第２章　確率分布を理解するための基本的なこと

この後学んでゆく様々な確率分布の理解に必要な確率に関する基本的なことをまとめている。この章では、本書のどこかで必要になる事項を中心に取り上げていて、逆に言うと、確率論を学ぶことには大事なことであっても、電波伝搬の確率分布理解に直接現れてこないことについては触れていない。（確率論の全体的な理解のためには、理工学両面から、それぞれのレベルに応じた良書が数多く出版されているので、身の丈にあったものを選んで学んでほしい。）

2.1　確率と確率密度

確率（probability）とは、偶然性を持つ事象について、その事象が起こることが期待される度合い、あるいは現れることが期待される割合のことを言い、0 から 1 までの値をとる。0 は絶対起こらない、1 は必ず起こるという極限値で、通常は、その中間の値になる。コイン投げでの表が出る確率は 1/2、サイコロでは、どの目の確率も 1/6 と言った値である。このように出現値が離散的なものであれば、上記のように、即座に数値で表すことができるが、連続的に分布する量では、このように簡単なことには行かない。例えば、「人間の身長が 160cm である確率は？」と問われても、160.000…とぴったりした値は起こりえず、問いに意味がない。強いて言えば「確率 0」である。それに対して、「160cm から 165cm の間にある確率は？」、あるいは、「170cm 以上になる確率は？」と言う問いには、前提が明確であれば、適切に答えることができる。電波伝搬で取り上げられる物理量は、その大部分が、後者の例に属するアナログ量（＝連続分布）になる。

確率的規則に従って連続的に変化する物理量（実現値）x があるとき、それを大文字表記 X に変えて**確率変数**（random variable）と呼ぶ。変数 X がその変化範囲の中の微小範囲 dx（すなわち、$x- dx/2 \leq X \leq x+dx/2$）に存在する確率が、関数 $f(x)$ を用いて $f(x)dx$ と表わされるとき、$f(x)$を**確率密度関数**（probability density function: PDF）という。また、変数 X が x_{min} （X の存在範囲の最小値）と x の範囲にある確率を**累積確率**（cumulative probability）、それを x の関数として表したものを**累積分布関数**（cumulative distribution function: CDF または**確率分布関数**：probability distribution function）といい、$F(x)$と書く。このように、連続分布の場合は、**確率密度**（probability density）と言う考え方が大事になる。サイコロのように、出現値が離散的である場合には、例えば、5 の目が出る確率密度は(1/6) $\delta(x-5)$（δは Dirac のデルタ関数）と表される。確率密度の場合は、確率と違って、その値は 0～∞の全範囲の値をとり得る。

変数 X の存在範囲を x_{min}～x_{max} とするとき、確率密度関数 $f(x)$には、連続分布・離散分布を問わず、以下の性質がある。関数 $f(x)$が確率密度関数であるための必要条件とも言える。

$$f(x) \geq 0, \qquad \int_{x_{\min}}^{x_{\max}} f(x)\,dx = 1 \tag{2.1}$$

累積分布関数 $F(x)$は次式で表わされる。

$$F(x) = \int_{x_{\min}}^{x} f(x)\,dx \tag{2.2}$$

上式より、$F(x_{min})=0, F(x_{max})=1$ となる。また、$F(x)$は、単調非減少の性質（$a<b$ ならば、$F(a)\leq F(b)$）を持つ。

図 2.1 は、後述するレイリー分布（式(4.6)：σ=1）を例に、確率密度関数と累積分布関数の関係を示している。累積確率は、時間的に変化する現象に対する確率の場合は**累積時間率**（あるいは単に時間率）、空間の場所的な確率の場合は**累積場所率**（あるいは場所率）と呼ばれている。時間率、場所率は、その確率を 100 倍して、パーセントで表わされることも多い。

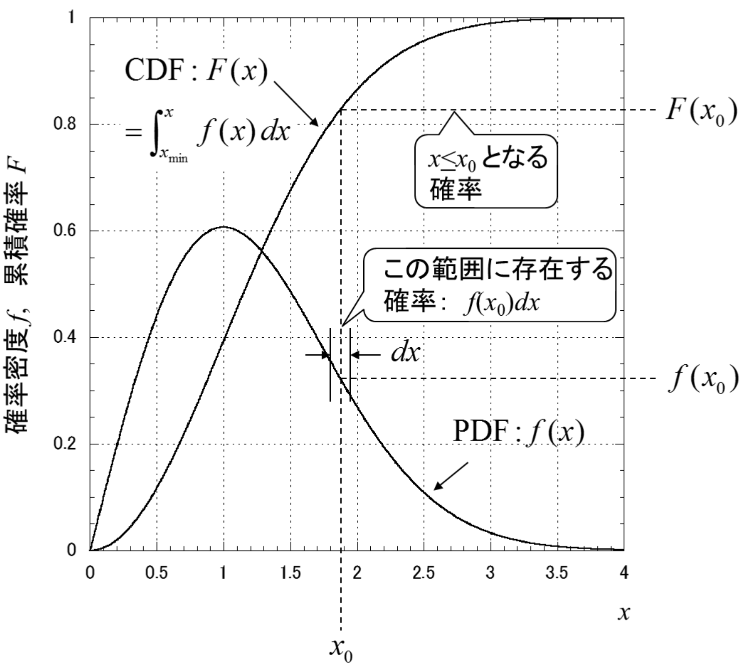


図 2.1　確率密度関数(PDF)と累積分布関数(CDF)　(曲線はレイリー分布(σ=1)の例)

2.2　積率と相関

2.2.1　平均・分散・積率（モーメント）

確率変数 X の関数で表される $h(X)$の**平均値**（mean value）$<h(X)>$は次式で与えられる。母集団 $h(X)$の確率的平均値（**集合平均**：ensemble average）であり、この意味では**期待値**（expected value）と呼ばれ、$E(h(X))$とも表される。本書では、期待値の場合も含めて、平均値 ＜・＞で表す。

$$\langle h(X) \rangle = \int_{x_{\min}}^{x_{\max}} h(x)\,f(x)\,dx \tag{2.3}$$

上式において、$h(X)=X$ では、変数 X の平均値を与える。$h(X)=X^n$（n=1,2,⋯）のとき、すなわち$<X^n>$値は n 次の**積率**（**モーメント**：moment）と呼ばれる。$\langle X\rangle=\mu$と置くとき、積率$<(X-\mu)^n>$は、平均値の周りの n 次の積率（**中心モーメント**：central moment）と呼ばれる。これらに関して、以下のような名前が付けられている。

① **平均値**（mean value, average）：μ（$E(X)$とも表現される）

$$\mu \equiv \langle X \rangle = \int_{x_{\min}}^{x_{\max}} x\, f(x)\, dx$$

② **分散**（variance）：σ^2（$V(X)$とも表現される）

$$\sigma^2 \equiv \left\langle (X-\mu)^2 \right\rangle = \int_{x_{\min}}^{x_{\max}} \left(x-\mu\right)^2 f(x)\, dx = \left\langle X^2 \right\rangle - \mu^2$$

③ **標準偏差**（standard deviation）：分散の平方根 σ

④ **中央値**（median）：$F(x) = 0.5$ となる x の値

⑤ **最頻値**（mode, most probable value）：$f(x)$が最大となる x の値

⑥ **自乗平均値**（mean square value）：$<X^2>$

$$\left\langle X^2 \right\rangle = \int_{x_{\min}}^{x_{\max}} x^2 f\left(x\right) dx = \mu^2 + \sigma^2$$

⑦ **rms 値**（root mean square value）：x_{rms}

$$x_{rms} \equiv \sqrt{\left\langle X^2 \right\rangle} = \sqrt{\mu^2 + \sigma^2}$$

⑧ **歪度**（わいど：skewness）：$<(X-\mu)^3>/\sigma^3$

⑨ **尖度**（せんど：kurtosis）：$<(X-\mu)^4>/\sigma^4$

分布の形状は、①、②、⑧、⑨の４つでほぼ特徴付けられる。

確率変数 X を、平均値μと標準偏差σで規格化して、$Y=(X-\mu)/\sigma$と置くと、$<Y>=0$, $<Y^2>=1$ となる。任意の正の数λに対して、$|Y|\geq\lambda$ となる確率、すなわち、$|X-\mu|\geq\lambda\sigma$となる確率は、λが大きいほど小さくなる。このとき、分布が何であっても、以下の式

$$\mathrm{Prob}\left\{ \left| X-\mu \right| \geq \lambda\sigma \right\} \leq 1/\lambda^2 \tag{2.4}$$

が成り立つ。これを**チェビシェフの不等式**と言う。[証明は、例えば、文献[1]の p. 63]。平均値から 3σ以上離れる確率は、どんな確率分布でも 1/9 以下であると言うことを示している。具体的な確率分布、例えば、正規分布では$\lambda=3$ において 0.2%（両側それぞれ 0.1%）となるように、(2.4)式よりははるかに収束が良いのであるが、全ての分布に対して押さえられるのが興味深い。

2.2.2　相関

時間 t に従って変化する現象を $x(t)$とする。同様に別の現象を $y(t)$とする。このとき、$x(t)$と $y(t)$の似通い具合（類似度）を表す指標として**相関**（correlation）が用いられる。人間の身長と体重のように、身長が増加すれば、体重も増加する傾向が強いようなものに対しては相関が強く、「風が吹くと桶屋が儲かる」的な因果関係の薄いものに対しては、相関が弱い、と言うように用いる。

以下、その定量化である。タイミング（時間である必要はない）を $i\,(=1, 2, \ldots, n)$として、そのときに生起する量を x_i, y_i, それぞれの平均値をμ_x, μ_y とする。このとき、**相関係数**（correlation coefficient）ρを次式で定義する。

$$\rho = \frac{\sum_{i=1}^{n}(x_i-\mu_x)(y_i-\mu_y)}{\sqrt{\sum_{i=1}^{n}(x_i-\mu_x)^2 \cdot \sum_{i=1}^{n}(y_i-\mu_y)^2}} \tag{2.5}$$

相関係数は、変化する量 x と y が完全相似形のとき１（**完全相関**）、正負反転で完全相似形のとき-１（**負の完全相関**）、独立なとき０（**無相関**）となり、この範囲（-1～1）に値を持つ。$|\rho|$が１を超えないことは、次式の**シュワルツの不等式**により保証されている（証明：略）。

$$\sum_{i}^{n}(x_i-\mu_x)^2\sum_{i}^{n}(y_i-\mu_y)^2 \geq \left|\sum_{i}^{n}(x_i-\mu_x)(y_i-\mu_y)\right|^2 \tag{2.6}$$

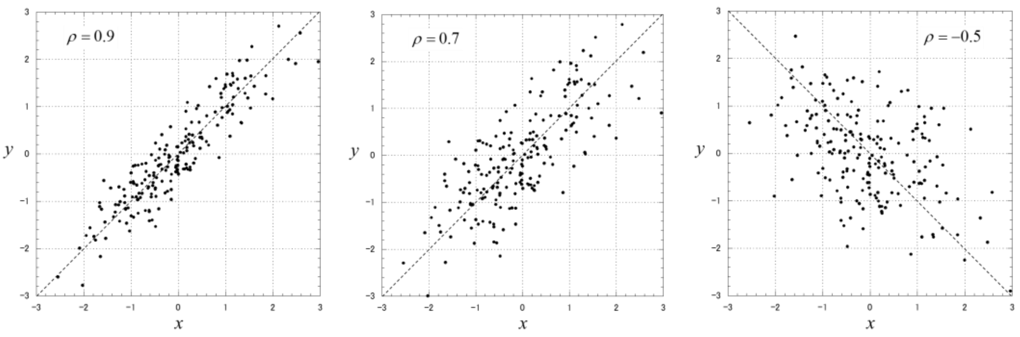
変数 X, Y の結合確率密度関数 $f(x,y)$ $(x_{min}\leq X\leq x_{max}, y_{min}\leq Y\leq y_{max})$を用いて、**共分散**（covariance）$\sigma_{xy}{}^2$を定義すると、次式となる。

$$\begin{aligned}\sigma_{xy}^2 &= \int_{y_{\min}}^{y_{\max}}\int_{x_{\min}}^{x_{\max}}(x-\mu_x)(y-\mu_y)\,f(x,y)\,dxdy \\ &= \langle (X-\mu_x)(Y-\mu_y)\rangle = \langle XY\rangle - \mu_x\mu_y\end{aligned} \tag{2.7}$$

相関係数ρは、X, Y のそれぞれの標準偏差をσ_x, σ_y とすると、(2.5)式に対応する形として、次式で表される。

$$\rho = \frac{\sigma_{xy}^2}{\sigma_x\sigma_y} \tag{2.8}$$

図 2.2 は、相関係数ρ=0.9, 0.7, -0.5 に対する x と y の散布図である。同図より相関係数の特徴が良くわかると思う。なお、ρ=0.7（正確には$1/\sqrt{2}$）は、X に対して、Y が X と X に独立な変動（例えば、雑音）が同じ平均電力で合成されるときの状況である。すなわち、X を所望波信号（=SN 比無限大）とすると、Y は SN 比が 1 (=0dB)の信号となる。



（a）　ρ = 0.9　　（b）　ρ = 0.7　　（c）　ρ = - 0.5

図 2.2　相関係数をパラメータとした散布図（標準正規分布乱数を使用）

2.3　確率変数の変換

2. 3. 1　１変数の変換

確率変数 X が確率密度関数 $f(x)$に従う時、$U=\phi(X)$とする新たな確率変数 U の確率密度関数 $g(u)$を求めてみよう。関数ϕは x に対して連続な１次導関数をもつとする。どのような変換によっても、$x_1 \le X \le x_2$, $\phi(x_1) \le U \le \phi(x_2)$の累積確率は同じであり、

$$\int_{x_1}^{x_2} f(x)\,dx = \int_{\phi(x_1)}^{\phi(x_2)} g(u)\,du \tag{2. 9}$$

の関係[あるいは $f(x)dx = g(u)du$]が満たされるので、

$$g(u) = f\left\{x(u)\right\}\left|\frac{dx}{du}\right| \tag{2. 10}$$

となる。電波伝搬のモデルに現れる例では、真数で与えられる確率変数の分布を dB 値の分布に変換する $U = 20\log_{10} X$、振幅の分布が与えられた場合の電力の分布を求める時の $U = X^2$ などがある。前者（真数値→dB 値変換）については、x の増加に対して u も単調に増加し、u の確率分布は、次式のようになる。

$$g(u) = \frac{1}{b}\exp\left(\frac{u}{b}\right) f\left\{\exp\left(\frac{u}{b}\right)\right\} \qquad (b \equiv 20\log_{10} e) \tag{2. 11}$$

後者（$X \to X^2$ 変換）の場合には、少し注意が必要になる。確率変数 X が正の領域に存在する場合は、同様の手順で式(2. 12a)となる。一方、$-\infty<X<\infty$で存在する場合には、$\pm x$ が同じ u 値になるので、これを考慮して式(2. 12b)となる。

$$g(u) = \frac{1}{2\sqrt{u}} f\left(\sqrt{u}\right) \qquad (0 \le X) \tag{2. 12a}$$

$$g(u) = \frac{1}{2\sqrt{u}}\left\{f\left(-\sqrt{u}\right) + f\left(\sqrt{u}\right)\right\} \qquad (-\infty < X < \infty) \tag{2. 12b}$$

上記の変換の考え方に基づく、既存の確率分布 f に従う乱数 x_i から、別の確率分布 g に従う乱数 u_i を生成することができる。ここでは、$0 \le x_i \le 1$ の一様分布乱数 $f(x_i)=1$ から、他の乱数を求める方法を示す。

ステップ１：一様乱数生成手法により乱数値 x_i を求める。

ステップ２：この累積確率値 $F(x_i)$を求める。上記一様乱数では、$F(x_i) = x_i$ である。

ステップ３：作りたい確率分布の累積確率 $G(u_i)=F(x_i)(=x_i)$となる u_i を求める。G が u の関数で与えられている場合には $u_i = G^{-1}(x_i)$で求める。

ステップ１に戻って、これを繰り返せば、u_i は求める確率分布 g に従う乱数列となる。図 2.3 はこの手順をまとめている。

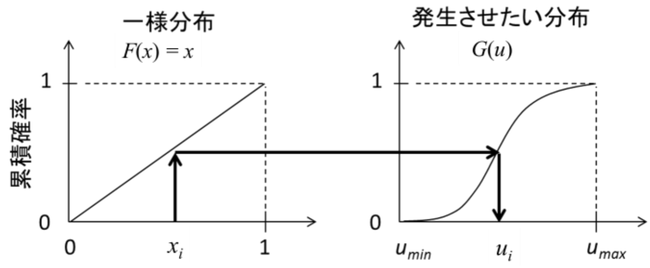


図 2.3　(0～1)の一様分布乱数から任意の確率分布を有する乱数の生成法

一般論としてはこの手順で良いのであるが、ステップ３の累積分布関数の逆関数 $G^{-1}(x_i)$が求められることが条件になる。分布によっては（＝大部分の分布では）それが簡単でなく、ケースバイケースでの工夫が必要である。以下、うまく行く例を示す。

一様乱数から**レイリー分布**乱数を発生させてみよう。レイリー分布の詳しい説明は 4,1 節で行うので、ここでは、分布の形 $g(u)$だけを見てほしい（パラメータσ^2はσ^2=1 としている）。

$$g(u)=u\exp\left(-\frac{u^2}{2}\right)\qquad (0\le u)$$

累積分布関数 $G(u)$は

$$G(u)=1-\exp\left(-\frac{u^2}{2}\right)$$

であるので、$G(u_i)=x_i$ より、

$$1-\exp\left(-\frac{u_i^2}{2}\right)=x_i \;\rightarrow\; u_i=\sqrt{-2\log(1-x_i)}$$

となり、右側の式でレイリー分布乱数が生成できる。なお、単に乱数を発生させたいだけならば、式中の $1-x_i$ は x_i と置きなおしても、生成される分布の形は変わらないので、

$$u_i=\sqrt{-2\log x_i}$$

としてもよい。さらに、4.1 節で学ぶように、位相角θが 0～2πの一様乱数と組み合わせて

$$z=u\cos\theta$$

なる量を考えると、確率変数 Z（すなわち、実現値 z_i）は、平均値 0、分散 1 の**標準正規分布**になる。ゆえに、独立な二つの一様乱数（0～1 区間での）x_i, y_i を用いて

$$z_i=\sqrt{-2\log x_i}\cos(2\pi y_i)$$

とすると、z_i は標準正規分布乱数になる（もちろん cos の部分は sin にしてもよい）。このように

して標準正規分布乱数を作る方法は**ボックス-ミュラー法**と呼ばれている。（ボックス-ミュラー法は便利な標準正規分布乱数生成方法（アルゴリズム）であるが、近年は、統計処理ソフトの中に、生成コマンドが標準装備されているので、実際に使う人は少なくなっているかもしれない）

2.3.2　２変数の変換

確率変数を X、Y、その結合確率密度関数を $f(x,y)$とする。また、変換後の確率変数を U, V、その結合確率密度関数を $g(u,v)$とする。X, Y と U, V の関係を与える

$$U=\phi_1(X,Y),\ \ V=\phi_2(X,Y) \tag{2.13}$$

における関数ϕ_1, ϕ_2, は X, Y に対して連続な１次偏導関数をもつとする。このとき、変換後の結合確率密度関数 $g(u, v)$は次式で与えられる。

$$g(u,v)=f\left\{x(u,v),\ y(u,v)\right\}\left|J\right| \tag{2.14}$$

ここで、J はヤコビアン

$$J\equiv\begin{vmatrix}\dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v}\\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v}\end{vmatrix} \tag{2.15}$$

であり、$|J|$はヤコビアンの絶対値である。当然ながら式(2.14)は１変数の場合の式(2.10)を含んだ一般化になっている。

２変数の変換では、直交座標(x,y)のから極座標(r,θ)への確率分布変換 ($f\rightarrow g$) が代表例である。

$$x=r\cos\theta,\ \ y=r\sin\theta \tag{2.16}$$

とおくと、

$$g(r,\theta)=rf\left(r\cos\theta,\ r\sin\theta\right) \tag{2.17}$$

となる。

確率変数が３以上でも、式(2.14), (2.15)の x, y, u, v を $x, y, z, \ldots, u, v, w, \ldots$としてそのまま拡張できる。

2.3.3　確率変数の和・差・積・商の分布

複数の確率変数の和や積など四則演算したものを新たな確率変数として分布を求めたい場合がある。ここでは、二つの確率変数 X, Y の和・差・積・商を Z とし、Z を確率変数とする分布の計算式を与える。一般的な和の分布の計算法は次節の積分変換（特性関数）の説明の中で示す。

以下に共通する記号の定義を与える。X, Y の結合確率密度関数を $f(x,y)$、Z の確率密度関数を

$g(z)$、X, Y が互いに独立なときの X, Y の確率密度関数を $f_1(x), f_2(y)$ とする。

（１）$Z= X+Y$ の分布

$$g(z)=\int_{x_{\min}}^{x_{\max}} f(x,z-x)dx=\int_{y_{\min}}^{y_{\max}} f(z-y,y)dy \tag{2.18}$$

X, Y が独立のときは、

$$g(z)=\int_{x_{\min}}^{x_{\max}} f_1(x)\,f_2(z-x)dx=\int_{y_{\min}}^{y_{\max}} f_1(z-y)f_2(y)dy \tag{2.19}$$

となる。これは二つの確率密度関数のたたみ込み積分である。

（２）$Z=X-Y$ の分布

$$g(z)=\int_{x_{\min}}^{x_{\max}} f(x,x-z)dx=\int_{y_{\min}}^{y_{\max}} f(z+y,y)dy \tag{2.20}$$

X, Y が独立のときは、式(2.21)になる。

$$g(z)=\int_{x_{\min}}^{x_{\max}} f_1(x)\,f_2(x-z)dx=\int_{y_{\min}}^{y_{\max}} f_1(z+y)\,f_2(y)dy \tag{2.21}$$

（３）$Z=X・Y$ の分布

$$g(z)=\int_{x_{\min}}^{x_{\max}} f\left(x,\frac{z}{x}\right)\frac{1}{|x|}dx=\int_{y_{\min}}^{y_{\max}} f\left(\frac{z}{y},y\right)\frac{1}{|y|}dy \tag{2.22}$$

X, Y が独立のときは、式(2.23) になる。

$$g(z)=\int_{x_{\min}}^{x_{\max}} f_1(x)\,f_2\left(\frac{z}{x}\right)\frac{1}{|x|}dx=\int_{y_{\min}}^{y_{\max}} f_1\left(\frac{z}{y}\right)f_2(y)\frac{1}{|y|}dy \tag{2.23}$$

（４）$Z= X/ Y$ の分布

$$g(z)=\int_{x_{\min}}^{x_{\max}} f\left(x,\frac{z}{x}\right)\frac{|x|}{z^2}dx=\int_{y_{\min}}^{y_{\max}} f(zy,y)|y|dy \tag{2.24}$$

X, Y が独立のときは、式(2.25) になる。

$$g(z)=\int_{x_{\min}}^{x_{\max}} f_1(x)f_2\left(\frac{x}{z}\right)\frac{|x|}{z^2}dx=\int_{y_{\min}}^{y_{\max}} f_1(zy)\,f_2(y)|y|dy \tag{2.25}$$

2.4　確率分布の積分変換

連続型分布を対象に、確率密度関数の積分変換の性質を示す。積分変換にはいくつかのタイプがあり、代表的なものは積率母関数（ラプラス変換形）、特性関数（フーリエ変換形）、ハンケル変換形特性関数（フーリエ・ベッセル変換形）である。ここではこの三つの積分変換とキュムラント母関数（積率母関数の対数をとったもの）について、定義と基本的な性質を説明する。いず

れも、多数の確率変数の和の分布を求めるときに有力な手段になる。

2.4.1　積率母関数

期待値が存在する確率変数 X において、確率密度関数 $f(x)$を以下のように積分変換して得られる関数は**積率母関数**（moment-generating function）と呼ばれる。

$$\Psi(s) \equiv \left\langle e^{sX} \right\rangle = \int_{-\infty}^{\infty} e^{sx} f(x) dx \tag{2.26}$$

電気数学等で習うラプラス変換の形である（指数部分に負符号が付いていないので逆ラプラス変換の形）。上式は、指数関数の s=0 を中心とするテイラー展開（＝マクローリン展開）を用いれば、以下のように、新しい変数 s のベキ級数に展開できる。

$$\Psi(s) = 1 + \left\langle X \right\rangle s + \frac{1}{2!}\left\langle X^2 \right\rangle s^2 + \frac{1}{3!}\left\langle X^3 \right\rangle s^3 + \cdots \tag{2.27}$$

式(2.27)は s に関するベキ展開の係数に、高次のモーメントが順次現れていることがわかる。このように、次々と高次のモーメントを生み出しているのが、**母関数**（generating function）と呼ばれる由縁である。

互いに独立な確率変数 $X_1, X_2, \cdots, X_n$ の積率母関数を $\Psi_1, \Psi_2, \cdots, \Psi_n$ とすると、X=X_1+X_2+ $\cdots$ +X_n の積率母関数 Ψは、定義式(2.26)から明らかなように、

$$\Psi(s) = \prod_{i=1}^{n} \Psi_i(s) \tag{2.28}$$

となり、それぞれの積率母関数の積になる。この性質が、複数の確率変数の和の分布を求めたい場合に役立つものになる。$\Psi(s)$から $f(x)$に戻す逆変換についてはラプラス変換に習ってほしい。

2.4.2　特性関数

積率母関数を一般化したものが**特性関数**（characteristic function）である。積率母関数は期待値が存在する分布のみに対して適用される変換であったが、特性関数はどのような分布に対しても必ず存在する。定義は次式である。

$$\Phi(t) \equiv \left\langle e^{jtx} \right\rangle = \int_{-\infty}^{\infty} e^{jtx} f(x) dx \tag{2.29}$$

定義式の形から、特性関数はフーリエ変換に対応していることがわかる。

さらに、特性関数も、(2. 27)式同様に、各次のモーメントを用いて次式で展開できる。

$$\Phi(t) = 1 + j\left\langle X \right\rangle t - \frac{1}{2!}\left\langle X^2 \right\rangle t^2 - j\frac{1}{3!}\left\langle X^3 \right\rangle t^3 + \cdots \tag{2.30}$$

互いに独立な確率変数 $X_1, X_2, \cdots, X_n$ の特性関数を $\Phi_1, \Phi_2, \cdots, \Phi_n$ とすると、X=X_1+X_2+ $\cdots$ +X_n の特性関数 Φは、積率母関数と同様、

$$\Phi(t) = \prod_{i=1}^{n} \Phi_i(t) \tag{2.31}$$

となり、それぞれの特性関数の積になる。電波伝搬の確率分布に表れる例では、ダイバーシチの最大比合成後の SN 比の分布を求める時にこの公式が利用される。

もとに戻す時は、反転公式が用いられる。確率分布関数 $F(x)$が x_1 と x_2 で連続であるならば、特性関数については、

$$F(x_2)-F(x_1)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-jx_1t}-e^{-jx_2t}}{jt}\Phi(t)dt \tag{2.32}$$

で求められる。この式(2.32)は**レウ″ィーの反転公式**（Levy's inversion formula）と呼ばれる。さらに、確率密度関数 $f(x)$が連続で、かつ、その特性関数が t に関して可積分であれば、フーリエ変換同様、次式が成り立つ。

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi(t)e^{-jxt}\,dt \tag{2.33}$$

電波伝搬に現れるほぼ全ての確率分布の反転公式には、(2.33)式が適用できる。

2.4.3　キュムラント母関数

積率母関数 Ψの対数をとったものを**キュムラント母関数**（cumulant generating function）と言い、次式で定義される。

$$K(s)=\log\Psi(s) \tag{2.34}$$

log (1+x)はテイラー展開により、

$$\log(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\cdots$$

であるので、キュムラント母関数は(2.27)式を用いて以下のように展開される。

$$K(s)=\log\left(1+\langle X\rangle s+\frac{1}{2!}\langle X^2\rangle s^2+\frac{1}{3!}\langle X^3\rangle s^3+\cdots\right)$$

$$=\langle X\rangle s+\frac{1}{2}\langle X^2\rangle s^2+\cdots-\frac{1}{2}\left(\langle X\rangle s+\frac{1}{2!}\langle X^2\rangle s^2+\cdots\right)^2+\frac{1}{3}\left(\langle X\rangle s+\frac{1}{2!}\langle X^2\rangle s^2+\cdots\right)^3-\cdots$$

$$=\sum_{m=1}^{\infty}\frac{s^m}{m!}\kappa_m \tag{2.35}$$

$$\kappa_1=\langle X\rangle=\mu$$

$$\kappa_2=\langle X^2\rangle-\langle X\rangle^2=\sigma^2$$

$$\kappa_3=\langle X^3\rangle-3\langle X^2\rangle\langle X\rangle$$

$$\kappa_4 = \langle X^4 \rangle - 4\langle X^3 \rangle \langle X \rangle - 3\langle X^2 \rangle^2 + 12\langle X^2 \rangle \langle X \rangle^2 - 6\langle X \rangle^4$$

$$\vdots$$

ここで、κ_m は **m 次のキュムラント**と言う。このキュムラントは３章 3.1.2 項の中心極限定理の証明の中で使われる。

2.4.4　ハンケル変換形特性関数

信号の包絡線など、正の値をとる確率変数（$R \geq 0$）に対しては、以下の式で与えられる**ハンケル**（Hankel）**変換形の特性関数**が便利である。

$$g(k) \equiv \langle J_0(kr) \rangle = \int_0^\infty J_0(kr) f(r) dr \tag{2.36}$$

逆変換による確率密度関数と累積分布関数は次式になる。

$$f(r) = r\int_0^\infty k J_0(rk) g(k) dk \tag{2.37}$$

$$F(r) = r\int_0^\infty J_1(rk) g(k) dk \tag{2.38}$$

ハンケル変換は**フーリエ-ベッセル変換**とも呼ばれる。ハンケル変換にはいくつかの変形があり、(2.36)式もその一つである。ハンケル変換に関する詳しい説明は本書の Appendix 2 で行っているので必要なときは、そちらを見てほしい。

この方法は多数の正弦信号が合成された信号の振幅（包絡線）の確率分布を求めたい場合に有力な手段になる。具体的には、

$$R = \left| R_1 e^{j\Theta_1} + R_2 e^{j\Theta_2} + \cdots + R_n e^{j\Theta_n} \right| \tag{2.39a}$$

$$f_{R_i}(r) = \delta(r - r_i), \quad f_{\Theta_i}(\theta) = \frac{1}{2\pi} \tag{2.39b}$$

のように、各正弦信号の振幅を表す確率変数 R_i は $R_i = r_i$ で一定値、位相を表す変数 Θ_i は $0 \sim 2\pi$ で一様に分布する場合である。n=1 の場合は明らかに、n=2 の場合は少し技巧を要す式変形を経て次式を得る。

$$g(k) = \begin{cases} J_0(kr_1) & (n=1) \\ J_0(kr_1)J_0(kr_2) & (n=2) \end{cases}$$

任意の n については Kluyver (1906)により導かれていて、

$$g(k) = \prod_{i=1}^{n} J_0(kr_i) \tag{2.40}$$

となる。

種々のケースについて得られた特性関数から確率分布に戻すには、(2.37)式や(2.38)式によるが、特殊関数を扱う数学公式集（例えば、[3]や[4]、あるいは数式処理ができるソフト）が必須になる。そして、多くの場合は、閉形式に導くのは難しく数値計算になろうが、特性関数さえ得られてしまえば、確率密度関数も累積分布関数も一重積分で求められるので、パソコン性能が十分なレベルにある今の時代においては、数値計算も大きな隘路にはならないであろう。ハンケル変換型特性関数は、４章で扱う仲上・ライス分布（仲上が導いた n 分布）（4.2 節）や複数定常波＋不規則波の分布（4.3 節）の解析に利用されている。

2.5　異なる分布の近似度評価指標

確率密度関数 $p(x)$を別の確率密度関数 $q(x)$で代用したいという場合がある。たとえば、ある確率分布が、他の関数と組み合わせて使う場合には、積分が残って扱いにくいために、解析性に優れた別の分布で近似したいというような場合である。あるいは、実測値の分布を得た場合に、それを特定の分布で近似したいというような場合である。そのときの２つの分布の差異を定量的に示す尺度が必要になる。

最も簡単な尺度は分布の距離を求めるノルムである。

$$\int \left| p(x) - q(x) \right| dx \qquad (L^1\text{ノルム}) \tag{2.41a}$$

$$\int \left\{ p(x) - q(x) \right\}^2 dx \qquad (L^2\text{ノルム}) \tag{2.41b}$$

上記の評価で十分な場合もあるが、分布の値の大きい部分の差が結果を支配するので、確率の小さい部分も含めて全体を見たい場合には不向きである。これに対して、次式で定義される**カルバック-ライブラー情報量**（Kullback-Leibler (KL) divergence；KL 情報量）がある。

$$D(p:q) = \int_{x_{\min}}^{x_{\max}} p(x) \ln \left\{ \frac{p(x)}{q(x)} \right\} dx \tag{2. 42}$$

２つの分布 p と q が同じ場合に、D=0 となり、ずれの程度に応じて値も大きくなる。置換目的であれば、置き換えたい分布 q のパラメータ値を調整して、D が最も小さくなるように最適化すればよい。

なお、この名前は、情報理論におけるカルバック-ライブラー情報量から来ている。直感的には二つの分布の距離のイメージであるが、上式で p と q を入れ替えたときの対称性が崩れているので、距離と呼びにくい事情がある。

(2.42)式で与えられるかるカルバック-ライブラー情報量は第９章で述べる AIC（赤池情報量規準）の構築に際して重要な働きをしている。

ティータイム　測度論的確率論

筆者の確率の知識は確率分布の伝搬モデルへの応用、すなわち、道具として利用するところから入っているので、数学的基礎が身についていない。確率を数学分野の一つとして体系的に理解するには測度論を学びなさい、といろいろの教科書に書いてある。測度論を解説している幾つかの本を読んでみると、我々が当たり前と思って受け入れていることも、根拠があいまいだということらしい。それを厳密な数学理論（公理から始めて、定理とその証明によって組み立てる理論体系）として現在の確率論の基礎を築いたのがコルモゴロフ（Andrei Nikolaevich Kolmogorov, 1933 年）で、それが測度論（集合の理論）に基づく公理主義的確率論である、と書かれている。

測度の解説書[5]を読んでのなまかじりの理解ではあるが、測度とは長さや面積に対応するもので、それを測るためには、構成する要素を見分けて数え上げることである。難しいのは範囲や数に無限大や不連続を含むようなときの扱いで、∞×0、0/0、∞/∞のようなことを議論するためには、この測度論が必須なのであると。このような例として、[5]では以下のような話題が取り上げられている。

【数学では、「点」の定義は位置以外の（大きさなどの）あらゆる特徴をもたないもの（ユーグリッドの「原論」では位置をもち部分をもたないもの)、「線」とは幅の無い長さであり、「直線」とはその上にある点について一様に横たわる線である、と決めている。その上に立って、以下の実数空間の長さを考えてみる。

１）0 と 1 の間にある実数全体を考える

２）実数（R）の数は無限にあるが、その並びを直線とみなし、長さを 1 としよう。
　　l_R=1 である。

３）実数は有理数（X）と無理数（Y）で構成される。

４）それぞれの長さを l_X, l_Y とすると、$l_X+l_Y=1$ である

５）では、それぞれの長さ l_X と l_Y はどう定められるか？

こういう難問に数学的（＝集合論的）答えを与えようとするのが測度論である。】

幸い（？)、我々が遭遇する伝搬現象での確率分布は、不連続なややこしさがない素直な特性である。ゆえに、測度論を持ち出さなくても（＝測度論の道具になるルベーグ積分を武器にしなくても）大丈夫だろうと思い、安心してしまうのである。それでよいのだろうか。

参考文献

[1] 伏見正則, *確率と確率過程*, 講談社, 1987.

[2] 唐沢好男, “伝搬環境のレイトレーシング：その後どうする？,” 技術レポート TR-YK-043:
http://www.radio3.ee.uec.ac.jp/ronbun/TR_YK_043_Hankel_Transform.pdf

[3] 森口繁一他, *岩波数学公式III*, 岩波書店, 1993.（新装版）
[4] 奥井重彦, *電子通信工学のための特殊関数とその応用*, 森北出版, pp. 154-157, 1997.
[5] 原啓介, *測度の考え方；測り測られることの数学*, 技術評論社, 2023.

目次のページはこちら