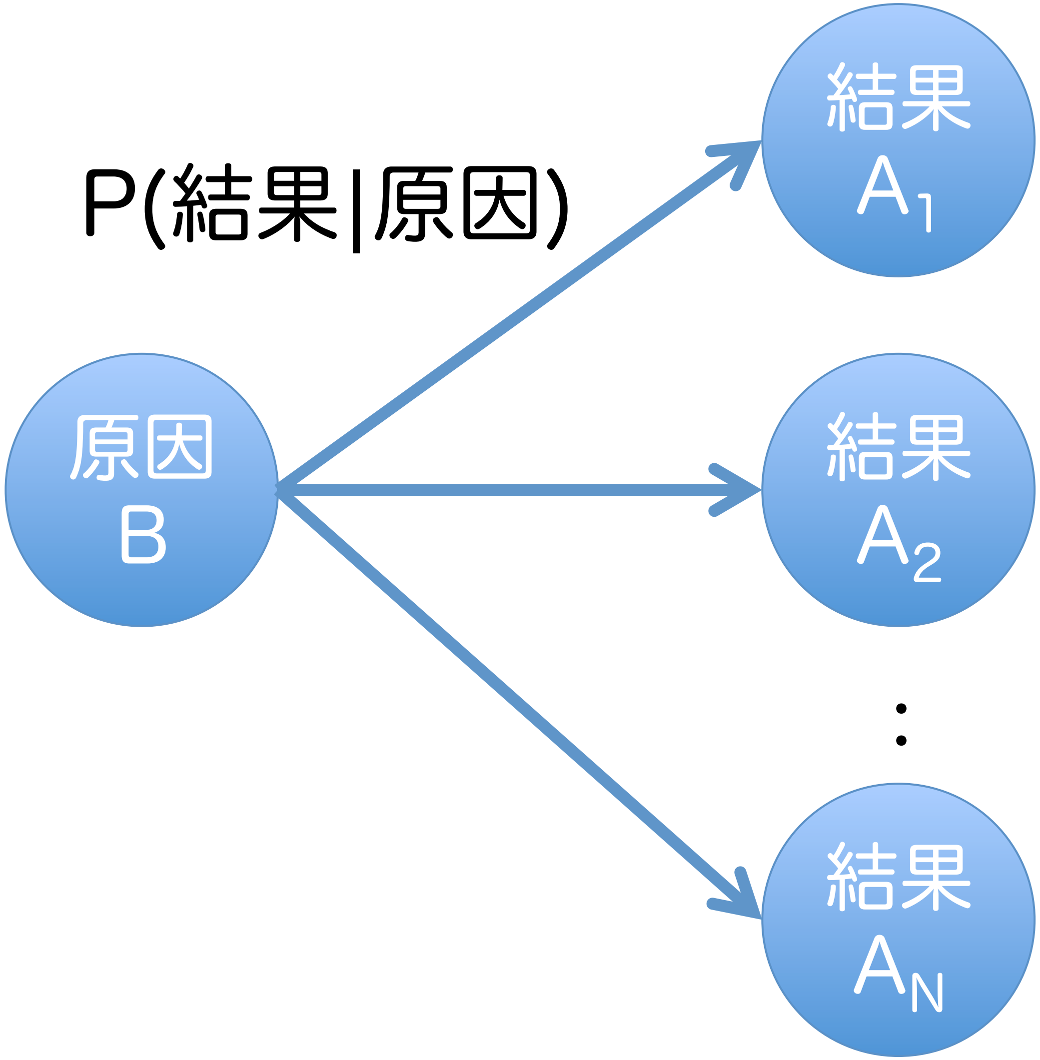
ベイズ統計

古谷知之

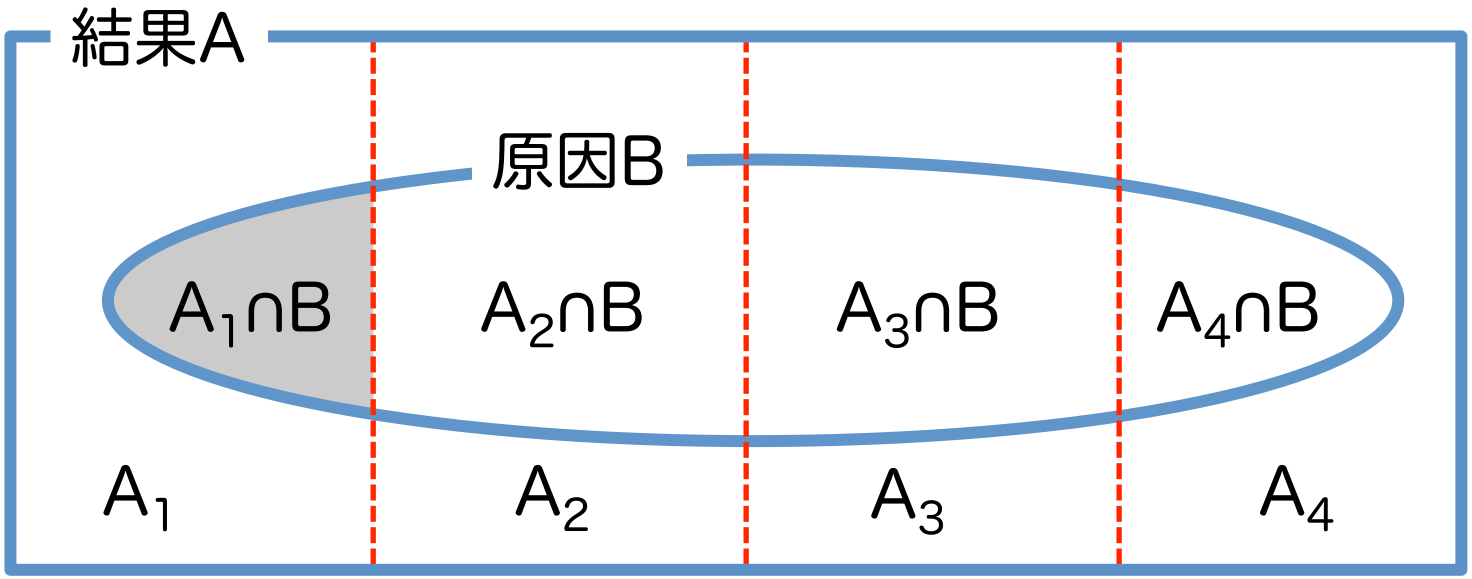
講義概要

- ベイズの定理と逆確率（復習）
- 事前確率、尤度、事後確率
- 確率変数と確率分布
- 一様分布
- 二項分布、ベータ分布
- ガンマ分布
- 正規分布

原因から結果の推定



原因と結果の関係



ベイズの定理（原因→結果）

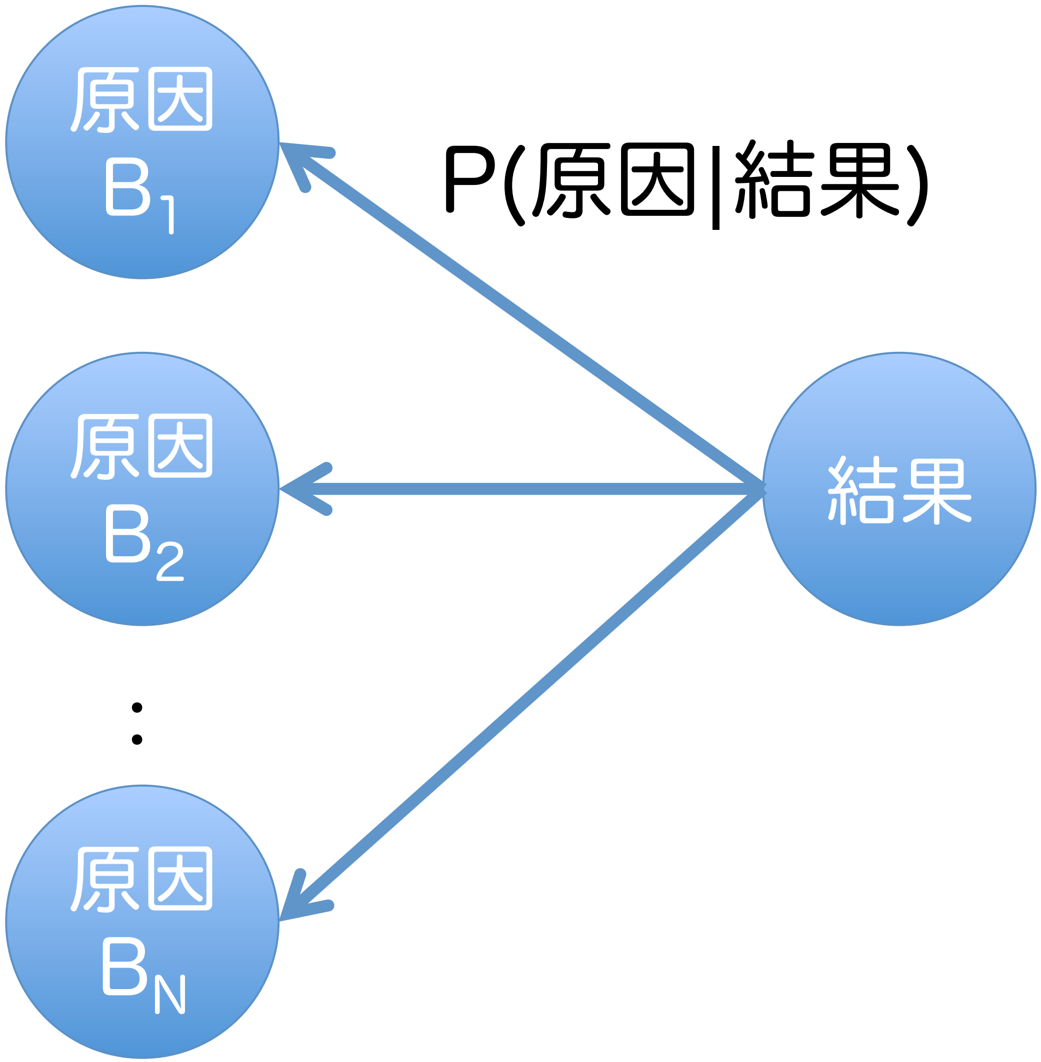
- 原因Bから結果A_1が生じうる確率

$$P(A_1 \mid B) = \frac{P(B \mid A_1)P(A_1)}{\sum_{i=1}^{N} P(B \cap A_i)}$$

$$= \frac{P(B \mid A_1)P(A_1)}{\sum_{i=1}^{N} P(B \mid A_i)P(A_i)}$$

$$= \frac{P(B \mid A)P(A)}{P(B)}$$

逆確率による原因の推定



逆確率（結果→原因）

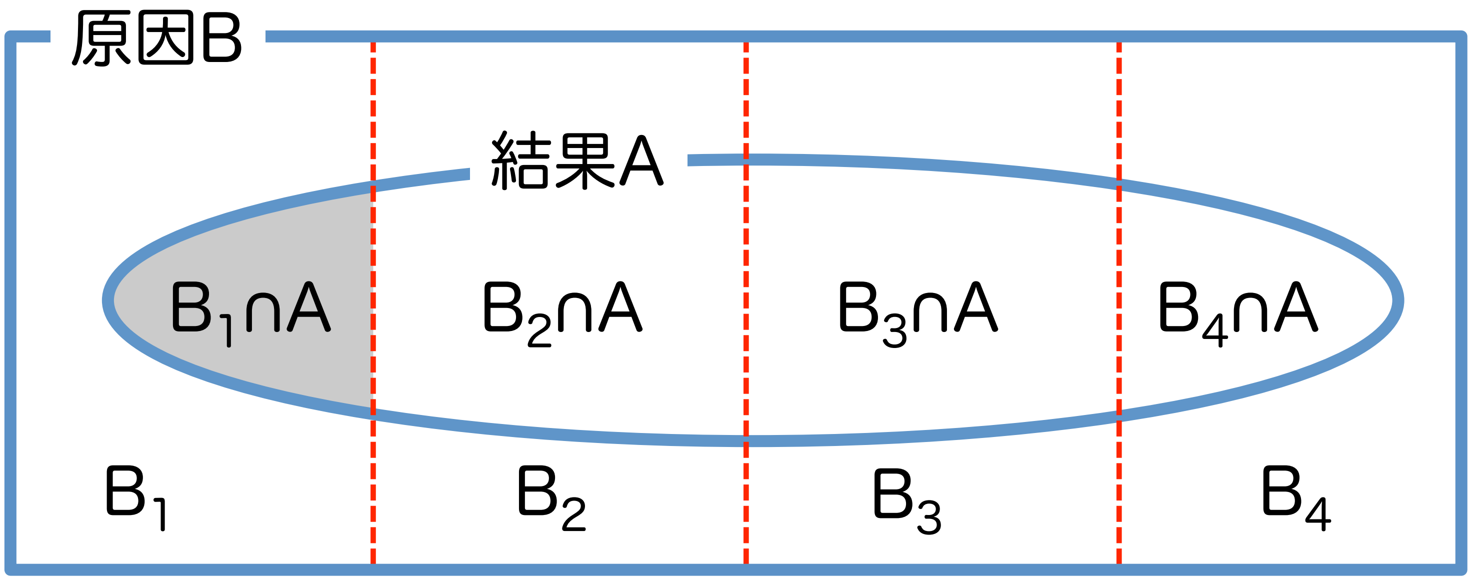
- 結果Aの原因BがB_1でありうる確率

$$P\left(B_1 \mid A\right) = \frac{P\left(A \mid B_1\right) P\left(B_1\right)}{\sum_{i=1}^{N} P\left(A \cap B_i\right)}$$

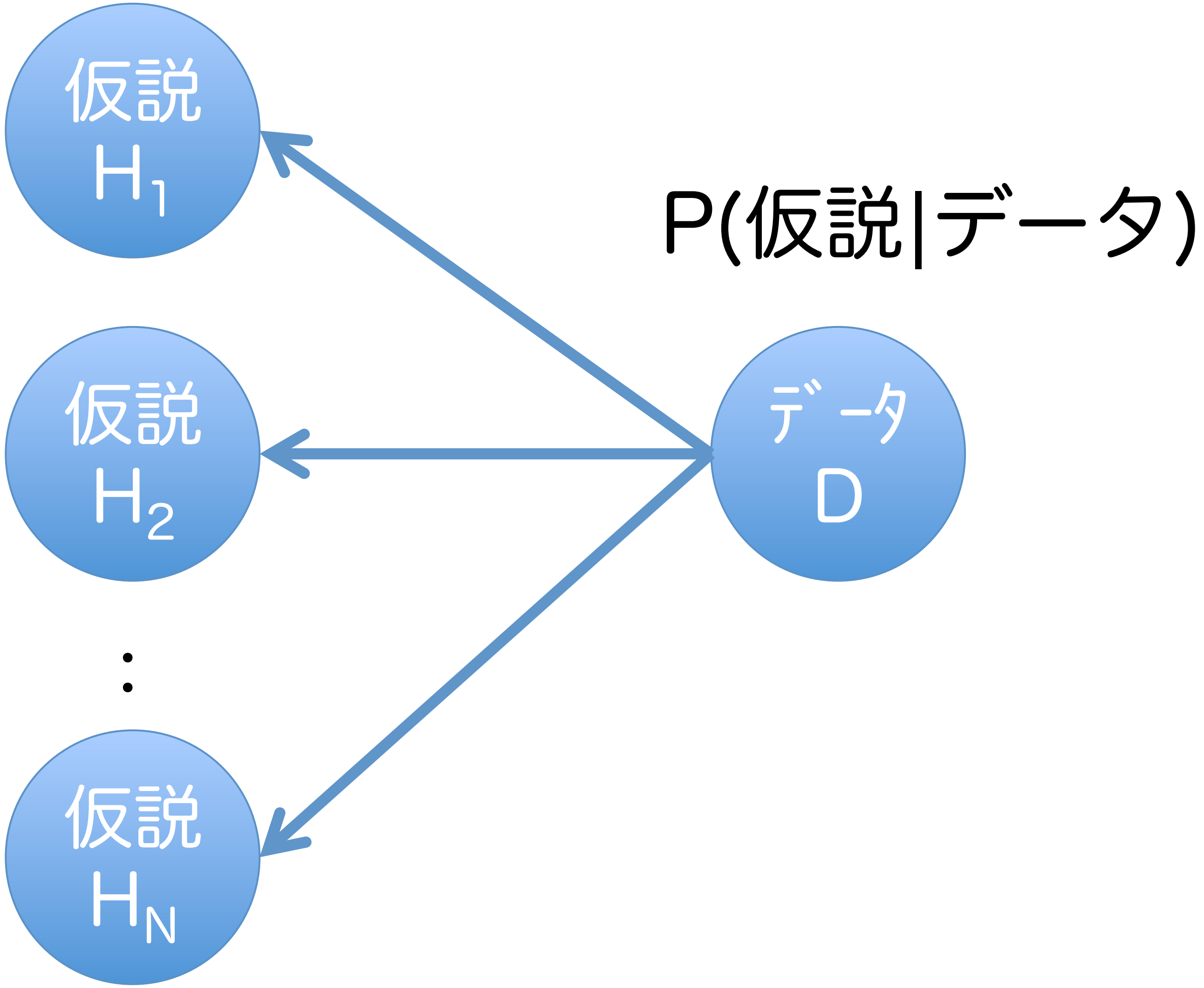
$$= \frac{P\left(A \mid B_1\right) P\left(B_1\right)}{\sum_{i=1}^{N} P\left(A \mid B_i\right) P\left(B_i\right)}$$

$$= \frac{P\left(A \mid B\right) P\left(B\right)}{P\left(A\right)}$$

逆確率における原因と結果の関係



データから仮説（モデル）を推定する



データから仮説（モデル）を推定する

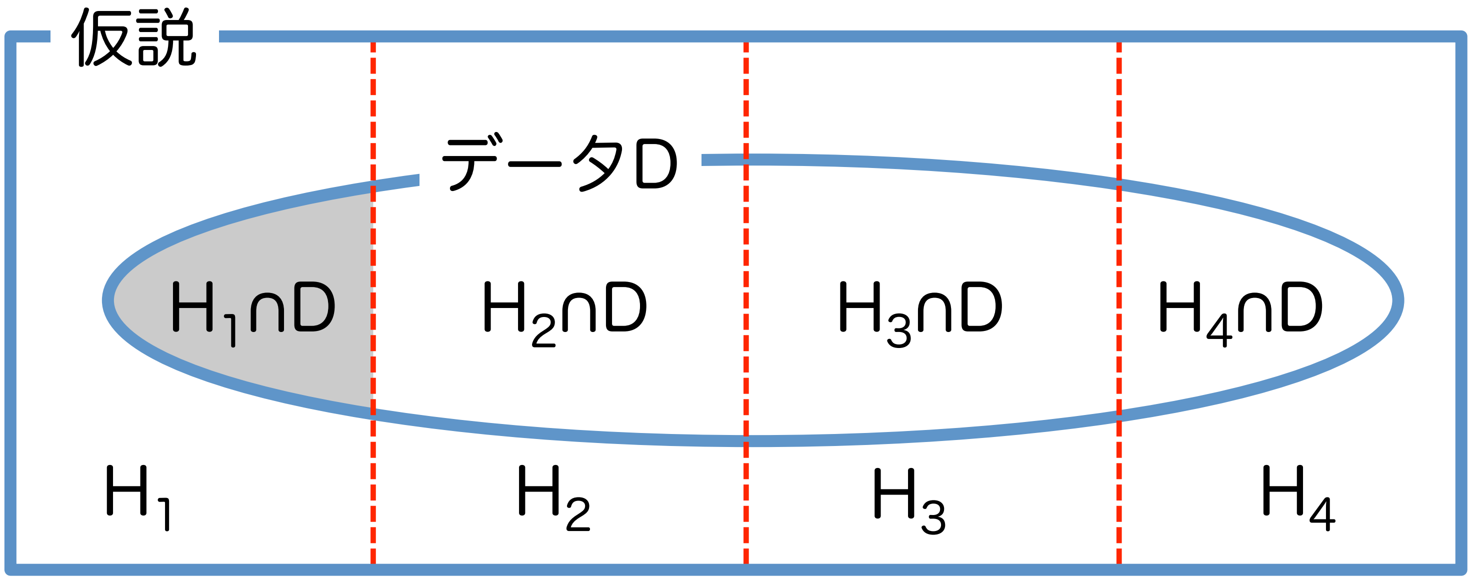
- データDから仮説H_1が成立する確率

$$P\left(H_1 \mid D\right) = \frac{P\left(D \mid H_1\right)P\left(H_1\right)}{\sum_{i=1}^{N} P\left(D \cap H_i\right)}$$

$$= \frac{P\left(D \mid H_1\right)P\left(H_1\right)}{\sum_{i=1}^{N} P\left(D \mid H_i\right)P\left(H_i\right)}$$

$$= \frac{P\left(D \mid H\right)P\left(H\right)}{P\left(D\right)}$$

データから仮説（モデル）を推定する



ベイズの定理とベイズ統計

- ベイズ統計でのモデル推定は、データが与えられた条件の下で仮定されたモデルの成立する確率P(仮説|データ)を求めていることに他ならない
- 確率的には与えられたデータのもとで様々なモデル（仮説）が成立する可能性がある

尤度（ゆうど）

- 原因から結果が生じる確率P(結果|原因)を尤度という
- 原因と結果を、データと仮説（モデル）と読み替えると、
- ある仮説（モデル）を所与としてデータが得られる確率P(データ|仮説)といえる
- 仮説がデータに当てはまる当てはまりやすさ（尤もらしさ）を尤度という

事前確率・事後確率・尤度

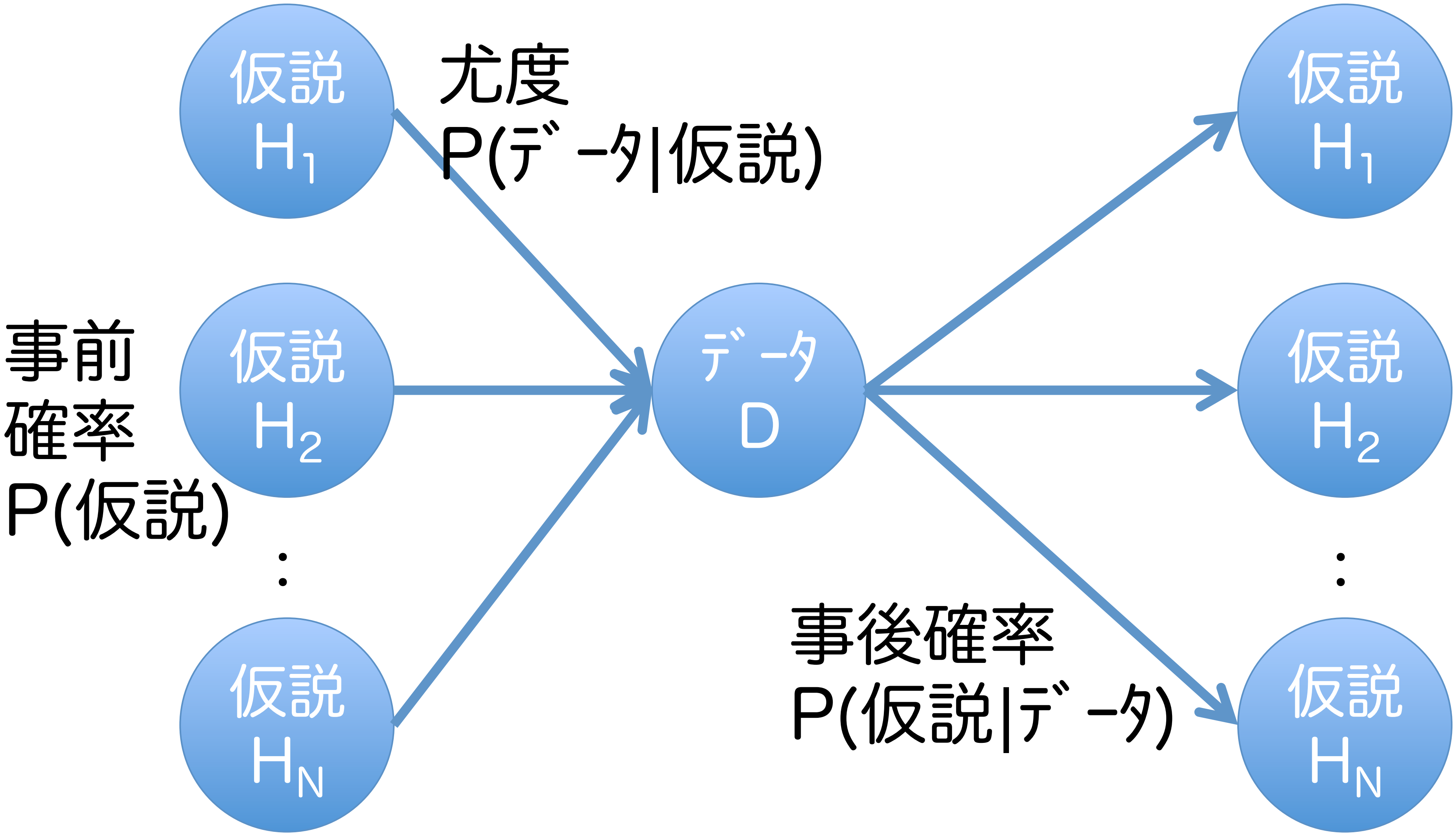
- ベイズの定理にもとづいて、P(H)を事前確率、P(D|H)を尤度、P(H|D)を事後確率という

尤度　事前確率

事後確率

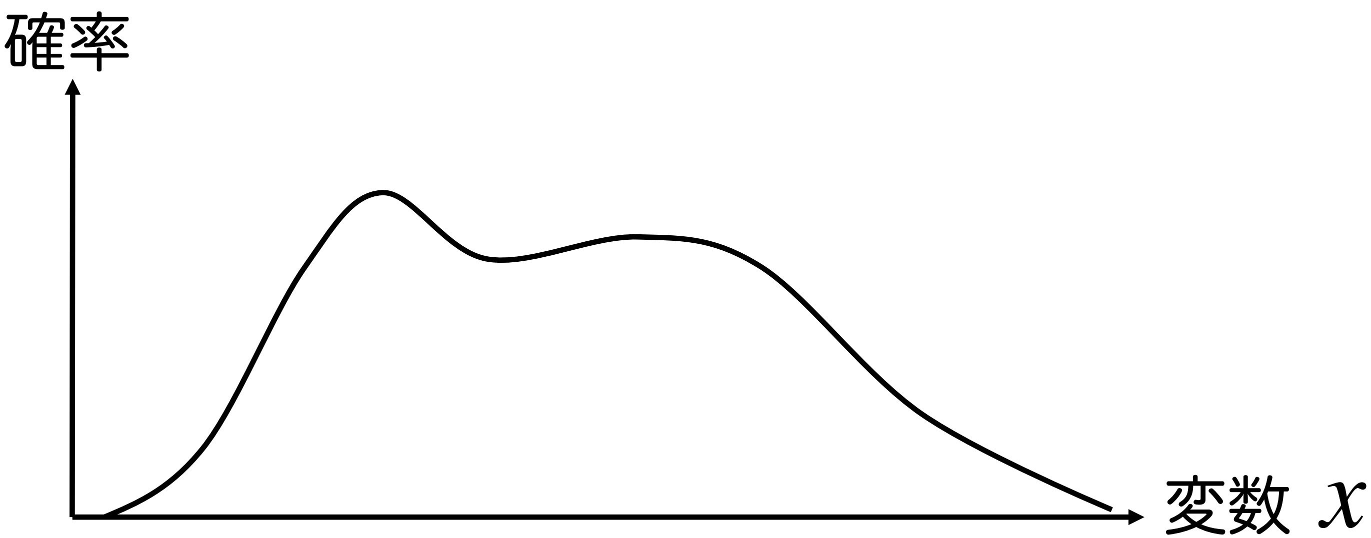
$$P(H \mid D) = \frac{P(D \mid H)\,P(H)}{P(D)}$$

事前確率・尤度・事後確率



事前確率・事後確率・尤度

- データや仮説が複数あるとき、事前確率、事後確率、尤度はデータと仮説の値に対応した確率値をとる



確率変数と確率分布

- 値の生じ易さに確率値を与えられるような変数を確率変数という
- 確率変数とそれに対応する確率値を確率変数
- 対応関係をまとめた表を確率分布表という

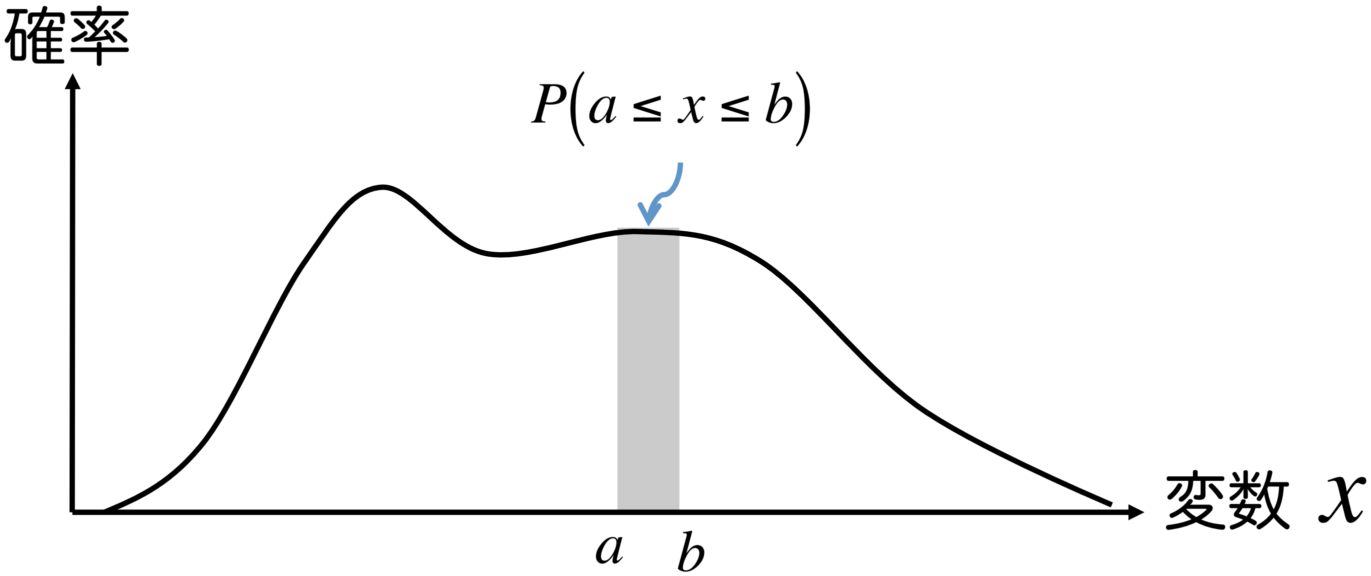
コイントス (確率変数)	表	裏
確率	1/2	1/2

サイコロの目 (確率変数)	1	2	3	4	5	6
確率	1/6	1/6	1/6	1/6	1/6	1/6

確率分布の種類

- 離散的な確率分布
  - 一様分布
  - 二項分布
  - ポアソン分布
- 連続的な確率分布
  - 正規分布
  - ベータ分布
  - ガンマ分布

連続的な確率分布



- 変数が連続値をとるときなどは確率変数の分布を確率密度関数で表現する

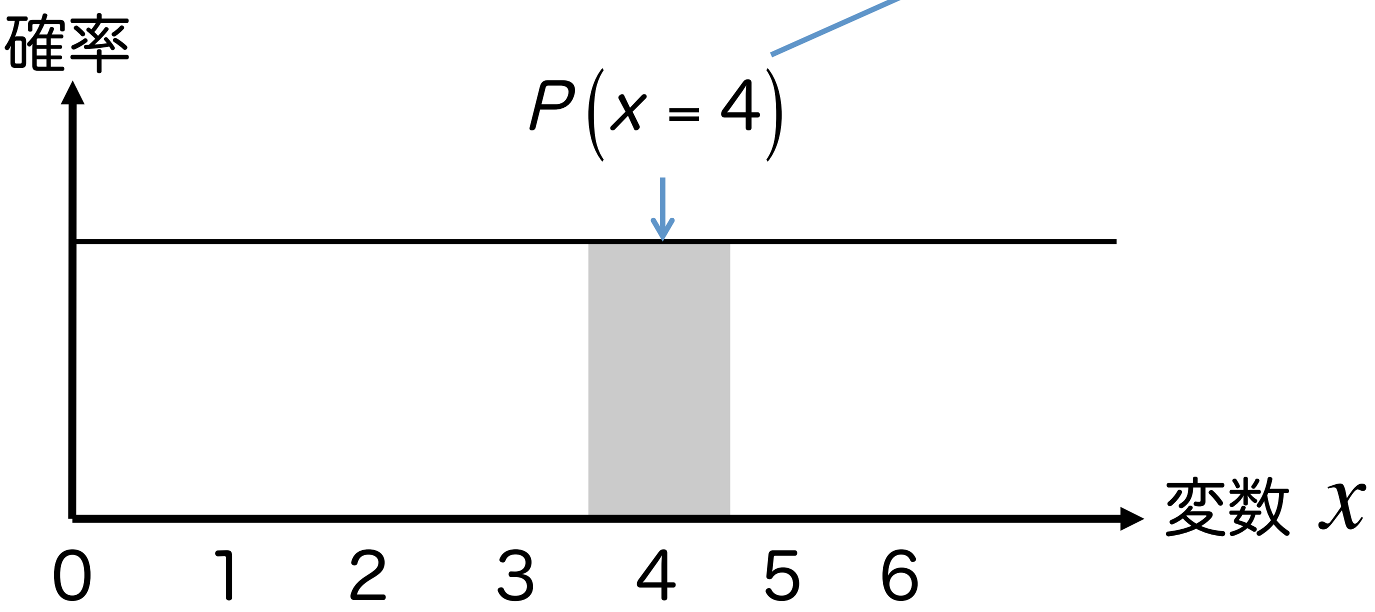
確率分布の平均・分散・標準偏差

確率変数X	x_1	x_2	x_3	…	x_n
確率p	p_1	p_2	p_3	…	p_n

- 平均(期待値) $\mu = x_1p_1 + x_2p_2 + \cdots + x_np_n$
- 分散 $\sigma^2 = (x_1 - \mu)p_1 + (x_2 - \mu)p_2 + \cdots + (x_n - \mu)p_n$
- 標準偏差 σ

サイコロの目の確率分布

サイコロの目 (確率変数)	1	2	3	4	5	6
確率	1/6	1/6	1/6	1/6	1/6	1/6



サイコロの目の確率分布

サイコロの目 (確率変数)	1	2	3	4	5	6
確率	1/6	1/6	1/6	1/6	1/6	1/6

- 平均(期待値) $\mu = 1 \times 1/6 + 2 \times 1/6 + \cdots + 6 \times 1/6 = 3.5$
- 分散 $\sigma^2 = (1-3.5)^2 \times (1/6) + (2-3.5)^2 \times (1/6) + \cdots + (6-3.5)^2 \times (1/6) = 35/12 \fallingdotseq 2.9$
- 標準偏差 $= \sigma \fallingdotseq 1.7$

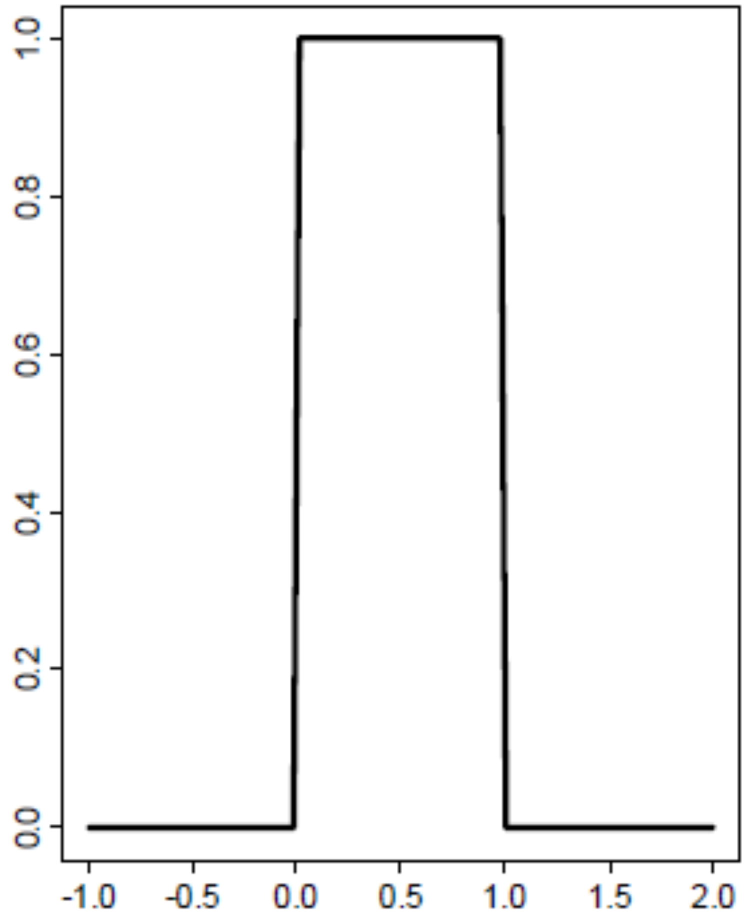
クイズ（1）

- 以下の確率分布表をもとに、平均・分散・標準偏差を計算しなさい

確率変数X	1	2	3	4	5
確率p	1/10	2/10	4/10	2/10	1/10

一様分布

- サイコロの目のように、どの変数に対しても同じ確率値をとる確率分布を一様分布という



コイントスの確率分布

- コイントスで表が出るか裏が出るかという試行は、ある行為が成功するか失敗するかという試行ととらえることができる

コイントス (確率変数)	表	裏
確率	1/2	1/2

ベルヌーイ試行と二項分布

- 経験的にコイントスをして表が出る確率がpだったとする
- １回目：コイントスで表とがでるのは${}_1C_1$回と${}_1C_0$回。従って表と裏が出る期待値は${}_1C_1$xp及び${}_1C_0$x(1-p)となる
- ２回目：コイントスで２回連続表が出る、１回目表(裏)で２回目裏(表）が出る、２回連続裏が出る期待値は、それぞれ${}_2C_2$xp^2、${}_2C_1$xp(1-p)、${}_2C_0$x$(1-p)^2$
- ３回目：…

コイントスとパスカルの三角形

表 ${}_{1}C_{1}xp$　　裏 ${}_{1}C_{0}x(1-p)$

表表 ${}_{2}C_{2}xp^{2}$　　表裏 ${}_{2}C_{1}xp(1-p)$　　裏裏 ${}_{2}C_{0}x(1-p)^{2}$

表表表 ${}_{3}C_{3}xp^{3}$　　表表裏 ${}_{3}C_{2}xp^{2}(1-p)$　　表裏裏 ${}_{3}C_{1}xp\,(1-p)^{2}$　　裏裏裏 ${}_{3}C_{0}x\,(1-p)^{3}$

ベルヌーイ試行と二項分布

- 0か1かしかない試行において、n回の試行でs回成功し、その期待値pがわかっているとき、実験が成功する確率は以下のベルヌーイ試行に従う
- ベルヌーイ試行の確率分布を二項分布といい、その分布は次式の確率密度関数に従う

$$Binom(n,p) = {}_nC_s \cdot p^n \cdot (1-p)^{n-s}$$

$$\approx p^n \cdot (1-p)^{n-s}$$

ベルヌーイ試行と二項分布

- ベルヌーイ試行の確率変数と確率値を下表のように与えた時、次式が成立する
- 平均μ=p
- 分散σ^2=p(1-p)

確率変数X	0	1
確率	1-p	p

二項分布と尤度関数

- 尤度関数がベルヌーイ試行に従うとする

$$L(p) = {}_nC_s p^s (1-p)^{n-s}$$

- 成功する確率がpで、全体で5回（=n）の試行中、3回（=s）成功する実験での尤度

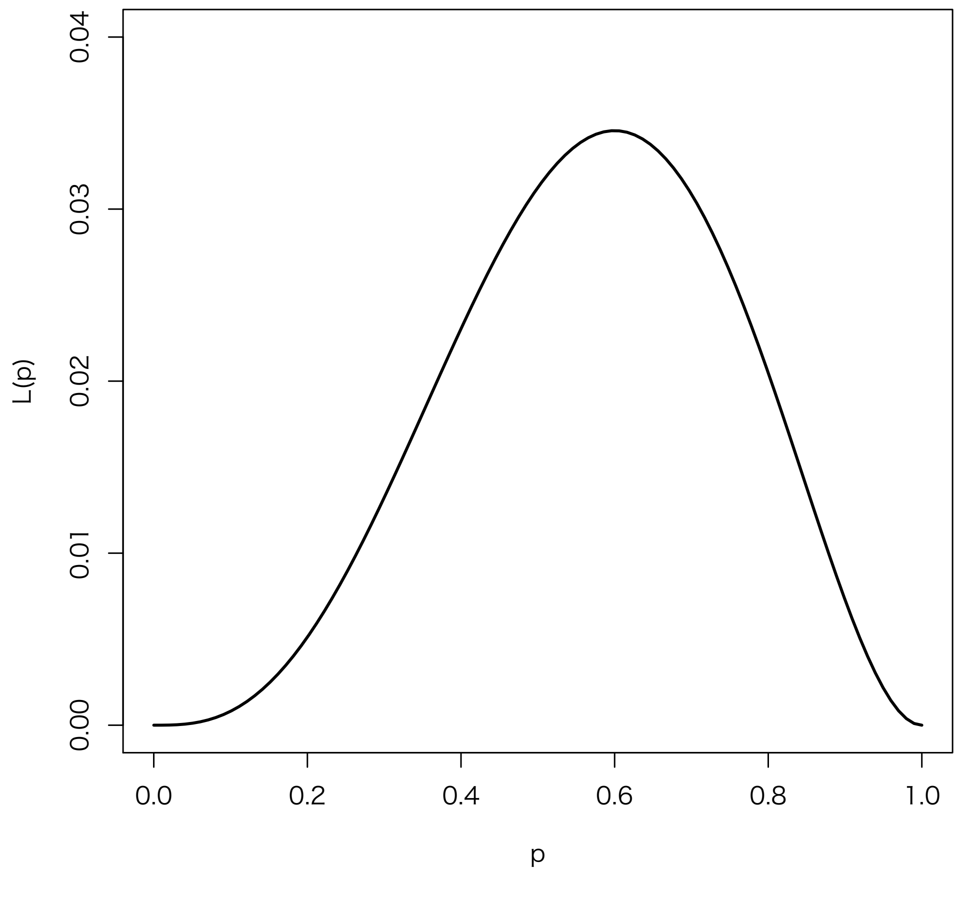
$$L(p) = {}_5C_3 p^3 (1-p)^2 \approx p^3 (1-p)^2$$

- 対数をとると便利

$$\log L(p) \approx \log\left[p^3 (1-p)^2\right] = 3\log p + 2\log(1-p)$$

尤度関数の計算例

- p=0.6で尤度関数L(p)は最大値0.03456となる



ベータ分布

- 二項分布を式変形すると次式のベータ関数に従うベータ分布となる

$n = \alpha - 1,\ n - s = \beta - 1$ とすると、

$$Binom\left(n, p\right) \approx p^{n} \cdot \left(1 - p\right)^{n-s}$$

$$= \frac{\left(\alpha - 1\right)!\left(\beta - 1\right)!}{\left(\alpha + \beta - 1\right)!}$$

$$B\left(\alpha, \beta\right) = k \cdot p^{\alpha-1} \cdot \left(1 - p\right)^{\beta-1}$$

$$0 < p < 1,\ 0 < \alpha,\ 0 < \beta$$

ベータ分布

- ベータ分布

$$B(\alpha,\beta) = k \cdot p^{\alpha-1} \cdot (1-p)^{\beta-1}$$

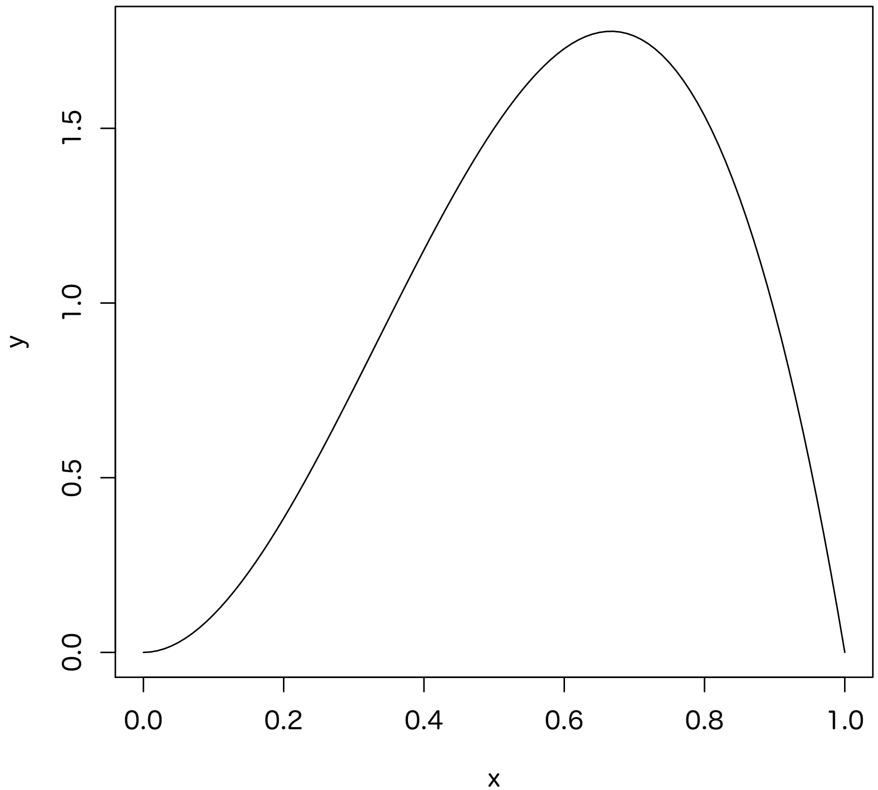
$$0 < p < 1,\ 0 < \alpha,\ 0 < \beta$$

- については、次式が成立する

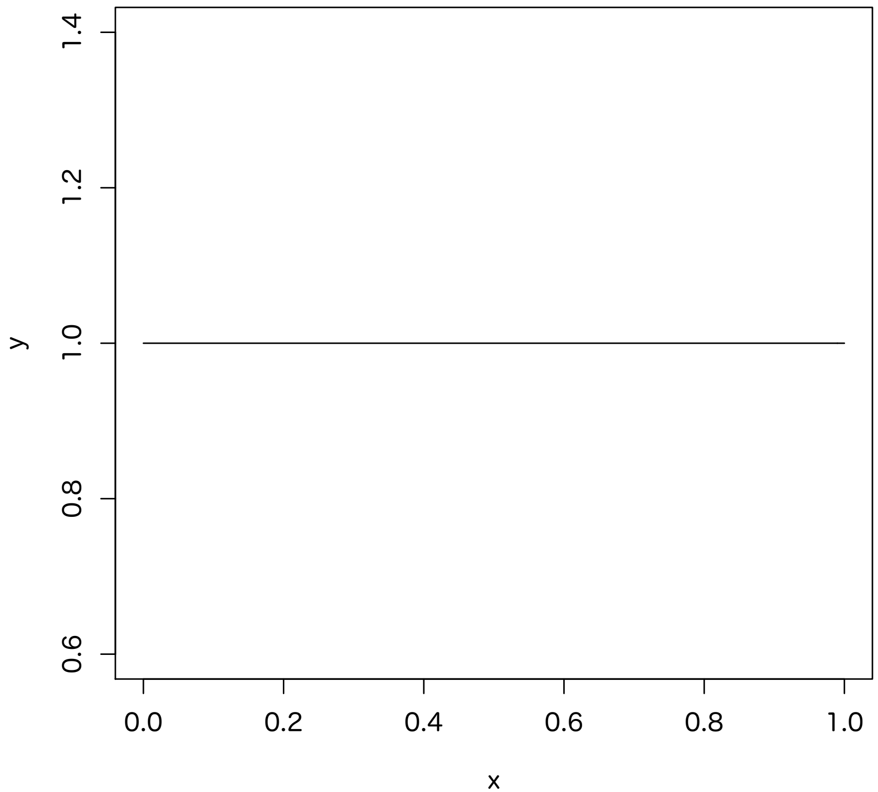
$$\mu = \frac{\alpha}{\alpha+\beta},\ \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

ベータ分布

B(3, 2)

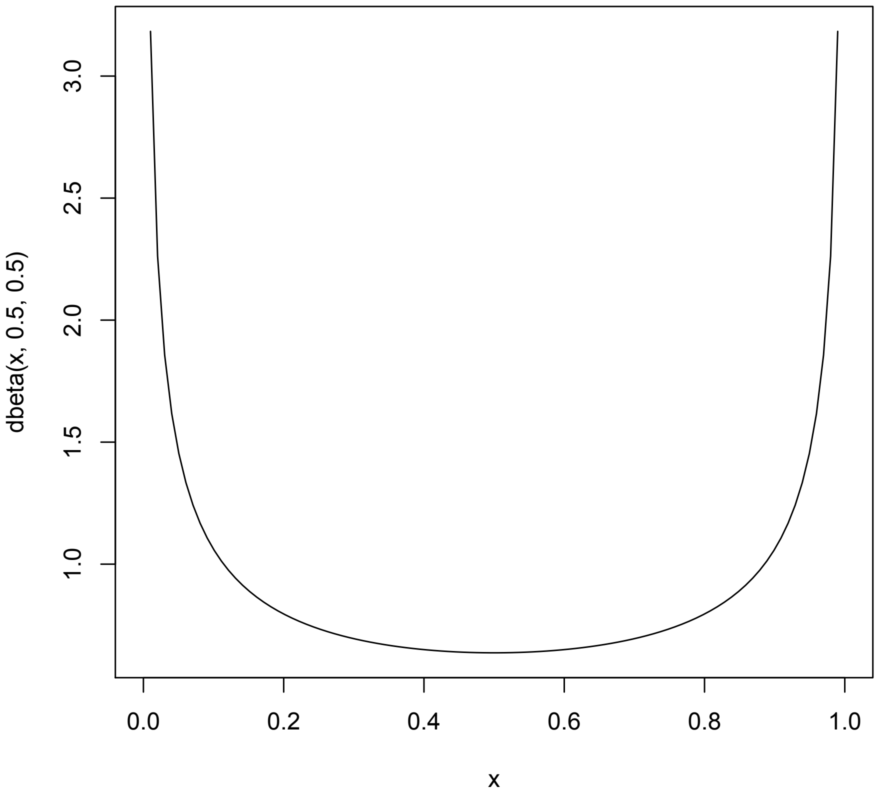


B(1, 1)=一様分布

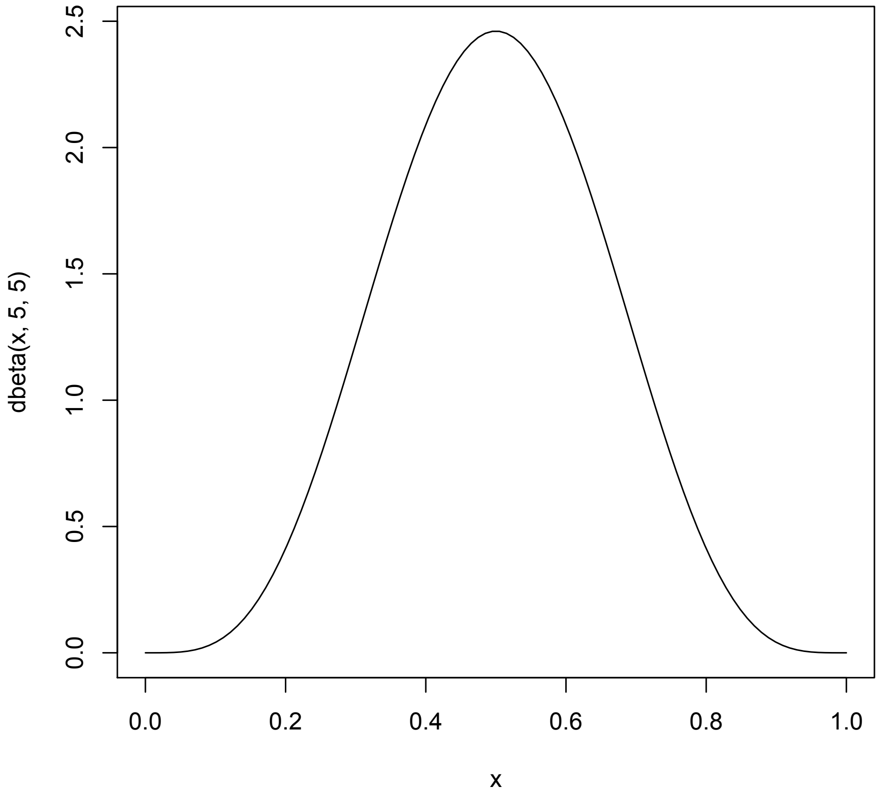


ベータ分布

B(0.5, 0.5)



B(5, 5)

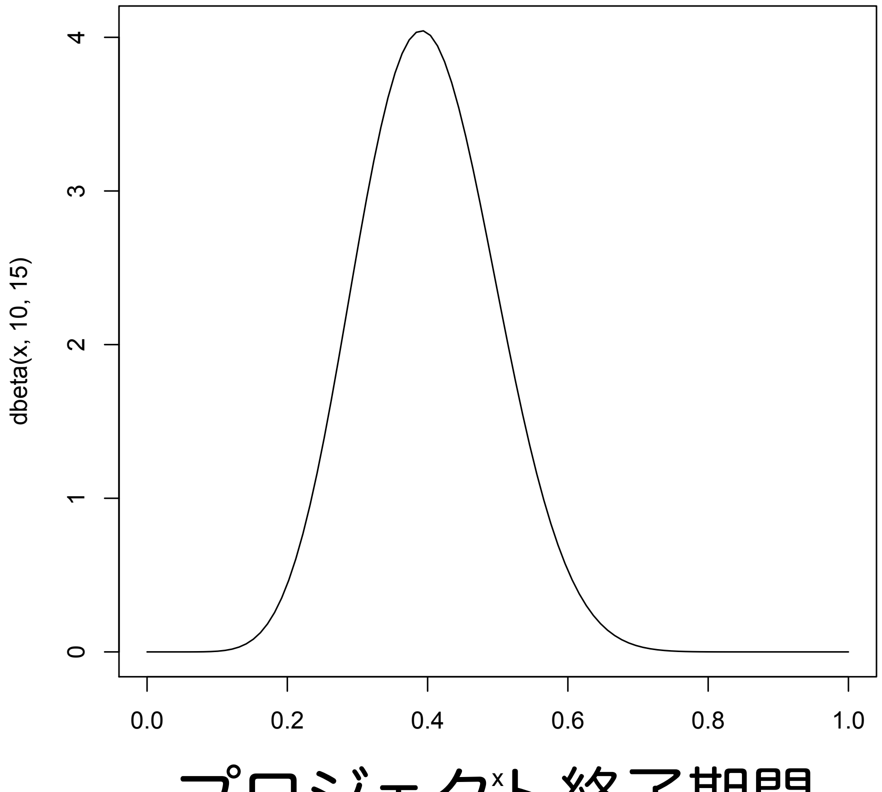


ベータ分布は万能な確率分布

- ベータ分布は以下の分布に変化できる
  - 一様分布、線形分布
  - 単調増加・単調減少分布
  - 単峰分布
  - 左右対称分布

ベータ分布

- 活用例は数少ないが、プロジェクトの時間管理などに用いられることがある
- あるプロジェクトがある期間内に終わることもあれば（確率は低いが）終了までとても時間がかかることがある



プロジェクト終了期間

ベータ関数とガンマ関数

- ベータ関数の確率密度関数は以下のように式変形できる

$$B(\alpha,\beta) \approx p^{\alpha-1} \cdot (1-p)^{\beta-1}$$

$$= \frac{(\alpha-1)!(\beta-1)!}{(\alpha+\beta-1)!}$$

$$= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

クイズ（2）

- 以下の値を求めなさい

1) Binom(n=10,p=0.4)
2) B(k=1,p=0.5,α=3,β=2)
3) B(α=5,β=4)の平均・分散

ガンマ関数

- ガンマ関数は次式で表される

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx$$

$$\alpha > 0$$

- ガンマ関数は以下のような性質を持つ

$$\Gamma(\alpha) = (\alpha - 1)!$$

$$\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$$

$$\Gamma(1/2) = \sqrt{\pi}$$

ガンマ分布

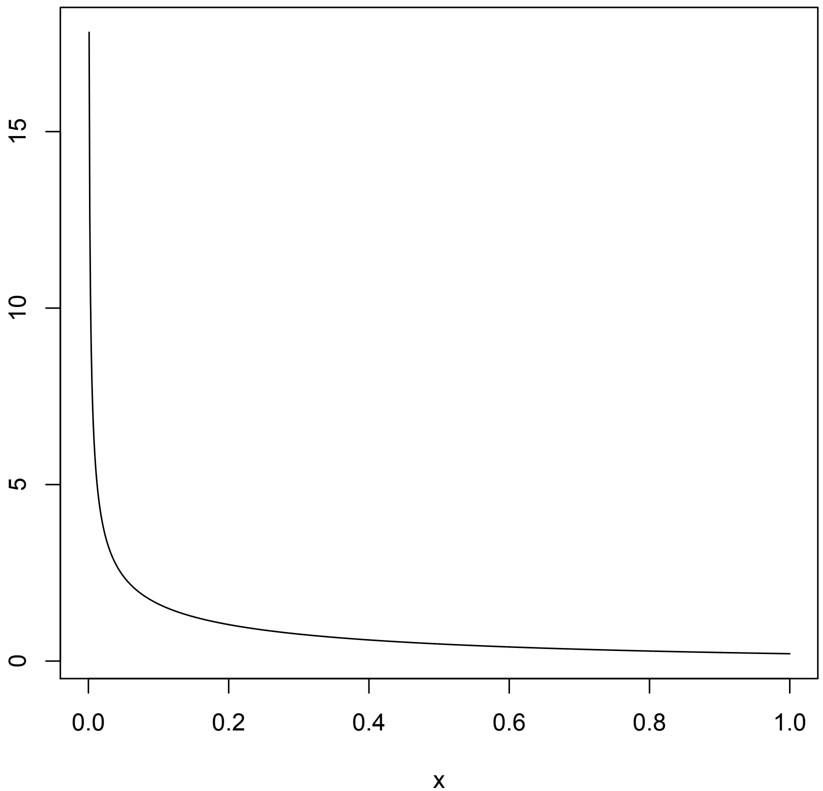
- ガンマ分布の確率密度関数とその性質は次式のようになる

$$f(x) = Ga(\alpha, \lambda) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x},\ x \geq 0$$

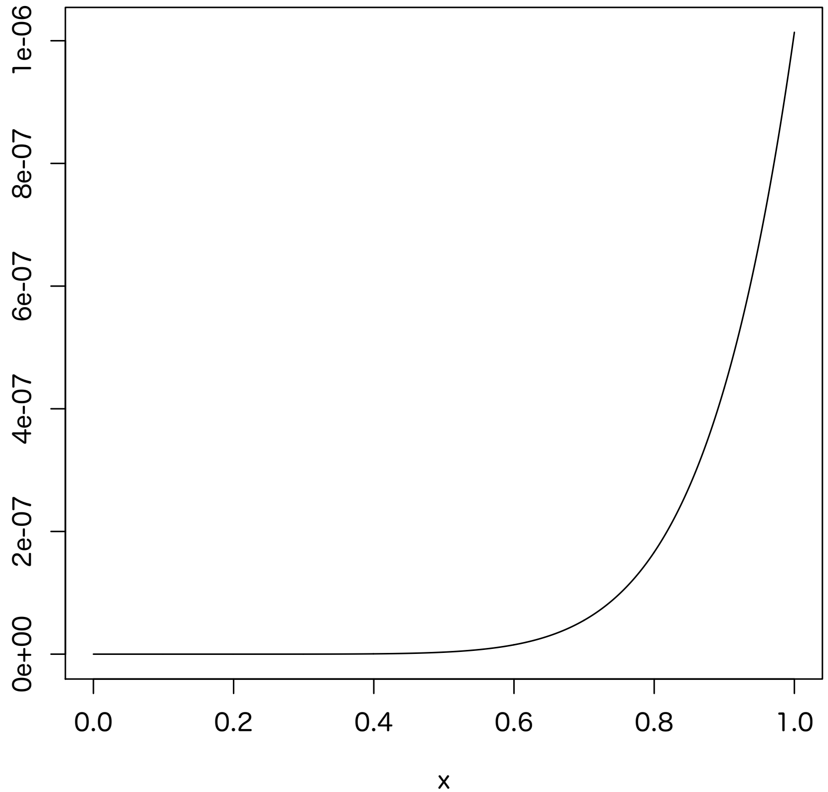
$$E(X) = \frac{\alpha}{\lambda},\ V(X) = \frac{\alpha}{\lambda^2}$$

ガンマ分布

Γ(0.5)



Γ(10)



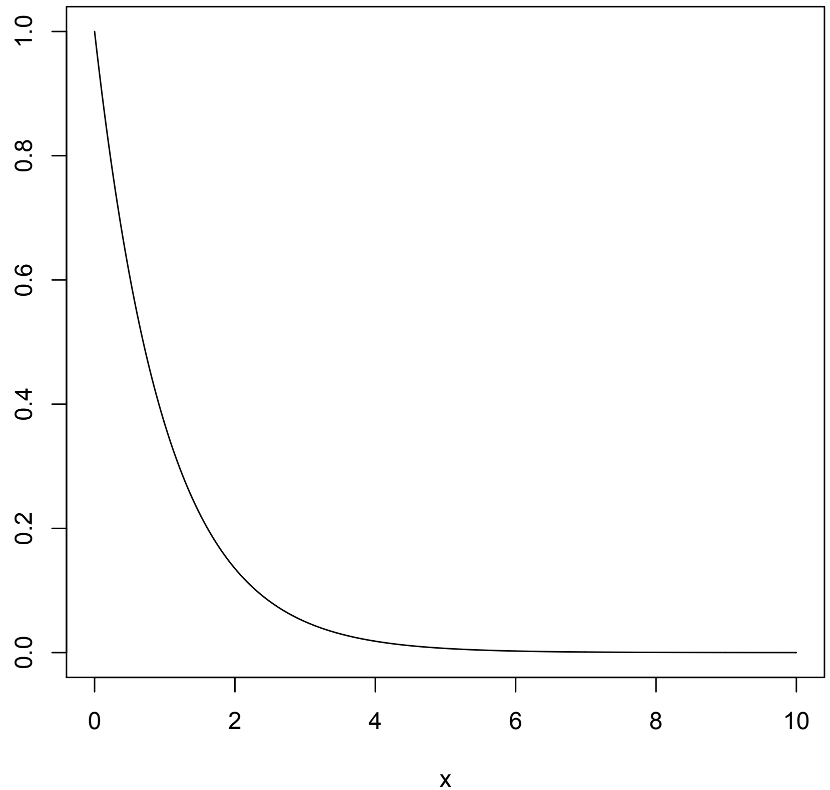
指数分布

- ガンマ分布の確率密度関数をα=1とすると指数関数となる

$$f(x) = \lambda e^{-\lambda x},\ x \geq 0$$

$$E(X) = \frac{1}{\lambda},\ V(X) = \frac{1}{\lambda^2}$$

指数分布



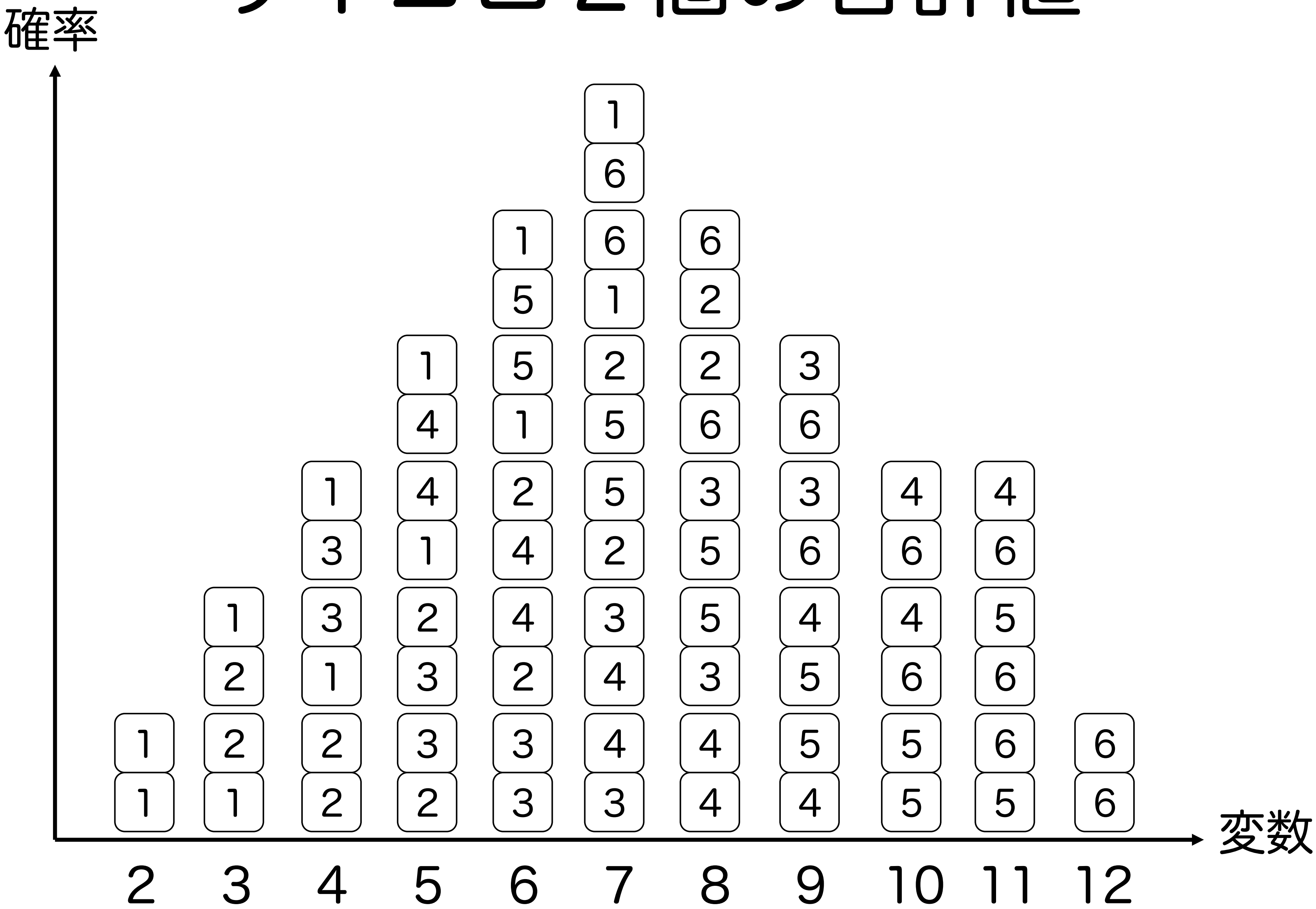
正規分布（ガウス分布）

- 平均μ、分散σ^2となる以下の確率密度関数に従う分布を正規分布という

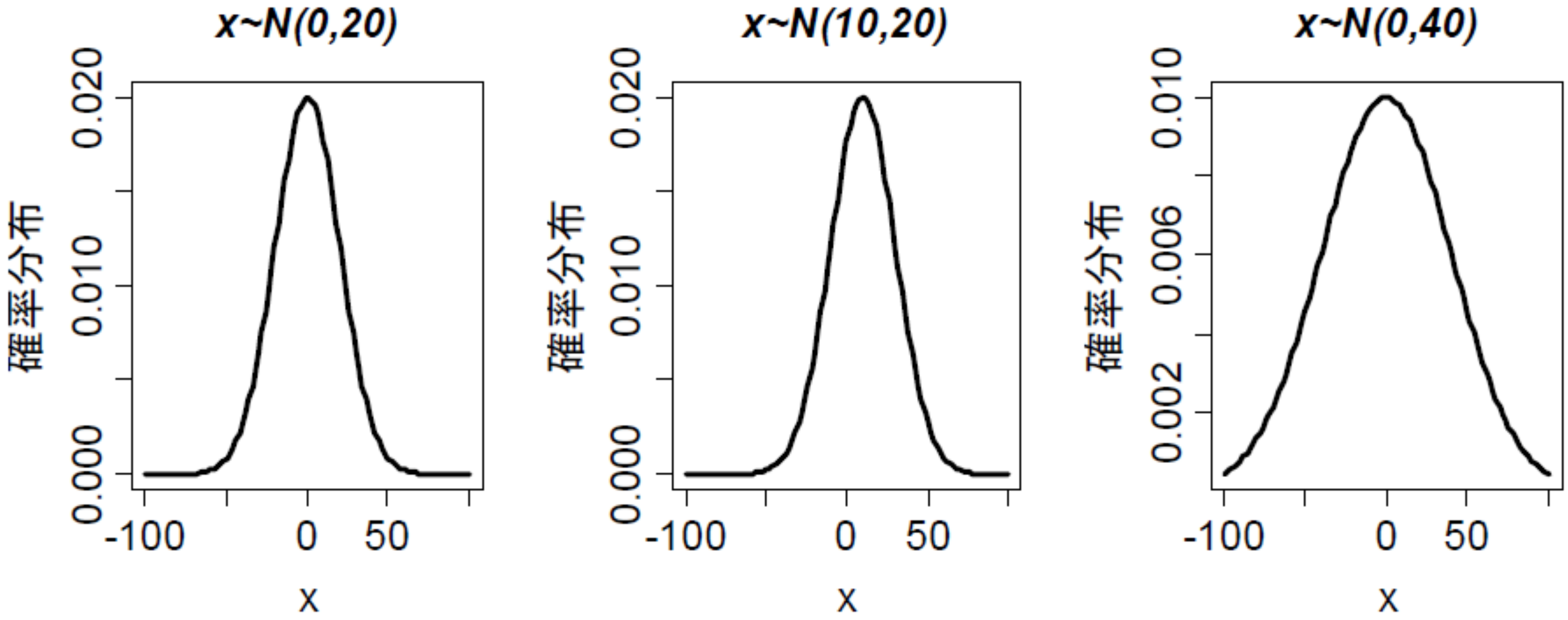
$$N\left(\mu,\sigma^2\right)=f\left(x\right)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{\left(x-\mu\right)^2}{2\sigma^2}\right)$$

$$E\left(Z\right)=\mu,\ V\left(Z\right)=\sigma^2$$

サイコロ２個の合計値



正規分布



x~N(0,20)
0.020
0.010
0.000
確率分布
-100
0
50
x
x~N(10,20)
0.020
0.010
0.000
確率分布
-100
0
50
x
x~N(0,40)
0.010
0.006
0.002
確率分布
-100
0
50
x

標準正規分布

- 平均μ=0、分散σ^2=1となる正規分布
- 標準化 $z=(x-\mu)/\sigma$
- 以下の確率密度関数に従う

$$N(0,1)=f(z)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right)$$

$$E(Z)=0,\ V(Z)=1$$

二項分布と正規分布

- 二項分布 $Binom(n,p) \approx p^n \cdot (1-p)^{n-s}$

- 二項分布の試行回数nが十分に大きい時、正規分布N(np,np(1-p))に近似できる

$$N\left(np, np(1-p)\right)$$

$$= \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left(-\frac{(x-np)^2}{2np(1-p)}\right)$$

ガンマ分布と正規分布

- ガンマ分布Ga(1/2,1/2)は標準正規分布（平均0,分散1の正規分布）に従う変数の二乗値z^2が従う分布となる

$$Ga\left(\frac{1}{2},\frac{1}{2}\right)=\frac{1}{\sqrt{2\pi x}}e^{-\frac{x}{2}}$$

$$x=z^2$$ とすると、

$$=\frac{1}{\sqrt{2\pi z^2}}\exp\left(-z^2/2\right)$$

ガンマ分布と正規分布

- ガンマ分布の平均と分散で確率変数を正規化する

$$z = \frac{x - (\alpha/\lambda)}{\sqrt{\alpha/\lambda^2}}$$

- $\alpha \to \infty$とするとガンマ分布は正規分布に近似できる