Introduction to mathematical statistics ゼミ (第？回)

担当: 伊藤真道

未定

3 Chap.3. Some Special Distributions

3.1 The Binomial and Related Distributions

まず，**ベルヌーイ実験** *(Bernoulli experiment)* と**ベルヌーイ試行** *(Bernoulli trials)* を紹介する．ベルヌーイ実験は，二つの互いに排反な結果 (例えば，成功と失敗，裏と表など) だけが起こりうる実験である．ベルヌーイ試行は，ベルヌーイ実験が独立に複数回行われた際の試行のことである．以下 p を成功の確率とする．

X をベルヌーイ試行に対する確率変数とした時，以下のように定義される．

$$X(\text{成功}) = 1 \ and \ X(\text{失敗}) = 0$$

つまり，二つの結果が $1, 0$ で表現される．X の確率質量関数 (probability mass function, pmf) は，

$$p(x) = p^x(1-p)^{1-x}, \ x = 0, 1 \tag{3.1.1}$$

と表され，これを，ベルヌーイ分布と呼ぶ．ベルヌーイ分布の平均，分散は，それぞれ，

$$\mu = E(X) = p$$
$$\sigma^2 = V(X) = p(1-p)$$

である．

証明 3.1. 簡単なので，実際に導出してみましょう．

長さ n のベルヌーイ試行の列にて，X_i を i 番目のベルヌーイ分布の確率変数とする．この列の観測値は，長さ n の $\{0\},\{1\}$ で構成された列となる．この時，多くのケースで興味の対象となるのは，成功回数の総数である．もし，確率変数 X が観測された成功回数を表すとすると，成功した試行の並び順は，

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

通り考えられる．試行は，全て独立に，成功確率 p，失敗確率 $1-p$ で行われているため，全てをまとめると，$p(x)$ は，

$$p(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 1, 2, \ldots, n \\ 0 & elsewhere \end{cases} \tag{3.1.2}$$

となる．また，n が正整数なら，

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} b^x a^{n-x}$$

であることに注意すると，

$$\begin{aligned} \sum_x p(x) &= \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} \\ &= [(1-p)+p]^{n-x+x} = 1 \end{aligned}$$

となり，これと，$p(x) \geq 0$ であることから，$p(x)$ は離散型の pmf であることがわかる．pmf として，この $p(x)$ を持つような離散型確率変数は，**二項分布** *(binomial distribution)* に従うという．二項分布は $b(n,p)$ や $Bin(n,p)$(あたしはこっちのが好き) のように表され，定数 n,p は二項分布のパラメータと呼ばれる．

二項分布の mgf，平均，分散

$$\begin{aligned} M(t) &= [(1-p)+pe^t]^n \\ \mu &= E[X] = np \\ \sigma^2 &= V[X] = np(1-p) \end{aligned}$$

となる．

[l]

証明 3.2. 簡単なので，導出してみましょう

成功確率 p の独立で連続して行われた実験を考える．Y を r 回成功するまでに失敗する回数を表す確率変数とする．この時，r 回成功するためには，$Y+r$ 回の試行を行わないといけない．実際には，$y+r-1$ 回目の試行まで考えれば良いので，この時の確率は，

$$\binom{y+r-1}{r-1}p^{r-1}(1-p)^y$$

であり，$y+r$ 番目には成功するため，Y の pmf は

$$p_Y(y) = \begin{cases} \binom{y+r-1}{r}p^{r-1}(1-p)^y & y = 1, 2, \ldots, \\ 0 & elsewhere \end{cases} \tag{3.1.4}$$

となる．この分布は**負の二項分布** *(negative binomial distribution)* と呼ばれる．この分布の mgf は $M(t) = p^r[1-(1-p)e^t]^{-r}, t < -log(1-p)$ であり，もし，$r = 1$ なら，

$$p_Y(y) = p(1-p)^y \tag{3.1.5}$$

となり，幾何分布 (geometric distribution) であることがわかる．

定理 3.1.1

$X_i \sim Bin(n_i, p), i = 1, 2, \ldots, m$ とし，各標本は互いに独立とする．この時，

$$Y = \sum_{i=1}^{m} X_i \sim Bin(\sum_{i=1}^{m} n_i, p)$$

である．

証明 3.1. X_i の *mgf* は $M_{X_i}(t) = (1-p+pe^t)^{n_i}$ である．独立性から，

$$M_Y(t) = \prod_{i=1}^{m}(1-p+pe^t)^{n_i} = (1-p+pe^t)^{\sum_{i=1}^{m} n_i}$$

であり，これは，$Bin(\sum_{i=1}^{m} n_i, p)$ の *mgf* である．よって $Y \sim Bin(\sum_{i=1}^{m} n_i, p)$ となる．

二項分布は，起こりうる事象が二つであった．起こりうる事象を k 個に拡張したものは，**多項分布** *(multinomial distribution)* と呼ばれる．試行の結果，起こりうる事象を C_i，それぞれの事象が起こる確率，回数をそれぞれ，p_i，x_i とする．ただし，$i = 1, \ldots, k, n = \sum_{i=1}^k x_i, \sum_{i=1}^k p_i = 1$ である．多項分布の pmf は，

$$\left(\frac{n!}{x_1! \cdots x_{k-1}! x_k!}\right) p_1^{x_1} \cdots p_{k-1}^{x_{k-1}} p_k^{x_k}$$

である．

$k = 3$ の時，$X = X_1, Y = X_2, n - X - Y = X_3$ とすると，X, Y は**三項分布** *(trinomial distribution)* に従うという．X, Y の同時 pmf は，

$$p(x, y) = \frac{n!}{x!y!(n-x-y)!} p_1^x p_2^y p_3^{n-x-y}$$

となる．ただし，x, y は非負整数であり，$x + y \leq n$ である．また，p_1, p_2, p_3 は正数であり，$p_1 + p_2 + p_3 = 1$ を満たす．$p(x, y)$ は二つの確率変数 X, Y の pmf である条件を満たしている．(簡単なので，確かめてみてください．答えはテキストの $p.144$ にあります．)

最後に**超幾何分布** *(hypergeometric distribution)* を紹介する．この分布の pmf は，

$$p(x) = \frac{\binom{N-D}{n-x}\binom{D}{x}}{\binom{N}{n}}$$

である．平均は

$$\begin{aligned}
E[X] &= \sum_{x=0}^{n} x p(x) = \sum_{x=1}^{n} x \binom{D}{x} \binom{N-D}{n-x} \binom{N}{n}^{-1} \\
&= \sum_{x=1}^{n} D \binom{N-D}{n-x} \binom{D-1}{x-1} \binom{N}{n}^{-1} \\
&= D \sum_{x=1}^{n} \frac{n}{N} \binom{N-D}{n-x} \binom{D-1}{x-1} \binom{N-1}{n-1}^{-1} \\
&= \frac{nD}{N} \sum_{x=1}^{n} \binom{N-1+1-D}{n-1+1-x} \binom{D-1}{x-1} \binom{N-1}{n-1}^{-1} \\
&= \frac{nD}{N} \sum_{x=1}^{n} \binom{N-1-(D-1)}{n-1-(x-1)} \binom{D-1}{x-1} \binom{N-1}{n-1}^{-1} \\
&= \frac{nD}{N}
\end{aligned}$$

となる．5 番目の等号の右辺は HG(N-1,D-1,n-1) の全範囲における和が 1 となることを利用している．分散は

$$Var[X] = n \frac{D}{N} \frac{N-D}{N} \frac{N-n}{N-1}$$

となる．これは，小寺平治「明解数理統計」という本に書いてある導出を参考にしてください．
¡note¿

3.2 The Poisson Distribution

いきなりだが，

$$1 + m + \frac{m^2}{2!} + \frac{m^3}{3!} + \cdots = \sum_{x=0}^{\infty} \frac{m^x}{x!}$$

は e^m に収束するということを思い出すと，$m > 0$ で定義される関数 $p(x)$

$$p(x) = \begin{cases} e^{-m} \frac{m^x}{x!} & x = 0, 1, 2, \dots \\ 0 & elsewhere \end{cases} \quad tag3.2.1$$

は，$p(x) \geq 0$ であり，さらに，

$$\sum_x p(x) = e^{-m} \sum_{x=0}^{\infty} \frac{m^x}{x!} = 1$$

となるため，pmf であるための性質を満たす．この $p(x)$ を pmf として持つ確率変数は**ポアソン分布** *(poisson distribution)* に従うという．平均 m のポアソン分布は $Poi(m)$ として表す．

ある固定された期間にて，起こる変化の数が生成される過程に**ポアソン過程** *(Poisson process)* がある．以下にポアソン過程が満たすいくつかの仮定を列挙する．$g(x, w)$ を x 個の変化が長さ w の期間で起こる確率とする．$o(h)$ を $\lim_{h \to 0}[o(h)/h] = 0$ なる関数を表すものとする．ポアソンは以下を仮定する．

1. $g(1, h) = \lambda h + o(h), \lambda\ is\ a\ positive\ constant,\ h > 0$
2. $\sum_{x=2}^{\infty} g(x, h) = o(h)$
3. 非重複区間における変化の数は独立である．

仮定 $1, 3$ は短い期間の一つの変化は，他の期間の変化と独立であり，近似的に期間の長さの部分的なものとなるということを表している．仮定 2 の内容は，同一の短い区間 h 内に二つ以上の変化がある確率は 0 になるということを示している．もし $x = 0$ なら，$g(0, 0) = 1$ となる．仮定 $1, 2$

から，区間 h に少なくとも一つの変化がある確率は

$$g(1,h)+\sum_{x=2}^{\infty}g(x,h)=\lambda h+o(h)+o(h)=\lambda h+o(h)$$

となる．よって，長さ h の区間にて一つも変化が起こらない確率は，$1-\lambda h+o(h)$ となる．ゆえに，$g(0,w+h)$ の確率は，仮定 3 から，$g(0,w)$ と長さ h の区間にて一つも変化が起こらない確率の積，

$$g(0,w+h)=g(0,w)[1-\lambda h+o(h)]$$

となる．この時，

$$\frac{g(0,w+h)-g(0,w)}{h}=-\lambda g(0,w)-\frac{o(h)g(0,w)}{h}$$

であり，$h\to 0$ の極限をとると，

$$D_w[g(0,w)]=-\lambda g(0,w) \tag{3.2.2}$$

という微分方程式が得られる．この DE の解は，

$$g(0,w)=ce^{-\lambda w}$$

となる．$g(0,0)=1$ という条件から，$c=1$ であり，結局，

$$g(0,w)=e^{\lambda w}$$

を得る．もし，x が正整数ならば，$g(x,0)=0$ である．上記の仮定群は，

$$g(x,w+h)=[g(x,w)][1-\lambda h-o(h)]+[g(x-1,w)][\lambda h+o(h)]+o(h)$$

であることを示唆し，結果として，$x=1,2,3,\dots$ について

$$\frac{g(x,w+h)-g(x,w)}{h}=-\lambda g(x,w)+\lambda g(x-1,w)+\frac{o(h)}{h}$$

と

$$D_w[g(x,w)]=-\lambda g(x,w)+\lambda g(x-1,w)$$

という DE を得る．これらの DE の解は，$g(x,0)=0$ という境界条件から，それぞれ，

$$g(x,w)=e^{\lambda w}\frac{(\lambda w)^x}{x!},\ x=1,2,3,\dots$$

となる．よって，長さ w の区間に起こる変化の数 X は $Poi(\lambda w)$ に従う．

ポアソン分布の mgf，平均，分散は，それぞれ，

ポアソン分布の mgf，平均，分散

$$\begin{aligned}M(t)&=e^{m(e^t-1)}\\ \mu&=E[X]=m\\ \sigma^2&=V[X]=m\end{aligned}$$

となる．

証明 3.2. 簡単なので，導出してみましょう．

定理 3.2.1

$X_i \sim Poi(m_i)$ とし，それぞれの X_i は独立とする．この時，

$$Y = \sum_{i=1}^{n} X_i \sim Poi(\sum_{i=1}^{n} m_i)$$

となる．

証明 3.2. 二項分布の時と同じ流れでできるので，各自埋めてください．

¡note¿

3.6 t- and F- Distribution

この節では，統計的推測決定問題において頻繁に利用される，t 分布，F 分布を紹介する．

3.6.1 The t-ditribution

$W \sim N(0,1), V \sim \chi^2(r)$ とし，W, V は独立とする．W, V の同時分布は，それぞれの pdf の積であるため，

$$h(w,v) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} \frac{1}{\Gamma(r/2)2^{r/2}} v^{r/2-1} e^{-v/2} & -\infty < w < \infty, 0 < v < \infty \\ 0 & elsewhere \end{cases}$$

となる．ここで，新たに，

$$T = \frac{W}{\sqrt{V/r}}$$

なる確率変数 T を定義する．T の pdf $g_1(t)$ を得るために変数変換を行う．

$$\begin{cases} t = \frac{w}{\sqrt{v/r}} \\ u = v \end{cases}$$

これを w, v についてとくと，

$$\begin{cases} t = \frac{w}{\sqrt{v/r}} \\ u = v \end{cases}$$

であり，定義域は $\{(w,v) : -\infty < w < \infty, 0 < v < \infty\}$ から，$\{(t,u) : -\infty < t < \infty, 0 < u < \infty\}$ に移される．この変換におけるヤコビアン J は

$$\begin{aligned} |J| &= \left| \frac{\partial(w,v)}{\partial(t,u)} \right| \\ &= \begin{vmatrix} \frac{\partial w}{\partial t} & \frac{\partial w}{\partial u} \\ \frac{\partial v}{\partial t} & \frac{\partial v}{\partial u} \end{vmatrix} \\ &= \begin{vmatrix} \sqrt{\frac{u}{r}} & \frac{1}{2}\frac{t}{\sqrt{ur}} \\ 0 & 1 \end{vmatrix} \\ &= \left| \sqrt{\frac{u}{r}} \right| = \sqrt{\frac{u}{r}} \\ & (\because u > 0) \end{aligned}$$

となる．よって，$(w,v)\to(t,u)$ への変換は，

$$
\begin{aligned}
g(t,u) &= h(t\sqrt{\frac{u}{r}},u)|J| \\
&= \begin{cases} \frac{1}{\sqrt{2\pi}\Gamma(r/2)2^{r/2}}u^{r/2-1}\exp\left[-\frac{u}{2}\left(1+\frac{t^2}{r}\right)\right]\sqrt{\frac{u}{r}} & |t|<\infty, 0<u<\infty \\ 0 & elsewhere \end{cases}
\end{aligned}
$$

となる．ここで，t について周辺化することで，T の pdf が得られるので，

$$
\begin{aligned}
g_1(t) &= \int_{-\infty}^{\infty} g(t,u)du \\
&= \int_0^{\infty} \frac{1}{\sqrt{2r\pi}\Gamma(r/2)2^{r/2}}u^{(r+1)/2-1}\exp\left[-\frac{u}{2}\left(1+\frac{t^2}{r}\right)\right]du
\end{aligned}
$$

ここで，$z=u[1+(t^2/r)]/2$ と変数変換すると，$u=2z/[1+(t^2/r)]$ であり，$du=2/[1+(t^2/r)]dz$，$u:0\to\infty$ の時，$z:0\to\infty$ となるから，

$$
\begin{aligned}
g_1(t) &= \int_0^{\infty}\frac{1}{\sqrt{2r\pi}\Gamma(r/2)2^{r/2}}\left(\frac{2z}{1+t^2/r}\right)^{(r+1)/2-1}e^{-z}\left(\frac{2}{1+t^2/r}\right)dz \\
&= \frac{1}{\sqrt{r\pi}\Gamma(r/2)\int_0^{\infty}2^{-(r+1)/2}}2^{(r+1)/2}z^{(r+1)/2-1}(1+t^2/r)^{-(r+1)/2}e^{-z}dz \\
&= \frac{1}{\sqrt{r\pi}\Gamma(r/2)(1+t^2/r)^{(r+1)/2}}\int_0^{\infty}\Gamma[(r+1)/2]\frac{1}{\Gamma[(r+1)/2]}z^{(r+1)/2-1}e^{-z}dz \\
&= \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\Gamma[r/2]}\frac{1}{(1+t^2/r)^{(r+1)/2}} \qquad (3.6.1)
\end{aligned}
$$

となる．ここで，三つ目の等号右辺から，4 つ目の等号の際に $\chi^2(r+1)$ の pdf の全区間での積分が 1 となることを利用した．よって，もし，$W\sim N(0,1), V\sim\chi(r)$ で W,V が独立なら，

$$
T=\frac{W}{\sqrt{V/r}} \qquad (3.6.2)
$$

は $g_1(t)$ に従うことがわかった．この $g_1(t)$ を pdf として持つ確率変数は，自由度 r の t **分布** *(t-distribution)* に従うといい，$T\sim t_r$ と表す．(3.6.1) にて，$\sqrt{\pi}=\Gamma(1/2), \Gamma((r+1)/2)=\Gamma(r/2+1/2), B(a,b)=\Gamma(a)\Gamma(b)/\Gamma(a+b)$ であることを利用すると，

$$
g_1(t)=\frac{1}{\sqrt{r}B(r/2,1/2)}\left(1+\frac{t^2}{r}\right)^{-(r+1)/2}
$$

とベータ関数 $B(r/2,1/2)$ を用いて表すことができる (こっちの方がよく見る気がします)．
¡note¿

3.6.2 The F-distribution

$U \sim \chi^2(r_1), V \sim \chi^2(r_2)$ とし，U, V は独立とする．同時 pdf は，

$$h(w, v) = \begin{cases} \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} u^{r_1/2-1} v^{r_2/2-1} e^{-(u+v)/2} & 0 < u, v < \infty \\ 0 & elsewhere \end{cases}$$

となる．新たに，

$$W = \frac{U/r_1}{V/r_2}$$

という確率変数を定義する．W の pdf を求めるために，以下の変数変換を行う．

$$\begin{cases} w = \frac{u/r_1}{v/r_2} \\ z = v \end{cases}$$

これを u, v についてとくと，

$$\begin{cases} u = \frac{r_1}{r_2} zw \\ v = z \end{cases}$$

であり，定義域は $\{(u, v) : 0 < u < \infty, 0 < v < \infty\}$ から，$\{(w, z) : 0 < w < \infty, 0 < z < \infty\}$ に移される．この変換に対応するヤコビアンは，

$$\begin{aligned} |J| &= \left| \frac{\partial(u, v)}{\partial(w, z)} \right| \\ &= \begin{vmatrix} \frac{\partial u}{\partial w} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial w} & \frac{\partial v}{\partial z} \end{vmatrix} \\ &= \begin{vmatrix} \frac{r_1}{r_2} z & \frac{r_1}{r_2} w \\ 0 & 1 \end{vmatrix} \\ &= \left| \frac{r_1}{r_2} z \right| = \frac{r_1}{r_2} z \\ & (\because z, r_1, r_2 > 0) \end{aligned}$$

となる．よって W, Z の同時 pdf は，台にて

$$\begin{aligned} g(w, z) &= h(\frac{r_1}{r_2} zw, z)|J| \\ &= \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \left(\frac{r_1}{r_2} zw \right)^{(r_1-2)/2} z^{(r_2-2)/2} \exp\left[-\frac{z}{2}(\frac{r_1}{r_2} w + 1) \right] \frac{r_1}{r_2} z \end{aligned}$$

となる．W の周辺密度は，

$$
\begin{aligned}
g_1(w) &= \int_{-\infty}^{\infty} g(w,z)dz \\
&= \int_0^{\infty} \frac{(r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} z^{(r_1+r_2)/2-1} \exp\left[-\frac{z}{2}(\frac{r_1}{r_2}w+1)\right] dz
\end{aligned}
$$

ここで，

$$
y = \frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right)
$$

と変換するならば，$z = 2y/(r_1 w/r_2+1)$ であり，$dz = 2/(r_1 w/r_2+1)dy$ である．また，$z : 0 \to \infty$ の時，$y : 0 \to \infty$ である．よって，

$$
\begin{aligned}
g_1(w) &= \int_0^{\infty} \frac{(r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \left(\frac{2y}{r_1 w/r_2+1}\right)^{(r_1+r_2)/2-1} e^{-y} \frac{2}{r_1 w/r_2+1} dy \\
&= \frac{(r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \int_0^{\infty} \frac{2^{(r_1+r_2)/2}}{(1+r_1 w/r_2)^{(r_1+r_2)/2}} \Gamma[(r_1+r_2)/2] \frac{1}{\Gamma[(r_1+r_2)/2]} y^{(r_1+r_2)/2-1} e^{-y} dy \\
&= \begin{cases} \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \frac{w^{r_1/2-1}}{(1+r_1 w/r_2)^{(r_1+r_2)/2}}, & 0 < w < \infty \\ 0 & elsewhere \end{cases}
\end{aligned}
$$

を得る．なお，先ほどの t 分布の場合と同様に，三つ目の等号の際に，$\chi^2(r_1+r_2)$ の全範囲での積分が 1 であることを利用した．以上から，$U \sim \chi^2(r_1), V \sim \chi^2(r_2)$ で U, V が独立の時，

$$
W = \frac{U/r_1}{V/r_2}
$$

は $g_1(w)$ を pdf として持つことがわかった．この pdf を持つ確率変数は F **分布** *(F-distribution)* に従うといい，$F(r_1, r_2)$ や $F_{r_1}^{r_2}$ のように表す．
¡note¿

3.6.3 Student's Theorem

定理 3.6.1

$X_i, i = 1, 2, \ldots, n$ を $N(\mu, \sigma^2)$ からの iid 標本とする．

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,\ S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

と確率変数を定義する．この時，

1. $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$
2. $\bar{X}, S^2$ は独立である．
3. $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$
4. 確率変数

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

は自由度 $n-1$ の t 分布に従う．

証明 3.6. $\mathbf{X} = (X_1, X_2, \ldots, X_n)'$ と定義すると，$\mathbf{X} \sim N(\mu\mathbf{1}, \sigma^2\mathbf{I})$ である．また $\mathbf{v} = (1/n, \ldots, 1/n)$ とすると，$\bar{X} = \mathbf{v}'\mathbf{X}$ である．さらに $\mathbf{Y}$ を $\mathbf{Y} = (X_1 - \bar{X}, \ldots, X_n - \bar{X})$ と定義する．まず，以下のような変換を考える．

$$\mathbf{W} = \begin{bmatrix} \bar{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{v}' \\ \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix} \mathbf{X} \tag{3.6.10}$$

$\mathbf{W}$ は多変量正規分布に従う確率変数ベクトルの線型変換であるため，これも平均

$$E[\mathbf{W}] = \begin{bmatrix} \mathbf{v}' \\ \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix} E[\mathbf{X}] = \begin{bmatrix} \mathbf{v}' \\ \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix} \mu\mathbf{1} = \begin{bmatrix} \mu \\ \mathbf{0}_n \end{bmatrix} \tag{3.6.11}$$

で，共分散行列が，

$$\begin{aligned} \mathbf{\Sigma} &= \begin{bmatrix} \mathbf{v}' \\ \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix} \sigma^2\mathbf{I} \begin{bmatrix} \mathbf{v}' \\ \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix}' \\ &= \sigma^2 \begin{bmatrix} \frac{1}{n} & \mathbf{0}' \\ \mathbf{0} & \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix} \end{aligned} \tag{3.6.12}$$

の多変量正規分布に従う．$\bar{X}$ は $\mathbf{W}$ の第 *1* 成分であるため，定理の 1 の結果を得る．次に，$\bar{X}$ と $\mathbf{Y}$ の共分散は 0 であるため，$\bar{X}, \mathbf{Y}$ は独立である．ここで，$S^2 = (n-1)^{-1}\mathbf{Y}'\mathbf{Y}$ であることに注意すると，$\bar{X}, S^2$ が独立だということも言える．よって定理の 2 がいえた．

確率変数

$$V = \sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2$$

を考える．$Z_i = (X_i - \mu)/\sigma \sim N(0,1)$ であり，この二乗は $\chi^2(1)$ に従う．また，Z_i は互いに独立であるため，$V = \sum_{i=1}^{n} Z_i \sim \chi^2(n)$ である．次の変換を考える．

$$\begin{aligned}
V &= \sum_{i=1}^{n} \left(\frac{(X_i - \bar{X}) + (\mu - \bar{X})}{\sigma} \right)^2 \\
&= \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 + n \left(\frac{\mu - \bar{X}}{\sigma} \right)^2 \\
&= \frac{(n-1)S^2}{\sigma^2} + \left(\frac{\mu - \bar{X}}{\sigma/\sqrt{n}} \right)^2 \qquad (3.6.13)
\end{aligned}$$

定理の 2 を利用すると，(3.6.13) の第 *1* 項と第 *2* 項は独立である．さらに第 *2* 項は標準正規分布に従う確率変数であり，その二乗は $\chi^2(1)$ に従う．両辺の *mgf* を考えると，

$$\begin{aligned}
&(1-2t)^{-n/2} = E[\exp\{t(n-1)S^2/\sigma^2\}](1-2t)^{-1/2} \qquad (3.6.14) \\
&\Leftrightarrow E[\exp\{t(n-1)S^2/\sigma^2\}] = (1-2t)^{-(n-1)/2}
\end{aligned}$$

であり，$(n-1)S^2/\sigma^2$ の *mgf* は $\chi^2(n-1)$ の *mgf* と同一である．*pdf* と *mgf* の一対一対応から，$(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$ である．よって定理の 3 を得る．最後に，定理の 4 であるが，

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{(\bar{X} - \mu)/(\sigma/\sqrt{n})}{\sqrt{(n-1)S^2/(\sigma^2(n-1))}}$$

となることから，t 分布に従う確率変数の成り立ちを思い出すと，これが t 分布に従うことがわかる．よって定理の結果を得る．

¡note¿