07/9/30,10/2

ー統計数学入門(不偏推定法)ー

❶結果から原因を解析、将来傾向を推定する方法＝**統計データ**:

<u>過去を知るは将来を予測する為にある</u>。しかしながら現実の複雑な動作原理を短時間でしかも専門知識もコストもかけずにまともに遣るのは難しい。そこで結果を測定した統計データを記録して表やグラフにして傾向を見るのが最も信頼性が高くしかも低コストでできる･

⑴サイコロの出目{1，2，3，4，5，6}の実現頻度(確率)をＮ回実際に振って調べる。

	n(1)		n(2)		n(3)		n(4)		n(5)		n(6)	
N=100	13	0.13	12	0.12	19	0.19	16	0.16	22	0.22	18	0.18
N=200	26	0.13	30	0.15	33	0.165	28	0.14	43	0.215	40	0.20
N=300	37	0.123	53	0.177	49	0.163	41	0.137	61	0.203	59	0.197
N=400	53	0.133	76	0.190	62	0.155	63	0.158	70	0.175	76	0.190
N=500	69	0.138	90	0.180	82	0.164	85	0.170	86	0.172	88	0.176
N=超大		0.1666		0.1666		0.1666		0.1666		0.1666		0.1666

⑵サイコロの出目{1，2，3，4，5，6}の**実現頻度(確率)≡個別事象の実現回数/全試行回数**はどれも等しく1/6＝0.1666……と一般に見なしてる。N=10回程度の試行では結構バラバラで1/6とはかなり外れた頻度になる。だがN=500回程度ではある落ちつた値に近ずく傾向が見えるだろう。もし精密なサイコロならばＮを超大にすれば1/6の等しい確率値が実現するだろう。

☞:上記表ではn(1)は小さく、n(6)は大き目、実は目1と目6は裏表にあり, 多分ここで使用されたサイコロでは重量や形状にバランスのズレがあるのだろう。

❷「なぜサイコロの出目は皆等しく1/6の確率なのか？」:**判らない事は等しい重みで計る！**。

もしどこかの目だけが特別に出易いという情報があるならば事情は異なるだろうが、

将来に起こりえる、想定される事象の集合E={a,b,c,････}のＮ個に何ら事前に確定した情報が無い場合はどれも等しく起こり得るとして扱うのが**最も妥当な推測**という事に反論はできない。

かような思想を**不偏推定**と呼ぶ。判らないと否定的に認める事が逆にどれにも等しい確率値を付与決定する逆説性に真髄がある事に留意！。

☞:宇宙はどうして生まれたか？、鶏の前には卵が、その前には鶏が、それは卵が、かくて原因→結果の因果律を遡る事になるが、宇宙が有から生まれるのならばその有はどうしてかの質問が可能になるので解答にならない。もし無から生まれるならば無は無しか生まない のが正常な因果律だろう。だとすると宇宙創始原点は判らない事になる。通信工学で言う**情報量無限大**[1)]になる。<u>情報量無限大では可能な力学的状態は全て**等確率**[1)]に成る</u>。実はこの条件を使うと宇宙初期条件が時間空間での特別な一点から生まれると言う**BIG.BANG理論**が得られ宇宙創始後の力学運動が逆に決定する事情がある。

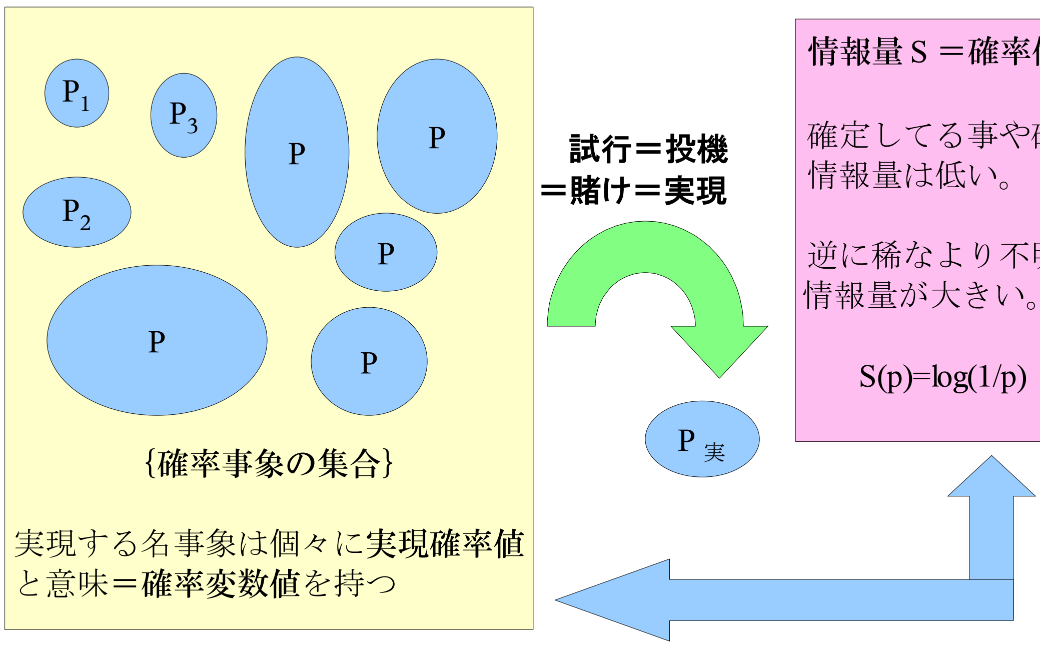
☞:本講座では小文字番号[1),2),3),....]を右肩につけて後に脚注補足を記載に注意！。

❸情報量の加算性質と情報測定量としての**エントロピー関数**:

10g と 20g の石を同時に乗せて計れば足し算の 30g になる。かような性質を加算性と言う。トランプでハートを引く確率＝P_H 、数字 7 を引く確率＝P_7、ハート 7 を引く確率 P_{H7}を考察。P_H＝1/4、P_7＝1/12、P_{H7}＝1/4X12＝1/48＝P_HXP_7.かように P_{H7} は 4X12=48 枚中の一枚の確率になる。掛け算に注意しよう。だとすると P_h の情報量と P_7 の情報量が加算的性質を備えるには掛け算が足し算になる**対数関数**[2)]が利用できる。Log(AxB)＝logA＋logB。

通信工学では情報を乗せた信号波に不可避的にランダム雑音が加算されるので信号検出は確率的になるので、特定の意味ある信号実現確率 p に付随して情報量 S(p)と言う確率値 p を変数とする関数概念が必要になった<***関数：数値 p を指定すると値 S が一つ決定の因果律***>。

S＝log(1/p) ＝−log(p)を「確率 p の情報量(エントロピー関数)」と言う。等確率では p＝1/W。W は実現可能な事象の全部の個数.かような関数を利用すると例えば平均値に関する情報だけがあると言う場合には確率(分布)が不偏推定原理(**S 値最大化**)で決定できると言う効用が発生し,実用に供されている。ただし所得分布は平均値だけでは非決定、だから他に重大な情報要因が他にもある事になる。

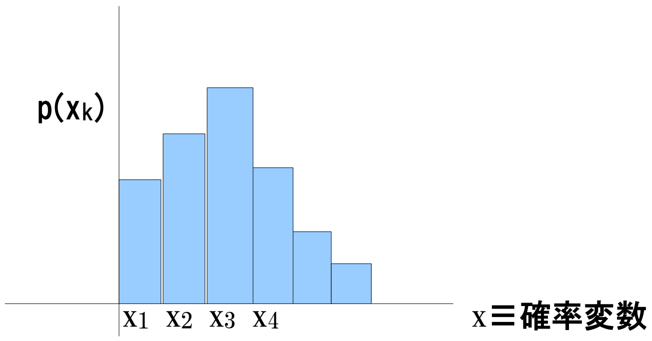


☞:**確率事象集合の正確な情報量の定義は個々の情報量の平均値となる。**

$S(p_1, p_2, p_3, \ldots\ldots, p_N) = \Sigma_{k=1}^{N} p_k \cdot \log(1/p_k)$.

$1 = \Sigma_{k=1}^{N} p_k \equiv p_1 + p_2 + p_3 \ldots\ldots + p_N$.

☞:確率変数 x の関数としての確率値≡**確率分布関数 p(x)**



例 1)サイコロ:確率変数 x は出目＝{ 1,　2,　3,　4,　5,　6},
確率値p(x)＝{1/6, 1/6, 1/6, 1/6, 1/6, 1/6}

例 2)電車に乗るあの人は一体幾らの所得に有るのだろうか？、その確率は？。
所得分布:確率変数 x は所得額層＝{0^100 万, 100^200, 200^300, ... }
分布値(確率値)　p(x)＝{ %,　. . . . . }

下記表は国税庁 2007/9 月公表の正式資料です。

所得額＝確率変数	平成 **14** 年度構成比率＝確率	平成 **18** 年度構成比率＝確率
100 万以下	1.9%	2.7%
100～200 万	5.0	6.9
200～300 万	10.9	12.0
300～400 万	18.2	17.7
400～500 万	17.8	17.2
500～600 万	14.1	12.9
600～700 万	9.7	9.1
700～800 万	7.2	6.6
800～900 万	4.7	4.5
900～1000 万	3.3	2.9
1000～1500 万	5.5	5.6
1500～2000 万	1.2	1.2
2000 万以上	0.6	0.7

因みに給与所得者総数 4485 万人＝100%. 財務省国民負担率推移表での H18 年国民所得＝374. 3 兆円、この二つの数字を使うと日本の所得構造が判るでしょう。問題は各所得層総額が正確に判らない。中間値で推計するしかないでしょう。

1):正確には情報量Ｓでなく,その(偏差値関数)^2＝$\Delta S^2 \equiv \Sigma_{k=1}{}^{N} p_k \langle \log(1/p_k) - S \rangle^2$.
Δs も確率値関数だが,確率値($p_1, p_2, p_3, \ldots, p_N$)のある値で ΔS→∞が証明可(**特異点**).
この時平均値Ｓが有限値だとしてもその偏差値 ΔS＝∞は**実質 S＝∞に等価**となる。
この時,運動量.エネルギ分散値も無限大,不確定性原理から創始の**点時間空間性**が判る.
<**統計集団**の概念と個々の実現である**統計標本**の概念の差異には注意要！>.
因みに宇宙は初期エネルギー値が正値だと流産宇宙に,負値だと宇宙創生になる。
これはゲージ場(一般化された Dalambert)方程式の成長消滅安定化を判断する
安定性基準量なる関係式から判別される<詳細は筆者の量子重力力学と超統一場論>。

2)指数関数は $y＝e^x$ で定義(P4, 5)される. 指数の性質から $e^{a+b}＝e^a \times e^b$。A＝e^a、B＝ e^b、
対数関数はその逆関数で x＝lny, だから a＋b＝ln($e^a \times e^b$)＝ln(AB)＝lnA＋lnB.

❹正規分布と中心極限定理の意味：

正規分布は名方面に頻出する分布関数の代表選手、なぜ代表選手になるかの一つ根拠が**中心極限定理**と言われる内容で、観測対象になる確率変数 x が各種多数の要素的確率変数の総計Ｎ個の加算総和である時、Ｎ→∞で成立すると言うのだ（証明は以下の❺）。酔歩問題と言うのが有名だが、これはふらふら一定幅前に進む確率と退く確率が等しいと初期位置から正規分布の偏差値が時間経過と共に増大する。水にインクをたらす拡散現象に類似で確率変数の極限的加算性がこの場合は容易に理解される。

⑴あるマクロに観測される物理量 x が要因変数 $\{x_1, x_2, x_3, \ldots, x_{N-1}, x_N\}$ に依存すると言うならばそれは一般に x は要因変数の因果律的な関数になる。

$$x = X(x_1, x_2, x_3, \ldots, x_{N-1}, x_N). \qquad \cdots\cdots\cdots\cdots\cdots\cdots (1)$$

勿論、x は一般には要素変数 $x = x_1 + x_2 + x_3 + \ldots + x_{N-1} + x_N$ の線形な総和になる事はない。**非線形依存**が一般と考えるのが常識だろう。

⑵ところが強い因果関係⑴の成立の下に**要素変数がある平衡点から微小に確率的変動**にあると仮定できると直に要素変数 x_k の総和ではないにせよその定数倍 $= A_k \cdot x_k \equiv y_k$ の確率変数に再定義すると総和的にする事が出来る。**固定平衡点**を中心にＸを Taylor 展開してその一次を取る.

$$x = X(x_1, x_2, x_3, \ldots, x_{N-1}, x_N) \equiv X(x_k) = X(x_k \equiv \delta x_k + a_k).$$
$$= X(a_k) + \Sigma_{k=1}^{N} \partial_k X(x_k = a_k) \cdot \delta x_k + \tfrac{1}{2}\Sigma_{j=1}^{N}{}_{,k=1}^{N}\, \partial_j \partial_k X \cdot \delta x_j \delta x_k + \cdots +$$
$$\doteqdot X(a_k) + \Sigma_{k=1}^{N} \partial_k X(x_k = a_k) \cdot \delta x_k \equiv \underbrace{X(a_k)}_{\text{平均値}} + \underbrace{\Sigma_{k=1}^{N} y_k}_{\text{ランダム総和}} \equiv \Sigma_{k=0}^{N} y_k. \qquad \cdots\cdots (2)$$

$x = \Sigma_{k=0}^{N} y_k$.　〈マクロ観測値〉.

$y_0 \equiv X(a_k)$ ＝平均値＝偏差値０の正規分布極限。.
$y_k \equiv \partial_k X \cdot \delta x_k$. ＝揺動値。

⑶この模型成立に必要な条件は中心極限定理本来の仮定であるＮ→∞で名確率変数の総和への寄与が０微小量に成る事から $\partial_k X < \infty$ である限り、$y_k \to 0$ が成立するはずである。但し上記仮定では $y_0 \equiv X(a_k)$ だけは有限値である。もしこの有限性が問題になるならば確率変数を $x \equiv x - X(a_k)$ に再規格化すればよかろう。y_0 が消去できる。

⑷現実の確率統計現場は**因果的要素に確率揺動要素が混合する系**と考えると上記非線形模型の近似条件下での線形表現はより自然でなかろうか？。

⑸巨視的化学系の人間身長や体重には**「ある明白な平均値」**が存在する事こそが $y_0 \equiv X(a_k)$ の有限値になる**強い因果律の存在**（身長体重では遺伝子）の意味でないのだろうか。なほ下の証明では名変数の分布性は平均値と偏差値が有限に成る以外に名確率変数の中心極限性だけしか仮定を必要としていない。Ｎ無限大の中心極限性も**平均値からの微小揺動値に変換できれば**無理のないＮ有限でも等価的に中心極限性仮定に相当するのでないか。仮定する微小変動とは展開一次で高精度近似できるの意味。

❺中心極限定理の簡易な証明法。

教科書やサイトを見ても判り易い証明法が見えないので補足する。
証明内容を知らないで定理応用には危険がある。内容納得で試用が望ましいが、
筆者自身も確率変数の超総和的現象の実解析経験がないので実感がよく判らない。

①正規分布と積率母関数、キュムラント関数：

⑴正規分布　：$f(x)=\exp[-(x-m)^2/2\sigma^2]/\sigma\sqrt{(2\pi)}$.

⑵積率母関数: $\phi(\omega)=\int_{-\infty}^{\infty}dx e^{i\omega x}f(x)=\exp[i\omega m-\omega^2\sigma^2/2]$.

⑶キュムラント関数：$\ln\phi(\omega)=i\omega m-\omega^2\sigma^2/2$.

②個々に**独立な平均値** m_k **と分散** σ_k^2 **が確定する分布関数** $f_k(x_k)$ が $k=1,2,...,N$ とある。
確率変数 x_k の**総和値** x を確率変数とする分布関数 $f(x)$ とその積率母関数 $\phi(\omega)$：

$$x=\Sigma_{k=1}^{N}x_k..$$
$$f(x)=\Pi_{k=1}^{N-1}\int dx_k f(x_1,x_2,....x_{N-1},x_N=x-\Sigma_{k=1}^{N-1}x_k)$$
$$=\Pi_{k=1}^{N-1}\int dx_k\cdot f_1(x_1)f_2(x_2)....f_{N-1}(x_{N-1})f_N(x_N=x-\Sigma_{k=1}^{N-1}x_k).$$

以下では関係 $dx_N=d(x-\Sigma_{k=1}^{N-1}x_k)=dx$ で積分できる事に留意.
確率変数和は**密度関数の畳込み**になり、Fourier 変換では単純に個々の**積**になる。

$$\phi(\omega)=\int_{-\infty}^{\infty}dx e^{i\omega x}f(x)$$
$$=\int_{-\infty}^{\infty}dx e^{i\omega x}\cdot\Pi_{k=1}^{N-1}\int dx_k\cdot f_1(x_1)f_2(x_2)....f_{N-1}(x_{N-1})f_N(x_N=x-\Sigma_{k=1}^{N-1}x_k)$$
$$=\int_{-\infty}^{\infty}dx\cdot\Pi_{k=1}^{N-1}\int dx_k\cdot e^{i\omega x_k}f_k(x_k)e^{i\omega x_N}f_N(x_N=x-\Sigma_{k=1}^{N-1}x_k)$$
$$=\int_{-\infty}^{\infty}dx_N\langle\Pi_{k=1}^{N-1}\int dx_k\cdot e^{i\omega x_k}f_k(x_k)\rangle e^{i\omega x_N}f_N(x_N)=\Pi_{k=1}^{N}\phi_k(\omega).$$

③ $\mathbf{x=\Sigma_{k=1}^{N}x_k/N}$. **の平均値で確率変数を再定義**してる事に注意。<..>≡**期待値**。
「以下ではN→∞とするので $(x_k/N)\to 0$ に成る事に注目すれば展開近似が利用できる」。
以下で見える事だが $(x_k/N)\to 0$ が本質的に効果してる。$e^{i\omega x_k/N}\equiv\exp[i\omega(x_k/N)]$.

＊　$e^x=1+x+(1/2)x^2+......$

＊　$\ln[1+x]=x-(1/2)x^2+...$

$$\psi(\omega)\equiv\Pi_{k=1}^{N}\psi_k(\omega)\equiv\Pi_{k=1}^{N}\int_{-\infty}^{\infty}d(x_k/N)e^{i\omega x_k/N}f_k(x_k/N)$$
$$\fallingdotseq\Pi_{k=1}^{N}\int_{-\infty}^{\infty}d(x_k/N)f_k(x_k/N)[1+i\omega(x_k/N)-(1/2)\omega^2(x_k/N)^2]$$
$$=\Pi_{k=1}^{N}[1+i\omega\langle x_k/N\rangle-(1/2)\omega^2\langle x_k/N\rangle^2].$$

$$\ln\psi(\omega)=\Sigma_{k=1}^{N}\ln[1+i\omega\langle x_k/N\rangle-(1/2)\omega^2\langle x_k/N\rangle^2]$$
$$\fallingdotseq\Sigma_{k=1}^{N}\ \{i\omega\langle x_k/N\rangle-(1/2)\omega^2\langle x_k/N\rangle^2-(1/2)[i\omega\langle x_k/N\rangle-1/2)\omega^2\langle x_k/N\rangle^2]^2\}$$
$$\fallingdotseq\Sigma_{k=1}^{N}\ \{i\omega\langle x_k/N\rangle-(1/2)\omega^2\langle x_k/N\rangle^2-(1/2)[-\omega^2\ \langle x_k/N\rangle^2-+...]\}$$
$$=\Sigma_{k=1}^{N}\{i\omega\langle x_k/N\rangle-(1/2)\omega^2[\langle x_k^2/N^2\rangle-\langle x_k/N\rangle^2]\}$$
$$=\Sigma_{k=1}^{N}[i\omega\langle x_k/N\rangle\rangle-\omega^2(\sigma_k^2/N^2)/2].$$

$\ln\psi(\omega)=i\omega\Sigma_{k=1}^{N}\langle x_k/N\rangle\rangle-(\omega^2/2)\Sigma_{k=1}^{N}(\sigma_k/N)^2=i\omega m-(\omega^2/2)\sigma^2$. <証明終わり>