

いろいろな確率分布

2022年7月(作成)
倉谷 隆博

1. 母集団 (population) と標本 (sample)

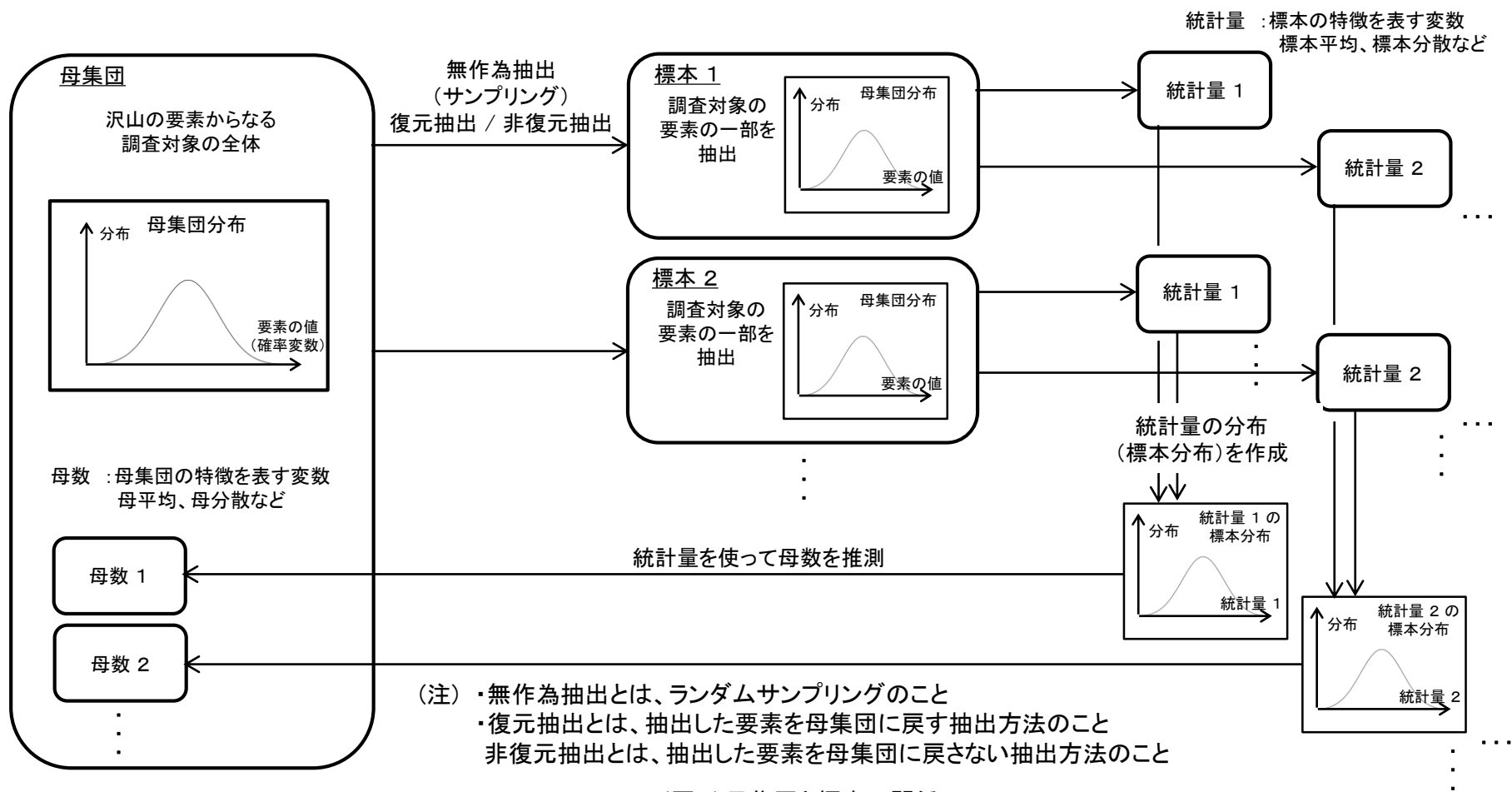
母集団と標本の関係を(図1)に整理しておく。

母集団は沢山の要素からなる調査対象の全体を示し、要素を確率変数として扱い、母集団を確率分布で表す。母集団の特徴は、母集団の平均(母平均)、母集団の分散(母分散)などの「**母数**」で表すことができる。「**母数**」は**母集団に固有のもので固定の値をとる**。

一方、標本は母集団から要素の一部を抽出したものであり、抽出した要素は母集団分布に従う。

標本を作る目的は、抽出した要素を使って母集団の母数を推測することである(推測統計学)。標本の特徴は、標本の要素から計算する「**統計量**」(標本平均、標本分散など)を使って表される。この「**統計量**」は、**標本ごとに異なる値を取る確率変数である**。統計量を使って、母集団の母数を推測する。例えば、標本平均が作る標本平均の確率分布(標本分布)をもとに母平均を推測する。

(注)なお、標本の統計量を使って母集団の特徴を示す、なじみのある統計学は、記述統計学と言われている。

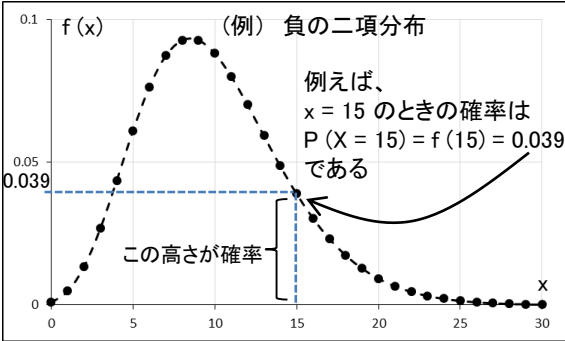
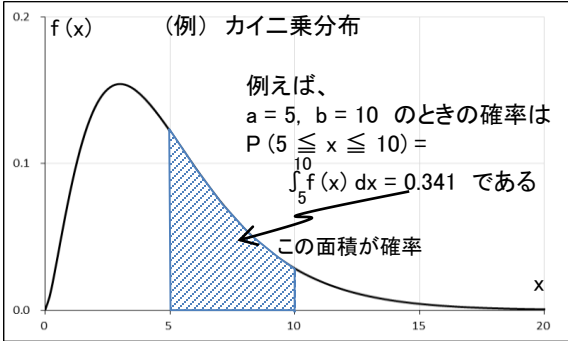


(図1) 母集団と標本の関係

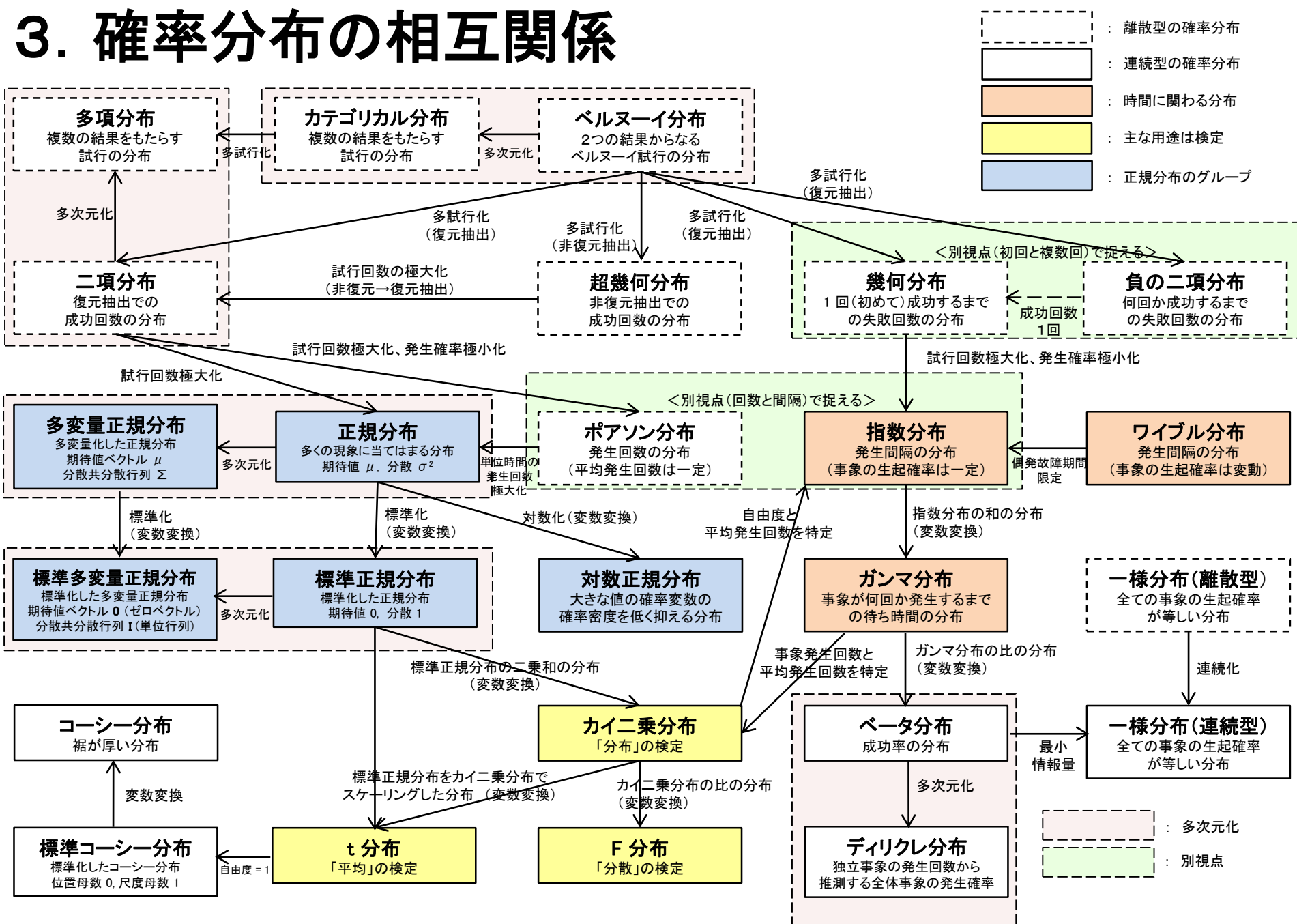
2. 確率質量関数と確率密度関数

確率質量関数は離散型確率変数の確率分布を表す。一方、確率密度関数は連続型確率変数の確率分布を表す。確率質量関数の「質量」という表現は確率が関数の値そのものであることに由来する。一方、確率密度関数の「密度」という表現は関数の値そのものが確率でなく、積分することによって確率が求まることに由来する。

(表1) 確率質量関数と確率密度関数

	確率質量関数 (probability mass function)	確率密度関数 (probability density function)
確率変数のタイプ	離散型 (例: 成功の回数)	連続型 (例: 待ち時間)
確率の算出	$P(X = x) = f(x)$ $f(x)$: 確率質量関数 $P(X = x)$ は、確率変数 X が実現値 x をとる確率を表す。 P は、Probability (確率) を表す。 この式が表すように、確率は関数 $f(x)$ の値そのものである。	$P(a \leq x \leq b) = \int_a^b f(x) dx$ $f(x)$: 確率密度関数 $P(a \leq x \leq b)$ は、確率変数 X の実現値 x が a 以上、 b 以下をとる確率を表す。 この式が表すように、確率は関数 $f(x)$ を区間積分して得られる(面積になる)。
関数の性質	$f(x) \geq 0$ $\sum_{\text{全ての } x} f(x) = 1$ x の取り得る全ての値を対象に $f(x)$ の総和をとる。	$f(x) \geq 0$ x の最大値 $\int_{x \text{ の最小値}} f(x) dx = 1$ $f(x)$ の積分区間は x の最小値($-\infty$)から x の最大値($+\infty$)まで、である。
確率を算出する仕組み	 (例) 負の二項分布 例えば、 $x = 15$ のときの確率は $P(X = 15) = f(15) = 0.039$ である この高さが確率 グラフは離散値(黒色ドット)になるが、 グラフ全体の傾向が分かるように、ドットを破線でつないでいる	 (例) カイ二乗分布 例えば、 $a = 5, b = 10$ のときの確率は $P(5 \leq x \leq 10) = \int_5^{10} f(x) dx = 0.341$ である この面積が確率 カイ二乗分布の面積(確率)の算出には、Excel の chidist 関数を利用

3. 確率分布の相互関係



(図2) 各種確率分布の関係

4-1. 確率分布(離散型)の一覧表

(表2-1) 確率分布(離散型)の一覧表

	名前	確率質量関数(X : 確率変数)	期待値	分散
1	一様分布	$P(X = x) = 1/N$ x : 事象(確率変数) N : 事象の数	$(N + 1) / 2$	$(N^2 - 1) / 12$
2	ベルヌーイ分布	$P(X = x) = p^x (1 - p)^{1-x}$ x : 事象 成功時 x = 1, 失敗時 x = 0 (確率変数) p : 成功確率	p	p (1 - p)
3	カテゴリカル分布	$P(X = x) = \prod_{i=1}^k p_i^{x_i}$ x : カテゴリベクトル 選択時に、 $x_i = 1$, 他の $x_i = 0$ (確率変数) p_i : カテゴリの選択確率 k : カテゴリの種類の数	p_i	$p_i (1 - p_i)$
4	二項分布	$P(X = x) = {}_n C_x p^x (1 - p)^{n-x}$ x : 成功回数(確率変数) n : 試行回数 p : 成功確率	n p	n p (1 - p)
5	多項分布	$P(X_i = x_i) = \{ n! / (x_1! x_2! \cdots x_m!) \} p_1^{x_1} p_2^{x_2} \cdots p_m^{x_m}$ m : 事象 A_i の数(確率変数) x_i : 試行結果が事象 A_i になる回数 $x_1 + x_2 + \cdots + x_m = n$ n : 試行回数 p_i : 試行結果が事象 A_i になる確率 $p_1 + p_2 + \cdots + p_m = 1$	n p_i	n $p_i (1 - p_i)$
6	負の二項分布	$P(X = x) = {}_{k+x-1} C_x p^k (1 - p)^x$ x : 失敗回数(確率変数) k : 成功回数 p : 成功確率	k q / p	k q / p ²
7	幾何分布	$P(X = x) = (1 - p)^{x-1} p$ x : x 回目の試行で初めて成功(確率変数) p : 成功確率	1 / p	(1 - p) / p ²
8	超幾何分布	$P(X = x) = {}_K C_x {}_{N-K} C_{n-x} / {}_N C_n$ x : 成功回数(確率変数) K : 成功状態の数 N : 母集団サイズ n : 抽出数	n K / N	n K (N - K) (N - n) / { N ² (N - 1) }
9	ポアソン分布	$P(X = x) = (\lambda^x / x!) e^{-\lambda}$ x : 発生回数(確率変数) λ : 単位時間当たりの平均発生回数	λ	λ

4-2. 確率分布(連続型)の一覧表(1/2)

(表2-2) 確率分布(連続型)の一覧表(1/2)

	確率分布の名前	確率密度関数(X : 確率変数)	期待値	分散
10	一様分布	$a \leq x \leq b$ のとき、 $f(x) = 1/(b - a)$ $x < a$ もしくは、 $b < x$ のとき、 $f(x) = 0$ x : 確率変数 a, b : 閉区間 [a, b]	$(a + b) / 2$	$(a - b)^2/12$
11	指数分布	$f(x) = \lambda e^{-\lambda x}$ x : 発生間隔、待ち時間(確率変数) λ : 単位時間当たりの平均発生回数	$1 / \lambda$	$1 / \lambda^2$
12	ワイブル分布	$f(x) = (b/a) (x/a)^{b-1} e^{-(x/a)^b}$ x : 経過時間(確率変数) a : 尺度パラメータ b : 形状パラメータ	$a \Gamma((b + 1)/b)$	$a^2 \{ \Gamma((b + 2)/b) - \Gamma^2((b + 1)/b) \}$
13	ガンマ分布	$f(x) = \lambda^k x^{k-1} e^{-\lambda x} / \Gamma(k)$ x : 待ち時間(確率変数) λ : 単位時間当たりの平均発生回数 k : 事象の発生回数	k / λ	k / λ^2
14	ベータ分布	$f(x) = x^{\alpha-1} (1 - x)^{\beta-1} / B(\alpha, \beta)$ x : 成功率(確率変数) α : 成功数 β : 失敗数	$\alpha / (\alpha + \beta)$	$\alpha \beta / \{ (\alpha + \beta)^2 (\alpha + \beta + 1) \}$
15	ディリクレ分布	$f(x) = \{ \Gamma(\sum_{i=1}^n \alpha_i) / \prod_{i=1}^n \Gamma(\alpha_i) \} \prod_{i=1}^n x_i^{\alpha_i-1}$ x : 全体の事象の発生確率(確率変数) x_i : 事象 i の発生確率 $\alpha_i - 1$: 事象 i の発生回数	$\alpha_i / \sum_{i=1}^n \alpha_i$	$\alpha_i (\sum_{i=1}^n \alpha_i - \alpha_i) / \{ (\sum_{i=1}^n \alpha_i)^2 (\sum_{i=1}^n \alpha_i + 1) \}$
16	正規分布	$f(x) = \{ 1 / (2 \pi \sigma^2)^{1/2} \} e^{-(x - \mu)^2/(2 \sigma^2)}$ x : 確率変数 μ : 期待値(平均) σ^2 : 分散	μ	σ^2
17	標準正規分布	$f(x) = (1/(2 \pi)^{1/2}) e^{-x^2/2}$ x : 確率変数	0	1

4-2. 確率分布(連続型)の一覧表(2/2)

(表2-2)確率分布(連続型)の一覧表(2/2)

	確率分布の名前	確率密度関数(X : 確率変数)	期待値	分散
18	多変量正規分布	$f(\mathbf{x}) = \{ 1 / ((2 \pi)^{n/2} \Sigma ^{1/2}) \} e^{-(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)/2}$ $\mathbf{x}$: 確率変数ベクトル μ : 期待値ベクトル(平均ベクトル) Σ : 分散共分散行列	μ	Σ
19	対数正規分布	$f(x) = e^{-(\ln x - \mu)^2/(2 \sigma^2)} / ((2 \pi \sigma^2)^{1/2} x)$ x : 確率変数 μ, σ : 定数	$e^{\mu + \sigma^2/2}$	$e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$
20	カイ二乗分布	$f(x) = (x^{n/2-1} e^{-x/2}) / (\Gamma(n/2) 2^{n/2})$ x : 確率変数 n : 自由度	n	$2 n$
21	F 分布	$f(x) = \{ 1/B(m/2, n/2) \} (m/n)^{m/2} x^{(m-2)/2} / \{ 1 + (m/n) x \}^{(m+n)/2}$ x : 確率変数 m, n : 自由度	$n > 2$ のとき、 $n / (n - 2)$	$n > 4$ のとき、 $\{ 2 n^2 (m + n - 2) \} / \{ m (n - 2)^2 (n - 4) \}$
22	t 分布	$f(x) = \{ \Gamma((n+1)/2) / ((n \pi)^{1/2} \Gamma(n/2)) \} (1 + x^2/n)^{-(n+1)/2}$ $f(x) = \{ 1 / (n^{1/2} B(n/2, 1/2)) \} (1 + x^2/n)^{-(n+1)/2}$ x : 確率変数 n : 自由度	0	$n / (n - 2)$
23	コーシー分布	$f(x) = (\gamma / \pi) / ((x - x_0)^2 + \gamma^2)$ x : 確率変数 x_0 : 位置母数 γ : 尺度母数	存在しない	存在しない
24	標準コーシー分布	$f(x) = (1 / \pi) / (x^2 + 1)$ x : 確率変数	存在しない	存在しない

5. 一様分布(Uniform distribution)

■概要

一様分布は、すべての事象が起こる確率が等しい現象を表すときの確率分布である。

言い換えれば、確率質量関数、もしくは確率密度関数が、ある区間内で常に一定である確率分布を一様分布という。

■確率分布

離散型と連続型の平均と分散の導出は、いずれも(付録25)に示す。

・離散型(確率質量関数)

確率変数 X が取りうる事象の数を N とすると、

確率変数 X が $X = x$ (但し、 $x = n, n + 1, \dots, n + N - 1$) となる確率質量関数 $P(X = x)$ は(式1)、(図3)で表される。

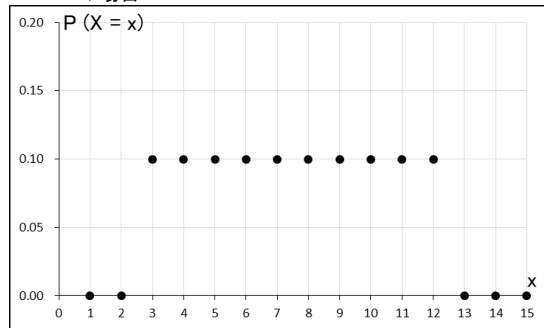
$$P(X = x) = 1/N \quad \dots \quad (式1)$$

・連続型(確率密度関数)

確率変数 X が連続区間 $a \leq X \leq b$ (a, b は実数)で一様分布に従うとき、確率密度関数 $f(x)$ は(式2)、(図4)で表される。

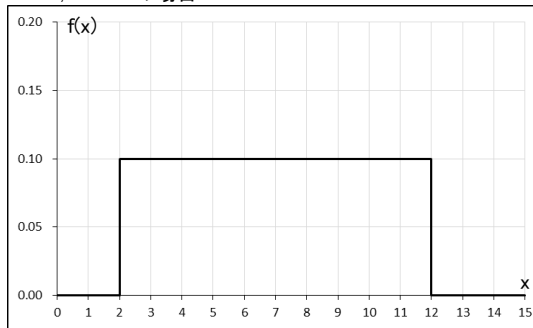
$$a \leq x \leq b \text{ のとき、} f(x) = 1/(b - a), \quad x < a \text{ もしくは } b < x \text{ のとき、} f(x) = 0 \quad \dots \quad (式2)$$

$N = 10$ の場合



(図3)離散型一様分布

$a = 2, b = 12$ の場合



(図4)連続型一様分布

6. ベルヌーイ分布(Bernoulli distribution)

■概要

「成功か失敗か」、「勝ちか負けか」など、何かを行ったときに起こる結果が2つしかない試行を、ベルヌーイ試行という。
このベルヌーイ試行の結果を、1 と 0 で表し、結果が 1 になる確率を p , 結果が 0 になる確率を $1 - p$ としたときの離散型確率分布をベルヌーイ分布という。

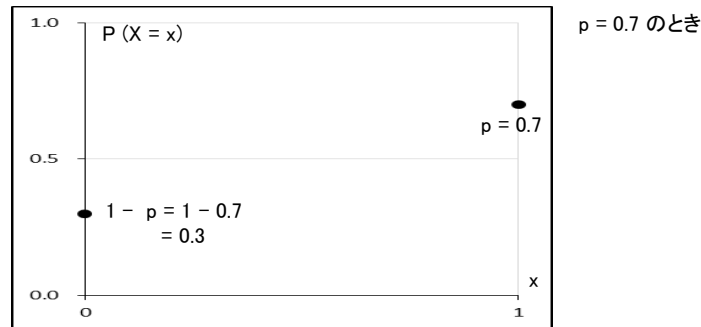
■確率質量関数

期待値と分散の導出は(付録26)に示す。

$$P(X = x) = p^x (1 - p)^{1-x} \quad \dots \text{ (式3)}$$

x : 事象 (成功時 $x = 1$, 失敗時 $x = 0$) , p : 成功確率 ($0 \leq p \leq 1$)

ベルヌーイ分布を(図5)に示す。



(図5)ベルヌーイ分布

7. カテゴリカル分布(Categorical distribution)

■概要

1 回の試行で複数の事象がそれぞれの確率で得られるとき、これらの事象全体が従う確率分布をカテゴリカル分布という。

1 回の試行で複数ある事象(ここでは、カテゴリと称す)からある確率で 1 つカテゴリを選ぶということに因んで、カテゴリカル分布という名前が付いている。

つまり、1 回の試行で k 種類のカテゴリ $X_1, X_2, \dots, X_k$ のなかから、それぞれ確率 $p_1, p_2, \dots, p_k$ で 1 つのカテゴリが選ばれるとき、確率変数ベクトル $\mathbf{X} = (X_1, X_2, \dots, X_k)$ が従う確率分布がカテゴリカル分布になる。

なお、カテゴリの種類が 2 つ、それぞれの確率が $p_1 = p, p_2 = 1 - p_1$ のとき、カテゴリカル分布はベルヌーイ分布になる。つまり、ベルヌーイ分布を多次元化した分布がカテゴリカル分布である。

■確率質量関数

期待値と分散の導出は(付録27)に示す。

$$P(\mathbf{X} = \mathbf{x}) = \prod_{i=1}^k p_i^{x_i} \quad \dots \quad (\text{式4})$$

$\mathbf{x}$: カテゴリベクトル(確率変数ベクトル $\mathbf{x} = (x_1, x_2, \dots, x_k)$ 1 つの要素 x_i が 1 で、他は 0)

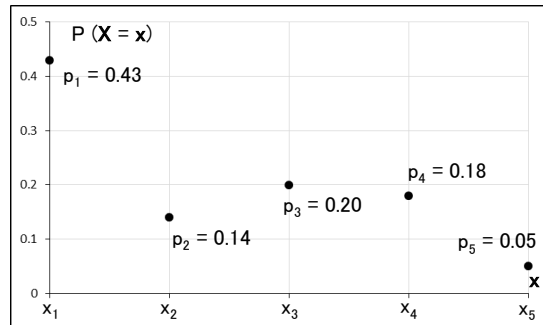
k : カテゴリの種類の数, p_i : カテゴリの選択確率 ($\sum_{i=1}^k p_i = 1, 0 \leq p_i \leq 1$)

(計算例)

$k = 5, p_3 = 0.2$ のときの $P(\mathbf{X} = \mathbf{x}_3)$ を求める。ここで、 $\mathbf{x}_3 = 1$, そして $\mathbf{x}_3$ 以外の $x_1, x_2, x_4, x_5 = 0$ である。

よって、 $P(\mathbf{X} = \mathbf{x}_3) = p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3} \cdot p_4^{x_4} \cdot p_5^{x_5} = p_1^0 \cdot p_2^0 \cdot 0.2^1 \cdot p_4^0 \cdot p_5^0 = 1 \cdot 1 \cdot 0.2 \cdot 1 \cdot 1 = 0.2$

カテゴリカル分布を(図6)に示す。



$p_1 = 0.43$
 $p_2 = 0.14$
 $p_3 = 0.20$
 $p_4 = 0.18$
よって
 $p_5 = 1 - (p_1 + p_2 + p_3 + p_4)$
 $= 1 - (0.43 + 0.14 + 0.20 + 0.18) = 0.05$
のとき

(図6)カテゴリカル分布

8. 二項分布(Binomial distribution)

■概要

母集団から標本として抜き取ったデータを母集団に戻すサンプリング方法(復元抽出)でベルヌーイ試行を繰り返し行い、試行回数を n 、試行の「成功」確率はすべて同じで p ($0 \leq p \leq 1$) として、成功回数(確率変数)の確率が従う確率分布が二項分布である。確率分布の式が二項定理(付録1)を使った式で表されるので、二項何分布という名前がついている。

■確率質量関数

二項分布の確率質量関数は次の式で表される。平均と分散の導出は(付録28)に示す。
 n 回の試行で、0 回、1 回、2 回、 $\dots$ 、 n 回となる成功回数 x が確率分布を構成する。

$$P(X = x) = {}_n C_x p^x (1 - p)^{n-x} \quad \dots \text{(式5)}$$

x : 成功回数(確率変数), n : 試行回数($n = 0, 1, \dots$), p : 成功確率($0 \leq p \leq 1$), 二項係数 ${}_n C_x = n! / \{x!(n-x)!\}$

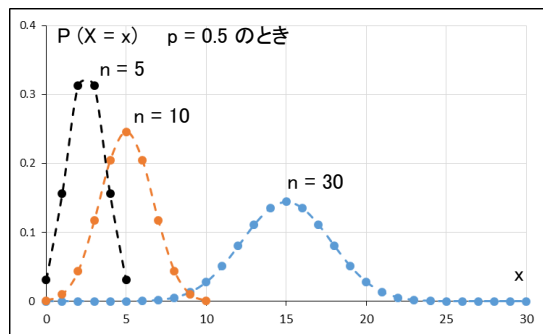
この式は、 n 回の試行のうち、確率 p で x 回成功して、確率 $1 - p$ で $n - x$ 回失敗するケースが ${}_n C_x$ 個になることをもとにして、成功確率を算出している。

(計算例)

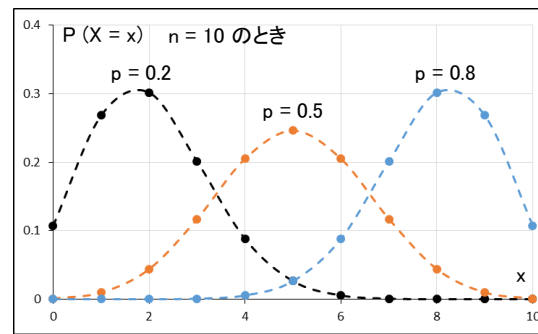
$n = 10$, $p = 0.8$ のとき、成功回数 $x = 7$ の確率は、(式5)を使って、

$$P(x = 7) = {}_{10} C_7 0.8^7 (1 - 0.8)^{10-7} = 10! / \{(10-7)! \cdot 7!\} \cdot 0.8^7 \cdot 0.2^3 = (10 \cdot 9 \cdot 8) / (3 \cdot 2 \cdot 1) \cdot 0.210 \cdot 0.008 = 0.201$$

二項分布を(図7)と(図8)に示す。



(図7)二項分布 (n を変化)



(図8)二項分布 (p を変化)

9. 多項分布(Multinomial distribution)

■概要

結果が2つだけの試行の確率分布である二項分布に対して、**多項分布は結果が複数ある試行の確率分布である**。従って、多項分布は二項分布を一般化した確率分布であると言える。

確率分布の式が多項定理(付録2)を使った式で表されるので、多項分布という名前がついている。

■確率質量関数

多項分布の確率質量関数は次の式で表される。期待値と分散の導出は(付録29)に示す。

1回の試行で m 通りの事象 $A_1, A_1, \dots, A_m$ が得られるとき、それぞれの事象の発生確率が $p_1, p_2, \dots, p_m$ であるとする。このような試行を n 回行って事象 A_i が x_i 回であるとき、この回数 x_i の従う確率分布が多項分布となる。

$$P(X_i = x_i) = \{ n! / (x_1! x_2! \dots x_m!) \} p_1^{x_1} p_2^{x_2} \dots p_m^{x_m} \quad \dots \text{ (式6)}$$

m : 試行結果である事象の数

n : 試行回数, x_i : 試行結果が事象 A_i になる回数(確率変数) $x_1 + x_2 + \dots + x_m = n$

p_i : 試行結果が事象 A_i になる確率 $p_1 + p_2 + \dots + p_m = 1, 0 \leq p_1, p_2, \dots, p_m \leq 1$

多項係数(同じ事象の数) : $n! / (x_1! x_2! \dots x_m!)$

この式は、発生確率 p_1 の事象 A_1 が x_1 回、発生確率 p_2 の事象 A_2 が x_2 回、 $\dots$ 、発生確率 p_m の事象 A_m が x_m 回、同時に発生する確率を表している。

(付録2)に示すように、まず試行回数 n 回のうち x_1 回 事象 A_1 が発生し、

次いで残りの試行回数 $n - x_1$ 回のうち x_2 回 事象 A_2 が発生し、 $\dots$

この試行を $x_1 + x_2 + \dots = n$ になるまで繰り返したものが組み合わせの数(多項係数)になり、この多項係数を使用して多項分布の確率質量関数が表される。

(計算例)

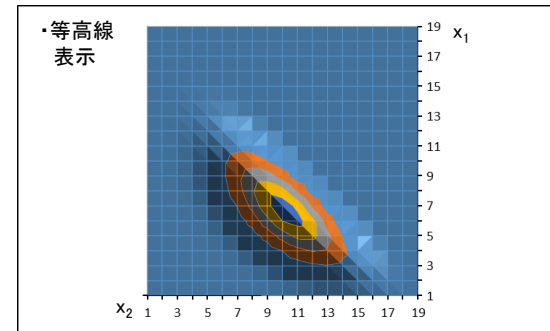
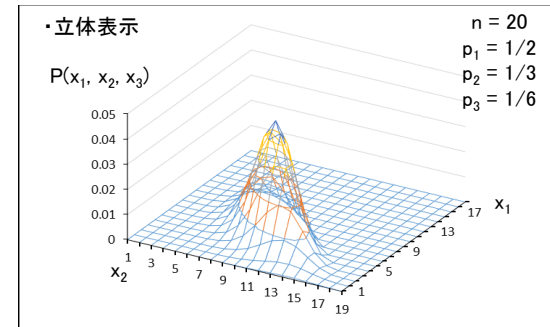
$n = 10$ で、 $p_1 = 1/2, p_2 = 1/3$ ($p_3 = 1 - 1/2 - 1/3 = 1/6$ となる) のとき、

$x_1 = 3, x_2 = 6$ ($x_3 = 10 - 3 - 6 = 1$ となる) である確率は、(式6)を使って、

$$P(x_1 = 3, x_2 = 6, x_3 = 1) = \{ 10! / (3! \cdot 6! \cdot 1!) \} \cdot (1/2)^3 \cdot (1/3)^6 \cdot (1/6)^1 = 0.024$$

試行回数 $n = 20$ で、確率変数の数が3つの場合の多項分布を(図9)に示す。

$x_1 + x_2 + x_3 = n$ の関係があるので、2つの確率変数を決めれば 残り1つの確率変数の値は自動的に決まる。よって、この場合の分布は $x_1 - x_2$ 平面上の分布として表すことができる。グラフ作成にはExcelの等高線機能を使用した。



(図9) 多項分布

10. 負の二項分布(Negative Binominal distribution)

■概要

ベルヌーイ試行を繰り返し行い、 k 回目の成功が得られるまでの失敗回数が従う確率分布の式である。なお、 $k = 1$ のとき、後述する幾何分布と等しくなる。モーメント母関数(付録21)を算出するとき負の値まで拡張した二項係数を使用するので、負の二項何分布という名前がついている。

■確率質量関数

負の二項分布の確率質量関数は次の式で表される。期待値と分散の導出は(付録30)に示す。

$$P(X = x) = {}_{k+x-1}C_x p^k (1-p)^x \quad \dots \text{(式7)}$$

x : 失敗回数(確率変数), k : 成功回数($k = 1, 2, \dots$), p : 成功確率($0 \leq p \leq 1$)

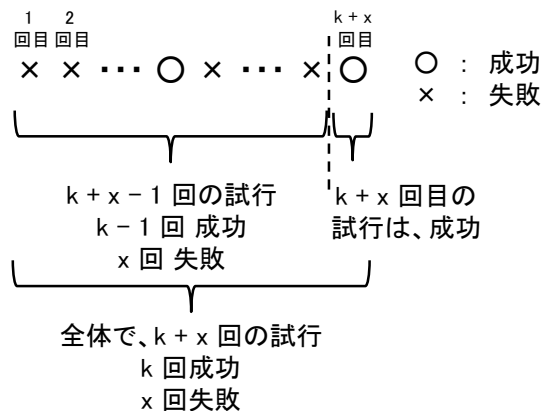
この式は、 $k + x$ 回の試行で、 $k + x$ 回目に k 回目の成功を収めること、その前の $k + x - 1$ 回の試行で x 回の失敗を重ねていること(その失敗パターンは、 ${}_{k+x-1}C_x$ 種類)、また、最終的に成功確率は p^k 、失敗確率は $(1-p)^x$ になることをもとに導出される(図10)。

(計算例)

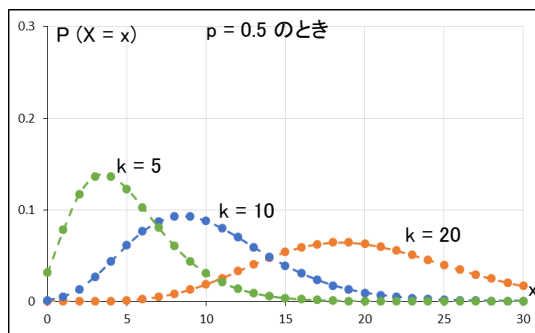
$k = 5$, $p = 0.7$ のとき、失敗回数 $x = 3$ の確率は、(式7)を使って、

$$P(X = 3) = {}_{5+3-1}C_3 0.7^5 (1-0.7)^3 = {}_7C_3 \cdot 0.7^5 \cdot 0.3^3 = 0.159 \quad (\text{なお、Excel の関数を使って、} {}_7C_3 = \text{combin}(7,3) = 35)$$

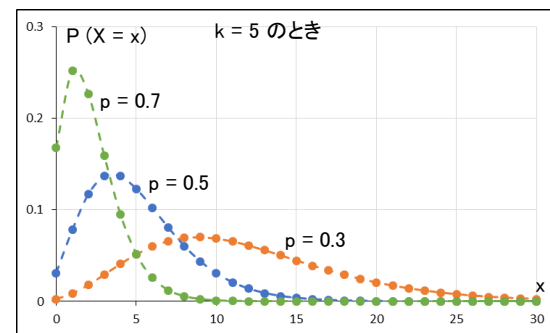
負の二項分布を(図11)と(図12)に示す。



(図10) 負の二項分布のプロセス



(図11) 負の二項分布 (k を変化)



(図12) 負の二項分布 (p を変化)

11. 幾何分布(Geometric distribution)

■概要

何かを行ったときに起こる結果が2つしかない試行を、ベルヌーイ試行という。

この2つの試行結果を「成功」と「失敗」としたとき、「成功」確率を p ($0 \leq p \leq 1$) として、初めて成功するまでの試行回数(確率変数)の確率が従う確率分布が、幾何分布である。言い換えれば、1回の成功を得るためには、平均 $1/p$ の時間を要するということを表している。例えば、 $p = 1$ であれば $1/p = 1$ の時間で成功し、 $p = 0.5$ であれば成功までに2倍の $1/p = 2$ の時間を必要とし、 $p = 0$ であれば $1/p = \infty$, つまり、無限大の時間が必要となって成功は永遠に訪れない、ということになる。なお、後述する指数分布は離散型の幾何分布を連続型にしたものになっている。

確率分布の式が等比数列(幾何数列)のようになっているので、幾何分布という名前がついている。

■確率質量関数

k 回目の試行で初めて成功する確率は、 $k - 1$ 回の失敗の後、 k 回目で成功することなので、幾何分布の確率質量関数は次の式で表される。期待値と分散の導出は(付録31)に示す。

$$P(X = x) = (1 - p)^{x-1} p \quad \dots \quad (\text{式8})$$

x : x 回目(確率変数)の試行で初めて成功, p : 成功確率 ($0 \leq p \leq 1$)

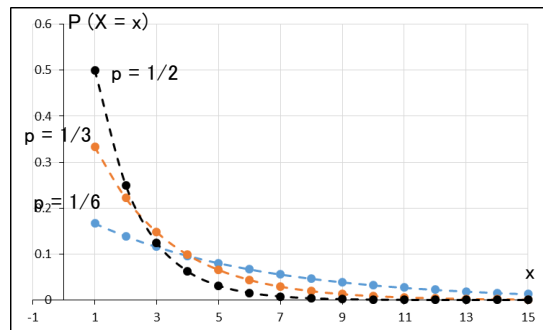
(計算例)

サイコロの目のでる確率はそれぞれ $1/6$ であるが、

サイコロを1の目がでるまで振り続けるとき、5回目で1の目が初めてでる確率は、(式8)を使って、

$$P(X = 5) = (1 - 1/6)^{5-1} (1/6) = (5/6)^4 (1/6) = 0.482 \cdot 0.167 = 0.08$$

幾何分布を(図13)に示す。



(図13)幾何分布

12. 超幾何分布(Hipergeometric distribution)

■概要

超幾何分布は、成功状態を持つ母集団からデータを非復元抽出でサンプリングしたときの成功状態の数が従う確率分布である。二項分布が「復元抽出」のベルヌーイ試行を対象とするのに対し、超幾何分布は「非復元抽出」のベルヌーイ試行を対象としている。超幾何関数という確率分布の名前は幾何分布との関連を想起させるが、直接の関係はない。名前の由来は、確率質量関数の式が超幾何関数と呼ばれる形をしていることに由来する。

■確率質量関数

超幾何分布の確率質量関数は(式9)、もしくは(式10)で表される。期待値と分散の導出は(付録32)に示す。

K 個の成功状態を持つ N 個の要素からなる母集団から、n 個の要素を非復元抽出したときに x 個の成功状態が含まれている確率を与える。

$$P(X = x) = \frac{{}_K C_x \cdot {}_{N-K} C_{n-x}}{{}_N C_n} \quad \dots \text{ (式9)}$$

$$P(X = x) = \frac{{}_n C_x \cdot {}_{N-n} C_{K-x}}{{}_N C_K} \quad \dots \text{ (式10)}$$

x : 成功回数(確率変数), N : 母集団のサイズ, K : 成功状態の数, n : 抽出数

(式9)の説明: 母集団の N 個の要素から n 個の要素を抽出する方法の組み合わせの数は、 ${}_N C_n$ である。このとき、抽出した n 個は、母集団のなかの成功状態 K 個からの x 個(この組み合わせの数は、 ${}_K C_x$)と母集団のなかの失敗状態 $N - K$ 個からの $n - x$ 個(この組み合わせの数は、 ${}_{N-K} C_{n-x}$)で構成される。

(式10)の説明: 母集団の N 個のなかの成功状態 K 個の組み合わせの数は ${}_N C_K$ である。この成功状態 K 個は、抽出した n 個のなかの成功状態 x 個(この組み合わせの数は、 ${}_n C_x$)と抽出しなかった $N - n$ 個のなかの残りの成功状態 $K - x$ 個(この組み合わせの数は、 ${}_{N-n} C_{K-x}$)で構成される。

母集団のサイズである N が大きくなると、非復元抽出は復元抽出と大差なくなる。つまり、母集団サイズが大きくなると、超幾何分布は二項分布に近づく。

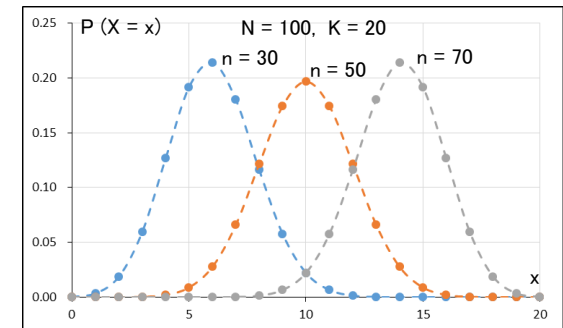
(計算例)

N = 100, K = 20, n = 30 のとき、成功回数 x = 7 の確率は、(式9)を使って、

$$\begin{aligned} P(X = 7) &= \frac{{}_{20} C_7 \cdot {}_{100-20} C_{30-7}}{100 C_{30}} = \frac{{}_{20} C_7 \cdot {}_{80} C_{23}}{100 C_{30}} \\ &= (7.75 \times 10^4) (6.83 \times 10^{19}) (2.94 \times 10^{25}) = 0.18 \end{aligned}$$

なお、組み合わせの計算には、Excel の combin 関数を利用した。

超幾何分布を(図14)に示す。



(図14)超幾何分布 (n を変化)

13. ポアソン分布 (Poisson distribution)

■概要

ある期間で平均してある回数起こる事象の発生回数(確率変数)が従う確率分布がポワソン分布である。この分布は、ポアソン過程(ランダムに発生する事象を確率変数を使って表したもの)の事象の発生回数に着目した確率分布である。頻繁には発生しない事象についての確率分布である。例えば、交通事故の発生件数、不良品の発生件数、一定時間内の来店客数などの確率分布が挙げられる。

■確率質量関数

ポアソン分布の確率質量関数は次の式で表される。期待値と分散の導出は(付録33)に示す。

$$P(X=x) = (\lambda^x / x!) e^{-\lambda} \quad \dots \text{(式11)}$$

x : 発生回数, λ : 単位時間(例えば、1 分間)あたりの平均発生回数

(計算例)

あるレストランでは、ランチタイムに平均して 20 分間に 4 組の来客がある。1 時間に 15 ~ 20 組の来客がある確率を計算する。

(注) 来客を組数で扱うことでポアソン過程を想定できる。人数で扱うと同じ組内の来客は一緒に来客となりポアソン過程には従わなくなる。

平均して 20 分間に 4 組の来客なので、平均して 1 時間に $\lambda = 4 \times (60/20) = 12$ 組の来客があることになる。

よって、1 時間に 15 組の来客がある確率は、 $P(X=15) = (12^{15}/15!) e^{-12} = 0.072$ となる。同様に、 $P(X=16) = (12^{16}/16!) e^{-12} = 0.054$,

$P(X=17) = 0.038$, $P(X=18) = 0.026$, $P(X=19) = 0.016$, $P(X=20) = 0.010$ (いずれの確率も小さい)

よって、1 時間に 15 ~ 20 組の来客がある確率は、

$P(15 \leq X \leq 20) = 0.072 + 0.054 + 0.038 + 0.016 + 0.010 = 0.216$ (21.6 %)

ポアソン分布を(図15)に示す。

(導出)

ある期間での事象の平均発生回数が λ である。この期間を n 等分し微小期間に分ける。

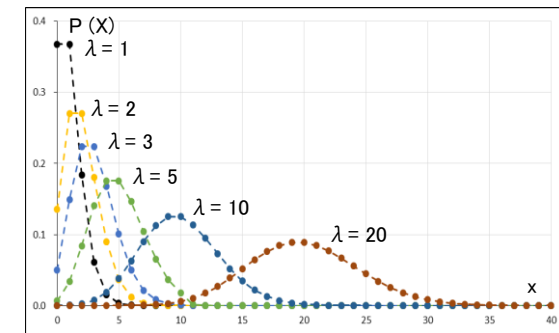
微小期間で事象が発生する機会が最大 1 ならば、ある期間内での事象の発生確率は λ/n になる。この事象は微小期間内で、発生するか、しないか、2者択一のベルヌーイ試行になるので、この確率分布は二項分布をもとに、 $n \rightarrow \infty$ (λ は定数、その結果、 $p \rightarrow 0$) としたものである。これがポアソン分布になる。

$$\begin{aligned} P(X=x) &= \lim_{n \rightarrow \infty} {}_n C_x (\lambda/n)^x (1 - \lambda/n)^{n-x} \\ &= \lim_{n \rightarrow \infty} \frac{n!}{x!(n-x)!} (\lambda/n)^x (1 - \lambda/n)^{n-x} \\ &= \lambda^x / x! \lim_{n \rightarrow \infty} \frac{n!}{(n-x)! n^x} (1 - \lambda/n)^{n-x} \\ &= \lambda^x / x! \lim_{n \rightarrow \infty} \left\{ \frac{n(n-1)(n-2) \cdots (n-x+1)}{n^x} \right\} (1 - \lambda/n)^n (1 - \lambda/n)^{-x} \\ &= \lambda^x / x! \lim_{n \rightarrow \infty} \{ 1(1-1/n)(1-2/n) \cdots (1-(x-1)/n) \} (1 - \lambda/n)^n (1 - \lambda/n)^{-x} \end{aligned}$$

ここで、(付録4)ネイピア数の、 $\lim_{n \rightarrow \infty} (1 - \lambda/n)^n = e^{-\lambda}$ の関係を使う。

また、 $n \rightarrow \infty$ で、 $1/n \rightarrow 0$, $2/n \rightarrow 0$, $\dots$, $(x-1)/n \rightarrow 0$, $\lambda/n \rightarrow 0$ であるので、

$$P(X=x) = (\lambda^x / x!) e^{-\lambda}$$



(図15)ポアソン分布

14. 指数分布(Exponential distribution)

■概要

ある期間で平均してある回数起こる事象の発生間隔(確率変数)が従う確率分布が指数分布である。ある事象が発生した後に、次の事象が発生するまでの時間間隔の確率分布である。指数分布はポアソン過程(ランダムに発生する事象を確率変数を使って表したもの)の事象の発生間隔に着目した確率分布である。一方、ポアソン分布は、ポアソン過程の発生回数に着目した確率分布である。また、指数分布では事象の生起確率が一定である(変化しない)としているが、後述するワイブル分布では、事象の生起確率が変化する場合は確率分布である。

確率分布の式に指数関数が使われているので、指数分布という名前がついている。

■確率密度関数

指数分布の確率密度関数は次の式で表される。期待値と分散の導出は(付録34)に示す。

$$f(x) = \lambda e^{-\lambda x} \quad \dots \text{(式12)} \quad x: \text{発生間隔(待ち時間)} \quad \lambda: \text{単位時間あたりの平均発生回数(期間 } 1/\lambda \text{ あたりに平均 1 回発生)}$$

$$\text{累積分布関数 } F(x) = \int_0^x f(x) dx = \lambda \int_0^x e^{-\lambda x} dx = \lambda \left[(-1/\lambda) e^{-\lambda x} \right]_0^x = 1 - e^{-\lambda x} \quad \dots \text{(式13)}$$

(計算例)

1 時間に平均 5 回起きる現象がある。その現象が起きてから、15 分以内にその現象が再び起きる確率を求める。 $\lambda = 5$, 発生間隔は $x = 15 \text{ 分} = 1/4 \text{ 時間}$ であるので、この確率は(式13)の累積分布関数を使って求めることができる。

$$F(1/4) = P(X \leq 1/4) = 1 - e^{-5 \cdot 1/4} = 1 - e^{-1.25} = 0.713$$

指数分布を(図16)に示す。

(導出)

① 現象が微小期間 Δx で発生する確率は $f(x) \Delta x$ (確率密度関数の定義)で表すことができる。

② 一方、この確率は(図17)に示すように、事象が x 期間で発生しないで、かつ、 $x + \Delta x$ の微小期間 Δx で新たに発生する場合の確率に等しい。

②-1 x 期間で発生しない確率は $1 - F(x)$ である。

②-2 一方、(図18)に示すように、ある期間を(ν)を n 等分した微小期間 $\Delta x = \nu/n$ を考える。

事象がこの微小期間で発生する機会が最大 1 回 であるとすれば、ある期間 ν で事象が m 回発生すればその発生確率は m/n である。

また、平均発生回数(単位期間あたりの発生回数) λ は、 $\lambda = m/\nu$ である。

よって、 $\lambda = m/\nu = m/(n \Delta x)$ より、微小期間 Δx での発生確率は、 $m/n = \lambda \Delta x$ となる。

よって①と②-1、②-2 から、 $f(x) \Delta x = (1 - F(x)) \lambda \Delta x$

両辺を Δx で割って、 $f(x) = (1 - F(x)) \lambda$ が求まる。

また、 $f(x) = d/dt F(x)$ であるので、 $d/dt F(x) + \lambda F(x) = \lambda \quad \dots \text{③}$ と表すことができる。

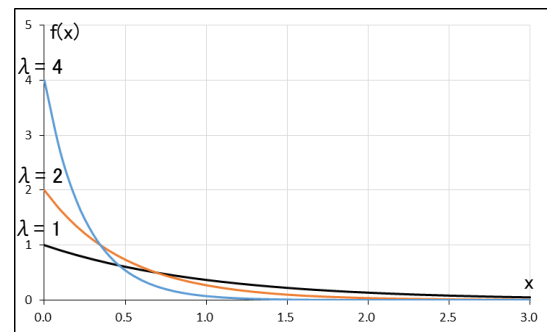
この微分方程式は、1階線形微分方程式であり、(付録6)の手順に従って解くと、

$F(x) = 1 + C e^{-\lambda x}$ (C は、不定積分定数) が求まる。

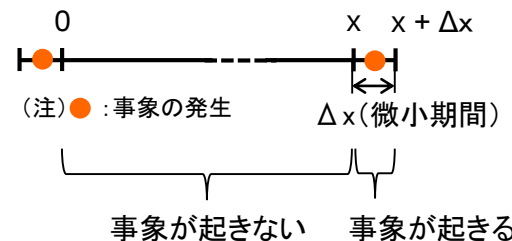
$F(x)$ は累積分布関数で $F(0) = 0$ なので、 $F(0) = 0 = 1 + C e^{-\lambda \cdot 0}$ より、 $C = -1$ である。

従って、微分方程式③の解は、 $F(x) = 1 - e^{-\lambda x}$ となる。

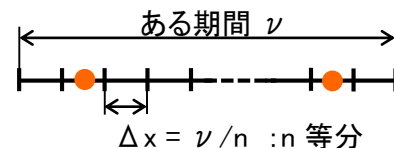
そして、この両辺を微分して、確率密度関数 $f(x) = \lambda e^{-\lambda x}$ が求まる。



(図16) 指数分布



(図17) 発生間隔の確率の考え方



(図18) 微小期間での発生確率

15. ワイブル分布(Weibull distribution)

■概要

ワイブル分布は事象の発生確率が変化する場合の事象の発生間隔の確率分布である。例えば、機械の故障は初期故障段階、偶発故障段階、摩耗故障段階を経ると言われるが、ワイブル分布は、このように故障の発生確率が経時変化する現象を表すことができる確率分布である。確率密度関数のパラメータを変えることで時間の経過に伴い故障率が小さくなる現象も、大きくなる現象も表すことができる。複数の環がつながった鎖を引っ張ったとき、最も弱い環の場所から壊れるという最弱リンクモデル使った確率分布であり、数学者のワイブル(スウェーデン)によって提案された。なお、ここでは(付録20)のハザード関数を使って確率密度関数を導出する。

■確率質量関数の導出

経過時間 X (確率変数)における故障発生率を表す確率密度関数は次の式で表される。ここではパラメータの数が2種類のワイブル分布の式を示す(他に位置パラメータがあるが、分布の形状には直接影響しない)。期待値と分散の導出は(付録35)に示す。また、ワイブル分布を(図19)に示す。

$$f(x) = (b/a) (x/a)^{b-1} e^{-(x/a)^b} \quad \dots \text{ (式14)}$$

x : 経過時間, a : 尺度パラメータ, b : 形状パラメータ

a について: $x = a$ 時間経過したとき、累積分布関数(式15)は、 $F(a) = 1 - e^{-(a/a)^b} = 1 - e^{-1} = 0.632$ となる。

つまり、 a 時間以内に約 63% が故障することを示している。

b について: (図19)から分かるように、 $b < 0$ で初期故障、 $b = 1$ で偶発故障、 $b > 0$ で摩耗故障に対応させることができる。

$a = 1/\lambda$, $b = 1$ (偶発故障) のとき、 $f(x) = (1/(1/\lambda)) (x/(1/\lambda))^{1-1} e^{-(x/(1/\lambda))^1} = \lambda e^{-\lambda x}$ となる。これは指数分布(式13)である。

(計算例)

$a = 3$, $b = 10$ のときの $x = 3$ の確率密度を求める。 $f(x) = (10/3) (3/3)^{10-1} e^{-(3/3)^{10}} = 1.226$

(導出)

ハザード関数を使った確率密度関数(付式55)に、ハザード関数(付式56)を代入する。

$$f(x) = h(x) e^{-\int_0^x h(x) dx} = \alpha \beta x^{\beta-1} e^{-\int_0^x \alpha \beta x^{\beta-1} dx} = \alpha \beta x^{\beta-1} e^{-[(\alpha \beta / \beta) x^\beta]_0^x} = \alpha \beta x^{\beta-1} e^{-\alpha x^\beta}$$

ここで、 $\beta = b$, $\alpha = 1/a^\beta = 1/a^b$ とおくと、ワイブル分布の確率密度関数が得られる。

$$f(x) = (1/a^b) b x^{b-1} e^{-(1/a^b) x^b} = (b/a) (x/a)^{b-1} e^{-(x/a)^b}$$

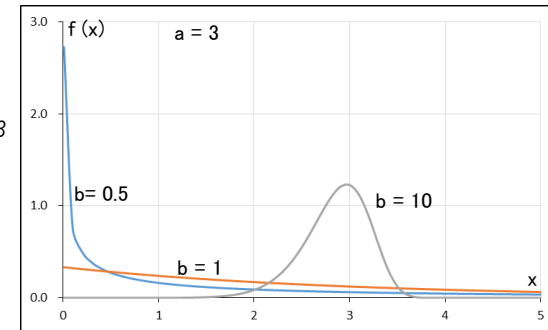
また、ハザード関数を使った累積分布関数(付式54)に、ハザード関数(付式56)を代入して、ワイブル分布の累積分布関数を求める。

$$F(x) = 1 - e^{-\int_0^x h(x) dx} = 1 - e^{-\int_0^x \alpha \beta x^{\beta-1} dx} = 1 - e^{-[(\alpha \beta / \beta) x^\beta]_0^x} = 1 - e^{-\alpha x^\beta}$$

ここで、上と同様に、 $\beta = b$, $\alpha = 1/a^\beta = 1/a^b$ とおくと、

ワイブル分布の累積分布関数が得られる。

$$F(x) = 1 - e^{-(1/a^b) x^b} = 1 - e^{-(x/a)^b} \quad \dots \text{ (式15)}$$



(図19)ワイブル分布

16. ガンマ分布(Gamma distribution)

■概要

ある期間で平均してある回数起こる事象があって、その事象が何回か発生するまでの待ち時間(確率変数)が従う確率分布がガンマ分布である。指数分布は、ある事象の発生後に次の事象が発生するまでの待ち時間の確率分布であるので、ガンマ分布は指数分布の和になる。ある事象が最初にかかるまでの時間の分布が指数分布で、複数回起きるまでの時間の分布がガンマ分布である。確率分布の式にガンマ関数(付録12)が使われている。

■確率密度関数

ガンマ分布の確率密度関数は次の式で表される。期待値と分散の導出は(付録36)に示す。

$$f(x) = \lambda^k x^{k-1} e^{-\lambda x} / \Gamma(k) \quad \cdots \text{(式16)} \quad \Gamma(k) \text{ はガンマ関数、}\Gamma \text{ は}\gamma \text{ (ガンマ)の大文字}$$

x : 待ち時間, λ : 単位期間あたりの平均発生回数(期間 $1/\lambda$ で平均 1 回), k : 事象の発生回数

(計算例)

$x = 4$ ヶ月間で、平均発生回数 $\lambda = 2$ 回のヒヤリ事故が $k = 5$ 回発生する確率密度を求める。

$$f(4) = 2^5 4^{5-1} e^{-2 \cdot 4} / \Gamma(5) = 2^5 4^4 e^{-8} / 4! = 0.115$$

ガンマ分布は、 $k = 1$ のとき、指数分布になる。また、 $k = n/2$, $\lambda = 1/2$ のとき、後述のカイ二乗分布になる。ガンマ分布を(図20)と(図21)に示す。

(導出)

確率変数 $X_1, X_2, \dots, X_k$ が互いに独立で、指数分布 $f(x_i) = \lambda e^{-\lambda x_i}$ に従うとき、新しい確率変数

$X = X_1 + X_2 + \dots + X_k$ が従う確率密度関数 $f_X(x)$ を求める。デルタ関数 $\rightarrow$ 畳み込み積分を用いて表すことができる(付録18)。

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \delta(x - (x_1 + \dots + x_k)) f(x_1) \cdots f(x_k) dx_1 \cdots dx_k \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x - (x_2 + \dots + x_k)) f(x_2) \cdots f(x_k) dx_2 \cdots dx_k \end{aligned}$$

なお、積分区間は、 $x = x_1 + x_2 + \dots + x_k$ より $x_1 = x - (x_2 + \dots + x_k) \geq 0$ となるので $x - (x_3 + \dots + x_k) \geq x_2 \geq 0$ 、以降同様に、 $x \geq x_k \geq 0$ が成立する。よって、

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \lambda e^{-\lambda(x - (x_2 + \dots + x_k))} (\lambda e^{-\lambda x_2}) \cdots (\lambda e^{-\lambda x_k}) dx_2 \cdots dx_k \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \lambda e^{-\lambda(x - (x_2 + \dots + x_k))} \lambda^{k-1} e^{-\lambda(x_2 + \dots + x_k)} dx_2 \cdots dx_k \\ &= \lambda^k e^{-\lambda x} \int_0^x dx_k \int_0^{x-x_k} dx_{k-1} \cdots \int_0^{x-(x_3 + \dots + x_k)} dx_2 \cdots \text{(式17)} \end{aligned}$$

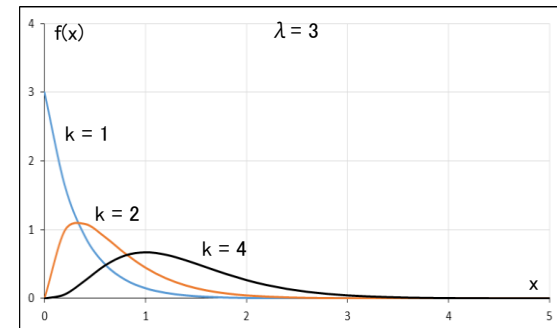
$$\text{ここで、} \int_0^{x-(x_3 + \dots + x_k)} dx_2 = \left[x_2 \right]_0^{x-(x_3 + \dots + x_k)} = x - (x_3 + \dots + x_k) - 0 = x - (x_4 + \dots + x_k) - x_3$$

$$\begin{aligned} \text{さらに、} \int_0^{x-(x_4 + \dots + x_k)} (x - (x_4 + \dots + x_k) - x_3) dx_3 &= \left[(x - (x_4 + \dots + x_k)) x_3 - (1/2) x_3^2 \right]_0^{x-(x_4 + \dots + x_k)} = \left[(x - (x_4 + \dots + x_k)) - (1/2) x_3 \right] x_3 \Big|_0^{x-(x_4 + \dots + x_k)} \\ &= \{x - (x_4 + \dots + x_k) - (1/2)(x - (x_4 + \dots + x_k))\} (x - (x_4 + \dots + x_k)) = (1/2) \{x - (x_4 + \dots + x_k)\}^2 = (1/2!) \{x - (x_5 + \dots + x_k) - x_4\}^2 \end{aligned}$$

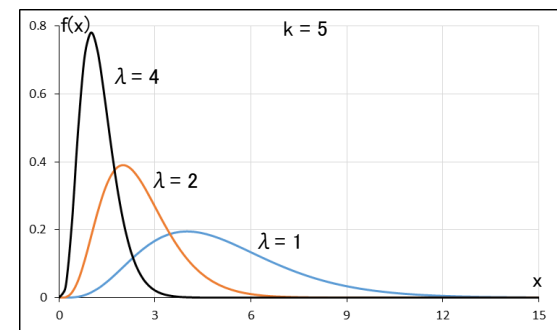
$$\text{同様にして、} (1/2!) \int_0^{x-(x_5 + \dots + x_k)} \{x - (x_5 + \dots + x_k) - x_4\}^2 dx_4 = (1/3!) \{x - (x_5 + \dots + x_k)\}^3 = (1/3!) \{x - (x_6 + \dots + x_k) - x_5\}^3$$

このように積分を繰り返すことによって、(式17)は、下式で表される。

$$\begin{aligned} f_X(x) &= \lambda^k e^{-\lambda x} \{1/(k-2)!\} \int_0^x (x - x_k)^{k-2} dx_k = \lambda^k e^{-\lambda x} \{1/(k-2)!\} \left[\{1/(k-1)\} (x - x_k)^{k-1} (-1) \right]_0^x = \lambda^k e^{-\lambda x} \{1/(k-2)!\} x^{k-1}/(k-1) = \lambda^k x^{k-1} e^{-\lambda x} / (k-1)! \\ &= \lambda^k x^{k-1} e^{-\lambda x} / \Gamma(k) \quad \text{ここで、} k=1 \text{ とおくと、当然だが、} f_X(x) = \lambda^1 x^{1-1} e^{-\lambda x} / \Gamma(1) = \lambda e^{-\lambda x} \text{ と指数分布になることを確認できる。} \end{aligned}$$



(図20)ガンマ分布



(図21)ガンマ分布

17. ベータ分布(Beta distribution)

■概要

ベータ分布は、成功数 α と失敗数 β が明らかとなときの成功率(確率変数)の確率分布である。確からしいと思われる成功率に「山」がある確率分布になる。この確からしさ(尤度:ゆうど、と言う)は試行回数が増えるに従って高くなる。 $\alpha = 1, \beta = 1$ のとき、つまり成功数に関する情報も失敗数に関する情報が最少であるときには、ベータ分布は一様分布になる。一方、成功数と失敗数の情報が増えるに伴い(α と β の数が大きくなるに伴い)、確率分布の「山」は高く、「幅」は狭くなる。確率分布の式にベータ関数(付録13)が使われているので、ベータ分布という名前がついている。

■確率密度関数

ベータ分布の確率密度関数は次の式で表される。期待値と分散の導出は(付録37)に示す。

$$f(x) = x^{\alpha-1} (1-x)^{\beta-1} / B(\alpha, \beta) \cdots (式18)$$

x : 成功率(確率変数), α : 成功回数, β : 失敗回数

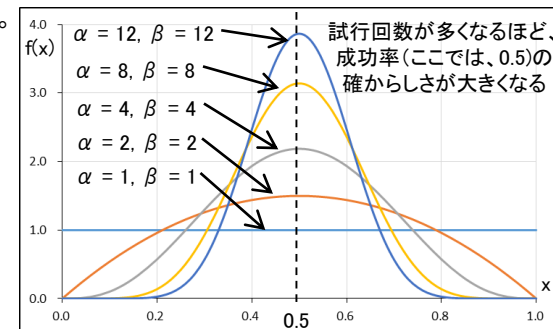
$B(\alpha, \beta)$ はベータ関数、 B は β (ベータ)の大文字

(計算例)

$\alpha = 2, \beta = 8$ のときの成功率 $x = 0.3$ の確率密度を求める(付録13. ベータ関数の公式を活用)。

$$f(0.3) = 0.3^{2-1} \cdot (1-0.3)^{8-1} / B(2, 8) = 0.3 \cdot 0.7^7 / 0.0139 = 1.78$$

但し、 $B(2, 8) = \Gamma(2) \Gamma(8) / \Gamma(10) = (1!) \cdot (7!) / (9!) = 0.0139$ (Excel の階乗関数 $7! = \text{fact}(7) = 5040$)



(図22)ベータ分布

ベータ分布を(図22)と(図23)に示す。

(導出)

ベータ分布の確率密度関数は、ガンマ分布に従う2つの確率変数の変数変換で導出される。

成功数 U の確率分布: (式16)で $\lambda = 1, k = \alpha$ において、 $g_U(u) = u^{\alpha-1} e^{-u} / \Gamma(\alpha) \cdots (式19)$

失敗数 V の確率分布: (式16)で $\lambda = 1, k = \beta$ において、 $g_V(v) = v^{\beta-1} e^{-v} / \Gamma(\beta) \cdots (式20)$

U と V の2変数を、(式21)と(式22)で示す X と Y への変数変換を行う。

(式21)は成功率(確率変数)を表し、(式22)は試行回数になる。

$$x = u/(u+v) \cdots (式21) \quad y = u+v \cdots (式22)$$

ここで、変数変換のためのヤコビアン $J(x, y)$ を求める。

(式21)と(式22)から、 $u(x, y) = xy$, $v(x, y) = y(1-x)$

$$J(x, y) = \partial(u, v) / \partial(x, y) = \begin{vmatrix} \partial u(x, y) / \partial x & \partial u(x, y) / \partial y \\ \partial v(x, y) / \partial x & \partial v(x, y) / \partial y \end{vmatrix} = \begin{vmatrix} \partial(xy) / \partial x & \partial(xy) / \partial y \\ \partial(y(1-x)) / \partial x & \partial(y(1-x)) / \partial y \end{vmatrix} = \begin{vmatrix} y & x \\ -y & 1-x \end{vmatrix} = y(1-x) - x(-y) = y$$

U と V は独立であるので、その同時確率分布は積で表すことができる。 $g_{UV}(u, v) = g_U(u) g_V(v) = \{u^{\alpha-1} e^{-u} / \Gamma(\alpha)\} \{v^{\beta-1} e^{-v} / \Gamma(\beta)\}$

ここで、ヤコビアンを使用して、 $g_{UV}(u, v)$ から $f_{XY}(x, y)$ を求める。 $f_{XY}(x, y) = g_{UV}(x, y, y(1-x)) \cdot y = \{(xy)^{\alpha-1} e^{-(xy)} / \Gamma(\alpha)\} \{(y(1-x))^{\beta-1} e^{-y(1-x)} / \Gamma(\beta)\} y$

次いで、 $f_{XY}(x, y)$ を変数 y で積分して変数 y を消し、 $f_X(x)$ を求める(周辺分布の確率密度関数を求める、周辺化するという)。

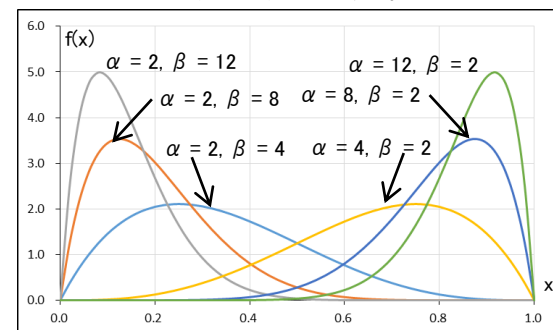
$$f_X(x) = \int_0^\infty f_{XY}(x, y) dy = \int_0^\infty \{(xy)^{\alpha-1} e^{-(xy)} / \Gamma(\alpha)\} \{(y(1-x))^{\beta-1} e^{-y(1-x)} / \Gamma(\beta)\} y dy = \int_0^\infty \{1 / (\Gamma(\alpha) \Gamma(\beta))\} x^{\alpha-1} y^{\alpha+\beta-1} (1-x)^{\beta-1} e^{-y} dy$$

$$= \left\{ \int_0^\infty \{1 / (\Gamma(\alpha + \beta))\} y^{\alpha+\beta-1} e^{-y} dy \right\} \cdot \{\Gamma(\alpha + \beta) / (\Gamma(\alpha) \Gamma(\beta))\} x^{\alpha-1} (1-x)^{\beta-1}$$

ここで、 $\{1 / \Gamma(\alpha + \beta)\} y^{\alpha+\beta-1} e^{-y}$ はガンマ分布の確率密度関数で、 $\int_0^\infty 1 / \Gamma(\alpha + \beta) y^{\alpha+\beta-1} e^{-y} dy = 1$ である。

よって、(式18)成功率の確率分布(ベータ分布)が以下の通り求まる。式変形では、付録13. ベータ関数の公式を用いる。

$$f_X(x) = \{\Gamma(\alpha + \beta) / (\Gamma(\alpha) \Gamma(\beta))\} x^{\alpha-1} (1-x)^{\beta-1} = x^{\alpha-1} (1-x)^{\beta-1} / B(\alpha, \beta)$$



(図23)ベータ分布

18. ディリクレ分布(Dirichlet distribution)

■概要

ベータ分布は2つの事象を対象にした確率分布であるが、**ディリクレ分布は対称事象を複数化した確率分布、つまり、ベータ分布を拡張した確率分布である。**互いに独立な事象の発生回数をもとに推測した全体の事象の発生確率を表す確率分布である。この分布は、ドイツの数学者ディリクレによって提唱された。ディシユレー分布とも、あるいは多変量ベータ分布とも呼ばれている。

■確率密度関数

ディリクレ分布の確率密度関数は次の式で表される。期待値と分散の導出は(付録38)に示す。

$$f(\mathbf{x}) = \left\{ \Gamma\left(\sum_{i=1}^n \alpha_i\right) / \prod_{i=1}^n \Gamma(\alpha_i) \right\} \cdot \prod_{i=1}^n x_i^{\alpha_i-1} \quad \text{ただし、} x_1 + x_2 + \dots + x_n = 1 \quad \dots \text{(式23)}$$

x_i : 事象 i の発生確率, $\alpha_i - 1$: 事象 i の発生回数, ベクトル $\mathbf{x} = (x_1, x_2, \dots, x_n)$: 事象の塊(かたまり)

(計算例)

$\alpha_1 - 1 = 7, \alpha_2 - 1 = 6, \alpha_3 - 1 = 5, x_1 = 0.55, x_2 = 0.15, x_3 = 0.3$ のときの確率密度 $f(x_1, x_2, x_3)$ を求める。

$$\begin{aligned} f(x_1, x_2, x_3) &= \left\{ \Gamma(\alpha_1 + \alpha_2 + \alpha_3) / (\Gamma(\alpha_1) \cdot \Gamma(\alpha_2) \cdot \Gamma(\alpha_3)) \right\} \cdot x_1^{\alpha_1-1} \cdot x_2^{\alpha_2-1} \cdot x_3^{\alpha_3-1} \\ &= \left\{ \Gamma(8 + 7 + 6) / (\Gamma(8) \cdot \Gamma(7) \cdot \Gamma(6)) \right\} \cdot 0.55^7 \cdot 0.15^6 \cdot 0.3^5 = 2.354 \end{aligned}$$

ベータ分布を(図24)に示す。グラフの作成には、Excel の等高線機能を使用した。

(導出)

互いに独立な事象 i の発生確率が x_i で、発生回数が $\alpha_i - 1$ であれば、事象の塊 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ の確率密度関数は、 $f(\mathbf{x}) = f(x_1, x_2, \dots, x_n) = C x_1^{\alpha_1-1} x_2^{\alpha_2-1} \dots x_n^{\alpha_n-1} \quad \dots \text{(式24)}$ で表される。

この式の積分(確率)が 1 になる定数 C が求まれば、(式24)がディリクレ分布の確率密度関数になる。

導出にあたって、次のガンマ関数の積を求めておく。ガンマ関数の定義に従って式変形する。

$$\begin{aligned} \prod_{i=1}^n \Gamma(\alpha_i) &= \Gamma(\alpha_1) \dots \Gamma(\alpha_n) = \int_0^\infty e^{-t_1} t_1^{\alpha_1-1} dt_1 \dots \int_0^\infty e^{-t_n} t_n^{\alpha_n-1} dt_n \\ &= \int_0^\infty \dots \int_0^\infty e^{-t_1 - \dots - t_n} (t_1^{\alpha_1-1} \dots t_n^{\alpha_n-1}) dt_1 \dots dt_n \end{aligned}$$

この重積分を求めるため、変数変換する。変数変換のために、ヤコビアン $J(u, y)$ (付録16)を求める。

変数変換 $u_1 + \dots + u_n = 1, t_1 = u_1 y, \dots, t_{n-1} = u_{n-1} y, t_n = u_n y = (1 - u_1 - \dots - u_{n-1}) y$

$$J(u, y) = \frac{\partial(t)}{\partial(u, y)} = \begin{vmatrix} \partial t_1 / \partial u_1 & \partial t_1 / \partial u_2 & \dots & \partial t_1 / \partial u_{n-1} & \partial t_1 / \partial y \\ \partial t_2 / \partial u_1 & \partial t_2 / \partial u_2 & \dots & \partial t_2 / \partial u_{n-1} & \partial t_2 / \partial y \\ \dots & \dots & \dots & \dots & \dots \\ \partial t_n / \partial u_1 & \partial t_n / \partial u_2 & \dots & \partial t_n / \partial u_{n-1} & \partial t_n / \partial y \end{vmatrix} = \begin{vmatrix} y & 0 & \dots & 0 & u_1 \\ 0 & y & \dots & 0 & u_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & y & u_{n-1} \\ -y & -y & \dots & -y & u_n \end{vmatrix} = \begin{vmatrix} y & 0 & \dots & 0 & u_1 \\ 0 & y & \dots & 0 & u_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & y & u_{n-1} \\ 0 & 0 & \dots & 0 & 1 \end{vmatrix} = y^{n-1} \quad \text{(三角行列の行列式は対角要素の積になる)}$$

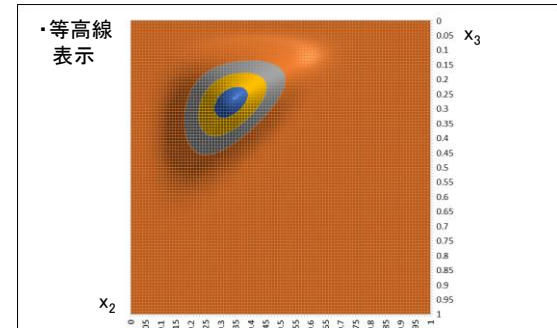
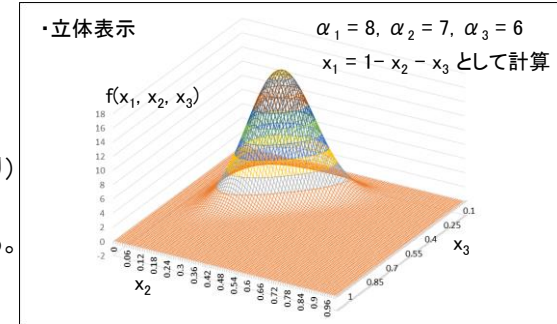
(n 行目を 1 行目から n 行目の和で置き換える)

$$\begin{aligned} \prod_{i=1}^n \Gamma(\alpha_i) &= \int_0^\infty \dots \int_0^\infty e^{-u_1 y - \dots - u_n y} ((u_1 y)^{\alpha_1-1} \dots (u_n y)^{\alpha_n-1}) y^{n-1} du_1 \dots du_n dy \\ &= \int_0^\infty e^{-y} (y^{\alpha_1-1} \dots y^{\alpha_n-1}) y^{n-1} dy \int_0^\infty \dots \int_0^\infty u_1^{\alpha_1-1} \dots u_n^{\alpha_n-1} du_1 \dots du_n = \int_0^\infty e^{-y} y^{\alpha_1 + \dots + \alpha_n - 1} dy \int_0^\infty \dots \int_0^\infty u_1^{\alpha_1-1} \dots u_n^{\alpha_n-1} du_1 \dots du_n \\ &= \Gamma(\alpha_1 + \alpha_2 + \dots + \alpha_n) \int_0^\infty \dots \int_0^\infty u_1^{\alpha_1-1} u_2^{\alpha_2-1} \dots u_n^{\alpha_n-1} du_1 du_2 \dots du_n = \Gamma\left(\sum_{i=1}^n \alpha_i\right) \int_0^\infty \dots \int_0^\infty u_1^{\alpha_1-1} u_2^{\alpha_2-1} \dots u_n^{\alpha_n-1} du_1 du_2 \dots du_n \end{aligned}$$

よって、 $\int_0^\infty \dots \int_0^\infty u_1^{\alpha_1-1} u_2^{\alpha_2-1} \dots u_n^{\alpha_n-1} du_1 du_2 \dots du_n = \frac{\prod_{i=1}^n \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^n \alpha_i)} \quad \dots \text{(式25)}$

従って、 $C \int_0^\infty \dots \int_0^\infty x_1^{\alpha_1-1} x_2^{\alpha_2-1} \dots x_n^{\alpha_n-1} dx_1 \dots dx_n = 1$ を成立させる定数 C は、 $C = \Gamma(\sum_{i=1}^n \alpha_i) / \prod_{i=1}^n \Gamma(\alpha_i) \quad \dots \text{(式26)}$ となる。

よって、(式24)に、この定数 C を代入すれば、ディリクレ分布の確率密度関数(式23)が求まる。



(図24)ディリクレ分布

19. 正規分布(Normal distribution, Gaussian distribution)

■概要

1733年、ド・モアブル(フランス、数学者)によって正規分布は誕生した。また、ガウス(ドイツ、数学者)の誤差理論が正規分布の基礎となったことから、ガウス分布とも呼ばれる。データが平均値(期待値)の付近に集まる分布であり、「normal (ありふれた)」という名前の通り、**多くの現象に当てはまる確率分布であり、幅広く使用されている。**

■確率密度関数

正規分布の確率密度関数は次の式で表される。期待値と分散の導出は(付録39)に示す。

$$f(x) = \{ 1 / (2\pi\sigma^2)^{1/2} \} e^{-(x-\mu)^2/(2\sigma^2)} \quad \dots \text{ (式27)}$$

x : 確率変数, μ : 期待値, σ^2 : 分散 (なお、分散 σ^2 がまず定義され、次いで標準偏差 σ が定義されている)

(計算例)

$$\mu = 1, \sigma^2 = 1, x = 0 \text{ のときの確率密度 } f(0) \text{ の値は、} f(0) = \{ 1 / (2\pi \cdot 1)^{1/2} \} e^{-(0-1)^2/(2 \cdot 1)} = 0.242$$

正規分布を(図25)に示す。

(導出)

正規分布の導出方法としては誤差関数法を使用する方法もあるが、ここでは二項分布の極限としての導出方法を記す。

二項分布 $B(n, p) = P(X = x) = {}_n C_x p^x (1-p)^{n-x} = n! / \{ x! (n-x)! \} p^x q^{n-x}$ 但し、 $p + q = 1$

ここで、スターリングの公式 $n! \doteq (2\pi n)^{1/2} (n/e)^n$ (付録8)を使って、階乗を指数関数で近似する。

$$B(n, p) \doteq \{ (2\pi n)^{1/2} (n/e)^n \} / \{ \{ (2\pi x)^{1/2} (x/e)^x \} \{ (2\pi (n-x))^{1/2} ((n-x)/e)^{n-x} \} \} p^x q^{n-x}$$

$$(2\pi)^{1/2} B(n, p) \doteq e^{-n+x+n-x} n^{n+1/2} p^x q^{n-x} / \{ x^{x+1/2} (n-x)^{n-x+1/2} \} = n^{n+1/2} p^x q^{n-x} / \{ x^{x+1/2} (n-x)^{n-x+1/2} \}$$

$$(2\pi npq)^{1/2} B(n, p) \doteq n^{n+1/2} p^{x+1/2} q^{n-x+1/2} / \{ x^{x+1/2} (n-x)^{n-x+1/2} \} \doteq (np/x)^{x+1/2} (nq/(n-x))^{n-x+1/2}$$

ここで、 $z = (x - np)/(npq)^{1/2}$ とおくと、二項分布であるので $n \rightarrow \infty$ で $p = x/n$, よって、 $z \rightarrow 0$

また、 $x = np + (npq)^{1/2} z$, $n - x = n - (np + (npq)^{1/2} z) = n(1 - p) - (npq)^{1/2} z = nq - (npq)^{1/2} z$

$$\begin{aligned} (2\pi npq)^{1/2} B(n, p) &\doteq (np/(np + (npq)^{1/2} z))^{np + (npq)^{1/2} z + 1/2} (nq/(nq - (npq)^{1/2} z))^{nq - (npq)^{1/2} z + 1/2} \\ &= (1 + ((npq)^{1/2} z)/(np))^{-np - (npq)^{1/2} z} (1 - ((npq)^{1/2} z)/(nq))^{-nq + (npq)^{1/2} z} \\ &= (1 + ((npq)^{1/2} z)/(np))^{-np - (npq)^{1/2} z} (1 - ((npq)^{1/2} z)/(nq))^{-nq + (npq)^{1/2} z} \end{aligned}$$

ここで、 n が十分に大きいとき、 $z \rightarrow 0$ であるので、 $(1 + ((npq)^{1/2} z)/(np))^{-1/2} \rightarrow 1$, $(1 - ((npq)^{1/2} z)/(nq))^{-1/2} \rightarrow 1$ となる。よって、

$$(2\pi npq)^{1/2} B(n, p) \doteq (1 + ((npq)^{1/2} z)/(np))^{-np - (npq)^{1/2} z} (1 - ((npq)^{1/2} z)/(nq))^{-nq + (npq)^{1/2} z}$$

この両辺の対数を取ると、

$$\begin{aligned} \log (2\pi npq)^{1/2} B(n, p) &\doteq - (np + (npq)^{1/2} z) \log (1 + (npq)^{1/2} z)/(np) - (nq - (npq)^{1/2} z) \log (1 - (npq)^{1/2} z)/(nq) \\ &\doteq - np (1 + (q^{1/2} z)/(np)^{1/2}) \log (1 + (q^{1/2} z)/(np)^{1/2}) - nq (1 - (p^{1/2} z)/(nq)^{1/2}) \log (1 - (p^{1/2} z)/(nq)^{1/2}) \end{aligned}$$

$u = (q^{1/2} z)/(np)^{1/2}$, $v = (p^{1/2} z)/(nq)^{1/2}$ とおき、 $\log$ 関数をマクローリン展開(付録3)する。

$$\log (2\pi npq)^{1/2} B(n, p) \doteq - np (1 + u) \log (1 + u) - nq (1 - v) \log (1 - v) = - np (1 + u) (u - u^2/2 + u^3/3 - \dots) - nq (1 - v) (-v - v^2/2 - v^3/3 - \dots)$$

ここで、 $npu = np (q^{1/2} z)/(np)^{1/2} = (npq)^{1/2} z$, $nqv = nq (p^{1/2} z)/(nq)^{1/2} = (npq)^{1/2} z$ であるので、 $-npu + nqv = 0$ となり、

$$\begin{aligned} \log (2\pi npq)^{1/2} B(n, p) &\doteq -npu - np \{ (-u^2/2 + u^3/3 - \dots) \} + nqv - nq \{ (-v^2/2 - v^3/3 - \dots) \} + (v^2 + v^3/2 + v^4/3 + \dots) \\ &\doteq np \{ (-1/2) u^2 + (1/6) u^3 + \dots \} + nq \{ (-1/2) v^2 + (-1/6) v^3 + \dots \} \end{aligned}$$

$u^2 = (q z^2)/(np)$, $u^3 = (q^{3/2} z^3)/(np)^{3/2}$, $u^4 = (q^2 z^4)/(np)^2$, ... であるが、 $n \rightarrow \infty$ で $z \rightarrow 0$ であり、 z^2 の項に比べて、 $z^3, z^4, \dots$ の項が速く 0 に近づき、 z^2 の項のみ残る。

$v^2, v^3, v^4, \dots$ についても同様であり、 $n \rightarrow \infty$ では、 z^2 の項のみ残る。従って、

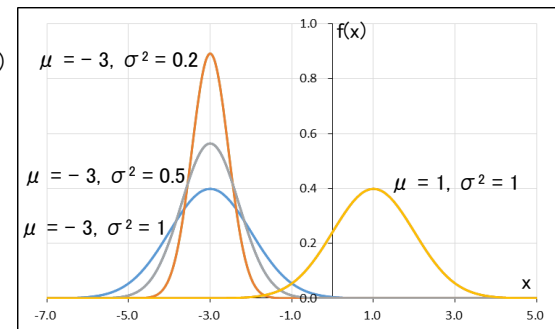
$$\lim_{n \rightarrow \infty} \log (2\pi npq)^{1/2} B(n, p) = np (-1/2) u^2 + nq (-1/2) v^2 = np (-1/2) (q z^2)/(np) + nq (-1/2) (p z^2)/(nq) = -1/2 (p + q) z^2 = -z^2/2 \quad (\text{注}) p + q = 1$$

$$\text{よって、} \lim_{n \rightarrow \infty} B(n, p) = e^{-z^2/2} / (2\pi npq)^{1/2}$$

ここで、二項分布の期待値を μ で表せば $\mu = np$ 、分散を σ^2 で表せば $\sigma^2 = np(1-p) = npq$ となるので、 $z = (x - np)/(npq)^{1/2} = (x - \mu)/\sigma$

$$f(x) = \lim_{n \rightarrow \infty} B(n, p) = e^{-z^2/2} / (2\pi \sigma^2)^{1/2} = \{ 1 / (2\pi \sigma^2)^{1/2} \} e^{-(x-\mu)^2/(2\sigma^2)}$$

これにより、二項分布の試行回数の極限が正規分布になることが確認できる。



(図25) 正規分布

20. 標準正規分布(Standard Normal distribution)

■概要

確率分布の標準化とは確率分布の期待値が 0、分散が 1 になるように変数変換すること(付録19)。標準正規分布とは、期待値が μ 、分散が σ^2 の正規分布を標準化した分布である。標準正規分布の平均は 0、分散は 1 になる。標準化により、データの単位の無次元化を実現でき、データ同士を比較することが可能になる。

■確率密度関数

正規分布の確率密度関数は次の式で表される。期待値は 0、分散は 1 である。

$$f(x) = (1/(2\pi)^{1/2}) e^{-x^2/2} \dots \text{(式28)}$$

(計算例)

確率密度 $f(1)$ の値は、 $f(1) = (1/(2\pi)^{1/2}) e^{-1^2/2} = 0.242$

標準正規分布を(図26)に示す。

(導出)

正規分布を標準化することにより標準正規分布が得られる。

正規分布の(式27)において、 $\mu = 0$ 、 $\sigma^2 = 1$ を代入して、(式28)の標準正規分布が得られる。

なお、正規分布を直接変数変換して標準正規分布が得られることを確認しておく。

まず、確率変数 U が正規分布 $g(u) = \{1 / (2\pi\sigma^2)^{1/2}\} e^{-(u-\mu)^2/(2\sigma^2)}$ に従うとき、新たな確率変数 $X = (aU + b)$ の確率密度関数を $f(x)$ とする。

$x = au + b$ より $u = (x - b)/a$ 、 $dx/du = a$ より $du/dx = 1/a$ (ヤコビアンになる)

よって、(付録14)の変数変換の式 $g(u) du = f(x) dx$ を使って、

$$f(x) = g(u) du/dx = \{1/(2\pi\sigma^2)^{1/2}\} \{ \exp(-(x-b)/a - \mu)^2/(2\sigma^2) \} (1/a) = (1/((2\pi)^{1/2}(a\sigma))) \exp(-(x - (a\mu + b))^2/(2(a\sigma)^2))$$

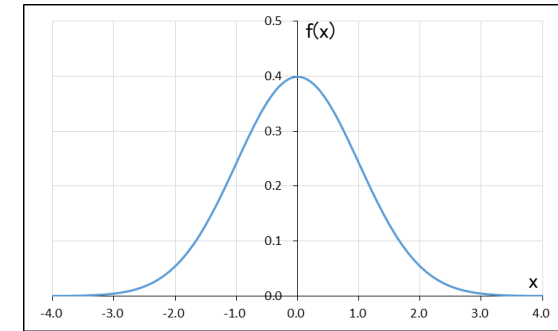
つまり、 $f(x)$ は、期待値 $a\mu + b$ 、分散 $a^2\sigma^2$ の正規分布に従うことが分かる。

ここで、 $a = 1/\sigma$ 、 $b = -\mu/\sigma$ とおくと、

$x = au + b = (1/\sigma)u - \mu/\sigma = (u - \mu)/\sigma$ なので、

$$\begin{aligned} f(x) &= (1/((2\pi)^{1/2}(1/\sigma)\sigma)) \exp(-(x - ((1/\sigma)\mu + (-\mu/\sigma)))^2/(2((1/\sigma)\sigma)^2)) \\ &= (1/(2\pi)^{1/2}) \exp(-x^2/2) \end{aligned}$$

この分布は、期待値 0、分散 1 の正規分布である。これは、標準正規分布である。



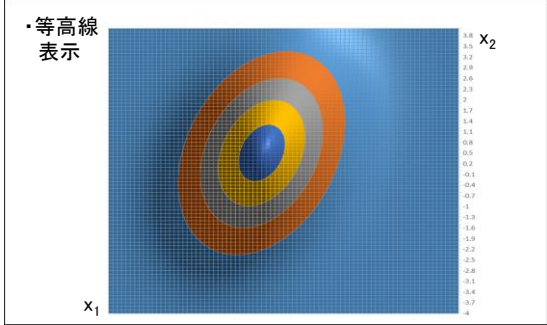
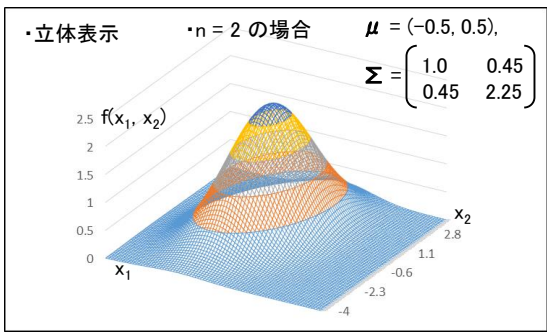
(図26)標準正規分布

21. 多変量正規分布(Multivariate Normal distribution)

■概要

多変量正規分布は、正規分布を n 次元に多変量化した確率分布である。

(表3)変数変換の比較



	正規分布 (1 次元)	多変量正規分布 (n 次元)
標準形 g(u)	$g(u) = (1/(2\pi)^{1/2}) e^{-u^2/2}$ 期待値 $E[U] = 0$ 分散 $V[U] = 1$	$g(u) = (1/(2\pi)^{n/2}) e^{-u^T u/2}$ 期待値 $E[U] = \mathbf{0}$ ゼロベクトル 分散 $V[U] = \mathbf{I}$ 単位行列
変数変換 g(u) → f(x)	$x = \sigma u + \mu$ つまり、 $u = \sigma^{-1}(x - \mu)$ (ヤコビアン) $du/dx = \sigma^{-1}$ 従って、 $f(x) = g(u) \sigma^{-1}$	$x = \mathbf{A}u + \mu$ つまり、 $u = \mathbf{A}^{-1}(x - \mu)$ (ヤコビアン) $ J = \mathbf{A}^{-1} $ 従って、 $f(x) = g(u) \mathbf{A}^{-1} $
一般形 f(x)	$f(x) = g(u) \sigma^{-1} = g(\sigma^{-1}(x - \mu)) \sigma^{-1}$ $= (1/(2\pi \sigma^2)^{1/2}) e^{-(x - \mu)^2/(2\sigma^2)}$ 期待値 $E[X] = \mu$, 分散 $V[X] = \sigma^2$	$f(x) = g(u) \mathbf{A}^{-1} = g(\mathbf{A}^{-1}(x - \mu)) \mathbf{A}^{-1} $ $= \{1/((2\pi)^{n/2} \Sigma ^{1/2})\} e^{-(x - \mu)^T \Sigma^{-1} (x - \mu)/2}$ 期待値 $E[X] = \mu$, 分散共分散 $V[X] = \Sigma$

■確率密度関数

多変量正規分布の確率密度関数は次の式で表される。
期待値ベクトルは μ , 分散共分散行列は Σ である。多変量正規分布を(図27)に示す。

$f(x) = \{1/((2\pi)^{n/2} |\Sigma|^{1/2})\} e^{-(x - \mu)^T \Sigma^{-1} (x - \mu)/2} \dots$ (式29)

(導出) 多変量正規分布は、正規分布に従う独立な確率変数を n 次元(多変量)に拡張したものである。
変数変換に焦点を当てて、1 次元の正規分布と多変量の正規分布の比較を(表3)に示す。

1) 標準多変量正規分布の確率密度関数の導出

確率変数ベクトル $U^T = (U_1, U_2, \dots, U_n)$, $U_1, U_2, \dots, U_n$ (T は転置)は互いに独立で標準正規分布に従い、実現値ベクトルを u とすれば、同時確率密度関数 $g(u)$ は 1 次元の標準正規分布の確率密度関数 $g(u_i) = (1/(2\pi)^{1/2}) e^{-u_i^2/2}$ (但し、期待値 $E[U_i] = 0$, 分散 $V[U_i] = 1$) の積で表される。

$$g(u) = g(u_1, u_2, \dots, u_n) = \prod_{i=1}^n g(u_i) = \prod_{i=1}^n (1/(2\pi)^{1/2}) e^{-u_i^2/2} = (1/(2\pi)^{n/2}) e^{-\sum_{i=1}^n (-u_i^2/2)} = (1/(2\pi)^{n/2}) e^{-(u_1^2 + u_2^2 + \dots + u_n^2)/2} = (1/(2\pi)^{n/2}) e^{-(u^T u)/2}$$

2) 次に、標準多変量正規分布の確率密度関数を変数変換して、多変量正規分布の一般形を求める。

確率密度関数(標準形) $g(u)$ を $X = \mathbf{A}U + \mu$ を使って確率密度関数(一般形) $f(x)$ へ変換する(但し、 $\mathbf{A}$ は正則行列、つまり逆行列が存在するとする)。

この変数変換によって下記の通り、 $f(x)$ は、期待値ベクトル $E[X] = \mu$, 分散共分散行列 $V[X] = \Sigma$ の確率密度関数になることが確認できる。

$E[X] = E[\mathbf{A}U + \mu] = \mathbf{A} E[U] + E[\mu] = \mathbf{A} \mathbf{0} + \mu = \mu$ ($\mathbf{0}$ は、ゼロベクトル)

$V[X] = E[(X - \mu)(X - \mu)^T] = E[(\mathbf{A}U)(\mathbf{A}U)^T] = E[\mathbf{A}U U^T \mathbf{A}^T] = \mathbf{A} E[U U^T] \mathbf{A}^T = \mathbf{A} \mathbf{I} \mathbf{A}^T = \mathbf{A} \mathbf{A}^T = \Sigma$ とおく。

なお、 U_i は互いに独立なので、 $i \neq j$ のとき $E[U_i U_j] = E[U_i] E[U_j] = 0 \cdot 0 = 0$ また、 $E[U_i^2] = V[U_i] + E[U_i]^2 = 1 + 0 = 1$ より、 $E[U U^T] = \mathbf{I}$ である。

この変換前後で確率は保存され、 $f(x) dx = g(u) du$ という関係が成立している。ここで、 dx と du の関係はヤコビアンで表される(付録16)。

ダミー変数 $y = v$ を用意して、 $u = \mathbf{A}^{-1}(x - \mu)$, $v = y$ より、ヤコビアンを求める。対角行列の行列式になるので、対角要素の積として求める。

$$J(x, y) = \partial(u, v) / \partial(x, y) = \begin{vmatrix} \partial u(x, y) / \partial x & \partial u(x, y) / \partial y \\ \partial v(x, y) / \partial x & \partial v(x, y) / \partial y \end{vmatrix} = \begin{vmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{vmatrix} = |\mathbf{A}^{-1}|$$

$f(x, y) = g(u, v) J(x, y) = g(u, v) |\mathbf{A}^{-1}|$ となるが、 $y = v$ であるので、 $f(x) = g(u) |\mathbf{A}^{-1}|$ となる。付録48を式変形の参考にして、

また、 $(\mathbf{A}^{-1})^T \mathbf{A}^{-1} = (\mathbf{A}^T)^{-1} \mathbf{A}^{-1} = (\mathbf{A} \mathbf{A}^T)^{-1} = \Sigma^{-1}$, $|\Sigma| = |\mathbf{A} \mathbf{A}^T| = |\mathbf{A}| |\mathbf{A}^T| = |\mathbf{A}| |\mathbf{A}|^T = |\mathbf{A}| |\mathbf{A}| = |\mathbf{A}|^2$ となるので、

$$f(x) = g(u) |\mathbf{A}^{-1}| = g(\mathbf{A}^{-1}(x - \mu)) |\mathbf{A}^{-1}| = (1/(2\pi)^{n/2}) e^{-(\mathbf{A}^{-1}(x - \mu))^T (\mathbf{A}^{-1}(x - \mu))/2} |\mathbf{A}^{-1}| = (1/(2\pi)^{n/2}) e^{-(x - \mu)^T (\mathbf{A}^{-1})^T \mathbf{A}^{-1} (x - \mu)/2} |\mathbf{A}^{-1}|$$

$$= \{1/((2\pi)^{n/2} |\Sigma|^{1/2})\} e^{-(x - \mu)^T \Sigma^{-1} (x - \mu)/2} = \{1/((2\pi)^{n/2} |\Sigma|^{1/2})\} e^{-(x - \mu)^T \Sigma^{-1} (x - \mu)/2}$$
 これで多変量正規分布(一般形)の確率密度関数が求まった。

22. 対数正規分布(Log-Normal distribution)

■概要

確率変数 U が正規分布に従うとき、確率変数 $X = e^U$ が従う確率分布を対数正規分布という（このとき、 $U = \ln X$ で、 $\ln$ は自然対数である）。確率密度関数のパラメータ設定によって、確率変数 X が大きな値をとるときその確率密度の減少度合いが小さな確率分布を作り出すことができる。このような分布は、年間所得の分布とかウィルスの潜伏期間の分布などに使用されている。

■確率密度関数

対数正規分布の確率密度関数は次の式で表される。期待値と分散の導出は(付録41)に示す。

$$f(x) = e^{-(\ln x - \mu)^2 / (2\sigma^2)} / ((2\pi\sigma^2)^{1/2} x) \quad \dots \text{ (式30)}$$

x : 確率変数 μ : パラメータ(定数), σ : パラメータ(定数)

(計算例)

$$\mu = 0.5, \sigma = 0.3, x = 1.5 \text{ のときの確率密度 } f(1.5) \text{ の値は、} f(1.5) = e^{-(\ln 1.5 - 0.5)^2 / (2 \cdot 0.3^2)} / ((2\pi \cdot 0.3^2)^{1/2} \cdot 1.5) = 0.844$$

対数正規分布を(図28)に示す。

(導出)

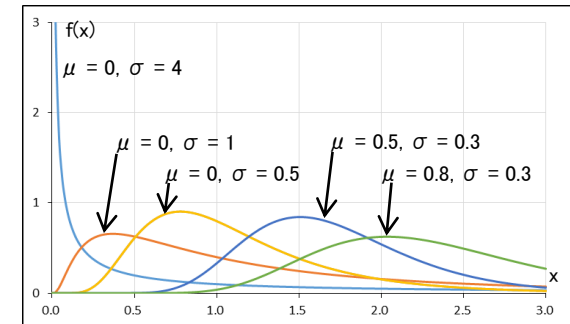
正規分布に従う確率変数 U を、 $X = e^U$ の関係を使用して、確率変数 X へ変数変換する。

確率変数 U は正規分布に従うので、 $g(u) = e^{-(u - \mu)^2 / (2\sigma^2)} / (2\pi\sigma^2)^{1/2}$

また、 $x = e^u$ より $u = \ln x$ そして、 $dx/du = e^u$ であるので $du/dx = 1/e^u$

(付録14)より、 $f(x) = g(u) du/dx$

$$f(x) = \{ e^{-(u - \mu)^2 / (2\sigma^2)} / (2\pi\sigma^2)^{1/2} \} / e^u = \{ e^{-(\ln x - \mu)^2 / (2\sigma^2)} / (2\pi\sigma^2)^{1/2} \} / e^{\ln x} \\ = \{ e^{-(\ln x - \mu)^2 / (2\sigma^2)} / (2\pi\sigma^2)^{1/2} \} / x = e^{-(\ln x - \mu)^2 / (2\sigma^2)} / ((2\pi\sigma^2)^{1/2} x)$$



(図28) 対数正規分布

23. カイニ乗分布(χ²乗分布 Chi-square distribution)

■概要

互いに独立な確率変数 $U_1, U_2, \dots, U_n$ が標準正規分布 ($\mu = 0, \sigma^2 = 1$) に従うとき、確率変数 $X = U_1^2 + U_2^2 + \dots + U_n^2 \dots$ (式31) が従う確率分布を、自由度 n (n は和の数の数に相当) のカイニ乗分布という。カイニ乗分布は確率変数の二乗和の確率分布であるので、分散(偏差の二乗和)の検定に使用されたり、構成比率が理論値に合っているかどうか(つまり、ズレの二乗和)を確認したい適合度検定などに使用する。確率変数の二乗和を扱うところから二乗和分布と言えは済むところを、統計でよく使用されている文字 X を避けて X に対応するギリシャ文字の χ (カイ)を使用して、カイニ乗分布と言われている。

■確率密度関数

カイニ乗分布の確率密度関数は次の式で表される。期待値と分散の導出は(付録42)に示す。

$f_n(x) = (x^{n/2-1} e^{-x/2}) / (\Gamma(n/2) 2^{n/2}) \dots$ (式32)

x : 確率変数 n : 自由度(二乗和の数)

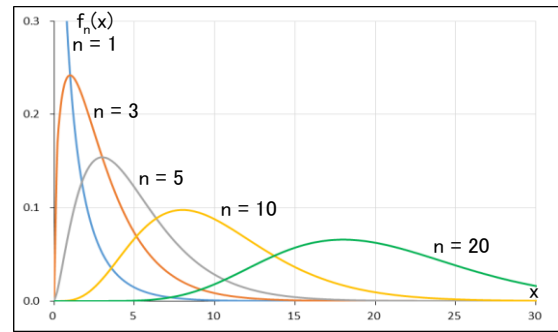
なお、自由度 $n = 2$ のとき、 $f_2(x) = (x^{2/2-1} e^{-x/2}) / (\Gamma(2/2) 2^{2/2}) = (x^0 e^{-x/2}) / (\Gamma(1) 2^1) = (1/2) e^{-x/2}$ となり、 $\lambda = 1/2$ のときの指数分布と一致する。また、 $k = n / 2, \lambda = 1/2$ のときのガンマ分布と一致する。

(計算例)

$n = 5, x = 3$ のときの確率密度 $f_5(3)$ の値は、 $f_5(3) = (3^{5/2-1} e^{-3/2}) / (\Gamma(5/2) 2^{5/2}) = 0.1024$

(注) $\Gamma(5/2)$ の値は、Excel のガンマ関数を使用して求めた。gamma(5/2) = 0.154

カイニ乗分布を(図29)に示す。



(図29)カイニ乗分布

(導出)

帰納法を使用して(式32)が成立することを確認する(なお、ヤコビアンを使った多重積分によって導出することもできる)。

(1) まず、 $n = 1$ で、(式32)が成立することを確認する。互いに独立な確率変数が従う標準正規分布は、(式28)で表される。

$f_1(u_1) = (1/(2\pi)^{1/2}) e^{-u_1^2/2}$

$x = u_1^2$ であり、 $f_1(u_1)$ は偶関数であるので、累積分布関数は、 $F_1(x) = P(X < x) = P(U_1^2 < x) = P(-x^{1/2} < u_1 < x^{1/2}) = \int_{-x^{1/2}}^{x^{1/2}} f_1(u_1) du_1 = 2 \int_0^{x^{1/2}} f_1(u_1) du_1$

累積分布関数を微分して確率密度関数を得る。ここで、関数の微分の公式 $d/dx \int_{p(x)}^{q(x)} f(t) dt = f(q(x)) q'(x) - f(p(x)) p'(x)$ と $\Gamma(1/2) = (\pi)^{1/2}$ (付録13)を使って、

$d/dx [F_1(x)] = 2 \{ f_1(x^{1/2}) (x^{1/2})' - f_1(0) \cdot 0 \} = f_1(x^{1/2}) x^{-1/2} = (1/(2\pi)^{1/2}) e^{-(x^{1/2})^2/2} x^{-1/2} = (1/(2\pi)^{1/2}) e^{-x/2} x^{-1/2} = (x^{1/2-1} e^{-x/2}) / (\Gamma(1/2) 2^{1/2})$

よって、 $n = 1$ のとき、(式32)は成り立つ。

(2) (式32)が、 $n = k$ で成立すると仮定して、 $n = k + 1$ のときも成立することを示す。

$x = u_1^2 + u_2^2 + \dots + u_k^2 + u_{k+1}^2 = y_1 + y_2 + \dots + y_k + y_{k+1}$

ここで、 $y = y_1 + y_2 + \dots + y_k$ とおけば、 $x = y + y_{k+1}$

(付録17)の畳み込み積分を使用して変数変換する。なお、積分範囲は、 $y_{k+1} = x - y \geq 0$ であるので、 $0 \leq y \leq x$ である。

$$f_{k+1}(x) = \int_0^x f_1(x-y) f_k(y) dy = \int_0^x \{ ((x-y)^{1/2-1} e^{-(x-y)/2}) / (\Gamma(1/2) 2^{1/2}) \} \{ (y^{k/2-1} e^{-y/2}) / (\Gamma(k/2) 2^{k/2}) \} dy$$
$$= \{ e^{-x/2} / (\Gamma(1/2) \Gamma(k/2) 2^{(k+1)/2}) \} \int_0^x (x-y)^{-1/2} y^{k/2-1} dy$$

ここで、 $y = xz$ として、置換積分を行う。 $z = y/x, x-y = x-xz = x(1-z), dy = x dz$, 積分範囲は $y : 0 \rightarrow x$ に対して、 $z : 0 \rightarrow 1$

$f_{k+1}(x) = \{ e^{-x/2} / (\Gamma(1/2) \Gamma(k/2) 2^{(k+1)/2}) \} \int_0^1 (x(1-z))^{-1/2} (xz)^{k/2-1} x dz = \{ x^{k/2-1/2} e^{-x/2} / (\Gamma(1/2) \Gamma(k/2) 2^{(k+1)/2}) \} \int_0^1 (1-z)^{-1/2} z^{k/2-1} dz$

(付録13)のベータ関数の定義、および $B(k/2, 1/2) = \Gamma(k/2) \Gamma(1/2) / \Gamma(k/2 + 1/2) = \Gamma(k/2) \Gamma(1/2) / \Gamma((k+1)/2)$ の関係を使う。

$f_{k+1}(x) = \{ x^{k/2-1/2} e^{-x/2} / (\Gamma(1/2) \Gamma(k/2) 2^{(k+1)/2}) \} B(k/2, 1/2) = \{ x^{k/2-1/2} e^{-x/2} / (\Gamma(1/2) \Gamma(k/2) 2^{(k+1)/2}) \} \{ \Gamma(k/2) \Gamma(1/2) / \Gamma((k+1)/2) \}$
$$= \{ x^{(k+1)/2-1} e^{-x/2} \} / \{ \Gamma((k+1)/2) 2^{(k+1)/2} \}$$

よって、(式32)が、 $n = k$ で成立すれば、 $n = k + 1$ のときも成立することが分かる。

従って、(1) と (2) から、(式32)が成立することを確認できる。

24. F 分布(F distribution)

■概要

自由度 m のカイニ乗分布に従う確率変数 U と自由度 n のカイニ乗分布に従う確率変数 V が互いに独立であるとき、確率変数 $X = (U / m) / (V / n) \cdots$ (式33) が従う確率分布を自由度 (m, n) の **F 分布**という。F 分布は、全体平均からのデータのズレがデータの属する群の全体平均からのズレと群内の平均からのデータのズレとで構成されるとし、そのズレ(二乗和)の割合を使って群間の差を検定する分散分析などで使われる。F 分布という名前は統計学者 R.A.フィッシャー(イギリス)に因んで付けられている。

■確率密度関数

F 分布の確率密度関数は次の式で表される。期待値と分散の導出は(付録43)に示す。

$$f(m, n) = f(x) = \{ 1/B(m/2, n/2) \} (m/n)^{m/2} x^{(m-2)/2} / \{ 1 + (m/n) x \}^{(m+n)/2} \cdots$$
 (式34)

$$f(m, n) = f(x) = \{ \Gamma((m+n)/2) / \Gamma(m/2) \Gamma(n/2) \} (m/n)^{m/2} x^{(m-2)/2} / \{ 1 + (m/n) x \}^{(m+n)/2} \cdots$$
 (式35)

x : F 分布の確率変数 u : カイニ乗分布の確率変数 (m : 自由度) v : カイニ乗分布の確率変数 (n : 自由度)

(計算例)

$m = 5, n = 2$ のときの確率密度 $f(2.5) = \{ \Gamma((5+2)/2) / \Gamma(5/2) \Gamma(2/2) \} (5/2)^{5/2} 2.5^{(5-2)/2} / \{ 1 + (5/2) \cdot 2.5 \}^{(5+2)/2} = 0.095$

(注) 例えば、 $\Gamma(7/2)$ の値は、Excel のガンマ関数を使用して求めた。gamma(7/2) = 3.32

なお、F 分布を(図30)に示す。(式34)で x を $1/x$ で置き換えると自由度の順番を入れ替えた $f(1/x) = F(n, m)$ となる。

(導出)

F 分布は、カイニ乗分布に従う 2 つの確率変数 U (自由度 m) と V (自由度 n) の比(式38)の確率分布である。

F 分布の確率密度関数は確率変数の変数変換によって導出される。

確立変数 U が従う自由度 m のカイニ乗分布: $g_U(u) = (u^{m/2-1} e^{-u/2}) / (\Gamma(m/2) 2^{m/2}) \cdots$ (式36)

確立変数 V が従う自由度 n のカイニ乗分布: $g_V(v) = (v^{n/2-1} e^{-v/2}) / (\Gamma(n/2) 2^{n/2}) \cdots$ (式37)

確率変数 U と V の 2 変数を、(式38)と(式39)で示す確率変数 X と Y へ変数変換を行う。

(式38)は自由度を加味した 2 つの確率変数 U と V の比であり、F 検定に使用される式である。(式39)は 2 変数変換のためのダミーの確率変数である。

$x = (u/m) / (v/n) \cdots$ (式38) $y = v \cdots$ (式39)

ここで、変数変換のためのヤコビアン $J(x, y)$ を求める。

(式38)と(式39)から、 $u(x, y) = (m/n) x y \cdots$ (式40) $v(x, y) = y \cdots$ (式41)

$$J(x, y) = \partial(u, v) / \partial(x, y) = \begin{vmatrix} \partial u(x, y) / \partial x & \partial u(x, y) / \partial y \\ \partial v(x, y) / \partial x & \partial v(x, y) / \partial y \end{vmatrix} = \begin{vmatrix} \partial((m/n)xy) / \partial x & \partial((m/n)xy) / \partial y \\ \partial y / \partial x & \partial y / \partial y \end{vmatrix} = \begin{vmatrix} (m/n)y & (m/n)x \\ 0 & 1 \end{vmatrix} = (m/n)y$$

U と V は独立であるので、その同時確率分布 $g_{UV}(u, v)$ は積で表すことができる。 $g_{UV}(u, v) = g_U(u) g_V(v) = \{ (u^{m/2-1} e^{-u/2}) / (\Gamma(m/2) 2^{m/2}) \} \{ (v^{n/2-1} e^{-v/2}) / (\Gamma(n/2) 2^{n/2}) \}$

ここで、ヤコビアンを使用して、 $g_{UV}(u, v)$ から $f_{XY}(x, y)$ を求める。

$f_{XY}(x, y) = g_{UV}((m/n)xy, y) (m/n)y = \{ (((m/n)xy)^{m/2-1} e^{-(m/n)xy/2}) / (\Gamma(m/2) 2^{m/2}) \} \{ (y^{n/2-1} e^{-y/2}) / (\Gamma(n/2) 2^{n/2}) \} (m/n)y$

次いで、 $f_{XY}(x, y)$ を変数 y で積分して $f_X(x)$ を求める(周辺分布の確率密度関数を求める、周辺化するという)。なお、カイニ乗分布の積分範囲は、 $0 \rightarrow \infty$ である。

$$f_X(x) = \int_0^\infty f_{XY}(x, y) dy = \int_0^\infty \{ (((m/n)xy)^{m/2-1} e^{-(m/n)xy/2}) / (\Gamma(m/2) 2^{m/2}) \} \{ (y^{n/2-1} e^{-y/2}) / (\Gamma(n/2) 2^{n/2}) \} (m/n)y dy$$

$$= \{ (m/n)^{m/2} x^{m/2-1} \} / \{ 2^{(m+n)/2} \Gamma(m/2) \Gamma(n/2) \} \int_0^\infty y^{(m+n)/2-1} e^{-y(mx+n)/(2n)} dy$$

ここで、 $t = y(mx+n)/(2n)$ とおく。 $y = 2nt/(mx+n)$, $dy = 2n/(mx+n) dt$ である。ガンマ関数の定義式(付録12)、ベータ関数とガンマ関数の関係(付録13)を使う。

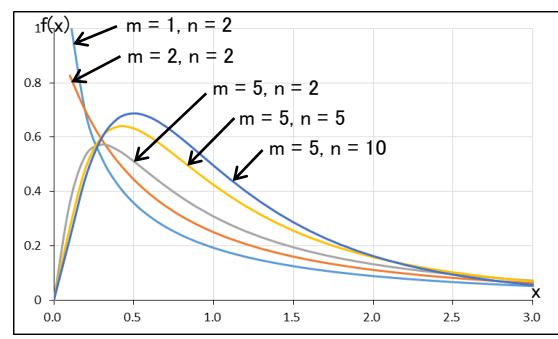
$$f_X(x) = \{ (m/n)^{m/2} x^{m/2-1} \} / \{ 2^{(m+n)/2} \Gamma(m/2) \Gamma(n/2) \} \int_0^\infty (2nt/(mx+n))^{(m+n)/2-1} e^{-t} 2n/(mx+n) dt$$

$$= \{ (m/n)^{m/2} x^{m/2-1} \} / \{ 2^{(m+n)/2} \Gamma(m/2) \Gamma(n/2) \} \{ 2n/(mx+n) \}^{(m+n)/2} \int_0^\infty t^{(m+n)/2-1} e^{-t} dt$$

$$= \{ (m/n)^{m/2} x^{m/2-1} \} / \{ 2^{(m+n)/2} \Gamma(m/2) \Gamma(n/2) \} \{ 2n/(mx+n) \}^{(m+n)/2} \Gamma((m+n)/2) = \{ \Gamma((m+n)/2) / (\Gamma(m/2) \Gamma(n/2)) \} (mx/n)^{m/2} \{ n/(mx+n) \}^{(m+n)/2} x^{-1}$$

$$= \{ 1 / B(m/2, n/2) \} (mx/n)^{m/2} \{ n/(mx+n) \}^{(m+n)/2} x^{-1} = \{ 1 / B(m/2, n/2) \} (m/n)^{m/2} x^{(m-2)/2} / \{ 1 + (m/n)x \}^{(m+n)/2}$$

これによって、F 分布の確率密度関数が得られた。



(図30) F 分布

25. t 分布 (t distribution)

■概要

確率変数 U が標準正規分布に従い、確率変数 V が自由度 n のカイニ乗分布に従うとき、確率変数 $X = U / (V/n)^{1/2}$... (式42) が従う確率分布を自由度 n の t 分布という。
2つの母集団の平均値の差の検定に使用される確率分布である。考案者である W.S.ゴセット(イギリス)の論文のペンネーム「スチューデント」に因んで、スチューデントの t 分布とも言われている。

■確率密度関数

t 分布の確率密度関数は次の式で表される。期待値と分散の導出は(付録44)に示す。

$$f(x) = \left\{ \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \frac{1}{\sqrt{n\pi}} \right\} (1 + x^2/n)^{-(n+1)/2} \dots \text{(式43)}$$

$$f(x) = \left\{ \frac{1}{n^{1/2} B(n/2, 1/2)} \right\} (1 + x^2/n)^{-(n+1)/2} \dots \text{(式44)}$$

x : t 分布の確率変数 n : 自由度

(計算例)

$$n = 2, x = 1.2 \text{ のときの確率密度 } f(1.2) = \left\{ \frac{\Gamma((2+1)/2)}{\Gamma(2/2)} \frac{1}{\sqrt{2\pi}} \right\} (1 + 1.2^2/2)^{-(2+1)/2} = 0.157$$

(注)例えば、 $\Gamma(3/2)$ の値は、Excel のガンマ関数を使用して求めた。gamma(3/2) = 0.886

t 分布を(図31)に示す(Excel の T.DIST 関数を使って確率密度を求めることもできる)。

(導出)

t 分布は、標準正規分布とカイニ乗分布に従う2つの確率変数から(式47)によって作られる新たな確率変数の確率分布である。 t 分布の確率密度関数は確率変数の変数変換によって導出される。

確立変数 U が従う標準正規分布: $g_U(u) = (1/(2\pi))^{1/2} e^{-u^2/2}$... (式45)

確立変数 V が従う自由度 n のカイニ乗分布: $g_V(v) = (v^{n/2-1} e^{-v/2}) / (\Gamma(n/2) 2^{n/2})$... (式46)

確率変数 U と V の2変数を、(式47)と(式48)で示す確率変数 X と Y へ変数変換を行う。

(式47)は差に関わる U を分散に関わる V で無次元化したような式であり、 t 検定に使用される式である。(式48)は2変数変換のためのダミーの確率変数である。

$$x = u / (v/n)^{1/2} \dots \text{(式47)} \quad y = v \dots \text{(式48)}$$

ここで、変数変換のためのヤコビアン $J(x, y)$ を求める。

(式47)と(式48)から、 $u(x, y) = (y/n)^{1/2} x$... (式49) $v(x, y) = y$... (式50)

$$J(x, y) = \frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} \frac{\partial u(x, y)}{\partial x} & \frac{\partial u(x, y)}{\partial y} \\ \frac{\partial v(x, y)}{\partial x} & \frac{\partial v(x, y)}{\partial y} \end{vmatrix} = \begin{vmatrix} \frac{\partial((y/n)^{1/2} x)}{\partial x} & \frac{\partial((y/n)^{1/2} x)}{\partial y} \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} (y/n)^{1/2} & 1/2 (y/n)^{-1/2} x \\ 0 & 1 \end{vmatrix} = (y/n)^{1/2}$$

U と V は独立であるので、その同時確率分布 $g_{UV}(u, v)$ は積で表すことができる。 $g_{UV}(u, v) = g_U(u) g_V(v) = \left\{ (1/(2\pi))^{1/2} e^{-u^2/2} \right\} \left\{ (v^{n/2-1} e^{-v/2}) / (\Gamma(n/2) 2^{n/2}) \right\}$

ここで、ヤコビアンを使用して、 $g_{UV}(u, v)$ から $f_{XY}(x, y)$ を求める。

$$f_{XY}(x, y) = g_{UV}((y/n)^{1/2} x, y) (y/n)^{1/2} = \left\{ (1/(2\pi))^{1/2} e^{-(y/n)x^2/2} \right\} \left\{ (y^{n/2-1} e^{-y/2}) / (\Gamma(n/2) 2^{n/2}) \right\} (y/n)^{1/2}$$

次いで、 $f_{XY}(x, y)$ を変数 y で積分して $f_X(x)$ を求める(周辺分布の確率密度関数を求める、周辺化するという)。

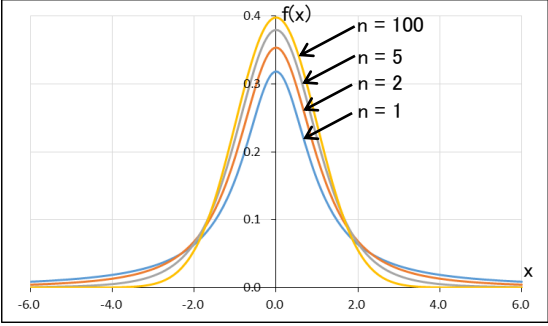
なお、カイニ乗の確率変数 $V \geq 0$ であるので、 y の積分範囲は、 $0 \rightarrow \infty$ である。

$$f_X(x) = \int_0^\infty f_{XY}(x, y) dy = \int_0^\infty \left\{ (1/(2\pi))^{1/2} e^{-(y/n)x^2/2} \right\} \left\{ (y^{n/2-1} e^{-y/2}) / (\Gamma(n/2) 2^{n/2}) \right\} (y/n)^{1/2} dy = (1/\{(2\pi)^{1/2} 2^{n/2} \Gamma(n/2)\}) \int_0^\infty e^{-y(x^2/n+1)/2} y^{(n-1)/2} dy$$

ここで、 $t = y(x^2/n + 1)/2$ とおく。 $y = 2t/(x^2/n + 1)$, $dy = 2/(x^2/n + 1) dt$ である。積分範囲は、 $y : 0 \rightarrow \infty$ に対して、 $t : 0 \rightarrow \infty$

また、ガンマ関数の定義式(付録12)とベータ関数とガンマ関数の関係(付録13)を使う。

$$\begin{aligned} f_X(x) &= \left\{ 1 / ((2\pi)^{1/2} 2^{n/2} \Gamma(n/2)) \right\} \int_0^\infty e^{-t} (2t/(x^2/n + 1))^{(n-1)/2} 2/(x^2/n + 1) dt = \left\{ 1 / ((2\pi)^{1/2} 2^{n/2} \Gamma(n/2)) \right\} (2/(x^2/n + 1))^{(n+1)/2} \int_0^\infty t^{(n-1)/2} e^{-t} dt \\ &= \left\{ 1 / ((2\pi)^{1/2} 2^{n/2} \Gamma(n/2)) \right\} (2/(x^2/n + 1))^{(n+1)/2} \int_0^\infty t^{(n+1)/2-1} e^{-t} dt = \left\{ 1 / ((\pi)^{1/2} \Gamma(n/2)) \right\} (x^2/n + 1)^{-(n+1)/2} \Gamma((n+1)/2) \\ &= \left\{ \Gamma((n+1)/2) / ((\pi)^{1/2} \Gamma(n/2)) \right\} (1 + x^2/n)^{-(n+1)/2} \quad \text{これが、(式43)である。} \\ &= \left\{ \Gamma((n+1)/2) / (\pi^{1/2} n^{1/2} \Gamma(n/2)) \right\} (1 + x^2/n)^{-(n+1)/2} = \left\{ \Gamma((n+1)/2) / (\Gamma(1/2) n^{1/2} \Gamma(n/2)) \right\} (1 + x^2/n)^{-(n+1)/2} \\ &= \left\{ 1 / (n^{1/2} B(n/2, 1/2)) \right\} (1 + x^2/n)^{-(n+1)/2} \quad \text{これが、(式44)である。} \end{aligned}$$



(図31) t 分布

26. コーシー分布 (Cauchy distribution)

■概要

標準コーシー分布は、t 分布の自由度を 1 と置くことで得られる。コーシー分布は期待値、分散、モーメント母関数が定義されない特殊な確率分布である(特性関数は存在する)。期待値が定義されないため、期待値の存在を前提とする大数の法則も、中心極限定理も成立しない。正規分布の形と似ているが、確率変数の値が大きくなっても小さくなくても正規分布のように指数的に確率密度の値が急激に減少しない。そこで、コーシー分布は裾が重い、裾が分厚い分布であると言われる。分布の名前はフランスの数学者 A.L.コーシーに因んで付けられている。

■確率密度関数

コーシー分布の確率密度関数は次の式で表される。期待値と分散が存在しないことを(付録45)に示す。

一般形 : $f(x) = (\gamma / \pi) / ((x - x_0)^2 + \gamma^2)$... (式51)

標準形 : $f(x) = (1 / \pi) / (x^2 + 1)$... (式52)

x : コーシー分布の確率変数 x_0 : 最頻値を与える位置母数

γ : 分布の広がり の程度を表す尺度母数(分布の最大値の 1/2 の値を示す幅の半分)

(計算例)

$x_0 = 0, \gamma = 1$ で、 $x = 1$ のときの確率密度 $f(1) = (1 / \pi) / ((1 - 0)^2 + 1^2) = 1 / (2 \pi) = 0.159$

コーシー分布を(図32)に示す。

(導出)

標準コーシー分布は、t 分布の(式43)において、自由度 $n = 1$ と置くことで得られる。

$\Gamma(1) = 1, \Gamma(1/2) = \pi^{1/2}$ (付録12) の関係を使って、

$$f(x) = \{ \Gamma((1+1)/2) / ((1 \cdot \pi)^{1/2} \Gamma(1/2)) \} (1 + x^2/1)^{-(1+1)/2} = \{ \Gamma(1) / ((\pi)^{1/2} \cdot \Gamma(1/2)) \} (1 + x^2)^{-1} = (1 / \pi) / (x^2 + 1)$$

一般形のコーシー分布は、標準コーシー分布を変数変換して得られる。

標準コーシー分布を、 $g(u) = (1 / \pi) / (u^2 + 1)$ と書き直して、確率変数 U と V の 2 変数を、(式53)と(式54)で示す確率変数 X と Y へ変数変換を行う。

(式54)は 2 変数変換のためのダミーの確率変数である。

$x = \gamma u + x_0$... (式53) $y = v$... (式54)

ここで、変数変換のためのヤコビアン $J(x, y)$ を求める。

(式53)と(式54)から、 $u(x, y) = (x - x_0) / \gamma$... (式55) $v(x, y) = y$... (式56)

ここで、ヤコビアン $J(x, y)$ を使用して、 $g(u, v)$ から $f(x, y)$ を求める。

$$J(x, y) = \partial(u, v) / \partial(x, y) = \begin{vmatrix} \partial u(x, y) / \partial x & \partial u(x, y) / \partial y \\ \partial v(x, y) / \partial x & \partial v(x, y) / \partial y \end{vmatrix} = \begin{vmatrix} 1/\gamma & 0 \\ 0 & 1 \end{vmatrix} = 1/\gamma$$

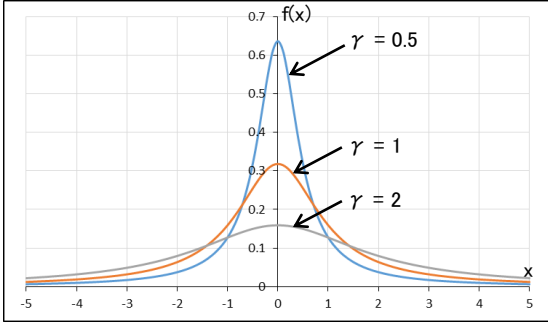
よって、 $f(x, y) = g(u, v) |J(x, y)| = (1 / \pi) / ((x - x_0)^2 / \gamma^2 + 1) (1 / \gamma) = (1 / \pi) / \{ ((x - x_0)^2 + \gamma^2) / \gamma^2 \} \gamma = (\gamma / \pi) / ((x - x_0)^2 + \gamma^2)$

次いで、 $f(x, y)$ を変数 y で積分して $f(x)$ を求める(周辺分布の確率密度関数を求める、周辺化するという)。

しかし、この式は、 $f(x, y)$ は y に依存しないことを示している。つまり、 $f(x) = f(x, y)$ である。よって、

$f(x) = (\gamma / \pi) / ((x - x_0)^2 + \gamma^2)$

となり、コーシー分布の一般形の確率密度関数が得られた。



(図32)コーシー分布

27. 大数(たいすう)の弱法則(Weak Law of Large Numbers)

大数の法則は、母集団からの無作為抽出の試行を無限に繰り返せば、母集団がどのような確率分布であっても、標本平均は母平均に近づくという法則である。大数の法則のイメージを(図33)に示す。大数の法則には、大数の弱法則と大数の強法則の2種類があるが、実用上は同一視でき、母平均を知りたいければ、試行回数を増やしたときの標本平均で推測できる。ただし、期待値が定義できないコーシー分布は対象外になる。

(大数の弱法則)

確率変数 $X_1, X_2, \dots$ が独立であり、母平均 μ と母分散 σ^2 が存在する場合、任意の正数 ε に対して、下式が成り立つ。

$$\lim_{n \rightarrow \infty} P(|(X_1 + X_2 + \dots + X_n)/n - \mu| < \varepsilon) = 1 \quad \dots \text{(式57)}$$

この法則は、同一の確率分布に従う独立な確率変数の標本平均 $\bar{X} = (X_1 + X_2 + \dots + X_n)/n$ と母平均 μ とのズレが、一定の値(ε)より小さい確率は、試行数 n が十分に大きければ1になること、逆に言えばズレが一定の値を超える確率は0であるということを意味する(確率収束するという)。試行回数を多くすれば、つまり、標本平均は母集団平均に近づく。

(大数の強法則)

確率変数 $X_1, X_2, \dots$ が独立であり、母平均 μ が存在するとき、下式が成り立つ。

$$P(\lim_{n \rightarrow \infty} (X_1 + X_2 + \dots + X_n)/n = \mu) = 1 \quad \dots \text{(式58)}$$

大数の強法則は、大数の弱法則より限定された強い条件下で成立する法則である。試行数 n が十分に大きければ、同一の確率分布に従う独立な確率数の標本平均 $\bar{X} = (X_1 + X_2 + \dots + X_n)/n$ と母平均 μ が一致する確率が1に近づくことを意味する(概収束するという)。

(大数の弱法則の確認)

標本平均 $\bar{X}$ の期待値 $E[\bar{X}] = E[(X_1 + X_2 + \dots + X_n)/n] = (E[X_1] + E[X_2] + \dots + E[X_n])/n = (\sum_{i=1}^n \mu)/n = (n \mu)/n = \mu$

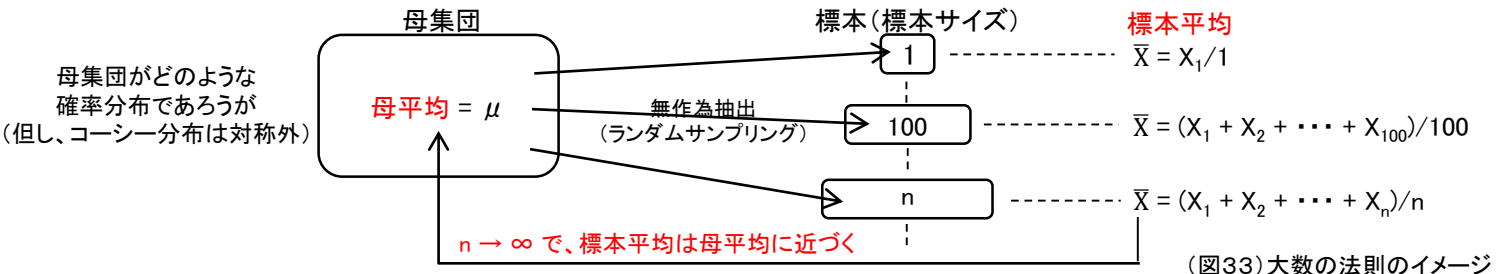
標本平均 $\bar{X}$ の分散 $V[\bar{X}] = V[(X_1 + X_2 + \dots + X_n)/n] = (V[X_1] + V[X_2] + \dots + V[X_n])/n^2 = (\sum_{i=1}^n \sigma^2)/n^2 = (n \sigma^2)/n^2 = \sigma^2/n$

ここで、(付録11)のチェビシェフの不等式(付式22)に上記の標本平均 $\bar{X}$ の期待値と分散を代入する。

$$P(|\bar{X} - \mu| \geq a) \leq (\sigma^2/n)/a^2$$

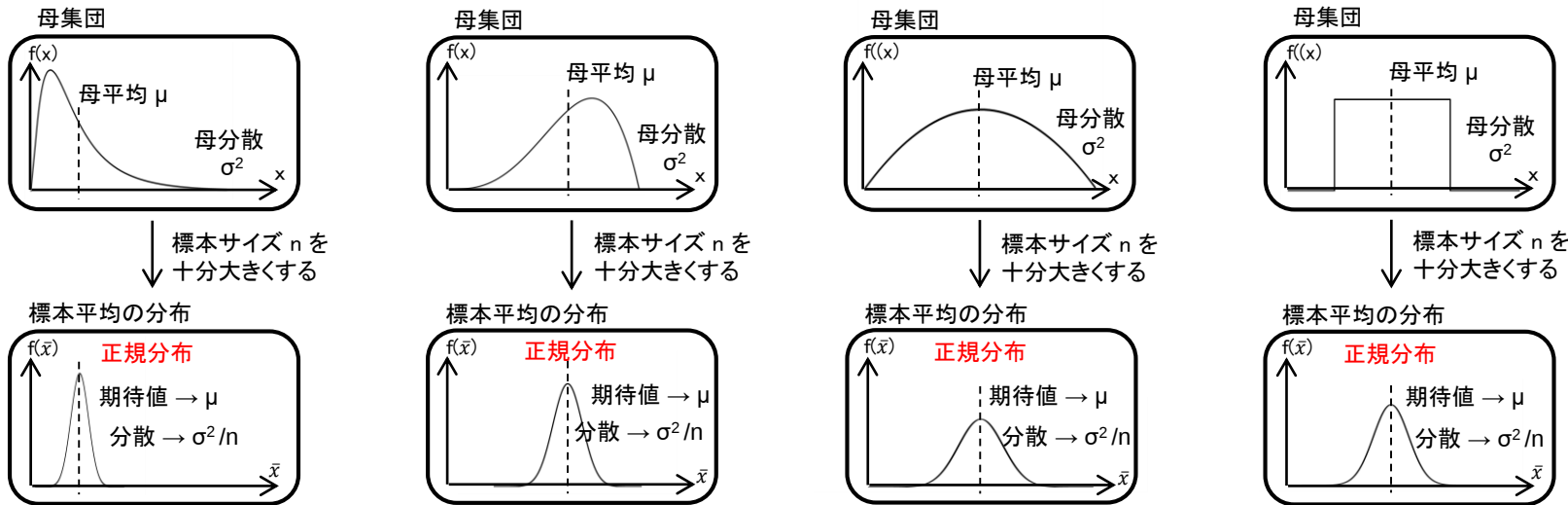
この式で、 $n \rightarrow \infty$ とすると、 $(\sigma^2/n)/a^2 \rightarrow 0$ となるので、 $\lim_{n \rightarrow \infty} P(|\bar{X} - \mu| \geq a) = 0$ が成立する。

領域 $|\bar{X} - \mu| \geq a$ の確率が0に近づくということは、これ以外の領域 $|\bar{X} - \mu| < a$ の確率が1に近づくことを意味する。つまり、 $\lim_{n \rightarrow \infty} P(|\bar{X} - \mu| < a) = 1$ が成立する。この式の a を ε に置き換えれば、(式57)を得る。



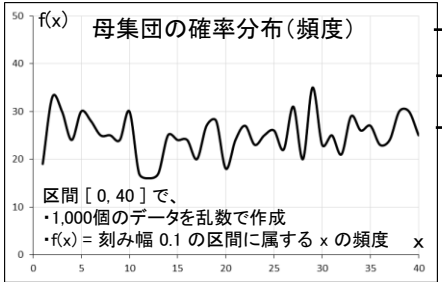
28. 中心極限定理(Central Limit Theorem)の概要

中心極限定理は、母集団からの無作為抽出の試行を十分に多く繰り返せば、母平均 μ と母分散 σ^2 が存在すれば、母集団がどのような確率分布であっても、標本平均の分布が期待値 μ 、分散 σ^2/n の正規分布に近づくという定理である(付録46)。中心極限定理のイメージを下図に示す。なお、標本サイズ n を無限大にすれば標本分散は 0 に近づく。これは大数の法則である。(図35)は簡単なシミュレーション結果である。



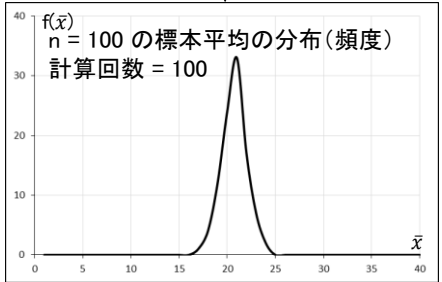
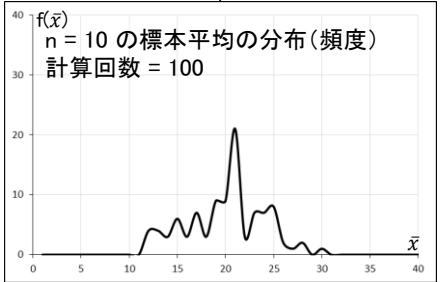
(図34) 中央極限定理のイメージ

母集団の確率分布は、母平均 μ と母分散 σ^2 さえ分かれば、どんな分布であってもよい
(よって、期待値が定義できないコーシー分布ではこの定理は成り立たない)
① 母集団から n 個の標本 $X_1, X_2, \dots, X_n$ を無作為抽出する
② 標本平均 $\bar{X} = (X_1 + X_2 + \dots + X_n)/n$ を求める。
③ ①と②を繰り返すことで、沢山の標本平均 $\bar{X}$ が計算され、確率変数 $\bar{X}$ としての確率分布が形成される
③ ここで、 n を十分大きくしたときの確率変数 $\bar{X}$ の確率分布が正規分布に近づく、というのが中央極限定理である。



標本平均の計算に使用する標本データの数 n を多くすると、標本平均の確率分布は正規分布に近づく

- ③ 標本サイズ n のとき、標本平均 $\bar{X} = (X_1 + X_2 + \dots + X_n)/n$
- ② 標本サイズ $n = 100$ のとき、標本平均 $\bar{X} = (X_1 + X_2 + \dots + X_{100})/100$
- ① 標本サイズ $n = 10$ のとき、標本平均 $\bar{X} = (X_1 + X_2 + \dots + X_{10})/10$



n を十分大きくすれば標本平均の確率分布は正規分布に近づく。
 $n = 100$ (左図)でも、正規分布に近づいているのが分かる

(図35) 中央極限定理のシミュレーション

付録1. 二項定理 (Binomial theorem)

・二項定理の公式

$$(x + y)^n = \sum_{k=0}^n {}_nC_k x^{n-k} y^k = {}_nC_0 x^n y^0 + {}_nC_1 x^{n-1} y^1 + {}_nC_2 x^{n-2} y^2 + \cdots + {}_nC_n x^0 y^n \quad \cdots \text{ (付式1)}$$

$$(x + y)^n = \sum_{k=0}^n {}_nC_k x^k y^{n-k} = {}_nC_0 x^0 y^n + {}_nC_1 x^1 y^{n-1} + {}_nC_2 x^2 y^{n-2} + \cdots + {}_nC_n x^n y^0$$

但し、 k は整数 $(0, 1, 2, \cdots)$, n は自然数 $(1, 2, \cdots)$

・説明

$x^{n-k} y^k$ の項は、 x を $(n - k)$ 個、 y を k 個選びそれらを掛け合わせたものである。
その組み合わせの数は、 ${}_nC_k$ 通りになり、これが各項の係数になる。

・例

$$\begin{aligned} (x + y)^5 &= {}_5C_0 x^5 y^0 + {}_5C_1 x^4 y^1 + {}_5C_2 x^3 y^2 + {}_5C_3 x^2 y^3 + {}_5C_4 x^1 y^4 + {}_5C_5 x^0 y^5 \\ &= 5! / (5! \cdot 0!) x^5 + 5! / (4! \cdot 1!) x^4 y + 5! / (3! \cdot 2!) x^3 y^2 + 5! / (2! \cdot 3!) x^2 y^3 + 5! / (1! \cdot 4!) x y^4 + 5! / (0! \cdot 5!) y^5 \\ &= x^5 + 5 x^4 y + 10 x^3 y^2 + 10 x^2 y^3 + 5 x y^4 + y^5 \end{aligned}$$

付録2. 多項定理 (Multinomial theorem)

$$(p_1 + p_2 + \cdots + p_m)^n = \sum_{x_1 + x_2 + \cdots + x_m = n} \{ n! / (x_1! x_2! \cdots x_m!) \} p_1^{x_1} p_2^{x_2} \cdots p_m^{x_m} \quad \cdots \text{ (付式2)}$$

但し、 m, n は自然数 $(1, 2, \cdots)$, $x_1 + x_2 + \cdots + x_m = n$

また、和 $\sum$ は、 $x_1 + x_2 + \cdots + x_m = n$ を満たす、非負の整数の全ての組み合わせの和を示す。

つまり、 $(p_1 + p_2 + \cdots + p_m)^n$ を展開したときの一般項は、下式で表される。

$$\{ n! / (x_1! x_2! \cdots x_m!) \} p_1^{x_1} p_2^{x_2} \cdots p_m^{x_m} \quad \cdots \text{ (付式3)}$$

(確認)

帰納法を使って、(付式2)を確認する。

(1) $m = 1$ のとき、

(付式2)より、左辺 = p_1^n また、 $x_1 = n$ であるので、右辺 = $\{ n! / x_1! \} p_1^{x_1} = \{ n! / n! \} p_1^n = p_1^n$

よって、 $m = 1$ のとき、(付式2)は成立する。

(2) $m = t$ のとき、(付式2)が成立すると仮定して、 $m = t + 1$ のときも(付式2)が成立することを確認する。

$m = t + 1$ のとき、(付式2)の左辺は、

$$\text{左辺} = (p_1 + p_2 + \cdots + p_{t+1})^n = ((p_1 + p_2 + \cdots + p_t) + p_{t+1})^n$$

ここで、 $x_1 + x_2 + \cdots + x_t = k$ とおき、二項定理(付録1)を使用して、また、 $m = t$ で(付式2)が成立すると仮定しているので、

$$\begin{aligned} \text{左辺} &= \sum_{k=0}^n {}^nC_k (p_1 + p_2 + \cdots + p_t)^k p_{t+1}^{n-k} = \sum_{k=0}^n {}^nC_k \{ \sum_{x_1 + x_2 + \cdots + x_t = k} (k! / (x_1! x_2! \cdots x_t!)) p_1^{x_1} p_2^{x_2} \cdots p_t^{x_t} \} p_{t+1}^{n-k} \\ &= \sum_{k=0}^n \sum_{x_1 + x_2 + \cdots + x_t = k} (n! / ((n-k)! k!)) (k! / (x_1! x_2! \cdots x_t!)) p_1^{x_1} p_2^{x_2} \cdots p_t^{x_t} \} p_{t+1}^{n-k} \\ &= \sum_{k=0}^n \sum_{x_1 + x_2 + \cdots + x_t = k} \{ n! / ((n-k)! (x_1! x_2! \cdots x_t!)) \} p_1^{x_1} p_2^{x_2} \cdots p_t^{x_t} p_{t+1}^{n-k} \end{aligned}$$

また、 $m = t + 1$ のとき、 $x_1 + x_2 + \cdots + x_t + x_{t+1} = n$ であるので、 $x_{t+1} = n - (x_1 + x_2 + \cdots + x_t) = n - k$ が成り立つ。

$$\text{左辺} = \sum_{x_1 + x_2 + \cdots + x_t + x_{t+1} = n} (n! / (x_1! x_2! \cdots x_t! x_{t+1}!)) p_1^{x_1} p_2^{x_2} \cdots p_t^{x_t} p_{t+1}^{x_{t+1}} = \text{右辺}$$

よって、 $m = t + 1$ のときも、(付式2)は成立する。

(3) 従って、(付式2)は、全ての m において成立する。

なお、(付式2)の係数は、下記に示すように、まず n 個から x_1 個選び $({}_nC_{x_1})$ 、次いで残りの $n - x_1$ 個から x_2 個を選ぶ $({}_{n-x_1}C_{x_2})$ 、という操作を繰り返すことで得られる。この式変形では、 $n - x_1 - \cdots - x_{m-1} = x_m$ および、 ${}_mC_m = 1$ であることを使う。

$$\begin{aligned} &({}_nC_{x_1})({}_{n-x_1}C_{x_2}) \cdots ({}_{n-x_1-\cdots-x_{m-2}}C_{x_{m-1}})({}_{n-x_1-\cdots-x_{m-1}}C_{x_m}) \\ &= ({}_nC_{x_1})({}_{n-x_1}C_{x_2}) \cdots ({}_{n-x_1-\cdots-x_{m-2}}C_{x_{m-1}})({}_{x_m}C_{x_m}) \\ &= ({}_nC_{x_1})({}_{n-x_1}C_{x_2}) \cdots ({}_{n-x_1-\cdots-x_{m-2}}C_{x_{m-1}}) \\ &= \{ n! / ((n-x_1)! x_1!) \} \{ (n-x_1)! / ((n-x_1-x_2)! x_2!) \} \cdots \{ (n-x_1-\cdots-x_{m-2})! / ((n-x_1-\cdots-x_{m-2}-x_{m-1})! x_{m-1}!) \} \\ &= \{ n! / ((n-x_1)! x_1!) \} \{ (n-x_1)! / ((n-x_1-x_2)! x_2!) \} \cdots \{ (n-x_1-\cdots-x_{m-2})! / (x_m! x_{m-1}!) \} \\ &= n! / (x_1! x_2! \cdots x_{m-1}! x_m!) \end{aligned}$$

付録3. テイラー展開(Taylor expansion) と マクローリン展開(McLaughlin expansion)

■テイラー展開

関数を(付式4)のように級数展開することを、テイラー展開という。

$$\begin{aligned} f(x) &= k_0 + k_1 (x - a) + k_2 (x - a)^2 + \cdots + k_n (x - a)^n + \cdots \\ &= f(a) + (f'(a)/1!) (x - a) + (f''(a)/2!) (x - a)^2 + (f'''(a)/3!) (x - a)^3 + \cdots + (f^{(n)}(a)/n!) (x - a)^n + \cdots \\ &= \sum_{n=0}^{\infty} (f^{(n)}(a)/n!) (x - a)^n \quad \cdots \text{ (付式4)} \end{aligned}$$

但し、 $a, k_0, k_1, \cdots, k_n$ は定数、また、 $f(x)$ は $x = a$ で微分可能

なお、' は 1 階微分、'' は 2 階微分、''' は 3 階微分を表す。また、 $f^{(n)}(x)$ は関数 $f(x)$ の n 階微分を表す。

■マクローリン展開

テイラー展開において、 $a = 0$ とおいた式が、マクローリン展開になる。

$$\begin{aligned} f(x) &= k_0 + k_1 (x - 0) + k_2 (x - 0)^2 + \cdots + k_n (x - 0)^n + \cdots \\ &= f(0) + (f'(0)/1!) (x - 0) + (f''(0)/2!) (x - 0)^2 + (f'''(0)/3!) (x - 0)^3 + \cdots + (f^{(n)}(0)/n!) (x - 0)^n + \cdots \\ &= f(0) + (f'(0)/1!) x + (f''(0)/2!) x^2 + (f'''(0)/3!) x^3 + \cdots + (f^{(n)}(0)/n!) x^n + \cdots \\ &= \sum_{n=0}^{\infty} (f^{(n)}(0)/n!) x^n \quad \cdots \text{ (付式5)} \quad \text{但し、} f(x) \text{ は、} x = 0 \text{ で微分可能} \end{aligned}$$

(例)

$$(1) f(x) = 1/(1 - x)$$

$$f'(x) = 1/(1 - x)^2, f''(x) = 2/(1 - x)^3, \cdots, f^{(n)}(x) = n/(1 - x)^{n+1}, \cdots$$

従って、

$$\begin{aligned} f(x) &= 1/(1 - 0) + (1/(1 - 0)^2/1!) x + (2/(1 - 0)^3/2!) x^2 + \cdots + (n/(1 - 0)^{n+1}/n!) x^n + \cdots \\ &= 1 + x + x^2 + \cdots + x^n + \cdots \end{aligned}$$

$$(2) f(x) = e^x$$

$$f'(x) = f''(x) = \cdots = f^{(n)}(x) = \cdots = e^x$$

従って、

$$f(x) = e^0 + (e^0/1!) x + (e^0/2!) x^2 + \cdots + (e^0/n!) x^n + \cdots = 1 + x/1! + x^2/2! + \cdots + x^n/n! + \cdots$$

$$(3) f(x) = \sin x$$

$$f'(x) = \cos x, f''(x) = -\sin x, f'''(x) = -\cos x, f^{(4)}(x) = \sin x, \cdots$$

従って、

$$\begin{aligned} f(x) &= \sin 0 + ((\cos 0)/1!) x + ((-\sin 0)/2!) x^2 + ((-\cos 0)/3!) x^3 + ((\sin 0)/4!) x^4 + \cdots \\ &= 0 + x/1! + 0 - x^3/3! + 0 + \cdots \\ &= x/1! - x^3/3! + x^5/5! - \cdots + (-1)^n x^{2n+1}/(2n+1)! + \cdots \end{aligned}$$

$$(4) f(x) = \log(1 + x)$$

$$f'(x) = 1/(1 + x), f''(x) = -1/(1 + x)^2, f'''(x) = 2/(1 + x)^3, f^{(4)}(x) = -3/(1 + x)^4, \cdots$$

従って、

$$\begin{aligned} f(x) &= \log(1 + 0) + 1(1 + 0)/1! x - 1/(1 + 0)^2/2! x^2 + 2/(1 + 0)^3/3! x^3 - 3/(1 + 0)^4/4! x^4 + \cdots \\ &= x - x^2/2 + x^3/3 - x^4/4 + \cdots \end{aligned}$$

付録4. ネイピア数 (Napier's constant)

ネイピア数は数学定数の一つであり、自然対数の底である。記号としては e が用いられる ($e = 2.718281828\cdots$)。ネイピア数の定義式はいくつかある。

・定義式

$$e = \lim_{n \rightarrow \infty} (1 + 1/n)^n \quad \cdots \text{ (付式6)}$$

ここで、 $n = 1/h$ とおくと、 $n \rightarrow \infty$ で $h \rightarrow 0$ となるので、この定義式は、 $e = \lim_{h \rightarrow 0} (1 + h)^{1/h} \quad \cdots \text{ (付式7)}$ と表すこともできる。

・ネイピア数の値を求める。

定義式を二項定理(付録1)を使って多項式で表す。

$$\begin{aligned} e &= \lim_{n \rightarrow \infty} (1 + 1/n)^n = \lim_{n \rightarrow \infty} \{ 1 + {}_nC_1 \cdot 1^{n-1} \cdot (1/n)^1 + {}_nC_2 \cdot 1^{n-2} \cdot (1/n)^2 + {}_nC_3 \cdot 1^{n-3} \cdot (1/n)^3 + \cdots \} \\ &= \lim_{n \rightarrow \infty} \{ 1 + n(1/n) + n(n-1)/2! (1/n)^2 + n(n-1)(n-2)/3! (1/n)^3 + \cdots \} \\ &= \lim_{n \rightarrow \infty} \{ 1 + 1 + (1 - 1/n)/2! + (1 - 1/n)(1 - 2/n)/3! + \cdots \} = 1 + 1 + 1/2! + 1/3! + \cdots = 2.718281828\cdots \end{aligned}$$

・ $y = a^x$ という指数関数の $x = 0$ のときの微分係数が 1 になるとき、 y の微分は a^x となり、微分しても式は変わらない。そして、このような a はネイピア数 e になる。 $y = e^x$ の微分は、 $dy/dx = e^x$ であることを、以下で確認する。

$$y = a^x, \quad dy/dx = \lim_{h \rightarrow 0} (a^{x+h} - a^x)/h = \lim_{h \rightarrow 0} a^x (a^h - 1)/h = a^x \lim_{h \rightarrow 0} (a^h - 1)/h \quad \cdots \text{ ①}$$

ここで、 $x = 0$ のとき、 $y = a^x$ の微分係数が 1 であれば、 $dy/dx = a^0 \lim_{h \rightarrow 0} (a^h - 1)/h = \lim_{h \rightarrow 0} (a^h - 1)/h = 1$ なので、①式は $dy/dx = a^x$ となる。

つまりこのような条件では、 $y = a^x$ は微分しても式は変わらないことが分かる。このような a の値を求める。

$y = a^x$ の逆関数 $y = \log_a x$ を微分すると、

$$dy/dx = \lim_{h \rightarrow 0} (\log_a (x+h) - \log_a x)/h = \lim_{h \rightarrow 0} (1/h) \log_a ((x+h)/x) = \lim_{h \rightarrow 0} (1/h) \log_a (1 + h/x)$$

逆関数 $y = \log_a x$ も、 $y = 0$ (つまり、 $x = 1$) のときの接線の傾きが 1 になるので、 $dy/dx = \lim_{h \rightarrow 0} (1/h) \log_a (1 + h) = 1$

ここで、 $(1/h) \log_a (1 + h) = \log_a (1 + h)^{1/h}$ であるので、 $\lim_{h \rightarrow 0} \log_a (1 + h)^{1/h} = 1$ よって、 $a = \lim_{h \rightarrow 0} (1 + h)^{1/h}$

従って、(付式7)より、 $a = \lim_{h \rightarrow 0} (1 + h)^{1/h} = e$ (ネイピア数)となる。

(注)逆関数 (inverse function)

関数 $y = f(x)$ に対して、 y の領域に含まれる任意の値に対して x の値がただ一つに決まるとき、 x は y の関数になり、これを $y = f^{-1}(x)$ で表す。この関数を $y = f(x)$ の逆関数であるという。直線 $y = x$ に対して対称になる。なお、指数関数の逆関数は対数関数である。

$$\lim_{n \rightarrow \infty} (1 + x/n)^n = e^x$$

$$\begin{aligned} \lim_{n \rightarrow \infty} (1 + x/n)^n &= \lim_{n \rightarrow \infty} \{ 1 + {}_nC_1 \cdot 1^{n-1} \cdot (x/n)^1 + {}_nC_2 \cdot 1^{n-2} \cdot (x/n)^2 + {}_nC_3 \cdot 1^{n-3} \cdot (x/n)^3 + \cdots \} \\ &= \lim_{n \rightarrow \infty} \{ 1 + n(x/n) + n(n-1)/2! (x/n)^2 + n(n-1)(n-2)/3! (x/n)^3 + \cdots \} \\ &= \lim_{n \rightarrow \infty} \{ 1 + x + (1 - 1/n)x^2/2! + (1 - 1/n)(1 - 2/n)x^3/3! + \cdots \} = 1 + x + x^2/2! + x^3/3! + \cdots \end{aligned}$$

(付録3)マクローリン展開の(例)より、右辺の級数展開は e^x に等しい、よって、 $\lim_{n \rightarrow \infty} (1 + x/n)^n = e^x \quad \cdots \text{ (付式8)}$

また、 $\lim_{n \rightarrow \infty} (1 + x/n)^n = \lim_{n \rightarrow \infty} ((1 + x/n)^{n/x})^x = (\lim_{n \rightarrow \infty} (1 + x/n)^{n/x})^x = e^x$ から導くこともできる。

付録5. ロピタルの定理(I' Hospital' s Rule)

ロピタルの定理とは、 $0/0$ とか $\pm\infty/\pm\infty$ といった不定形の極限を、微分を使って計算できる可能性があるという定理である。

・ロピタルの定理

ロピタルの定理が使用できる条件

- ① $-\infty \leq a \leq \infty$ であるとし、関数 $f(x)$, $g(x)$ が、 $x = a$ を含む区間(区間 I とする)で微分可能な関数である。
- ② $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$ または、 $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = \pm\infty$ である。つまり、 $\lim_{x \rightarrow a} f(x)/g(x)$ が不定形である。
- ③ 区間 I から a を除くすべての点で、 $g'(x) \neq 0$ である。(注)' は 1 階微分を表す。
- ④ 極限 $\lim_{x \rightarrow a} f'(x)/g'(x) = A$ が存在する。

これらの条件がすべて満たされたときに限って、ロピタルの定理(付式9)が成立する。

$$\lim_{x \rightarrow a} f(x)/g(x) = \lim_{x \rightarrow a} f'(x)/g'(x) = A \quad \cdots \text{(付式9)}$$

・ロピタルの定理が使える例

$f(x) = 1 - \cos x$, $g(x) = x$ のとき、 $\lim_{x \rightarrow 0} f(x)/g(x) = \lim_{x \rightarrow 0} (1 - \cos x)/x$ を求める。

使用条件の確認

- ① $x = 0$ を含む全領域で、 $f(x)$, $g(x)$ とも微分可能である。条件 ① を満たす。
- ② $\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} (1 - \cos x) = 0$, $\lim_{x \rightarrow 0} g(x) = \lim_{x \rightarrow 0} x = 0$ であり、 $\lim_{x \rightarrow 0} f(x)/g(x) = 0/0$ の不定形である。条件 ② を満たす。
- ③ $g'(x) = 1$ であるので、全区間において $g'(x) \neq 0$ である。条件 ③ を満たす。
- ④ $\lim_{x \rightarrow 0} f'(x)/g'(x) = \lim_{x \rightarrow 0} (1 - \cos x)' / (x)' = \lim_{x \rightarrow 0} \sin x / 1 = 0$ 収束する。条件 ④ を満たす。

よって、 $\lim_{x \rightarrow 0} (1 - \cos x)/x = 0$

・ロピタルの定理が使えない例

$f(x) = x + \sin x$, $g(x) = x$ のとき、 $\lim_{x \rightarrow \infty} f(x)/g(x) = \lim_{x \rightarrow \infty} (x + \sin x)/x$ を求める。

使用条件の確認

- ① $x = 0$ を含む全領域で、 $f(x)$, $g(x)$ とも微分可能である。条件 ① を満たす。
- ② $\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} (x + \sin x) = \infty$, $\lim_{x \rightarrow \infty} g(x) = \lim_{x \rightarrow \infty} x = \infty$ であり、 $\lim_{x \rightarrow \infty} f(x)/g(x) = \infty/\infty$ の不定形である。条件 ② を満たす。
- ③ $g'(x) = 1$ であるので、全区間において $g'(x) \neq 0$ である。条件 ③ を満たす。
- ④ しかし、 $\lim_{x \rightarrow \infty} f'(x)/g'(x) = \lim_{x \rightarrow \infty} (x + \sin x)' / (x)' = \lim_{x \rightarrow \infty} (1 + \cos x) / 1$ これは、振動して収束しない。条件 ④ を満たさない。

よって、 $\lim_{x \rightarrow \infty} f'(x)/g'(x)$ は存在しないので、ロピタルの定理は使用できない。

付録6. 一階線形微分方程式の解法

(付式10)に示す一階線形微分方程式の一般解(積分定数が残る解)を求める方法としては、定数変化法と積分因子法が知られている。

$$dy/dx + P(x)y = Q(x) \quad \cdots \text{(付式10)}$$

例として、一階線形微分方程式 $dy/dx + \lambda y = \lambda$ (λ は定数) $\cdots$ (付式11) の一般解を求める。(付式11)を満足する y を求める。

(1) 定数変化法

この方法は初めに、 $dy/dx + P(x)y = 0$ を解き、次いで、この解の不定積分の定数に関数を導入してもとの微分方程式の一般解を求める方法である。

まず、 $dy/dx + \lambda y = 0 \quad \cdots \text{①}$ を解く。

式①より、 $(1/y) dy = -\lambda dx$ であるが、両辺を積分すると、 $\int (1/y) dy = -\lambda \int dx \quad \cdots \text{②}$ になる。

式②より、 $\log |y| = -\lambda x + C_1$ (C_1 は積分定数) $\cdots \text{③}$ が求まる。

式③より、 $y = \pm e^{C_1} e^{-\lambda x}$ となる。この式の定数項をまとめて、 $C_2 = \pm e^{C_1}$ とおいて、式①の一般解である式④が求まる。

$$y = C_2 e^{-\lambda x} \quad (C_2 \text{は積分定数}) \quad \cdots \text{④}$$

ここで、 $C_2 = u(x)$ とおき、これを式④に代入した $y = u(x) e^{-\lambda x} \quad \cdots \text{⑤}$ が(付式11)の一般解であると予想して(付式11)に代入する。

$$\text{つまり、} \{u(x) e^{-\lambda x}\}' + \lambda \{u(x) e^{-\lambda x}\} = \lambda$$

この式の第1項を微分して、 $\{u'(x) e^{-\lambda x} - \lambda u(x) e^{-\lambda x}\} + \lambda \{u(x) e^{-\lambda x}\} = \lambda$ を得る。

この式を整理すると、 $u'(x) e^{-\lambda x} = \lambda$ つまり、 $du(x)/dx e^{-\lambda x} = \lambda \quad \cdots \text{⑥}$ が求まる。

この微分方程式⑥を解く。

式⑥から、 $du(x) = \lambda e^{\lambda x} dx$ となるが、式⑥は変数分離形の微分方程式である。

そこで、両辺を積分すると、 $\int du(x) = \int \lambda e^{\lambda x} dx$ となり、従って、 $u(x) = \lambda (e^{\lambda x} / \lambda) + C = e^{\lambda x} + C$ (C は積分定数) が求まる。

この $u(x)$ を式⑤に代入すれば、(付式11)の一般解である(付式12)が求まる。

$$y = u(x) e^{-\lambda x} = (e^{\lambda x} + C) e^{-\lambda x} = 1 + C e^{-\lambda x} \quad \cdots \text{(付式12)}$$

(2) 積分因子法

(付式10)の $P(x)$ を使った $e^{\int P(x) dx}$ を(付式10)の積分因子という。この積分因子を使用すると、微分方程式が簡素化される。

(付式11)より、 $P(x) = \lambda$ である。よって、積分因子は、 $e^{\int P(x) dx} = e^{\int \lambda dx} = e^{\lambda x + C_1} = C_2 e^{\lambda x}$ (C_1, C_2 は積分定数) $\cdots \text{⑦}$ となる。

この積分因子を(付式11)の両辺にかける。

$$dy/dx C_2 e^{\lambda x} + \lambda y C_2 e^{\lambda x} = \lambda C_2 e^{\lambda x}$$

$$\text{積分定数 } C_2 \text{ を消去し、簡略化して、} dy/dx e^{\lambda x} + \lambda y e^{\lambda x} = \lambda e^{\lambda x} \quad \cdots \text{⑧}$$

ところで、 $\{y e^{\lambda x}\}' = (dy/dx) e^{\lambda x} + y \lambda e^{\lambda x}$ であり、これは式⑧の左辺そのものである(これが、積分因子法の狙いである)。

$$\text{よって、} \{y e^{\lambda x}\}' = \lambda e^{\lambda x} \quad \cdots \text{⑨} \text{ が成り立つ。}$$

この式⑨を積分すると、 $y e^{\lambda x} = \lambda \int e^{\lambda x} dx = e^{\lambda x} + C$ (C は積分定数) が求まる。

$$\text{よって、} y \text{ の一般解 } y = (e^{\lambda x} + C) / e^{\lambda x} = 1 + C e^{-\lambda x} \quad \cdots \text{(付式13) が求まる。}$$

付録7. ウォリスの公式(Wallis formulas)

ウォリスの公式は無限積の公式である。ウォリスの公式を(付式13)に示す。
また、この式は二重階乗(値を1つ飛ばしにした階乗、記号は!! 例えば $7!! = 7 \cdot 5 \cdot 3 \cdot 1 = 105$)を使用して表すことができる(付式14)。

$$\prod_{n=1}^{\infty} (2n)^2 / \{ (2n-1)(2n+1) \} = \{ (2 \cdot 2)(4 \cdot 4) \cdots \} / \{ (1 \cdot 3)(3 \cdot 5) \cdots \} = \pi/2 \cdots \text{(付式13)}$$

$$\lim_{n \rightarrow \infty} \{ (2n)!! \}^2 / \{ (2n-1)!! (2n+1)!! \} = \pi/2 \cdots \text{(付式14)}$$

(参考)部分積分(f' は1階微分を表す)
$$\int_a^b f(x) g'(x) dx = [f(x) g(x)]_a^b - \int_a^b f'(x) g(x) dx$$

(確認)

$$I_n = \int_0^{\pi/2} \sin^n x dx \cdots \text{① とおく。部分積分(参考)を使って、}$$

$$\begin{aligned} I_n &= \int_0^{\pi/2} (\sin x) (\sin^{n-1} x) dx = \int_0^{\pi/2} (-\cos x)' (\sin^{n-1} x) dx = [(-\cos x) \sin^{n-1} x]_0^{\pi/2} - \int_0^{\pi/2} (-\cos x) (n-1) (\sin^{n-2} x) (-\cos x) dx \\ &= 0 + (n-1) \int_0^{\pi/2} (\cos^2 x) (\sin^{n-2} x) dx = (n-1) \int_0^{\pi/2} (1 - \sin^2 x) (\sin^{n-2} x) dx = (n-1) \int_0^{\pi/2} \sin^{n-2} x dx - (n-1) \int_0^{\pi/2} \sin^n x dx = (n-1) I_{n-2} - (n-1) I_n \end{aligned}$$

よって、 $I_n = \{ (n-1)/n \} I_{n-2}$ (漸化式が求まる)

(1) n が奇数のとき

$$\begin{aligned} I_n &= \{ (n-1)/n \} I_{n-2} = \{ (n-1)/n \} \{ (n-2-1)/(n-2) \} I_{n-2-2} = \{ (n-1)/n \} \{ (n-3)/(n-2) \} I_{n-4} \\ &= \{ (n-1)/n \} \{ (n-3)/(n-2) \} \cdots \{ 2/3 \} I_1 = (n-1)!! / n!! \quad (\text{注、} I_1 = \int_0^{\pi/2} \sin^1 x dx = [-\cos x]_0^{\pi/2} = 1) \end{aligned}$$

(2) n が偶数のとき

$$I_n = \{ (n-1)/n \} I_{n-2} = \{ (n-1)/n \} \{ (n-3)/(n-2) \} \cdots \{ 1/2 \} I_0 = (\pi/2) (n-1)!! / n!! \quad (\text{注、} I_0 = \int_0^{\pi/2} \sin^0 x dx = [x]_0^{\pi/2} = \pi/2)$$

以上の結果から、次の式が得られる。

$$I_{2n+2} = (\pi/2) (2n+2-1)!! / (2n+2)!! = (\pi/2) (2n+1)!! / (2n+2)!! \cdots \text{②}$$

$$I_{2n+1} = (2n+1-1)!! / (2n+1)!! = (2n)!! / (2n+1)!! \cdots \text{③}$$

$$I_{2n} = (\pi/2) (2n-1)!! / (2n)!! \cdots \text{④}$$

ところで、 $0 \leq x \leq \pi/2$ では、 $0 \leq \sin x \leq 1$ であるので、 $\sin^{2n+2} x \leq \sin^{2n+1} x \leq \sin^{2n} x$ となる。

つまり、①より、 $I_{2n+2} \leq I_{2n+1} \leq I_{2n}$ の関係が成立する。この関係式に、②、③、④を代入すると、

$$(\pi/2) (2n+1)!! / (2n+2)!! \leq (2n)!! / (2n+1)!! \leq (\pi/2) (2n-1)!! / (2n)!!$$

各辺に、 $(2n)!! / (2n-1)!!$ をかける。 $(2n+1)!! = (2n+1)(2n-1)!!$ 、 $(2n+2)!! = (2n+2)2n!!$ であることに注意して、

$$(\pi/2) (2n+1)(2n-1)!! / (2n+2)(2n)!! \leq (2n)!! / (2n+1)!! \leq (\pi/2) (2n-1)!! / (2n)!!$$

$$\{ (2n)!! / (2n-1)!! \} \{ (\pi/2) (2n+1) (2n-1)!! / (2n+2) (2n)!! \} \leq \{ (2n)!! / (2n-1)!! \} \{ (2n)!! / (2n+1)!! \} \leq \{ (2n)!! / (2n-1)!! \} \{ (\pi/2) (2n-1)!! / (2n)!! \}$$

$$(\pi/2) (2n+1) / (2n+2) \leq \{ (2n)!! \}^2 / \{ (2n-1)!! (2n+1)!! \} \leq \pi/2$$

ここで、 $n \rightarrow \infty$ とすると、 $\lim_{n \rightarrow \infty} (\pi/2) (2n+1) / (2n+2) = \pi/2$ となるので、 $\pi/2 \leq \lim_{n \rightarrow \infty} \{ (2n)!! \}^2 / \{ (2n-1)!! (2n+1)!! \} \leq \pi/2$

結局、不等式のはさみ打ちによって、次のウォリスの公式(付式14)が成立する。

$$\lim_{n \rightarrow \infty} \{ (2n)!! \}^2 / \{ (2n-1)!! (2n+1)!! \} = \pi/2$$

付録8. スターリングの公式(Stirling's approximation)

(付式15)のスターリングの公式(近似式)は、階乗 $n!$ を指数関数で近似する式である。スターリングの公式は、ウォリスの公式から導かれる。
 $n! \doteq (2\pi n)^{1/2} (n/e)^n \cdots$ (付式15)

(付式13)のウォリスの公式の左辺を変形する。まず、 $k = 1$ から n の範囲で変形し、そのあとで $n \rightarrow \infty$ へと展開する。

$$\begin{aligned} \prod_{k=1}^n (2k)^2 / \{ (2k-1)(2k+1) \} &= \{ (2 \cdot 2) (4 \cdot 4) \cdots ((2n)(2n)) \} / \{ (1 \cdot 3) (3 \cdot 5) (5 \cdot 7) \cdots ((2n-1)(2n+1)) \} \\ &= \{ (2 \cdot 4 \cdots (2n)) (2 \cdot 4 \cdots (2n)) \} / \{ (1 \cdot 3 \cdot 5 \cdots (2n-1)) (3 \cdot 5 \cdot 7 \cdots (2n+1)) \} \\ &= (2 \cdot 4 \cdots (2n))^2 \{ (2 \cdot 4 \cdots (2n)) / (1 \cdot 2 \cdot 3 \cdots (2n-1) \cdot (2n)) \} \{ (2 \cdot 4 \cdots (2n)) / (1 \cdot 2 \cdot 3 \cdots (2n-1) \cdot (2n) \cdot (2n+1)) \} \\ &= ((2 \cdot 1) (2 \cdot 2) (2 \cdot 3) \cdots (2n))^4 / \{ (1 \cdot 2 \cdot 3 \cdots (2n-1) \cdot (2n))^2 (2n+1) \} \\ &= 2^{4n} (n!)^4 / \{ ((2n)!)^2 (2n+1) \} = \{ (n!)^4 / ((2n)!)^2 \} \{ 2^{4n} / (2n+1) \} \end{aligned}$$

ウォリスの公式(付式14)は式①の通りであるが、

$$\lim_{n \rightarrow \infty} \{ (n!)^4 / ((2n)!)^2 \} \{ 2^{4n} / (2n+1) \} = \pi/2 \cdots \text{①}$$

ここで、

$$A = n! e^n / (n^n n^{1/2}) \cdots \text{②}$$

$$\text{とおくと、} A^4 = (n!)^4 e^{4n} / (n^{4n} n^2) \cdots \text{③}$$

$$\text{また、式②は、} A = (2n)! e^{2n} / ((2n)^{2n} (2n)^{1/2}) \cdots \text{④}$$

$$\text{と表すこともでき、} A^2 = ((2n)!)^2 e^{4n} / ((2n)^{4n} (2n)) \cdots \text{⑤}$$

ここで、③/⑤を計算すると、

$$\begin{aligned} A^2 &= A^4 / A^2 = \{ (n!)^4 e^{4n} / (n^{4n} n^2) \} / \{ ((2n)!)^2 e^{4n} / ((2n)^{4n} (2n)) \} \\ &= \{ (n!)^4 e^{4n} ((2n)^{4n} (2n)) \} / \{ (n^{4n} n^2) ((2n)!)^2 e^{4n} \} = \{ (n!)^4 e^{4n} 2^{4n+1} n^{4n+1} \} / \{ n^{4n+2} ((2n)!)^2 e^{4n} \} \\ &= \{ (n!)^4 / ((2n)!)^2 \} (2^{4n+1} / n) \cdots \text{⑥} \end{aligned}$$

$$\text{よって、} (n!)^4 / ((2n)!)^2 = A^2 n / 2^{4n+1}$$

この式を、式①の左辺に代入して式変形すると、

$$\lim_{n \rightarrow \infty} \{ (n!)^4 / ((2n)!)^2 \} \{ 2^{4n} / (2n+1) \} = \lim_{n \rightarrow \infty} \{ A^2 n / 2^{4n+1} \} \{ 2^{4n} / (2n+1) \} = A^2 \lim_{n \rightarrow \infty} n / \{ 2(2n+1) \} = A^2 \lim_{n \rightarrow \infty} 1 / \{ 2(2 + 1/n) \} = A^2/4$$

$$\text{よって、} A^2/4 = \pi/2$$

$A > 0$ であるので、 $A = (2\pi)^{1/2}$ となる。

$$\text{式②を使って、} \lim_{n \rightarrow \infty} A = \lim_{n \rightarrow \infty} n! e^n / (n^n n^{1/2}) = (2\pi)^{1/2} \text{ となる。そして、この式を変形すると、} \lim_{n \rightarrow \infty} n! / \{ (2\pi n)^{1/2} (n/e)^n \} = 1$$

従ってこの式から、スターリングの公式(近似式) $n! \doteq (2\pi n)^{1/2} (n/e)^n$ が求まる。

なお、(付式15)の対数を取ると、スターリングの公式(近似式)は(付式16)のように表すこともできる。

$$\log n! \doteq \log (2\pi n)^{1/2} + \log n^n + \log e^{-n}$$

$$\text{よって、} \log n! \doteq (1/2) \log (2\pi n) + n(\log n - 1) \cdots \text{(付式16)}$$

付録9. ガウス積分(Gaussian integral)

ガウス関数 $f(t) = e^{-at^2}$ の積分の演算結果を(付式17)に示す。この式はよく使用される。

$$\int_{-\infty}^{\infty} e^{-at^2} dt = \sqrt{\pi/a} \quad \dots \text{(付式17)}$$

ここで、 a は正の定数 $a > 0$

(確認) (付式17)の積分を計算する。

$$A = \int_{-\infty}^{\infty} e^{-at^2} dt \text{ とおくと、}$$

$$A^2 = \left\{ \int_{-\infty}^{\infty} e^{-ax^2} dx \right\} \left\{ \int_{-\infty}^{\infty} e^{-ay^2} dy \right\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-a(x^2+y^2)} dx dy \quad \dots \text{①}$$

この積分は、微小面積 $dS = dx dy \dots \text{②}$ での $e^{-(x^2+y^2)}$ の総和になる。

ここで、 $x = r \cos \theta$, $y = r \sin \theta$ において極座標変換する。

$$x^2 + y^2 = r^2 \quad \dots \text{③} \text{ が成立する。}$$

積分範囲は x, y とも $-\infty \sim \infty$ と全空間になるので、 θ は $0 \sim 2\pi$, r は $0 \sim \infty \quad \dots \text{④}$ になる。

また、極座標系での微小面積は、(付図1)に示す通り、

$$dS = \pi(r+dr)^2 (d\theta/(2\pi)) - \pi r^2 (d\theta/(2\pi)) = 1/2 ((r^2 + 2r dr + (dr)^2) - r^2) d\theta = (r dr + (1/2)(dr)^2) d\theta$$

ここで、 $(dr)^2$ の項は他の項に比べて極めて小さく無視できるので、 $dS = r dr d\theta \quad \dots \text{⑤}$ となる。

式②、③、④、⑤ を(付式18)に代入すると、

$$A^2 = \int_0^{2\pi} \int_0^{\infty} e^{-ar^2} r dr d\theta = \int_0^{2\pi} d\theta \int_0^{\infty} r e^{-ar^2} dr = 2\pi \left[\left\{ -1/(2a) \right\} e^{-ar^2} \right]_0^{\infty} = -\pi \left[(1/a) e^{-ar^2} \right]_0^{\infty} = -\pi/a (0 - 1) = \pi/a$$

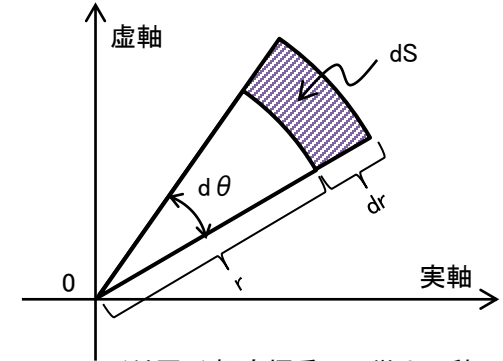
A は正の値をとるので、 $A = \sqrt{\pi/a}$ となり、(付式17)が求まる。

なお、次式の積分結果も、変数 t を b だけ平行移動したに過ぎないので、(付式17)と同じく、 $\sqrt{\pi/a}$ となる(置換積分でも確認できる)。

$$\int_{-\infty}^{\infty} e^{-a(t-b)^2} dt = \sqrt{\pi/a} \quad \dots \text{(付式18)}$$

また、(付式17)を a で微分すると、(付式19)が求まる。

$$\int_{-\infty}^{\infty} t^2 e^{-at^2} dt = 1/2 \sqrt{\pi/a^3} \quad \dots \text{(付式19)}$$



(付図1)極座標系での微小面積

付録10. マルコフの不等式(Markov's inequality)

マルコフの不等式は確率分布に関する不等式であり、確率分布の種類に関係なく成立する不等式である。

マルコフの不等式は、チェビシェフの不等式の導出に使用される。

・マルコフの不等式

X を非負の値 ($X > 0$) をとる確率変数であるとしたとき、任意の実数 c に対して、

$$P(X \geq c) \leq E[X]/c \quad \dots \text{ (付式20) } \quad \text{が成り立つ。}$$

(式の持つ意味)

確率変数 X がある任意の定数 c より大きい確率 $P(X \geq c)$ (は、 X の期待値(母平均)の $1/c$ 倍以下である、ということを表している。

(確認)

ここでは、 X が連続型の確率変数のときを確認する。確率変数 X が従う確率分布の確率密度関数を $f(x)$ とする。
期待値の定義をもとに、積分区間を 2 分割する。

$$E[X] = \int_0^{\infty} x f(x) dx = \int_0^c x f(x) dx + \int_c^{\infty} x f(x) dx$$

ここで、 $x > 0$ であることから、 $\int_0^c x f(x) dx \geq 0$

よって、 $\int_0^c x f(x) dx = E[X] - \int_c^{\infty} x f(x) dx \geq 0$ である。

従って、 $E[X] \geq \int_c^{\infty} x f(x) dx$ が成り立つ。

また、この不等式の右辺の非積分関数 $x f(x)$ の x の項を、積分範囲 $c \leq x \leq \infty$ の最小値 c に置き換えれば、

$$E[X] \geq c \int_c^{\infty} f(x) dx \quad \text{が成り立つ。}$$

この不等式の右辺は、確率密度関数を $f(x)$ を積分範囲 $c \leq x \leq \infty$ で積分したものであり、これは確率 $P(X \geq c)$ になる。 よって、

$$E[X] \geq c P(X \geq c)$$

この式を変形すれば、(付式20)のマルコフの不等式が得られる。

付録11. チェビシェフの不等式(Chebyshev's inequality)

チェビシェフの不等式は、マルコフの不等式と同じく確率分布に関する不等式であり、確率分布の種類に関係なく成立する不等式である。
チェビシェフの不等式は、大数の弱法則の導出に使用される。

・チェビシェフの不等式

$E[X] = \mu$, $V[X] = \sigma^2$ とするとき、任意の実数 $a > 0$ に対して、

$$P(|X - \mu| \geq a\sigma) \leq 1/a^2 \quad \dots \text{ (付式21) } \quad \text{が成り立つ。}$$

(式の持つ意味)

不等式の左辺は、偏差 $|X - \mu|$ が標準偏差 σ の定数倍 (a 倍) となる確率 P を表す。

一方、不等式の右辺は、定数 a が大きくなるほど、小さくなることを表している。

つまり、確率変数 X がとる値が期待値から離れるほど、確率変数 X がその値をとる確率は小さくなることを表している。

(確認)

マルコフの不等式(付式20)において、

$$X = (Y - \mu)^2 \geq 0, \quad c = a^2 \sigma^2 \text{ とおく。}$$

$$P((Y - \mu)^2 \geq a^2 \sigma^2) \leq E[(Y - \mu)^2] / (a^2 \sigma^2)$$

ここで、不等式の左辺は、 a と σ はいずれも非負であることを考慮して平方根の絶対値で置き換える。

また、不等式の右辺は、 $E[(Y - \mu)^2] = \sigma^2$ であることから、

$$P(|Y - \mu| \geq a\sigma) \leq \sigma^2 / (a^2 \sigma^2) = 1/a^2$$

となり、(付式21)のチェビシェフの不等式が得られる。

なお、(付式21)は下記のようにも表すことができる。

マルコフの不等式(付式20)において、

$$X = (Y - \mu)^2 \geq 0, \quad c = a^2 \text{ とおく。}$$

同様に、チェビシェフの不等式として、(付式22)を得ることができる。

$$P(|Y - \mu| \geq a) \leq \sigma^2 / a^2 \quad \dots \text{ (付式22) }$$

付録12. ガンマ関数(Γ関数)の定義と性質

ガンマ関数は、階乗を正の実数に拡張した関数である。(付図2)にガンマ関数のグラフを示す。

・定義式

$$s > 0 \text{ に対して、} \Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx \quad \cdots \text{ (付式23)}$$

(注)文字「Γ」は、ギリシャ文字「γ」の大文字

・性質

① $\Gamma(1) = 1 \quad \cdots \text{ (付式24)}$

$$\Gamma(1) = \int_0^{\infty} x^{1-1} e^{-x} dx = \int_0^{\infty} e^{-x} dx = [-e^{-x}]_0^{\infty} = -(0 - 1) = 1$$

② $\Gamma(s) = (s-1) \Gamma(s-1) \quad \cdots \text{ (付式25)}$

部分積分(参考)を使用して(ここで、' は1階微分を表す)、

$$\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx = \int_0^{\infty} x^{s-1} (-e^{-x})' dx = [x^{s-1} (-e^{-x})]_0^{\infty} - \int_0^{\infty} (x^{s-1})' (-e^{-x}) dx = 0 + (s-1) \int_0^{\infty} x^{s-2} e^{-x} dx = (s-1) \Gamma(s-1)$$

③ n が正の整数のとき、 $\Gamma(n) = (n-1)!$ $\cdots \text{ (付式26)}$

(付式25)より、 $\Gamma(n) = (n-1) \Gamma(n-1) = (n-1)(n-2) \Gamma(n-2) = \cdots = (n-1)(n-2) \cdots 1 \Gamma(1)$

ここで、(付式24)より、 $\Gamma(1) = 1$ であるので、

$$\Gamma(n) = (n-1)(n-2) \cdots 1 = (n-1)!$$

④ $\Gamma(1/2) = \sqrt{\pi} \quad \cdots \text{ (付式27)}$

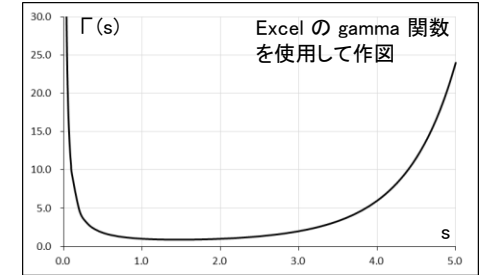
置換積分を行う。 $x = t^2$ と置くと、 $dx = 2t dt$ 、積分範囲は $x: 0 \rightarrow \infty$ に対して、 $t: 0 \rightarrow \infty$ である。

ガウス積分の公式(付録9)を使用する。また、非積分関数が偶関数であることを利用する。

$$\int_{-\infty}^{\infty} e^{-t^2} dt = 2 \int_0^{\infty} e^{-t^2} dt = \sqrt{\pi}$$

よって

$$\Gamma(1/2) = \int_0^{\infty} x^{-1/2} e^{-x} dx = \int_0^{\infty} t^{-1} e^{-t^2} 2t dt = 2 \int_0^{\infty} e^{-t^2} dt = 2(\sqrt{\pi}/2) = \sqrt{\pi}$$



(付図2)ガンマ関数

付録13. ベータ関数(B 関数)の定義と性質(1/2)

ベータ関数の定義式は以下の通りである。なお、この積分が収束することは確認されている。

・定義式

$$s, t > 0 \text{ に対して、} B(s, t) = \int_0^1 x^{s-1} (1-x)^{t-1} dx \quad \cdots \text{ (付式28)}$$

・性質

① $B(s, t) = B(t, s) \quad \cdots \text{ (付式29)}$

$u = 1 - x$ として、置換積分を行う。 $x = 1 - u$, $dx/du = -1$,
積分範囲は $x : 0 \rightarrow 1$ に対して、 $u : 1 \rightarrow 0$ である。

$$B(s, t) = \int_1^0 (1-u)^{s-1} u^{t-1} (-1) du = \int_0^1 u^{t-1} (1-u)^{s-1} du = B(t, s)$$

②-1 $B(s, t+1) = (t/s) B(s+1, t) \quad \cdots \text{ (付式30)}$

$$\begin{aligned} \text{部分積分を使用して、} B(s, t+1) &= \int_0^1 x^{s-1} (1-x)^t dx = \int_0^1 \left\{ (1/s) x^s \right\}' (1-x)^t dx = \left[(1/s) x^s (1-x)^t \right]_0^1 - \int_0^1 (1/s) x^s (-t(1-x)^{t-1}) dx \\ &= 0 + (t/s) \int_0^1 x^s (1-x)^{t-1} dx = (t/s) B(s+1, t) \end{aligned}$$

②-2 $B(s+1, t) = \{ s/(s+t) \} B(s, t) \quad \cdots \text{ (付式31)}$

$$\text{(付式30)より、} B(s+1, t) = (s/t) B(s, t+1) = (s/t) \int_0^1 x^{s-1} (1-x)^t dx$$

ここで、 $(1-x)^t = (1-x)^{t-1} (1-x)$ を代入する。

$$\begin{aligned} B(s+1, t) &= (s/t) \int_0^1 x^{s-1} (1-x)^t dx = (s/t) \int_0^1 x^{s-1} (1-x)^{t-1} (1-x) dx = (s/t) \left\{ \int_0^1 x^{s-1} (1-x)^{t-1} dx - \int_0^1 x^s (1-x)^{t-1} dx \right\} \\ &= (s/t) \{ B(s, t) - B(s+1, t) \} = (s/t) B(s, t) - (s/t) B(s+1, t) \end{aligned}$$

$$\text{よって、} (1 + s/t) B(s+1, t) = (s/t) B(s, t) \quad \text{従って、} B(s+1, t) = \{ (s/t)/(1 + s/t) \} B(s, t) = \{ s/(s+t) \} B(s, t)$$

②-3 $B(s, t+1) = \{ t/(s+t) \} B(s, t) \quad \cdots \text{ (付式32)}$

$$\text{(付式30)と(付式31)より、} B(s, t+1) = (t/s) B(s+1, t) = (t/s) \{ s/(s+t) \} B(s, t) = \{ t/(s+t) \} B(s, t)$$

③ $B(1, 1) = 1 \quad \cdots \text{ (付式33)}$

$$B(1, 1) = \int_0^1 x^{1-1} (1-x)^{1-1} dx = \int_0^1 dx = [x]_0^1 = 1$$

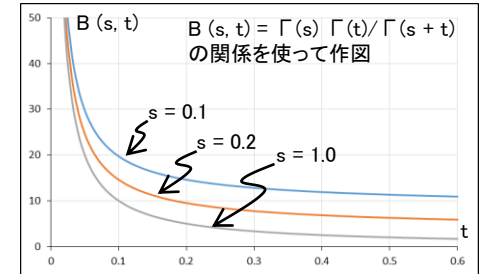
④ $B(1/2, 1/2) = \pi \quad \cdots \text{ (付式34)}$

$$B(1/2, 1/2) = \int_0^1 x^{1/2-1} (1-x)^{1/2-1} dx = \int_0^1 \{ x(1-x) \}^{-1/2} dx = \int_0^1 \{ -(x-1/2)^2 + 1/4 \}^{-1/2} dx$$

ここで、 $x = 1/2 + (1/2) \sin \theta$ として、置換積分を行う。

$dx/d\theta = (1/2) \cos \theta$, 積分範囲は $x : 0 \rightarrow 1$ に対して、 $\theta : -\pi/2 \rightarrow \pi/2$ である。

$$B(1/2, 1/2) = \int_{-\pi/2}^{\pi/2} \{ -(1/2 + (1/2) \sin \theta - 1/2)^2 + 1/4 \}^{-1/2} (1/2) \cos \theta d\theta = \int_{-\pi/2}^{\pi/2} \{ (1/4) (-\sin^2 \theta + 1) \}^{-1/2} (1/2) \cos \theta d\theta = \int_{-\pi/2}^{\pi/2} d\theta = \pi$$



(付図3) ベータ関数

付録13. ベータ関数(B 関数)の定義と性質(2/2)

・性質(つづき)

$$\textcircled{5} \quad B(s, t) = 2 \int_0^{\pi/2} \sin^{2s-1} \theta \cos^{2t-1} \theta d\theta \quad \cdots \text{ (付式35)}$$

$x = \sin^2 \theta$ として、置換積分を行う。 $dx/d\theta = 2 \sin \theta \cos \theta$, 積分範囲は $x : 0 \rightarrow 1$ に対して、 $\theta : 1 \rightarrow \pi/2$ である。

$$B(s, t) = \int_0^{\pi/2} (\sin^2 \theta)^{s-1} (1 - \sin^2 \theta)^{t-1} 2 \sin \theta \cos \theta d\theta = 2 \int_0^{\pi/2} \sin^{2s-1} \theta \cos^{2t-1} \theta d\theta$$

$$\textcircled{6} \quad B(s, t) = \Gamma(s) \Gamma(t) / \Gamma(s+t) \quad \cdots \text{ (付式36)}$$

ガンマ関数の定義から、 $\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx$

$x = p^2$ ($p > 0$) として、置換積分を行う。 $dx/dp = 2p$, 積分範囲は $x : 0 \rightarrow \infty$ に対して、 $p : 0 \rightarrow \infty$ である。

$$\Gamma(s) = \int_0^{\infty} (p^2)^{s-1} e^{-p^2} 2p dp \quad \text{同様に、} x = q^2 \text{ (} q > 0 \text{) として置換積分を行えば、} \Gamma(t) = \int_0^{\infty} (q^2)^{t-1} e^{-q^2} 2q dq$$

$$\text{よって、} \Gamma(s) \Gamma(t) = 4 \left\{ \int_0^{\infty} (p^2)^{s-1} e^{-p^2} p dp \right\} \left\{ \int_0^{\infty} (q^2)^{t-1} e^{-q^2} q dq \right\} = 4 \int_0^{\infty} \int_0^{\infty} p^{2s-1} q^{2t-1} e^{-(p^2+q^2)} dp dq$$

ここで、 $p = r \cos \theta$, $q = r \sin \theta$ と置いて、極座標変換を行う。

(付録16)の計算例に示すように、この変数変換(極座標変換)のヤコビアンは r なので、 $dp dq = r dr d\theta$ である。

また、積分範囲は、 $p : 0 \rightarrow \infty$, $q : 0 \rightarrow \infty$ に対して、 $r : 0 \rightarrow \infty$, $\theta : 0 \rightarrow \pi/2$ である。

よって、ガンマ関数の定義と(付式35)から、

$$\begin{aligned} \Gamma(s) \Gamma(t) &= 4 \int_0^{\pi/2} \int_0^{\infty} (r \cos \theta)^{2s-1} (r \sin \theta)^{2t-1} e^{-r^2} r dr d\theta = \left\{ 2 \int_0^{\infty} e^{-r^2} r^{2(s+t)-1} dr \right\} \left\{ 2 \int_0^{\pi/2} \sin^{2t-1} \theta \cos^{2s-1} \theta d\theta \right\} \\ &= \Gamma(s+t) B(t, s) = \Gamma(s+t) B(s, t) \end{aligned}$$

従って、 $B(s, t) = \Gamma(s) \Gamma(t) / \Gamma(s+t)$ となる。

$$\textcircled{7} \quad s, t \text{ が自然数、に対して、} B(s, t) = (s-1)!(t-1)! / (s+t-1)! \quad \cdots \text{ (付式37)}$$

$\Gamma(s+1) = s!$ であるので、

$$B(s, t) = \Gamma(s) \Gamma(t) / \Gamma(s+t) = (s-1)!(t-1)! / (s+t-1)!$$

$$\textcircled{8} \quad B(s, t) = \int_0^{\infty} x^{s-1} (1+x)^{-(s+t)} dx \quad \cdots \text{ (付式38)}$$

ベータ関数の定義式(付式28)で、 $x = p/(1+p)$ として、置換積分を行う。

$p = x/(1-x)$, $dx/dp = 1/(1+p) - p/(1+p)^2 = 1/(1+p)^2$, 積分範囲は $x : 0 \rightarrow 1$ に対して、 $p : 0 \rightarrow \infty$ である。

$$\begin{aligned} B(s, t) &= \int_0^1 x^{s-1} (1-x)^{t-1} dx = \int_0^{\infty} \left\{ p/(1+p) \right\}^{s-1} \left\{ 1 - p/(1+p) \right\}^{t-1} \left\{ 1/(1+p)^2 \right\} dp = \int_0^{\infty} p^{s-1} (1+p)^{-s+1} (1+p)^{-t+1} (1+p)^{-2} dp \\ &= \int_0^{\infty} p^{s-1} (1+p)^{-(s+t)} dp = \int_0^{\infty} x^{s-1} (1+x)^{-(s+t)} dx \end{aligned}$$

付録14. 確率変数(1変数)の変数変換

確率変数の変数変換とは、既知の確率変数から新しい確率変数を作り出すことをいう。
ここでは、1つの確率変数から新しい1つの確率変数を作る場合を考え、この新しい確率変数の確率分布(確率密度関数)を求める。
既知の確率変数を U とし、その確率密度関数を $g(u)$ とする。そして、新しい確率変数を X とし、その確率密度関数を $f(x)$ とする。確率変数 U を X に変換したとき、確率密度関数 $f(x)$ は u を使ってどう表されるか？ これが、確率変数(1変数)の変数変換になる。

この変数変換では変換の前後で確率の大きさが保存されることを利用する(付図4)。
確率密度関数で表した確率分布では、図で示すように確率変数の区間に相当する面積が確率を表し、この確率は変数変換の前後で変わらない。

つまり、 $g(u) \Delta u = f(x) \Delta x$ ($\Delta u > 0, \Delta x > 0$) が成り立つ。
ここで、 $\Delta u \rightarrow 0, \Delta x \rightarrow 0$ の極限を考える。

すると、 $g(u) du = f(x) dx$
よって、 $f(x) = g(u) du/dx$... (付式39) が成り立つ。
この(付式39)を使って、1変数の変数変換を行う。

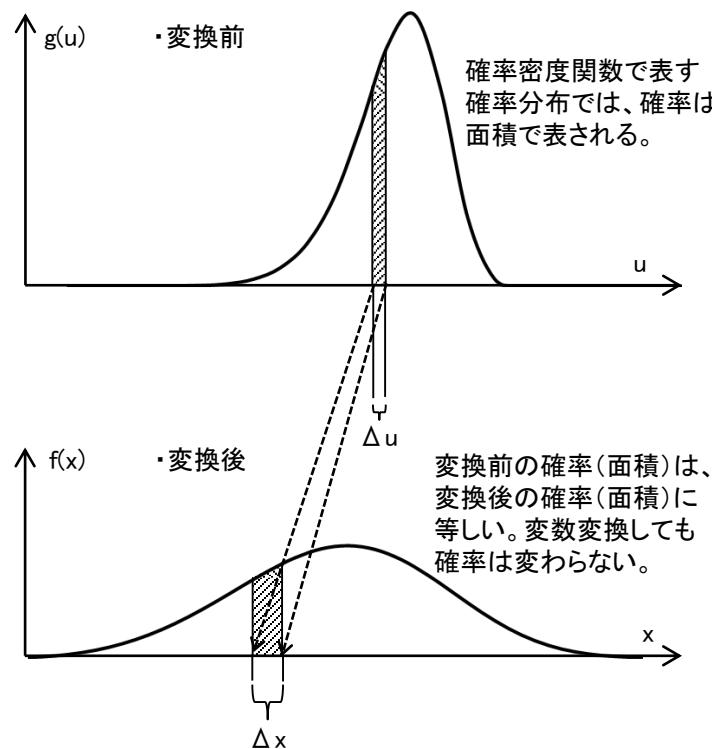
(例) 確率密度関数 $g(u)$ に従う確率変数 U を、1次変換 $X = aU + b$ を使って新しい確率密度変数 X を作る。そのとき、この新しい確率変数 X が従う確率密度関数 $f(x)$ を求める。

$x = au + b$ より、 $u = (1/a)(x - b)$
従って、 $du = (1/a) dx$ つまり、 $du/dx = 1/a$
よって、確率変数 x の確率密度関数は(付式39)を使って、
 $f(x) = g((1/a)(x - b)) (1/a)$

なお、 $g(u)$ の期待値を μ_U として、
変換後の $f(x)$ の期待値 μ_X を求めることができる。

$$\begin{aligned} \mu_X &= \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{\infty} (au + b) g(u) du/dx dx = \int_{-\infty}^{\infty} (au + b) g(u) du \\ &= a \int_{-\infty}^{\infty} u g(u) du + b \int_{-\infty}^{\infty} g(u) du = a \mu_U + b \end{aligned}$$

なお、この式変形では、 $\int_{-\infty}^{\infty} g(u) du = 1$ の関係を使っている。



(付図4) 変数変換に伴う確率の保存

付録15. 2行2列の行列式は平行四辺形の面積

(付図5)に示すような、行列 A を使った、図形の1次変換を考える。

x 軸方向の基本ベクトル e_1 と y 軸方向の基本ベクトル e_2 で囲まれた図形(面積 = 1) が、この変換行列 A によって、ベクトル a_1 とベクトル a_2 で囲われた図形(面積 = S)に変換された場合を考える。

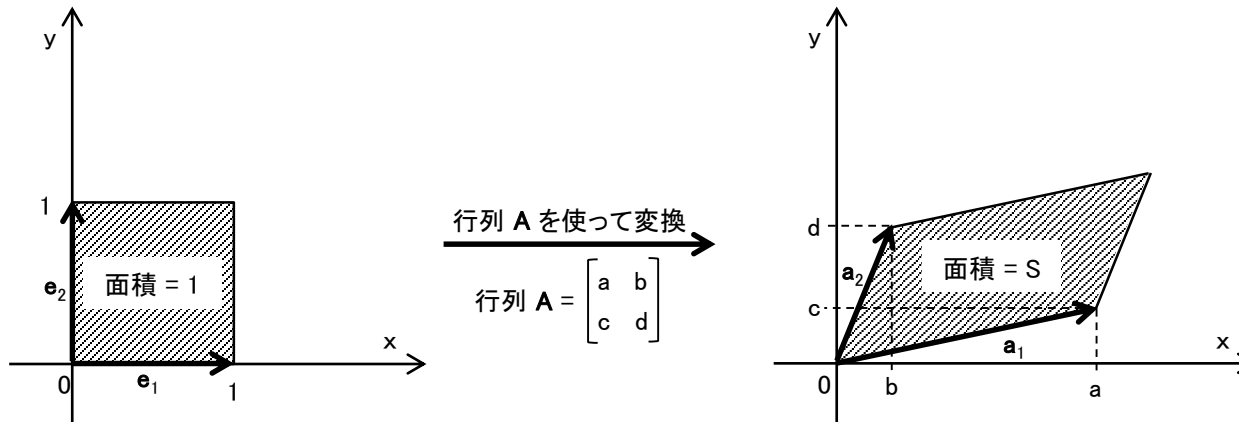
具体的には、

$$\text{行列 } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad a_1 = A e_1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} \quad a_2 = A e_2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$

面積 S は平行四辺形の面積である。ベクトルの内積の関係を使って式を変形する。

$$\begin{aligned} S &= |a_1| |a_2| \sin \theta && \text{(注) } \theta \text{ は、ベクトル } a_1 \text{ とベクトル } a_2 \text{ の角度} \\ &= |a_1| |a_2| (1 - \cos^2 \theta)^{1/2} \\ &= |a_1| |a_2| \{1 - (a_1 \cdot a_2)^2 / (|a_1|^2 |a_2|^2)\}^{1/2} && \text{(注) ベクトルの内積は、}(a_1 \cdot a_2) = |a_1| |a_2| \cos \theta = a b + c d \\ &= \{ |a_1|^2 |a_2|^2 - (a_1 \cdot a_2)^2 \}^{1/2} \\ &= \{ (a^2 + c^2)(b^2 + d^2) - (a b + c d)^2 \}^{1/2} \\ &= (a^2 d^2 + b^2 c^2 - 2 a b c d)^{1/2} \\ &= a d - b c \\ &= \det A && \text{(注) 行列式 } \det A = a d - b c \end{aligned}$$

よって、行列 A を使った1次変換によって、面積は $\det A$ 倍に拡大される。つまり、 $\det A$ は面積拡大率であると言える。



(付図5)図形変換

付録16. 変数変換に使うヤコビアン (Jacobian determinant)

2重積分は、① 微小面積 と ② 高さで作られる微小体積を積分区間全域で積分して求める。

$$\int \int_D \underbrace{g(u, v)}_{\text{② 高さ}} \underbrace{du dv}_{\text{① 微小面積}} \quad D : \text{積分領域} \quad \cdots \text{ (付式40)}$$

重積分の解法として置換積分(変数変換)が知られており、変数 (u, v) を 変数 (x, y) に置換して、関数 g(u, v) を関数 g(u(x, y), v(x, y)) で表す。しかし、この変数変換では、微小面積に関して $du dv = dx dy$ の関係が成立しない。つまり、変数変換によって微小面積の大きさが変化する。ヤコビアン J (x, y) はその変化率を表す。ヤコビアンという変化率を加味して、 $du dv = |J(x, y)| dx dy$ となる。このヤコビアンは以下の手順で求めることができる。

(付図6)に示す図形OACBの面積を計算する。

- 点 O の座標 = (u(x, y), v(x, y))
- 点 A の座標 = (u(x + Δx, y), v(x + Δx, y))
- 点 B の座標 = (u(x, y + Δy), v(x, y + Δy))

依って、ベクトル OA とベクトル OB は以下のように表すことができる。

ベクトル OA = (u(x + Δx, y) - u(x, y), v(x + Δx, y) - v(x, y))
ベクトル OB = (u(x, y + Δy) - u(x, y), v(x, y + Δy) - v(x, y))

よって、図形 OACB を平行四辺形で近似すれば、図形 OACB の領域 S の面積はベクトル OA とベクトル OB の座標を使って計算することができる(付録15)。

領域 S の面積 = $du dv = \{ |(u(x + \Delta x, y) - u(x, y))(v(x, y + \Delta y) - v(x, y)) - (v(x + \Delta x, y) - v(x, y))(u(x, y + \Delta y) - u(x, y))| / (dx dy) \} (dx dy)$

ここで、偏微分の定義より、

$(u(x + \Delta x, y) - u(x, y)) / (dx dy) = \partial u(x, y) / \partial x, \quad (u(x, y + \Delta y) - u(x, y)) / (dx dy) = \partial u(x, y) / \partial y$
 $(v(x + \Delta x, y) - v(x, y)) / (dx dy) = \partial v(x, y) / \partial x, \quad (v(x, y + \Delta y) - v(x, y)) / (dx dy) = \partial v(x, y) / \partial y$

よって、領域 S の面積 = $du dv = |(\partial u(x, y) / \partial x)(\partial v(x, y) / \partial y) - (\partial v(x, y) / \partial x)(\partial u(x, y) / \partial y)| dx dy$

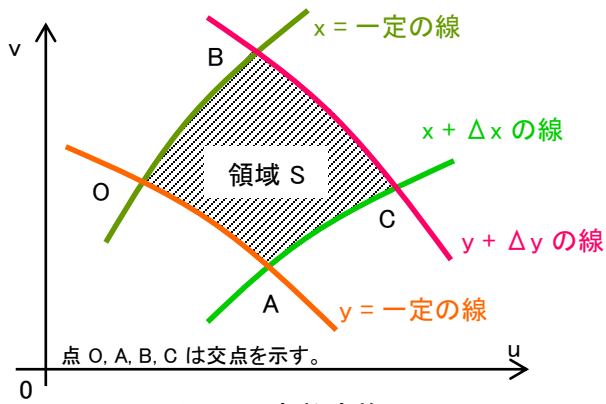
ここで、

$J(x, y) = \partial(u, v) / \partial(x, y) = \begin{vmatrix} \partial u(x, y) / \partial x & \partial u(x, y) / \partial y \\ \partial v(x, y) / \partial x & \partial v(x, y) / \partial y \end{vmatrix} \quad \cdots \text{ (付式41)}$

とおけば、変数変換した領域 S の面積 = $du dv = |J(x, y)| dx dy$ となる。
この式から、ヤコビアン J (x, y) が面積の変化率を表していることが分かる。

以上から、変数変換の重積分の式が求まる。

$\int \int_D g(u, v) du dv = \int \int_S g(u(x, y), v(x, y)) J(x, y) dx dy \quad \cdots \text{ (付式42)}$



(付図6) 変数変換

■ 計算例

重積分でよく使う極座標変換でのヤコビアンを求める。

変数 (u, v) を変数 (r, θ) に変換する。

$u = r \cos \theta, \quad v = r \sin \theta$

$J(r, \theta) = \begin{vmatrix} \partial u / \partial r & \partial u / \partial \theta \\ \partial v / \partial r & \partial v / \partial \theta \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r \cos^2 \theta + r \sin^2 \theta = r$

よって、 $du dv = r dr d\theta$ となる。

付録17. ヤコビアンを使った確率変数の変数変換(一例)

確率変数の変数変換とは、既知の確率変数から新しい確率変数を作り出すことをいう。
ここでは、2つの確率変数の和という新しい確率変数を作る場合を考え、この新しい確率変数の確率分布(確率密度関数)を求める。
既知の 2つの確率変数を u と v とし、その同時確率密度関数を $g(u, v)$ とすると、この2変数の確率分布は、微小面積に相当する $du dv$ に、高さに相当する $g(u, v)$ を掛けた微小体積を u と v の全領域で積分することによって計算できる(付式40)。
ここに、新しい確率変数を2つ用意して、(付式41)でヤコビアンを求めれば、既存の2つの確率変数の重積分の式を新しい2つの確率変数の重積分の式に置き換えることができる。つまり、(付式42)を使用すれば確率変数の和の確率分布を求めることができる。

ここで、問題を整理しておく。
 u と v : 既知の確率変数、 $g(u, v)$: 確率変数 u と v の同時確率密度関数
 x と y : 新しい確率変数、 $x = u + v$ (既知の確率変数の和)、 $y = v$ (2変数の変数変換のためにダミーの確率変数を用意する)
これらの条件のもとで、確率変数 x (つまり確率変数の和)の確率密度関数 $f(x)$ を求める。

新しい確率変数 x と y の同時確率密度関数は、(付式42)に示すように、既存の確率変数の確率密度関数 $g(u, v)$ とヤコビアンを使って、 $g(u(x, y), v(x, y)) |J(x, y)|$ と表すことができる。この同時確率密度関数を使って、新しい確率変数の確率密度関数を求める。
具体的には、新しい確率変数 x の確率密度関数 $f(x)$ は、確率変数 x と y の同時確率密度関数から確率変数 y の存在を消した(確率変数 y に依存しない)確率密度関数(周辺確率密度関数という) $\int_{-\infty}^{\infty} g(u(x, y), v(x, y)) |J(x, y)| dy$ として求めることができる。

以上の手順に従って、まず同時確率密度関数 $g(u(x, y), v(x, y)) |J(x, y)|$ を求める。
ヤコビアン $J(x, y)$ は、(付式41)を使って求める。ここでは、 $u(x, y) = x - v = x - y$ 、 $v(x, y) = y$ であるので、

$$J(x, y) = \partial(u, v) / \partial(x, y) = \begin{vmatrix} \partial u(x, y) / \partial x & \partial u(x, y) / \partial y \\ \partial v(x, y) / \partial x & \partial v(x, y) / \partial y \end{vmatrix} = \begin{vmatrix} \partial(x - y) / \partial x & \partial(x - y) / \partial y \\ \partial y / \partial x & \partial y / \partial y \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1$$

$J(x, y) = 1$ であるので、(付式42)から、 $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(u, v) du dv = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(u(x, y), v(x, y)) dx dy$ となり、

非積分関数に着目すれば、 $g(u, v) = g(u(x, y), v(x, y))$ が成り立つ。

従って、**確率変数 x の確率密度関数 $f(x)$ は、下式のように周辺確率密度関数として求まる。**

$$f(x) = \int_{-\infty}^{\infty} g(u(x, y), v(x, y)) dy = \int_{-\infty}^{\infty} g(x - y, y) dy \quad \cdots \text{ (付式43)}$$

なお、確率変数 u と v が互いに独立している場合は、 $g(u, v) = g_u(u) g_v(v)$ が成立するので、下式のように表すことができる。

$$f(x) = \int_{-\infty}^{\infty} g(x - y, y) dy = \int_{-\infty}^{\infty} g_u(x - y) g_v(y) dy \quad \cdots \text{ (付式44)}$$

(付式44)は、畳み込み積分の形をしている。
つまり、**独立した2つの確率変数の和の確率密度関数は、畳み込み積分で表すことができる。**

畳み込み積分は、確率密度関数 $g_u(x - y)$ と $g_v(y)$ との積を全 y の領域で積分したものである。
なお、畳み込み積分については、別途資料「生産技術 - 古典制御理論の基礎知識 - 動的システムと伝達関数」でも取り上げている。

付録18. デルタ関数を使った確率変数の変数変換

確率変数の変数変換とは、既知の確率変数から新しい確率変数を作り出すことをいう。
ここでは、デルタ関数を使用して、新しい確率変数の確率分布(確率密度関数)を求める方法を示す。

・デルタ関数の定義と性質

$$\int_{-\infty}^{\infty} g(u) \delta(u) du = g(0) \quad \cdots \text{ (付式45)} \quad g(u) \text{ は、任意の実連続関数}$$

この式は、**δ関数(デルタ関数)**は、 $g(u)$ から、 $u = 0$ のときの $g(u)$ の値、つまり、 $g(0)$ を取り出す機能を持っていると解釈できる。

$$\text{また、一般に、} \int_{-\infty}^{\infty} g(u) \delta(u - a) du = g(a) \quad \cdots \text{ (付式46)}$$

この式は、**δ関数(デルタ関数)**は、 $g(u)$ から、 $u = a$ のときの $g(u)$ の値、つまり、 $g(a)$ を取り出す機能を持っていると解釈できる。

$$\text{なお、(付式45)で、} g(x) = 1 \text{ と置けば、} \int_{-\infty}^{\infty} \delta(x) dx = g(0) = 1 \quad \cdots \text{ (付式47) である。}$$

・デルタ関数と確率密度関数

(付式47)の性質から、デルタ関数は確率密度関数であると捉えることができる。

また、(付式46)の性質から、デルタ関数を使用すると、離散的な確率分布を表すことができることが分かる。

例えば、サイコロを振って出る目の確率密度関数 $g(u)$ は、 $g(u) = (1/6) \delta(1 - x) + (1/6) \delta(2 - x) + \cdots + (1/6) \delta(6 - x)$ と表すことができる。

(付録17)と同様に、互いに独立した2つの確率変数を U と V 、確率密度関数をそれぞれ $g_U(u)$ と $g_V(v)$ としたときに、 $X = U + V$ という新しい確率変数が従う確率密度関数を、デルタ関数を使用して求める。なお、ダミー変数として、 $Y = V$ を用意する。

確率変数 U, V は互いに独立であるので、確率変数 U が u という値をとり、確率変数 V が v という値をとる同時確率密度関数は $g_U(u) g_V(v)$ になる。ここで、確率密度関数 $f(x) = f(u + v)$ は、 $g_U(u) g_V(v)$ からデルタ関数 $\delta(x - (u + v))$ の取り出し機能を使って求めることができる。つまり、この機能によって、 $x = u + v$ の確率密度関数得ることができる。この関係は次式のように表される。

$$f(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x - (u + v)) g_U(u) g_V(v) du dv \quad \cdots \text{ (付式48)}$$

$t = x - (u + v) = x - (u + y)$ として置換積分を行う。 $u = x - t - y$, $du/dt = -1$, 積分範囲は $u : -\infty \rightarrow \infty$ に対して、 $t : \infty \rightarrow -\infty$ である。

$$f(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(t) g_U(x - t - y) g_V(y) (-dt) dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(t) g_U(x - t - y) g_V(y) dt dy$$

$t = 0$ のとき、 $\delta(0)$ の抽出機能を使用して、 $\int_{-\infty}^{\infty} \delta(t) g_U(x - t - y) dt = g_U(x - y)$ が抽出されて、(付式49)が求まる

$$f(x) = \int_{-\infty}^{\infty} g_U(x - y) g_V(y) dy \quad \cdots \text{ (付式49)}$$

この式は、(付式44)と同じく、**畳み込み積分**の式である。

次いで、 n 個の確率変数の和について確認しておく。

確率変数 $U_1, U_2, \cdots, U_n$ の確率密度関数がそれぞれ、 $g_1(u_1), g_2(u_2), \cdots, g_n(u_n)$ であるとき、それらの和 $X = \sum_{i=1}^n U_i$ の確率密度関数を求める。

(付式48)を拡張して(付式50)が求まる。また、(付式49)を拡張して畳み込み積分の形式の(付式51)が求まる。

$$f(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \delta(x - \sum_{i=1}^n u_i) \prod_{i=1}^n g_i(u_i) du_1 du_2 \cdots du_n \quad \cdots \text{ (付式50)}$$

$$f(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g_1(x - \sum_{i=2}^n u_i) \prod_{i=2}^n g_i(u_i) du_2 du_3 \cdots du_n \quad \cdots \text{ (付式51)}$$

但し、 Π は総乗を表す。(例) $\prod_{i=1}^n a_i = a_1 \times a_2 \times \cdots \times a_n$

付録19. 標準化という変数変換

期待値が $E[X] = \mu$ で、分散が $V[X] = \sigma^2$ である確率分布は、 $Z = (X - \mu)/\sigma \cdots$ (付式52)
という変数変換によって、平均 = 0, 分散 = 1 の確率分布に変換できる。この変数変換を標準化という。

以下の式変形では、(付録21)と(付録22)に示した式を使う。

$$E[Z] = E[(X - \mu)/\sigma] = (1/\sigma)(E[X] - E[\mu]) = (1/\sigma)(\mu - \mu) = 0$$

よって、確率変数 Z の確率分布において、期待値 = 0 を確認できる。

$$\text{また、} E[Z^2] = E[(X - \mu)^2/\sigma^2] = E[(X - E[X])^2]/\sigma^2 = V[X]/\sigma^2 = \sigma^2/\sigma^2 = 1$$

$$\text{従って、} V[Z] = E[Z^2] - E[Z]^2 = 1 - 0 = 1$$

よって、確率変数 Z の確率分布において、分散 = 1 を確認できる。

(参考)

期待値 = μ , 分散 = σ^2 の確率分布の標本平均の確率分布を標準化してみる。

まず、標本平均の確率分布の期待値と分散を求める。

確率変数を X_i とすれば、標本平均は、 $\bar{X} = (\sum_{i=1}^n X_i)/n$ であるので、標本平均の期待値と分散はそれぞれ、

$$E[\bar{X}] = E[(\sum_{i=1}^n X_i)/n] = E[\sum_{i=1}^n X_i]/n = \sum_{i=1}^n E[X_i]/n = \sum_{i=1}^n \mu/n = (n \mu)/n = \mu$$

$$V[\bar{X}] = V[(\sum_{i=1}^n X_i)/n] = V[\sum_{i=1}^n X_i]/n^2 = \sum_{i=1}^n V[X_i]/n^2 = \sum_{i=1}^n \sigma^2/n^2 = (n \sigma^2)/n^2 = \sigma^2/n$$

よって、この標本平均は、期待値 = μ , 分散 = σ^2/n の確率分布に従う。

次いで、この期待値 = μ , 分散 = σ^2/n の確率分布を、 $Z = n^{1/2}(\bar{X} - \mu)/\sigma$ で変数変換する。

分散が、 σ^2/n であるので、 $Z = (\bar{X} - \mu)/\sigma$ という変数変換ではなく、 $n^{1/2}$ を加味した $Z = n^{1/2}(\bar{X} - \mu)/\sigma$ で変数変換する。

$$E[Z] = E[n^{1/2}(\bar{X} - \mu)/\sigma] = (n^{1/2}/\sigma)E[\bar{X} - \mu] = (n^{1/2}/\sigma)(E[\bar{X}] - \mu) = (n^{1/2}/\sigma)(\mu - \mu) = 0$$

$$V[Z] = V[n^{1/2}(\bar{X} - \mu)/\sigma] = (n/\sigma^2)V[\bar{X} - \mu] = (n/\sigma^2)V[\bar{X}] = (n/\sigma^2)(\sigma^2/n) = 1$$

よって、この変数変換を行った(標準化した)標本平均は、期待値 = 0, 分散 = 1 の確率分布に従うことが分かる。

付録20. ハザード関数を使った確率密度関数

ハザード関数はある時点における故障率を表す関数である。ある時点まで正常に動作していて、次の瞬間に故障する確率を表す。

X を故障が起こる時間(確率変数)とする。また、この X の確率密度関数を $f(x)$ 累積分布関数を $F(x)$ とする。 $F(x) = \int_{-\infty}^x f(x) dx$ である。

さて、ある時間 x までは正常動作していて、次の瞬間 $x + \Delta x$ の間で故障する確率 P は、 $X \geq x$ という条件のもとで(条件付確率として)、

$$P(x < X \leq x + \Delta x | X > x)$$

この確率を Δx で割れば、短時間当たりの故障確率(故障率)が求まり、 $\Delta x \rightarrow 0$ としたものをハザード関数 $h(x)$ という。

$$h(x) = \lim_{\Delta x \rightarrow 0} P(x < X \leq x + \Delta x | X > x) / \Delta x$$

ここで、確率密度関数を $f(x)$ 、累積分布関数を $F(x)$ とすれば、 $F'(x) = f(x)$ (ここで、' は1階微分を表す)の関係があり、また、確率を表すのに累積分布関数を使えば、確率は累積分布関数の差分になるので、

$P(x < X \leq x + \Delta x) = F(x + \Delta x) - F(x)$ であり、 $P(X) = 1 - F(x)$ である。この関係を活用すれば、

$$P(x < X \leq x + \Delta x | X > x) = P(x < X \leq x + \Delta x) / P(X > x) = (F(x + \Delta x) - F(x)) / (1 - F(x))$$

$$\begin{aligned} \text{よって、ハザード関数 } h(x) &= \lim_{\Delta x \rightarrow 0} \{ P(x < X \leq x + \Delta x | X > x) / \Delta x \} = \lim_{\Delta x \rightarrow 0} \{ (F(x + \Delta x) - F(x)) / (1 - F(x)) / \Delta x \} \\ &= \lim_{\Delta x \rightarrow 0} \{ ((F(x + \Delta x) - F(x)) / \Delta x) / (1 - F(x)) \} = F'(x) / (1 - F(x)) = f(x) / (1 - F(x)) \quad \dots \text{(付式53)} \end{aligned}$$

このハザード関数で累積分布関数と確率密度関数を表す。 $\{ \log(1 - F(x)) \}' = (1 - F(x))' / (1 - F(x)) = -f(x) / (1 - F(x))$ であることを使って、

$$h(x) \text{ の式を積分すると、} F(0) = 0 \text{ であるので、} \int_0^x h(x) dx = \int_0^x f(x) / [1 - F(x)] dx = [-\log(1 - F(x))]_0^x = -\log(1 - F(x))$$

$$\text{よって、} 1 - F(x) = e^{-\int_0^x h(x) dx}$$

$$\text{従って、累積分布関数は、} F(x) = 1 - e^{-\int_0^x h(x) dx} \quad \dots \text{(付式54)}$$

$$\text{確率密度関数は、} f(x) = F'(x) = h(x) e^{-\int_0^x h(x) dx} \quad \dots \text{(付式55)}$$

■ハザード関数の例(バスタブ曲線)

次のようなハザード関数を考える。

$$h(x) = \alpha \beta x^{\beta-1} \quad \alpha > 0, \beta > 0 \quad \dots \text{(付式56)}$$

(付式56)を図にしたのが、(付図7)と(付図8)である。

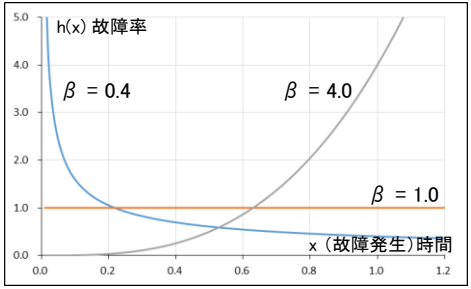
2つの図から分かることは、

まず、 α の値によって故障率の大きさが変化することが分かる。

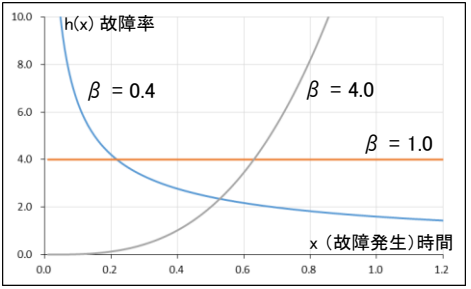
次いで、 β の値によって故障率の形状が変化することが分かる。

そして、 $\beta < 1$ (図では、 $\beta = 0.4$) のときは早い時間帯での故障率が高い。これは初期故障の期間を表していると言える。 $\beta = 1$ のときは故障率が一定しており、偶発故障の期間を表していると言える。 $\beta > 1$ (図では、 $\beta = 4.0$) のときは時間が経過するほど故障率が上昇しており、これは劣化(摩耗)に伴い故障率が上がる期間を表していると言える。

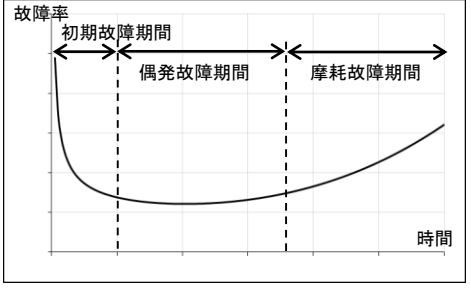
このように、(付式56)のハザード関数は、(付図9)に示すような、バスタブ型の故障率を表すことができる関数であると言える。



(付図7)ハザード関数($\alpha = 1$)



(付図8)ハザード関数($\alpha = 4$)



(付図9)バスタブ曲線

付録21. 関数の期待値の算出方法

■関数の期待値

確率変数が定義されているときには、確率変数の関数もまた確率変数になるので、関数の期待値を定義することができる。

確率変数を X としたときの、関数 $\phi(X)$ の期待値(定義)

・離散確率変数の場合 $E[\phi(X)] = \sum_{i=1}^{\infty} \phi(x_i) f(x_i) \cdots$ (付式57) ・連続確率変数の場合 $E[\phi(X)] = \int_{-\infty}^{\infty} \phi(x) f(x) dx \cdots$ (付式58)

ただし、 $f(x)$ が離散型のとき確率質量関数で $\sum_{i=1}^{\infty} f(x_i) = 1$, $f(x)$ が連続型のとき確率密度関数で $\int_{-\infty}^{\infty} f(x) dx = 1$

■モーメント(moment)

モーメントは確率変数 X が従う確率分布の特徴を数値化するものであり、積率とも言われる(モーメントの名前の由来は力学のモーメントの計算式に似ていることによる)。(表2)にモーメントの定義と1次と2次のモーメントを使用した確率分布の期待値と分散の計算式を載せる。

(付表1)モーメント

	説明	離散型確率変数の場合	連続型確率変数の場合
モーメントの定義	$\phi(X) = X^k$ のときの期待値を、確率変数 X の k 次モーメントという。	$E[\phi(X)] = E[X^k] = \sum_{i=1}^{\infty} x_i^k f(x_i)$	$E[\phi(X)] = E[X^k] = \int_{-\infty}^{\infty} x^k f(x) dx$
1次モーメント	$k = 1$, $E[X^k] = E[X]$ となる。 確率変数 X の平均(確率分布を加味した加重平均)を表す。	期待値 $E[X] = \sum_{i=1}^{\infty} x_i f(x_i)$	期待値 $E[X] = \int_{-\infty}^{\infty} x f(x) dx$
2次モーメント	$k = 2$, $E[X^k] = E[X^2]$ となる。 分散 $V[X]$ は、1次モーメント $E[X]$ と2次モーメント $E[X^2]$ を使って表すことができる。	分散 $V[X] = \sum_{i=1}^{\infty} (x_i - E[X])^2 f(x_i)$ $= \sum_{i=1}^{\infty} x_i^2 f(x_i) - 2E[X] \sum_{i=1}^{\infty} x_i f(x_i) + E[X]^2 \sum_{i=1}^{\infty} f(x_i)$ $= E[X^2] - 2E[X]^2 + E[X]^2 = E[X^2] - E[X]^2$ (なお、 $\sum_{i=1}^{\infty} f(x_i) = 1$ である)	分散 $V[X] = \int_{-\infty}^{\infty} (x - E[X])^2 f(x) dx$ $= \int_{-\infty}^{\infty} x^2 f(x) dx - 2E[X] \int_{-\infty}^{\infty} x f(x) dx + E[X]^2 \int_{-\infty}^{\infty} f(x) dx$ $= E[X^2] - 2E[X]^2 + E[X]^2 = E[X^2] - E[X]^2$ (なお、 $\int_{-\infty}^{\infty} f(x) dx = 1$ である)

■モーメント母関数(積率母関数、moment generating function) (付録24) 参照

ある確率分布の確率変数の k 次モーメントは、定義式を使って求める方法の他に多項式を用意して求める方法がある。この多項式が、**モーメント母関数** $M_X(t) = E[e^{tX}] \cdots$ (付式59) である。このモーメント母関数は、マクローリン級数展開すると(付60)で表すことができる。

$$M_X(t) = E[e^{tX}] = E[1 + tX + 1/2! (tX)^2 + 1/3! (tX)^3 + \cdots] = 1 + t E[X] + 1/2! t^2 E[X^2] + 1/3! t^3 E[X^3] + \cdots \cdots \text{(付式60)}$$

ここで、 t の k 次の係数を C_k と置くと、 $C_k = 1/k! E[X^k] \cdots$ (付式61) となり、この係数 C_k を求めることができれば、 $E[X^k] = k! C_k \cdots$ (式62) から、確率変数 X の k 次モーメント $E[X^k]$ を求めることができる。

係数 C_k は容易に求めることができる。つまり、(付式61)を(付式60)に代入すれば、 $M_X(t) = 1 + C_1 t + C_2 t^2 + C_3 t^3 + \cdots \cdots$ (付式63) について、(付式63)を k 階微分し(付式62)を代入すれば、 $M_X^{(k)}(t) = d^k M_X(t)/dt^k = k! C_k + (t \text{ の多項式}) = E[X^k] + (t \text{ の多項式}) \cdots$ (付式64)

よって、(付式64)の両辺に $t = 0$ を代入すれば、 $(t \text{ の多項式}) = 0$ となるので、 $M_X^{(k)}(t)|_{t=0} = d^k M_X(t)/dt^k|_{t=0} = E[X^k] \cdots$ (付式65) となり、モーメントの関数 $M_X(t)$ から、 k 次のモーメント $E[X^k]$ を求めることができる。

(付録22)に、期待値の計算事例として、定義式に従って計算した事例とモーメント母関数を使用して求めた事例を載せる。

付録22. 期待値の計算例

例として、 $x \geq 0$ のとき $f(x) = x e^{-x}$, $x < 0$ のとき $f(x) = 0$ である確率密度関数 $f(x)$ (付図10)の平均と分散を求める。

(1) 期待値の定義の式から求める方法

この確率密度関数の期待値と分散を計算する前に、以下の積分の値を部分積分を使って求めておく。

$$\lim_{N \rightarrow \infty} \int_0^N x e^{-x} dx = \lim_{N \rightarrow \infty} \int_0^N x (-e^{-x})' dx = \lim_{N \rightarrow \infty} \left[x (-e^{-x}) \right]_0^N - \int_0^N (-e^{-x}) dx = \lim_{N \rightarrow \infty} \left[-N e^{-N} - [e^{-x}]_0^N \right] = \lim_{N \rightarrow \infty} (-N e^{-N} - e^{-N} + 1) = 0 + 0 + 1 = 1$$

$$\lim_{N \rightarrow \infty} \int_0^N x^2 e^{-x} dx = \lim_{N \rightarrow \infty} \int_0^N x^2 (-e^{-x})' dx = \lim_{N \rightarrow \infty} \left[x^2 (-e^{-x}) \right]_0^N - 2 \int_0^N x (-e^{-x}) dx = \lim_{N \rightarrow \infty} (-N^2 e^{-N}) + 2 \lim_{N \rightarrow \infty} \int_0^N x e^{-x} dx = 0 + 2 \cdot 1 = 2$$

$$\lim_{N \rightarrow \infty} \int_0^N x^3 e^{-x} dx = \lim_{N \rightarrow \infty} \int_0^N x^3 (-e^{-x})' dx = \lim_{N \rightarrow \infty} \left[x^3 (-e^{-x}) \right]_0^N - 3 \int_0^N x^2 (-e^{-x}) dx = \lim_{N \rightarrow \infty} (-N^3 e^{-N}) + 3 \lim_{N \rightarrow \infty} \int_0^N x^2 e^{-x} dx = 0 + 3 \cdot 2 = 6$$

ここでの計算では、下記に示す極限値の計算にロピタルの定理(付録5)を使用している。

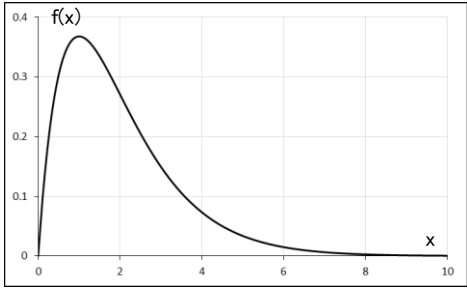
$$\lim_{N \rightarrow \infty} (-N^3 e^{-N}) = \lim_{N \rightarrow \infty} \{ -(N^3)' / (e^N)' \} = \lim_{N \rightarrow \infty} \{ -3N^2 / e^N \} = \lim_{N \rightarrow \infty} \{ -3(N^2)' / (e^N)' \} = \lim_{N \rightarrow \infty} \{ -6N / e^N \} = \lim_{N \rightarrow \infty} \{ -6(N)' / (e^N)' \} = \lim_{N \rightarrow \infty} \{ -6 / e^N \} = 0$$

・期待値は、確率変数 X の期待値である。

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x^2 e^{-x} dx = \lim_{N \rightarrow \infty} \int_0^N x^2 e^{-x} dx = 2$$

・分散は $(X - \mu)^2$ の期待値である。上記で求めた、期待値 = 2 を使う。

$$\begin{aligned} E[(X - \mu)^2] &= \int_{-\infty}^{\infty} (x - E[X])^2 f(x) dx = \int_0^{\infty} (x - 2)^2 x e^{-x} dx \\ &= \lim_{N \rightarrow \infty} \int_0^N (x^3 - 4x^2 + 4x) e^{-x} dx = 6 - 4 \cdot 2 + 4 \cdot 1 = 2 \end{aligned}$$



(付図10) 確率密度関数

(2) モーメント母関数を使う方法

モーメント母関数を求める。求める際の極限値の計算には、 $t = 0$ 近傍(つまり、 $1 - t > 0$)でロピタルの定理を使用する。また、部分積分を行う。

$$\begin{aligned} M_x(t) &= E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx = \int_0^{\infty} e^{tx} x e^{-x} dx = \int_0^{\infty} x e^{-(1-t)x} dx = \lim_{N \rightarrow \infty} \int_0^N x e^{-(1-t)x} dx = \lim_{N \rightarrow \infty} \int_0^N x \{ -e^{-(1-t)x} / (1-t) \}' dx \\ &= \lim_{N \rightarrow \infty} \left[x (-e^{-(1-t)x} / (1-t)) \right]_0^N - \int_0^N (-e^{-(1-t)x} / (1-t)) dx = \lim_{N \rightarrow \infty} \left[-N e^{-(1-t)N} / (1-t) - [e^{-(1-t)x} / (1-t)^2]_0^N \right] \\ &= \lim_{N \rightarrow \infty} \left[-N e^{-(1-t)N} / (1-t) - e^{-(1-t)N} / (1-t)^2 + 1 / (1-t)^2 \right] = \lim_{N \rightarrow \infty} \left[(-N / (1-t) - 1 / (1-t)^2) / e^{(1-t)N} \right] + 1 / (1-t)^2 \\ &= \lim_{N \rightarrow \infty} \left[(-N / ((1-t) e^{(1-t)N})) \right] + \lim_{N \rightarrow \infty} \left[-1 / ((1-t)^2 e^{(1-t)N}) \right] + 1 / (1-t)^2 = \lim_{N \rightarrow \infty} \left[(-N / ((1-t) e^{(1-t)N})) \right] + 0 + 1 / (1-t)^2 \\ &= \lim_{N \rightarrow \infty} (-N)' / \{ (1-t) e^{(1-t)N} \}' + 1 / (1-t)^2 = \lim_{N \rightarrow \infty} (-1 / (-e^{(1-t)N} + (1-t) (-N) e^{(1-t)N})) + 1 / (1-t)^2 = 0 + 1 / (1-t)^2 = (1-t)^{-2} \end{aligned}$$

よって、 $M_x'(t) = 2(1-t)^{-3}$, $M_x''(t) = 6(1-t)^{-3}$, ... となり、1 次モーメント $M_x'(0) = 2$, 2 次モーメント $M_x''(0) = 6$ などが求まる。

なお、 $M_x(t) = (1-t)^{-2} = 1 + 2t + 3t^2 + 4t^3 + \dots$ とマクローリン展開した上で微分して、 $M_x'(t) = 2 + 6t + 12t^2 + \dots$, $M_x''(t) = 6 + 24t + \dots$ を求め、次いでそれぞれの式で、 $t = 0$ において $M_x'(0) = 2$, $M_x''(0) = 6$ を求める方法もある。

最後に、これらのモーメントを使用して期待値と分散を求める。

$$\text{期待値} = E[X] = M_x'(0) = 2 \quad \text{分散} = V[X] = E[X^2] - (E[X])^2 = M_x''(0) - (M_x'(0))^2 = 6 - 2^2 = 2$$

付録23. 期待値と分散の公式

■期待値(expected value)

ある確率密度分布に従う確率変数の期待値は、確率変数を取るすべての値に対して夫々の確率の重みを掛けた加重平均として定義される。このように定義された期待値は、確率変数がとると期待される値になる。このことから、期待値という名前が付けられている。期待値と分散のそれぞれの和、定数倍、定数の加算の公式は離散型と連続型で共通である。

(付表2) 期待値と分散の公式

説明	期待値	分散
定義(離散型)	$P(X = x_i)$ を、確率変数 X が実現値 x_i を取る確率(確率質量関数)とすれば、期待値 $E[X] = \sum_{i=1}^n x_i P(X = x_i)$	$P(X = x_i)$ を、確率変数 X が実現値 x_i を取る確率(確率質量関数)とすれば、分散 $V[X] = E[(X - E[X])^2] = \sum_{i=1}^n (x_i - E[X])^2 P(X = x_i)$
定義(連続型)	$f(x)$ を確率変数 X の確率密度関数であるとすれば、期待値 $E[X] = \int_{-\infty}^{\infty} x f(x) dx$	$f(x)$ を確率変数 X の確率密度関数であるとすれば、分散 $V[X] = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (x - E[X])^2 f(x) dx$
和 (注) $i = 1 \sim m$ $j = 1 \sim n$ の総和範囲を記載省略	$E[X + Y] = E[X] + E[Y]$ (離散型で確認) $X = \{x_1, x_2, \dots, x_m\}, Y = \{y_1, y_2, \dots, y_n\}$ とすれば、 $X = x_i, Y = y_j$ となる同時確率は、 $P(X = x_i, Y = y_j)$ であるので、 $E[X + Y] = \sum \sum (x_i + y_j) P(X = x_i, Y = y_j)$ $= \sum \sum x_i P(X = x_i, Y = y_j) + \sum \sum y_j P(X = x_i, Y = y_j)$ $= \sum x_i P(X = x_i) + \sum y_j P(Y = y_j)$ $= E[X] + E[Y]$	$V[X + Y] = V[X] + V[Y] + 2 \text{Cov}[X, Y]$ 但し、Cov は共分散 (離散型で確認) $X = \{x_1, x_2, \dots, x_m\}, Y = \{y_1, y_2, \dots, y_n\}$ とすれば、 $X = x_i, Y = y_j$ となる同時確率は、 $P(X = x_i, Y = y_j)$ であるので、 $V[X + Y] = \sum \sum ((x_i + y_j) - E[X + Y])^2 P(X = x_i, Y = y_j)$ $= \sum \sum ((x_i + y_j) - (E[X] + E[Y]))^2 P(X = x_i, Y = y_j)$ $= \sum \sum (x_i - E[X])^2 P(X = x_i, Y = y_j)$ $+ \sum \sum (y_j - E[Y])^2 P(X = x_i, Y = y_j)$ $+ 2 \sum \sum (x_i - E[X]) (y_j - E[Y]) P(X = x_i, Y = y_j)$ $= \sum (x_i - E[X])^2 P(X = x_i) + \sum (y_j - E[Y])^2 P(Y = y_j)$ $+ 2 \sum \sum (x_i - E[X]) (y_j - E[Y]) P(X = x_i, Y = y_j)$ $= V[X] + V[Y] + 2 \text{Cov}[X, Y]$ 但し、 $\text{Cov}[X, Y] = \sum \sum (x_i - E[X]) (y_j - E[Y]) P(X = x_i, Y = y_j)$
定数倍 (注) $i = 1 \sim n$ の総和範囲を記載省略	$E[cX] = c E[X]$ c は定数 (離散型で確認) $E[cX] = \sum (c x_i) P(X = x_i)$ $= c \sum x_i P(X = x_i)$ $= c E[X]$	$V[cX] = c^2 V[X]$ c は定数 (離散型で確認) $V[cX] = \sum (c x_i - E[cX])^2 P(X = x_i)$ $= \sum (c x_i - c E[X])^2 P(X = x_i)$ $= c^2 \sum (x_i - E[X])^2 P(X = x_i) = c^2 V[X]$
定数の加算 (注) $i = 1 \sim n$ の総和範囲を記載省略	$E[X + t] = E[X] + t$ t は定数 (離散型で確認) $E[X + t] = \sum (x_i + t) P(X = x_i) = \sum x_i P(X = x_i) + t \sum P(X = x_i)$ ここで、 $\sum P(X = x_i) = 1$ なので、 $E[X + t] = E[X] + t$	$V[X + t] = V[X]$ t は定数 (離散型で確認) $V[X + t] = \sum (x_i + t - E[X + t])^2 P(X = x_i)$ $= \sum (x_i + t - E[X] - t)^2 P(X = x_i)$ $= \sum (x_i - E[X])^2 P(X = x_i) = V[X]$

付録24. モーメント母関数と特性関数 (Characteristic function)

モーメント母関数(積率母関数)と特性関数、それぞれの特徴を比較して(付表3)に整理する。
特性関数はモーメント母関数を虚数領域に拡大した関数であり、全ての確率分布に対して存在し、確率密度関数と1:1の関係があることから、中心極限定理の証明にも使用されている。

(付表3)モーメント母関数と特性関数の比較

	モーメント母関数 $M_X(t)$	特性関数 $\phi_X(t)$
定義式	e^{tX} の期待値として定義される $M_X(t) = E[e^{tX}]$	モーメント母関数を虚数領域に拡大して、 e^{itX} の期待値として定義される (j は虚数単位) $\phi_X(t) = E[e^{itX}]$
マクローリン展開	$M_X(t) = E[e^{tX}] = E[1 + tX + (1/2!) (tX)^2 + (1/3!) (tX)^3 + \dots]$ $= E[1] + E[tX] + E[(1/2!) (tX)^2] + E[(1/3!) (tX)^3] + \dots]$ $= E[1] + t^1/1! E[X] + t^2/2! E[X^2] + t^3/3! E[X^3] + \dots]$	$\Phi_X(t) = E[e^{itX}] = E[1 + j tX + (1/2!) (j tX)^2 + (1/3!) (j tX)^3 + \dots]$ $= E[1] + E[j tX] + E[(1/2!) (j tX)^2] + E[(1/3!) (j tX)^3] + \dots]$ $= E[1] + (j t)^1/(1!) E[X] + (j t)^2/(2!) E[X^2] + (j t)^3/(3!) E[X^3] + \dots]$
1 階微分	$d M_X(t)/dt = M_X'(t) = E[X] + (1/1!) t E[X^2] + 1/(2!) t^2 E[X^3] + \dots$	$d \phi_X(t)/dt = \phi_X'(t) = j E[X] + (j^2/1!) t E[X^2] + (j^3/2!) t^2 E[X^3] + \dots$
2 階微分	$d^2 M_X(t)/dt^2 = M_X''(t) = E[X^2] + 1/(1!) t E[X^3] + \dots$	$d^2 \phi_X(t)/dt^2 = \phi_X''(t) = j^2 E[X^2] + (j^3/(1!)) t E[X^3] + \dots$
n 階微分	$d^n M_X(t)/dt^n = M_X^{(n)}(t) = E[X^n] + 1/(1!) t E[X^{n+1}] + \dots$	$d^n \phi_X(t)/dt^n = \phi_X^{(n)}(t) = j^n E[X^n] + j^{n+1}/(1!) t E[X^{n+1}] + \dots$
1 次の期待値	$M_X'(0) = E[X]$ より、 $E[X] = M_X'(0)$	$\phi_X'(0) = j E[X]$ より、 $E[X] = j^{-1} \phi_X'(0)$
2 次の期待値	$M_X''(0) = E[X^2]$ より、 $E[X^2] = M_X''(0)$	$\phi_X''(0) = j^2 E[X^2]$ より、 $E[X^2] = j^{-2} \phi_X''(0)$
n 次の期待値	$M_X^{(n)}(0) = E[X^n]$ より、 $E[X^n] = M_X^{(n)}(0)$	$\phi_X^{(n)}(0) = j^n E[X^n]$ より、 $E[X^n] = j^{-n} \phi_X^{(n)}(0)$
特徴	・確率分布の母数(平均、分散など)を求めるときに使う ・コーシー分布のように、モーメント母関数が存在しない確率分布がある	・確率分布の母数(平均、分散など)を求めるときに使う ・特性関数は、全ての確率分布に対して存在する ・確率分布の特性関数は確率密度関数と1:1の関係がある 特性関数が一致すれば確率密度関数(確率分布)は同じになる - この性質は、中心極限定理の証明で使用される ・特性関数の定義式は、確率密度関数をフーリエ変換した式の形になっている

付録25. 期待値と分散の導出(一様分布)

■離散型の一様分布

・期待値

$$E[X] = \sum_{x=1}^N x P(X=x) = \sum_{x=1}^N x (1/N) = \{N(N+1)/2\} (1/N) = (N+1)/2$$

・分散

(付録2)より、 $V[X] = E[X^2] - E[X]^2$

つまり、 $E[X^2]$ が求まれば、この式を使って $V[X]$ が求まる(この頁の下段、(参考)2乗和の計算)。

$$E[X^2] = \sum_{x=1}^N x^2 (1/N) = \{N(N+1)(2N+1)/6\} (1/N) = (2N^2 + 3N + 1)/6$$

よって、

$$V[X] = E[X^2] - E[X]^2 = \{(2N^2 + 3N + 1)/6\} - \{(N+1)/2\}^2 = (2N^2 + 3N + 1)/6 - (N^2 + 2N + 1)/4 = (N^2 - 1)/12$$

■連続型の一様分布

・期待値

$x < a, b < x$ の区間で、 $f(x) = 0$ であるので、

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^a x f(x) dx + \int_a^b x f(x) dx + \int_b^{\infty} x f(x) dx = 0 + \int_a^b x \{1/(b-a)\} dx + 0 \\ &= \{1/(b-a)\} [x^2/2]_a^b = \{1/(b-a)\} (b^2 - a^2)/2 = (a+b)/2 \end{aligned}$$

・分散

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{-\infty}^a x^2 f(x) dx + \int_a^b x^2 f(x) dx + \int_b^{\infty} x^2 f(x) dx = 0 + \int_a^b x^2 \{1/(b-a)\} dx + 0 = \int_a^b x^2 (1/(b-a)) dx \\ &= \{1/(b-a)\} [x^3/3]_a^b = \{1/(b-a)\} (b^3 - a^3)/3 = \{1/(b-a)\} (b-a)(a^2 + ab + b^2)/3 \\ &= (a^2 + ab + b^2)/3 \end{aligned}$$

よって、

$$\begin{aligned} V[X] &= E[X^2] - E[X]^2 \\ &= (a^2 + ab + b^2)/3 - \{(a+b)/2\}^2 = (4a^2 + 4ab + 4b^2 - 3a^2 - 6ab - 3b^2)/12 = (a^2 - 2ab + b^2)/12 = (a-b)^2/12 \end{aligned}$$

(参考)2乗和の計算

$$\text{・2乗和} \quad \sum_{k=1}^N k^2 = N(N+1)(2N+1)/6$$

確認 $(k+1)^3 = k^3 + 3k^2 + 3k + 1$ の関係を利用する。

$$\sum_{k=1}^N (3k^2 + 3k + 1) = 3 \sum_{k=1}^N k^2 + 3 \{N(N+1)/2\} + N \quad \cdots \text{①}$$

$$\sum_{k=1}^N \{(k+1)^3 - k^3\} = (2^3 - 1^3) + (3^3 - 2^3) + \cdots + (N^3 - (N-1)^3) + ((N+1)^3 - N^3) = (N+1)^3 - 1^3 = N^3 + 3N^2 + 3N \quad \cdots \text{②}$$

$$\text{ここで、①} = \text{②} \text{ であるので、} 3 \sum_{k=1}^N k^2 + 3 \{N(N+1)/2\} + N = N^3 + 3N^2 + 3N$$

$$\begin{aligned} \text{従って、} \sum_{k=1}^N k^2 &= \{N^3 + 3N^2 + 3N - 3 \{N(N+1)/2\} - N\} / 3 = \{N^3 + 3N^2 + 2N - (N^2 + N)/2\} / 3 \\ &= (2N^3 + 6N^2 + 4N - N^2 - N) / 6 = N(N+1)(2N+1)/6 \end{aligned}$$

付録26. 期待値と分散の導出(ベルヌーイ分布)

・期待値

$$E[X] = \sum_{x=0}^1 x P(X=x) = \sum_{x=0}^1 x p^x (1-p)^{1-x} = 0 p^0 (1-p)^{1-0} + 1 p^1 (1-p)^{1-1} = 0 + p = p$$

・分散

$$E[X^2] = \sum_{x=0}^1 x^2 P(X=x) = \sum_{x=0}^1 x^2 p^x (1-p)^{1-x} = 0^2 p^0 (1-p)^{1-0} + 1^2 p^1 (1-p)^{1-1} = 0 + p = p$$

よって、

$$\begin{aligned} V[X] &= E[X^2] - E[X]^2 \\ &= p - p^2 = p(1-p) \end{aligned}$$

付録27. 期待値と分散の導出(カテゴリカル分布)

・期待値

$x_i = 1$, 他の $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_k = 0$ であるとして、

$$\begin{aligned} E[X_i] &= \sum_{i=0}^k x_i P(X = x_i) = 0 \cdot p_1^{x_1} + 0 \cdot p_2^{x_2} + \dots + 0 \cdot p_{i-1}^{x_{i-1}} + 1 \cdot p_i^{x_i} + 0 \cdot p_{i+1}^{x_{i+1}} + \dots + 0 \cdot p_k^{x_k} \\ &= p_i^{x_i} = p_i^1 = p_i \end{aligned}$$

・分散

同様に、 $x_i = 1$, 他の $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_k = 0$ であるとして、

$$\begin{aligned} E[X_i^2] &= \sum_{i=0}^k x_i^2 P(X = x_i) = 0^2 \cdot p_1^{x_1} + 0^2 \cdot p_2^{x_2} + \dots + 0^2 \cdot p_{i-1}^{x_{i-1}} + 1^2 \cdot p_i^{x_i} + 0^2 \cdot p_{i+1}^{x_{i+1}} + \dots + 0^2 \cdot p_k^{x_k} \\ &= p_i^{x_i} = p_i^1 = p_i \end{aligned}$$

よって、

$$\begin{aligned} V[X_i] &= E[X_i^2] - E[X_i]^2 \\ &= p_i - p_i^2 = p_i(1 - p_i) \end{aligned}$$

付録28. 期待値と分散の導出(二項分布)

二項分布のモーメント母関数 $M_x(t)$ (付録21)を求める。

$$M_x(t) = E[e^{tx}] = \sum_{x=0}^n e^{tx} \{ {}_n C_x p^x (1-p)^{n-x} \} = \sum_{x=0}^n {}_n C_x (e^t p)^x (1-p)^{n-x}$$

二項定理(付録1)より、

$$M_x(t) = (e^t p + (1-p))^n$$

このモーメント母関数を使用して、二項分布の期待値と分散を求める。

・期待値

$$E[X] = d/dt \{ M_x(t) \} \Big|_{t=0} = d/dt \{ (e^t p + 1 - p)^n \} \Big|_{t=0} = \{ n (e^t p + 1 - p)^{n-1} p e^t \} \Big|_{t=0} = \mathbf{n p}$$

・分散

$$\begin{aligned} E[X^2] &= d^2/dt^2 \{ M_x(t) \} \Big|_{t=0} = d^2/dt^2 \{ (e^t p + 1 - p)^n \} \Big|_{t=0} = d/dt \{ n (e^t p + 1 - p)^{n-1} p e^t \} \Big|_{t=0} \\ &= \{ n (n-1) (e^t p + 1 - p)^{n-2} p^2 e^{2t} + n (e^t p + 1 - p)^{n-1} p e^t \} \Big|_{t=0} \\ &= n (n-1) (p + 1 - p)^{n-2} p^2 + n (p + 1 - p)^{n-1} p = n (n-1) p^2 + n p \end{aligned}$$

よって、

$$\begin{aligned} V[X] &= E[X^2] - E[X]^2 \\ &= \{ n (n-1) p^2 + n p \} - (n p)^2 = \mathbf{n p (1 - p)} \end{aligned}$$

付録29. 期待値と分散の導出(多項分布)

・期待値

確率変数 X_i の平均を求める。式変形には、(付録2)の多項定理を使用する。

$$\begin{aligned} E[X_i] &= \sum_{x_1+x_2+\dots+x_m=n} x_i \{ n! / (x_1! x_2! \dots x_m!) \} p_1^{x_1} p_2^{x_2} \dots p_m^{x_m} \\ &= \sum_{x_1+x_2+\dots+x_m=n} x_i \{ (n(n-1)!) p_1^{x_1} p_2^{x_2} \dots p_{i-1}^{x_{i-1}} p_i^{(x_i-1)+1} p_{i+1}^{x_{i+1}} \dots p_m^{x_m} / (x_1! x_2! \dots x_{i-1}! (x_i(x_i-1)!) x_{i+1}! \dots x_m!) \} \\ &= n p_i \sum_{x_1+x_2+\dots+x_m=n-1} (n-1)! p_1^{x_1} p_2^{x_2} \dots p_{i-1}^{x_{i-1}} p_i^{x_i-1} p_{i+1}^{x_{i+1}} \dots p_m^{x_m} / (x_1! x_2! \dots x_{i-1}! (x_i-1)! x_{i+1}! \dots x_m!) \\ &= n p_i (p_1 + p_2 + \dots + p_m)^{n-1} \quad (\text{注}) x_1 + x_2 + \dots + x_{i-1} + (x_i - 1) + x_{i+1} + \dots + x_m = (x_1 + x_2 + \dots + x_m) - 1 = n - 1 \end{aligned}$$

ここで、 $p_1 + p_2 + \dots + p_m = 1$ であるので、

$$E[X_i] = n p_i$$

・分散

$$\begin{aligned} E[X_i^2] &= \sum_{x_1+x_2+\dots+x_m=n} x_i^2 \{ n! / (x_1! x_2! \dots x_m!) \} p_1^{x_1} p_2^{x_2} \dots p_m^{x_m} \\ &= \sum_{x_1+x_2+\dots+x_m=n} \{ x_i(x_i-1) + x_i \} n! p_1^{x_1} p_2^{x_2} \dots p_m^{x_m} / (x_1! x_2! \dots x_m!) \\ &= \sum_{x_1+x_2+\dots+x_m=n} x_i(x_i-1) n! p_1^{x_1} p_2^{x_2} \dots p_m^{x_m} / (x_1! x_2! \dots x_m!) + \sum_{x_1+x_2+\dots+x_m=n} x_i n! p_1^{x_1} p_2^{x_2} \dots p_m^{x_m} / (x_1! x_2! \dots x_m!) \\ &= \sum_{x_1+x_2+\dots+x_m=n} x_i(x_i-1) n! p_1^{x_1} p_2^{x_2} \dots p_m^{x_m} / (x_1! x_2! \dots x_m!) + n p_i \\ &= \sum_{x_1+x_2+\dots+x_m=n} x_i(x_i-1) n(n-1)(n-2)! p_1^{x_1} p_2^{x_2} \dots p_i^{(x_i-2)+2} \dots p_m^{x_m} / (x_1! x_2! \dots (x_i(x_i-1)(x_i-2)!) \dots x_m!) + n p_i \\ &= n(n-1) p_i^2 \sum_{x_1+x_2+\dots+x_m=n-2} (n-2)! p_1^{x_1} p_2^{x_2} \dots p_{i-1}^{x_{i-1}} p_i^{x_i-2} p_{i+1}^{x_{i+1}} \dots p_m^{x_m} / (x_1! x_2! \dots x_{i-1}! (x_i-2)! x_{i+1}! \dots x_m!) + n p_i \\ &= n(n-1) p_i^2 (p_1 + p_2 + \dots + p_m)^{n-2} + n p_i \quad (\text{注}) x_1 + x_2 + \dots + x_{i-1} + (x_i - 2) + x_{i+1} + \dots + x_m = (x_1 + x_2 + \dots + x_m) - 2 = n - 2 \end{aligned}$$

ここで、 $p_1 + p_2 + \dots + p_m = 1$ であるので、

$$E[X_i^2] = n(n-1) p_i^2 + n p_i$$

従って、

$$\begin{aligned} V[X_i] &= E[X_i^2] - E[X_i]^2 \\ &= n(n-1) p_i^2 + n p_i - (n p_i)^2 = n^2 p_i^2 - n p_i^2 + n p_i - n^2 p_i^2 = n p_i (1 - p_i) \end{aligned}$$

なお、ある事象が起こり、その他複数の事象は起こらないとひとまとめにして考えれば、多項分布を二項分布と捉えることができる。よって、多項分布の個々の確率変数の平均と分散は、二項分布と同じになる。

また、カテゴリカル分布における試行を n 回繰り返したとき、それぞれの事象の出現回数が従う分布が多項分布になることから、カテゴリカル分布の期待値、分散を n 倍したものが多項分布の期待値と分散になっている。

付録30. 期待値と分散の導出(負の二項分布)

負の二項分布のモーメント母関数 $M_x(t)$ (付録21)を求める。

負の値まで拡張した二項係数を準備する。

$$\begin{aligned} {}_{-k}C_x &= (-k)(-k-1)\cdots(-k-(x-1)) \frac{(-k-x)!}{\{(-k-x)!x!\}} = (-k)(-k-1)\cdots(-k-(x-1)) / x! \\ &= (-1)^x k(k+1)\cdots(k+x-1) / x! = (-1)^x (k+x-1)\cdots(k+1)k \frac{(k-1)!}{\{(k-1)!x!\}} \\ &= (-1)^x (k+x-1)! / \{((k+x-1)-x)!x!\} = (-1)^x {}_{k+x-1}C_x \end{aligned}$$

この負の二項係数を使って、モーメント母関数を表す。

$q = 1 - p$ とおく。また、下式のなかで、 $1 = 1^{-k-x}$ は、二項定理を使うための式変形である。

$$M_x(t) = E[e^{tx}] = \sum_{x=-\infty}^{\infty} e^{tx} \{ {}_{k+x-1}C_x p^k q^x \} = \sum_{x=-\infty}^{\infty} e^{tx} \{ (-1)^x {}_{-k}C_x p^k q^x \} = p^k \sum_{x=-\infty}^{\infty} {}_{-k}C_x (-q e^t)^x = p^k \sum_{x=-\infty}^{\infty} {}_{-k}C_x (-q e^t)^x 1^{-k-x}$$

二項定理(付録1)より、

$$M_x(t) = p^k (-q e^t + 1)^{-k} = p^k (1 - q e^t)^{-k}$$

このモーメント母関数を使用して、二項分布の期待値と分散を求める。

なお、 $p + q = 1$ の関係を使う。

・期待値

$$\begin{aligned} E[X] &= d/dt \{ M_x(t) \} \Big|_{t=0} = d/dt \{ p^k (1 - q e^t)^{-k} \} \Big|_{t=0} = \{ p^k (-k) (1 - q e^t)^{-k-1} (-q e^t) \} \Big|_{t=0} = \{ k q p^k e^t (1 - q e^t)^{-k-1} \} \Big|_{t=0} \\ &= k q p^k e^0 (1 - q e^0)^{-k-1} = k q p^k (1 - q)^{-k-1} = k q p^k p^{-k-1} = k q p^{-1} = k q / p \end{aligned}$$

・分散

$$\begin{aligned} E[X^2] &= d^2/dt^2 \{ M_x(t) \} \Big|_{t=0} = d^2/dt^2 \{ p^k (1 - q e^t)^{-k} \} \Big|_{t=0} = d/dt \{ k q p^k e^t (1 - q e^t)^{-k-1} \} \Big|_{t=0} \\ &= \{ (k q p^k \exp(t))' (1 - q \exp(t))^{-(k+1)} + k q p^k \exp(t) ((1 - q \exp(t))^{-(k+1)})' \} \Big|_{t=0} \\ &= \{ k q p^k \exp(t) (1 - q \exp(t))^{-(k+1)} + k q p^k \exp(t) (-(k+1)) (1 - q \exp(t))^{-(k+1)-1} (-q \exp(t)) \} \Big|_{t=0} \\ &= \{ k q p^k \exp(t) (1 - q \exp(t))^{-(k+1)} + k(k+1) q^2 p^k \exp(2t) (1 - q \exp(t))^{-(k+2)} \} \Big|_{t=0} \\ &= k q p^k \exp(0) (1 - q \exp(0))^{-(k+1)} + k(k+1) q^2 p^k \exp(2 \cdot 0) (1 - q \exp(0))^{-(k+2)} \\ &= k q p^k (1 - q)^{-(k+1)} + k(k+1) q^2 p^k (1 - q)^{-(k+2)} \\ &= k q p^k p^{-(k+1)} + k(k+1) q^2 p^k p^{-(k+2)} \\ &= k q p^{-1} + k(k+1) q^2 p^{-2} \end{aligned}$$

よって、

$$\begin{aligned} V[X] &= E[X^2] - E[X]^2 \\ &= (k q p^{-1} + k(k+1) q^2 p^{-2}) - (k q / p)^2 \\ &= (1/p^2) \{ k q p + k(k+1) q^2 - (k q)^2 \} \\ &= (1/p^2) (k q p + k q^2) = (1/p^2) k q (p + q) \quad (\text{注}) p + q = 1 \\ &= k q / p^2 \end{aligned}$$

付録31. 期待値と分散の導出(幾何分布)

・期待値

$$E[X] = \sum_{x=1}^{\infty} x P(X=x) = \sum_{x=1}^{\infty} x (1-p)^{x-1} p = p \sum_{x=1}^{\infty} x (1-p)^{x-1}$$

ところで、 $1/(1-z)$ をマクローリン展開(付録3)すると、

$$1/(1-z) = 1 + z + z^2 + \cdots = \sum_{x=0}^{\infty} z^x$$

両辺を z で微分すると、

$$1/(1-z)^2 = \sum_{x=1}^{\infty} x z^{x-1}$$

ここで、 $z = 1-p$ とおくと、

$$1/p^2 = \sum_{x=1}^{\infty} x (1-p)^{x-1}$$

ゆえに、

$$E[X] = p (1/p^2) = 1/p$$

・分散

$$E[X^2] = \sum_{x=1}^{\infty} x^2 P(X=x) = \sum_{x=1}^{\infty} x^2 (1-p)^{x-1} p = p \sum_{x=1}^{\infty} x^2 (1-p)^{x-1}$$

ところで、 $1/(1-z)$ をマクローリン展開すると、

$$1/(1-z) = 1 + z + z^2 + \cdots = \sum_{x=0}^{\infty} z^x$$

両辺を z で微分すると、

$$1/(1-z)^2 = \sum_{x=1}^{\infty} x z^{x-1}$$

両辺に z をかけると、

$$z/(1-z)^2 = \sum_{x=1}^{\infty} x z^x$$

さらに、両辺を z で微分すると、

$$(1-z)^{-2} + 2z(1-z)^{-3} = \sum_{x=1}^{\infty} x^2 z^{x-1}$$

ここで、 $z = 1-p$ とおくと、

$$1/p^2 + 2(1-p)/p^3 = \sum_{x=1}^{\infty} x^2 (1-p)^{x-1} = E[X^2]/p$$

よって、

$$E[X^2] = p \{1/p^2 + 2(1-p)/p^3\} = p/p^2 + 2(1-p)/p^2 = (p + 2(1-p))/p^2 = (2-p)/p^2$$

従って、

$$V[X] = E[X^2] - E[X]^2$$

$$= (2-p)/p^2 - (1/p)^2 = (1-p)/p^2$$

付録32. 期待値と分散の導出(超幾何分布)

・期待値

$$E[X] = \sum_{x=0}^n x P(X=x) = \sum_{x=0}^n x \frac{{}_K C_x {}_{N-K} C_{n-x}}{{}_N C_n}$$

この式で、 $x=0$ で $E[X]=0$ となるので、上式の和は $x=1 \sim n$ の和(次式)で表される。

$$E[X] = \sum_{x=1}^n x \frac{{}_K C_x {}_{N-K} C_{n-x}}{{}_N C_n}$$

ここで、 ${}_K C_x / {}_N C_n$ の式を式変形する。

$$\begin{aligned} {}_K C_x / {}_N C_n &= \{K! / ((K-x)! x!)\} / \{N! / ((N-n)! n!)\} \\ &= \{K(K-1)! / (((K-1)-(x-1))! x!)\} / \{N(N-1)! / (((N-1)-(n-1))! n!)\} \\ &= \{K(K-1)! / (((K-1)-(x-1))! x!)\} / \{N(N-1)! / (((N-1)-(n-1))! n!)\} \\ &= (n/x) (K/N) \{K(K-1)! / (((K-1)-(x-1))! (x-1)!)\} / \{(N-1)! / (((N-1)-(n-1))! (n-1)!)\} \\ &= (n/x) (K/N) {}_{K-1} C_{x-1} / {}_{N-1} C_{n-1} \end{aligned}$$

$$\text{よって、} E[X] = \sum_{x=1}^n x (n/x) (K/N) {}_{K-1} C_{x-1} {}_{N-K} C_{n-x} / {}_{N-1} C_{n-1} = n (K/N) \sum_{x=1}^n {}_{K-1} C_{x-1} {}_{N-K} C_{n-x} / {}_{N-1} C_{n-1}$$

超幾何関数の確率質量関数 ${}_K C_x {}_{N-K} C_{n-x} / {}_N C_n$ において、 K を $K-1$ に、 x を $x-1$ に、 N を $N-1$ に、 n を $n-1$ に置き換えて考えると、 ${}_{K-1} C_{x-1} {}_{N-K} C_{n-x} / {}_{N-1} C_{n-1}$ は超幾何関数の確率質量関数である。

よって、その総和は、 $\sum_{x=1}^n {}_{K-1} C_{x-1} {}_{N-K} C_{n-x} / {}_{N-1} C_{n-1} = 1$ である。

従って、 $E[X] = nK/N$

・分散

$$V[X] = E[X^2] - E[X]^2 = E[X^2 - X + X] - E[X]^2 = E[X^2 - X] + E[X] - E[X]^2 = E[X(X-1)] + E[X] - E[X]^2$$

$E[X]$ は上記で求まっているので、 $E[X(X-1)]$ が分かれば $V[X]$ を導出できる。

$$E[X(X-1)] = \sum_{x=0}^n x(x-1) \frac{{}_K C_x {}_{N-K} C_{n-x}}{{}_N C_n} = \sum_{x=1}^n x(x-1) \frac{{}_K C_x {}_{N-K} C_{n-x}}{{}_N C_n}$$

ここで、 ${}_K C_x / {}_N C_n$ の式を式変形する。

$$\begin{aligned} {}_K C_x / {}_N C_n &= \{K! / ((K-x)! x!)\} / \{N! / ((N-n)! n!)\} \\ &= \{K(K-1)(K-2)! / (((K-2)-(x-2))! x(x-1)!)\} / \{N(N-1)(N-2)! / (((N-2)-(n-2))! n(n-1)(n-2)!)\} \\ &= \{n(n-1)/(x(x-1))\} \{K(K-1)/(N(N-1))\} \{K(K-2)! / (((K-2)-(x-2))! (x-2)!)\} / \{(N-2)! / (((N-2)-(n-2))! (n-2)!)\} \\ &= \{n(n-1)/(x(x-1))\} \{K(K-1)/(N(N-1))\} {}_{K-2} C_{x-2} / {}_{N-2} C_{n-2} \end{aligned}$$

$$\begin{aligned} \text{よって、} E[X(X-1)] &= \sum_{x=1}^n x(x-1) \{n(n-1)/(x(x-1))\} \{K(K-1)/(N(N-1))\} {}_{K-2} C_{x-2} {}_{N-K} C_{n-x} / {}_{N-2} C_{n-2} \\ &= \sum_{x=1}^n n(n-1) K(K-1) / (N(N-1)) {}_{K-2} C_{x-2} {}_{N-K} C_{n-x} / {}_{N-2} C_{n-2} = n(n-1) K(K-1) / (N(N-1)) \sum_{x=1}^n {}_{K-2} C_{x-2} {}_{N-K} C_{n-x} / {}_{N-2} C_{n-2} \end{aligned}$$

この式で、上記と同じく、総和の項は超幾何分布の確率質量関数の総和であるので、 $\sum_{x=1}^n {}_{K-2} C_{x-2} {}_{N-K} C_{n-x} / {}_{N-2} C_{n-2} = 1$ である。

よって、 $E[X(X-1)] = n(n-1) K(K-1) / (N(N-1))$

$$\begin{aligned} \text{従って、} V[X] &= E[X(X-1)] + E[X] - E[X]^2 = n(n-1) K(K-1) / (N(N-1)) + nK/N - (nK/N)^2 \\ &= \{Nn(n-1)K(K-1) + nKN(N-1) - n^2K^2(N-1)\} / \{N^2(N-1)\} = \{-NnK^2 - Nn^2K + nKN^2 + n^2K^2\} / \{N^2(N-1)\} \\ &= nK(-NK - Nn + N^2 + nK) / \{N^2(N-1)\} = nK(N-K)(N-n) / \{N^2(N-1)\} \end{aligned}$$

付録33. 期待値と分散の導出(ポアソン分布)

ポアソン分布のモーメント母関数 $M_x(t)$ を求める。

$$M_x(t) = E[e^{tx}] = \sum_{x=0}^{\infty} e^{tx} P(X=x) = \sum_{x=0}^{\infty} e^{tx} \{ (\lambda^x / x!) e^{-\lambda} \} = e^{-\lambda} \sum_{x=0}^{\infty} \{ (\lambda e^t)^x / x! \}$$

ここで、 e^t のマクローリン展開(付録3)は、 $e^t = \sum_{i=0}^{\infty} t^i / i!$ であることを使うと、

$$e^{\lambda e^t} = \sum_{x=0}^{\infty} (\lambda e^t)^x / x! \text{ となる。}$$

よって、モーメント母関数は、

$$M_x(t) = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

このモーメント母関数を使用して、ポアソン分布の平均と分散を求める(下段の関数の微分を参考に)。

・期待値

$$E[X] = d/dt \{ M_x(t) \} \Big|_{t=0} = d/dt \{ e^{\lambda(e^t - 1)} \} \Big|_{t=0} = e^{\lambda(e^t - 1)} \lambda e^t \Big|_{t=0} = e^0 \lambda e^0 = \lambda$$

・分散

$$\begin{aligned} E[X^2] &= d^2/dt^2 \{ M_x(t) \} \Big|_{t=0} = d^2/dt^2 \{ e^{\lambda(e^t - 1)} \} \Big|_{t=0} = d/dt \{ e^{\lambda(e^t - 1)} \lambda e^t \} \Big|_{t=0} = \{ e^{\lambda(e^t - 1)} (\lambda e^t)' + \{ e^{\lambda(e^t - 1)} \}' (\lambda e^t) \} \Big|_{t=0} \\ &= (e^{\lambda(e^t - 1)} \lambda e^t + \lambda e^t e^{\lambda(e^t - 1)} \lambda e^t) \Big|_{t=0} = \{ e^{\lambda(e^t - 1)} \lambda e^t (1 + \lambda e^t) \} \Big|_{t=0} = e^0 \lambda e^0 (1 + \lambda e^0) = \lambda(1 + \lambda) \end{aligned}$$

よって、

$$\begin{aligned} V[X] &= E[X^2] - E[X]^2 \\ &= \lambda(1 + \lambda) - \lambda^2 = \lambda \end{aligned}$$

(参考) 関数の微分

$$\{ e^{g(t)} \}' = e^{g(t)} \{ g(t) \}'$$

$$\{ f(t) g(t) \}' = f'(t) g(t) + f(t) g'(t)$$

付録34. 期待値と分散の導出(指数分布)

指数分布のモーメント母関数 $M_x(t)$ を求める。

$$M_x(t) = E[e^{tx}] = \int_0^{\infty} e^{tx} \{ \lambda e^{-\lambda x} \} dx = \lambda \int_0^{\infty} e^{(t-\lambda)x} dx = \lambda \left[\frac{1}{t-\lambda} e^{(t-\lambda)x} \right]_0^{\infty} = \lambda / (\lambda - t)$$

ここで、 t はモーメント母関数の定義より、 0 に近い値である。一方、 λ は指数分布の定義より正の値をとるパラメータである。つまり、 $t - \lambda$ は負の値になる。よって、 $x \rightarrow \infty$ で、 $e^{(t-\lambda)x} \rightarrow 0$ である。

このモーメント母関数を使用して、二項分布の期待値と分散を求める。

・期待値

$$E[X] = d/dt \{ M_x(t) \} \big|_{t=0} = d/dt \{ \lambda / (\lambda - t) \} \big|_{t=0} = \lambda / (\lambda - t)^2 \big|_{t=0} = 1 / \lambda$$

・分散

$$E[X^2] = d^2/dt^2 \{ M_x(t) \} \big|_{t=0} = d^2/dt^2 \{ \lambda / (\lambda - t) \} \big|_{t=0} = d/dt \{ \lambda / (\lambda - t)^2 \} \big|_{t=0} = -\lambda (-2 / (\lambda - t)^3) \big|_{t=0} = 2 / \lambda^2$$

よって、

$$V[X] = E[X^2] - E[X]^2$$

$$= (2 / \lambda^2) - (1 / \lambda)^2 = 1 / \lambda^2$$

付録35. 期待値と分散の導出(ワイブル分布)

・期待値

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x f(x) dx = \int_0^{\infty} x \left\{ (b/a) (x/a)^{b-1} e^{-(x/a)^b} \right\} dx \quad (0 \leq x)$$

$t = (x/a)^b$ として、置換積分を行う。 $x = a t^{1/b}$, $dx/dt = (a/b) t^{1/b-1}$, 積分範囲は $x : 0 \rightarrow \infty$ に対して、 $t : 0 \rightarrow \infty$

$$\begin{aligned} E[X] &= \int_0^{\infty} a t^{1/b} (b/a) t (x/a)^{-1} e^{-t} (a/b) t^{1/b-1} dt = \int_0^{\infty} a t^{1/b} (b/a) t (a t^{1/b}/a)^{-1} e^{-t} (a/b) t^{1/b-1} dt \\ &= a \int_0^{\infty} t^{1/b} e^{-t} dt = a \int_0^{\infty} t^{(b+1)/b-1} e^{-t} dt = a \Gamma((b+1)/b) \quad (\text{注) 付録12. ガンマ関数の定義 (付式23)より}) \end{aligned}$$

・分散

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^{\infty} x^2 f(x) dx = \int_0^{\infty} x^2 \left\{ (b/a) (x/a)^{b-1} e^{-(x/a)^b} \right\} dx \quad (0 \leq x)$$

同じく、 $t = (x/a)^b$ として、置換積分を行う。 $x = a t^{1/b}$, $dx/dt = (a/b) t^{1/b-1}$, 積分範囲は $x : 0 \rightarrow \infty$ に対して、 $t : 0 \rightarrow \infty$

$$\begin{aligned} E[X^2] &= \int_0^{\infty} (a t^{1/b})^2 (b/a) t (x/a)^{-1} e^{-t} (a/b) t^{1/b-1} dt = \int_0^{\infty} a t^{1/b} (b/a) t (a t^{1/b}/a)^{-1} e^{-t} (a/b) t^{1/b-1} dt \\ &= a^2 \int_0^{\infty} t^{2/b} e^{-t} dt = a^2 \int_0^{\infty} t^{(b+2)/b-1} e^{-t} dt = a^2 \Gamma((b+2)/b) \end{aligned}$$

よって、

$$V[X] = E[X^2] - E[X]^2$$

$$= a^2 \Gamma((b+2)/b) - a \Gamma^2((b+1)/b) = a^2 \{ \Gamma((b+2)/b) - \Gamma^2((b+1)/b) \}$$

付録36. 期待値と分散の導出(ガンマ分布)

ガンマ分布のモーメント母関数 $M_x(t)$ を求める。

$$\begin{aligned} M_x(t) &= E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} \{ (\lambda^k x^{k-1} e^{-\lambda x}) / \Gamma(k) \} dx = \int_{-\infty}^{\infty} \lambda^k x^{k-1} e^{-\lambda x + tx} / \Gamma(k) dx \\ &= (\lambda - t)^k (\lambda - t)^{-k} \int_{-\infty}^{\infty} \lambda^k x^{k-1} e^{-(\lambda - t)x} / \Gamma(k) dx = \lambda^k (\lambda - t)^{-k} \int_{-\infty}^{\infty} (\lambda - t)^k (x^{k-1} e^{-(\lambda - t)x}) / \Gamma(k) dx \end{aligned}$$

ここで、非積分関数 $(\lambda - t)^k (x^{k-1} e^{-(\lambda - t)x}) / \Gamma(k)$ は ガンマ分布の確率密度関数になっており、

$\int_{-\infty}^{\infty} (\lambda - t)^k (x^{k-1} e^{-(\lambda - t)x}) / \Gamma(k) dx = 1$ であるので、

$$M_x(t) = \lambda^k (\lambda - t)^{-k}$$

このモーメント母関数を使用して、ガンマ分布の期待値と分散を求める。

・期待値

$$E[X] = d/dt \{ M_x(t) \} \Big|_{t=0} = d/dt \{ \lambda^k (\lambda - t)^{-k} \} \Big|_{t=0} = -k \lambda^k (\lambda - t)^{-k-1} (-1) \Big|_{t=0} = k \lambda^k (\lambda - t)^{-k-1} \Big|_{t=0} = k \lambda^k (\lambda - 0)^{-k-1} = k / \lambda$$

・分散

$$\begin{aligned} E[X^2] &= d^2/dt^2 \{ M_x(t) \} \Big|_{t=0} = d^2/dt^2 \{ \lambda^k (\lambda - t)^{-k} \} \Big|_{t=0} = d/dt \{ k \lambda^k (\lambda - t)^{-k-1} \} \Big|_{t=0} \\ &= -(k+1) k \lambda^k (\lambda - t)^{-k-2} (-1) \Big|_{t=0} = (k+1) k \lambda^k (\lambda - t)^{-k-2} \Big|_{t=0} = (k+1) k \lambda^k (\lambda - 0)^{-k-2} = (k+1) k / \lambda^2 \end{aligned}$$

よって、

$$V[X] = E[X^2] - E[X]^2$$

$$= (k+1) k / \lambda^2 - (k / \lambda)^2 = k / \lambda^2$$

付録37. 期待値と分散の導出(ベータ分布)

・期待値

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x f(x) dx = \int_0^1 x \{ (x^{\alpha-1} (1-x)^{\beta-1}) / B(\alpha, \beta) \} dx = \int_0^1 (x^{\alpha} (1-x)^{\beta-1}) / B(\alpha, \beta) dx \quad (0 \leq x \leq 1)$$

ここで、 $\alpha = a - 1$ とおき、(付式31)から導かれる $B(s-1, t) = \{ (s-1+t) / (s-1) \} B(s, t)$ の関係を使用すると(注1)、

$$\begin{aligned} E[X] &= \int_0^1 (x^{a-1} (1-x)^{\beta-1}) / B(a-1, \beta) dx = \int_0^1 (x^{a-1} (1-x)^{\beta-1}) / \{ \{ (a-1+\beta) / (a-1) \} B(a, \beta) \} dx \\ &= \{ (a-1) / (a-1+\beta) \} \int_0^1 (x^{a-1} (1-x)^{\beta-1}) / B(a, \beta) dx \end{aligned}$$

ここで、非積分関数 $(x^{a-1} (1-x)^{\beta-1}) / B(a, \beta)$ ベータ分布の確率密度関数になっており、

$$\int_0^1 (x^{a-1} (1-x)^{\beta-1}) / B(a, \beta) dx = 1 \quad \text{であるので、} a = \alpha + 1 \text{ と戻せば、ベータ分布の期待値が求まる。}$$

$$E[X] = (a-1) / (a-1+\beta) = (\alpha+1-1) / (\alpha+1-1+\beta) = \alpha / (\alpha + \beta)$$

・分散

$$E[X^2] = \int_0^1 x^2 f(x) dx = \int_0^1 x^2 \{ (x^{\alpha-1} (1-x)^{\beta-1}) / B(\alpha, \beta) \} dx = \int_0^1 (x^{\alpha+1} (1-x)^{\beta-1}) / B(\alpha, \beta) dx \quad (0 \leq x \leq 1)$$

ここで、 $\alpha + 1 = a - 1$ 、つまり、 $\alpha = a - 2$ とおき、(付式31)から導かれる(注2)の式を使用する。

$$\begin{aligned} E[X^2] &= \int_0^1 (x^{a-1} (1-x)^{\beta-1}) / B(a-2, \beta) dx = \int_0^1 (x^{a-1} (1-x)^{\beta-1}) / \{ \{ (a-2+\beta)(a-1+\beta) / ((a-2)(a-1)) \} B(a, \beta) \} dx \\ &= (a-2)(a-1) / \{ (a-2+\beta)(a-1+\beta) \} \int_0^1 (x^{a-1} (1-x)^{\beta-1}) / B(a, \beta) dx \end{aligned}$$

ここで、非積分関数 $(x^{a-1} (1-x)^{\beta-1}) / B(a, \beta)$ は、ベータ分布の確率密度関数になっており、

$$\int_0^1 (x^{a-1} (1-x)^{\beta-1}) / B(a, \beta) dx = 1 \quad \text{であるので、} a = \alpha + 2 \text{ と戻せば、} E[X^2] \text{ の期待値が求まる。}$$

$$\begin{aligned} E[X^2] &= (a-2)(a-1) / \{ (a-2+\beta)(a-1+\beta) \} = (\alpha+2-2)(\alpha+2-1) / \{ (\alpha+2-2+\beta)(\alpha+2-1+\beta) \} \\ &= \alpha(\alpha+1) / \{ (\alpha+\beta)(\alpha+\beta+1) \} \end{aligned}$$

よって、

$$\begin{aligned} V[X] &= E[X^2] - E[X]^2 \\ &= \alpha(\alpha+1) / \{ (\alpha+\beta)(\alpha+\beta+1) \} - \{ \alpha / (\alpha+\beta) \}^2 = \{ \alpha(\alpha+1)(\alpha+\beta) - \alpha^2(\alpha+\beta+1) \} / \{ (\alpha+\beta)^2(\alpha+\beta+1) \} \\ &= \alpha \beta / \{ (\alpha+\beta)^2(\alpha+\beta+1) \} \end{aligned}$$

(注1)付録13. の(付式31)は、 $B(s+1, t) = \{ s / (s+t) \} B(s, t)$

この式から、 $B(s, t) = \{ (s+t) / s \} B(s+1, t)$ が求まる。ここで、 s を $s-1$ に置き換えると、

$B(s-1, t) = \{ (s-1+t) / (s-1) \} B(s, t)$ となる。

(注2)(注1)を使って、 $B(s-2, t) = \{ (s-2+t) / (s-2) \} B(s-1, t) = \{ (s-2+t) / (s-2) \} \{ (s-1+t) / (s-1) \} B(s, t)$

$$= \{ (s-2+t)(s-1+t) / ((s-2)(s-1)) \} B(s, t)$$

付録38. 期待値と分散の導出(ディリクレ分布)

・期待値

$$\begin{aligned} E[X_i] &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i C x_1^{\alpha_1-1} x_2^{\alpha_2-1} \cdots x_i^{\alpha_i-1} \cdots x_n^{\alpha_n-1} dx_1 dx_2 \cdots dx_n \\ &= C \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{\alpha_1-1} x_2^{\alpha_2-1} \cdots x_{i-1}^{\alpha_{i-1}-1} x_i^{\alpha_i} x_{i+1}^{\alpha_{i+1}-1} \cdots x_n^{\alpha_n-1} dx_1 dx_2 \cdots dx_n \end{aligned}$$

この式に、

$$(式26)より、C = \Gamma(\sum_{i=1}^n \alpha_i) / \prod_{i=1}^n \Gamma(\alpha_i) = \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n) / (\Gamma(\alpha_1) \Gamma(\alpha_2) \cdots \Gamma(\alpha_n))$$

$$\begin{aligned} (式25)より、\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{\alpha_1-1} x_2^{\alpha_2-1} \cdots x_{i-1}^{\alpha_{i-1}-1} x_i^{\alpha_i} x_{i+1}^{\alpha_{i+1}-1} \cdots x_n^{\alpha_n-1} dx_1 dx_2 \cdots dx_n \\ = \Gamma(\alpha_1) \Gamma(\alpha_2) \cdots \Gamma(\alpha_{i-1}) \Gamma(\alpha_i + 1) \Gamma(\alpha_{i+1}) \cdots \Gamma(\alpha_n) / \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{i-1} + (\alpha_i + 1) + \alpha_{i+1} + \cdots + \alpha_n) \end{aligned}$$

を代入して、 $\Gamma(s) = (s-1) \Gamma(s-1)$ (付録12)の関係を使用すると、

$$\begin{aligned} E[X_i] &= \{ \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n) / (\Gamma(\alpha_1) \Gamma(\alpha_2) \cdots \Gamma(\alpha_{i-1}) \Gamma(\alpha_i) \Gamma(\alpha_{i+1}) \cdots \Gamma(\alpha_n)) \} \\ &\quad \cdot \{ \Gamma(\alpha_1) \Gamma(\alpha_2) \cdots \Gamma(\alpha_{i-1}) \Gamma(\alpha_i + 1) \Gamma(\alpha_{i+1}) \cdots \Gamma(\alpha_n) / \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{i-1} + (\alpha_i + 1) + \alpha_{i+1} + \cdots + \alpha_n) \} \\ &= \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n) \Gamma(\alpha_i + 1) / \{ \Gamma(\alpha_i) \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{i-1} + \alpha_i + \alpha_{i+1} + \cdots + \alpha_n + 1) \} \\ &= \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n) \alpha_i \Gamma(\alpha_i) / \{ \Gamma(\alpha_i) (\alpha_1 + \alpha_2 + \cdots + \alpha_n) \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n) \} \\ &= \alpha_i / (\alpha_1 + \alpha_2 + \cdots + \alpha_n) \\ &= \alpha_i / \sum_{i=1}^n \alpha_i \end{aligned}$$

・分散

$$\begin{aligned} E[X_i^2] &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i^2 f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i^2 C x_1^{\alpha_1-1} x_2^{\alpha_2-1} \cdots x_i^{\alpha_i-1} \cdots x_n^{\alpha_n-1} dx_1 dx_2 \cdots dx_n \\ &= C \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{\alpha_1-1} x_2^{\alpha_2-1} \cdots x_{i-1}^{\alpha_{i-1}-1} x_i^{\alpha_i+1} x_{i+1}^{\alpha_{i+1}-1} \cdots x_n^{\alpha_n-1} dx_1 dx_2 \cdots dx_n \end{aligned}$$

この式に、

$$(式26)より、C = \Gamma(\sum_{i=1}^n \alpha_i) / \prod_{i=1}^n \Gamma(\alpha_i) = \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n) / (\Gamma(\alpha_1) \Gamma(\alpha_2) \cdots \Gamma(\alpha_n))$$

$$\begin{aligned} (式25)より、\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{\alpha_1-1} x_2^{\alpha_2-1} \cdots x_{i-1}^{\alpha_{i-1}-1} x_i^{\alpha_i+1} x_{i+1}^{\alpha_{i+1}-1} \cdots x_n^{\alpha_n-1} dx_1 dx_2 \cdots dx_n \\ = \Gamma(\alpha_1) \Gamma(\alpha_2) \cdots \Gamma(\alpha_{i-1}) \Gamma(\alpha_i + 2) \Gamma(\alpha_{i+1}) \cdots \Gamma(\alpha_n) / \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{i-1} + (\alpha_i + 2) + \alpha_{i+1} + \cdots + \alpha_n) \end{aligned}$$

を代入して、 $\Gamma(s+2) = \Gamma(s+1+1) = (s+1) \Gamma(s+1) = s(s+1) \Gamma(s)$ (付録12)の関係を使用すると、

$$\begin{aligned} E[X_i^2] &= \{ \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n) / (\Gamma(\alpha_1) \Gamma(\alpha_2) \cdots \Gamma(\alpha_{i-1}) \Gamma(\alpha_i) \Gamma(\alpha_{i+1}) \cdots \Gamma(\alpha_n)) \} \\ &\quad \cdot \{ \Gamma(\alpha_1) \Gamma(\alpha_2) \cdots \Gamma(\alpha_{i-1}) \Gamma(\alpha_i + 2) \Gamma(\alpha_{i+1}) \cdots \Gamma(\alpha_n) / \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{i-1} + (\alpha_i + 2) + \alpha_{i+1} + \cdots + \alpha_n) \} \\ &= \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n) \Gamma(\alpha_i + 2) / \{ \Gamma(\alpha_i) \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_{i-1} + \alpha_i + \alpha_{i+1} + \cdots + \alpha_n + 2) \} \\ &= \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n) \alpha_i (\alpha_i + 1) \Gamma(\alpha_i) / \{ \Gamma(\alpha_i) (\alpha_1 + \alpha_2 + \cdots + \alpha_n) (\alpha_1 + \alpha_2 + \cdots + \alpha_n + 1) \Gamma(\alpha_1 + \alpha_2 + \cdots + \alpha_n) \} \\ &= \alpha_i (\alpha_i + 1) / \{ (\alpha_1 + \alpha_2 + \cdots + \alpha_n) (\alpha_1 + \alpha_2 + \cdots + \alpha_n + 1) \} \\ &= \alpha_i (\alpha_i + 1) / \{ (\sum_{i=1}^n \alpha_i) (\sum_{i=1}^n \alpha_i + 1) \} \end{aligned}$$

よって、

$$\begin{aligned} V[X_i] &= E[X_i^2] - E[X_i]^2 \\ &= \alpha_i (\alpha_i + 1) / \{ (\sum_{i=1}^n \alpha_i) (\sum_{i=1}^n \alpha_i + 1) \} - (\alpha_i / \sum_{i=1}^n \alpha_i)^2 = \alpha_i (\sum_{i=1}^n \alpha_i - \alpha_i) / \{ (\sum_{i=1}^n \alpha_i)^2 (\sum_{i=1}^n \alpha_i + 1) \} \end{aligned}$$

付録39. 期待値と分散の導出(正規分布)

正規分布のモーメント母関数 $M_x(t)$ を求める。

$$\begin{aligned} M_x(t) &= E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} \{ 1/(2\pi\sigma^2)^{1/2} e^{-(x-\mu)^2/(2\sigma^2)} \} dx = \int_{-\infty}^{\infty} 1/(2\pi\sigma^2)^{1/2} e^{(1/(-2\sigma^2))(x^2 - 2\mu x + \mu^2 - 2\sigma^2 t x)} dx \\ &= \int_{-\infty}^{\infty} 1/(2\pi\sigma^2)^{1/2} e^{(-1/(2\sigma^2))\{ (x - (\mu + \sigma^2 t))^2 - 2\mu\sigma^2 t - \sigma^4 t^2 \}} dx = \int_{-\infty}^{\infty} 1/(2\pi\sigma^2)^{1/2} e^{(-1/(2\sigma^2))(x - (\mu + \sigma^2 t))^2 + \mu t + (1/2)\sigma^2 t^2} dx \\ &= e^{(\mu t + \sigma^2 t^2/2)} \int_{-\infty}^{\infty} 1/(2\pi\sigma^2)^{1/2} e^{(-1/(2\sigma^2))(x - (\mu + \sigma^2 t))^2} dx \end{aligned}$$

ここで、非積分関数 $1/(2\pi\sigma^2)^{1/2} e^{(-1/(2\sigma^2))(x - (\mu + \sigma^2 t))^2}$ は 平均 $= \mu + \sigma^2 t$, 分散 σ^2 の正規分布の確率密度関数になっており、

$\int_{-\infty}^{\infty} 1/(2\pi\sigma^2)^{1/2} e^{(-1/(2\sigma^2))(x - (\mu + \sigma^2 t))^2} dx = 1$ であるので、

$$M_x(t) = e^{(\mu t + \sigma^2 t^2/2)}$$

このモーメント母関数を使用して、正規分布の期待値と分散を求める。

・期待値

$$E[X] = d/dt \{ M_x(t) \} \big|_{t=0} = d/dt \{ e^{(\mu t + \sigma^2 t^2/2)} \} \big|_{t=0} = (\mu + \sigma^2 t) e^{(\mu t + \sigma^2 t^2/2)} \big|_{t=0} = (\mu + \sigma^2 \cdot 0) e^{(\mu \cdot 0 + \sigma^2 \cdot 0^2/2)} = \mu$$

・分散

$$\begin{aligned} E[X^2] &= d^2/dt^2 \{ M_x(t) \} \big|_{t=0} = d^2/dt^2 \{ e^{(\mu t + \sigma^2 t^2/2)} \} \big|_{t=0} = d/dt \{ (\mu + \sigma^2 t) e^{(\mu t + \sigma^2 t^2/2)} \} \big|_{t=0} \\ &= \sigma^2 e^{(\mu t + \sigma^2 t^2/2)} + (\mu + \sigma^2 t)^2 e^{(\mu t + \sigma^2 t^2/2)} \big|_{t=0} = \sigma^2 e^{(\mu \cdot 0 + \sigma^2 \cdot 0^2/2)} + (\mu + \sigma^2 \cdot 0)^2 e^{(\mu \cdot 0 + \sigma^2 \cdot 0^2/2)} = \sigma^2 + \mu^2 \end{aligned}$$

よって、

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

付録40. 期待値と分散の導出(多変量正規分布)

多変量正規分布のベクトルと分散共分散行列は、多変量正規分布の導出のプロセスの中で既に確認されている。
ここでは、多変量正規分布のモーメント母関数 $M_x(t)$ を求め、期待値ベクトルと分散共分散行列を再確認する。

確率変数ベクトル U が n 次元の標準多変量正規分布 $N(0, I)$ に従うとき、 $x = Au + \mu = \Sigma^{1/2}u + \mu \quad \cdots \textcircled{1}$ という変数変換によって生み出される新しい確率変数ベクトル X は n 次元の多変量正規分布 $N(\mu, \Sigma)$ に従う。

(注) $N(\mu, \Sigma)$: μ は期待値ベクトルで各々の要素が期待値を示す。 Σ は分散共分散行列で対角要素は分散、非対角要素は共分散を示す。
 $N(0, I)$: 0 はゼロベクトルで各々の期待値 $= 0$ を示す。 I は単位行列で対角要素の分散 $= 1$, 非対角要素の共分散 $= 0$ を示す。

1) まず、標準多変量正規分布のモーメント母関数 $M_u(t)$ を求める。(注) T は転置ベクトル、転置行列を表す。 $^{-1}$ は逆行列を表す。
正規分布(1次元)のモーメント母関数は付録39に示す通り、 $M_u(t) = E[e^{tu}] = e^{(\mu t + (\sigma^2 t^2)/2)}$ である。

よって、標準正規分布のモーメント母関数は、 $\mu = 0, \sigma^2 = 1$ において、 $M_u(t) = E[e^{tu}] = e^{t^2/2}$ である。この関係を使って、

$$M_u(t) = E[e^{t^T u}] = E[e^{t_1 u_1 + \cdots + t_n u_n}] = \prod_{i=1}^n E[e^{t_i u_i}] = \prod_{i=1}^n e^{t_i^2/2} = e^{t^T t/2} \quad \cdots \textcircled{2}$$

2) 次に、多変量正規分布のモーメント母関数 $M_x(t)$ を求める。

$M_x(t)$ の式に、式 $\textcircled{1}$ を代入する。また、分散共分散行列 Σ は対称行列である($\Sigma = \Sigma^T$)。

$$M_x(t) = E[e^{t^T x}] = E[e^{t^T (\Sigma^{1/2} u + \mu)}] = e^{t^T \mu} E[e^{t^T \Sigma^{1/2} u}] = e^{t^T \mu} E[e^{t^T (\Sigma^{1/2})^T u}] = e^{t^T \mu} E[e^{(\Sigma^{1/2} t)^T u}] = e^{t^T \mu} M_u(\Sigma^{1/2} t)$$

ここで、式 $\textcircled{2}$ を代入すると、

$$\begin{aligned} M_x(t) &= e^{t^T \mu} M_u(\Sigma^{1/2} t) = e^{t^T \mu} e^{(\Sigma^{1/2} t)^T (\Sigma^{1/2} t)/2} = e^{t^T \mu} e^{(t^T (\Sigma^{1/2})^T \Sigma^{1/2} t)/2} = e^{t^T \mu} e^{(t^T \Sigma^{1/2} \Sigma^{1/2} t)/2} = e^{t^T \mu} e^{(t^T \Sigma t)/2} \\ &= e^{(t^T \mu + (t^T \Sigma t)/2)} \end{aligned}$$

このモーメント母関数 $M_x(t)$ を使用して、多変量正規分布の期待値ベクトルと分散共分散行列を求める(確認になる)。

・期待値ベクトル

$$E[X] = d/dt \{ M_x(t) \} |_{t=0} = d/dt \{ e^{(t^T \mu + (t^T \Sigma t)/2)} \} |_{t=0} = (\mu + \Sigma t) e^{(t^T \mu + (t^T \Sigma t)/2)} |_{t=0} = (\mu + \Sigma 0) e^{(0^T \mu + (0^T \Sigma 0)/2)} = \mu$$

・分散共分散行列

$$\begin{aligned} E[X^2] &= d^2/dt^2 \{ M_x(t) \} |_{t=0} = d^2/dt^2 \{ e^{(t^T \mu + (t^T \Sigma t)/2)} \} |_{t=0} = d/dt \{ (\mu + \Sigma t) e^{(t^T \mu + (t^T \Sigma t)/2)} \} |_{t=0} \\ &= \Sigma e^{(t^T \mu + (t^T \Sigma t)/2)} + (\mu + \Sigma t)^T (\mu + \Sigma t) e^{(t^T \mu + (t^T \Sigma t)/2)} |_{t=0} \\ &= \Sigma e^{(0^T \mu + (0^T \Sigma 0)/2)} + (\mu + \Sigma 0)^T (\mu + \Sigma 0) e^{(0^T \mu + (0^T \Sigma 0)/2)} = \Sigma + \mu^T \mu \end{aligned}$$

よって、分散共分散行列は、

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

付録41. 期待値と分散の導出(対数正規分布)

正規分布 $g(u)$ の確率変数を U とすると、対数正規分布 $f(x)$ の確率変数 X とする。 U と X との関係は、 $X = e^U$ である。

(付録14)の変数変換を適用する。また、(付録9)のガウス積分を使う。

・期待値

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{\infty} e^u g(u) du = \int_{-\infty}^{\infty} e^u \left\{ \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-(u-\mu)^2/(2\sigma^2)} \right\} du \\ &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{u-(u-\mu)^2/(2\sigma^2)} du = \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{-1/(2\sigma^2)(u^2 - 2(\mu+\sigma^2)u + \mu^2)} du \\ &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{-1/(2\sigma^2)((u-(\mu+\sigma^2))^2 - (\mu+\sigma^2)^2 + \mu^2)} du = \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{-1/(2\sigma^2)((u-(\mu+\sigma^2))^2 - 2\mu\sigma^2 - \sigma^4)} du \\ &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{-(u-(\mu+\sigma^2))^2/(2\sigma^2) + \mu + \sigma^2/2} du \end{aligned}$$

$t = u - (\mu + \sigma^2)$ として、置換積分を行う。よって、 $dt = du$, 積分範囲は $u : -\infty \rightarrow \infty$ に対して、 $t : -\infty \rightarrow \infty$

$$\begin{aligned} E[X] &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{-t^2/(2\sigma^2) + \mu + \sigma^2/2} dt = \left(e^{\mu + \sigma^2/2} / (2\pi\sigma^2)^{1/2} \right) \int_{-\infty}^{\infty} e^{-t^2/(2\sigma^2)} dt = \left(e^{\mu + \sigma^2/2} / (2\pi\sigma^2)^{1/2} \right) (2\pi\sigma^2)^{1/2} \\ &= e^{\mu + \sigma^2/2} \end{aligned}$$

・分散

$X^2 = (e^u)^2 = e^{2u}$ である。

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{-\infty}^{\infty} e^{2u} g(u) du = \int_{-\infty}^{\infty} e^{2u} \left\{ \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-(u-\mu)^2/(2\sigma^2)} \right\} du \\ &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{2u-(u-\mu)^2/(2\sigma^2)} du = \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{-1/(2\sigma^2)(u^2 - 2(\mu+2\sigma^2)u + \mu^2)} du \\ &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{-1/(2\sigma^2)((u-(\mu+2\sigma^2))^2 - (\mu+2\sigma^2)^2 + \mu^2)} du = \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{-1/(2\sigma^2)((u-(\mu+2\sigma^2))^2 - 4\mu\sigma^2 - 4\sigma^4)} du \\ &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{-(u-(\mu+2\sigma^2))^2/(2\sigma^2) + 2\mu + 2\sigma^2} du \end{aligned}$$

$t = u - (\mu + 2\sigma^2)$ として、置換積分を行う。よって、 $dt = du$, 積分範囲は $u : -\infty \rightarrow \infty$ に対して、 $t : -\infty \rightarrow \infty$

$$\begin{aligned} E[X^2] &= \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} e^{-t^2/(2\sigma^2) + 2\mu + 2\sigma^2} dt = \left(e^{2\mu + 2\sigma^2} / (2\pi\sigma^2)^{1/2} \right) \int_{-\infty}^{\infty} e^{-t^2/(2\sigma^2)} dt = \left(e^{2\mu + 2\sigma^2} / (2\pi\sigma^2)^{1/2} \right) (2\pi\sigma^2)^{1/2} \\ &= e^{2\mu + 2\sigma^2} \end{aligned}$$

よって、

$$\begin{aligned} V[X] &= E[X^2] - E[X]^2 \\ &= e^{2\mu + 2\sigma^2} - (e^{\mu + \sigma^2/2})^2 = e^{2\mu + 2\sigma^2} - (e^{2\mu + \sigma^2}) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1) \end{aligned}$$

付録42. 平均と分散の導出(カイ2乗分布)

カイ2乗分布 (χ^2 乗分布) のモーメント母関数 $M_x(t)$ を求める。

$$\begin{aligned} M_x(t) &= E[e^{tx}] = \int_0^{\infty} e^{tx} \{ (x^{(n/2-1)} e^{-x/2}) / (\Gamma(n/2) 2^{n/2}) \} dx = \int_0^{\infty} (x^{(n/2-1)} e^{-(1-2t)x/2}) / (\Gamma(n/2) 2^{n/2}) dx \\ &= \int_0^{\infty} (x^{(n/2-1)} e^{-(1-2t)x/2}) / \{ \Gamma(n/2) (1-2t)^{n/2} (2/(1-2t))^{n/2} \} dx = (1-2t)^{-n/2} \int_0^{\infty} (x^{(n/2-1)} e^{-(1-2t)x/2}) / \{ \Gamma(n/2) (2/(1-2t))^{n/2} \} dx \end{aligned}$$

ここで、非積分関数を計算する。

$$\eta = 2/(1-2t) \text{ とおくと、 } (x^{(n/2-1)} e^{-(1-2t)x/2}) / \{ \Gamma(n/2) (2/(1-2t))^{n/2} \} = (x^{(n/2-1)} e^{-x/\eta}) / \{ \Gamma(n/2) \eta^{n/2} \}$$

ここで、 n を $2n$, η を $1/\lambda$ に置き換える。

$$(x^{(n/2-1)} e^{-x/\eta}) / \{ \Gamma(n/2) \eta^{n/2} \} = (x^{(n-1)} e^{-\lambda x}) / \{ \Gamma(n) \lambda^{-n} \} = \lambda^n x^{(n-1)} e^{-\lambda x} / \Gamma(n)$$

この非積分関数は、ガンマ分布の確率密度関数 $\lambda^n x^{(n-1)} e^{-\lambda x} / \Gamma(n)$ になっており、 $\int_0^{\infty} \lambda^n x^{(n-1)} e^{-\lambda x} / \Gamma(n) dx = 1$ である。

よって、 $\int_0^{\infty} (x^{(n/2-1)} e^{-(1-2t)x/2}) / \{ \Gamma(n/2) (2/(1-2t))^{n/2} \} dx = 1$ となる。

従って、カイ2乗分布のモーメント母関数が次のように求まる。

$$M_x(t) = (1-2t)^{-n/2}$$

このモーメント母関数を使用して、カイ2乗分布の期待値と分散を求める。

・平均

$$E[X] = d/dt \{ M_x(t) \} \big|_{t=0} = d/dt \{ (1-2t)^{-n/2} \} \big|_{t=0} = (-n/2) (1-2t)^{-n/2-1} (-2) \big|_{t=0} = n (1-2t)^{-n/2-1} \big|_{t=0} = n (1-2 \cdot 0)^{-n/2-1} = \mathbf{n}$$

・分散

$$\begin{aligned} E[X^2] &= d^2/dt^2 \{ M_x(t) \} \big|_{t=0} = d^2/dt^2 \{ (1-2t)^{-n/2} \} \big|_{t=0} = d/dt \{ n (1-2t)^{-n/2-1} \} \big|_{t=0} = n (-n/2-1) (1-2t)^{-n/2-2} (-2) \big|_{t=0} \\ &= \{ n (-n/2-1) (1-2 \cdot 0)^{-n/2-2} (-2) \} = n (n+2) \end{aligned}$$

よって、

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

付録43. 期待値と分散の導出(F分布)

・期待値

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x f(x) dx = \int_0^{\infty} x \{ B(m/2, n/2) \}^{-1} (m/n)^{m/2} x^{(m-2)/2} \{ 1 + (m/n)x \}^{-(m+n)/2} dx \quad (\text{確率変数 } X \geq 0)$$

$$= \{ B(m/2, n/2) \}^{-1} (m/n)^{m/2} \int_0^{\infty} x^{m/2} \{ 1 + (m/n)x \}^{-(m+n)/2} dx$$

$t = (m/n)x$ として、置換積分を行う。 $x = (n/m)t$, $dx/dt = n/m$, 積分範囲は $x: 0 \rightarrow \infty$ に対して、 $t: 0 \rightarrow \infty$

$$E[X] = \{ B(m/2, n/2) \}^{-1} (m/n)^{m/2} \int_0^{\infty} \{ (n/m)t \}^{m/2} (1+t)^{-(m+n)/2} (n/m) dt$$

$$= (n/m) \{ B(m/2, n/2) \}^{-1} \int_0^{\infty} t^{m/2} (1+t)^{-(m+n)/2} dt = (n/m) \{ B(m/2, n/2) \}^{-1} \int_0^{\infty} t^{[(m/2)+1]-1} (1+t)^{-[(m/2)+1+(n/2-1)]} dt$$

ここで、ベータ関数(付録13)の(付式38)を使用して、

$$E[X] = (n/m) \{ B(m/2, n/2) \}^{-1} B(m/2+1, n/2-1)$$

ついで、ベータ関数(付録13)の(付式31)と(付式32)を使用して、

$$E[X] = (n/m) \{ B(m/2, n/2) \}^{-1} \{ (m/2)/(m/2+n/2-1) \} B(m/2, n/2-1) = (n/m) \{ B(m/2, n/2) \}^{-1} \{ m/(m+n-2) \} B(m/2, n/2-1)$$

$$= (n/m) \{ B(m/2, n/2) \}^{-1} \{ m/(m+n-2) \} \{ (m/2+n/2-1)/(n/2-1) \} \{ B(m/2, n/2) \}$$

$$= (n/m) \{ B(m/2, n/2) \}^{-1} \{ m/(m+n-2) \} \{ (m+n-2)/(n-2) \} B(m/2, n/2) = n/(n-2) \quad (n > 2)$$

・分散

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^{\infty} x^2 f(x) dx = \int_0^{\infty} x^2 \{ B(m/2, n/2) \}^{-1} (m/n)^{m/2} x^{(m-2)/2} \{ 1 + (m/n)x \}^{-(m+n)/2} dx$$

$$= \{ B(m/2, n/2) \}^{-1} (m/n)^{m/2} \int_0^{\infty} \{ 1 + (m/n)x \}^{-(m+n)/2} x^{m/2+1} dx$$

$t = (m/n)x$ として、置換積分を行う。 $x = (n/m)t$, $dx/dt = n/m$, 積分範囲は $x: 0 \rightarrow \infty$ に対して、 $t: 0 \rightarrow \infty$

$$E[X^2] = \{ B(m/2, n/2) \}^{-1} (m/n)^{m/2} \int_0^{\infty} (1+t)^{-(m+n)/2} \{ (n/m)t \}^{m/2+1} (n/m) dt$$

$$= (n/m)^2 \{ B(m/2, n/2) \}^{-1} \int_0^{\infty} (1+t)^{-(m+n)/2} t^{m/2+1} dt = (n/m)^2 \{ B(m/2, n/2) \}^{-1} \int_0^{\infty} t^{[(m/2)+2]-1} (1+t)^{-[(m/2)+2+(n/2-2)]} dt$$

ここで、ベータ関数(付録13)の(付式38)を使用して、

$$E[X^2] = (n/m)^2 \{ B(m/2, n/2) \}^{-1} B(m/2+2, n/2-2) \quad (n > 4)$$

ついで、ベータ関数(付録13)の(付式31)と(付式32)を使用して、

$$E[X^2] = (n/m)^2 \{ B(m/2, n/2) \}^{-1} B(m/2+2, n/2-2)$$

$$= (n/m)^2 \{ B(m/2, n/2) \}^{-1} \{ (m/2+1)/(m/2+1+n/2-2) \} B(m/2+1, n/2-2)$$

$$= (n/m)^2 \{ B(m/2, n/2) \}^{-1} \{ (m+2)/(m+n-2) \} \{ (m/2)/(m/2+n/2-2) \} B(m/2, n/2-2)$$

$$= (n/m)^2 \{ B(m/2, n/2) \}^{-1} \{ (m+2)/(m+n-2) \} \{ m/(m+n-4) \} \{ (m/2+n/2-2)/(n/2-2) \} B(m/2, n/2-2)$$

$$= (n/m)^2 \{ B(m/2, n/2) \}^{-1} \{ (m+2)/(m+n-2) \} \{ m/(m+n-4) \} \{ (m+n-4)/(n-4) \} \{ (m/2+n/2-1)/(n/2-1) \} B(m/2, n/2)$$

$$= (n/m)^2 \{ B(m/2, n/2) \}^{-1} \{ (m+2)/(m+n-2) \} \{ m/(m+n-4) \} \{ (m+n-4)/(n-4) \} \{ (m+n-2)/(n-2) \} B(m/2, n/2)$$

$$= \{ n^2(m+2) \} / \{ m(n-2)(n-4) \}$$

$$\text{よって、} V[X] = E[X^2] - E[X]^2 = \{ n^2(m+2) \} / \{ m(n-2)(n-4) \} - \{ n/(n-2) \}^2 = \{ 2n^2(m+n-2) \} / \{ m(n-2)^2(n-4) \} \quad (n > 4)$$

付録44. 期待値と分散の導出(t 分布)

・期待値

t 分布の確率密度関数 $f(x)$ は偶関数であるので、確率変数との積である $x f(x)$ は奇関数になる。

奇関数は、積分区間 $-\infty$ から ∞ で積分すると 0 になる。

つまり、 $E[X] = \int_{-\infty}^{\infty} x f(x) dx = 0$

・分散

$x^2 f(x)$ は偶関数であるので、積分区間 $-\infty$ から ∞ の積分の値は、積分区間 0 から ∞ の積分の値の 2 倍になる。

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) dx = 2 \int_0^{\infty} x^2 f(x) dx = 2 \{ n^{1/2} B(n/2, 1/2) \} \int_0^{\infty} x^2 (1 + x^2/n)^{-(n+1)/2} dx$$

$1/t = 1 + x^2/n$ として、置換積分を行う。 $t = n/(n + x^2)$, $x = n^{1/2} (1/t - 1)^{1/2}$, $dx/dt = 1/2 n^{1/2} (1/t - 1)^{-1/2} (-t^{-2})$

積分範囲は $x : 0 \rightarrow \infty$ に対して、 $t : 1 \rightarrow 0$

$$\begin{aligned} E[X^2] &= 2 \{ n^{1/2} B(n/2, 1/2) \} \int_1^0 \{ n^{1/2} (1/t - 1)^{1/2} \}^2 \{ 1 + (n^{1/2} (1/t - 1)^{1/2})^2/n \}^{-(n+1)/2} 1/2 n^{1/2} (1/t - 1)^{-1/2} (-t^{-2}) dt \\ &= \{ 1/B(n/2, 1/2) \} n^{-1/2+1+1/2} \int_0^1 (1/t - 1)^{1-1/2} (1/t)^{-(n+1)/2} t^{-2} dt \\ &= \{ 1/B(n/2, 1/2) \} n \int_0^1 \{ (1-t)/t \}^{1/2} t^{(n+1)/2-2} dt = \{ 1/B(n/2, 1/2) \} n \int_0^1 (1-t)^{1/2} t^{(n+1)/2-2-1/2} dt \\ &= \{ 1/B(n/2, 1/2) \} n \int_0^1 (1-t)^{1/2} t^{n/2-2} dt = \{ 1/B(n/2, 1/2) \} n \int_0^1 t^{(n/2-1)-1} (1-t)^{3/2-1} dt \\ &= \{ n/B(n/2, 1/2) \} B(n/2 - 1, 3/2) \end{aligned}$$

ここで、ベータ関数(付録13)の(付式36)を利用して、

$$\begin{aligned} E[X^2] &= n \{ \Gamma((n/2 + 1/2) / (\Gamma(n/2) \Gamma(1/2))) \} \{ \Gamma(n/2 - 1) \Gamma(3/2) / \Gamma(n/2 - 1 + 3/2) \} \\ &= n \{ \Gamma((n+1)/2) \Gamma(n/2 - 1) \Gamma(3/2) \} / \{ \Gamma(n/2) \Gamma(1/2) \Gamma((n+1)/2) \} \\ &= n \{ \Gamma(n/2 - 1) \Gamma(3/2) \} / \{ \Gamma(n/2) \Gamma(1/2) \} \end{aligned}$$

ここで、ガンマ関数(付録12)の(付式25)を利用して、

$$E[X^2] = n \{ \Gamma(n/2 - 1)^{1/2} \Gamma(1/2) \} / \{ (n/2 - 1) \Gamma(n/2 - 1) \Gamma(1/2) \} = n / \{ 2(n/2 - 1) \} = n / (n - 2)$$

よって、

$$V[X] = E[X^2] - E[X]^2 = n / (n - 2) - 0 = n / (n - 2)$$

付録45. 期待値と分散の導出(コーシー分布)

コーシー分布には期待値と分散が存在しないことを、標準コーシー分布で確認する。

- ・期待値

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{\infty} x ((1/\pi)/(x^2 + 1)) dx = (1/\pi) \int_{-\infty}^{\infty} x/(x^2 + 1) dx \\ &= \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} (1/\pi) \int_a^b x/(x^2 + 1) dx = \lim_{b \rightarrow \infty} (1/2) \log(1 + b^2) - \lim_{a \rightarrow -\infty} (1/2) \log(1 + a^2) = \infty - \infty \end{aligned}$$

よって、期待値は存在しない。

- ・分散

期待値が存在しないので、分散も存在しない。

付録46. 中心極限定理の導出

標本平均の確率分布が標本のサンプルサイズを大きくしたときに近づく確率分布の特性関数と、正規分布の特性関数が一致することを確認し、これをもって両者の確率密度関数が一致する、すなわち、標本のサンプルサイズを大きくすれば標本平均の確率分布が正規分布に近づくことを、標準正規分布を使って確認する。(注) 以下、 $\sum_{i=1}^n$ は省略して Σ で表す。

(1) 標本平均の確率分布の平均と分散

母平均 $E[X_i] = \mu$ 、母分散 $V[X_i] = \sigma^2$ の母集団からサンプリングして作る標本平均 $\bar{X} = (\Sigma X_i) / n$ の確率分布の平均と分散を求める。

$$E[\bar{X}] = E[(\Sigma X_i) / n] = (1/n) E[\Sigma X_i] = (1/n) \Sigma E[X_i] = (1/n) (n \mu) = \mu$$

$$V[\bar{X}] = V[(\Sigma X_i) / n] = (1/n^2) V[\Sigma X_i] = (1/n^2) \Sigma V[X_i] = (1/n^2) (n \sigma^2) = \sigma^2 / n$$

つまり、確率変数 $\bar{X}$ は、平均 μ 、分散 σ^2 / n の確率分布に従うことが分かる。よって、平均 μ 、分散 σ^2 の確率分布を標準化するときの変数変換を参考にして(付録19)、平均 μ 、分散 σ^2 / n の確率分布を標準化するために、 $Y = n^{1/2} \bar{Z}$ ただし、 $\bar{Z} = (\bar{X} - \mu) / \sigma$ という変数変換を行う。

この変数変換で標準化できていること、つまり、平均 0、分散 σ^2 の確率分布になることを確認しておく。

$$E[Y] = E[n^{1/2} (\bar{X} - \mu) / \sigma] = (n^{1/2} / \sigma) E[\bar{X} - \mu] = (n^{1/2} / \sigma) (E[\bar{X}] - \mu) = (n^{1/2} / \sigma) (\mu - \mu) = 0$$

$$V[Y] = V[n^{1/2} (\bar{X} - \mu) / \sigma] = (n / \sigma^2) V[\bar{X} - \mu] = (n / \sigma^2) (V[\bar{X}] - V[\mu]) = (n / \sigma^2) (\sigma^2 / n - 0) = 1$$

(2) 標本のサンプルサイズを大きくしたときの、標準化した標本平均の確率分布(確率変数 $Y = n^{1/2} \bar{Z}$)の特性関数(付録24)を求める。虚数単位を、 j で表す。

$$\phi_Y(t) = E[e^{j t Y}] = E[e^{j t n^{1/2} (Z_1 + Z_2 + \dots + Z_n) / n}] = E[e^{j t n^{1/2} Z_1 / n}] E[e^{j t n^{1/2} Z_2 / n}] \dots E[e^{j t n^{1/2} Z_n / n}] = E[e^{j Z_1 t n^{-1/2}}] E[e^{j Z_2 t n^{-1/2}}] \dots E[e^{j Z_n t n^{-1/2}}]$$

i 番目の項に着目して、 $n \rightarrow \infty$ で $t n^{-1/2} \rightarrow 0$ と収束するので、 $t n^{-1/2}$ をひと塊の変数であると捉えて、 $f(t n^{-1/2}) = e^{j Z_i t n^{-1/2}}$ は、マクローリン展開(付録3)できる。

$$f(t n^{-1/2}) = e^{j Z_i t n^{-1/2}} = f(0) + f'(0) (t n^{-1/2}) / 1! + f''(0) (t n^{-1/2})^2 / 2! + f'''(0) (t n^{-1/2})^3 / 3! + \dots$$

$$\text{ここで、} f(0) = e^{j Z_i t n^{-1/2}} \Big|_{t n^{-1/2} = 0} = e^0 = 1$$

$$f'(0) = d f(t n^{-1/2}) / d(t n^{-1/2}) \Big|_{t n^{-1/2} = 0} = j Z_i e^{j Z_i t n^{-1/2}} \Big|_{t n^{-1/2} = 0} = j Z_i e^0 = j Z_i$$

$$f''(0) = d^2 f(t n^{-1/2}) / d(t n^{-1/2})^2 \Big|_{t n^{-1/2} = 0} = -Z_i^2 e^{j Z_i t n^{-1/2}} \Big|_{t n^{-1/2} = 0} = -Z_i^2 e^0 = -Z_i^2$$

$$f'''(0) = d^3 f(t n^{-1/2}) / d(t n^{-1/2})^3 \Big|_{t n^{-1/2} = 0} = -j Z_i^3 e^{j Z_i t n^{-1/2}} \Big|_{t n^{-1/2} = 0} = -j Z_i^3 e^0 = -j Z_i^3$$

$$\text{よって、} E[e^{j t n^{1/2} Z_i / n}] = E[f(0) + f'(0) (t n^{-1/2}) / 1! + f''(0) (t n^{-1/2})^2 / 2! + f'''(0) (t n^{-1/2})^3 / 3! + \dots]$$

$$= E[1 + j Z_i (t n^{-1/2}) / 1! - Z_i^2 (t n^{-1/2})^2 / 2! - j Z_i^3 (t n^{-1/2})^3 / 3! + \dots]$$

ここで、 Z_i^3 以上の高次の項は、 $n \rightarrow \infty$ で、急速に 0 になるので無視できる。

また、 $E[Z_i]$ と $V[Z_i]$ は次のように求まる。

$$E[Z_i] = E[(X_i - \mu) / \sigma] = (1 / \sigma) E[X_i - \mu] = (1 / \sigma) (E[X_i] - E[\mu]) = (1 / \sigma) (\mu - \mu) = 0$$

$$V[Z_i] = V[(X_i - \mu) / \sigma] = (1 / \sigma^2) V[X_i - \mu] = (1 / \sigma^2) (V[X_i] - V[\mu]) = (1 / \sigma^2) (\sigma^2 - 0) = 1$$

$$E[Z_i^2] = V[Z_i] + E[Z_i]^2 = 1 + 0 = 1$$

$$\text{よって、} \phi_{Y_i}(t) = E[e^{j t Y_i}] = E[e^{j t n^{1/2} Z_i / n}] \doteq E[1 + j Z_i (t n^{-1/2}) / 1! - Z_i^2 (t n^{-1/2})^2 / 2!]$$

$$= E[1] + j (t n^{-1/2}) E[Z_i] - ((t n^{-1/2})^2 / 2) E[Z_i^2] = 1 + j (t n^{-1/2}) \cdot 0 - (t n^{-1/2})^2 \cdot 1 = 1 - t^2 / (2n)$$

以上から、標本サイズ $n \rightarrow \infty$ にしたときの、標準化した標本平均の確率分布の特性関数 $\phi_W(t)$ が求まる(式変形には、(付録4)のネイピア数の公式を使用する)。

$$\phi_W(t) = \lim_{n \rightarrow \infty} \phi_Y(t) = \lim_{n \rightarrow \infty} \left\{ \prod_{i=1}^n \phi_{Y_i}(t) \right\} = \lim_{n \rightarrow \infty} (1 - t^2 / (2n))^n = \lim_{n \rightarrow \infty} (1 + (-t^2 / 2) / n)^n = e^{-t^2 / 2}$$

(3) 標準正規分布の特性関数を求める。

標準正規分布の確率密度関数は、 $f(x) = (1 / (2\pi))^{1/2} e^{-x^2 / 2}$ である。この標準正規分布の特性関数は、以下のようにして求まる。

$$\begin{aligned} \phi_X(t) &= E[e^{j t X}] = \int_{-\infty}^{\infty} e^{j t x} (1 / (2\pi))^{1/2} e^{-x^2 / 2} dx = (1 / (2\pi))^{1/2} \int_{-\infty}^{\infty} e^{-x^2 / 2} e^{j t x} dx = (1 / (2\pi))^{1/2} \int_{-\infty}^{\infty} e^{-x^2 / 2 + j t x} dx = (1 / (2\pi))^{1/2} \int_{-\infty}^{\infty} e^{-(1/2)(x - j t)^2 - (1/2) t^2} dx \\ &= (1 / (2\pi))^{1/2} e^{-(1/2) t^2} \int_{-\infty}^{\infty} e^{-(1/2)(x - j t)^2} dx = (1 / (2\pi))^{1/2} e^{-t^2 / 2} \int_{-\infty}^{\infty} e^{-(1/2)(x - j t)^2} dx \end{aligned}$$

ここで、 $u = x - j t$ と置けば、 $dx = du$ 、積分区間は、 $-\infty$ から ∞ のままで同じで、式変形に(付録9)のガウス積分を使うと、

$$\phi_X(t) = (1 / (2\pi))^{1/2} e^{-t^2 / 2} \int_{-\infty}^{\infty} e^{-(1/2) u^2} du = (1 / (2\pi))^{1/2} e^{-t^2 / 2} (2\pi)^{1/2} = e^{-t^2 / 2}$$

(4) よって、(2)と(3)の特性関数が一致したので、(2)サンプルサイズを大きくしたときの標準化した標本平均の確率密度関数と(3)標準正規分布の確率密度関数が一致する。

従って、母集団からの無作為抽出の試行を無限に繰り返せば、母集団がどのような確率分布であっても、期待値と分散さえ存在すれば、標本平均の分布が正規分布に近づくという中心極限定理が導かれた。

付録47. 行列式の定義

(1) 置換

1, 2, ..., n を並び替える操作を、n 次の置換と言う。置換の数は n! 個になる。

(例) 1, 2, 3, 4, 5 を 3, 5, 2, 1, 4 のように並び替える操作を、元の並びを上にも、並び替えたものを下にして、次のように表す。

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 1 & 5 \end{bmatrix}$$

この例では、置換 σ によって、1 は 2 に、2 は 3 に、... 移されることを示している。
ここで、1 を 2 に移すことを $\sigma(1) = 2$ 、2 を 3 に移すことを $\sigma(2) = 3$ のように表記する。

(2) 互換

置換のなかで、2 つの要素だけを交換するものを、互換と言う。互換が偶数回行われる置換を偶置換、奇数回行われる置換を奇置換と言う。

(例) 1, 2, 3, 4, 5 を 1, 2, 4, 3, 5 に並び替える操作は、3 と 4 の 2 つの要素を並び替える互換の操作である。この操作は $\sigma = (3, 4)$ と表す。

(例) 1, 2, 3, 4, 5 を 1, 2, 4, 3, 5 に並び替え、さらに 1, 2, 4, 5, 3 と並び替える操作は $\sigma = (3, 5)(3, 4)$ と表す。

(3) 行列式の定義

互換の操作を使用して、n × n の行列 $A = [a_{ij}]$ の行列式 $|A|$ を表す(行列式は $\det A$ とも表す)。

$$|A| = \sum_{\text{全ての置換を対象に}} \text{sgn}(\sigma) \prod_{i=1}^n a_{i, \sigma(i)} = \sum_{\text{全ての置換を対象に}} \text{sgn}(\sigma) a_{1, \sigma(1)} a_{2, \sigma(2)} \cdots a_{n, \sigma(n)} \quad \cdots \text{(付式66)}$$

但し、 $\text{sgn}(\sigma) = 1$ (偶置換のとき)、 $\text{sgn}(\sigma) = -1$ (奇置換のとき)

(例)

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

の行列式を、(付式66)に従って求める。

$$\sigma = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$$

のとき、置換回数 = 0 (偶置換) で $\text{sgn}(\sigma) = 1$ 、 $\sigma(1) = 1$ 、 $\sigma(2) = 2$ 、 $\sigma(3) = 3$ であるので、 $a_{1, \sigma(1)} a_{2, \sigma(2)} a_{3, \sigma(3)} = a_{11} a_{22} a_{33}$

$$\sigma = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

のとき、置換回数 = 2 (偶置換) で $\text{sgn}(\sigma) = 1$ 、 $\sigma(1) = 2$ 、 $\sigma(2) = 3$ 、 $\sigma(3) = 1$ であるので、 $a_{1, \sigma(1)} a_{2, \sigma(2)} a_{3, \sigma(3)} = a_{12} a_{23} a_{31}$

$$\sigma = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$$

のとき、置換回数 = 2 (偶置換) で $\text{sgn}(\sigma) = 1$ 、 $\sigma(1) = 3$ 、 $\sigma(2) = 1$ 、 $\sigma(3) = 2$ であるので、 $a_{1, \sigma(1)} a_{2, \sigma(2)} a_{3, \sigma(3)} = a_{13} a_{21} a_{32}$

$$\sigma = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$$

のとき、置換回数 = 1 (奇置換) で $\text{sgn}(\sigma) = -1$ 、 $\sigma(1) = 3$ 、 $\sigma(2) = 2$ 、 $\sigma(3) = 1$ であるので、 $a_{1, \sigma(1)} a_{2, \sigma(2)} a_{3, \sigma(3)} = a_{13} a_{22} a_{31}$

$$\sigma = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}$$

のとき、置換回数 = 1 (奇置換) で $\text{sgn}(\sigma) = -1$ 、 $\sigma(1) = 2$ 、 $\sigma(2) = 1$ 、 $\sigma(3) = 3$ であるので、 $a_{1, \sigma(1)} a_{2, \sigma(2)} a_{3, \sigma(3)} = a_{12} a_{21} a_{33}$

$$\sigma = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}$$

のとき、置換回数 = 1 (奇置換) で $\text{sgn}(\sigma) = -1$ 、 $\sigma(1) = 1$ 、 $\sigma(2) = 3$ 、 $\sigma(3) = 2$ であるので、 $a_{1, \sigma(1)} a_{2, \sigma(2)} a_{3, \sigma(3)} = a_{11} a_{23} a_{32}$

従って、A の行列式 $|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{13} a_{22} a_{31} - a_{12} a_{21} a_{33} - a_{11} a_{23} a_{32}$

付録48. 行列式と行列の性質(補足)

(1) 行列 A が正方行列のとき、 $|A^T| = |A|$ が成り立つ。
行列 A の行列式は、付録47の(付式66)で表される(下記に再掲載)。

$$|A| = \sum_{\text{全ての置換を対象に}} \text{sgn}(\sigma) \prod_{i=1}^n a_{i\sigma(i)} = \sum_{\text{全ての置換を対象に}} \text{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \quad \cdots \text{(付式66)}$$

よって、行列 A の転置行列 A^T の行列式は、

$$|A^T| = \sum_{\text{全ての置換を対象に}} \text{sgn}(\sigma) \prod_{i=1}^n a_{i\sigma(i)} = \sum_{\text{全ての置換を対象に}} \text{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

置換操作 σ は置換元(射影元)と置換先(射影先)が 1 対 1 対応しているので、逆置換(逆射影) σ^{-1} が存在する。よって、 $\sigma^{-1}(\sigma(i)) = i$ が成り立つ。
また、 $\{\sigma(1) \sigma(2) \cdots \sigma(n)\}$ は、 $\{1 \ 2 \cdots n\}$ の順番を入れ替えたものであり、 $\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma)$ であるので、上式は以下のように書き表せる。

$$\begin{aligned} |A^T| &= \sum_{\text{全ての置換を対象に}} \text{sgn}(\sigma) \prod_{i=1}^n a_{i\sigma(i)} = \sum_{\text{全ての置換を対象に}} \text{sgn}(\sigma) a_{\sigma(1)\sigma^{-1}(\sigma(1))} a_{\sigma(2)\sigma^{-1}(\sigma(2))} \cdots a_{\sigma(n)\sigma^{-1}(\sigma(n))} \\ &= \sum_{\text{全ての置換を対象に}} \text{sgn}(\sigma^{-1}) a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)} \end{aligned}$$

従って、 $\xi = \sigma^{-1}$ とおくと、 $|A^T| = \sum_{\text{全ての置換を対象に}} \text{sgn}(\xi) a_{1\xi(1)} a_{2\xi(2)} \cdots a_{n\xi(n)}$ となる。この式は(付式65)に等しい。よって、 $|A^T| = |A|$ が成り立つ。

(2) 行列 A が正則行列のとき(つまり、逆行列が存在するとき)、 $(AB)^{-1} = B^{-1}A^{-1}$ が成り立つ。

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

従って、 $B^{-1}A^{-1}$ は AB の逆行列であることが分かる。つまり、 $(AB)^{-1} = B^{-1}A^{-1}$ が成り立つ。

(3) 行列 A が正則行列のとき(つまり、逆行列が存在するとき)、 $(A^{-1})^T = (A^T)^{-1}$ が成り立つ。

一般に、 $(AA^{-1})^T = (A^{-1})^T A^T$ が成り立つ。

左辺 $= (AA^{-1})^T = I$ (単位行列) より、 $(A^{-1})^T A^T = I$

従って、 $(A^{-1})^T$ は、 A^T の逆行列であることが分かる。つまり、 $(A^{-1})^T = (A^T)^{-1}$ が成り立つ。

なお、 A が対称行列であれば、 $A = A^T$ であるので、 $(A^{-1})^T = (A^T)^{-1} = A^{-1}$ が成り立つ。

また、 $(AA^T)^T = (A^T)^T A^T = AA^T$ であるので、行列 A と A^T の積 AA^T は対称行列になる。よって、 $((AA^T)^{-1})^T = (AA^T)^{-1}$ が成り立つ。

同様に、 $((A^T A)^{-1})^T = (A^T A)^{-1}$ が成り立つ。

参考にした書籍

涌井良幸、涌井貞美：「統計学の図鑑」、技術評論社

Academaid：「これなら分かる！ はじめての数理統計学」、[数理統計 | Academaid \(academ-aid.com\)](#)

Fukusukeの数学メモ：「テイラーの定理」、[テイラーの定理～具体例からわかりやすく解説！ 剰余項の由来はコーシーの平均値の定理！？～ | Fukusukeの数学めも \(mathsuke.jp\)](#)
おいしい数学：「多項定理」、[多項定理 | おいしい数学 \(hiraocafe.com\)](#)

基礎からの数学入門：「一階線形微分方程式」、[一階線形微分方程式の解き方 - 一階常微分方程式の解法 - 微分方程式 - 基礎からの数学入門 \(keicode.com\)](#)

理数アラカルト：「転置行列の性質と公式」、[転置行列の性質と公式集 \(具体例と証明付\) - 理数アラカルト - \(risalc.info\)](#)

数理アラカルト：「ネイピア数」、[ネイピアの定数 e とは？ \(定義・例・性質\) - 理数アラカルト - \(risalc.info\)](#)

数理アラカルト：「ロピタルの定理」、[ロピタルの定理とは？ ～証明と具体例～ - 理数アラカルト - \(risalc.info\)](#)

数理アラカルト：「ガンマ関数」、[ガンマ関数の定義と性質 - 理数アラカルト - \(risalc.info\)](#)

数理アラカルト：「ベータ関数」、[ベータ関数とは？ ～性質と公式～ \(証明付\) - 理数アラカルト - \(risalc.info\)](#)

しまね統計情報データベース：「母集団と標本」、[母集団と標本：しまね統計情報データベース \(shimane-toukei.jp\)](#)

数理アラカルト：「期待値と分散の公式」、[期待値と分散の公式 \(証明と具体例\) - 理数アラカルト - \(risalc.info\)](#)

AVILEN AI Trend：「積率母関数」、[積率母関数とは？モーメントの求め方も解説 | AVILEN AI Trend \(ai-trend.jp\)](#)

エルミの勉強部屋：「特性関数」、[特性関数①「概要」 - YouTube](#)

マサムネの部屋：「ハザード関数」、[ハザード関数とモデル選択の話 | マサムネの部屋 \(masamunetogotoge.com\)](#)

EMANの物理学：「デルタ関数」、[デルタ関数 - EMANの物理数学 \(eman-physics.net\)](#)

予備校のノリで学ぶ「大学の数学・物理」：「行列式の幾何学的意味」、[行列式の幾何学的意味 - YouTube](#)

予備校のノリで学ぶ「大学の数学・物理」：「ヤコビアン」、[【大学数学】重積分④\(ヤコビアン\)/全4回【解析学】 - YouTube](#)

早川美徳(東北大学)：「Pythonプログラミング(ステップ7・確率密度関数とその計算)」、<https://wagtail.cds.tohoku.ac.jp/coda/python/p-7-array-stat-pdf.html#header>

高校数学の美しい物語：「ウォリスの公式」、[ウォリスの公式とその2通りの証明 | 高校数学の美しい物語 \(manabitimes.jp\)](#)

高校数学の美しい物語：「スターリングの公式」、[スターリングの公式とその証明 | 高校数学の美しい物語 \(manabitimes.jp\)](#)

安野嘉晃(筑波大学)：「確率論 全10回」、[確率論 講義 #001 確率変数と確率分布関数 - YouTube](#)

具体例で学ぶ数学：「標準化」、[統計における標準化の意味と目的 - 具体例で学ぶ数学 \(mathwords.net\)](#)

いちばんやさしい、医療統計：「一様分布」、[一様分布とは？サイコロの例でわかりやすく！期待値\(平均\)や分散の出し方も | いちばんやさしい、医療統計 \(best-biostatistics.com\)](#)

AVILEN AI Trend：「ベルヌーイ分布」、[ベルヌーイ分布のわかりやすいまとめ | AVILEN AI Trend \(ai-trend.jp\)](#)

ゆっくりキカイガクシュウ：「カテゴリカル分布」、[カテゴリカル分布と多項分布 期待値・分散・共分散の求め方 | Yukkuri Machine Learning \(laid-back-scientist.com\)](#)

統計WEB (SSRI)：「二項分布」、[13-1. 二項分布 | 統計学の時間 | 統計WEB \(bellcurve.jp\)](#)

データ科学便覧：「負の二項分布」、[負の二項分布 \(data-science.gr.jp\)](#)

いちばんやさしい、医療統計：「幾何分布」、[幾何分布とは？期待値\(平均\)や分散の証明はどうなるか例題を用いて解説 | いちばんやさしい、医療統計 \(best-biostatistics.com\)](#)

初心者からはじめる統計学：「多項分布」、[多項分布の期待値・分散の求め方【証明付きで解説】 | 初心者からはじめる統計学 \(develop-chronos.com\)](#)

AVILEN AI Trend：「超幾何分布」、[超幾何分布の確率密度関数からの期待値と分散の導出 | AVILEN AI Trend \(ai-trend.jp\)](#)

予備校のノリで学ぶ「大学の数学・物理」：「指数分布」、[【大学数学】指数分布\(具体例やその意味、ポアソン分布との関係\)【確率統計】 - YouTube](#)

データ科学便覧：「ワイブル分布」、[ワイブル分布 \(data-science.gr.jp\)](#)

エルミの勉強部屋：「ポアソン分布」、[統計学⑨「ポアソン分布」 - YouTube](#)

Hatsudy：総合学習サイト：「ポアソン分布」、[ポアソン分布：正規分布との違いや期待値・確率の計算方法 | Hatsudy: 総合学習サイト](#)

緑川章一(青森大学)：「ガンマ分布」、<https://shoichimidorikawa.github.io/Lec/ProbDistr/gammaDist.pdf>

バナナでもわかる話：「ベータ分布」、[統計検定1級対策問題集`ベータ分布編`～バナナでもわかる話 \(bananarian.net\)](#)

高校数学の美しい物語：「ディリクレ分布」、[ディリクレ分布の意味と正規化、平均などの計算 | 高校数学の美しい物語 \(manabitimes.jp\)](#)

長谷川浩司(東北大学)：「正規分布」、[確率分布導出のまとめ\(東北大学\) 200905toukei.print.pdf](#)

山下雄史(京都大学)：「多変量正規分布」、[【統計学入門】多変量正規分布～多次元の正規分布～ - YouTube](#)

高校数学の美しい物語：「対数正規分布」、[対数正規分布の例と平均、分散 | 高校数学の美しい物語 \(manabitimes.jp\)](#)

biostatistics：「カイニ乗分布」、[カイニ乗分布 | 標準正規分布に従う確率変数の2乗和はカイニ乗分布に従う \(biopapyrus.jp\)](#)

AVILEN AI Trend：「F 分布」、[F分布の確率密度関数をカイニ乗分布を用いて導出 | AVILEN AI Trend \(ai-trend.jp\)](#)

みつのきチャンネル：「t 分布」、[t分布\(スチューデント分布\)の式の導出【統計学 数学】 - YouTube](#)

ウサギさんの統計学サロン：「コーシー分布」、[【統計学】コーシー分布の確率密度関数 \(multivariate-statistics.com\)](#)

予備校のノリで学ぶ「大学の数学・物理」：「マルコフの不等式」、「チェビシェフの不等式」、「大数の弱法則」、[【大学数学】チェビシェフの不等式【確率統計】 - YouTube](#)

はじめての統計学：中心極限定理、[中心極限定理の証明をわかりやすく基礎から解説！ - YouTube](#)