参考

$V_1, V_2, \cdots, V_n$ が，標準正規分布 $N(0, 1)$ に従うとき，$V_1^2+V_2^2+\cdots+V_n^2$ は自由度 n の χ^2 分布に従うんだったね。(P125)

よって，$\displaystyle\sum_{i=1}^{n}\left(\boxed{\frac{X_i-\mu}{\sigma}}\right)^2$ は，自由度 n の χ^2 分布に従う。

これは標準正規分布 $N(0, 1)$ に従う

ここで，この μ の代わりに $\overline{X}$ を代用すると，

$\displaystyle\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2$ は，自由度 $(n-1)$ の χ^2 分布に従う。何故だかわかる？

このとき，$\dfrac{X_1-\overline{X}}{\sigma}+\dfrac{X_2-\overline{X}}{\sigma}+\cdots+\dfrac{X_n-\overline{X}}{\sigma}=0$ が成り立つので，

$\dfrac{X_i-\overline{X}}{\sigma}$ $(i=1, 2, \cdots, n)$ は線形従属となって，μ のときに比べて $\overline{X}$ では，自由度が 1 つ減るからだ。

以上②, ③より，①の U は

$$U=\frac{\boxed{Y}}{\sqrt{\dfrac{\boxed{Z}}{n-1}}}$$

$N(0, 1)$ に従う。

自由度 $(n-1)$ の χ^2 分布に従う。

となって，

自由度 $(n-1)$ の t 分布に従う。

これで，話がすべて明らかになっただろう？

t 分布の定義 (P132):
Y と Z が独立で，
Y は $N(0, 1)$ に従い，
Z が自由度 n の χ^2 分布に従うとき，
$X=\dfrac{Y}{\sqrt{\dfrac{Z}{n}}}$ は，自由度 n の t 分布に従う。
$t_n(x)=K_n\cdot\left(\dfrac{x^2}{n}+1\right)^{-\frac{n+1}{2}}$

後は，前に話した計算手順に従って，t 分布の数表も利用して，実際に母平均 μ の信頼区間を算出してみればいいんだね。後に例題と演習問題で練習しよう。

● 母比率の区間推定は正規分布を利用する！

例えば，ある県の全世帯の自動車の保有率や，ある国の全有権者の X 政党への支持率のように，ある性質をもつものの全体に対する割合を，母集団の場合は**母比率**（ぼひりつ）と呼び p で，また，大きさ n の標本の場合は**標本比率**（ひょうほんひりつ）と呼び $\overline{p}$ で表すことにする。そして，n と $\overline{p}$ を用いて，母比率 p の区間推定

を行うことができる。

車の保有率や政党支持率など，ある性質 A に対して母比率 p をもつ非常に大きな母集団から，十分な大きさ n の標本を無作為に抽出する場合，1 つ 1 つ

これは，母集団の大きさ N が非常に大きいものとして，非復元でも復元と考えていい。

の標本を n 回抽出すると考えると，これは事象 A が n 回中 X 回起こる反復試行の確率 $P_B(x)$ を求めることと同様なことに気付くはずだ。

つまり，1 回の試行 (抽出) で事象 A の起こる確率が母比率 p，起こらない確率が $q\,(=1-p)$ であり，n 回中 X 回だけ事象 A の起こる確率と同様に，n 個の標本中 X 個だけ A の性質をもつ確率 $P_B(x)$ は，

$P_B(x)={}_n\mathrm{C}_x\,p^x q^{n-x}\quad(x=0,\ 1,\ 2,\ \cdots,\ n)$ となるんだね。(P20, P86 参照)

よって，確率変数 X は二項分布 $B(n,\ p)$ に従い，さらに n が十分に大きければ，

この平均は np，分散は $npq=np(1-p)$ だね。

これは近似的に平均 np，分散 $np(1-p)$ の正規分布 $N(np,\ np(1-p))$ に従うことになるんだね。さらに，n が十分に大きいときは，分散 $np(1-p)$ の p を近似的に標本比率 $\overline{p}$ でおきかえてもよいことが大数の法則 (P110) から導ける。よって，X は正規分布 $N(np,\ n\overline{p}(1-\overline{p}))$ に従うと言える。よって，標準化変数 $Z\left(=\dfrac{X-np}{\sqrt{n\overline{p}(1-\overline{p})}}\right)$ は標準正規分布 $N(0,\ 1)$ に従うので，Z の信頼係数 $1-\alpha$ の信頼区間を $-z\left(\dfrac{\alpha}{2}\right)\leqq Z\leqq z\left(\dfrac{\alpha}{2}\right)$ とおくと，

$P\left(-z\left(\dfrac{\alpha}{2}\right)\leqq Z\leqq z\left(\dfrac{\alpha}{2}\right)\right)=1-\alpha$ となる。したがって，

$-z\left(\dfrac{\alpha}{2}\right)\leqq\dfrac{X-np}{\sqrt{n\overline{p}(1-\overline{p})}}\leqq z\left(\dfrac{\alpha}{2}\right)$，$-z\left(\dfrac{\alpha}{2}\right)\sqrt{n\overline{p}(1-\overline{p})}\leqq X-np\leqq z\left(\dfrac{\alpha}{2}\right)\sqrt{n\overline{p}(1-\overline{p})}$

$np\leqq X+z\left(\dfrac{\alpha}{2}\right)\sqrt{n\overline{p}(1-\overline{p})}$　　$X-z\left(\dfrac{\alpha}{2}\right)\sqrt{n\overline{p}(1-\overline{p})}\leqq np$

$X-z\left(\dfrac{\alpha}{2}\right)\cdot\sqrt{n\overline{p}(1-\overline{p})}\leqq np\leqq X+z\left(\dfrac{\alpha}{2}\right)\cdot\sqrt{n\overline{p}(1-\overline{p})}$　この両辺を n で割ると，

n が十分に大きいとき，$\dfrac{X}{n}=\overline{p}$ とおけるので，p の信頼係数 $1-\alpha$ の信頼区間は，

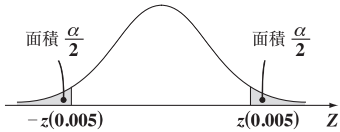
$$\overline{p}-z\left(\frac{\alpha}{2}\right)\cdot\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}\leqq p\leqq\overline{p}+z\left(\frac{\alpha}{2}\right)\cdot\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$$ となるんだね。

それでは，ここで，例題を解いておこう。

(2) 非常に大きな数 N の母集団があり，この母集団の X 政党への支持率（母比率）が p であるものとする。この母集団から標本数 $n=1600$ の標本を無作為に抽出して，X 政党への支持率を調べた結果，$\overline{p}=0.2$ $(=20\%)$ であった。このとき，母集団の X 政党への支持率 p の 99% 信頼区間を小数第 3 位まで求めよ。

(2) 母比率 p の 99% 信頼区間より，

$1-\alpha=0.99$，すなわち有意水準 $\alpha=0.01$ となる。また，標本の大きさ $n=1600$，標本比率 $\overline{p}=0.2$ より，



P223 の標準正規分布表から，

$z\left(\frac{\alpha}{2}\right)=z(0.005)$ の値を求めると，

$z(0.005)=2.58$ となる。

標準正規分布表

z	……	0.08
⋮		⋮
2.5	……	0.00494

以上より，母比率 p の 99% 信頼区間は，

$$0.2-2.58\cdot\sqrt{\frac{0.2(1-0.2)}{1600}}\leqq p\leqq 0.2+2.58\cdot\sqrt{\frac{0.2(1-0.2)}{1600}}$$ より，

$$\left[\ \overline{p}-z\left(\frac{\alpha}{2}\right)\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}\leqq p\leqq \overline{p}+z\left(\frac{\alpha}{2}\right)\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}\ \right]$$

$$\sqrt{\frac{0.2(1-0.2)}{1600}}=\sqrt{\frac{0.2\times0.8}{1600}}=\sqrt{\frac{0.16}{1600}}=\sqrt{\frac{1}{10^4}}=\frac{1}{100}$$

$$\underbrace{0.2-0.0258}_{0.1742}\leqq p\leqq\underbrace{0.2+0.0258}_{0.2258}$$

$\therefore\ 0.174\leqq p\leqq 0.226$ となる。……………………………………………(答)

● 母分散 σ^2 の区間推定は χ^2 分布が決め手だ！

母分散 σ^2 の区間推定では，標本 $X_1, X_2, \cdots, X_n$ から，新たな確率変数 V を「$V=\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2$ と定義する。これが，自由度 $(n-1)$ の χ^2 分布に従う」ことから σ^2 の区間推定が可能になる。

また，同じことが出てきたね！