確率の基礎

用語の定義

多くの場合，ひとつの物理量の測定を繰り返してもさまざまな理由で一定の値とはならない。たとえば，定常電流を測定するとしよう。ある一定時間内に検出器に流れ込む電荷は一定の平均値のまわりに揺らぐだろう。このとき電荷を表す変数を X とすれば，X を確率変数（random variable）という。X の値は，測定のたびに異なる値をとることになるだろうが，その値の分布について一定の規則があれば，それを X が従う確率分布（probability distribution）という。X を連続として考えて良いばあいもあるが，たとえば電子の個数になおさなくてはならない場合には不連続量となる。

X が x と $x+dx$ の値として実現される確率が $f(x)dx$ のとき $f(x)$ を X の確率密度関数という。X が x 以下の任意の値で実現する確率は

$$F(x)=\int_{-\infty}^{x} f(x)dx$$

であり，累積分布関数（cumulative distribution function）という。不連続な分布のときは X が x_i という値をとる確率を p_i とすると

$$F\left(x\right)=\sum_{x_i \le x} p_i$$

である。

X のとりうる値の平均値は期待値ともよばれ

$$E\left[X\right]=\sum_i x_i p_i=\int_{-\infty}^{\infty} xf(x)dx$$

である。一方，平均値のまわりのどれぐらい範囲にとりうる値がばらつくかは

$$V\left[X\right]=E\left[\left(X-E\left[X\right]\right)^2\right]$$

で評価できる。この値を分散という。

期待値の性質として

$$E\left[aX\right]=\sum_i\left(ax_i\right)p_i=\int_{-\infty}^{\infty} axf\left(x\right)dx=aE\left[X\right]$$

である。ただし a は実数の定数である。したがって

$$V\left[X\right]=\sum\left(x_i^{\,2}-2E\left[X\right]\cdot x_i+E\left[X\right]^2\right)p_i=\sum x_i^2 p_i-2E\left[X\right]\cdot\sum x_i p_i+E\left[X\right]^2=\sum x_i^2 p_i-E\left[X\right]^2$$

である。連続分布の場合にもまったく同じである。

２項分布とは

コインを１回投げて表がでる（値を１とする）か裏がでる（０とする）かを測定する場合，

確率変数は０と１をとる。それぞれが出現する確率が p, q のとき，n 回投げ（試行という）このＮを非常に大きくすると，それぞれが実現する回数は pn, qn に近づくと予想できる。これを大数の法則といい，確率分布を実験から推定する方法の基礎となる。確率は $0 \le p \le 1, 0 \le q \le 1$ であり，$p+q=1$ を満たす。

２項分布の形

n 回の試行で現れた０と１を全て加えたものも確率変数であり，これを X としよう。X が k という値をとる（表が k 回出る）確率は

$$P(X=k)=\frac{n!}{k!(n-k)!}p^k q^{n-k}$$

である。これを２項分布と呼ぶ。もちろん

$$1=(p+q)^n=\sum_{k=0}^{n}\frac{n}{k!(n-k)!}p^k q^{n-k}$$

２項分布の期待値

２項分布の期待値が

$$E[X]=pn$$

であることはすぐ上に述べた大数の法則から明らかだが

$$\sum_{k=0}^{n}k\frac{n}{k!(n-k)!}p^k q^{n-k}=pn\sum_{k=1}^{n}\frac{(n-1)}{(k-1)!((n-1)-(k-1))!}p^{k-1}q^{(n-1)-(k-1)}$$
$$=pn\sum_{h=0}^{m}\frac{m}{k!(m-h)!}p^h q^{m-h}=pn$$

のように計算しても簡単に求まる。

期待値の性質（確率変数の和について：２項分布を例として）

期待値を求めるための別の方法もあるが少し準備を要する。しかし後の計算で利用する考え方だから丁寧に見ることにしよう。たとえば n 回の試行で得た結果が

$$1,0,0,0,0,1,0,1,\cdots,\underset{\hat{\lambda}}{0},\cdots1,0,1,1,1$$

であったとする。すなわち λ 番目の試行の確率変数 $X^{(\lambda)}$ の測定値が $x^{(\lambda)}=0$ である。λ 番目でこの測定値を得る確率は p である。こうして，上の例が実現する確率は

$$q\times p^3\times q\times p\times q\times\cdots\times p\times\cdots\times q\times p\times q^3$$

となっている。$q=p_1, p=p_2$ および $0=x_1, 1=x_2$ と書き直しさらに λ 番目の試行であることを上つき添え字で書き加える。n 回の試行を終えたとき１が出た回数は、その期待値が

$$\text{期待値}=\sum_{\text{全ての場合}}\left(\text{確率変数がとりうる値}\right)\times\left(\text{その値が実現する確率}\right)$$

$$=$$

$$\sum_{\substack{i_1=ura,omote\\ i_2=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(x_{i_1}^{(1)}+x_{i_2}^{(2)}+\cdots+x_{i_n}^{(n)}\right)\times\left(p_{i_1}^{(1)}\cdot p_{i_2}^{(2)}\cdots p_{i_n}^{(n)}\right)$$

$$=\sum_{\substack{i_1=ura,omote\\ i_2=ura,omote\\ i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}x_{i_1}^{(1)}\times\left(p_{i_1}^{(1)}\cdot p_{i_2}^{(2)}\cdots p_{i_n}^{(n)}\right)$$

$$+\sum_{\substack{i_1=ura,omote\\ i_2=ura,omote\\ i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(x_{i_2}^{(2)}+\cdots+x_{i_n}^{(n)}\right)\times\left(p_{i_1}^{(1)}\cdot p_{i_2}^{(2)}\cdots p_{i_n}^{(n)}\right)$$

$$\text{第 1 項}=\sum_{\substack{i_2=ura,omote\\ i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}0\times\left(p\cdot p_{i_2}^{(2)}\cdots p_{i_n}^{(n)}\right)+\sum_{\substack{i_2=ura,omote\\ i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}1\times\left(q\cdot p_{i_2}^{(2)}\cdots p_{i_n}^{(n)}\right)$$

$$=0\times p\times\sum_{\substack{i_2=ura,omote\\ i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(p_{i_2}^{(2)}\cdots p_{i_n}^{(n)}\right)+1\times q\times\sum_{\substack{i_2=ura,omote\\ i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(p_{i_2}^{(2)}\cdots p_{i_n}^{(n)}\right)$$

$$=0\times p\times\sum_{\substack{i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(p\cdot p_{i_2}^{(3)}\cdots p_{i_n}^{(n)}\right)+0\times p\times\sum_{\substack{i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(q\cdot p_{i_2}^{(3)}\cdots p_{i_n}^{(n)}\right)$$

$$+1\times q\times\sum_{\substack{i_2=ura,omote\\ i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(p_{i_2}^{(2)}\cdots p_{i_n}^{(n)}\right)$$

$$=0\times p\times\left(p+q\right)\times\sum_{\substack{i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(p_{i_2}^{(3)}\cdots p_{i_n}^{(n)}\right)+1\times q\times\sum_{\substack{i_2=ura,omote\\ i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(p_{i_2}^{(2)}\cdots p_{i_n}^{(n)}\right)$$

$$=0\times p\times 1\times\sum_{\substack{i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(p_{i_2}^{(3)}\cdots p_{i_n}^{(n)}\right)+1\times q\times\sum_{\substack{i_2=ura,omote\\ i_3=ura,omote\\ \vdots\\ i_n=ura,omote}}\left(p_{i_2}^{(2)}\cdots p_{i_n}^{(n)}\right)$$

$$=0\times p\times 1^{n-1}+1\times q\times 1^{n}=0\times p+1\times q=1\text{回目の試行の期待値}$$

よって

“和である確率変数の期待値は１回の試行の期待値×試行回数”

となる。重大な仮定は各試行が独立だとして，確率の積を用いている点である。

期待値の性質（確率変数の和について：一般的な場合）

以上の議論は一般に成立し，独立な確率変数の和については

$$E\left[X^{(1)}+X^{(2)}+\cdots+E^{(n)}\right]=E\left[X^{(1)}\right]+E\left[X^{(2)}\right]+\cdots E\left[X^{(n)}\right]$$

となる。また，独立な確率変数の積については

$$E\left[X\cdot Y\right]=\sum_{i,j}\left(x_i y_j\right)\left(p_i^{(x)}p_j^{(y)}\right)=\sum_i\sum_j x_i y_j p_i^{(x)}p_j^{(y)}=\sum_i x_i p_i^{(x)}\sum_j y_j p_j^{(y)}$$
$$=E\left[X\right]\cdot E\left[Y\right]$$

である。

分散値の性質（確率変数の和について：一般的な場合）

分散は

$$E\left[\left(\sum_{i=1,n}X_i^{(1)}-E\left[X_i^{(1)}\right]\right)^2\right]=E\left[\sum\left(X_i^{(1)}-E\left[X_i^{(1)}\right]\right)^2\right]+E\left[\sum\left(X_i^{(1)}-E\left[X_i^{(1)}\right]\right)\left(X_j^{(1)}-E\left[X_j^{(1)}\right]\right)\right]$$

である。まず第２項が 0 となること：和の期待値が期待値の和となること，独立な確率変数の積の期待値は期待値の積となること，そして期待値のまわりの期待値は 0 となることから

$$E\left[\sum\left(X_i^{(1)}-E\left[X_i^{(1)}\right]\right)\left(X_j^{(1)}-E\left[X_j^{(1)}\right]\right)\right]=\sum E\left[\left(X_i^{(1)}-E\left[X_i^{(1)}\right]\right)\left(X_j^{(1)}-E\left[X_j^{(1)}\right]\right)\right]$$
$$=\sum E\left[\left(X_i^{(1)}-E\left[X_i^{(1)}\right]\right)\right]E\left[\left(X_j^{(1)}-E\left[X_j^{(1)}\right]\right)\right]=0$$

である。よって

$$E\left[\left(\sum_{i=1,n}X_i^{(1)}-E\left[X_i^{(1)}\right]\right)^2\right]=\sum E\left[\left(X_i^{(1)}-E\left[X_i^{(1)}\right]\right)^2\right]$$

すなわち，確率変数の和に対する分散は個々の確率変数の分散の和である。

２項分布の分散（確率変数の和について）

２項分布の分散については，１回の試行の分散が

$$V\left[X^{(1)}\right]=\sum x_i^2 p_i-E\left[X^{(1)}\right]^2=\left(0^2\times p+1^2\times q\right)-q^2=q\left(1-q\right)=pq$$

よって n 回の試行の合計に対する分散は

$$V[X]=npq$$

となり、１回の試行のときのｎ倍である。

ポアソン分布

ある時間１秒間に平均λ回起きる現象がある。たとえば定常電流が流れるときに伝導電子の個数を測定する場合を想定すればよい。1秒を十分に大きな数nの区間に分割し、小区間内では現象が 1 回以上起きない程度にする。では、この小区間で現象がおきる確率はどのようになるだろうか（めったに起きない確率的現象）。

コインを投げて表がでたら（確率をqとする）この小区間で現象が起きることにすれば、すべての区間であわせてλ回現象が起きる様子を表すものが 2 項分布になることは明らかだろう。2 項分布の期待値と 1 秒内の現象の数λは同じだから

$$np=\lambda$$

である。この数λは平均値であり、1秒内に実際に観測される現象の個数は2項分布に従う。1 秒内にk回観測する確率は

$$P_{\mathrm{B}}=\frac{n!}{(n-k)!k!}p^k q^{n-k}$$

である。

1 秒内の回数の期待値λを一定に保ち、分割の数nを大きくする。このとき $p=\dfrac{\lambda}{n}$ で小さくなることに注意して 2 項分布を書き直すと

$$P_{\mathrm{B}}=\frac{n!}{(n-k)!k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}=\left(\frac{\lambda^k}{k!}\right)\left(1-\frac{\lambda}{n}\right)^n\cdot\left(\frac{n!}{(n-k)!}\right)\left(\frac{1}{n^k}\right)\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$\lim_{n\to\infty}\left(\frac{n!}{(n-k)!}\right)\left(\frac{1}{n^k}\right)\left(1-\frac{\lambda}{n}\right)^{-k}=1$$

$$\because\ \frac{n!}{(n-k)!n^k}=\frac{(n-k+1)\cdot(n-k+2)\cdots n}{n^k}=\left(1-\frac{k-1}{n}\right)\left(1-\frac{k-2}{n}\right)\cdots\left(1-\frac{1}{n}\right)\cdot 1\to 1$$

$$\left(1-\frac{\lambda}{n}\right)^k\to 1$$

$$P_{\mathrm{B}}\to\lim_{n\to\infty}\left(\frac{\lambda^k}{k!}\right)\left(\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right)^{-\lambda}=\left(\frac{\lambda^k}{k!}\right)e^{-\lambda}$$

となる。これがポアソン分布である。

ポアソン分布の性質

ポアソン分布は離散変数$k=0,1,2,\ldots$を持つ。

確率分布だからkについて和をとると１になる。実際

$$\sum_{k=0,\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0,\infty}\frac{\lambda^k}{k!}$$

とすれば、右辺の総和記号が指数関数e^{λ}の定義であり、$e^{-\lambda}e^{\lambda}=1$となる。

期待値はλである。これは定義から明らか。しかし計算で示すなら

$$\sum_{k=0,\infty} k\left(e^{-\lambda}\frac{\lambda^k}{k!}\right) = \sum_{k=1,\infty} e^{-\lambda}\frac{\lambda^k}{(k-1)!} = \lambda\sum_{k=1,\infty} e^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!} = \lambda\sum_{k'=0,\infty} e^{-\lambda}\frac{\lambda^{k'}}{k'!} = \lambda$$

である。

分散を求める：

$$\sum_{k=1,\infty} k\cdot(k-1)\left(e^{-\lambda}\frac{\lambda^k}{k!}\right) = \sum_{k=2,\infty} e^{-\lambda}\frac{\lambda^k}{(k-2)!} = \lambda^2\sum_{k=2,\infty} e^{-\lambda}\frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

$$\sum_{k=1,\infty} k^2\left(e^{-\lambda}\frac{\lambda^k}{k!}\right) - \sum_{k=1,\infty} k\left(e^{-\lambda}\frac{\lambda^k}{k!}\right) = \sum_{k=0,\infty} k^2\left(e^{-\lambda}\frac{\lambda^k}{k!}\right) - \lambda = \lambda^2$$

$$E\left[K^2\right] = \lambda^2 + \lambda$$

$$V[K] = E\left[\left(K-E[K]\right)^2\right] = E\left[K^2\right] - E[K]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

となり、分散は平均と同じ値となる。

まとめると，ポアソン分布は 1 回の試行では非常に小さな確率でおきる現象を、非常に多数回寄せ集めた結果を与え平均値と分散が等しい。

ランダムウォーク

最初 x 軸上原点にある点が，コインを投げて表がでれば正方向に１，裏が出れば動かないとしよう。この運動を１次元のランダムウオークという（普通は１と０ではなく＋１と－１にして左右に動くようにしている)。n回の試行の後，この点がkに見いだされる確率は２項分布で与えられる。言いかえると，１と０の一様ランダム分布（裏と表が等確率）をn個用意し，その一様にランダムな確率変数の和をとると２項分布に従う。ランダムウォークはたとえば衝突を繰り返しながら拡散していく粒子の運動など，自然現象の様々な場面で現れる。また，測定誤差のように，さまざまな要因で小さなランダム現象が積み重なって起きる場合も，測定結果はランダムウォークの結果といえる（これらの場合はいずれも±１の移動で中心位置が不変)。

n 回の試行によるランダムウォークの結果現れる２項分布のグラフは，中心が $\frac{n}{2}$ 原点にある。n に比例して分散が大きくなるとともに，期待値にある確率が小さくなる（山がつぶれる）。注意：ポアソン分布では n が大きくなるとき現象が起きる確率が反比例して小さくなったが，今度の場合には確率は変化しない。

正規分布（ガウス分布）

ランダムウォークの場合を想定して，２項分布で n を無限に大きくしたときの極限を求めよう。２項分布は $p=q=\frac{1}{2}$ として

$$P(X=k)=\frac{n!}{k!(n-k)!}\frac{1}{2^n}$$

である。ここで $k=\frac{n}{2}+h$ （$n-k=\frac{n}{2}-h$ ）なる h を導入するが，これは k を $\frac{n}{2}$ からの差で表したことに相当する。このとき

$$P=\frac{n!}{\left(\frac{n}{2}+h\right)!\left(\frac{n}{2}-h\right)!}\frac{1}{2^n}$$

である。両辺の対数（底を e とする自然対数なので log natural の頭文字 ln）をとり

$$\ln P=\ln n!-\ln\left(\frac{n}{2}+h\right)!-\ln\left(\frac{n}{2}-h\right)!-n\ln 2$$

さらに $n\gg 1$ としてスターリングの公式

$$\ln n!\approx n\ln n-n$$

を用いる（この公式は $\ln n!=\sum_{k=1,n}\ln k\approx\int_1^n \ln x\,dx$ より導かれる）。こうして

$$\ln P\approx(n\ln n-n)-\left(\left(\frac{n}{2}+h\right)\ln\left(\frac{n}{2}+h\right)-\left(\frac{n}{2}+h\right)\right)-\left(\left(\frac{n}{2}-h\right)\ln\left(\frac{n}{2}-h\right)-\left(\frac{n}{2}-h\right)\right)-n\ln 2$$

$$=n\ln n-\left(\frac{n}{2}+h\right)\ln\left(\frac{n}{2}+h\right)-\left(\frac{n}{2}-h\right)\ln\left(\frac{n}{2}-h\right)-n\ln 2$$

となる。第２および第３項を

$$\left(\frac{n}{2}\pm h\right)\ln\left(\frac{n}{2}\pm h\right)=\left(\frac{n}{2}\right)\cdot\left(1\pm\frac{h}{n/2}\right)\ln\left(\frac{n}{2}\right)\cdot\left(1\pm\frac{h}{n/2}\right)=\left(\frac{n}{2}\right)\cdot\left(1\pm\frac{h}{n/2}\right)\left(\ln\left(\frac{n}{2}\right)+\ln\left(1\pm\frac{h}{n/2}\right)\right)$$

と変形する。$1\gg x\Rightarrow\ln(1+x)\approx x-\frac{1}{2}x^2$ と展開できるから，$n\gg h$ のとき $x=\frac{h}{n/2}$ とおけ

ば$(1+x)\ln(1+x)+(1-x)\ln(1-x)\approx x^2$である。この近似を用いて

$$\ln P \approx n\ln n-\frac{n}{2}\left(\left(1+\frac{h}{n/2}\right)+\left(1-\frac{h}{n/2}\right)\right)\ln\frac{n}{2}-\frac{n}{2}\left(\frac{h}{n/2}\right)^2-n\ln 2$$
$$=n\ln n-n\ln\frac{n}{2}-\frac{h^2}{\left(\frac{n}{2}\right)}-n\ln 2=-\frac{h^2}{\left(\frac{n}{2}\right)}$$

を得る。対数をもとにもどすと，期待値$\dfrac{n}{2}$からhだけ異なる位置に見いだされる確率は

$$P\approx e^{-2\frac{h^2}{n}}=e^{-2\frac{h^2}{n}}=e^{-\frac{h^2}{2n\cdot\frac{1}{2}\cdot\frac{1}{2}}}=e^{-\frac{h^2}{2(npq)}}$$

となる。これは、いくつかの近似計算を経た結果であって，とくに対数をとってから級数展開していることを考慮すると，最後の式で係数が 1 とは異なる可能性がある。これは、確率の和（積分）が 1 になるための規格化を行う必要があることを示唆する。

変数をhからxに書き直すと

$$P(X)=A\cdot e^{-\frac{x^2}{2\sigma^2}}$$

の形なので，規格化因子Aは

$$\left(\int_{-\infty}^{\infty}e^{-\alpha x^2}dx\right)\left(\int_{-\infty}^{\infty}e^{-\alpha y^2}dy\right)=\iint e^{-\alpha\left(x^2+y^2\right)}dxdy=\iint e^{-\alpha r^2}rdrd\theta=2\pi\int_0^{\infty}e^{-\alpha r^2}rdr=\frac{2\pi}{2\alpha}\int_0^{\infty}e^{-z}dz=\frac{\pi}{\alpha}$$

より

$$A=\frac{1}{\sqrt{2\pi}\cdot\sigma}$$

となる。この確率分布

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

を正規分布（ガウス分布）という。

この分布の特徴を見ると, まず$x\to -x$により分布の形は変わらないので, $x=0$が中心であり$E[X]=0$である。つぎに$\sigma^2=npq$注目すると、npqがもとの 2 項分布の分散であったことから、σ^2が正規分布の分散となることが推察される（証明略）。