二項分布からガウス分布：中心極限定理, ポアソン分布

Masahiro Yamamoto

Modified on May 22, 2017

イカ京なら必ず持っているといわれる名著「確率・統計入門」小針アキ宏　(岩波書店, 1973) を参考に以下記述した。この本は，Caltech の大栗氏も「名物教授だった森毅に確率や統計を教わった時の教科書で，広中氏の序文が泣かせる」と書いている。https://ooguri.caltech.edu/japanese/mathematics/references

1　二項分布：binominal distribution

図 1 のパチンコ玉が左に行く確率を p，右に行く確率を q とすると，N 回の試行において，左に r 回，右に $N-r$ 回いく確率は，二項分布 $B_{N,p}(r)$ で表される。

$$B_{N,p}(r) \quad = \quad {}_N\mathrm{C}_r p^r q^{N-r} = \frac{N!}{r!(N-r)!} p^r q^{N-r}, \qquad p+q=1 \tag{1}$$

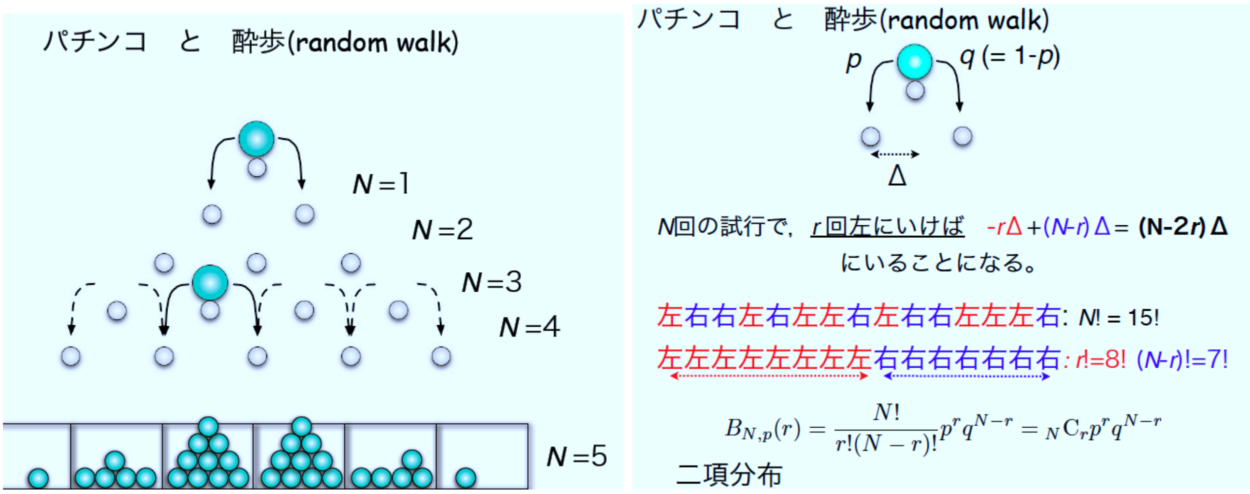

Figure 1: パチンコと酔歩 (random walk)

二項分布 $B_{N,p}(r)$ は，確率の総和が 1 になることは以下のように得られる。

$$\sum_{r=0}^{N} B_{N,p}(r) \quad = \quad \sum_{r=0}^{N} {}_N\mathrm{C}_r p^r q^{N-r} = (p+q)^N = 1, \qquad p+q=1 \tag{2}$$

r(左に行く回数) の平均値 $\langle r\rangle$ と分散 σ_r^2 は

$$\begin{aligned}
\langle r\rangle = \sum_{r=0}^{N} rB_{N,p}(r) &= \sum_{r=0}^{N} {}_N\mathrm{C}_r r p^r q^{N-r} && (3)\\
p\frac{\mathrm{d}}{\mathrm{d}p}(p+q)^N &= pN(p+q)^{N-1} = pN = \sum_{r=0}^{N} {}_N\mathrm{C}_r r p^r q^{N-r} = \langle r\rangle && (4)\\
\sigma_r^2 &= \langle r-\langle r\rangle\rangle^2 = \langle r^2\rangle - 2\langle r\rangle^2 + \langle r\rangle^2 = \langle r^2\rangle - \langle r\rangle^2 && (5)\\
p\frac{\mathrm{d}}{\mathrm{d}p}\left[pN(p+q)^{N-1}\right] &= pN(p+q)^{N-1} + p^2N(N-1)(p+q)^{N-2} = pN + p^2N(N-1) = \sum_{r=0}^{N} {}_N\mathrm{C}_r r^2 p^r q^{N-r} && \\
\langle r^2\rangle &= \sum_{r=0}^{N} {}_N\mathrm{C}_r r^2 p^r q^{N-r} = pN + p^2N(N-1) && (6)\\
\sigma_r^2 &= pN + p^2N(N-1) - p^2N^2 = pN - p^2N = Np(1-p) = Npq && (7)\\
\langle r\rangle &= pN, \quad \sigma_r^2 = Npq && (8)
\end{aligned}$$

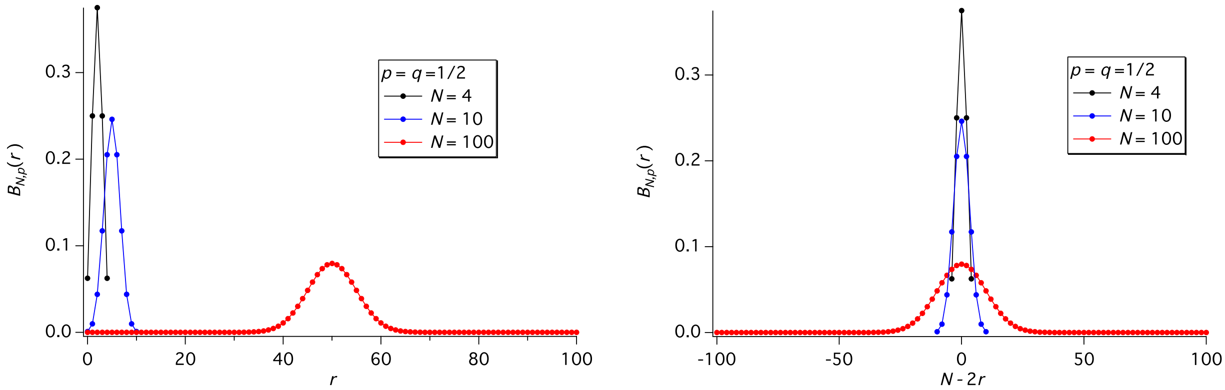

Figure 2: パチンコと酔歩 (random walk) $p = q = 1/2$ の場合

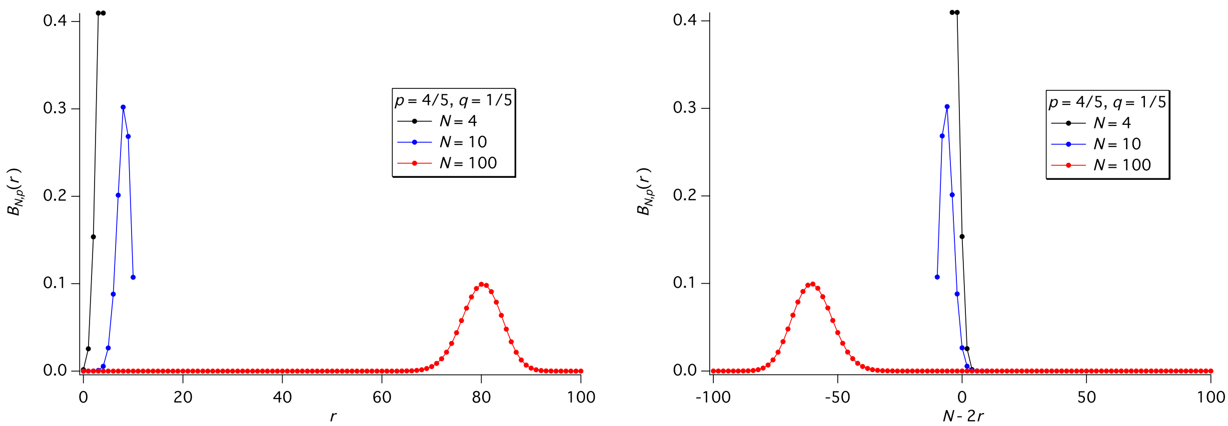

Figure 3: パチンコと酔歩 (random walk) $p = 4/5, q = 1/5$ の場合

2　変数変換

いま r(左に行く回数) の二項分布 $B(r)$ から，ある変数 $t = t(r)$ の分布関数 $P(t)$ に変換することを考える。Fig.3 に示すように, 短冊状のところの確率を考えると

$$B(r)\mathrm{d}r \quad = P(t)\mathrm{d}t \tag{9}$$

となる。従って，$P(t)$ を求めるには，$\mathrm{d}r/\mathrm{d}t$ を求めればよい。

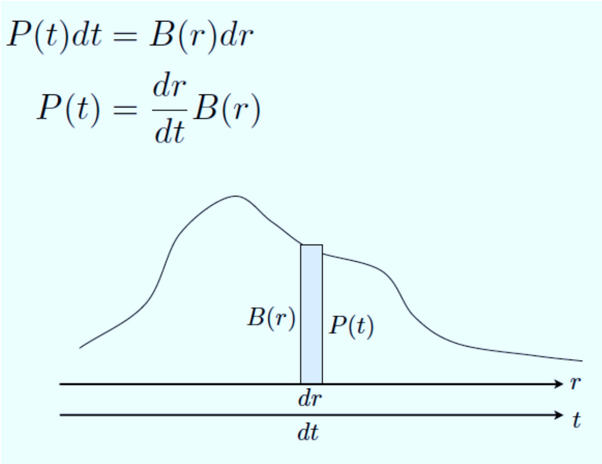

Figure 4: 確率分布関数の変数変換

3　大数の法則: Law of large numbers: $t = r/N$

二項分布においては，N 回の試行のうち r 回左にいくという確率を求めたわけであるが，その確率 r/N の平均値および分散 $\sigma^2_{r/N}$ はどう分布するであろうか。$t = r/N$ として $P_N(t)$ を求め，平均と分散を求めることもできるが，以下のように容易に求めることも可能である。

$$\langle \frac{r}{N} \rangle = \frac{1}{N}\langle r \rangle = \frac{Np}{N} = p \tag{10}$$

$$\sigma^2_{r/N} = \left\langle \left(\frac{r}{N} - \langle \frac{r}{N} \rangle \right)^2 \right\rangle = \frac{1}{N^2}\left(\langle r^2 \rangle - \langle r \rangle^2\right) = \frac{Npq}{N^2} = \frac{pq}{N} \tag{11}$$

ここで，$\langle (r - \langle r \rangle)^2 \rangle = \langle r^2 - 2r\langle r \rangle + \langle r \rangle^2 \rangle = \langle r^2 \rangle - 2\langle r \rangle^2 + \langle r \rangle^2 = \langle r^2 \rangle - \langle r \rangle^2$ を使った。Fig.4 に，$p = 1/2$ の時の，N の依存性を示す。分布の幅 $\sigma_{r/N}$ は，

$$\sigma_{r/N} = \sqrt{\frac{pq}{N}} = \sqrt{\frac{p(1-p)}{N}} \tag{12}$$

となり，N が小さい時はかなり広がるが，N が無限大になるときは，幅が無限小の δ 関数になる。

変数変換により，分布関数 $P_N(t)$ は

$$P_N(t) = \frac{\mathrm{d}r}{\mathrm{d}t} B_{N,p}(r) = NB_{N,p}(Nt) \tag{13}$$

一次元のブラウン運動では，あるサイズの円板（面は一次元の運動方向に垂直）に水分子が左から衝突するのかあるいは右から衝突するのかという衝突分子数の揺らぎが運動の本質である。ブラウン運動で，微粒子がゆらぐのが観察できるのは微粒子の直径が 1 μm 以下の時であるがそれは微粒子が運動する時間レベル (速度の減衰）での溶媒の衝突数 ($N = 10^{12}$) のゆらぎが $1/\sqrt{N} = 10^{-6}$ と ppm オーダーになるからである。ppm オーダーのゆらぎになれば，ブラウン運動することを示すことができる。これより粒子が大きいと揺らぎが $1/\sqrt{N}$ で小さくなり観測できない。また，光学顕微鏡では 1 μm が測定の限界であり，これより小さい粒子は観測出来ない。10 ns の時間刻みで，場所と速度を測定した実験が 2014 年に報告された。(S. Kheifets et al. Science, 343, 149 (2014))

ブラウン運動は，単なる微粒子の酔歩の問題ではない。アインシュタインの理論的予測をペランが詳細に実験で証明して，皆が原子・分子の存在を最終的に認めたのである。ペランが書いた Les Atomes「原子」では，原子論の勝利を宣言しており，本は 1913 年の出版されたので，原子の存在の論争は 100 年ほどの前まで続いていたのである。

4　ガウス分布：中心極限定理（統計学において中心となる極限定理という意味）

二項分布 $B_{N,p}(r)$ において，N が無限大の時に，$t = (r - \mu)/\sigma$ は，ガウス分布になることを示す。 ここで，$\mu(= Np)$ は平均値，$\sigma^2(= Npq)$ は分散である。

$$t = \frac{r - Np}{\sqrt{Npq}}, \quad r = \sqrt{Npq}t + Np, \quad \frac{dr}{dt} = \sqrt{Npq} \tag{14}$$

$$F(t) = B(r)\frac{dr}{dt} = \sqrt{Npq}B(r) = \sqrt{Npq}\frac{N!}{r!(N-r)!}p^r q^{N-r} \tag{15}$$

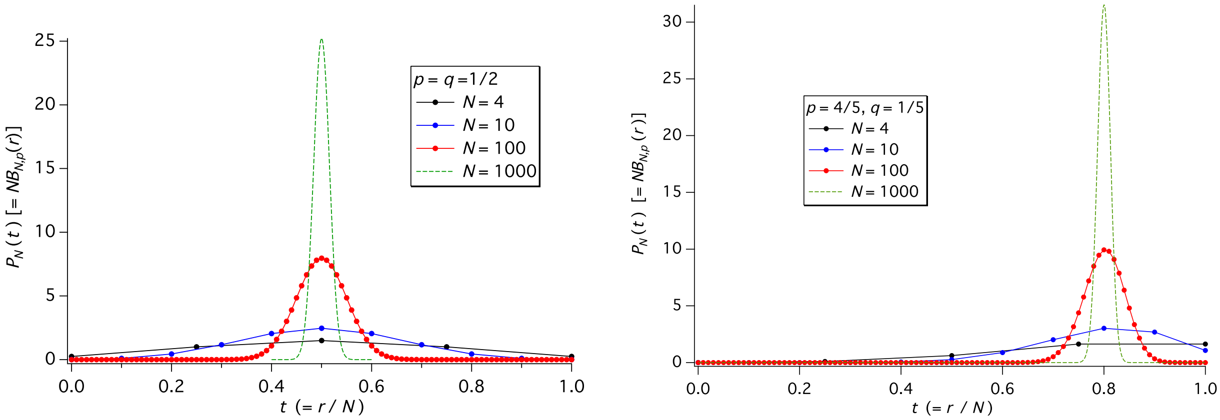

Figure 5: 大数の法則: 左 $p = q = 1/2$, 右図 $p = 4/5, q = 1/5$ の場合　$N = 1000$ の時，２項定数は Lanzos の方法を用いて計算した (Appendix 参照)。

ここで，スターリングの公式（きびしいバージョン：証明は Appendix2）を使って N が大きいときの近似をおこなう。

$$N! \;=\; \sqrt{2\pi N}\left(\frac{N}{e}\right)^N \tag{16}$$

従って，

$$\begin{aligned}F(t) &= \sqrt{Npq}\frac{N!}{r!(N-r)!}p^r q^{N-r} = \sqrt{Npq}\frac{\sqrt{2\pi N}N^N e^{-N}}{\sqrt{2\pi r}r^r e^{-r}\sqrt{2\pi(N-r)}(N-r)^{N-r}e^{-(N-r)}}p^r q^{N-r} \\ &= \frac{1}{\sqrt{2\pi}}\frac{N^{N+1}}{r^{r+1/2}(N-r)^{N-r+1/2}}p^{r+1/2}q^{N-r+1/2} = \frac{1}{\sqrt{2\pi}}\left(\frac{Np}{r}\right)^{r+1/2}\left(\frac{Nq}{N-r}\right)^{N-r+1/2}\end{aligned}$$

Eq.(14) より

$$\frac{r}{Np} = 1 + \sqrt{\frac{q}{Np}}t \tag{17}$$

$$N - r = N - Np - \sqrt{Npq}t = Nq - \sqrt{Npq}t, \quad \frac{N-r}{Nq} = 1 - \sqrt{\frac{p}{Nq}}t \tag{18}$$

$$F(t) = \frac{1}{\sqrt{2\pi}}\left(1+\sqrt{\frac{q}{Np}}t\right)^{-r-1/2}\left(1-\sqrt{\frac{p}{Nq}}t\right)^{-N+r-1/2} \tag{19}$$

$$-\ln[\sqrt{2\pi}F(t)] = \left(r+\frac{1}{2}\right)\ln\left(1+\sqrt{\frac{q}{Np}}t\right) + \left(N-r+\frac{1}{2}\right)\ln\left(1-\sqrt{\frac{p}{Nq}}t\right) \tag{20}$$

$$r = Np + \sqrt{Npq}t, \quad N - r = Nq - \sqrt{Npq}t \tag{21}$$

$x << 1$ で，

$$\begin{aligned}\ln(1\pm x) &\simeq \ln(1) \pm \left(\frac{1}{1\pm x}\right)_{x=0} x - \frac{1}{2!}\left(\frac{1}{(1\pm x)^2}\right)_{x=0} x^2 + O(x^3) \\ &= \pm x - \frac{1}{2}x^2 + O(x^3)\end{aligned} \tag{22}$$

となる。$\sqrt{q/Npt} << 1$, $\sqrt{p/Nqt} << 1$ なので，

$$\begin{aligned}
-\ln[\sqrt{2\pi}F(t)] &= \left(r+\frac{1}{2}\right)\left(\sqrt{\frac{q}{Np}}t-\frac{1}{2}\frac{q}{Np}t^2+O(N^{-3/2})\right) \\
&\quad +\left(N-r+\frac{1}{2}\right)\left(-\sqrt{\frac{p}{Nq}}t-\frac{1}{2}\frac{p}{Nq}t^2+O(N^{-3/2})\right) \\
&= (Np+\sqrt{Npq}t+\frac{1}{2})\left(\sqrt{\frac{q}{Np}}t-\frac{1}{2}\frac{q}{Np}t^2\right)+(Nq-\sqrt{Npq}t+\frac{1}{2})\left(-\sqrt{\frac{p}{Nq}}t-\frac{1}{2}\frac{p}{Nq}t^2\right)+ \\
&\quad +O(N^{-1/2}) \qquad (23) \\
&= \sqrt{Npq}t-\frac{1}{2}qt^2+qt^2-\frac{1}{2}\sqrt{\frac{q^3}{Np}}t^3+\frac{1}{2}\sqrt{\frac{q}{Np}}t-\frac{1}{4}\frac{q}{Np}t^2 \qquad (24) \\
&\quad -\sqrt{Npq}t-\frac{1}{2}pt^2+pt^2+\frac{1}{2}\sqrt{p^3}Nqt^3-\frac{1}{2}\sqrt{\frac{p}{Nq}}t-\frac{1}{4}\frac{p}{Nq}t^2+O(N^{-1/2}) \qquad (25) \\
&= \frac{1}{2}t^2+O(N^{-1/2}) \qquad (26) \\
F(t) &= \frac{1}{\sqrt{2\pi}}\exp(-t^2/2)=G_{0,1}(t) \qquad (27) \\
\text{where} \quad G_{\mu,\sigma}(x) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] \qquad (28)
\end{aligned}$$

また，もとの $B(r)$ は，

$$B_{N,p}(r)=\frac{F(t)}{\sigma}=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[\frac{-(r-\mu)^2}{2\sigma^2}\right], \quad \mu=Np, \quad \sigma=\sqrt{Npq} \qquad (29)$$

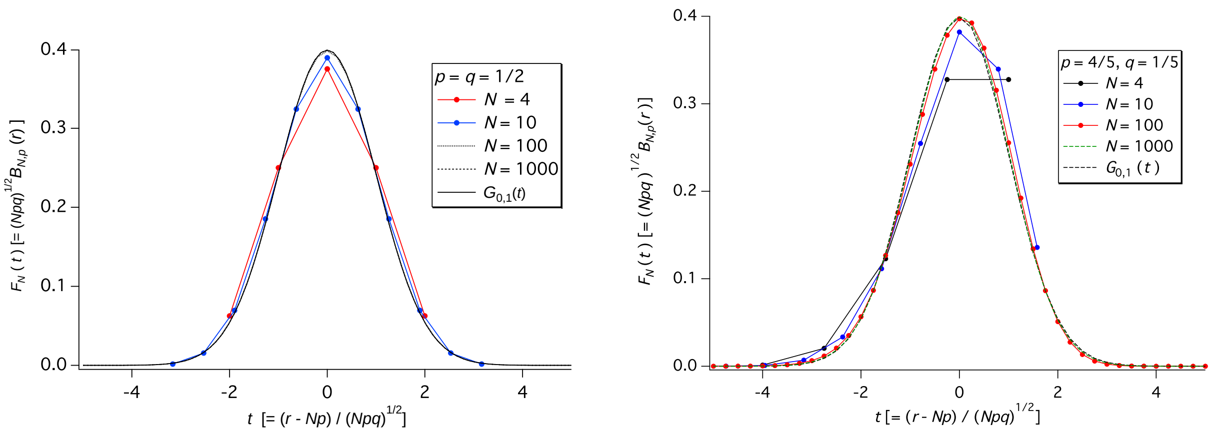

Figure 6: 中心極限定理： 左 $p = q = 1/2$, 右図 $p = 4/5, q = 1/5$ の場合　$N = 1000$ の時，２項定数は Lanzos の方法を用いて計算した (Appendix 参照)。

5　Poisson 分布

$Np = \lambda$ 一定で，$N \to \infty$ の時を考える。$p = \lambda/N, q = 1 - p = 1 - \lambda/N$.

$$\begin{aligned}
B_{N,p} &= \frac{N!}{r!(N-r)!}\left(\frac{\lambda}{N}\right)^r\left(1-\frac{\lambda}{N}\right)^{N-r} \qquad (30) \\
&= \frac{\lambda^r}{r!}\frac{N(N-1)...(N-r+1)}{N^r}\left(1-\frac{\lambda}{N}\right)^{-r}\left(1-\frac{\lambda}{N}\right)^N \qquad (31) \\
&= \frac{\lambda^r}{r!}\underbrace{1(1-\frac{1}{N})(1-\frac{2}{N})...(1-\frac{r-1}{N})}_{\lim_{N\to\infty}\simeq 1}\underbrace{\left(1-\frac{\lambda}{N}\right)^{-r}}_{\lim_{N\to\infty}\simeq 1}\left[\underbrace{\left(1-\frac{\lambda}{N}\right)^{-N/\lambda}}_{\lim_{N\to\infty}\simeq e}\right]^{-\lambda} \qquad (32) \\
&= \frac{\lambda^r}{r!}e^{-\lambda} \equiv P_\lambda(r) \qquad (33)
\end{aligned}$$

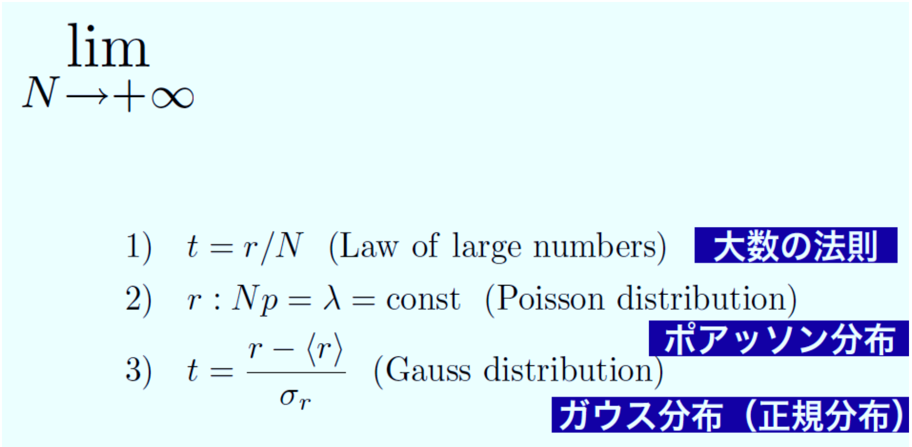

Figure 7: パチンコと酔歩 (random walk)

ここで，r はゼロおよび自然数である。確率分布かどうか調べるためにその総和をとると，

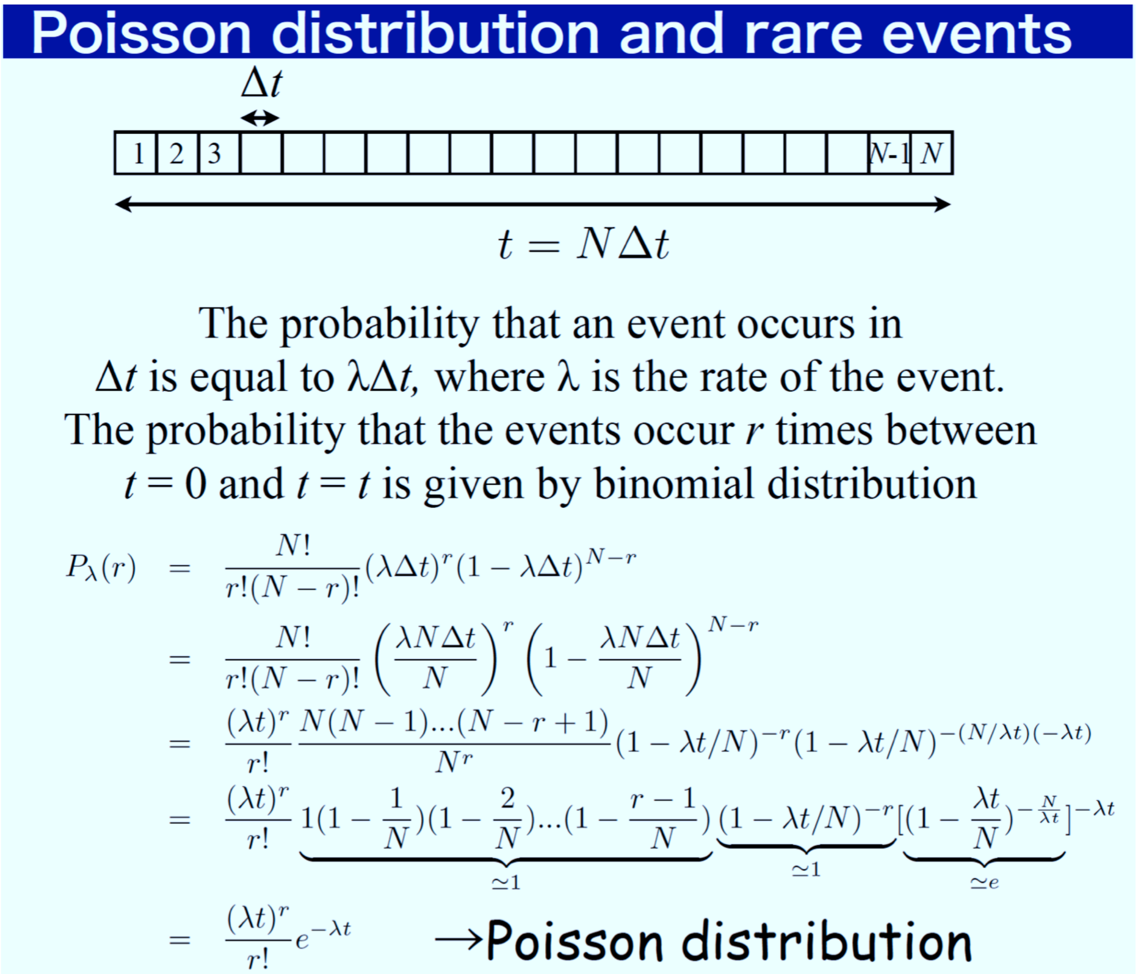

Figure 8: Poisson 分布の導入

$$\sum_{r=0}^{\infty} P_\lambda(r) = \sum_{r=0}^{\infty}\frac{\lambda^r}{r!}e^{-\lambda} = e^{-\lambda}\underbrace{\sum_{r=0}^{\infty}\frac{\lambda^r}{r!}}_{=e^\lambda} = 1 \tag{34}$$

となる。平均値は,

$$\langle r\rangle = \sum_{r=0}^{\infty} rP_\lambda(r) = \sum_{r=0}^{\infty} r\frac{\lambda^r}{r!}e^{-\lambda} \tag{35}$$

$$e^\lambda = \sum_{r=0}^{\infty}\frac{\lambda^r}{r!},\quad \frac{d}{d\lambda}e^\lambda = e^\lambda = \sum_{r=0}^{\infty}\frac{r\lambda^{r-1}}{r!},\quad \lambda e^\lambda = \sum_{r=0}^{\infty}\frac{r\lambda^r}{r!},\quad \lambda = \sum_{r=0}^{\infty}\frac{r\lambda^r}{r!}e^{-\lambda} \tag{36}$$

$$\langle r\rangle = \lambda \tag{37}$$

となる。分散は,

$$\langle r^2\rangle = \sum_{r=0}^{\infty} r^2P_\lambda(r) = \sum_{r=0}^{\infty} r^2\frac{\lambda^r}{r!}e^{-\lambda} \tag{38}$$

$$\frac{d}{d\lambda}(\lambda e^\lambda) = e^\lambda + \lambda e^\lambda = \sum_{r=0}^{\infty}\frac{r^2\lambda^{r-1}}{r!},\quad (\lambda+\lambda^2)e^\lambda = \sum_{r=0}^{\infty}\frac{r^2\lambda^r}{r!},\quad \lambda+\lambda^2 = \sum_{r=0}^{\infty}\frac{r^2\lambda^r}{r!}e^{-\lambda} \tag{39}$$

$$\sigma = \langle(r-\langle r\rangle)^2\rangle = \langle r^2\rangle - \langle r\rangle^2 = \lambda+\lambda^2-\lambda^2 = \lambda \tag{40}$$

となり，平均値も分散も λ となる。小針の本では，ブドウが貴重であった時代に，ブドウパンのなかに３個未満しかブドウがないパンが 5%を超えると苦情が来てその確率をポアソン分布で計算するという例が出ている。本がでた 1973 年当時は p が小さかったのである。

6　平均の平均，平均の分散

実験を繰り返すことによって,「平均の平均」は母集団の平均に近づき,「平均の分散」は実験の回数の平方根に反比例して減少する。すなわち実験回数を多く行えば実験結果の信頼性は増加するという自明の事実をここでは、独立な多変数分布を使って定式化する。

「平均の平均」に対する中心極限定理：**任意の分布をもつ母集団**が，平均 μ, 分散 σ^2 をもつとする。**標本サイズ N が十分大きいとき**，標本平均 $\bar{x}$ $(=\sum_i^N x_i/N)$ は，近似的に正規分布 $G_{\mu,\sigma/\sqrt{N}}(\bar{x})$ に従う。

いま，(x,y) という 2 変数の確率分布を $p(x,y)$ を考える[1]。x と y の任意の関数 $h(x,y)$ の (例えば，$h(x,y)=x+y$ である。) この分布における期待値 $\langle h\rangle$ [2] は，は,

$$\langle h\rangle = \int\int h(x,y)p(x,y)dxdy \tag{43}$$

で与えられる。 x と y が独立であれば[3]，図 9 に示したような分布関数 $p(x,y)$ は,

$$p(x,y) = f(x)g(y) \tag{44}$$

という２つの分布関数の積で表される。[4]　分布が独立でない場合を，図 10 に示すが，ここでは考慮しない。

---

[1] x と y は例えば身長と体重でもよいし，体重と試験の成績でもよい。ある母集団で身長 175 cm で体重 65 kg の人がいる確率が $p(175,65)$ である。

[2] 確率分布 $p(x)$ が与えられている時，任意の関数 $h(x)$ の期待値は,

$$\langle h\rangle = \int h(x)p(x)dx \tag{41}$$

で与えられる。x がとびとびの値 x_i をもつときは,

$$\langle h\rangle = \sum_i h(x_i)p(x_i) \tag{42}$$

である。後者の例としてある宝くじを考えてみよう。x_i は 1 等，前後賞，組違い賞,,，8 等，はずれを意味し，それぞれどの確率で当選るのかが決められている。例えば，$p(1\text{等}) = 10^{-7}$ で,...,$p(\text{はずれ}) = 0.8888693$ である。$h(x_i)$ を当選金（1 等 2 億円,..., 8 等 300 円）にすると，宝くじ 1 枚あたりどの程度当選金があたるのかすなわち宝くじの当選金の期待値が求めることができる。ちなみに，1 枚 300 円の宝くじでは $\langle h\rangle = 141.99$ 円である。。

同様にして，2 確率変数分布の期待値も式 (43) から求めることができる。x を身長，y を体重にして，$h(x,y)=yx^{-2}$ とおけば，式 (43) から，ボディ・マス・インデックス (BMI) の期待値が得られる。

[3] x と y が体重と成績の場合は独立になると考えられるが，身長と体重の場合は明らかに独立でない。

[4] $f(x), g(y)$ は次のようにあらわされる。

$$f(x) = \int p(x,y)dy \;\; [= f(x)\underbrace{\int g(y)dy}_{=1}],\qquad g(y) = \int p(x,y)dx \;\; [= g(y)\underbrace{\int f(x)dx}_{=1}]$$

である。

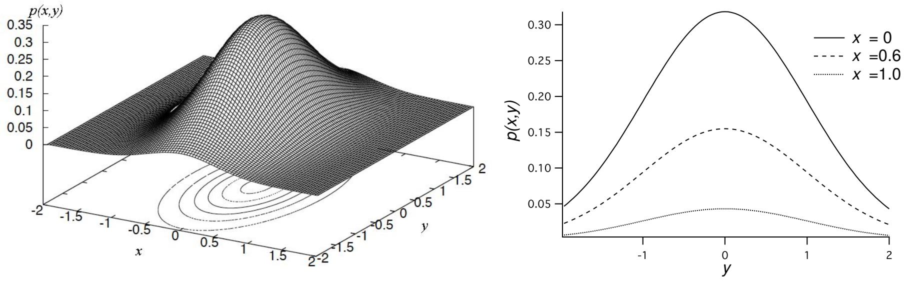

Figure 9: x と y が独立であるときの分布関数 $p(x,y)$。x が一定の値をもつ時の $p(x,y)$ のピークの断面は，すべて相似形 [定数倍を除いて同じ関数] になる。

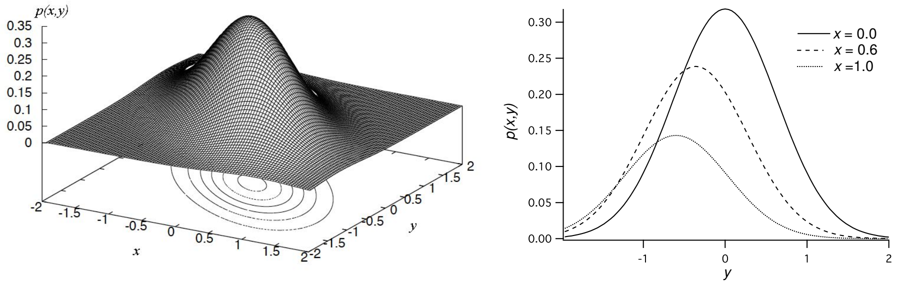

Figure 10: x と y が独立でない時の分布関数 $p(x,y)$。x と y が独立であるときと異なり，x が一定の値をもつ時の $p(x,y)$ のピークの断面は，同じ関数にならない。右図では y の最大値の位置が x に依存している。

いま,$h(x,y)=x+y$ で，x と y が独立であるとしよう。

$$\begin{aligned}\langle x+y\rangle &= \int\int(x+y)f(x)g(y)dxdy=\int xf(x)dx\underbrace{\int g(y)dy}_{=1}+\int yg(y)dy\underbrace{\int f(x)dx}_{=1}\\ &= \langle x\rangle+\langle y\rangle \end{aligned} \tag{45}$$

同様にして, 以下の式が得られる。

$$\frac{1}{N}\langle x_1+x_2+x_3+...+x_N\rangle=\frac{\langle x_1\rangle+\langle x_2\rangle+\langle x_3\rangle+...+\langle x_N\rangle}{N} \tag{46}$$

$\langle x_i\rangle=\mu$(母集団の平均) であるとすれば，上式の左辺を「平均の平均」$\langle\bar{x}\rangle$ と定義できるので、

$$\langle\bar{x}\rangle\equiv\langle\frac{x_1+x_2+x_3+...+x_N}{N}\rangle=\frac{N\mu}{N}=\mu \tag{47}$$

となる[5]。

次に，$h(x,y)=[(x+y)-\langle x+y\rangle]^2$ を考える。x と y は独立であるとしよう。

$$\begin{aligned}\langle[(x+y)-\langle x+y\rangle]^2\rangle &= \int\int[(x+y)-\langle x+y\rangle]^2f(x)g(y)dxdy\\ =\int x^2f(x)dx+\int y^2g(y)dy&+2\int xf(x)dx\int yg(y)dy-2\langle x+y\rangle^2+\langle x+y\rangle^2\\ =\langle x^2\rangle+\langle y^2\rangle+2\langle x\rangle\langle y\rangle&-\langle x\rangle^2-\langle y\rangle^2-2\langle x\rangle\langle y\rangle=\langle x^2\rangle-\langle x\rangle^2+\langle y^2\rangle-\langle y\rangle^2\end{aligned}$$

従って，以下の式が得られる。

$$\begin{aligned}\sigma_{x+y}^2 &= \sigma_x^2+\sigma_y^2 \qquad (48)\\ \sigma_x^2 &= \int(x-\langle x\rangle)^2f(x)dx=\langle x^2\rangle-\langle x\rangle^2,\qquad \sigma_y^2=\int(y-\langle y\rangle)^2g(y)dy=\langle y^2\rangle-\langle y\rangle^2\end{aligned}$$

同様にして

$$\langle[(\frac{x_1+x_2+x_3+...+x_N}{N})-\langle\frac{x_1+x_2+x_3+...+x_N}{N}\rangle]^2\rangle=\frac{\langle x_1^2\rangle-\langle x_1\rangle^2+....+\langle x_N^2\rangle-\langle x_N\rangle^2}{N^2} \tag{49}$$

[5] $\langle x\rangle$ と $\bar{x}$ の意味は違うことに注意。

$\langle x_i^2\rangle - \langle x_i\rangle^2 = \sigma^2$(母集団の分散）であるとすれば，「平均の分散」として以下の極めて重要な式が得られる。

$$\sigma_{\bar{x}}^2 \;=\; \frac{N\sigma^2}{N^2} = \frac{\sigma^2}{N} \tag{50}$$

この証明から以下のことが示せた。「母集団分布が (μ, σ^2) の分布であるとき，大きさ N の標本の平均 $\bar{x} = \dfrac{1}{N}\displaystyle\sum_{i-1}^{N} x_i$ は $(\mu, \sigma^2/N)$ の分布に従う。」

7　標本の分散の平均

実際の実験においては，母集団の分散 σ^2 を事前に知ることは少ない。従って，標本から推定することになる。

まず，標本分散 $s^2 = \dfrac{1}{N}\displaystyle\sum_{i}^{N}(x_i - \bar{x})^2$ の平均をもとめよう。この場合も，N 個の標本から s^2 を計算することを 1 回の操作と考えて，その操作を多数回繰り返せば s^2 に関する分布ができ，その平均・分散が定義できる。s^2 を以下のように変形して考える。

$$\begin{aligned} s^2 &= \frac{1}{N}\sum_{i}^{N}(x_i - \bar{x})^2 = \frac{1}{N}\sum_{i}^{N}(x_i - \mu + \mu - \bar{x})^2 \\ &= \frac{1}{N}\sum_{i}^{N}(x_i - \mu)^2 - (\mu - \bar{x})^2 \end{aligned} \tag{51}$$

ここで，$\bar{x}, s^2$ は標本平均，標本分散である。また，母集団は，平均 μ, 分散 σ^2(未知の量）をもつとしよう。平均 $\langle s^2\rangle$ をとると，

$$\langle s^2\rangle = \langle\frac{1}{N}\sum_{i}^{N}(x_i - \mu)^2 - (\mu - \bar{x})^2\rangle = \frac{N\sigma^2}{N} - \frac{\sigma^2}{N} = \frac{N-1}{N}\sigma^2 \tag{52}$$

ここで，平均 $\bar{x}$ の分散 $= \sigma^2/N$ を第二項に対して使った。標本平均の平均は母集団の平均 μ と一致したが，標本分散の平均は母集団の分散 σ^2 とは一致しない。

ここで，標本分散 s^2 を変形して

$$u^2 = \frac{N}{N-1}s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 \tag{53}$$

を新たに定義する。u^2 を**不偏分散** (unbiased variance) という[6]。また，**標準偏差** (**SD: Standard Deviation**) は u^2 の平方根 u で定義される。標本分散 s^2 と不偏分散 u^2 に違いに注意すること。

8　ガンマ関数

ガンマ関数 $\Gamma(\alpha)$ を以下のように定義する。

$$\Gamma(\alpha) \;=\; \int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x \tag{54}$$

$\Gamma(\alpha+1)$ を求める。

$$\Gamma(\alpha+1) \;=\; \int_0^\infty x^{\alpha}e^{-x}\mathrm{d}x \tag{55}$$

[6] N 個の標本から，$\sum_{i=1}^{N}(x_i - \bar{x})^2$ 求める際，$\bar{x} = (1/N)\sum_{i=1}^{N} x_i$ の条件が付くため，$(x_i - \bar{x})^2$ は N 個の独立な項の和ではなく，独立な項は $N-1$ 個である。この数を**自由度**という。
u^2 の定義で分母が $N-1$ になっているのはこのためである。標本分散 s^2 より u^2 は大きい値にをもつので，u^2 を使うのが "正直" ともいえる。$u/\bar{x}$ は変動係数 (coefficient of variance: CV) ともよばれ，ばらつきを無次元化して比較するときに便利である。

$(fg)' = f'g + fg', \int f'g = fg - \int fg', f' = e^{-x}, f = -e^{-x}$

$$\Gamma(\alpha+1) = \int_0^\infty x^\alpha e^{-x} \mathrm{d}x = [-e^{-x}x^\alpha]_0^\infty + \alpha \int_0^\infty e^{-x} x^{\alpha-1} \tag{56}$$

$$= \alpha\Gamma(\alpha) \tag{57}$$

α が自然数の時を考えよう。

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = n(n-1)(n-2)....2 \times 1 \times \Gamma(1) \tag{58}$$

$$= n! \int_0^\infty x^0 e^{-x} \mathrm{d}x = n![-e^{-x}]_0^\infty = n! \tag{59}$$

$\Gamma(\frac{1}{2})$ を計算しよう。ガンマ関数 $\Gamma(\alpha)$ を以下のように定義する。

$$\Gamma(\frac{1}{2}) = \int_0^\infty x^{-1/2} e^{-x} \mathrm{d}x \tag{60}$$

$$x^{1/2} = X, x = X^2, \mathrm{d}x = 2X\mathrm{d}X \tag{61}$$

$$\Gamma(\frac{1}{2}) = \int_0^\infty X^{-1} e^{-X^2} 2X \mathrm{d}X = 2\int_0^\infty e^{-X^2} \mathrm{d}X = 2\frac{1}{2}\sqrt{\pi} = \sqrt{\pi} \tag{62}$$

$\Gamma(\frac{n}{2})$ を計算しよう。n が偶数の時 $(n = 2m)$, n が奇数の時 $(n = 2m+1)$ にわけて考えると，

$$\Gamma(\frac{n}{2}) = \begin{cases} \Gamma(m) = (m-1)!, & n = 2m \\ \Gamma(m+\frac{1}{2}) = (m-1)!\Gamma(\frac{1}{2}) = (m-1)!\sqrt{\pi}, & n = 2m+1 \end{cases} \tag{63}$$

9 ベータ関数

ベーター関数の定義は，$\alpha > 0, \beta > 0$ に対して，

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} \mathrm{d}x \tag{64}$$

である。$0 < \alpha < 1$ の時 0 の近くで, $0 < \beta < 1$ のとき 1 の近くでは発散するが，それ以外では積分は収束する。

$$B(\alpha, \beta) = B(\beta, \alpha) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{65}$$

が成立することを証明しよう。$(1-x) = y, \mathrm{d}x = \mathrm{d}y$ とおくと，

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x = -\int_1^0 (1-y)^{\alpha-1} y^{\beta-1} \mathrm{d}y = \int_0^1 y^{\beta-1}(1-y)^{\alpha-1}\mathrm{d}y \tag{66}$$

$$= B(\beta, \alpha) \tag{67}$$

次に，$x = u^2, y = v^2, \mathrm{d}x = 2u\mathrm{d}u, \mathrm{d}y = 2v\mathrm{d}v$ とおくと，

$$\Gamma(\alpha)\Gamma(\beta) = \int_0^\infty x^{\alpha-1}e^{-x}\mathrm{d}x \int_0^\infty y^{\beta-1}e^{-y}\mathrm{d}y \tag{68}$$

$$= 4\int_0^\infty u^{2\alpha-1}e^{-u^2}\mathrm{d}u \int_0^\infty v^{2\beta-1}e^{-v^2}\mathrm{d}v = 4\int_0^\infty\int_0^\infty u^{2\alpha-1}v^{2\beta-1}e^{-(u^2+v^2)}\mathrm{d}u\mathrm{d}v \tag{69}$$

$u = r\cos\theta, v = r\sin\theta, \mathrm{d}u\mathrm{d}v = r\mathrm{d}r\mathrm{d}\theta$ と変数変換すると，

$$\Gamma(\alpha)\Gamma(\beta) = 4\int_0^{\pi/2}\int_0^\infty r^{2\alpha-1}\cos^{2\alpha-1}\theta r^{2\beta-1}\sin^{2\beta-1}\theta e^{-r^2} r\mathrm{d}r\mathrm{d}\theta \tag{70}$$

$$= 4\int_0^{\pi/2}\cos^{2\alpha-1}\theta\sin^{2\beta-1}\theta\mathrm{d}\theta \int_0^\infty r^{2(\alpha+\beta)-1}e^{-r^2} r\mathrm{d}r \tag{71}$$

$r^2 = R, \mathrm{d}R = 2r\mathrm{d}r$ とすると，r についての積分は

$$2\int_0^\infty r^{2(\alpha+\beta)-1}e^{-r^2}r\mathrm{d}r = \int_0^\infty R^{\alpha+\beta-1}e^{-R}\mathrm{d}R = \Gamma(\alpha+\beta) \tag{72}$$

$t = \sin^2\theta, \mathrm{d}t = 2\sin\theta\cos\theta\mathrm{d}\theta, \mathrm{d}\theta = (1/2)\sin^{-1}\theta\cos^{-1}\theta\mathrm{d}t, \cos^{2\alpha-2}\theta = (\cos^2\theta)^{\alpha-1} = (1-\sin^2\theta)^{\alpha-1} = (1-t)^{\alpha-1}$ と変数変換すると，

$$2\int_0^{\pi/2}\cos^{2\alpha-1}\theta\sin^{2\beta-1}\theta\mathrm{d}\theta = \int_0^1 t^{\beta-1}(1-t)^{\alpha-1}\mathrm{d}t = B(\beta,\alpha) \tag{73}$$

従って，

$$\Gamma(\alpha)\Gamma(\beta) = \Gamma(\alpha+\beta)B(\beta,\alpha) = \Gamma(\alpha+\beta)B(\alpha,\beta) \tag{74}$$

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{75}$$

10　χ^2-分布

自由度 n の χ^2（カイ二乗）-分布は，

$$T_n(x) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(\frac{n}{2})}x^{(n-2)/2}e^{-x/2}, & x > 0 \\ 0, & x \le 0 \end{cases} \tag{76}$$

ここで，n は正の整数である。

　$G_{\mu,\sigma}$ の分布をする正規母集団から N 個の標本 $x_1, x_2, x_3, ..., x_N$ を任意に選び

$$y = \frac{1}{\sigma^2}\sum_{i=1}^N (x_i - \bar{x})^2 = \frac{Ns^2}{\sigma^2} \tag{77}$$

を作るとき，これは自由度 $N-1$ の χ^2 分布に従う。また，y と $\bar{x}$ は独立である。(証明略,　小針「確率・統計入門」（岩波） p.173-176 参照のこと）

11　t-分布

$-\infty < t < +\infty$ の t に対して，確率密度が

$$f_N(t) = \frac{1}{\sqrt{N}B\left(\frac{1}{2},\frac{N}{2}\right)}\left(1+\frac{t^2}{N}\right)^{-(N+1)/2} \tag{78}$$

で定義される分布を，自由度 N の t-分布（またはスチューデント分布）という。

　正規母集団 $G_{\mu,\sigma}$ から大きさ N の独立な標本を選び，

$$t = \frac{(\bar{x}-\mu)\sqrt{N}}{u} \tag{79}$$

$$\bar{x} = \frac{1}{N}\sum i = 1^N x_i, \quad u^2 = \frac{1}{N-1}\sum_{i=1}^N (x_i-\bar{x})^2 \tag{80}$$

とすれば，t は自由度 $N-1$ の t-分布に従う。証明では，$(\bar{x}-\mu)\sqrt{N}$ は $G_{0,\sigma}$ に従い，$\dfrac{(N-1)}{\sigma^2}u^2 = \dfrac{1}{\sigma^2}\sum_{i=1}^N (x_i-\bar{x})^2$ は自由度 $N-1$ の χ^2-分布に従い，$\bar{x}$ と u は独立であることを用いる。(証明略, 小針「確率・統計入門」（岩波） p.278-280 参照のこと）

12 Appendix

２項分布の計算 FORTRAN プログラムその１（$N < 100$ の場合)

```
c234567  bbinominal_dist.f

      integer  n,r,nr,ii
      real(8) :: enfac,rfac,rnfac,p,q
      real(8) :: prob,en,pi,enfac_sr


      pi=acos(-1.0d0)
      p=0.8d0
      q=1.0d0-p

      n=100
      en=real(n)
      enfac_sr=en*log(en)-en
      enfac=real(n)
      do i=1, n
       ii=n-i
       if (ii == 1) exit
       enfac=enfac*real(ii)
      enddo
      write (6,*) n,enfac,enfac_sr

      do r=0, n
       nr=n-r
       if (r <= 1) then
        rfac=1.0d0
        go to 10
       endif
       rfac=real(r)
       do i=1, r
        ii=r-i
        if (ii == 1) exit
        rfac=rfac*real(ii)
       enddo

 10    if (nr <= 1) then
        rnfac=1.0d0
        go to 20
       endif
       rnfac=real(nr)
       do i=1, nr
        ii=nr-i
        if (ii == 1) exit
        rnfac=rnfac*real(ii)
       enddo

c20    write (6,*) r,rfac,nr,rnfac
 20    prob=enfac/rfac/rnfac*p**r*q**nr
       write (6,*) r,n-2*r,prob,r/en,en*prob,
     &  (real(r)-en*p)/sqrt(en*p*q),sqrt(en*p*q)*prob

      enddo
```

```
      stop
      end
```

２項分布の計算 FORTRAN プログラムその２（N が大きい場合)

```
c234567  bbinominal_dist.f

      integer  n,r,nr,ii
      real(8) :: enfac,rfac,rnfac,p,q
      real(8) :: prob,en,pi,enfac_sr,gammln


      pi=acos(-1.0d0)
      p=0.8d0
      q=1.0d0-p

      n=1000
      en=real(n)

      do r=1, n-1
       er=real(r)
       x=gammln(en+1.0d0)-gammln(er+1.0d0)-gammln(en-er+1.0d0)
       y=er*log(p)+(en-er)*log(q)

       prob=exp(x+y)
       write (6,*) r,n-2*r,prob,r/en,en*prob,
     &  (real(r)-en*p)/sqrt(en*p*q),sqrt(en*p*q)*prob

      enddo
      stop
      end

      FUNCTION gammln(xx)
      DOUBLE PRECISION  gammln,xx
      INTEGER j
      DOUBLE PRECISION ser,stp,tmp,x,y,cof(6)
      SAVE cof,stp
      DATA cof,stp/76.18009172947146d0,-86.50532032941677d0,
     *24.01409824083091d0,-1.231739572450155d0,.1208650973866179d-2,
     *-.5395239384953d-5,2.5066282746310005d0/
      x=xx
      y=x
      tmp=x+5.5d0
      tmp=(x+0.5d0)*log(tmp)-tmp
      ser=1.000000000190015d0
      do 11 j=1,6
        y=y+1.d0
        ser=ser+cof(j)/y
11    continue
      gammln=tmp+log(stp*ser/x)
      return
      END
```

13 Appendix2

14 スターリングの公式: Stirling's formula ゆるいバージョン (通常はこれで OK)

スターリングの公式 $\ln N! \simeq N \ln N - N, (N >> 1)$ [7]を導く。

$$\ln N! \;=\; \ln[N(N-1)(N-2)...321] = \sum_{k=1}^{N} \ln k \tag{81}$$

図に示すように $\ln x$ は x の単調増加関数である。図中の 2 つの長方形の面積と積分の関係より,

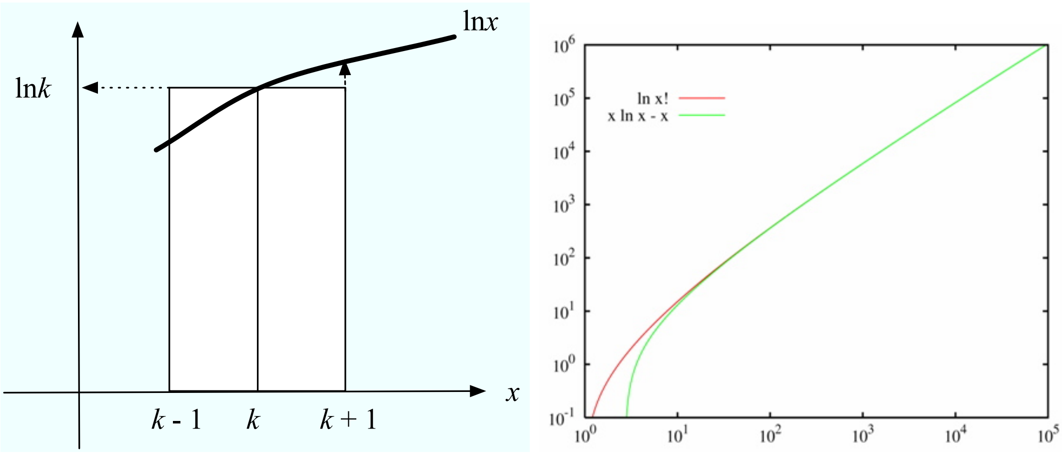

Figure 11: k vs $\ln k$(left), $\ln N!$ vs $N \ln N - N$(right)

k を整数として，以下の関係が成り立つことがわかる。

$$\int_{k-1}^{k} dx \ln x \;\leq\; [k-(k-1)]\ln k, \quad [(k+1)-k]\ln k \leq \int_{k}^{k+1} dx \ln x \tag{82}$$

$$\int_{k-1}^{k} dx \ln x \;\leq\; \ln k \leq \int_{k}^{k+1} dx \ln x \tag{83}$$

$k=1$ から $k=N$ まで和をとると

$$\int_{0}^{N} dx \ln x \;\leq\; \sum_{k=1}^{N} \ln k \leq \int_{1}^{N+1} dx \ln x \tag{84}$$

部分積分をつかうと[8],

$$\int_{0}^{N} dx \ln x \;=\; \int_{0}^{N} dx (x)' \ln x = [x \ln x]_0^N - \int_0^N x(1/x)dx = N \ln N - N \tag{85}$$

$$\int_{1}^{N+1} dx \ln x \;=\; [x \ln x]_1^{N+1} - \int_1^{N+1} x(1/x)dx = (N+1)\ln(N+1) - N \tag{86}$$

故に，$N >> 1$ の時,

$$\ln N! \simeq N \ln N - N \tag{87}$$

以下のようにもあらわすことができる。

$$e^{\ln N!} \;=\; e^{\ln N^N} e^{-N} \tag{88}$$

$$N! \;=\; N^N e^{-N} = \left(\frac{N}{e}\right)^N \tag{89}$$

[7]ゆるい版の式。詳しくは田崎晴明「統計力学 1」の付録を参照のこと。
[8]$(fg)' = f'g + fg', fg = \int f'g + \int fg', \int f'g = fg - \int fg'$

15　厳しいバージョン：中央極限定理の証明の際に用いる（小針，確率・統計入門）

以上の近似式はより正確には，

$$N! = \sqrt{2\pi N}\left(\frac{N}{e}\right)^N \tag{90}$$

となるがこれを示そう。

$$N\ln N - N \leq \ln N! \leq (N+1)\ln(N+1) - N \tag{91}$$

今，$\ln N! \simeq (N+1/2)\ln N - N$ と「あたり」をつけ，そのずれを以下の様に定義する。

$$d_N \equiv \ln N! - [(N+1/2)\ln N - N] \tag{92}$$

$N+1$ の時との差をとると

$$\begin{aligned}
d_{N+1} - d_N &= \ln(N+1) - (N+3/2)\ln(N+1) + (N+1) + (N+1/2)\ln N - N & (93)\\
&= -(N+1/2)\ln\frac{N+1}{N} + 1 = -(N+1/2)\ln(1+\frac{1}{N}) + 1 & (94)\\
&= -(N+1/2)[1/N - 1/(2N^2) + 1/(3N^3) - 1/(4N^4)...] + 1 & (95)\\
&= -[1 + 1/(2N) - 1/(2N) - 1/(4N^2) + 1/(3N^2) + O(N^{-3})] + 1 & (96)\\
&= -\frac{1}{12N^2} + O(N^{-3}) & (97)\\
d_{N+1} &= d_N - \frac{1}{12N^2} + O(N^{-3}) & (98)
\end{aligned}$$

差 d_N は，$N \to +\infty$ である値 C に収束する。従って，

$$\begin{aligned}
C &= \lim_{N\to+\infty}\{\ln N! - [(N+1/2)\ln N - N]\} & (99)\\
e^C &= \lim_{N\to+\infty}\frac{N!}{N^{N+1/2}e^{-N}} & (100)
\end{aligned}$$

$e^C \equiv A$ とすると，

$$\begin{aligned}
A^2 &= \lim_{N\to+\infty}\frac{(N!)^2}{N^{2N+1}e^{-2N}} & (101)\\
\frac{1}{A} &= \lim_{N\to+\infty}\frac{(2N)^{2N+1/2}e^{-2N}}{(2N)!} & (102)
\end{aligned}$$

この二つをかけると

$$\begin{aligned}
A &= \lim_{N\to+\infty}\frac{(N!)^2 2^{2N+1/2}N^{2N+1/2}}{(2N)!N^{2N+1}} & (103)\\
&= \lim_{N\to+\infty}\frac{2^{2N}\sqrt{2}}{\sqrt{N}[(2N)!/(N!N!)]} = \sqrt{2}\lim_{N\to+\infty}\frac{2^{2N}}{\sqrt{N}{}_{2N}\mathrm{C}_N} & (104)
\end{aligned}$$

ここで ${}_N\mathrm{C}_r = N!/[r!(N-r)!]$ の二項係数である。$A = \sqrt{2\pi}$ となるが，まずワリスの公式を導こう。

$$\begin{aligned}
S_n &= \int_0^{\pi/2} \sin^n x \mathrm{d}x = \left[-\cos x \sin^{n-1} x\right]_0^{\pi/2} + (n-1)\int_0^{\pi/2} \sin^{n-2} x \cos^2 x \mathrm{d}x & (105)\\
(fg)' &= f'g + fg', \quad \int f'g = fg - \int fg', f' = \sin x, f = -\cos x, g = \sin^{n-1} x & (106)\\
S_n &= (n-1)\int_0^{\pi/2} \sin^{n-2} x(1-\sin^2 x)\mathrm{d}x = (n-1)S_{n-2} - (n-1)S_n & (107)\\
S_n &= \frac{n-1}{n} S_{n-2} & (108)\\
S_{2n} &= \frac{2n-1}{2n}\frac{2n-3}{2n-2}...\frac{3}{4}\frac{1}{2}S_0, \quad S_{2n+1} = \frac{2n}{2n+1}\frac{2n-2}{2n-1}...\frac{4}{5}\frac{2}{3}S_1 & (109)\\
S_0 &= \int_0^{\pi/2} \mathrm{d}x = \frac{\pi}{2}, \quad S_1 = \int_0^{\pi/2} \sin x \mathrm{d}x = [-\cos x]_0^{\pi/2} = -(0-1) = 1 & (110)\\
\frac{S_{2n+1}}{S_{2n}} &= \left(\frac{2n}{2n+1}\frac{2n-2}{2n-1}...\frac{4}{5}\frac{2}{3}S_1\right) / \left(\frac{2n-1}{2n}\frac{2n-3}{2n-2}...\frac{3}{4}\frac{1}{2}S_0\right) & (111)\\
&= \frac{2n}{2n+1}\frac{2n}{2n-1}\frac{2n-2}{2n-1}\frac{2n-2}{2n-3}...\frac{4}{5}\frac{4}{3}\frac{2}{3}\frac{2}{1}\frac{2}{\pi} & (112)
\end{aligned}$$

$0 < x < \pi/2$ で，

$$\begin{aligned}
0 &< \sin^{2n+1} x < \sin^{2n} x < \sin^{2n-1} x & (113)\\
0 &< S_{2n+1} < S_{2n} < S_{2n-1} & (114)\\
1 &< \frac{S_{2n}}{S_{2n+1}} < \frac{S_{2n-1}}{S_{2n+1}} = \frac{S_{2n-1}}{\frac{2n}{2n+1}S_{2n-1}} = \frac{2n+1}{2n} & (115)\\
\lim_{n\to\infty} \frac{S_{2n}}{S_{2n+1}} &= 1 & (116)\\
\lim_{n\to\infty} \frac{2\cdot 2\cdot 4\cdot 4\cdots(2n-2)(2n-2)2n\cdot 2n}{1\cdot 3\cdot 3\cdot 5\cdots(2n-1)(2n-1)(2n-1)(2n+1)} &= \frac{\pi}{2} & (117)
\end{aligned}$$

最後の式をワリスの公式と呼ぶ。その式の左辺は，

$$\begin{aligned}
\prod_{n=1}^{\infty}\left[\frac{2n\cdot 2n}{(2n-1)(2n+1)}\right] = \prod_{n=1}^{\infty}\left[\frac{(2n)^2}{(2n)^2-1}\right] &= \frac{\pi}{2} & (118)\\
\prod_{n=1}^{\infty}\left[1-\frac{1}{(2n)^2}\right] &= \frac{2}{\pi} & (119)
\end{aligned}$$

また，式 (29) より

$$\begin{aligned}
S_{2n}S_{2n+1} &= \frac{2n-1}{2n}\frac{2n-3}{2n-2}...\frac{3}{4}\frac{1}{2}S_0\frac{2n}{2n+1}\frac{2n-2}{2n-1}...\frac{4}{5}\frac{2}{3}S_1 & (120)\\
&= \frac{2n}{2n+1}\frac{2n-1}{2n}\frac{2n-2}{2n-1}\frac{2n-3}{2n-2}...\frac{4}{5}\frac{3}{4}\frac{2}{3}\frac{1}{2}S_1S_0 = \frac{1}{2n+1}\frac{\pi}{2} & (121)\\
\sqrt{n}S_{2n+1}\sqrt{\frac{S_{2n}}{S_{2n+1}}} &= \sqrt{\frac{n}{2n+1}}\sqrt{\frac{\pi}{2}} & (122)\\
\lim_{n\to\infty}\frac{S_{2n}}{S_{2n+1}} &= 1, \quad \lim_{n\to\infty}\frac{n}{2n+1} = \frac{1}{2} & (123)\\
\lim_{n\to\infty}\sqrt{n}S_{2n+1} &= \frac{\sqrt{\pi}}{2} & (124)\\
S_{2n+1} &= \frac{2n}{2n+1}\frac{2n-2}{2n-1}...\frac{4}{5}\frac{2}{3}S_1 = \frac{(2n)^2(2n-2)^2...4^2\cdot 2^2}{(2n+1)!} & (125)\\
&= \frac{2^2(n)^2 2^2(n-1)^2...2^2 2^1\cdot 2^2 2^0}{(2n+1)(2n)!} = \frac{(n!)^2 2^{2n}}{(2n+1)(2n)!} = \frac{2^{2n}}{(2n+1)\,{}_{2n}\mathrm{C}_n} & (126)
\end{aligned}$$

式 (44) より,

$$\sqrt{\pi} = \lim_{n\to\infty} 2\sqrt{n} S_{2n+1} = \lim_{n\to\infty} \frac{2\sqrt{n}2^{2n}}{(2n+1)_{2n}\mathrm{C}_n} \tag{127}$$

$$\lim_{n\to\infty} \frac{2\sqrt{n}\sqrt{n}}{2n+1} = 1 \tag{128}$$

$$\sqrt{\pi} = \lim_{n\to\infty} \frac{2^{2n}}{\sqrt{n}_{2n}\mathrm{C}_n} \tag{129}$$

従って, (20),(24) 式の $e^C = \sqrt{2\pi}$ となる。スターリングの公式は,

$$\lim_{N\to\infty} \frac{N!}{\sqrt{2\pi N} N^N e^{-N}} = 1 \tag{130}$$

$$N! \simeq \sqrt{2\pi} N^{N+1/2} e^{-N} \tag{131}$$

$$\ln N! \simeq (N + \frac{1}{2}) \ln N - N + \ln \sqrt{2\pi} \tag{132}$$