対数正規分布での平均値と CV 値

データには「ばらつき（分布）」があり、その形は統計分布関数で説明される。薬物動態解析では血中濃度や薬物動態パラメータの分布は正規分布あるいは対数正規分布で仮定されることが多い。それぞれの分布関数（確率密度関数）から平均値（期待値）、分散、標準偏差、変動係数が計算される。

(1) 正規分布（ガウス分布とも呼ばれる）。平均値を中心に対称に分布し、確率密度関数は次式。

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

データ $X = (x_1, x_2, x_3, \ldots, x_n)$ に対して、平均値（算術平均）、分散、変動係数はそれぞれ次式。

$$\mu = \frac{1}{n}\sum_{k=1}^{n} x_k$$

$$\sigma^2 = \frac{1}{n-1}\sum_{k=1}^{n}(x_k-\mu)^2$$

$$CV = \frac{\sqrt{\sigma^2}}{\mu} = \frac{\sigma}{\mu}$$

(2) 対数正規分布。対数変換したときに正規分布の形状となり、確率密度関数は次式。

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

データの対数をとったとき、すなわち $y_k = log(x_k)$ としたとき、平均値、（不偏）分散は次式：（上記(1)と μ 、σ^2 の定義は異なる）。

$$\mu = \frac{1}{n}\sum_{k=1}^{n} y_k = \frac{1}{n}\sum_{k=1}^{n} log(x_k)$$

$$\sigma^2 = \frac{1}{n-1}\sum_{k=1}^{n}(y_k-\mu)^2 = \frac{1}{n-1}\sum_{k=1}^{n}\{log(x_k)-\mu\}^2$$

平均値、分散、変動係数はそれぞれ次式。

$$exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

$$exp(2\mu+\sigma^2)\,[exp(\sigma^2)-1]$$

$$\frac{\sqrt{exp(2\mu+\sigma^2)\,[exp(\sigma^2)-1]}}{exp\left(\mu+\frac{1}{2}\sigma^2\right)} = \sqrt{exp(\sigma^2)-1}$$

（参考：指数関数の定義より、$a^{n/m} = \sqrt[m]{a^n}$ なので、

$$exp\left(\mu+\frac{1}{2}\sigma^2\right) = exp\frac{(2\mu+\sigma^2)}{2} = \sqrt{exp(2\mu+\sigma^2)}$$

で上式が成立する。

（続きます）

関係論文（S. A. Julious et al, *J. Biopharm. Stat.*, **10(1)** 55-71 (2000).）に同様の記載があり、特に、テイラー展開すると、分散が小さい場合において次式が成立することが触れられている。

$$CV = \sqrt{exp(\sigma^2) - 1} = \sqrt{\left(1 + \sigma^2 + \frac{\sigma^4}{2} + \ldots\right) - 1} \cong \sqrt{\sigma^2} = \sigma$$

データ $X = (x_1, x_2, x_3, \ldots, x_n)$ に対し算術平均、幾何平均は次で定義される。

$$m_A = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$m_G = \left\{\prod_{i=1}^{n} X_i\right\}^{\frac{1}{n}}$$

すなわち、

$$log(m_G) = \frac{1}{n}\sum_{1}^{n} log\{X_i\}$$

EXCEL による計算例を示す。

	A	B	C	D
1				
2			真値	対数値（ln）
3		1	4.6	1.526
4		2	5.4	1.686
5		3	5.1	1.629
6		4	3.5	1.253
7		5	4.7	1.548
8		6	5.2	1.649
9		7	5.8	1.758
10		8	4.1	1.411
11		そのまま平均	4.80	1.557
12		そのまま分散	0.549	0.027
13		変動係数	0.114	
14		真値の幾何平均	4.747	4.747
15		対数正規分布仮定時の平均値		4.810
16		対数正規分布仮定時の分散		0.622
17		対数正規分布仮定時の変動係数		0.164
18		〃		0.164

	A	B	C	D
1				
2			真値	対数値（ln）
3		1	4.6	=LN(C3)
4		2	5.4	=LN(C4)
5		3	5.1	=LN(C5)
6		4	3.5	=LN(C6)
7		5	4.7	=LN(C7)
8		6	5.2	=LN(C8)
9		7	5.8	=LN(C9)
10		8	4.1	=LN(C10)
11		そのまま平均	=AVERAGE(C3:C10)	=AVERAGE(D3:D10)
12		そのまま分散	=VAR(C3:C10)	=VAR(D3:D10)
13		変動係数	=C12/C11	
14		真値の幾何平均	=GEOMEAN(C3:C10)	=EXP(D11)
15		対数正規分布仮定時の平均値		=EXP(D11+D12/2)
16		対数正規分布仮定時の分散		=EXP(2*D11+D12)*(EXP(D12)-1)
17		対数正規分布仮定時の変動係数		=SQRT(D16)/D15
18		〃		=SQRT(EXP(D12)-1)