数理統計学 補足プリント

2009 年 5 月, 長谷川浩司

1 準備 1:スターリングの公式等

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n. \qquad (1)$$

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - + \cdots. \quad (|x|<1) \qquad (2)$$

$$(\textbf{Stirling の公式})\ n! \simeq \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \quad (n\to\infty) \qquad (3)$$

ただし以下 $\simeq$ は、左辺/右辺 $\overset{n\to\infty}{\longrightarrow} 1$ の意味とする。

2 ガウス積分 $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$.

$\because$ 左辺を I と置き、その 2 乗を計算すると

$$\begin{aligned} I^2 &= \int_{x\text{ 直線}}\int_{y\text{ 直線}} e^{-x^2}e^{-y^2}dxdy \\ &= \int\int_{\text{全平面}} e^{-r^2}dxdy \quad (r^2=x^2+y^2) \\ &= \int_{r=0}^{\infty} e^{-r^2}\times 2\pi r dr = 2\pi\left[\frac{e^{-r^2}}{-2}\right]_0^{\infty} = \pi. \end{aligned} \qquad (4)$$

3 正規分布の再生性

$X \sim N(\mu, \sqrt{A}^2), Y \sim N(\nu, \sqrt{B}^2)$ (独立) [1] のとき、$X+Y \sim N(\mu+\nu, \sqrt{A}^2+\sqrt{B}^2)$ を示す。

(X,Y) の同時分布を $C_t : X+Y=t$(一定) 上で加え,

$$\int_{t=x+y:\text{一定}} \frac{e^{-\frac{x^2}{2A}}}{\sqrt{2\pi A}}dx \frac{e^{-\frac{y^2}{2B}}}{\sqrt{2\pi B}}dy = \int_{C_t} e^{-\frac{B(t-y)^2+Ay^2}{2AB}} \frac{dxdy}{2\pi\sqrt{AB}}$$

が和 $X+Y$ の分布を与える。C_t 上の積分を変数 y で行うと、$dxdy = dtdy$, また

$B(t-y)^2 + Ay^2 = (A+B)y^2 - 2Byt + Bt^2 = (A+B)(y-\frac{Bt}{A+B})^2 + \frac{AB}{A+B}t^2$ なので, 上の積分は

$$\begin{aligned} &\int_{y=-\infty}^{\infty} \frac{e^{-\{(A+B)(y-\frac{Bt}{A+B})^2+\frac{AB}{A+B}t^2\}/2AB}}{2\pi\sqrt{AB}}dtdy \\ &= \frac{e^{-\frac{t^2/2}{A+B}}}{\sqrt{2\pi(A+B)}}dt = N(0, A+B). \end{aligned}$$

[1] 本の σ_A, σ_B を $\sqrt{A}, \sqrt{B}$ と書く。尚 $\mu, \nu = 0$ で十分。

4 2 項分布 $\overset{n\to\infty}{\longrightarrow}$ 正規分布について

2 項分布 $Bin(n,p)$ $(q=1-p)$ について

$$P(X=k) \overset{n\to\infty}{\simeq} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\frac{1}{\sqrt{npq}} \quad (z = \frac{k-np}{\sqrt{npq}})$$

を示す。$P(X=k) = \frac{n!}{k!(n-k)!}p^kq^{n-k}$ に Stirling の公式 (3) を使うと, n が大のとき ($pq=s^2$ とおく)

$$\begin{aligned} \sqrt{2\pi npq}P(X=k) &\simeq \sqrt{\frac{n^2pq}{k(n-k)}}\frac{n^np^kq^{n-k}}{k^k(n-k)^{n-k}} \\ &= \left(\frac{np}{k}\right)^{k+\frac12}\left(\frac{nq}{n-k}\right)^{n-k+\frac12}, \qquad k = np+\sqrt{n}sz \\ &= \left(\frac{np}{np+\sqrt{n}sz}\right)^{np+\sqrt{n}sz+\frac12}\left(\frac{nq}{nq-\sqrt{n}sz}\right)^{nq-\sqrt{n}sz+\frac12} \\ &\simeq \left(1+\frac{sz}{\sqrt{np}}\right)^{-np-\sqrt{n}sz}\left(1-\frac{sz}{\sqrt{nq}}\right)^{-nq+\sqrt{n}sz} \end{aligned} \qquad (5)$$

この対数をとり、テイラー展開 (2) を使い極限をとると

$$\begin{aligned} \log((5)\text{ 式}) &= -np(1+\frac{sz}{\sqrt{np}})\log\left(1+\frac{sz}{\sqrt{np}}\right) \\ &\quad - nq(1-\frac{sz}{\sqrt{nq}})\log\left(1-\frac{sz}{\sqrt{nq}}\right) \\ &= -np(1+\frac{sz}{\sqrt{np}})\left(\frac{sz}{\sqrt{np}} - \frac{(\frac{sz}{\sqrt{np}})^2}{2} + \frac{(\frac{sz}{\sqrt{np}})^3}{3} - \cdots\right) \\ &\quad +nq(1-\frac{sz}{\sqrt{nq}})\left(\frac{sz}{\sqrt{nq}} + \frac{(\frac{sz}{\sqrt{nq}})^2}{2} + \frac{(\frac{sz}{\sqrt{nq}})^3}{3} + \cdots\right) \\ &= -np\left\{(Z-\frac{Z^2}{2}+\frac{Z^3}{3}-\cdots)+(Z^2-\frac{Z^3}{2}+\cdots)\right\} \\ &\quad +nq\left\{(Z'+\frac{Z'^2}{2}+\frac{Z'^3}{3}+\cdots)-(Z'^2+\frac{Z^3}{2}+\cdots)\right\} \end{aligned}$$

$(Z = \frac{sz}{\sqrt{np}}, Z' = \frac{sz}{\sqrt{nq}}.)$ $n\to\infty$ では z^2 の項だけ残り、

$$\longrightarrow -p\frac{(\frac{sz}{p})^2}{2} - q\frac{(\frac{sz}{q})^2}{2} \overset{s^2=pq}{=} -\frac{qz^2}{2} - \frac{pz^2}{2} = -\frac{z^2}{2}$$

となる。$\therefore (5) \overset{n\to\infty}{\longrightarrow} e^{-z^2/2}$.

5 ポアソン分布の導出

$$P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda} \quad (k=0,1,2,\cdots)$$

である離散確率分布をポアソン分布という ($\lambda > 0$ とする)。その期待値と分散は、[p63] にあるように

$$E[X] = \lambda,\ V[X] = \lambda.$$

ポアソン分布は、p が小さいが n が大きい場合 ($np = \lambda$ が一定) の 2 項分布 (の極限) として得られる:

$$\begin{aligned}
&\because \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\
&\overset{np=\lambda}{=} \frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!}\cdot\frac{n(n-1)\cdots(n-k+1)}{n^k}\cdot\frac{(1-\lambda/n)^n}{(1-\lambda/n)^k} \\
&\overset{n\to\infty}{\longrightarrow}= \frac{\lambda^k}{k!}\cdot 1\cdot\frac{e^{-\lambda}}{1}.
\end{aligned}$$

6 ポアソン分布 → 正規分布

ポアソン分布も $\lambda \to \infty$ の極限で正規分布となる。

$$\frac{\lambda^k}{k!}e^{-\lambda} \overset{\lambda\to\infty}{\simeq} \frac{e^{-y^2/2}}{\sqrt{2\pi\lambda}} \quad (\text{ただし } y = \frac{k-\lambda}{\sqrt{\lambda}})^2 \qquad (6)$$

示すには、(5) のときと同様に対数をとる：$\log\left(\frac{\lambda^k}{k!}e^{-\lambda}\right)$

$$\begin{aligned}
&\overset{(3)}{\simeq} \log\left(\frac{\lambda^{\sqrt{\lambda}y+\lambda}e^{-\lambda}}{(\sqrt{\lambda}y+\lambda)^{\sqrt{\lambda}y+\lambda}}\frac{1}{e^{-\sqrt{\lambda}y-\lambda}\sqrt{2\pi(\sqrt{\lambda}y+\lambda)}}\right) \\
&= \log\left((1+\frac{y}{\sqrt{\lambda}})^{-(\sqrt{\lambda}y+\lambda)}\frac{e^{\sqrt{\lambda}y}}{\sqrt{2\pi(\sqrt{\lambda}y+\lambda)}}\right) \\
&= -\lambda(1+\frac{y}{\sqrt{\lambda}})\log(1+\frac{y}{\sqrt{\lambda}})+\sqrt{\lambda}y-\frac{\log(2\pi\lambda(1+\frac{y}{\sqrt{\lambda}}))}{2} \\
&\overset{(2)}{=} -\lambda(1+\frac{y}{\sqrt{\lambda}})\left\{\frac{y}{\sqrt{\lambda}}-\frac{1}{2}(\frac{y}{\sqrt{\lambda}})^2+\frac{1}{3}(\frac{y}{\sqrt{\lambda}})^3-\cdots\right\} \\
&+\lambda\frac{y}{\sqrt{\lambda}}-\frac{\log(2\pi\lambda(1+\frac{y}{\sqrt{\lambda}}))}{2} \overset{\lambda\to\infty}{\simeq} -\frac{y^2+\log\sqrt{2\pi\lambda}}{2}.
\end{aligned}$$

7 準備 2：ガンマ関数

ガンマ関数は次で定義される。

$$\Gamma(x) = \int_0^\infty t^x e^{-t}\frac{dt}{t} \qquad (7)$$

部分積分により

$$\begin{aligned}
&\int_0^\infty t^x e^{-t}dt = [-t^x e^{-t}]_{t=0}^\infty + \int_0^\infty x t^{x-1}e^{-t}dt \\
\therefore\ &\Gamma(x+1) = x\Gamma(x) \qquad (8)
\end{aligned}$$

である。また次が成り立つ。

$$\begin{aligned}
\Gamma(1) &= \int_0^\infty e^{-t}dt = [-e^{-t}]_0^\infty = 1, \qquad (9) \\
\Gamma(\frac{1}{2}) &= \int_0^\infty e^{-t}\frac{dt}{\sqrt{t}} \overset{t=x^2}{=} 2\int_0^\infty e^{-x^2}dx \overset{(4)}{=} \sqrt{\pi} \qquad (10)
\end{aligned}$$

よって $x = n$ (自然数) のとき $\Gamma(n+1) = n!$ であり、$x = n + \frac{1}{2}$ (半整数) のときの値もわかる。

8 χ^2 分布の導出

χ^2 分布の定義より、$X_1, \cdots, X_n \sim N(0,1)$ のとき

$$\int_0^r f(x)dx = P\left[0 \le \sum_{i=1}^n X_i^2 \le r\right]$$

となる f が、$\chi^2(n)$ の分布関数である。

$$\text{右辺} = \int_{x_1^2+\cdots+x_n^2\le r}\frac{e^{-x_1^2/2}}{\sqrt{2\pi}}\cdots\frac{e^{-x_n^2/2}}{\sqrt{2\pi}}dx_1\cdots dx_n$$

$$= \int_{R^2\le r}\frac{e^{-R^2/2}}{\sqrt{2\pi}^n}dx_1\cdots dx_n \quad (R = \sqrt{x_1^2+\cdots+x_n^2})$$

であり、$dx_1\cdots dx_n = R^{n-1}dRd\Sigma$ ($d\Sigma$ は $n-1$ 次元単位「球面」S^{n-1} の「面素」) なので

$$= \int_{S^{n-1}}d\Sigma\int_{R=0}^{\sqrt{r}}\frac{e^{-R^2/2}}{\sqrt{2\pi}^n}R^{n-1}dR \overset{R^2=x}{=} K\int_0^r e^{-\frac{x}{2}}x^{\frac{n-1}{2}}\frac{dx}{2\sqrt{x}}$$

$\therefore f(x) = \frac{K}{2}e^{-\frac{x}{2}}x^{\frac{n}{2}-1}$. 定数 $\frac{K}{2}$ は「全確率 1」で決まり

$$\begin{aligned}
1 &= \frac{K}{2}\int_0^\infty e^{-\frac{x}{2}}x^{\frac{n}{2}-1}dx \overset{x=2t}{=} \frac{K}{2}\int_0^\infty e^{-t}\sqrt{2t}^n\frac{dt}{t} \\
&= \frac{K}{2}2^{\frac{n}{2}}\Gamma(\frac{n}{2}), \quad \therefore f(x) = \frac{e^{-x/2}}{2\Gamma(n/2)}\left(\frac{x}{2}\right)^{\frac{n}{2}-1}. \qquad (11)
\end{aligned}$$

9 χ^2 分布の再生性など

定義より、次の**再生性**は明らか。

$$X \sim \chi^2(n), Y \sim \chi^2(m) \Rightarrow X+Y \sim \chi^2(m+n) \quad (12)$$

実際、$X = X_1^2+\cdots+X_n^2, Y = Y_1^2+\cdots+Y_m^2 (X_i, Y_i \sim N(0,1))$ とすれば、$X+Y = X_1^2+\cdots+X_n^2+Y_1^2+\cdots+Y_m^2$ も正規確率変数の 2 乗和だからである。

[2] このとき $dy \simeq \frac{1}{\sqrt{\lambda}}$ であり、(6) は $P(X{=}k){=}P(|X-k|{<}\frac{1}{2}) = P(y-\frac{dy}{2} < Y < y+\frac{dy}{2}) = \frac{e^{-y^2/2}dy}{\sqrt{2\pi}}$ ともいえる。

同様に考え、$X \sim \chi^2(n)$ の期待値と分散もわかる：

$$E[X] = E[X_1^2 + \cdots + X_n^2] = E[X_1^2] + \cdots + E[X_n^2] = n \tag{13}$$

($\because X_i \sim N(0,1)$ のとき、$E[X_i^2] = V[X_i] = 1$.) また

$$V[X] = V[X_1^2] + \cdots + V[X_n^2] = 2n. \tag{14}$$

($\because V[X_i^2] = E[X_i^4] - E[X_i^2]^2 \overset{(18)}{=} 3 - 1^2 = 2$.)

10 t 分布の導出

定義より、2 次元分布 $(X, Y) \sim (N(0,1), \chi^2(n))$ の密度を、$T = \frac{X}{\sqrt{Y/n}}$: 一定の上で足し上げればよい。

$$(X \text{ の密度}) \cdot (Y \text{ の密度}) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \frac{e^{-y/2}(\frac{y}{2})^{n/2}}{\Gamma(n/2)} \frac{dy}{y}$$

を $(x,y) \to (t = \frac{x}{\sqrt{y/n}}, y)$ と変数変換し y で積分する: [3]

$$\begin{aligned}
\int_{t\text{ 一定},y>0} &= dt \int_0^\infty \frac{e^{-\frac{t^2 y}{2n} - \frac{y}{2}}(\frac{y}{2})^{n/2}}{\sqrt{2\pi}\Gamma(n/2)} \frac{dy}{\sqrt{ny}} \\
&\overset{\frac{y}{2}=\eta}{=} \frac{dt}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \int_{\eta=0}^\infty e^{-(1+\frac{t^2}{n})\eta} \eta^{\frac{n+1}{2}} \frac{d\eta}{\eta} \\
&\overset{(1+\frac{t^2}{n})\eta=s}{=} \frac{dt}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \int_{s=0}^\infty e^{-s} \left(\frac{s}{1+\frac{t^2}{n}}\right)^{\frac{n+1}{2}} \frac{ds}{s} \\
&= \frac{\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} \Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} dt = (t(n) \text{ の密度}) \cdot dt.
\end{aligned}$$

11 t 分布の分散とベータ関数

$T \sim t(n)$ のとき、$E[T] = 0, V[T] = \frac{n}{n-2}$ である。

期待値 $= 0$ は左右対称ゆえだが、分散の計算はベータ関数 $B(p,q)$ が必要である。その定義は

$$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1} dt \overset{t=\frac{1}{x+1}}{=} \int_0^\infty \frac{x^{p-1}dx}{(1+x)^{p+q}}.$$

次の有名な公式が成り立つ。

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \tag{15}$$

すると

$$\begin{aligned}
V[T] = E[T^2] &= \int_{-\infty}^\infty t^2 \frac{\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} \Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} dt \\
&\overset{\frac{t^2}{n}=x}{=} \frac{2\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \int_0^\infty \frac{\sqrt{n^3}}{2} \frac{\sqrt{x}dx}{(1+x)^{\frac{n+1}{2}}} \\
&= \frac{2\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \cdot \frac{\sqrt{n^3}}{2} B\left(\frac{3}{2}, \frac{n}{2} - 1\right) \\
&\overset{(15)}{=} \frac{2\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \cdot \frac{\sqrt{n^3}}{2} \frac{\Gamma(\frac{3}{2})\Gamma(\frac{n}{2}-1)}{\Gamma(\frac{n+1}{2})} = \frac{n\Gamma(\frac{3}{2})\Gamma(\frac{n}{2}-1)}{\sqrt{\pi}\Gamma(\frac{n}{2})} \\
&\overset{(8)}{=} \frac{n \cdot \frac{1}{2}\Gamma(\frac{1}{2})}{\sqrt{\pi}(\frac{n}{2}-1)} \overset{(10)}{=} \frac{n}{n-2}.
\end{aligned}$$

12 F 分布の導出

$X \sim \chi^2(m), Y \sim \chi^2(n)$ のとき

$$T = \frac{X/m}{Y/n} \sim F(m,n)$$

が F 分布の定義であった。t 分布のときと同様に考え、

$$(\chi^2(m) \text{ の密度})(\chi^2(n) \text{ の密度}) = \frac{e^{-\frac{x}{2}} x^{\frac{m}{2}-1} dx}{\sqrt{2^m}\Gamma(\frac{m}{2})} \frac{e^{-\frac{y}{2}} y^{\frac{n}{2}-1} dy}{\sqrt{2^n}\Gamma(\frac{n}{2})}$$

を $\frac{x}{y} = s$:一定 の上で y 積分すると $(t = \frac{n}{m}s)$、[4]

$$\begin{aligned}
&\int_{s\text{ 一定},y>0} \frac{e^{-\frac{x}{2}} x^{\frac{m}{2}-1} dx}{\sqrt{2^m}\Gamma(\frac{m}{2})} \frac{e^{-\frac{y}{2}} y^{\frac{n}{2}-1} dy}{\sqrt{2^n}\Gamma(\frac{n}{2})} \\
&= \frac{s^{-1}ds}{\sqrt{2}^{m+n}\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \int_{y=0}^\infty e^{-\frac{ys}{2}} (ys)^{\frac{m}{2}} e^{-\frac{y}{2}} y^{\frac{n}{2}} \frac{dy}{y} \\
&= \frac{s^{\frac{m}{2}-1}ds}{\sqrt{2}^{m+n}\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \int_{y=0}^\infty e^{-\frac{1+s}{2}y} y^{\frac{m+n}{2}} \frac{dy}{y} \\
&= \frac{s^{\frac{m}{2}-1}ds}{\sqrt{2}^{m+n}\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \frac{\Gamma(\frac{m+n}{2})}{(\frac{1+s}{2})^{\frac{m+n}{2}}} \\
&\overset{\frac{n}{m}s=t}{=} \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \frac{(\frac{m}{n}t)^{\frac{m}{2}}}{(1+\frac{m}{n}t)^{\frac{m+n}{2}}} \frac{dt}{t} = (F \text{ の密度}) \cdot dt.
\end{aligned}$$

13 準備3:分配関数

確率変数 X は、X の任意の関数 $g(X)$ の期待値 $E[g(X)]$ で決まる。($g(x)$ が $a \le x \le b$ で 1, これ以外では 0, という関数のときの期待値 $E[g(X)]$ が、$P(a \le X \le b)$ である。) $g(x)$ をテーラー展開して考

[3] $x^2 = \frac{t^2}{y}$, $2xdx = \frac{2tdt \cdot y + t^2 dy}{n}$ $\therefore dx \wedge \frac{dy}{y} = \frac{1}{2x}\frac{2tdt}{n}dy = \frac{dtdy}{\sqrt{ny}}$.

[4] $\frac{dx}{x} = \frac{yds+sdy}{ys}$, $\therefore \frac{dx}{x}\frac{dy}{y} = \frac{dyds}{ys}$.

えれば、$E[X], E[X^2], \cdots$ (n 次モーメント全体) で X が決まると言っても良い。これらを足し上げた

$$E[X^0] + tE[X^1] + \frac{t^2}{2}E[X^2] + \frac{t^3}{3!}E[X^3] + \cdots = E[e^{tX}] \tag{16}$$

を X の分配関数と言う。$t = 0$ で n 回微分すれば $E[X^n]$ が復元されるから、X は分配関数 $E[e^{tX}]$ で定まると言ってもよいことになる。

2 次元確率変数についても同様で、(X, Y) は

$$E[e^{sX+tY}] = E[1] + E[sX+tY] + \frac{1}{2}E[(sX+tY)^2] + \cdots$$

で完全に決まる。k 次元の $(X_1, \cdots, X_k)$ でも同様。

例：$X \sim N(0,1)$ の分配関数は

$$\begin{aligned} E[e^{tX}] &= \int_{-\infty}^{\infty} e^{tx} f(x) dx \\ &= \int_{-\infty}^{\infty} e^{tx} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \\ &= \int_{-\infty}^{\infty} e^{(-(x-t)^2+t^2)/2} \frac{dx}{\sqrt{2\pi}} \overset{(4)}{=} e^{t^2/2}. \end{aligned} \tag{17}$$

特に、これを展開して $\frac{t^4}{4!}$ の係数をみれば

$$E[e^{tX}] = 1 + \frac{t^2}{2} + \frac{1}{2}\left(\frac{t^2}{2}\right)^2 + \cdots, \ \therefore E[X^4]{=}3. \tag{18}$$

14　k 項分布 $\overset{n\to\infty}{\longrightarrow}$ $k-1$ 次元正規分布

$p_1 + \cdots + p_k = 1$ とする。k 項分布とは、$n_1 + \cdots + n_k = n$ である自然数 $n_1, \cdots, n_k$ に対し

$$P(X_1 = n_1, \cdots, X_k = n_k) := \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k}$$

というもの。これを分配関数で表して極限を見よう。

$$\begin{aligned} &E[X_1^{m_1} \cdots X_k^{m_k}] \\ &= \sum_{n_1+\cdots+n_k=n} n_1^{m_1} \cdots n_k^{m_k} P(X_1 = n_1, \cdots, X_k = n_k) \end{aligned}$$

であるから

$$\begin{aligned} &E[e^{t_1X_1} \cdots e^{t_kX_k}] \\ =& \sum_{m_1,\cdots,m_k \geq 0} \frac{t_1^{m_1} \cdots t_k^{m_k}}{m_1! \cdots m_k!} E[X_1^{m_1} \cdots X_k^{m_k}] \\ =& \sum_{m_1,\cdots,m_k \geq 0} \sum_{n_1+\cdots+n_k=n} \frac{t_1^{m_1} \cdots t_k^{m_k}}{m_1! \cdots m_k!} n_1^{m_1} \cdots n_k^{m_k} \\ & \times \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k} \\ =& \sum_{n_1+\cdots+n_k=n} e^{n_1t_1} \cdots e^{n_kt_k} \frac{n!}{n_1! \cdots n_k!} p_1^{n_1} \cdots p_k^{n_k} \\ =& (e^{t_1}p_1 + \cdots + e^{t_k}p_k)^n. \end{aligned}$$

$$\begin{aligned} \therefore \quad E\left[e^{\frac{t_1X_1+\cdots+t_kX_k}{\sqrt{n}}}\right] &= \left(e^{\frac{t_1}{\sqrt{n}}}p_1 + \cdots + e^{\frac{t_k}{\sqrt{n}}}p_k\right)^n \\ &= \left(1 + \frac{t_1p_1 + \cdots + t_kp_k}{\sqrt{n}} + \frac{t_1^2p_1 + \cdots + t_k^2p_k}{2\sqrt{n}^2} + \cdots\right)^n \\ &\overset{n\to\infty}{\longrightarrow} e^{(t_1^2p_1+\cdots+t_k^2p_k)/2} \quad (t_1p_1 + \cdots + t_kp_k = 0 \text{ とする}) \end{aligned}$$

よって、$n \to \infty$ のとき $X_i/\sqrt{n} \to z_i$ とすれば、

$$E[e^{t_1z_1+\cdots+t_kz_k}] = e^{(t_1^2p_1+\cdots+t_k^2p_k)/2}$$

(ただし $t_1p_1 + \cdots + t_kp_k = 0$.) これは $k-1$ 次元正規分布の分配関数に等しく、実際フーリエ変換により

$$\frac{e^{-\frac{z_1^2/p_1+\cdots+z_k^2/p_k}{2}}}{\sqrt{(2\pi)^{n-1}p_1 \cdots p_k}} \quad (z_1 + \cdots + z_k = Const.)$$

を得るから、$(z_1, \cdots, z_k)$ は $k-1$ 次元正規分布に従う。

これより、次も判る [**p66 性質 5**]:

$$\sum_{i=1}^{k} \left(\frac{X_i - np_i}{\sqrt{np_i}}\right)^2 \simeq \sum_{i=1}^{k} (\frac{z_i}{\sqrt{p_i}})^2 \sim \chi^2(k-1). \tag{19}$$

15　中心極限定理の証明

$E[X]{=}0$ としてよい ($\because V[X] = E[X^2] = \sigma^2$)。[5] すると $\bar{X}{=}\frac{X_1+\cdots+X_n}{n}$ ($X_i{\sim}X$) について

$$\begin{aligned} &\sum_{m=0}^{\infty} \frac{t^m}{m!} E\left[\left(\frac{\overline{X}}{\sigma/\sqrt{n}}\right)^m\right] = E\left[e^{\frac{t\sqrt{n}}{\sigma}\bar{X}}\right] = E\left[e^{t\frac{X_1+\cdots+X_n}{\sqrt{n}\sigma}}\right] \\ &= E\left[e^{\frac{tX_1}{\sqrt{n}\sigma}}\right] \cdots E\left[e^{\frac{tX_n}{\sqrt{n}\sigma}}\right] = E\left[e^{\frac{tX}{\sqrt{n}\sigma}}\right]^n \\ &= \left\{1 + (\frac{tX}{\sqrt{n}\sigma})E[X] + \frac{1}{2}(\frac{tX}{\sqrt{n}\sigma})^2E[X^2] + \cdots\right\}^n \\ &\overset{n\to\infty}{\longrightarrow} e^{t^2/2} = E[e^{tZ}], \ Z \sim N(0,1). \ (\because (1)) \end{aligned}$$

$$\therefore \frac{\bar{X}}{\sigma/\sqrt{n}} \overset{n\to\infty}{\longrightarrow} Z \sim N(0,1). \tag{20}$$

16　参考書

高木貞治「解析概論」岩波書店

田栗正章「統計学とその応用」放送大学

[5] 更に、$E[X^m]$ が全ての m について存在すると仮定する。