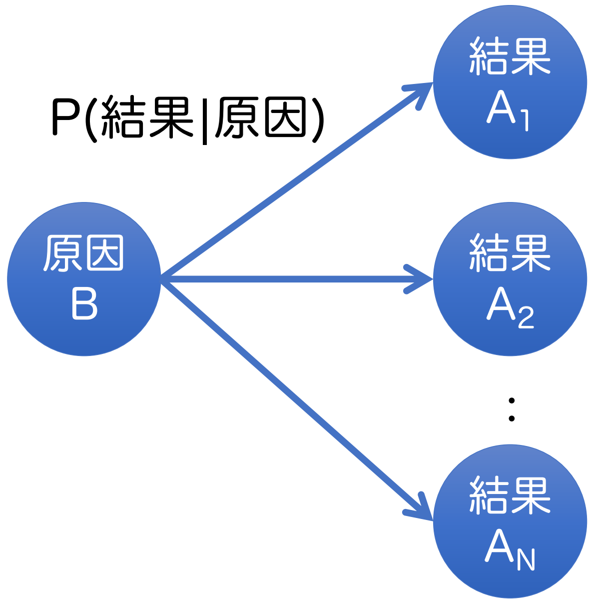
ベイズ統計

古谷知之

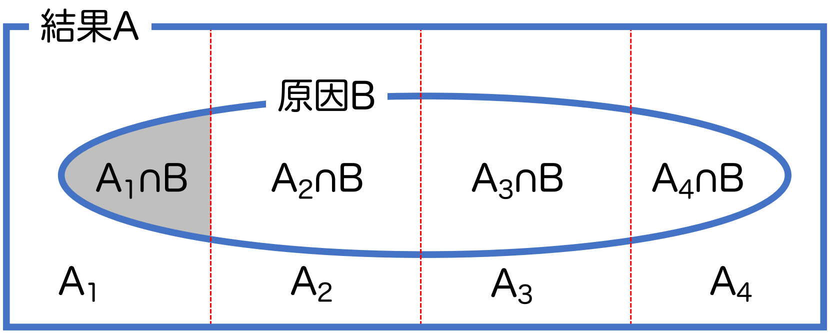
講義概要

- 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）

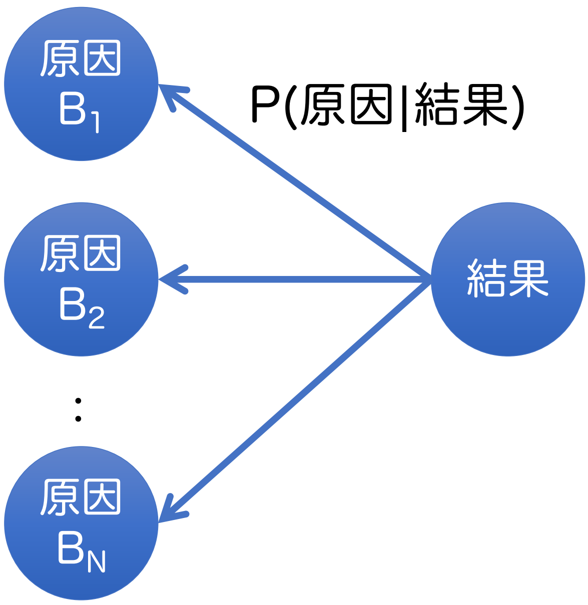
- 原因Bから結果A_1が生じうる確率

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

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

$$= \frac{P(B|A_1)P(A_1)}{P(B)}$$

逆確率による原因の推定



逆確率（結果→原因）

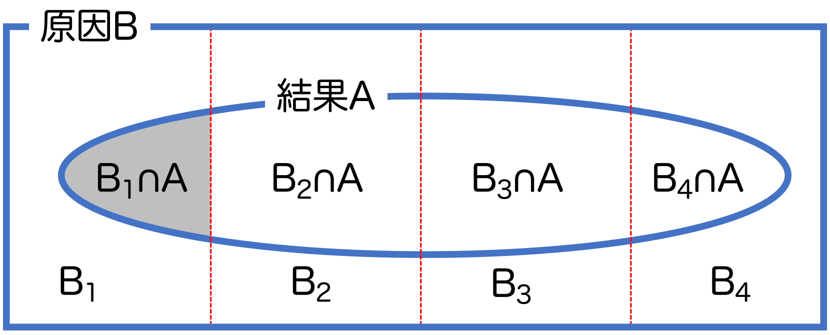
- 結果Aの原因BがB_1でありうる確率

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{\sum_{i=1}^{n} P(A \cap B_i)}$$

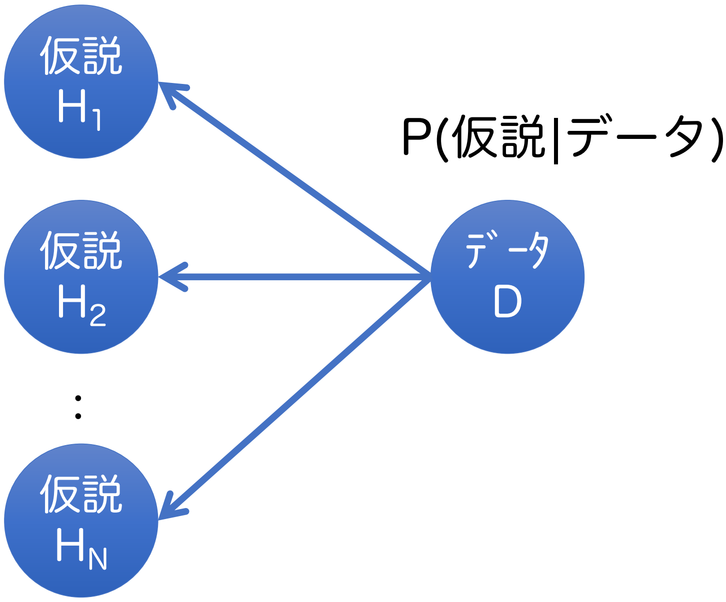
$$= \frac{P(A|B_1)P(B_1)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)}$$

$$= \frac{P(A|B_1)P(B_1)}{P(A)}$$

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

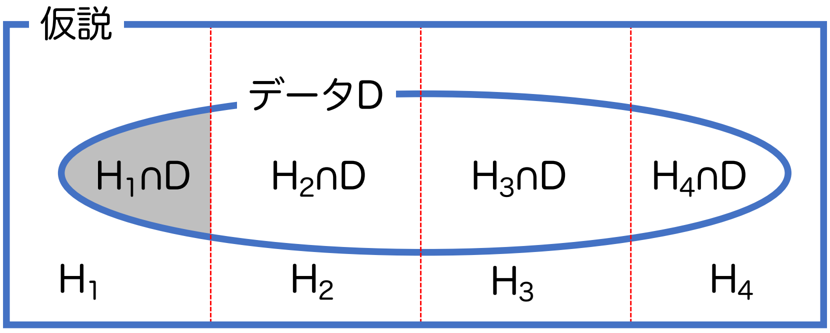
- データDから仮説H_1が成立する確率は以下のようになる

$$P(H_1|D) = \frac{P(D|H_1)P(H_1)}{\sum_{i=1}^{n} P(D \cap H_i)}$$

$$= \frac{P(D|H_1)P(H_1)}{\sum_{i=1}^{n} P(D|H_i)P(H_i)}$$

$$= \frac{P(D|H_1)P(H_1)}{P(D)}$$

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

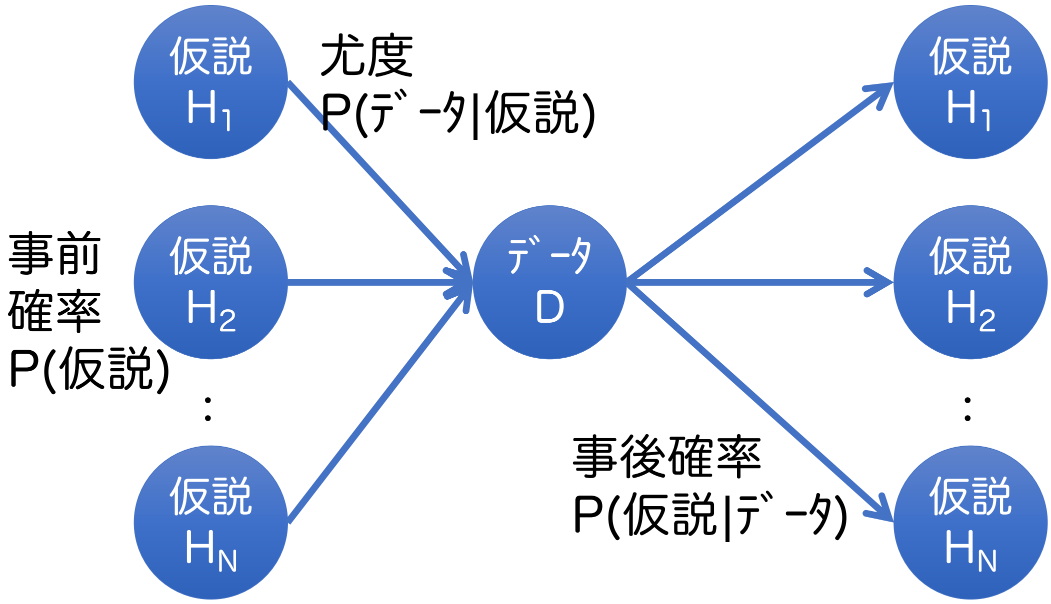
- 仮説がデータに当てはまる当てはまりやすさ（尤もらしさ）を尤度という

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

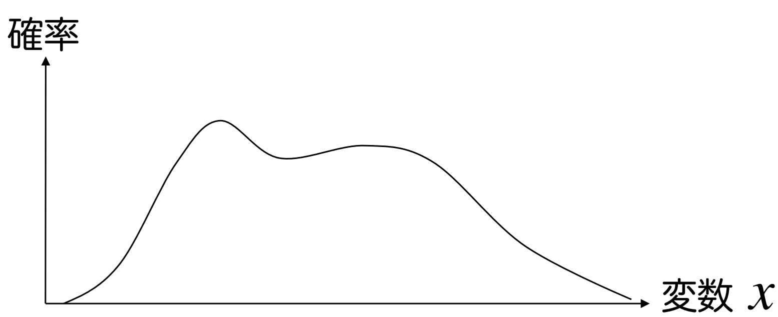
$$\overset{\text{事後確率}}{P(H|D)} = \frac{\overset{\text{尤度}}{P(D|H)}\overset{\text{事前確率}}{P(H)}}{P(D)}$$

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

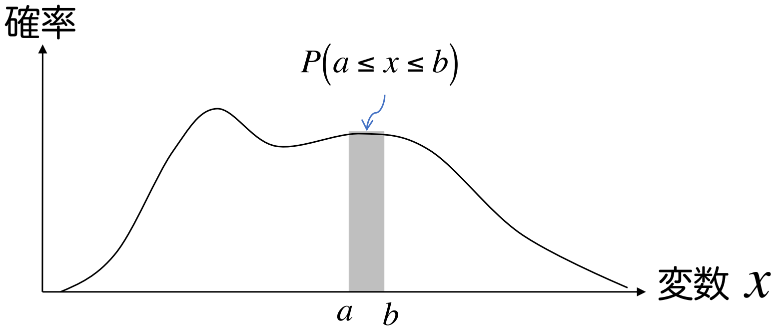
コイントス（確率変数）	表	裏
確率	1/2	1/2

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

連続的な確率変数と確率分布



- 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

離散型の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

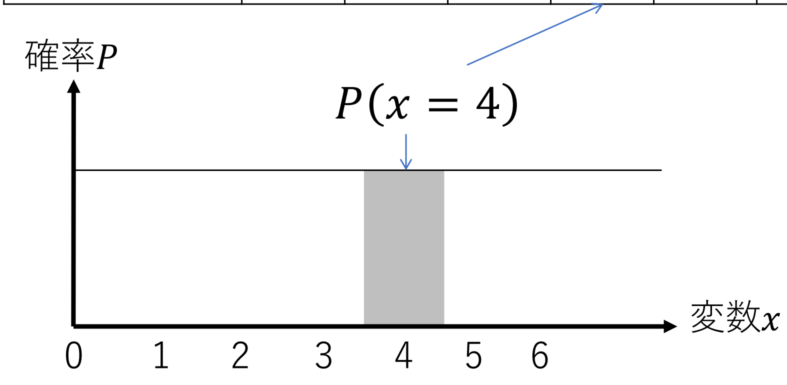
離散型確率分布の平均・分散・標準偏差

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

- 確率変数Xの値x_iと確率p_iが上表のように分布しているとき、確率変数Xの平均μ、分散σ^2 、標準偏差σは以下のようになる
- 平均(期待値)$\mu = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n$
  →確率変数がどの辺りに集中しているかを示す
- 分散$\sigma^2 = (x_1 - \mu)^2 p_1 + (x_2 - \mu)^2 p_2 + \cdots + (x_n - \mu)^2 p_n$
- 標準偏差σ
  →分散と標準偏差は、確率変数がどの程度散らばっているかを示す

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



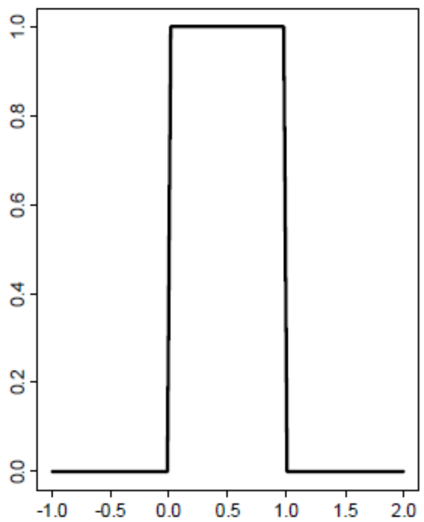
サイコロの目の確率分布

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$\sigma^2 = (1-3.5)^2\times(1/6) + (x_2-\mu)^2\times(1/6) + \cdots + (x_n-\mu)^2\times(1/6) \cong 2.9$
- 標準偏差$\sigma \cong 1.7$

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
- 区間[0, 1]での一様分布は右図のようになる



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：コイントスで表と裏がでるのは ${}_1C_1$ 回と ${}_1C_0$ 回。従って表と裏が出る期待値は ${}_1C_1 \times p$及び ${}_1C_0 \times (1-p)$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 ${}_2C_2 \times p^2$ ${}_2C_2 \times p^2$ 、 ${}_2C_1 \times p(1-p)$、 ${}_2C_2 \times (1-p)^2$
- ３回目：…

コイントスとパスカルの三角形

表 　　裏

${}_1C_1\ \times p$ 　 ${}_1C_0\ \times(1-p)$

表表 　　表裏 　　裏裏

${}_2C_2\ \times p^2$ 　 ${}_2C_1\ \times p(1-p)$ 　 ${}_2C_0\ \times(1-p)^2$

表表表 　　表表裏 　　表裏裏 　　裏裏裏

${}_3C_3\ \times p^3$ 　 ${}_3C_2\ \times p^2(1-p)$ 　 ${}_3C_1\ \times p(1-p)^2$ 　 ${}_3C_0\ \times(1-p)^3$

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、その確率pがわかっているとき($s = n\times p$)、実験が成功する期待値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

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

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：$\mu = np$
- 分散：$\sigma^2 = np(1-p)$

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$L(p)$がベルヌーイ試行に従うとすると、

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

- 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は、以下のようになる

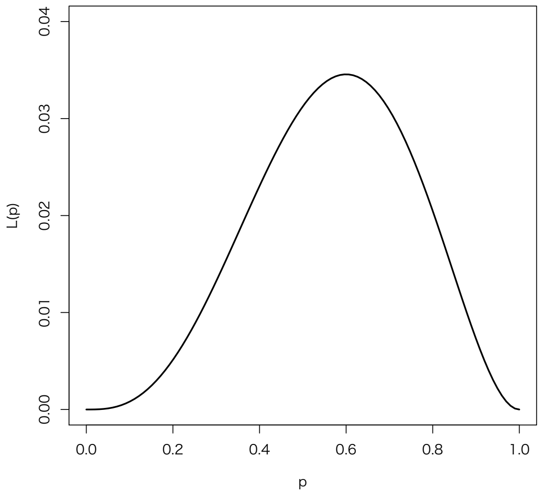
$$L(p) = {}_5C_3 \cdot p^3 \cdot (1-p)^2 \propto p^3 \cdot (1-p)^2$$

- 対数尤度関数は以下のようになる

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

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



二項分布

- $s = \alpha - 1, n - s = \beta - 1$とすると、二項分布$Binom(n, p)$は次式のように変形できる

$$
\begin{aligned}
Binom(n, p) &\propto p^s \cdot (1-p)^{n-s} \\
&= p^{\alpha-1} \cdot (1-p)^{\beta-1} \\
&= \frac{(\alpha-1)!\,(\beta-1)!}{(\alpha+\beta-1)!}
\end{aligned}
$$

ベータ分布

- 二項分布の式変形の結果から、ベータ分布の確率密度関数 $f(\alpha,\beta,p)$ を関係づけることができる

$$f(p;\alpha,\beta) = Beta(\alpha,\beta) = k\cdot p^{\alpha-1}\cdot(1-p)^{\beta-1}$$
$$= k\cdot\frac{(\alpha-1)!\,(\beta-1)!}{(\alpha+\beta-1)!}$$

$$0 < p < 1, 0 < \alpha, 0 < \beta$$
$$k = \frac{1}{\int_0^1 p^{\alpha-1}\cdot(1-p)^{\beta-1}dp}$$

ベータ分布

- ベータ分布 $Beta(\alpha,\beta)$ の平均 μ と分散 σ^2 は以下のようになる

$$\mu = \frac{\alpha}{\alpha+\beta}$$

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

ベータ関数

- 一般に、次式で表される関数をベータ関数$B(\alpha, \beta)$という

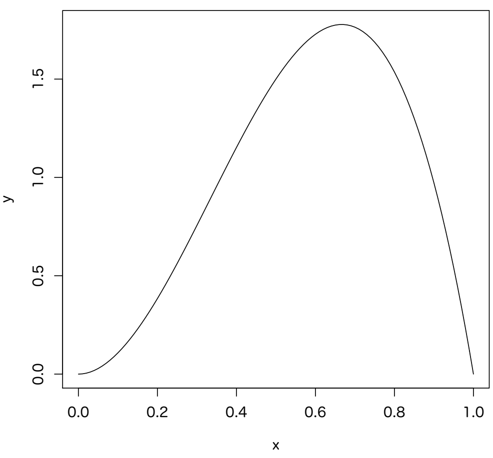
$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} \cdot (1-x)^{\beta-1} dx$$

- ベータ関数は以下の性質を持つ

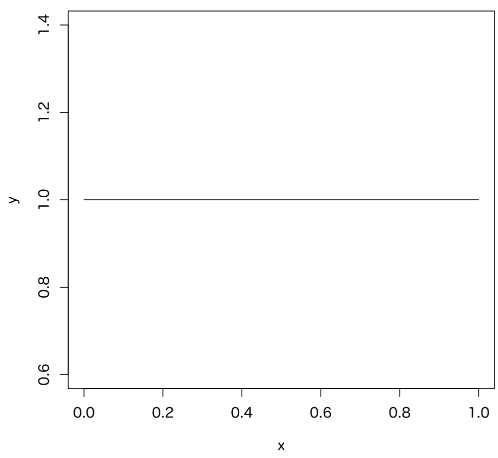
$$B(\alpha, \beta) = B(\beta, \alpha)$$
$$\alpha B(\alpha, \beta+1) = \beta B(\alpha+1, \beta)$$
$$B(\alpha, \beta) = B(\alpha, \beta+1) + B(\alpha+1, \beta)$$

ベータ分布

$Beta(3,2)$

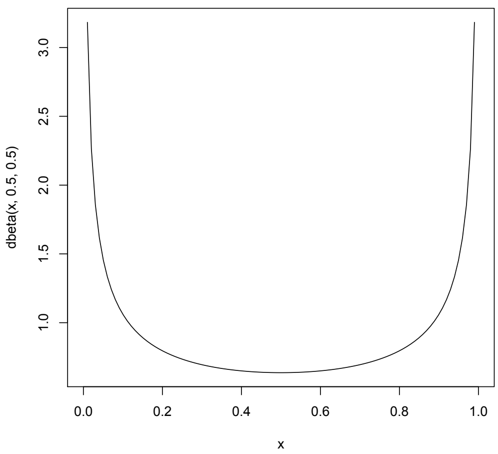


$Beta(1,1)$=一様分布

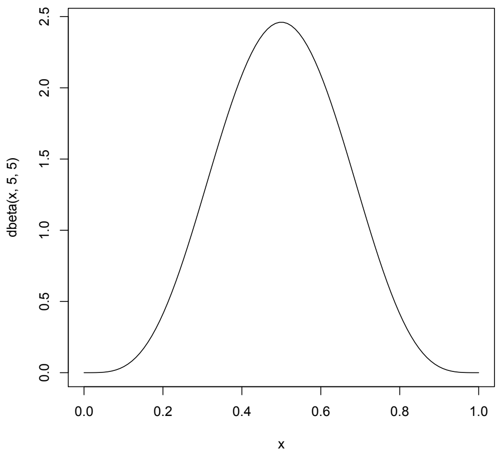


ベータ分布

$Beta(0.5, 0.5)$



$Beta(5, 5)$

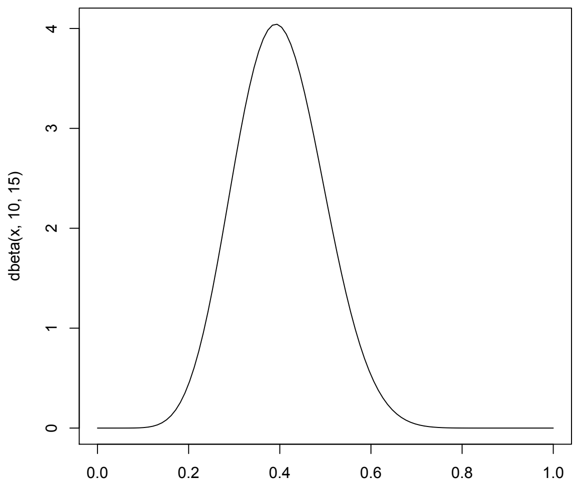


ベータ分布は万能な確率分布

- 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

ベータ分布の応用例

- 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

- ベータ関数は以下のように式変形できる

$$B(\alpha,\beta) = \int_0^1 p^{\alpha-1} \cdot (1-p)^{\beta-1} dp$$
$$= \frac{(\alpha-1)!\,(\beta-1)!}{(\alpha+\beta-1)!}$$
$$= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

- ここで$\Gamma(\alpha)$はガンマ関数という

ガンマ関数

- ガンマ関数$\Gamma(\alpha)$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$\Gamma(\alpha)$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$f(x)$は次式のようになる

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

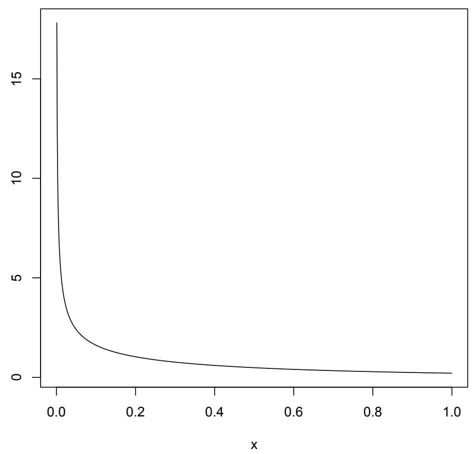
- ガンマ分布の平均$E(X)$と分散$V(X)$は以下の通り

$$E(X) = \frac{\alpha}{\lambda}$$

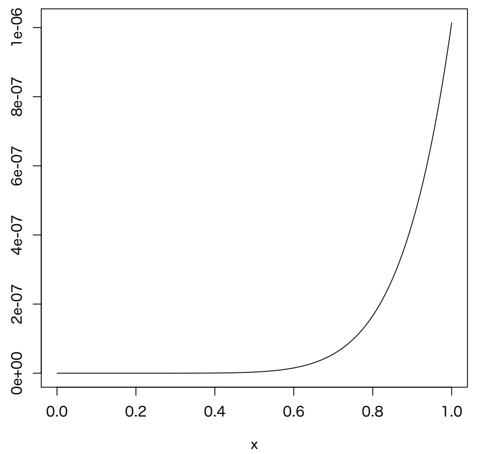
$$E(X) = \frac{\alpha}{\lambda^2}$$

ガンマ分布

$Ga(0.5, 1)$



$Ga(10, 1)$



指数分布

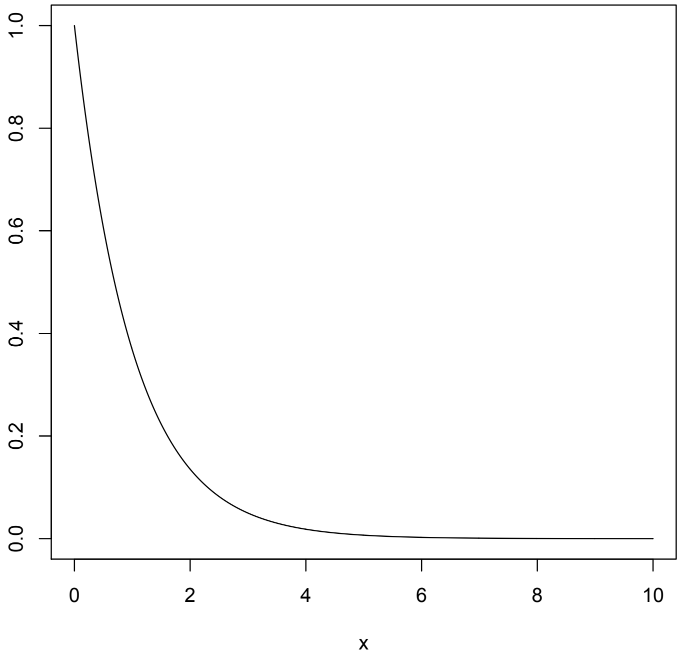
- ガンマ分布の確率密度関数について、$\alpha = 1$とすると指数関数となる

$$f(x) = \lambda e^{-\lambda x}, x \geq 1$$

$$E(x) = \frac{1}{\lambda}$$

$$V(x) = \frac{1}{\lambda^2}$$

指数分布



正規分布（ガウス分布）

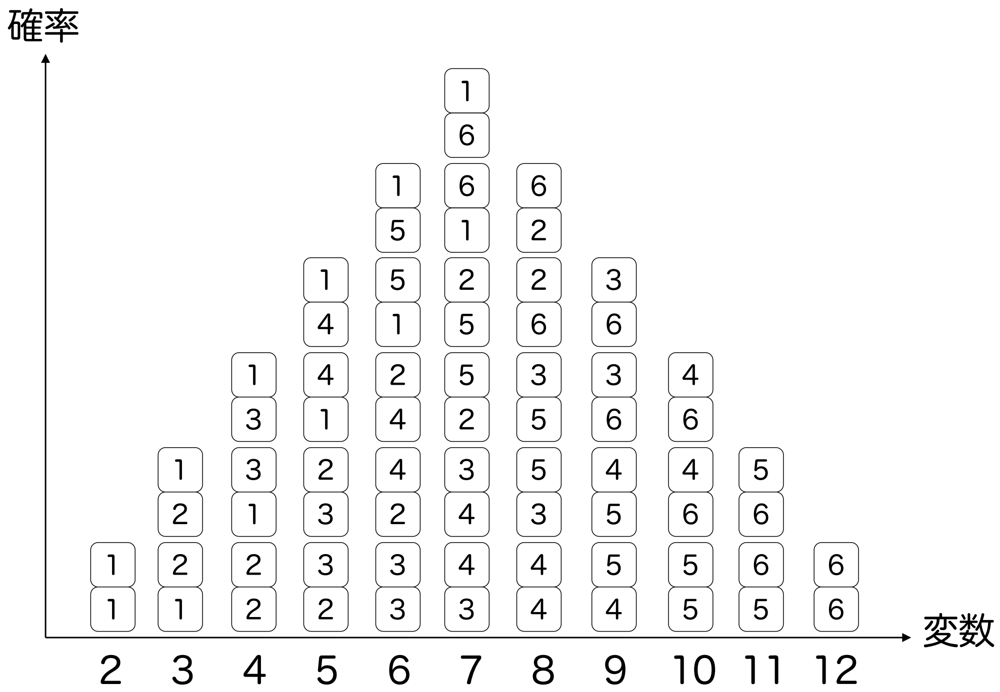
- 平均μ、分散σ^2となる確率変数xについて、以下の確率密度関数に従う分布を正規分布$N(\mu, \sigma^2)$という

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

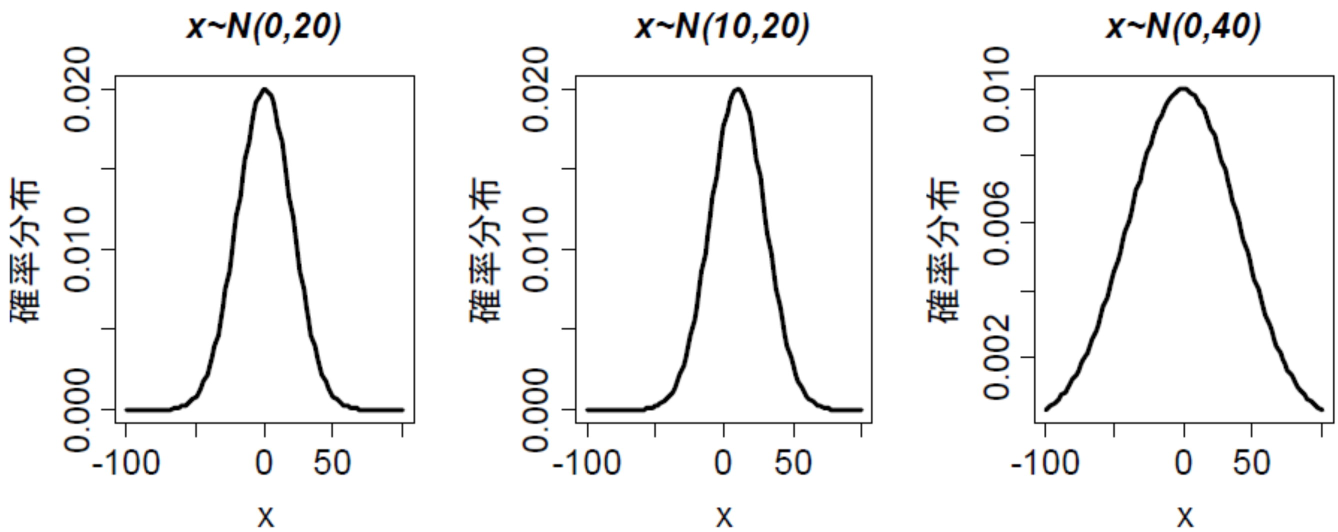
- 正規分布の平均$E(X)$と分散$V(X)$は以下のとおり

$$E(X) = \mu$$
$$V(X) = \sigma^2$$

サイコロ２個の合計値



正規分布



標準正規分布

- 確率変数xを次式により標準化する
$$z = (x - \mu)/\sigma$$
- このとき確率変数zは、平均$\mu = 0$、分散$\sigma^2 = 1$となる標準正規分布$N(0,1)$となる

- 標準正規分布$N(0,1)$の確率密度関数$f(z)$は以下のように表せる

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

二項分布と正規分布

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

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

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

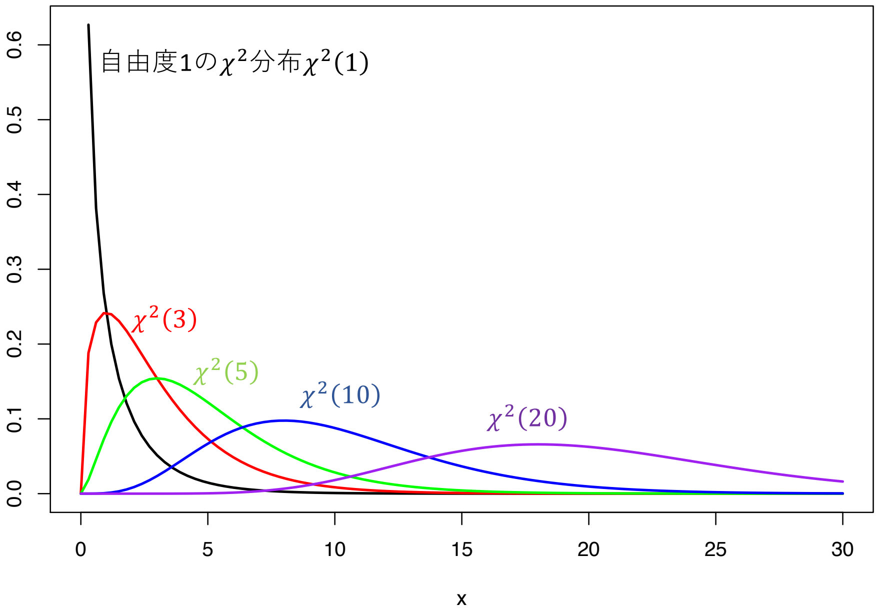
χ^2分布

- 確率変数X_iが平均μ、分散σ^2の正規分布に従う$X_i \sim N(\mu, \sigma^2)$とき、その標準化された値$Z_i = \frac{X_i - \mu}{\sigma}$は標準正規分布に従う
$$Z_i \sim N(0,1)$$

- このときZ_i^2は自由度1のχ^2分布$\chi^2(1)$に従うことが知られている
$$Z_i^2 = \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(1)$$

- $W = \sum_{i=1}^{n} Z_i^2$は自由度nのχ^2分布$\chi^2(n)$に従う
$$W = \sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$$

χ^2分布



χ^2分布

- 自由度kのχ^2分布$\chi^2(k)$の確率密度分布$f(x;k)$は$0 \leq x$の範囲で次式で表すことができる($x < 0$のときは0)

$$f(x;k) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} e^{-\frac{x}{2}} x^{\frac{k}{2}-1}$$

- ここで$\Gamma(\alpha)$はガンマ関数
- 自由度kのχ^2分布の確率密度関数について、期待値と分散はそれぞれkおよび$2k$となる

ガンマ分布とχ^2分布

- ガンマ分布$Ga(k/2\,, 2)$ は自由度kのχ^2分布$\chi^2(k)$となる

$$Ga\left(\frac{k}{2}, 2\right) = \chi^2(k)$$

- ここで、$x = z^2$とすると、ガンマ分布$Ga(1/2\,, 1/2)$は標準正規分布に式変形できる

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