

Gauss積分, ガンマ関数, ベータ関数

- 黒木玄
- 2022-04-10~2022-04-12

このノートではJulia言語 (<https://julialang.org/>)を使用している:

- [Julia言語のインストールの仕方の一例](https://nbviewer.org/github/genkuroki/msfd28/blob/master/install.ipynb) (<https://nbviewer.org/github/genkuroki/msfd28/blob/master/install.ipynb>)

自明な誤りを見つけたら, 自分で訂正して読んで欲しい. 大文字と小文字の混同や書き直しが不完全な場合や符号のミスは非常によくある.

このノートの内容よりもさらに詳しいノートを自分で作ると勉強になるだろう. 膨大な時間を取られることになるが, このノートの内容に関係することで飯を食っていく可能性がある人にはそのためにかけた時間は無駄にならないと思われる.

目次

- ▼ [1 Gauss積分](#)
 - [1.1 Gauss積分の定義](#)
 - [1.2 問題: 正規分布での確率の総和](#)
 - [1.3 正規分布の定義](#)
 - [1.4 \$p\(x|\theta\)\$ という記号法について](#)
 - [1.5 問題: Gauss積分の拡張とガンマ関数の関係](#)
- ▼ [2 ガンマ関数](#)
 - [2.1 ガンマ関数の定義](#)
 - [2.2 問題: ガンマ関数はこの形でよく使われる](#)
 - [2.3 ガンマ分布の定義](#)
 - [2.4 指数分布の定義](#)
 - [2.5 ガンマ関数の函数等式](#)
 - [2.6 問題: 正の半整数でのガンマ関数の値](#)
 - [2.7 問題: 標準正規分布の偶数次のモーメント](#)
 - [2.8 Stirlingの公式](#)
 - [2.9 問題: 小さな \$n\$ でもStirlingの公式による \$n!\$ の近似精度がそう悪くないこと](#)
 - [2.10 Stirlingの公式とガンマ分布の正規分布近似\(中心極限定理\)の関係](#)
 - [2.11 問題: \$n-1\$ 次元単位球面の面積のガンマ関数表示](#)
 - [2.12 問題: \$n\$ 次元単位球体の面積のガンマ関数表示](#)
- ▼ [3 ベータ関数](#)
 - [3.1 ベータ関数の定義](#)
 - [3.2 ベータ分布の定義](#)
 - [3.3 ベータ関数は \$\alpha\$ と \$\beta\$ を交換しても不変](#)
 - [3.4 問題: ベータ分布の三角関数を使った表示](#)
 - [3.5 ベータ関数の別の基本的表示](#)
 - [3.6 問題: \$t\$ 分布の基礎になる公式](#)
 - [3.7 \$t\$ 分布の定義](#)
 - [3.8 問題: \$t\$ 分布の自由度無限大の極限で標準正規分布が得られること](#)
 - [3.9 ベータ関数の極限としてガンマ関数を表せること](#)
- ▼ [4 ベータ関数のガンマ関数表示](#)
 - [4.1 積分 \$J\[f\]\$ を導入とベータ関数のガンマ関数表示の証明](#)
 - [4.2 積分 \$J\[f\]\$ に関する公式の証明](#)
 - [4.3 問題: ガンマ分布とベータ分布の関係](#)
 - [4.4 問題: ガンマ関数のGaussの乗法公式](#)
 - [4.5 問題: ガンマ関数の相反公式\(reflection formula\)](#)
 - [4.6 ディガンマ関数, トリガンマ関数, ポリガンマ関数, Hurwitzのゼータ関数](#)
- ▼ [5 Dirichlet積分=多変数ベータ関数](#)
 - [5.1 Dirichlet積分=多変数ベータ関数の定義とそのガンマ関数表示](#)
 - [5.2 Dirichlet分布の定義](#)
 - [5.3 多変数ベータ関数のガンマ関数表示の証明](#)
 - [5.4 問題: Dirichlet \(1839\)の公式の証明](#)

```
In [1]: 1 ENV["LINES"], ENV["COLUMNS"] = 100, 100
2 using Distributions
3 using StatsPlots
4 default(fmt = :png, titlefontsize = 10, size = (400, 250))
5 using Random
6 Random.seed!(4649373)
7 using StatsBase
8 using QuadGK
9 using SymPy
10 using SpecialFunctions
11 using Printf
```

1 Gauss積分

1.1 Gauss積分の定義

次の積分 I を **Gauss積分** と呼ぶ:

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

この積分の値が実際に円周率の平方根になることを以下で示そう.

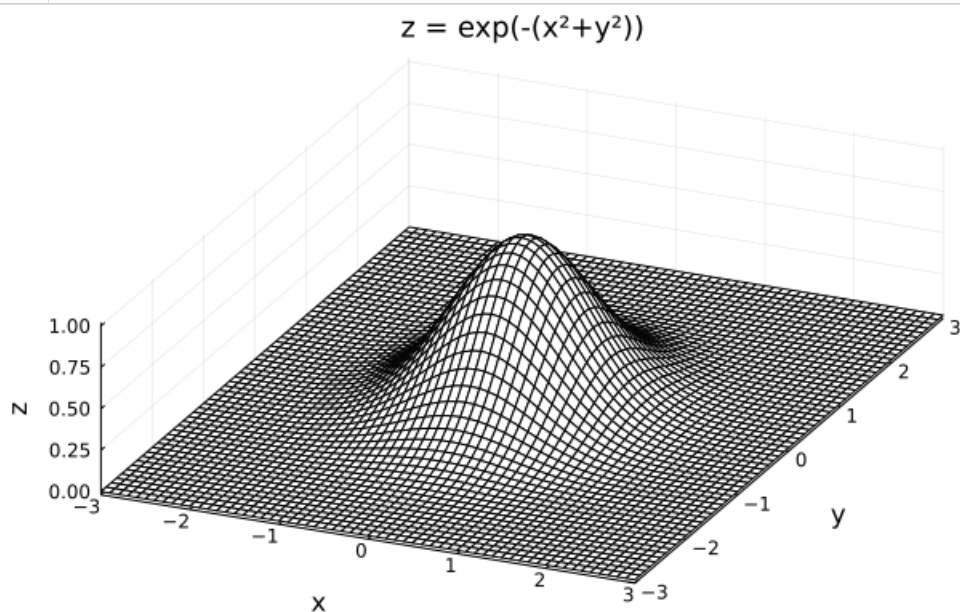
Gauss積分の2乗

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2} dx \cdot \int_{-\infty}^{\infty} e^{-y^2} dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy$$

は次のグラフで示された高さ 1 の山の体積になる.

```
In [2]: 1 x = y = range(-3, 3, 61)
2 wireframe(x, y, (x, y) -> exp(-(x^2+y^2))); camera=(30, 70), colorbar=false)
3 plot!(); zlim=(-0.02, 1.02), size=(600, 400))
4 plot!(); xlabel="x", ylabel="y", zlabel="z")
5 title!("z = exp(-(x^2+y^2))"; titlefontsize=12)
```

Out[2]:

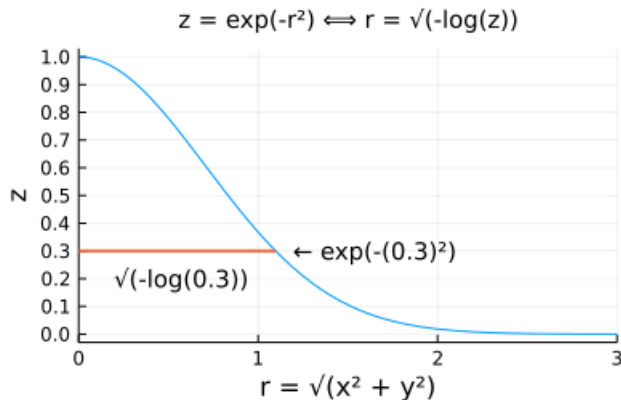


この山は z 軸 $x = y = 0$ を中心にして回転対称である. 次のグラフは横軸を半径 $r = \sqrt{x^2 + y^2}$ とした場合のこの山の高さのグラフである.

In [3]:

```
1 plot(r -> exp(-r^2), 0, 3; label="", title="z = exp(-r^2) ⇔ r = √(-log(z))")
2 plot!(); xlabel="r = √(x^2 + y^2)", ylabel="z")
3 plot!(); ytick=0:0.1:1
4 z = 0.3
5 plot!([(0, z), (√(-log(z)), z)]; label="", lw=2)
6 plot!(); xlim=(0, 3))
7 annotate!(0.2, 0.2, ("√(-log(0.3))", 10, :left))
8 annotate!(1.2, 0.3, ("← exp(-(0.3)^2)", 10, :left))
```

Out[3]:



この図より、高さ $0 < z \leq 1$ での山の切断面は、半径 $\sqrt{-\log z}$ の円盤になる。その円盤の面積 $\pi(-\log z)$ を $0 < z \leq 1$ について積分すれば上の山の体積(= Gauss積分の2乗 I^2 の値)が得られる:

$$I^2 = \int_0^1 \pi(-\log z) dz = -\pi[z \log z - z]_0^1 = -\pi(-1) = \pi.$$

ゆえに

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

1.2 問題: 正規分布での確率の総和

$\mu, \sigma \in \mathbb{R}, \sigma > 0$ であると仮定する。次を示せ:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = 1.$$

解答例: 左辺の積分部分が $\sqrt{2\pi\sigma^2}$ に等しいことを示せばよい。

$$\int_{-\infty}^{\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy = \sqrt{2\sigma^2} \int_{-\infty}^{\infty} e^{-z^2} dz = \sqrt{2\sigma^2} \sqrt{\pi} = \sqrt{2\pi\sigma^2}$$

最初の等号で $x = y + \mu$ とおき、2つめの等号で $y = \sqrt{2\sigma^2} z$ とおき、3つめの等号で Gauss積分を使った。

解答終

1.3 正規分布の定義

$\mu, \sigma \in \mathbb{R}, \sigma > 0$ であると仮定する。

次の確率密度関数によって定義される連続分布を平均 μ 、分散 σ^2 (もしくは標準偏差 σ) の **正規分布** (normal distribution) と呼び $\text{Normal}(\mu, \sigma)$ と表す:

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

特に $\mu = 0, \sigma = 1$ の正規分布を **標準正規分布** (standard normal distribution)と呼び, `Normal()` と表すことにする:

$$p(x|0, 1) = \frac{1}{\sqrt{2\pi}} e^{-x^2}.$$

1.4 $p(x|\theta)$ という記号法について

パラメータ θ を持つ x に関する確率密度関数を $p(x|\theta)$ と書く.

この記号法によって, 縦棒 $|$ の左側の変数 x について確率密度関数になっているが, 縦棒 $|$ の右側の変数 θ については単なる関数で確率密度関数ではないことを表す.

パラメータ θ を持つ x に関する確率質量関数 $P(x|\theta)$ についても同様である.

1.5 問題: Gauss積分の拡張とガンマ関数の関係

$\alpha \in \mathbb{R}, \alpha > 0$ であると仮定する. 次を示せ:

$$\int_{-\infty}^{\infty} e^{-x^2} |x|^{2\alpha-1} dx = \int_0^{\infty} e^{-y} y^{\alpha-1} dy$$

解答例: $x = \sqrt{y}$ とおくと, $dx/x = d \log x = (1/2)d \log y = dy/(2y)$ なので,

$$\int_{-\infty}^{\infty} e^{-x^2} |x|^{2\alpha-1} dx = 2 \int_0^{\infty} e^{-x^2} x^{2\alpha} \frac{dx}{x} = 2 \int_0^{\infty} e^{-y} y^{\alpha} \frac{dy}{2y} = \int_0^{\infty} e^{-y} y^{\alpha-1} dy.$$

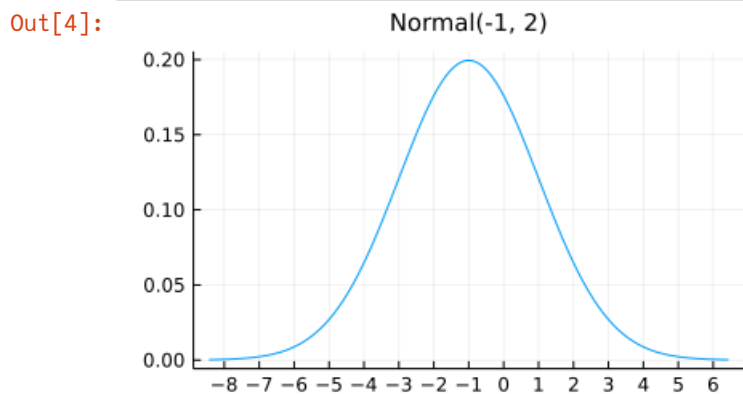
最初の等号で被積分関数が偶関数であることを用い, 第2の等号で $x = \sqrt{y}$ とおいた.

解答終

この問題では単なる $\exp(-x^2)$ の積分であるGauss積分をそれに $|x|$ のべきをかけたものの積分に拡張した. 実はこの拡張は次の節で扱うガンマ関数の定義そのものになっている. すなわち, ガンマ関数はGauss積分の拡張の一種になっている.

正規分布の確率密度関数は釣鐘型(bell-shaped)になる.

```
In [4]: 1 plot(Normal(-1, 2); label="", title="Normal(-1, 2)", xtick = -100:100)
```



2 ガンマ関数

2.1 ガンマ関数の定義

$\alpha > 0$ について, ガンマ関数 $\Gamma(\alpha)$ を次のように定義する:

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

前節の問題の結果より, これは

$$\Gamma(\alpha) = \int_{-\infty}^{\infty} e^{-x^2} |x|^{2\alpha-1} dx$$

とGauss積分の拡張形でも表される. 特に

$$\Gamma(1/2) = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

このようにガンマ関数 $\Gamma(\alpha)$ は Gauss 積分 $\Gamma(1/2)$ の拡張になっている.

$\Gamma(1) = 1$ となることも容易に示せる:

$$\Gamma(1) = \int_0^{\infty} e^{-x} dx = [-e^{-x}]_0^{\infty} = -(-1) = 1.$$

注意: Gauss 積分は正規分布の定義の基礎になっている. ここで扱っているような Gauss 積分の拡張形は平均が 0 の正規分布の絶対モーメントになっており, 正規分布に関係した計算には自然にガンマ関数が出て来る.

2.2 問題: ガンマ関数はこの形でよく使われる

$\alpha, \theta > 0$ について次の公式が成立することを示せ:

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

解答例: $x = \theta y$ とおくと, $dx/x = \theta dy/(\theta y) = dy/y$ なので,

$$\int_0^{\infty} e^{-x/\theta} x^{\alpha-1} dx = \int_0^{\infty} e^{-x/\theta} x^{\alpha} \frac{dx}{x} = \int_0^{\infty} e^{-y} (\theta y)^{\alpha} \frac{dy}{y} = \theta^{\alpha} \int_0^{\infty} e^{-y} y^{\alpha-1} dy = \theta^{\alpha} \Gamma(\alpha).$$

2つめの等号で $x = \theta y$ を代入して, 4つめの等号でガンマ関数の定義を使った.

解答終

ガンマ関数は主にこの形で使う と覚えておくと便利だと思う.

例えば, この問題の結果から,

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

という公式が得られる. これは「ガンマ分布において確率の総和が1になる」という意味を持つ.

2.3 ガンマ分布の定義

$\alpha, \theta > 0$ であると仮定する.

次の確率密度関数によって定義される連続分布を形状パラメータ α , スケールパラメータ θ の **ガンマ分布** (Gamma distribution) と呼び $\text{Gamma}(\alpha, \theta)$ と表す:

$$p(x|\alpha, \theta) = \frac{1}{\theta^{\alpha} \Gamma(\alpha)} e^{-x/\theta} x^{\alpha-1} \quad (x > 0).$$

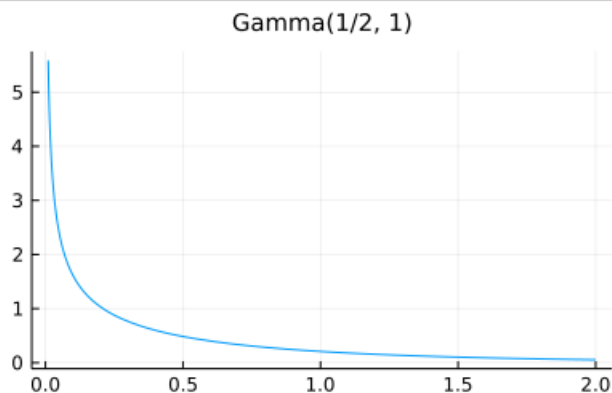
ガンマ分布は正規分布から派生して出て来る分布の中で非常に基本的なものになっている.

負の二項分布の連続極限としてガンマ分布を出すことによって, ガンマ分布には「1回起こるまでにかかる時間の期待値が θ であるようなイベントが α 回起こるまでにかかる時間の分布」だと解釈できる. この点については後で別のノートで説明する.

$\alpha = 1/2$ のガンマ分布の確率密度関数は次の形になる.

```
In [5]: 1 plot(Gamma(1/2, 1), 0.01, 2; label="", title="Gamma(1/2, 1)")
```

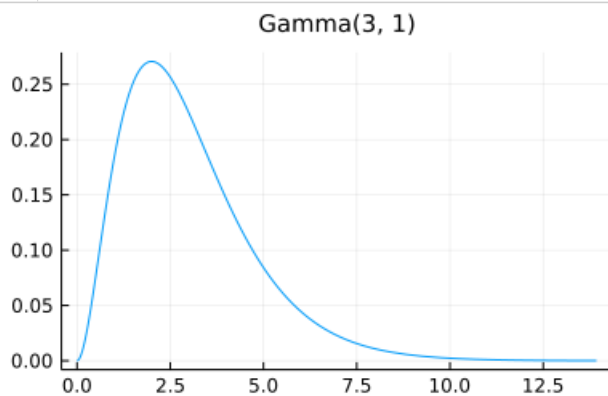
Out[5]:



$\alpha = 3$ のガンマ分布の確率密度関数は次の形になる.

```
In [6]: 1 plot(Gamma(3, 1); label="", title="Gamma(3, 1)")
```

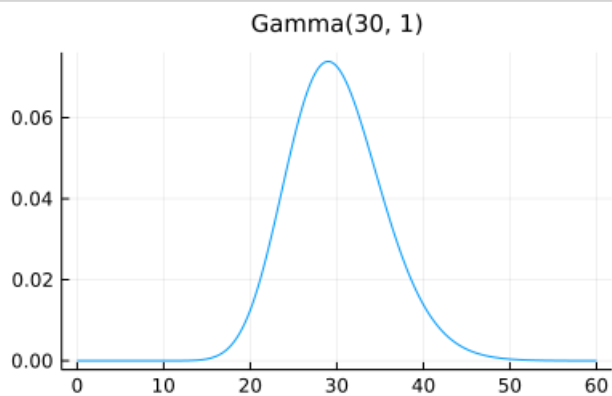
Out[6]:



$\alpha = 30$ のガンマ分布の確率密度関数は次の形になる.

```
In [7]: 1 plot(Gamma(30, 1), 0, 60; label="", title="Gamma(30, 1)")
```

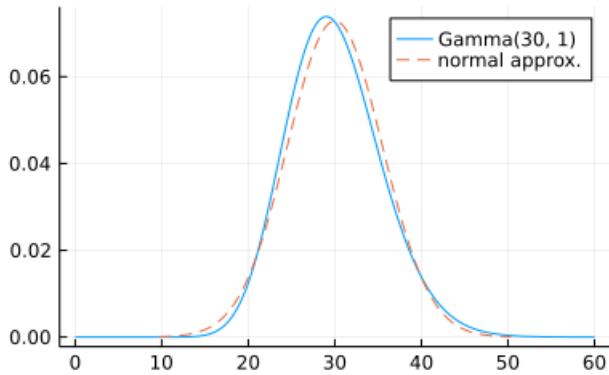
Out[7]:



形状パラメータ α が大きなガンマ分布は正規分布でよく近似される.

```
In [8]: 1 gam = Gamma(30, 1)
2 plot(gam, 0, 60; label="Gamma(30, 1)")
3 plot!(Normal(mean(gam), std(gam)); label="normal approx.", ls=:dash)
```

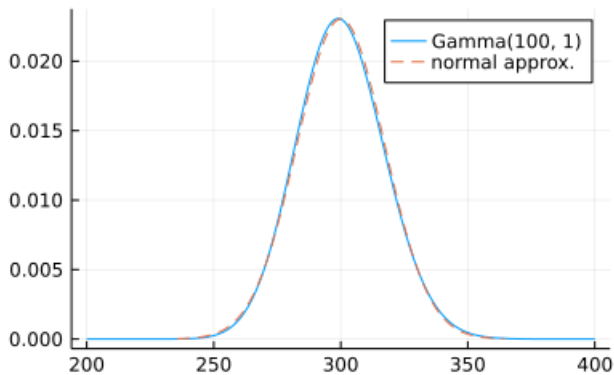
Out[8]:



ガンマ分布の確率密度関数は $(\alpha - 1)\theta$ で最大になることを示せる。それに対して、それを近似する正規分布の期待値はもとのガンマ分布の期待値 $\alpha\theta$ に等しくなるので、 $\alpha = 30, \theta = 1$ の場合にはその分だけずれが生じている。

```
In [9]: 1 gam = Gamma(300, 1)
2 plot(gam, 200, 400; label="Gamma(100, 1)")
3 plot!(Normal(mean(gam), std(gam)); label="normal approx.", ls=:dash)
```

Out[9]:



注意: 形状パラメータ α が大きなガンマ分布は正規分布でよく近似されることは本質的に階乗に関するStirlingの(近似)公式に同等である。

2.4 指数分布の定義

$\alpha = 1$ のガンマ分布の密度関数は

$$p(x|1, \theta) = \frac{1}{\theta} e^{-x/\theta} \quad (x > 0)$$

の形になる。この形の確率密度関数で定義される連続確率分布を **指数分布** (exponential distribution)と呼び、 $\text{Exponential}(\theta)$ と表す。

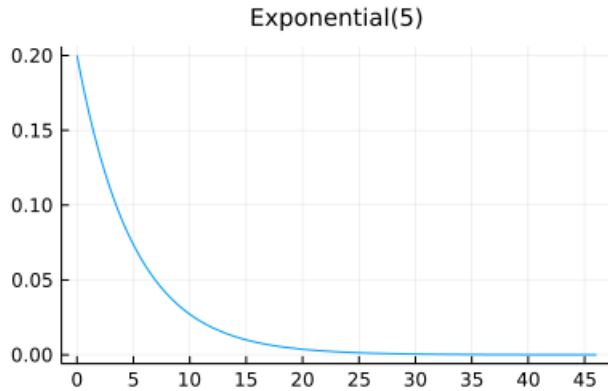
指数分布は「1回起こるまでにかかる時間の期待値が θ であるようなイベントが起こるまでにかかる時間の分布」だと解釈可能である。この点については後で別のノートで説明する。

家計が保有する金融資産の分布は指数分布に近い形の分布になる。

指数分布の確率密度関数は次の形になる。

```
In [10]: 1 plot(Exponential(5); label="", title="Exponential(5)", xtick=0:5:100)
```

Out[10]:



2.5 ガンマ関数の函数等式

ガンマ関数の函数等式

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha) \quad (\alpha > 0)$$

を示そう. $(-e^{-x})' = e^{-x}$ に関する部分積分を行うと,

$$\begin{aligned} \Gamma(\alpha + 1) &= \int_0^\infty e^{-x} x^\alpha dx = \int_0^\infty (-e^{-x})' x^\alpha dx \\ &= [-e^{-x} x^\alpha]_0^\infty + \int_0^\infty e^{-x} \alpha x^{\alpha-1} dx = \alpha \int_0^\infty e^{-x} x^{\alpha-1} dx = \alpha \Gamma(\alpha). \end{aligned}$$

3つめの等号で部分積分を行った. 4つめの等号では $\alpha > 0$ より $\lim_{x \rightarrow \infty} e^{-x} x^\alpha = 0$ となることと, $\lim_{x \rightarrow 0} e^{-x} x^\alpha = 0$ となることを使った. 最初と最後の等号でガンマ関数の定義を使った.

ガンマ関数の函数等式と易しく証明できる $\Gamma(1) = 1$ から,

$$\Gamma(n + 1) = n! \quad (n = 0, 1, 2, \dots)$$

となることを示せる. 実際, n が 0 以上の整数のとき,

$$\Gamma(n + 1) = n \Gamma(n) = n(n - 1) \Gamma(n - 1) = \dots = n(n - 1) \dots 2 \cdot 1 \Gamma(1) = n!.$$

ガンマ関数の函数等式を n 回使い, 最後に $\Gamma(1) = 1$ を使った.

2.6 問題: 正の半整数でのガンマ関数の値

次を示せ:

$$\Gamma\left(\frac{2n+1}{2}\right) = \frac{1 \cdot 3 \cdots (2n-1)}{2^n} \sqrt{\pi} = \frac{(2n)!}{2^{2n} n!} \sqrt{\pi} \quad (n = 0, 1, 2, \dots)$$

解答例: n が 0 以上の整数であるとき,

$$\begin{aligned} \Gamma\left(\frac{2n+1}{2}\right) &= \frac{2n-1}{2} \Gamma\left(\frac{2n-1}{2}\right) = \frac{2n-1}{2} \frac{2n-3}{2} \Gamma\left(\frac{2n-3}{2}\right) = \dots \\ &= \frac{2n-1}{2} \frac{2n-3}{2} \cdots \frac{3}{2} \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{1 \cdot 3 \cdots (2n-1)}{2^n} \sqrt{\pi}. \end{aligned}$$

ガンマ関数の函数等式を n 回使い, 最後に Gauss 積分 $\Gamma(1/2) = \sqrt{\pi}$ を使った.

さらに, 分子分母に $2 \cdot 4 \cdots (2n) = 2^n n!$ をかけることによって,

$$\Gamma\left(\frac{2n+1}{2}\right) = \frac{1 \cdot 3 \cdots (2n-1)}{2^n} \sqrt{\pi} = \frac{1 \cdot 3 \cdots (2n-1)}{2^n} \frac{2 \cdot 4 \cdots (2n)}{2^n n!} \sqrt{\pi} = \frac{(2n)!}{2^{2n} n!} \sqrt{\pi}.$$

解答終

2.7 問題: 標準正規分布の偶数次のモーメント

次を示せ:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} x^{2n} dx = \frac{2^n}{\sqrt{\pi}} \Gamma\left(\frac{2n+1}{2}\right) = \frac{(2n)!}{2^n n!} = 1 \cdot 3 \cdots (2n-1) \quad (n = 0, 1, 2, \dots)$$

解答例: 被積分関数が偶関数であることを使ってから, $x = \sqrt{y}$ とおくと $dx = y^{-1/2} dy/2$ となることより,

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} x^{2n} dx &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-x^2/2} x^{2n} dx = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-y/2} y^{n-1/2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-y/2} y^{(2n+1)/2-1} dy = \frac{1}{\sqrt{2\pi}} 2^{(2n+1)/2} \Gamma\left(\frac{2n+1}{2}\right) \\ &= \frac{2^n}{\sqrt{\pi}} \Gamma\left(\frac{2n+1}{2}\right) = \frac{(2n)!}{2^n n!} = 1 \cdot 3 \cdots (2n-1). \end{aligned}$$

最後から4つめの等号でガンマ函数のよく使われる形 $\int_0^{\infty} e^{-x/\theta} x^{\alpha-1} dx = \theta^{\alpha} \Gamma(\alpha)$ を使い, 終わりの2つの等号で正の半整数でのガンマ函数の値を使った.

解答終

注意: 一般に確率変数 X について $E[X^k]$ ($k = 0, 1, 2, \dots$) を X もしくは X が従う分布の **モーメント** (積率, moment) と呼ぶ. 上の結果は標準正規分布に従う確率変数 X のモーメントについて,

$$E[X^{2n}] = \frac{(2n)!}{2^n n!} = 1 \cdot 3 \cdots (2n-1) \quad (n = 0, 1, 2, \dots)$$

が成立することを意味している. x^{2n+1} は奇関数になるので $E[X^{2n+1}] = 0$ となる. これらの結果を使うと,

$$E[e^{tX}] = \sum_{n=0}^{\infty} \frac{t^{2n} E[X^{2n}]}{(2n)!} = \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!} \frac{(2n)!}{2^n n!} = \sum_{n=0}^{\infty} \frac{(t^2/2)^n}{n!} = e^{t^2/2}$$

となることもわかる(この公式を上の結果を使わずに直接証明した方が実は易しい). これを標準正規分布の **モーメント母函数** (moment generating function, mgf) と呼ぶ. さらに

$$\log E[e^{tX}] = \frac{t^2}{2}$$

を標準正規分布の **キュムラント母函数** (cumulant generating function, cgf) と呼ぶ. 標準正規分布のキュムラント母函数がこのようにシンプルな形になることは, 中心極限定理(正規分布による普遍的な近似定理)や正規分布で近似したときの誤差の見積もりなどで使われる.

2.8 Stirlingの公式

次の階乗の近似公式を **Stirlingの公式** と呼ぶ:

$$n! = n^n e^{-n} \sqrt{2\pi n} (1 + \varepsilon_n) \quad (\varepsilon_n \rightarrow 0 \text{ as } n \rightarrow \infty)$$

Stirlingの公式は後で空気のごとく使われることになる.

Stirlingの公式はガンマ函数の応用として以下のように示される.

$n!/(n^n e^{-n} \sqrt{n})$ が $n \rightarrow \infty$ で $\int_{-\infty}^{\infty} e^{-y^2/2} dy = \sqrt{2\pi}$ に収束することを示せばよい.

$$n! = \Gamma(n+1) = \int_0^{\infty} e^{-x} x^n dx \text{ で}$$

$$x = n + \sqrt{n} y = n \left(1 + \frac{y}{\sqrt{n}} \right)$$

とおくと(これがポイント!),

$$n! = \int_{-\sqrt{n}}^{\infty} e^{-n} e^{-\sqrt{n}y} n^n \left(1 + \frac{y}{\sqrt{n}}\right)^n \sqrt{n} dy = n^n e^{-n} \sqrt{n} \int_{-\sqrt{n}}^{\infty} e^{-\sqrt{n}y} \left(1 + \frac{y}{\sqrt{n}}\right)^n dy.$$

そして,

$$\begin{aligned} \log\left(e^{-\sqrt{n}y} \left(1 + \frac{y}{\sqrt{n}}\right)^n\right) &= n \log\left(1 + \frac{y}{\sqrt{n}}\right) - \sqrt{n}y \\ &= n \left(\frac{y}{\sqrt{n}} - \frac{y^2}{2n} + \frac{y^3}{3n\sqrt{n}} - \dots\right) - \sqrt{n}y = -\frac{y^2}{2} + \frac{y^3}{3\sqrt{n}} - \dots \\ &\rightarrow -\frac{y^2}{2} \quad (\text{as } n \rightarrow \infty) \end{aligned}$$

ゆえに, $e^{-\sqrt{n}y} \left(1 + y/\sqrt{n}\right)^n \rightarrow e^{-y^2/2} (n \rightarrow \infty)$.

したがって, $n \rightarrow \infty$ のとき,

$$\frac{n!}{n^n e^{-n} \sqrt{n}} = \int_{-\sqrt{n}}^{\infty} e^{-\sqrt{n}y} \left(1 + \frac{y}{\sqrt{n}}\right)^n dy \rightarrow \int_{-\infty}^{\infty} e^{-y^2/2} dy = \sqrt{2\pi}.$$

これで示すべきことが示された.

注意: 極限と積分の交換可能性が気になる読者はLebesgueの収束定理の適用を考えてみよ. もっと初等的な方法でも極限と積分の交換可能性を示せる. この手の細かいことを気にならない人はそのまま気にする必要はないが, 関連のグラフを大量に描いてみたり, 数値積分を試してみたり, 様々な試行錯誤によって理解を深める努力をした方がよい.

2.9 問題: 小さな n でもStirlingの公式による $n!$ の近似精度がそう悪くないこと

$n = 1, 2, 3, \dots, 10$ について $n!$ と

$$\text{stirling}(n) = n^n e^{-n} \sqrt{2\pi n}$$

の両方を計算し, 前者の後者による近似の相対誤差を求めよ. ここで $a > 0$ の b による近似の相対誤差の定義は $b/a - 1$ であるとする. $\text{stirling}(n)$ については小数点以下1桁以上かつ有効桁4桁以上を, 相対誤差について有効桁3桁以上を計算せよ.

拡張版解答例+解説: $1/(12n)$ によって補正した近似公式

$$\text{stirling}_1(n) = n^n e^{-n} \sqrt{2\pi n} \left(1 + \frac{1}{12n}\right)$$

についてもあわせて計算してみよう.

| n | $n!$ | $\text{stirling}(n)$ | relative error | $n!$ | $\text{stirling}_1(n)$ | relative error |
|-----|---------|----------------------|----------------|---------|------------------------|----------------|
| 1 | 1 | 0.9221 | -0.0779 | 1 | 0.99898 | -0.00102 |
| 2 | 2 | 1.919 | -0.0405 | 2 | 1.9990 | -0.000519 |
| 3 | 6 | 5.836 | -0.0273 | 6 | 5.9983 | -0.000279 |
| 4 | 24 | 23.51 | -0.0206 | 24 | 23.996 | -0.000171 |
| 5 | 120 | 118.0 | -0.0165 | 120 | 119.99 | -0.000115 |
| 6 | 720 | 710.1 | -0.0138 | 720 | 719.94 | -0.0000828 |
| 7 | 5040 | 4980.4 | -0.0118 | 5040 | 5039.7 | -0.0000623 |
| 8 | 40320 | 39902.4 | -0.0104 | 40320 | 40318.0 | -0.0000485 |
| 9 | 362880 | 359536.9 | -0.00921 | 362880 | 362865.9 | -0.0000388 |
| 10 | 3628800 | 3598695.6 | -0.00830 | 3628800 | 3628684.7 | -0.0000318 |

この結果を見れば, $n!$ のStirling公式による近似の相対誤差の絶対値は $n = 2$ ですでに 5% を切っており, $n = 9$ で 1% を切っていることがわかる. $1/(12n)$ で補正した場合には $n = 1$ の段階ですでに相対誤差の絶対値はほぼ 0.1% になっていて非常に小さい. この事実は

$$\frac{13\sqrt{2\pi}}{12e} \approx 0.9989817596371048$$

が非常に 1 に近いことが偶然ではないことも意味している. このように, $1/(12n)$ による補正で誤差が非常に小さくなるという事実は, Stirlingの公式による階乗の近似の相対誤差の大きさがほぼ $1/(12n)$ になることを意味している.

階乗は確率の記述でよく使われるので, 階乗のシンプルな近似式であるStirlingの公式は確率が絡む議論において非常に有用である.

拡張版解答終

```
In [11]: 1 stirling(n) = n^n * exp(-n) * √(2π*n)
2 stirling1(n) = n^n * exp(-n) * √(2π*n) * (1 + 1/(12n))
3 for n in 1:10
4     a = factorial(n)
5     b = stirling(n)
6     c = stirling1(n)
7     println("$n, $a, $b, $(b/a - 1), $c, $(c/a - 1)")
8 end
```

```
1, 1, 0.9221370088957891, -0.07786299110421091, 0.9989817596371048, -0.0010182403628952175
2, 2, 1.9190043514889832, -0.04049782425550841, 1.9989628661343577, -0.0005185669328211517
3, 6, 5.836209591345864, -0.027298401442355957, 5.998326524438804, -0.000278912593532632
4, 24, 23.50617513289329, -0.02057603612944625, 23.995887114828566, -0.0001713702154764185
5, 120, 118.01916795759008, -0.01650693368674927, 119.98615409021657, -0.00011538258152854475
6, 720, 710.0781846421849, -0.013780299108076544, 719.9403816511041, -8.280326235543534e-5
7, 5040, 4980.395831612461, -0.01182622388641652, 5039.686258179276, -6.225036125484529e-5
8, 40320, 39902.3954526567, -0.010357255638474672, 40318.04540528855, -4.847705137533964e-5
9, 362880, 359536.87284194824, -0.00921276223008094, 362865.9179608552, -3.880632480379731e-5
10, 3628800, 3.5986956187410364e6, -0.008295960443938433, 3.6286847488972116e6, -3.1760114304502096e-5
```

```
In [12]: 1 3.5986956187410364e6 ▷ x → round(x; digits=1)
```

Out[12]: 3.5986956e6

```
In [13]: 1 3.6286847488972116e6 ▷ x → round(x; digits=1)
```

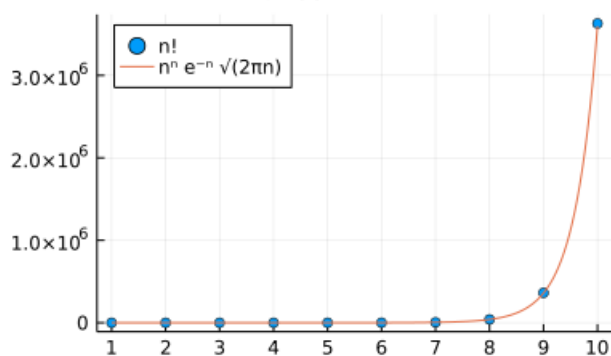
Out[13]: 3.6286847e6

```
In [14]: 1 13√(2π)/(12e)
```

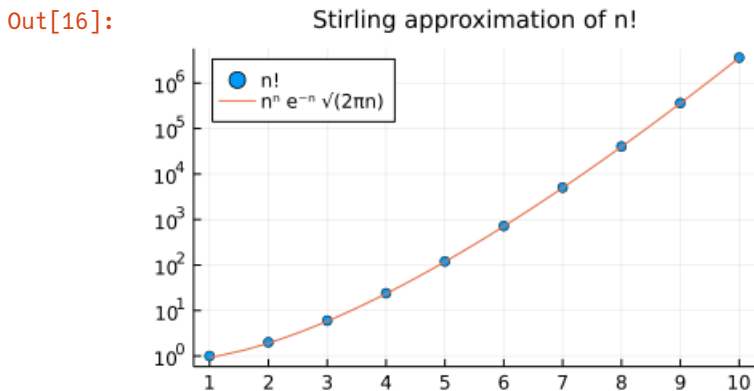
Out[14]: 0.9989817596371048

```
In [15]: 1 N = 10
2 plot(; legend=:topleft, xtick=1:N)
3 scatter!(1:N, factorial; label="n!")
4 plot!(stirling, 1, N; label="n^n e^-n √(2πn)")
5 title!("Stirling approximation of n!")
```

Out[15]: Stirling approximation of n!



```
In [16]: 1 N = 10
2 plot(; legend=:topleft, xtick=1:N)
3 scatter!(1:N, factorial; label="n!")
4 plot!(stirling, 1, N; label="n^n e^{-n} \sqrt{(2\pi n)}")
5 title!("Stirling approximation of n!")
6 plot!(;yscale=:log10)
```



2.10 Stirlingの公式とガンマ分布の正規分布近似(中心極限定理)の関係

この節では, Stirlingの公式は本質的にガンマ分布の正規分布近似(中心極限定理ともいう)から出て来ることを説明する.

スケールパラメータが $\theta = 1$ で形状パラメータが $\alpha = n + 1$ のガンマ分布 $\text{Gamma}(n + 1, 1)$ の確率密度関数は

$$p_n(x) = \frac{e^{-x} x^n}{\Gamma(n + 1)} = \frac{e^{-x} x^n}{n!} \quad (x > 0)$$

という形になる. この $p_n(x)$ の分子の対数を $f_n(x)$ と書く:

$$f_n(x) = \log(e^{-x} x^n) = n \log x - x.$$

$f_n(x)$ を $x = n$ でTaylor展開してみよう.

$$f_n(n) = n \log n - n, \quad f'_n(x) = \frac{n}{x} - 1, \quad f'_n(n) = 0, \quad f''_n(x) = -\frac{n}{x^2}, \quad f''_n(n) = -\frac{1}{n}.$$

これより, x が n に近いとき,

$$f_n(x) \approx n \log n - n - \frac{(x - n)^2}{2n}.$$

ゆえに, x が n に近いとき, $f_n(x)$ の $x = n$ でのTaylor展開より,

$$e^{-x} x^n = \exp(f_n(x)) \approx \exp\left(n \log n - n - \frac{(x - n)^2}{2n}\right) = n^n e^{-n} \exp\left(-\frac{(x - n)^2}{2n}\right).$$

n が大きくなると, $e^{-x} x^n$ も $\exp(-(x - n)^2/(2n))$ も x が n から離れると急激に 0 に近づくので, それぞれの $x > 0$, $-\infty < x < \infty$ における積分は n に近い x における積分で近似される. したがって,

$$\begin{aligned} 1 &= \int_0^\infty p_n(x) dx = \frac{1}{n!} \int_0^\infty e^{-x} x^n dx \\ &\approx \frac{1}{n!} \int_{\text{around } n} e^{-x} x^n dx \approx \frac{n^n e^{-n}}{n!} \int_{\text{around } n} \exp\left(-\frac{(x - n)^2}{2n}\right) dx \\ &\approx \frac{n^n e^{-n}}{n!} \int_{-\infty}^\infty \exp\left(-\frac{(x - n)^2}{2n}\right) dx = \frac{n^n e^{-n} \sqrt{2\pi n}}{n!}. \end{aligned}$$

これで $n! \approx n^n e^{-n} \sqrt{2\pi n}$ が再び示された.

以上の計算によってガンマ分布 $\text{Gamma}(n + 1, 1)$ の密度関数 $p_n(x)$ が平均 n , 分散 n の正規分布の密度関数で近似されることも示している:

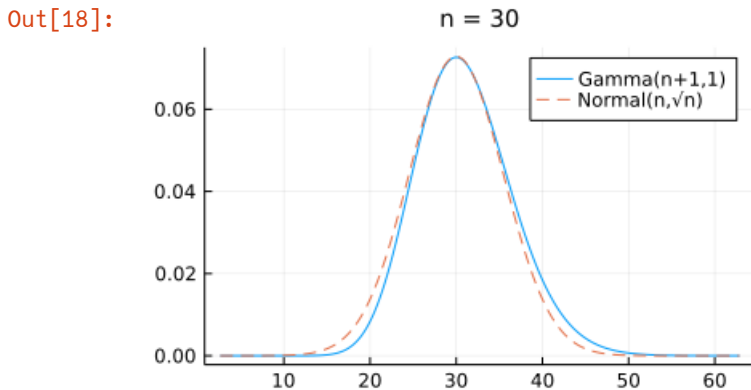
$$p_n(x) = \frac{e^{-x} x^n}{n!} \approx \frac{1}{\sqrt{2\pi n}} \exp\left(-\frac{(x - n)^2}{2n}\right).$$

以上はかなり雑な議論だが, 議論の全体を正確に作り直すこともできるし, 以下で示すようにグラフを描くことによっても結果が正しいことを確認できる.

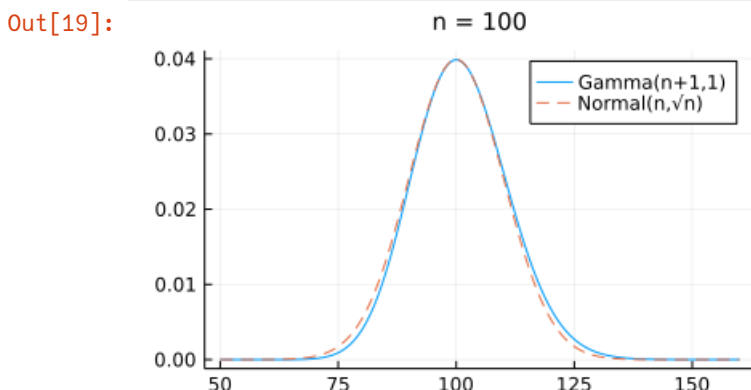
```
In [17]: 1 p(n, x) = x > 0 ? exp(n*log(x) - x - loggamma(n+1)) : zero(x)
2 f(n, x) = 1/sqrt(2*pi*n) * exp(-(x - n)^2/(2n))
```

Out[17]: f (generic function with 1 method)

```
In [18]: 1 n = 30
2 x = range(n - 5*sqrt(n), n + 6*sqrt(n), 1000)
3 plot(x, x -> p(n, x); label="Gamma(n+1,1)")
4 plot!(x, x -> f(n, x); label="Normal(n,sqrt(n))", ls=:dash)
5 title!("n = $n")
```



```
In [19]: 1 n = 100
2 x = range(n - 5*sqrt(n), n + 6*sqrt(n), 1000)
3 plot(x, x -> p(n, x); label="Gamma(n+1,1)")
4 plot!(x, x -> f(n, x); label="Normal(n,sqrt(n))", ls=:dash)
5 title!("n = $n")
```



注意: 以上のグラフにおけるガンマ分布と正規分布の密度関数の差の大部分は $x - n$ に関する奇関数になっている. 積分するとその奇関数成分は消える.

2.11 問題: $n-1$ 次元単位球面の面積のガンマ関数表示

$n-1$ 次元単位球面 S^{n-1} を

$$S^{n-1} = \{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1^2 + \dots + x_n^2 = 1 \}$$

と定める(S はsphereの頭文字). 例えば, S^1 は平面上の半径が1の円周(曲線なので1次元の図形)になり, S^2 は3次元空間上の半径が1の球面(曲面なので2次元の図形)になる. S^{n-1} の面積を A_{n-1} がガンマ関数によって次のように表されることを示せ:

$$A_{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

解答例: Gauss積分の n 乗

$$I^n = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-(x_1^2 + \dots + x_n^2)} dx_1 \dots dx_n$$

を二通りに計算することによって, $n-1$ 次元単位球面 S^{n-1} の面積 A_{n-1} を求めよう. まず, これがGauss積分の n 乗であることから,

$$I^n = \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 = (\sqrt{\pi})^2 = \pi^{n/2}.$$

$n-1$ 次元球面上の微小領域の面積を $d\omega$ と書くと、半径 r の $n-1$ 次元球面上の対応する微小領域の面積はその r^{n-1} 倍の $r^{n-1} d\omega$ になり、さらにそれに dr の厚みを持たせて作られる n 次元の微小領域の体積はほぼそれらの積 $r^{n-1} dr d\omega$ に等しくなる。 $y = r^2 = x_1^2 + \cdots + x_n^2$ とおくと、 $dr = d\sqrt{y} = y^{-1/2} dy/2$ なので、

$$r^{n-1} dr d\omega = \frac{1}{2} y^{(n-1)/2-1/2} dy d\omega = \frac{1}{2} y^{n/2-1} dy d\omega.$$

$dx_1 \cdots dx_n$ に関する積分は $\frac{1}{2} y^{n/2-1} dy d\omega$ に関する積分に書き換えられるので、

$$\pi^{n/2} = I^n = \int_{S^{n-1}} \left(\int_0^\infty e^{-y} \frac{1}{2} y^{n/2-1} dy \right) d\omega = \frac{A_{n-1}}{2} \int_0^\infty e^{-y} y^{n/2-1} dy = \frac{A_{n-1}}{2} \Gamma(n/2).$$

すなわち、 $A_{n-1} = 2\pi^{n/2}/\Gamma(n/2)$.

解答終

注意: 半径 r の $n-1$ 次元球面の面積は単位球面の面積の r^{n-1} 倍になるので

$$A_{n-1} r^{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)} r^{n-1}$$

になる。例えば、

$$\Gamma(1/2) = \sqrt{\pi}, \quad \Gamma(1) = 1, \quad \Gamma(3/2) = \frac{1}{2}\sqrt{\pi}, \quad \Gamma(2) = 1, \quad \Gamma(5/2) = \frac{3}{4}\sqrt{\pi}$$

なので、

$$\begin{aligned} A_0 &= \frac{2\pi^{1/2}}{\Gamma(1/2)} = 2, & A_1 r &= \frac{2\pi^{2/2}}{\Gamma(2/2)} r = 2\pi r, & A_2 r^2 &= \frac{2\pi^{3/2}}{\Gamma(3/2)} r^2 = 4\pi r^2, \\ A_3 r^3 &= \frac{2\pi^{4/2}}{\Gamma(4/2)} r^3 = 2\pi^2 r^3, & A_4 r^4 &= \frac{2\pi^{5/2}}{\Gamma(5/2)} r^4 = \frac{8\pi^2}{3} r^4. \end{aligned}$$

2 は集合 $S^0 = \{-1, 1\}$ の点の個数である。 $2\pi r$ や $4\pi r^2$ は高校生にとってもおなじみの公式である。

2.12 問題: n 次元単位球体の面積のガンマ函数表示

n 次元単位球体 B^n を

$$B^n = \{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 \leq 1 \}$$

と定める(B はballの頭文字)。例えば、 B^1 は直線上の線分 $[-1, 1]$ になり、 B^2 は2次元平面上の半径が1の円盤になり、 B^3 は3次元空間内の半径1の単位球体になる。 B^n の体積を V_n がガンマ函数によって次のように表されることを示せ:

$$V_n = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)}.$$

解答例: 半径 r の $n-1$ 次元球面の面積

$$A_{n-1} r^{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)} r^{n-1}$$

を r について0から1まで積分すれば、 n 次元単位球体の体積が得られる:

$$V_n = \int_0^1 A_{n-1} r^{n-1} dr = \frac{A_{n-1}}{n} = \frac{2\pi^{n/2}}{n\Gamma(n/2)} = \frac{\pi^{n/2}}{(n/2)\Gamma(n/2)} = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)}.$$

最後の等号でガンマ函数の函数等式を使った。

解答終

注意: 半径 r の n 次元球体の体積は単位球体の体積の r^n 倍になるので

$$V_n r^n = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)} r^n$$

になる。例えば、

$$\Gamma(3/2) = \frac{1}{2}\sqrt{\pi}, \quad \Gamma(2) = 1, \quad \Gamma(5/2) = \frac{3}{4}\sqrt{\pi}\Gamma(3) = 2\sqrt{\pi}\Gamma(7/2) = \frac{15}{8}\sqrt{\pi}$$

なので,

$$V_1 r = \frac{\pi^{1/2}}{\Gamma(1/2 + 1)} r = 2r, \quad V_2 r^2 = \frac{\pi^{2/2}}{\Gamma(2/2 + 1)} r^2 = \pi r^2, \quad V_3 r^3 = \frac{\pi^{3/2}}{\Gamma(3/2 + 1)} r^3 = \frac{4\pi^2 r^3}{3},$$

$$V_4 r^4 = \frac{\pi^{4/2}}{\Gamma(4/2 + 1)} r^4 = \frac{\pi^2 r^4}{2}, \quad V_5 r^5 = \frac{\pi^{5/2}}{\Gamma(5/2 + 1)} r^5 = \frac{8\pi^2 r^5}{15}.$$

$2r$ は線分 $[-r, r]$ の長さである. πr^2 や $4\pi r^3/3$ は高校生にとってもおなじみの公式だろう.

注意: m は0以上の整数であるとし, $n = 2m$ または $n = 2m + 1$ と仮定する. このとき, n 次元球体の球体の体積の公式に π^m が含まれる理由は, $n = 2m, 2m + 1$ 次元の空間では互いに可換な回転軸を m 本取れるからである. 1本の回転軸ごとに1つの π が出て来る. 例えば, $m = 2, n = 5$ のとき, 5次元空間 $\mathbb{R}^5 = \{(x_1, x_2, x_3, x_4, x_5) \mid x_1, x_2, x_3, x_4, x_5 \in \mathbb{R}\}$ における互いに可換な回転軸として, (x_1, x_2) 平面の回転軸と (x_3, x_4) 平面の回転軸の2本が取れて, 5次元球体の体積の公式には π^2 が含まれている.

3 ベータ函数

3.1 ベータ函数の定義

$\alpha, \beta > 0$ について, ベータ函数 $B(\alpha, \beta)$ を次のように定める:

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt.$$

ベータ函数のこの表示は通常の(第1種)ベータ分布だけではなく, F 分布(Fisherの F 分布)とも関係している. ベータ函数の様々な表示が正規分布に関連した重要な分布達の基礎を与える.

3.2 ベータ分布の定義

$\alpha, \beta > 0$ と仮定する.

次の確率密度函数によって定義される連続分布をパラメータ α, β の **ベータ分布** (Beta distribution)と呼び, $Beta(\alpha, \beta)$ と表す:

$$p(t|\alpha, \beta) = \frac{1}{B(\alpha, \beta)} t^{\alpha-1} (1-t)^{\beta-1} \quad (0 < t < 1)$$

二項分布がBernoulli試行(値が1と0のみの乱数生成の繰り返し)から得られたのと同じように, パラメータが正の整数の場合のベータ分布は一様乱数の生成の繰り返しから得られる. この事実については後で別のノートで解説する予定である. この段階ではベータ分布にも後で説明されるように直観的な意味付けがあることを覚えておけばよい. 実用的に使用される確率分布の世界は豊かかつ複雑であり, それらを単に暗記しようとするには無理がある. 直観的な意味付けを十分に理解して先に進む方が効率的な場合が多い.

3.3 ベータ函数は α と β を交換しても不変

$t = 1 - u$ とおくと,

$$B(\beta, \alpha) = \int_0^1 t^{\beta-1} (1-t)^{\alpha-1} dt = \int_1^0 (1-u)^{\beta-1} u^{\alpha-1} (-du) = \int_0^1 u^{\alpha-1} (1-u)^{\beta-1} du = B(\alpha, \beta).$$

3.4 問題: ベータ分布の三角函数を使った表示

次を示せ:

$$B(\alpha, \beta) = 2 \int_0^{\pi/2} (\cos \theta)^{2\alpha-1} (\sin \theta)^{2\beta-1} d\theta \quad (\alpha, \beta > 0).$$

解答例: $\alpha, \beta > 0$ と仮定する. $t = \cos^2 \theta$ とおくと, $dt = -2 \cos \theta \sin \theta d\theta$ なので,

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt = \int_{\pi/2}^0 (\cos \theta)^{2(\alpha-1)} (\sin \theta)^{2(\beta-1)} (-2 \cos \theta \sin \theta d\theta) =$$

$$= 2 \int_0^{\pi/2} (\cos \theta)^{2\alpha-1} (\sin \theta)^{2\beta-1} d\theta.$$

解答終

注意: 特に

$$B(1/2, 1/2) = 2 \int_0^{\pi/2} d\theta = \pi.$$

となることがわかる. 次の積分は **Wallis積分** と呼ばれている:

$$\int_0^{\pi/2} \sin^n \theta d\theta = \frac{1}{2} B(1/2, (n+1)/2).$$

3.5 ベータ函数の別の基本的表示

統計学的にはベータ函数の次の表示も重要である:

$$B(\alpha, \beta) = \int_0^\infty \frac{u^{\alpha-1} du}{(1+u)^{\alpha+\beta}} = \int_0^\infty \frac{u^{\beta-1} du}{(1+u)^{\alpha+\beta}}.$$

この表示は t 分布(Studentの t 分布)と関係している.

上の表示は以下のようにして証明される.

$t = u/(1+u) = 1 - 1/(1+u)$ とおくと, $dt = du/(1+u)^2$ となるので,

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt = \int_0^\infty \left(\frac{u}{1+u}\right)^{\alpha-1} \left(\frac{1}{1+u}\right)^{\beta-1} \frac{du}{(1+u)^2} = \int_0^\infty \frac{u^{\alpha-1} du}{(1+u)^{\alpha+\beta}}$$

となる. ベータ函数は α と β の交換で不変なので被積分函数の分子の α を β で置き換えても等しくなる.

注意: $0 < t < 1$ と $0 < u < \infty$ と $-\infty < x < \infty$ のあいだの変数変換

$$t = \frac{u}{1+u} = \frac{1}{1+e^{-x}} = \text{logistic}(x), \quad \frac{t}{1-t} = u = e^x, \quad \text{logit}(t) = \log \frac{t}{1-t} = \log u = x$$

は統計学に関連した計算で非常によく現れる鉄板の基本パターンの一つである. t を確率とみなすとき, $u = t/(1-t)$ をその **オッズ** と呼び, $x = \log(t/(1-t))$ をその **ロジット変換** と呼ぶ.

3.6 問題: t 分布の基礎になる公式

次を示せ:

$$\int_{-\infty}^\infty \left(1 + \frac{t^2}{v}\right)^{-(v+1)/2} dt = \sqrt{v} B(1/2, v/2) \quad (v > 0).$$

解答例: $v > 0$ と仮定する.

$$\begin{aligned} \int_{-\infty}^\infty \left(1 + \frac{t^2}{v}\right)^{-(v+1)/2} dt &= 2 \int_0^\infty \left(1 + \frac{t^2}{v}\right)^{-(v+1)/2} dt = 2\sqrt{v} \int_0^\infty (1+s^2)^{-(v+1)/2} ds \\ &= \sqrt{v} \int_0^\infty (1+u)^{-(v+1)/2} u^{-1/2} du = \sqrt{v} \int_0^\infty \frac{u^{1/2-1} du}{(1+u)^{1/2+v/2}} = \sqrt{v} B(1/2, v/2). \end{aligned}$$

最初の等号で被積分函数が偶函数であることを使った. 2つめの等号で $t = \sqrt{v}s$ とおいた. 3つめの等号で $s = \sqrt{u}$, $ds = u^{-1/2} du/2$ とおいた. 最後の等号でベータ函数の別の基本的表示を使った.

解答終

注意: 上の計算での積分変数の変換では $t = \sqrt{vu}$, $u = t^2/v$ となっている.

3.7 t 分布の定義

$v > 0$ と仮定する.

次の確率密度函数によって定義される連続分布を **自由度** (degree of freedom) が v の **t 分布** (Student の t 分布, Student's t -distribution) と呼び, $\text{TDist}(v)$ と表す:

$$p(t|\nu) = \frac{1}{\sqrt{\nu} B(1/2, \nu/2)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}.$$

t 分布は正規分布モデルによる平均に関する検定や推定で基本的な役目を果たすことになる. この点については後で別のノートで詳しく説明する.

3.8 問題: t 分布の自由度無限大の極限で標準正規分布が得られること

t 分布の確率密度関数 $p(t|\nu)$ について次を示せ:

$$\lim_{\nu \rightarrow \infty} p(t|\nu) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

解答例: $t \in \mathbb{R}$ であるとする.

$$p(t|\nu) = \frac{1}{\sqrt{\nu} B(1/2, \nu/2)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

なので, $\nu \rightarrow \infty$ のとき, $\sqrt{\nu} B(1/2, \nu/2) \rightarrow \sqrt{2\pi}$ かつ $(1 + t^2/\nu)^{-(\nu+1)/2} \rightarrow e^{-t^2/2}$ となることを示せばよい.

$N \rightarrow \infty$ のとき $(1 + x/N)^N \rightarrow e^x$ より, $\nu \rightarrow \infty$ とすると,

$$\left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2} = \left(1 + \frac{t^2/2}{\nu/2}\right)^{-\nu/2} \left(1 + \frac{t^2/2}{\nu/2}\right)^{-1/2} \rightarrow e^{-t^2/2}.$$

$u = x/(\nu/2)$ とおくと, $du/u = dx/x$ なので,

$$\begin{aligned} \sqrt{\nu} B(1/2, \nu/2) &= \sqrt{\nu} \int_0^\infty \frac{u^{1/2-1} du}{(1+u)^{1/2+\nu/2}} = \frac{\sqrt{\nu}}{(\nu/2)^{1/2}} \int_0^\infty \frac{x^{1/2-1} dx}{(1+x/(\nu/2))^{1/2+\nu/2}} \\ &= \sqrt{2} \int_0^\infty \frac{x^{1/2-1} dx}{(1+x/(\nu/2))^{1/2+\nu/2}} \end{aligned}$$

なので $\nu \rightarrow \infty$ とすると,

$$\sqrt{\nu} B(1/2, \nu/2) \rightarrow \sqrt{2} \int_0^\infty \frac{x^{1/2-1} dx}{e^x} = \sqrt{2} \int_0^\infty e^{-x} x^{1/2-1} dx = \sqrt{2} \Gamma(1/2) = \sqrt{2\pi}.$$

解答終

注意: t 分布という新しい複雑に見える確率分布を定義したが, 自由度が大きな場合には, t 分布の代わりに正規分布を使った議論と実践的には(近似的には)同じ結果が得られることになる. 自由度が大体 30 程度で実践的には t 分布はほぼ正規分布だと思ってよい場合が多い. t 分布の自由度はほぼ標本サイズになることが多いので, 標本サイズが大きな場合には複雑に見える t 分布の代わりに正規分布を使っても大差ないことになる. このように, ざくっとどんぶり勘定することによって, 複雑に見える議論を単純化することは統計学を実用的に使うときに重要である.

3.9 ベータ関数の極限としてガンマ関数を表せること

$\alpha > 0, b \in \mathbb{R}$ であると仮定する. このとき, 次が成立している:

$$\lim_{n \rightarrow \infty} n^\alpha B(\alpha, n+b) = \Gamma(\alpha).$$

積分変数を $t = x/n$ で x に変換すると,

$$\begin{aligned} n^\alpha B(\alpha, n+b) &= n^\alpha \int_0^1 t^{\alpha-1} (1-t)^{n+b-1} dt \\ &= n^\alpha \int_0^n \left(\frac{x}{n}\right)^{\alpha-1} \left(1 - \frac{x}{n}\right)^{n+b-1} \frac{dx}{n} = \int_0^n x^{\alpha-1} \left(1 - \frac{x}{n}\right)^{n+b-1} dx \end{aligned}$$

なので, $n \rightarrow \infty$ のとき, $(1 - x/n)^{n+b-1} \rightarrow e^{-x}$ となることを使うと,

$$n^\alpha B(\alpha, n+b) \rightarrow \int_0^\infty x^{\alpha-1} e^{-x} dx = \Gamma(\alpha).$$

例: t 分布の自由度無限大の極限の計算で使った $\sqrt{\nu} B(1/2, \nu/2)$ の $\nu \rightarrow \infty$ の極限は上の結果を使うと以下のようにして計算される:

$$\sqrt{\nu} B(1/2, \nu/2) = \sqrt{2} \sqrt{\nu/2} B(1/2, \nu/2) \rightarrow \sqrt{2} \Gamma(1/2) = \sqrt{2\pi}.$$

4 ベータ函数のガンマ函数表示

ベータ函数をガンマ函数で表せることから、ガンマ分布とベータ分布の関係が得られる。

$\alpha, \beta > 0$ について次が成立している:

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

これを示すためには

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

を示せばよい。

4.1 積分 $J[f]$ を導入とベータ函数のガンマ函数表示の証明

後で、ガンマ分布とベータ分布を関係付けるために、函数 $f(z, t)$ について次の積分を考える:

$$J[f] = \int_0^\infty \int_0^\infty f\left(x + y, \frac{x}{x + y}\right) e^{-(x+y)} x^{\alpha-1} y^{\beta-1} dx dy.$$

このとき、次が成立している:

$$J[f] = \int_0^\infty \left(\int_0^1 f(z, t) e^{-z} z^{\alpha+\beta-1} t^{\alpha-1} (1-t)^{\beta-1} dt \right) dz$$

もしもこの公式が証明されたならば、 $J[f]$ の定義より、 $f(z, t) = 1$ のとき、

$$\begin{aligned} J[1] &= \int_0^\infty \int_0^\infty e^{-(x+y)} x^{\alpha-1} y^{\beta-1} dx dy \\ &= \int_0^\infty e^{-x} x^{\alpha-1} dx \cdot \int_0^\infty e^{-y} y^{\beta-1} dy = \Gamma(\alpha)\Gamma(\beta) \end{aligned}$$

となり、上の公式より、

$$\begin{aligned} J[1] &= \int_0^\infty \left(\int_0^1 e^{-z} z^{\alpha+\beta-1} t^{\alpha-1} (1-t)^{\beta-1} dt \right) dz \\ &= \int_0^\infty e^{-z} z^{\alpha+\beta-1} dz \cdot \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt = \Gamma(\alpha + \beta)B(\alpha, \beta). \end{aligned}$$

以上を比較すれば $\Gamma(\alpha)\Gamma(\beta) = \Gamma(\alpha + \beta)B(\alpha, \beta)$ が得られる。

4.2 積分 $J[f]$ に関する公式の証明

前節で説明した $J[f]$ に関する公式は以下のようにして示される。積分変数 y を $y = z - x$ によって z に変換すると、 $0 < y < \infty$ と $x < z < \infty$ のあいだの一一対応が得られるので、

$$\begin{aligned} J[f] &= \int_0^\infty \left(\int_0^\infty f(x + y, x/(x + y)) e^{-(x+y)} x^{\alpha-1} y^{\beta-1} dy \right) dx \\ &= \int_0^\infty \left(\int_x^\infty f(z, x/z) e^{-z} x^{\alpha-1} (z - x)^{\beta-1} dz \right) dx. \end{aligned}$$

この二重積分で (x, z) の動く範囲は $0 < x < z < \infty$ である。その範囲で (x, z) を動かすためには、まず $0 < z < \infty$ を満たす z を与えてから、 $0 < x < z$ を満たす x を与えればよい。ゆえにその二重積分の積分順序を次のように交換できる:

$$J[f] = \int_0^\infty \left(\int_0^z f(z, x/z) e^{-z} x^{\alpha-1} (z - x)^{\beta-1} dx \right) dz.$$

積分変数 x を $x = zt$ によって t に変換すると、 $0 < x < z$ と $0 < t < 1$ のあいだの一一対応が得られるので、

$$J[f] = \int_0^\infty \left(\int_0^1 f(z, t) e^{-z} (zt)^{\alpha-1} (z(1-t))^{\beta-1} z dt \right) dz$$

この二重積分の被積分函数は z, t のそれぞれの単独変数函数の積の形になっており、次のように整理される:

$$J[f] = \int_0^\infty \left(\int_0^1 f(z, t) e^{-z} z^{\alpha+\beta-1} t^{\alpha-1} (1-t)^{\beta-1} dt \right) dz.$$

これで示したい公式が示された。

4.3 問題: ガンマ分布とベータ分布の関係

$\alpha, \beta > 0$ と仮定する。正の実数に値を持つ確率変数 X, Y の組 (X, Y) の確率密度関数は、ガンマ分布 $\text{Gamma}(\alpha, 1)$, $\text{Gamma}(\beta, 1)$ の確率密度関数の積になっていると仮定する:

$$p(x, y) = \frac{e^{-x} x^{\alpha-1}}{\Gamma(\alpha)} \frac{e^{-y} y^{\beta-1}}{\Gamma(\beta)} = \frac{e^{-(x+y)} x^{\alpha-1} y^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \quad (x, y > 0)$$

この $p(x, y)$ を確率変数 X, Y の **同時確率密度関数** (simultaneous probability density function) と呼ぶ。同時確率密度関数がこのように各々の確率変数の確率密度関数の積になっているとき、それらの確率変数達は **独立** (independent) であるという。以上の設定を次のように書くことにする:

$$X \sim \text{Gamma}(\alpha, 1), \quad Y \sim \text{Gamma}(\beta, 1) \quad (\text{independent})$$

このとき,

$$X + Y \sim \text{Gamma}(\alpha + \beta, 1), \quad \frac{X}{X + Y} \sim \text{Beta}(\alpha, \beta) \quad (\text{independent})$$

となることを示せ。すなわち, $Z = X + Y, T = X/(X + Y) = X/Z$ とおき, 分布 $\text{Gamma}(\alpha + \beta, 1)$ の確率密度関数を $p_Z(z)$ と書き, 分布 $\text{Beta}(\alpha, \beta)$ の確率密度関数を $p_T(t)$ と書くとき, 次が成立することを示せ:

$$E[f(Z, T)] = \int_0^\infty \left(\int_0^1 f(z, t) p_Z(z) p_T(t) dt \right) dz.$$

解答例: 分布 $\text{Gamma}(\alpha + \beta, 1)$ の確率密度関数 $p_Z(z)$ と分布 $\text{Beta}(\alpha, \beta)$ の確率密度関数を $p_T(t)$ はそれらの定義より次の形をしている:

$$p_Z(z) = \frac{e^{-z} z^{\alpha+\beta-1}}{\Gamma(\alpha + \beta)}, \quad p_T(t) = \frac{t^{\alpha-1} (1-t)^{\beta-1}}{B(\alpha, \beta)} \quad (z > 0, 0 < t < 1)$$

確率変数の期待値の定義と前節の証明で使った $J[f]$ の定義を使うと,

$$\begin{aligned} E[f(Z, T)] &= \int_0^\infty \int_0^1 f(x + y, x/(x + y)) p(x, y) dx dy \\ &= \int_0^\infty \int_0^1 f(x + y, x/(x + y)) \frac{e^{-(x+y)} x^{\alpha-1} y^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} dx dy = \frac{J[f]}{\Gamma(\alpha)\Gamma(\beta)} =. \end{aligned}$$

$\Gamma(\alpha)\Gamma(\beta) = \Gamma(\alpha + \beta)B(\alpha, \beta)$ と前節で示した $J[f]$ に関する公式と $p_Z(z), p_T(t)$ の定義を使うと,

$$\begin{aligned} E[f(Z, T)] &= \frac{J[f]}{\Gamma(\alpha + \beta)B(\alpha, \beta)} \\ &= \int_0^\infty \left(\int_0^1 f(z, t) \frac{e^{-z} z^{\alpha+\beta-1}}{\Gamma(\alpha + \beta)} \frac{t^{\alpha-1} (1-t)^{\beta-1}}{B(\alpha, \beta)} dt \right) dz \\ &= \int_0^\infty \left(\int_0^1 f(z, t) p_Z(z) p_T(t) dt \right) dz. \end{aligned}$$

これで示すべきことが示された。

解答終

注意: 上の問題の結果はガンマ分布 $\text{Gamma}(\alpha, 1), \text{Gamma}(\beta, 1)$ に従う独立な確率変数 X, Y からガンマ分布 $\text{Gamma}(\alpha + \beta, 1)$ に従う確率変数を $X + Y$ の形で作れてかつ, ベータ分布 $\text{Beta}(\alpha, \beta)$ に従う確率変数を $X/(X + Y)$ の形で作れることを示した。

これは, ガンマ分布 $\text{Gamma}(\alpha, 1), \text{Gamma}(\beta, 1)$ に従う乱数達から, ガンマ分布 $\text{Gamma}(\alpha + \beta, 1)$ に従う乱数とベータ分布 $\text{Beta}(\alpha, \beta)$ に従う乱数を作れることを意味している。

上のように積分の煩雑な計算をするだけだと何をやっているのかがピンとこない人が多く出て来るかもしれないが, 確率変数達の以上のような関係は乱数の生成法のあいだの関係のことだと了解しておけば結果的に何をやっているのかわかりやすくなる。

そして, 以下のようにコンピュータによる計算によって数学的結果の正しさを数値的に確認することもできる。

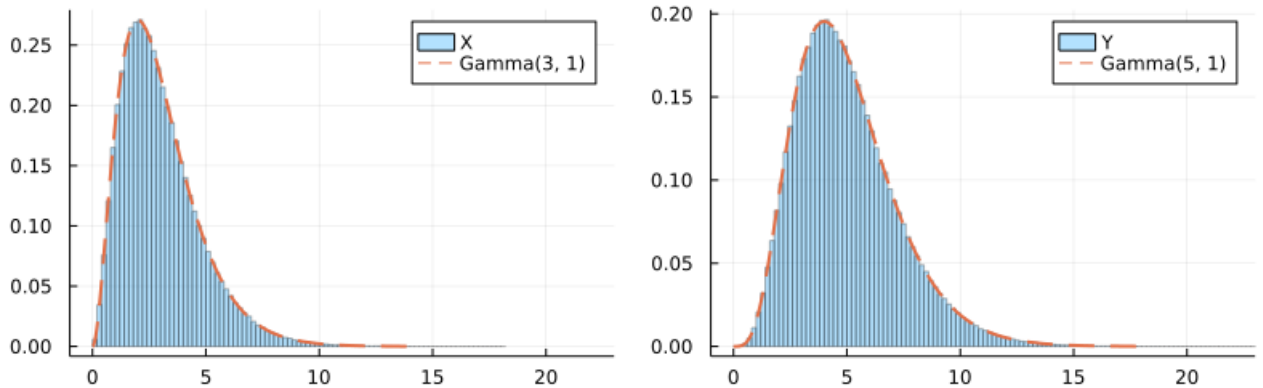
ガンマ分布達の乱数を大量に生成。

```
In [20]: 1 α, β = 3, 5
2 n = 10^6
3 X = rand(Gamma(α, 1), n)
4 Y = rand(Gamma(β, 1), n);
```

ガンマ関数達の乱数達 X, Y の分布の確認.

```
In [21]: 1 P1 = histogram(X; norm=true, alpha=0.3, bin=140, xlim=(-1, 23), label="X")
2 plot!(Gamma(α, 1); label="Gamma($α, 1)", lw=2, ls=:dash)
3
4 P2 = histogram(Y; norm=true, alpha=0.3, bin=140, xlim=(-1, 23), label="Y")
5 plot!(Gamma(β, 1); label="Gamma($β, 1)", lw=2, ls=:dash)
6
7 plot(P1, P2; size=(800, 250))
```

Out[21]:



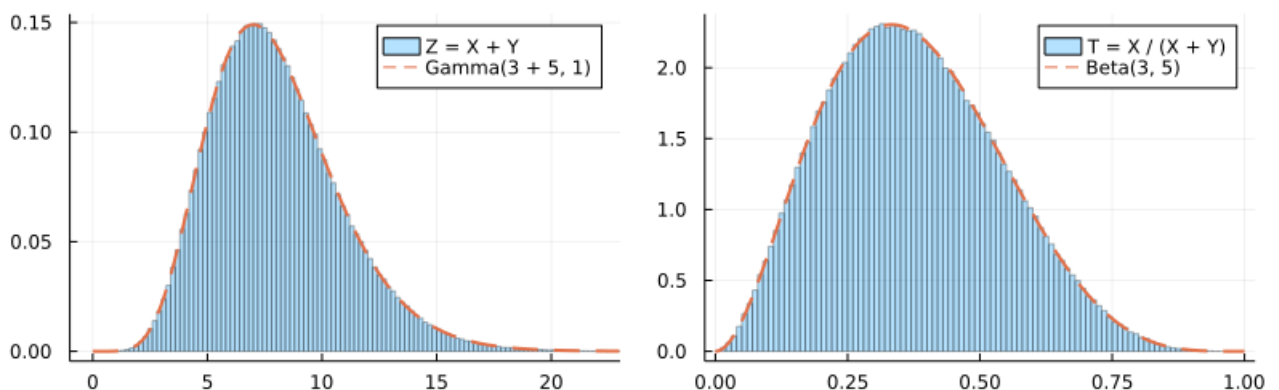
ガンマ関数の乱数達から Z と T を作成.

```
In [22]: 1 Z = @. X + Y
2 T = @. X / Z;
```

Z, T の分布達がそれぞれガンマ分布とベータ分布になっていることの確認.

```
In [23]: 1 P3 = histogram(Z; norm=true, alpha=0.3, bin=140, xlim=(-1, 23), label="Z = X + Y")
2 plot!(Gamma(α + β, 1); label="Gamma($α + $β, 1)", lw=2, ls=:dash)
3
4 P4 = histogram(T; norm=true, alpha=0.3, bin=140, xlim=(-0.02, 1.02), label="T = X / (X + Y)")
5 plot!(Beta(α, β); label="Beta($α, $β)", lw=2, ls=:dash)
6
7 plot(P3, P4; size=(800, 250))
```

Out[23]:



4.4 問題: ガンマ関数のGaussの乗法公式

ベータ関数のガンマ関数表示とガンマ関数の函数等式とベータ関数の極限でガンマ関数を表せることを用いて, 次の公式を証明せよ:

$$\lim_{n \rightarrow \infty} \frac{n^\alpha n!}{\alpha(\alpha+1) \cdots (\alpha+n)} = \Gamma(\alpha).$$

解答例: ガンマ関数の関数等式より $\Gamma(\alpha + n + 1) = (\alpha + n) \cdots (\alpha + 1)\alpha\Gamma(\alpha)$ なので, $\Gamma(n + 1) = n!$ とベータ関数のガンマ関数表示より,

$$n^\alpha B(\alpha, n + 1) = \frac{n^\alpha \Gamma(\alpha) \Gamma(n + 1)}{\Gamma(\alpha + n + 1)} = \frac{n^\alpha \Gamma(\alpha) n!}{(\alpha + n) \cdots (\alpha + 1)\alpha\Gamma(\alpha)} = \frac{n^\alpha n!}{\alpha(\alpha + 1) \cdots (\alpha + n)}.$$

ベータ関数の極限でガンマ関数が表せることより, $n \rightarrow \infty$ のとき,

$$\frac{n^\alpha n!}{\alpha(\alpha + 1) \cdots (\alpha + n)} = n^\alpha B(\alpha, n + 1) \rightarrow \Gamma(\alpha)$$

が得られる.

解答終

注意: 上の公式をガンマ関数に関する **Gaussの乗法公式** と呼ぶ. 上の公式の極限は α が0以下の整数以外の任意の複素数であっても収束している.

4.5 問題: ガンマ関数の相反公式(reflection formula)

sin の無限積表示

$$\sin(\pi\alpha) = \pi\alpha \prod_{k=1}^{\infty} \left(1 - \frac{\alpha^2}{k^2}\right)$$

を認めて使って, ガンマ関数のGaussの乗法公式から次の公式を導け:

$$B(\alpha, 1 - \alpha) = \Gamma(\alpha)\Gamma(1 - \alpha) = \frac{\pi}{\sin(\pi\alpha)}.$$

これをガンマ関数の **相反公式** (reflection formula)と呼ぶ.

解答例: 相反公式の両辺に α をかけて, $\alpha\Gamma(\alpha) = \Gamma(1 + \alpha)$ を使った後に, 逆数を取ることで得られる次の公式を示せば十分である:

$$\frac{1}{\Gamma(1 + \alpha)\Gamma(1 - \alpha)} = \frac{\sin(\pi\alpha)}{\pi\alpha}.$$

sin の無限積表示を認めて使うと, この公式は次と同値である:

$$\frac{1}{\Gamma(1 + \alpha)\Gamma(1 - \alpha)} = \prod_{k=1}^{\infty} \left(1 - \frac{\alpha^2}{k^2}\right).$$

これをGaussの乗法公式から導けばよい.

Gaussの情報公式より, 次の式は $n \rightarrow \infty$ で $1/(\Gamma(1 + \alpha)\Gamma(1 - \alpha))$ に収束する:

$$\begin{aligned} & \frac{(1 + \alpha)(2 + \alpha) \cdots (n + 1 + \alpha)}{n^{1+\alpha} n!} \frac{(1 - \alpha)(2 - \alpha) \cdots (n + 1 - \alpha)}{n^{1-\alpha} n!} \\ &= \frac{(n + 1)^2}{n^2} \left(1 + \frac{\alpha}{1}\right) \left(1 + \frac{\alpha}{2}\right) \cdots \left(1 + \frac{\alpha}{n + 2}\right) \left(1 - \frac{\alpha}{1}\right) \left(1 - \frac{\alpha}{2}\right) \cdots \left(1 - \frac{\alpha}{n + 2}\right) \\ &= \frac{(n + 1)^2}{n^2} \prod_{k=1}^{n+1} \left(1 - \frac{\alpha^2}{k^2}\right) \end{aligned}$$

最初の等号で, 分母の n^α と $n^{-\alpha}$ をキャンセルさせて, 分子分母を $(n + 1)! \times (n + 1)!$ で割った.

上の最後の式は $n \rightarrow \infty$ で次に

$$\prod_{k=1}^{\infty} \left(1 - \frac{\alpha^2}{k^2}\right)$$

に収束する. これで示すべきことが示された.

解答終

4.6 ディガンマ関数, トリガンマ関数, ポリガンマ関数, Hurwitzのゼータ関数

統計学関係の計算ではガンマ関数やベータ関数の導関数達が必要になる場合がある。ベータ関数はガンマ関数で表示されるのでベータ関数の偏導関数の計算はガンマ関数の導関数の計算に帰着する。ガンマ関数の導関数が必要になる場合として、ガンマ分布の最尤法や二項分布モデルやPoisson分布モデルのベイズ統計における情報量規準の計算などがある。

コンピュータでの基本特殊関数ライブラリの中にはガンマ関数の導関数の計算のために必要な **ディガンマ関数** (digamma function), **トリガンマ関数** (trigamma function), **ポリガンマ関数** (polygamma function)が含まれている。それらはガンマ関数そのものではなく、ガンマ関数の対数の導関数として定義される。ディガンマ関数 $\psi(\alpha)$ は

$$\psi(\alpha) = \frac{d}{d\alpha} \log \Gamma(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$$

と定義される。この導関数

$$\psi'(\alpha) = \frac{\Gamma(\alpha)\Gamma''(\alpha) - \Gamma'(\alpha)^2}{\Gamma(\alpha)^2}$$

をトリガンマ関数と呼び、高階の導関数 $\psi^{(k)}(\alpha)$ をポリガンマ関数と呼ぶ。

ガンマ関数の函数等式 $\gamma(\alpha + 1) = \alpha\Gamma(\alpha)$ の両辺の対数を微分することによって、

$$\psi(\alpha + 1) = \psi(\alpha) + \frac{1}{\alpha}, \quad \psi'(\alpha + 1) = \psi'(\alpha) - \frac{1}{\alpha^2}, \quad \psi''(\alpha + 1) = \psi''(\alpha) + \frac{2}{\alpha^3}, \quad \dots \quad (*)$$

ガンマ関数のGaussの乗法公式

$$\lim_{n \rightarrow \infty} \frac{n^\alpha n!}{\alpha(\alpha + 1) \cdots (\alpha + n)} = \Gamma(\alpha).$$

と

$$\frac{d}{d\alpha} \log \frac{n^\alpha n!}{\alpha(\alpha + 1) \cdots (\alpha + n)} = \log n - \sum_{k=0}^n \frac{1}{\alpha + k}$$

より、

$$\psi(\alpha) = \lim_{n \rightarrow \infty} \left(\log n - \sum_{k=0}^n \frac{1}{\alpha + k} \right), \quad \psi'(\alpha) = \sum_{k=0}^{\infty} \frac{1}{(\alpha + k)^2}, \quad \psi''(\alpha) = -2 \sum_{k=0}^{\infty} \frac{1}{(\alpha + k)^3}, \quad \dots \quad (**)$$

この(**)から上の(*)が成立することを確認することもできる。

調和級数 $\sum_{n=1}^N 1/n$ は $N \rightarrow \infty$ で無限大に発散するが、それから $\log N$ を引いたものは有限の値 $\gamma \approx 0.5772$ に収束することが知られている。この γ を **Eulerの γ** (オイラーの γ , Euler's gamma)と呼ぶことがある。(**)の最初の式より、

$$-\psi(1) = \lim_{N \rightarrow \infty} \left(\sum_{n=1}^{N+1} \frac{1}{n} - \log N \right) = \lim_{N \rightarrow \infty} \left(\sum_{n=1}^N \frac{1}{n} - \log N \right) = \gamma \approx 0.5772.$$

これと(*)の最初の式から、次が得られる:

$$\psi(n + 1) = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} - \gamma \quad (n = 0, 1, 2, \dots)$$

ガンマ関数は階乗 $\Gamma(n + 1) = n!$ の一般化になっていたが、ディガンマ関数はEulerの γ の情報を含む調和級数の一般化になっていると考えられる。

Hurwitzのゼータ関数 (フルヴィッツのゼータ関数)が

$$\zeta(s, \alpha) = \sum_{k=0}^{\infty} \frac{1}{(\alpha + k)^s}$$

と定義される。トリガンマ関数以降のポリガンマ関数達は、定数倍の違いを除いて、Hurwitzのゼータ関数で s を2以上の整数にした場合になっている:

$$\psi^{(k)}(\alpha) = (-1)^{k+1} k! \zeta(k + 1, \alpha) \quad (k = 1, 2, 3, \dots)$$

$\zeta(s) = \zeta(s, 1) = \sum_{n=1}^{\infty} n^{-s}$ は **Riemannのゼータ関数** (リーマンのゼータ関数)と呼ばれている。ゼータ関数達は数論の基本的研究対象である。ガンマ関数のよく使われる形から得られる

$$\frac{1}{(\alpha + k)^s} = \frac{1}{\Gamma(s)} \int_0^{\infty} e^{-(\alpha+k)t} t^{s-1} dt$$

という公式を使えば、Hurwitzのゼータ関数はガンマ関数のような積分表示を持つこともわかる:

5.1 Dirichlet積分=多変数ベータ関数の定義とそのガンマ関数表示

Dirichlet積分 は次の積分のことである:

$$B(\alpha_1, \dots, \alpha_{n+1}) = \int \cdots \int_{\Delta_n} t_1^{\alpha_1-1} \cdots t_n^{\alpha_n-1} (1 - (t_1 + \cdots + t_n))^{\alpha_{n+1}-1} dt_1 \cdots dt_n \quad (\alpha_i > 0)$$

ここで Δ_n は次の n 次元領域である:

$$\Delta_n = \{ (t_1, \dots, t_n) \in \mathbb{R}_{>0}^n \mid t_1 + \cdots + t_n < 1 \}.$$

$B(\alpha_1, \dots, \alpha_{n+1})$ を **多変数ベータ関数** とも呼ぶことにする. 多変数ベータ関数は次のガンマ関数表示を持つ.

$$B(\alpha_1, \dots, \alpha_{n+1}) = \frac{\Gamma(\alpha_1) \cdots \Gamma(\alpha_{n+1})}{\Gamma(\alpha_1 + \cdots + \alpha_{n+1})}.$$

5.2 Dirichlet分布の定義

$\alpha_1, \dots, \alpha_r > 0$ と仮定する. このとき,

$$\Delta_{r-1} = \{ (t_1, \dots, t_{r-1}) \in \mathbb{R}_{>0}^{r-1} \mid t_1 + \cdots + t_{r-1} < 1 \}$$

上の確率密度関数を

$$p(t_1, \dots, t_{r-1} \mid \alpha_1, \dots, \alpha_r) = \frac{t_1^{\alpha_1-1} \cdots t_{r-1}^{\alpha_{r-1}-1} (1 - (t_1 + \cdots + t_{r-1}))^{\alpha_r-1}}{B(\alpha_1, \dots, \alpha_r)}$$

と定めることができる. この密度関数が定める多変量連続分布をDirichlet分布と呼び,

$$\text{Dirichlet}(\alpha_1, \dots, \alpha_r)$$

と表す. 対称性のためには

$$t_r = 1 - (t_1 + \cdots + t_{r-1})$$

とにおいて,

$$\Delta^{r-1} = \{ (t_1, \dots, t_r) \in \mathbb{R}_{>0}^r \mid t_1 + \cdots + t_r = 1 \}$$

上の確率分布だとみなした方がよい. これによって, Dirichlet分布は, カテゴリカル分布

$$\text{Categorical}(t_1, \dots, t_r) \quad (t_i > 0, t_1 + \cdots + t_r = 1)$$

および多項分布

$$\text{Multinomial}(n, t_1, \dots, t_r) \quad (t_i > 0, t_1 + \cdots + t_r = 1)$$

のパラメータ (t_1, \dots, t_r) に関する確率分布だとみなされる. Dirichlet分布はカテゴリカル分布や多項分布の共役事前分布として重要である.

5.3 多変数ベータ関数のガンマ関数表示の証明

$z = x_1 + \cdots + x_{n+1}$, $t_i = x_i/z$ とおく. 関数 $f(z, t_1, \dots, t_n)$ について次の積分を考える:

$$J[f] = \int_0^\infty \cdots \int_0^\infty f(z, t_1, \dots, t_n) e^{-(x_1 + \cdots + x_{n+1})} x_1^{\alpha_1-1} \cdots x_{n+1}^{\alpha_{n+1}-1} dx_1 \cdots dx_{n+1}.$$

本当は z, t_i に $z = x_1 + \cdots + x_{n+1}$, $t_i = x_i/z$ を代入した式を書きたかったのだが, 式が煩雑になり過ぎたのでそうしていない. この式の z, t にはそのように代入してから積分を行うという意味だと解釈する必要がある. そのとき

$$J[f] = \int_0^\infty \cdots \int_0^\infty \left(\int_0^\infty f(z, t_1, \dots, t_n) e^{-(x_1 + \cdots + x_{n+1})} x_1^{\alpha_1-1} \cdots x_{n+1}^{\alpha_{n+1}-1} dx_{n+1} \right) dx_1 \cdots dx_n$$

における内側の積分変数 x_{n+1} を $x_{n+1} = z - (x_1 + \cdots + x_n)$ を使って z に変換すると, z の動く範囲は $x_1 + \cdots + x_n < z < \infty$ になるので,

$$J[f] = \int_0^\infty \cdots \int_0^\infty \left(\int_{x_1 + \cdots + x_n}^\infty f(z, t_1, \dots, t_n) e^{-z} x_1^{\alpha_1-1} \cdots x_n^{\alpha_n-1} (z - (x_1 + \cdots + x_n))^{\alpha_{n+1}-1} dz \right) dx_1 \cdots dx_n.$$

この $n+1$ 重積分において (x_1, \dots, x_n, z) は $x_1, \dots, x_n > 0$ かつ $x_1 + \cdots + x_n < z$ を満たす範囲を動く. ゆえに,

$$\Delta_n(z) = \{ (x_1, \dots, x_n) \in \mathbb{R}_{>0}^n \mid x_1 + \cdots + x_n < z \}$$

とおくと, 上の $n+1$ 重積分は次のように積分順序を交換できる:

$$J[f] = \int_0^\infty \left(\int \cdots \int_{\Delta_n(z)} f(z, t_1, \dots, t_n) e^{-z} x_1^{\alpha_1-1} \cdots x_n^{\alpha_n-1} (z - (x_1 + \cdots + x_n))^{\alpha_{n+1}-1} dx_1 \cdots dx_n \right) dz.$$

積分変数 x_1, \dots, x_n を $x_i = zt_i$ によって t_1, \dots, t_n に変換すると,

$$J[f] = \int_0^\infty \left(\int \cdots \int_{\Delta_n} f(z, t_1, \dots, t_n) e^{-z} (zt_1)^{\alpha_1-1} \cdots (zt_n)^{\alpha_n-1} (z(1 - (t_1 + \cdots + t_n)))^{\alpha_{n+1}-1} (z dt_1) \cdots (z dt_n) \right) dz.$$

これは次のように整理される:

$$J[f] = \int_0^\infty \left(\int \cdots \int_{\Delta_n} f(z, t_1, \dots, t_n) e^{-z} z^{\alpha_1+\cdots+\alpha_{n+1}-1} t_1^{\alpha_1-1} \cdots t_n^{\alpha_n-1} (1 - (t_1 + \cdots + t_n))^{\alpha_{n+1}-1} dt_1 \cdots dt_n \right) dz.$$

これが $J[f]$ について示したかった公式である.

特に $f = 1$ のとき, $J[f]$ の定義より

$$J[1] = \Gamma(\alpha_1) \cdots \Gamma(\alpha_{n+1})$$

であることがわかり, 上で示した公式より,

$$J[1] = \Gamma(\alpha_1 + \cdots + \alpha_{n+1}) B(\alpha_1, \dots, \alpha_{n+1})$$

となることがわかる. ゆえに

$$\Gamma(\alpha_1) \cdots \Gamma(\alpha_{n+1}) = \Gamma(\alpha_1 + \cdots + \alpha_{n+1}) B(\alpha_1, \dots, \alpha_{n+1}).$$

すなわち

$$B(\alpha_1, \dots, \alpha_{n+1}) = \frac{\Gamma(\alpha_1) \cdots \Gamma(\alpha_{n+1})}{\Gamma(\alpha_1 + \cdots + \alpha_{n+1})}.$$

5.4 問題: Dirichlet (1839)の公式の証明

$a_i, p_i, \alpha_i > 0$ と仮定する. 領域 X を

$$X = \left\{ (x_1, \dots, x_n) \in \mathbb{R}_{\geq 0}^n \mid \left(\frac{x_1}{a_1} \right)^{p_1} + \cdots + \left(\frac{x_n}{a_n} \right)^{p_n} \leq 1 \right\}$$

と定める. 次の積分を求めよ:

$$V = \int \cdots \int_X x_1^{\alpha_1-1} \cdots x_n^{\alpha_n-1} dx_1 \cdots dx_n.$$

解答例: X の定義を

$$X = \left\{ (x_1, \dots, x_n) \in \mathbb{R}_{> 0}^n \mid \left(\frac{x_1}{a_1} \right)^{p_1} + \cdots + \left(\frac{x_n}{a_n} \right)^{p_n} < 1 \right\}$$

に置き換えて構わない. (等号部分の体積は 0 なので積分に寄与しない.) このとき,

$$t_i = \left(\frac{x_i}{a_i} \right)^{p_i} \iff x_i = a_i t_i^{1/p_i}$$

という変換によって, 領域 X と領域 Δ_n のあいだの一対一対応が得られる. ゆえに

$$\begin{aligned} V &= \int \cdots \int_{\Delta_n} (a_1 t_1^{1/p_1})^{\alpha_1-1} \cdots (a_n t_n^{1/p_n})^{\alpha_n-1} \left(\frac{a_1}{p_1} t_1^{1/p_1-1} dt_1 \right) \cdots \left(\frac{a_n}{p_n} t_n^{1/p_n-1} dt_n \right) \\ &= \frac{a_1^{\alpha_1} \cdots a_n^{\alpha_n}}{p_1 \cdots p_n} \int \cdots \int_{\Delta_n} t_1^{\alpha_1/p_1-1} \cdots t_n^{\alpha_n/p_n-1} dt_1 \cdots dt_n = \frac{a_1^{\alpha_1} \cdots a_n^{\alpha_n}}{p_1 \cdots p_n} B(\alpha_1/p_1, \dots, \alpha_n/p_n, 1) \\ &= \frac{a_1^{\alpha_1} \cdots a_n^{\alpha_n}}{p_1 \cdots p_n} \frac{\Gamma(\alpha_1/p_1) \cdots \Gamma(\alpha_n/p_n) \Gamma(1)}{\Gamma(\alpha_1/p_1 + \cdots + \alpha_n/p_n + 1)} = \frac{a_1^{\alpha_1} \cdots a_n^{\alpha_n}}{p_1 \cdots p_n} \frac{\Gamma(\alpha_1/p_1) \cdots \Gamma(\alpha_n/p_n)}{\Gamma(\alpha_1/p_1 + \cdots + \alpha_n/p_n + 1)}. \end{aligned}$$

解答終

