P(Y=2|X=1) = P(Y=3|X=1) = 1/2$
  - $P(Y=2|X=2) = 0$, $P(Y=3|X=2) = 1$
  - $P(Y=2|X=3) = 1$, $P(Y=3|X=3) = 0$

等確率で選ぶ

当たりのドアは選べない

- モンティはドア2を選んだとする

ドア2はハズレ

- プレーヤーはドア3に変えた方が得する
  - $P(X=1|Y=2) = P(Y=2|X=1)P(X=1)/P(Y=2) = 1/3$
  - $P(X=3|Y=2) = P(Y=2|X=3)P(X=3)/P(Y=2) = 2/3$

$P(Y=2) = P(Y=2|X=1)P(X=1) + P(Y=2|X=3)P(X=3) = 1/2$

演習

- 1年間のうち晴れの日に学生Bが出席する確率が

  P(Y=出｜X=晴)=0.75

  また，学生Bが出席した日が晴れである確率が

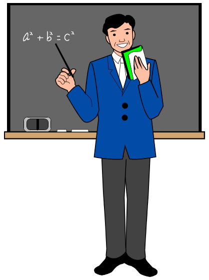
  P(X=晴｜Y=出)=0.8

  であるとする．さらに，晴れかつ学生Bが出席する確率が

  P(X=晴，Y=出)=0.6

  であるとする．

  1. P(X=晴)およびP(Y=出)を求めよ．
  2. P(X=晴，Y=欠), P(X=雨，Y=出), P(X=雨，Y=欠)をそれぞれ求めよ．
  3. 学生Bにおいて天気と出欠は独立か？

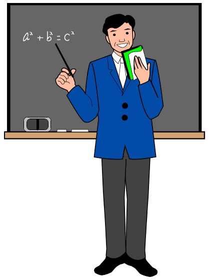


講義の流れ

1. 確率変数と確率分布
2. 確率分布の性質を表す指標
3. 同時確率
4. 条件付き確率

まとめ

- 確率分布：確率質量関数，確率密度関数，累積分布関数
- 確率変数の性質を表わす指標：期待値，中央値，最頻値，分散，標準偏差，歪度，尖度，積率と積率母関数
- 同時確率，周辺確率，共分散，相関，独立性，無相関性，畳み込み
- 条件付き確率，ベイズの定理



次回の予告

- 離散型確率分布の例:
  - 一様分布, 二項分布, 超幾何分布, ポアソン分布, 負の二項分布
- 連続型確率分布の例:
  - 正規分布, ガンマ分布, ベータ分布, コーシー分布

宿題1

$-\infty < a < b < \infty$

1. $[a,b]$ 上に定義された確率密度関数 $f(x)$ を考える.

A) 次の二乗誤差 $J_1(y)$ を最小にする y を y_1 で表す:

$$y_1 = \operatorname*{argmin}_y J_1(y) \qquad J_1(y) = \int_a^b (x-y)^2 f(x) dx$$

このとき，y_1 は X の期待値(つまり, $y_1 = E[X]$)であることを示せ.

B) 次の絶対誤差 $J_2(y)$ を最小にする y を y_2 で表す:

$$y_2 = \operatorname*{argmin}_y J_2(y) \qquad J_2(y) = \int_a^b |x-y| f(x) dx$$

このとき，y_2 は X の中央値(つまり, $F(y_2) = 1/2$)であることを示せ.

$\operatorname*{argmin}_y J(y)$ ： $J(y)$を最小にする y

宿題2

- 相関係数

$$\rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$$

が

$$-1 \leq \rho_{XY} \leq 1$$

を満たすことを証明せよ.

- ヒント: $X' = X/\sqrt{V(X)},\ Y' = Y/\sqrt{V(Y)}$ とおいて $X' + Y'$ や $X' - Y'$ の分散を調べる

宿題3

- 二つの確率変数 X と Y が独立のとき，次式が成り立つことを証明せよ．

i. $E[XY] = E[X]E[Y]$

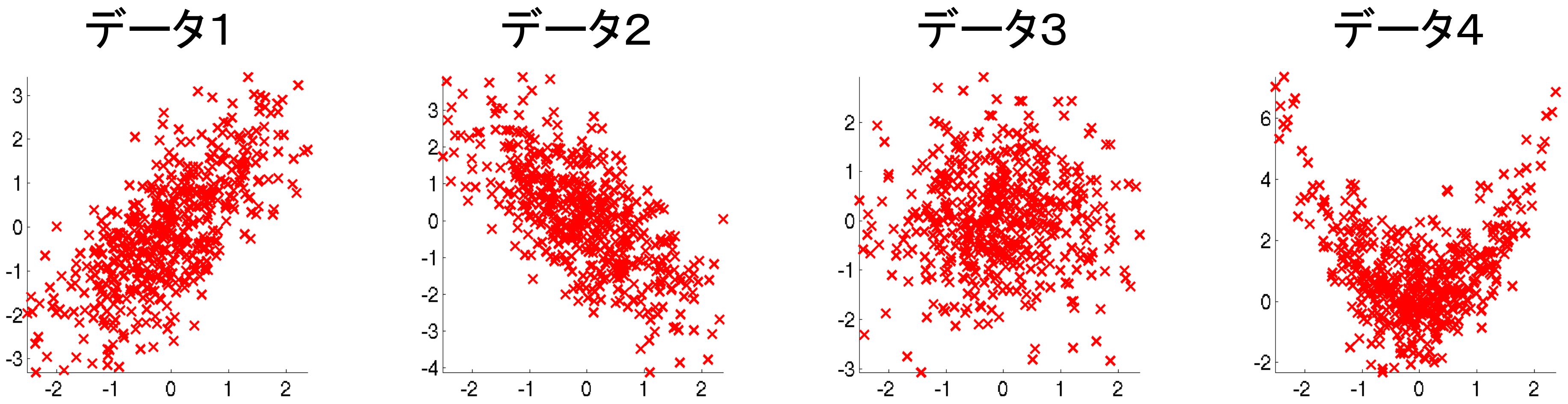
ii. $M_{X+Y}(t) = M_X(t)M_Y(t)$

iii. $\mathrm{Cov}(X, Y) = 0$

$$M_X(t) = E[e^{tX}]$$

宿題4

Octaveなどを用いて，ITC-LMSに置いてある次の４つのデータの相関係数を計算せよ



そして，その結果からわかったことを論ぜよ.