①

□ 正規分布の式 $y=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ をどう導くか。

飯島光治

はじめに　以下「統計法とその教育上の応用」（半田正吉，矢野建治* 共立出版 昭和19刊）の内容より。理由は全体的に高校数学の範囲で，取り扱われていることより。

なお この内容を始め読んだ時，よくわからなかった箇所につき，その後私なりに理解した所は，このレポートに追加しました。

* 2人の著者共，当時 奈良女子高等師範学校教授。

0）前提 ―誤差の法則について―

イ）誤差の定義 ＝ 測定値－真の値 ①

しかし 真の値は 得られないので，そのかわりに最確値（ふつう 平均値 M を用いる。）を用いる。

測定値－M ②　①を 真の誤差，②を みかけの誤差 という。

ここで 測定の回数を 多くすればする程 ②は，①に近づくと 考えられる。

ロ）誤差には
(1) 定誤差
(2) 過失による誤差
(3) 偶然による誤差

がある。(1)，(2)は 避けられ，(3)は 避けられない誤差といわれる。(1)には 器具の一定の狂い，目盛りのつけ方が 不完全，測定者のくせ。

外部からの原因がわかる防害による等の誤差。(2)は計算ミス，数の読みまちがい等。(3)は一陣の突風による，温度の突然の昇降による器具の伸縮に基く，その他全く未知の外部からの原因による等の。

ハ) 誤差の法則

種々のデータからの誤差の分布表より

(1) 大きな誤差は小さな誤差より起こる回数が少ない。

(2) 誤差の分布は，対称的である。即ち同じ大きさの正の誤差と負の誤差とは同じ回数だけ起こる。

(3) 余り大きい誤差は起こらない。

以下 偶然の誤差について，考える。

I) 計算による(1)

ある量の真値を g，N 個の測定値を $X_1, X_2, \cdots, X_N$ とし，誤差を $x_1, x_2, \cdots, x_N$ $(x_k = X_k - g)$ とする。また各誤差の起こる確率を $y_1, y_2, \cdots, y_N$ とする。

ここにいう誤差 x_k の起こる確率とは，x_k を含むある小範囲内の誤差のおこる確率とする。

さてある範囲内の誤差とその起こる度数との間には，一定の関数関係があるようにみえる。そこで誤差 x とその起こる確率 y との間には，$y = \varphi(x)$ なる関数関係があるものとして，未知の関数 $\varphi(x)$ の形が，どうなるかを，目的とする。

真の値 q は、決して知る事ができない。
そこで q に多くの試みの値を仮定し、その結果得られる誤差の試みの中から、上記誤差の法則から考えて、最も自然的分布と思われるものを選択することによって $\varphi(x)$ の形を定める。

この意味で qを変数として考えると、誤差 $x_1, x_2, \cdots, x_N$ は q の関数となる。従って誤差の確率 $y_1, y_2, \cdots, y_N$ もまた q の関数となる。

ここで最も自然的な誤差の分布とは試みの一通りの誤差 $x_1, x_2, \cdots, x_N$ の全体的にみた起こる確率が（ある1つの一通りが同時に起こる確率が）最大になる場合になると考えて以下を進める。

ここで測定の誤差 $x_1, x_2, \cdots, x_N$ は決して同時に得られるものではないが、これを同時に起こるものと考えて良い。なぜなら偶然による誤差より $x_1, x_2, \cdots, x_N$ は独立より。

よって互いに独立な事象が同時におこる確率は、

$$P = y_1 y_2 \cdots y_N = \varphi(x_1)\varphi(x_2)\cdots\varphi(x_N) \cdots (1)$$

で、P が最大となるように $\varphi(x)$ の形を定める。

以下 P は q の関数として、q について微分可能で、唯1つの極大をもち、極小はないものとする。

この仮定により P は、$\dfrac{dP}{dq} = 0$ をみたす q の値に対して最大となる。

以上、以下の計算の方針です。

< 計算(1) >

p3の(1)の両辺に、底をe(自然対数)とする対数をとれば、(1)の右辺は、積が和になり、次に両辺をqについて微分すると、合成関数の微分法により

$$\frac{1}{p}\frac{dp}{dq} = \frac{d}{dx_1}\left[\log \varphi(x_1)\right]\frac{dx_1}{dq} + \frac{d}{dx_2}\left[\log \varphi(x_2)\right]\frac{dx_2}{dq} + \cdots + \frac{d}{dx_N}\left[\log \varphi(x_N)\right]\frac{dx_N}{dq} \quad \text{(注)}$$

(注) 誤差 $x_1, x_2, \cdots, x_N$ は始め定数としてであったが、q を変数とみることによって、$x_1, x_2, \cdots, x_N$ も変数となる。

ここで $x_K = X_K - q$ より $\dfrac{dx_K}{dq} = -1$ $\left(\begin{array}{l}K=1,2,3\\ \cdots, n\end{array}\right)$

従って $\psi(x) \overset{\text{(注)}}{=} \dfrac{d}{dx}\left[\log \varphi(x)\right]$ とおくと

(注) $\psi(x)$ と $\varphi(x)$ は混同しやすい。ψ はプシー、φ はファイ。(or フィー)

$\dfrac{dp}{dq} = 0$ ならば

$$-\frac{1}{\varphi(x_1)} - \frac{1}{\varphi(x_2)} - \cdots - \frac{1}{\varphi(x_N)} = 0 \quad \text{より}$$

$$\underline{\psi(x_1) + \psi(x_2) + \cdots + \psi(x_N) = 0} \quad (1)$$

また測定数 N が十分大きければ、誤差の法則(2)(2頁)より

$$\underline{x_1 + x_2 + \cdots + x_N = 0} \quad (2)$$

次に $x_1, x_2, \cdots, x_N$ を定数と考え、

$x'_r = x_r + \varepsilon$, $x'_s = x_s - \varepsilon$ (ε は極めて絶対値の小さい数とする。)

とおくと、$x_1, x_2, \cdots, x'_r, \cdots, x'_s \cdots x_N$ は これらの和は0をみたし、誤差の一系列とみなせる。従って

$$\psi(x_1) + \psi(x_2) + \cdots + \psi(x'_r) + \cdots + \psi(x'_s) + \cdots + \psi(x_N) = 0$$

も満足しなくてはならない。

(1)−(3)をすると $\psi(x_r)-\psi(x_r')+\psi(x_s)-\psi(x_s')=0$

$$\therefore \psi(x_r)+\psi(x_s)=\psi(x_r')+\psi(x_s')$$

$$\therefore \psi(x_r)+\psi(x_s)=\psi(x_r+\varepsilon)+\psi(x_s-\varepsilon)$$

$$\therefore \frac{\psi(x_r+\varepsilon)-\psi(x_r)}{\varepsilon}=\frac{\psi(x_s)-\psi(x_s-\varepsilon)}{\varepsilon}=\frac{\psi(x_s-\varepsilon)-\psi(x_s)}{-\varepsilon}$$

$\varepsilon\to 0$ とすると

$$\left[\frac{d}{dx}\psi(x)\right]_{x=x_r}=\left[\frac{d}{dx}\psi(x)\right]_{x=x_s}$$

x_r, x_s は $x_K\ (K=1,2,\cdots,n)$ の中の任意の2つより

$$\left[\frac{d}{dx}\psi(x)\right]_{x=x_1}=\left[\frac{d}{dx}\psi(x)\right]_{x=x_2}=\cdots=\left[\frac{d}{dx}\psi(x)\right]_{x=x_N}$$

以上の与えられた誤差の系列については

$\psi(x)$ の導関数の値は一定。

ここで $x_1, x_2, \cdots, x_N$ は誤差の系列の任意のものと考えてよいから $\psi(x)$ の導関数の値は、x のすべての値に対して一定となり

$$\frac{d}{dx}\psi(x)=K\ (K\text{は定数})\quad (4)$$

<コメント> 以上 $x_r+\varepsilon$, $x_s-\varepsilon$ を考えるから始まりその後の展開は、仲々思いつきません。

(4)より $\psi(x)=Kx+C'$ (C'は定数)

これを4頁の(1)に代入すると

$$K(x_1+x_2+\cdots+x_N)+NC'=0$$

4頁の(2)より $C'=0$ $\quad\therefore \psi(x)=Kx$

$$\therefore \frac{d}{dx}[\log\varphi(x)]=Kx \rightarrow \log\varphi(x)=\frac{1}{2}Kx^2+C''$$

(C''は定数)

$$\therefore\ \varphi(x)=e^{\frac{1}{2}kx^2+c''}=ce^{\frac{1}{2}kx^2}\quad(c=e^{c''}>0)$$

すると$\varphi(x)$は，偶関数（$\varphi(x)=\varphi(-x)$）より
$x\geqq 0$で考えればよく，また誤差の法則より
大きな誤差は，小さな誤差より少ないから
x（＝誤差）を大きくすると，$\varphi(x)$は減少する。

よって　$\underline{k<0}$

∵）$\varphi(0)=c\ (>0)$で　$\varphi'(x)=ce^{\frac{1}{2}kx^2}\times kx$
$x\geqq 0$で，$\varphi(x)$が減少関数のためには
$\varphi'(x)\leqq 0$　　　$\therefore k<0$　〃

＜コメント＞　ここの所，上記の本には"小さな誤差は，大きな誤差より多く起こるのであるから，kは負でなければならない。"とあり始めわかりませんでした。

よって　$\frac{1}{2}k=-h^2$　（$\underline{h>0}$（注））　とおくと

$$\underline{\underline{\varphi(x)=ce^{-h^2x^2}}}$$ の形となる。

（注）なぜ$h>0$かは，この段階ではよくわからず。
次の計算（2）にくると，$A=\int_{-\infty}^{\infty}ce^{-h^2\xi^2}d\xi=\frac{c\sqrt{\pi}}{h}$
を用いますが，Aはこの時面積より$A>0$
かつ$c\ (=e^{c''})>0$より　　$h>0$　となります。

（Ⅱ）計算（2）

次の課題は，上記の式で具体的にc，hの式を決めていく。　になります。

ここで以上（Ⅰ）までは抽象的に真の誤差の法則についてですが，実際問題では，みかけの誤差を知るのみなので以下はみかけの誤差について考える。

ところで，みかけの誤差の分布と真の誤差の分布とはみかけの誤差で，Nを十分大きくすると，全く同形であると考えられるから，みかけの誤差ξ（クシー，グザイ）の確率を

y_kとすれば　$y = ce^{-h^2\xi^2}$ $(c>0, h>0)$ (5) とかける。 ⑦

さて　ある量N個の測定値を　区分の大きさΔなる等しい階級に分けたとして、それらの階級の

中央値を $x_1, x_2, \cdots, x_n$、度数を $f_1, f_2, \cdots, f_n$

$(\sum f_k = N)$ とする。　測定値の平均値をMとすれば

$\xi_1 = x_1 - M, \quad \xi_2 = x_2 - M, \quad \cdots, \quad \xi_n = x_n - M$

ここで　ξ_kの起こる確率y_kは$\dfrac{f_k}{N}$と考えられる。

以後　Nは十分大きく、また$f_k \neq 0$　従って$y_k \neq 0$ とする。

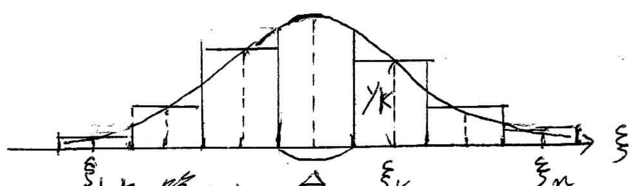
$y_1 = \dfrac{f_1}{N}, \quad y_2 = \dfrac{f_2}{N}, \quad \cdots, \quad y_n = \dfrac{f_n}{N}$

即ち　$\sum y_k = 1$　　従って　この頁の(5)より

$\sum y_k = \sum ce^{-h^2\xi_k^2} = 1$

いいかえると　$\dfrac{1}{\Delta}\sum y_k \Delta = 1$

ここで　$\sum y_k \Delta$は、以下のグラフのヒストグラムの面積を表す。(＊)



(注)　上記の本では、計算(1)の前で

$\varphi(x) = ce^{-h^2x^2}$ の $c = \dfrac{1}{\sqrt{2\pi}}$, $h = \dfrac{1}{\sqrt{2}}$ として

$\varphi(x)$のグラフを、xの値を与えて$\varphi(x)$の値を求め点(座標)をとって、上記のグラフを描いている。

ただ　c, hの値をなぜ上のようにおくかは、この段階では、わからない。

(＊)の続き　$\Delta \to 0$とすると、曲線の面積Aに等しくなる。(上記の)

次に $A=\int_{-\infty}^{\infty} ce^{-h^2\xi^2}d\xi=\dfrac{c\sqrt{\pi}}{h}$ (注) がいえる。(*)

$\left(\because \int_0^{\infty} e^{-x^2}dx=\dfrac{\sqrt{\pi}}{2}\right.$ がいえ、この結果を用いると

この時 $h\xi=t$ とおくと $\xi=\dfrac{t}{h}$ $\therefore d\xi=\dfrac{1}{h}dt$

$$\int_0^{\infty} e^{-h^2\xi^2}d\xi=\int_0^{\infty} e^{-t^2}\cdot\frac{1}{h}dt \quad (h>0\text{より})$$

$$=\frac{1}{h}\int_0^{\infty} e^{-t^2}dt=\frac{\sqrt{\pi}}{2h}$$

この時、$\int_{-\infty}^{\infty}$ は $\int_0^{\infty}$ の2倍より

$$\left.A=\int_{-\infty}^{\infty} ce^{-h^2\xi^2}d\xi=\frac{c\sqrt{\pi}}{h}\right)$$

(コメント) この本の内容で唯一(ゆいいつ) $\int_0^{\infty} e^{-x^2}dx=\dfrac{\sqrt{\pi}}{2}$ が高校の範囲外です。これを証明するのは易しくありません。が微積分、他の本に載っています。

(*)に続く。 従って $\sum y_k\Delta \fallingdotseq \dfrac{c\sqrt{\pi}}{h}$

$\dfrac{1}{\Delta}\sum y_k\Delta=1$ より $\dfrac{1}{\Delta}\cdot\dfrac{c\sqrt{\pi}}{h}\fallingdotseq 1 \to c\fallingdotseq\dfrac{h\Delta}{\sqrt{\pi}}$

$$\therefore y_k\fallingdotseq\frac{h\Delta}{\sqrt{\pi}}e^{-h^2\xi_k^2}\quad (k=1,2,\cdots,n)$$

次に hの値を定めようとする。

今 ξ_1なるf_1コの誤差, ξ_2なるf_2コの誤差, …… ξ_nなるf_nコの誤差が、同時に起こる確率

$$p=(\underbrace{y_1y_1\cdots y_1}_{f_1\text{コ}})(\underbrace{y_2y_2\cdots y_2}_{f_2\text{コ}})\cdots\cdots(\underbrace{y_ny_n\cdots y_n}_{f_n\text{コ}})$$

$=y_1^{f_1}y_2^{f_2}\cdots y_n^{f_n}$ が最大になるようにする。

これを書き直すと

$$p=\left(\frac{h\Delta}{\sqrt{\pi}}e^{-h^2\xi_1^2}\right)^{f_1}\left(\frac{h\Delta}{\sqrt{\pi}}e^{-h^2\xi_2^2}\right)^{f_2}\cdots\cdots\left(\frac{h\Delta}{\sqrt{\pi}}e^{-h^2\xi_n^2}\right)^{f_n}$$

$$=\left(\frac{h\Delta}{\sqrt{\pi}}\right)^N e^{-h^2\sum f_k\xi_k^2}\overset{(*)}{=}\left(\frac{\Delta}{\sqrt{\pi}}\right)^N h^N e^{-Nh^2\sigma^2}\ (**)$$

(＊) ∵) $\left(e^{-h^2\xi_1^2}\right)^{f_1} = \underbrace{\left(e^{-h^2\xi_1^2}\right)\left(e^{-h^2\xi_1^2}\right)\cdots\left(e^{-h^2\xi_1^2}\right)}_{f_1 \text{コ}}$

$= e^{-h^2(\underbrace{\xi_1^2+\xi_1^2+\cdots+\xi_1^2}_{f_1 \text{コ}})} = e^{-h^2 f_1 \xi_1^2}$

以下同様。するとeの指数は

$-h^2(f_1\xi_1^2+f_2\xi_2^2+\cdots+f_n\xi_n^2)$ より。

(＊＊) ∵) $\sigma^2 = \frac{1}{N}\sum f_K \xi_K^2 \quad (\xi_K = x_K - M)$ より。

このPを極大ならしめる（この時同時に最大）hの値は，

$\frac{dP}{dh} = 0$

〈コメント〉hは始め定数であったが（$\frac{1}{2}K = -h^2$）であったがここでは変数としている。こういう文字の用法はたとえば $f(a)$ の a。aは始め定数。次に変数とみて，導関数がえられた。

この時 $\frac{dP}{dh} = \left(\frac{\Delta}{\sqrt{\pi}}\right)^N e^{-Nh^2\sigma^2} N h^{N-1}(1-2h^2\sigma^2) = 0$

（積の微分法の公式より）

$\therefore 1-2h^2\sigma^2 = 0 \rightarrow h = \frac{1}{\sqrt{2}\sigma}$ （$h>0$ より）

$c = \frac{h\Delta}{\sqrt{\pi}}$ (注) より $c = \frac{\Delta}{\sqrt{2\pi}\,\sigma}$

(注) $c \fallingdotseq \frac{h\Delta}{\sqrt{\pi}}$（8頁）であったが $n \rightarrow \infty$ にすることにより。

P.7の1行目の式(5)に代入して．

$\therefore y = \frac{d}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{\xi}{\sigma}\right)^2}$ （$\Delta = d$ とした。）

ここで $Y = \frac{y\sigma}{d}$ (＊) とおくと $y = \frac{d}{\sigma}Y$ より，また $x = \frac{\xi}{\sigma}$

（大文字のY）とおき，

$$Y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

がえられる。

（ここで改めて $Y \rightarrow y$ としてよい。但しすぐ上のyとはちがう意味として）

(＊) このおきかえは，私が。

（付記1）

上記の本では、ガウス（独 1777〜1855）の1809年の論文（誤差の研究）の方法により とありますが、

羽鳥裕久先生の「Gaussは、どのように正規分布を導いたか」（「数学の小さな旅」（近代科学社 1992刊 P139〜P142）をみますと だいぶ展開の仕方が違います。それで 以上のこのレポートの内容は、上記の2人の著者の方々が考えたのではないかと思われます。

（付記2）

正規分布の式を導くのに、二項分布→正規分布への説明があります。誤差の法則からより この方が、多いと思います。

たとえば「二項分布から正規分布へ」（田島一郎、数学セミナー 1974.3月号。後に「数学セミナーリーディングス 確率・統計＋近似・誤差」日本評論社 1979刊 所収）

（2009.3.27 記）

（付記3）

「正規分布の

$$y=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}$$

の式と P9↑7行目の $y=\dfrac{d}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{\xi}{\sigma}\right)^2}$ の式とを比べると、$\xi=x-m$ より 1とdの違いに。

即ち $d=1$ としたものが、正規分布の式になります。$\underbrace{\text{以上の}}_{\text{レポートの}}$内容からいいますと。