P.94 の定理 6.3 から，

$$\frac{\sum_{i=1}^{n_1}(X_{1i}-\overline{X}_1)^2}{\sigma_1^2}=\frac{(n_1-1)S_1^2}{\sigma_1^2}\sim\chi^2(n_1-1)$$

ただし，$\overline{X}_1=\dfrac{1}{n_1}\displaystyle\sum_{i=1}^{n_1}X_{1j}$，$S_1^2=\dfrac{1}{n_1-1}\displaystyle\sum_{i=1}^{n_1}(X_{1i}-\overline{X}_1)^2$

$$\frac{\sum_{j=1}^{n_2}(X_{2j}-\overline{X}_2)^2}{\sigma_2^2}=\frac{(n_2-1)S_2^2}{\sigma_2^2}\sim\chi^2(n_2-1)$$

ただし，$\overline{X}_2=\dfrac{1}{n_2}\displaystyle\sum_{j=1}^{n_2}X_{2j}$，$S_2^2=\dfrac{1}{n_2-1}\displaystyle\sum_{j=1}^{n_2}(X_{2j}-\overline{X}_2)^2$

定理 6.6 から，

$$\frac{\dfrac{(n_1-1)S_1^2}{\sigma_1^2}\Big/(n_1-1)}{\dfrac{(n_2-1)S_2^2}{\sigma_2^2}\Big/(n_2-1)} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$$

となる。

帰無仮説： $H_0: \ \sigma_1^2 = \sigma_2^2$ 　**（2 つの分散が等しい，または，等分散）**

対立仮説： $H_1: \ \sigma_1^2 \neq \sigma_2^2$ 　**（2 つの分散が異なる，または，不等分散）**

帰無仮説 H_0 **のもとで（帰無仮説が正しいとき），**

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2}{S_2^2} \sim F(n_1-1, n_2-1)$$

となる（P.99，定理 6.7）。

S_1^2，S_2^2 をそれぞれデータ s_1^2，s_2^2 で置き換えて，$\dfrac{s_1^2}{s_2^2}$ と $F(n_1-1, n_2-1)$ の比較

$\dfrac{s_1^2}{s_2^2} > F_{\alpha/2}(n_1-1, n_2-1)$，または，$\dfrac{s_1^2}{s_2^2} < \dfrac{1}{F_{\alpha/2}(n_2-1, n_1-1)}$ のとき，有意水準 α で H_0 を棄却する。

8.8　母比率の検定 (P.153)

n 回の試行 (実験)

i 回目の試行で，

成功のとき $X_i = 1$，失敗のとき $X_i = 0$ とする。

$$R = \sum_{i=1}^{n} X_i$$

とする。

n 回の実験で成功する回数 $= R$

1 回の試行で成功する確率 $= p \implies P(X_i = 1) = p$

X_i の平均・分散を求める。

$$\mathbf{E}(X_i) = 1 \times P(X_i = 1) + 0 \times P(X_i = 0) = p$$

$$\begin{aligned}\mathbf{V}(X_i) &= (1-p)^2 \times P(X_i = 1) \\ &\qquad +(0-p)^2 \times P(X_i = 0) \\ &= p(1-p)\end{aligned}$$

離散型確率変数の平均・分散の求め方は，P.53，54 の (4.12)，(4.14) を参考に。

一方，

比率 p の推定量 $\overline{X}$ を，

$$\overline{X} = \frac{R}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad (= \hat{p})$$

とする。テキストでは，$\overline{X}$ を $\hat{p}$ としている。

すなわち，

$$\mathbf{E}(X_1) = \mathbf{E}(X_2) = \cdots = \mathbf{E}(X_n) = p$$

$$\mathbf{V}(X_1) = \mathbf{V}(X_2) = \cdots = \mathbf{V}(X_n) = p(1-p)$$

$\overline{X}$ は n 個の確率変数 $X_1, X_2, \cdots, X_n$ の標本平均である。

$\Longrightarrow$ 中心極限定理 (定理 6.1, P.90) の適用

$\mathbf{E}(\overline{X}) = p$, $\mathbf{V}(\overline{X}) = \dfrac{p(1-p)}{n}$ なので，n が大きいとき，

$$Z = \frac{\overline{X} - \mathbf{E}(\overline{X})}{\sqrt{\mathbf{V}(\overline{X})}} = \frac{\overline{X} - p}{\sqrt{p(1-p)/n}} \sim N(0, 1)$$

として近似できる。

● ケース 1 (片側検定)

帰無仮説　$H_0 : p = p_0$

対立仮説　$H_1 : p < p_0$

帰無仮説 H_0 が正しいもとで，近似的に，

$$\frac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1)$$

なので，

$$P\left(\frac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)/n}} < -z_\alpha\right) = \alpha$$

となる。

このとき，$\dfrac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)/n}}$ を検定統計量と呼ぶ。

もし $\dfrac{\overline{x} - p_0}{\sqrt{p_0(1-p_0)/n}} < -z_\alpha$ ならば，帰無仮説 $H_0 : p = p_0$ が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ **有意水準 α で，母比率 p は p_0 よりも小さいと判断する。**

テキストでは，

$$\frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} < -z_\alpha$$ **のとき，有意水準 α で，母比率 p は p_0 よりも小さいと判断する。**

● ケース 2 (片側検定)

帰無仮説　$H_0 : p = p_0$

対立仮説　$H_1 : p > p_0$

帰無仮説 H_0 が正しいもとで，近似的に，

$$\frac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1)$$

なので，

$$P(\frac{\overline{X} - p_0}{\sqrt{p_0(1 - p_0)/n}} > z_\alpha) = \alpha$$

となる。

このとき，$\dfrac{\overline{X} - p_0}{\sqrt{p_0(1 - p_0)/n}}$ を検定統計量と呼ぶ。

もし $\dfrac{\overline{x} - p_0}{\sqrt{p_0(1 - p_0)/n}} > z_\alpha$ ならば，帰無仮説 $H_0 : p = p_0$ が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ 有意水準 α で，母比率 p は p_0 よりも大きいと判断する。

テキストでは，

$$\frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} > z_\alpha$$ **のとき，有意水準 α で，母比率 p は p_0 よりも大きいと判断する。**

● ケース 3 (両側検定)

帰無仮説　$H_0 : p = p_0$

対立仮説　$H_1 : p \neq p_0$

帰無仮説 H_0 が正しいもとで，近似的に，

$$\frac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0, 1)$$

なので，

$$P\left(\left|\frac{\overline{X}-p_0}{\sqrt{p_0(1-p_0)/n}}\right| > z_{\alpha/2}\right) = \alpha$$

となる。

このとき，$\dfrac{\overline{X}-p_0}{\sqrt{p_0(1-p_0)/n}}$ を検定統計量と呼ぶ。

もし

$$\frac{\overline{x}-p_0}{\sqrt{p_0(1-p_0)/n}} < -z_{\alpha/2} \text{ または } \frac{\overline{x}-p_0}{\sqrt{p_0(1-p_0)/n}} > z_{\alpha/2}$$

ならば，帰無仮説 $H_0 : p = p_0$ が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ 有意水準 α で，母比率 p と p_0 は異なると判断する。

テキストでは，

$\dfrac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} < -z_{\alpha/2}$ または $\dfrac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}} > z_{\alpha/2}$ のとき，有意水準 α で，母比率 p は p_0 と異なると判断する。

例題：　我が国の総人口に対する 65 歳以上の高齢者の占める割合は 1987 年に 10.9 % であった。気候が平均寿命に影響するかどうかを調べるために，気候の温暖な四国地方から無作為に 1350 人を選び，65 歳以上の人口数を調べたところ，170 人であった。四国地方の高齢者比率は全国より高いといえるか？

解答： n **が大きいとき，**

$$Z = \frac{\overline{X} - p}{\sqrt{p(1-p)/n}} \sim N(0,1)$$

として近似できる。

帰無仮説 $H_0 : p = p_0$

対立仮説 $H_1 : p > p_0$

帰無仮説 H_0 **が正しいもとで，近似的に，**

$$\frac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1)$$

なので，

$$P(\frac{\overline{X} - p_0}{\sqrt{p_0(1 - p_0)/n}} > z_\alpha) = \alpha$$

となる。

このとき，$\dfrac{\overline{X} - p_0}{\sqrt{p_0(1 - p_0)/n}}$ を検定統計量と呼ぶ。

もし $\dfrac{\overline{x} - p_0}{\sqrt{p_0(1 - p_0)/n}} > z_\alpha$ ならば，帰無仮説 $H_0 : p = p_0$ が起こる確率は低いということになり，有意水準 α で H_0 を棄却する (H_1 を採択する)。

$\Longrightarrow$ 有意水準 α で，母比率 p は p_0 よりも大きいと判断する。

$p_0 = 0.109$

$\alpha = 0.05$ **のとき，** $z_\alpha = 1.645$

$$n = 1350,\ \overline{x} = \frac{170}{1350} = 0.126$$

$$\frac{\overline{x} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.126 - 0.109}{\sqrt{0.109(1-0.109)/1350}} = 2.00 > z_\alpha = 1.645$$

有意水準 0.05 で， H_0 **を棄却する。四国地方の高齢者比率は全国平均より高いと言える。**

(*) n **が大きいとき，**

$$Z = \frac{\overline{X} - p}{\sqrt{\overline{x}(1-\overline{x})/n}} \sim N(0,1)$$

として近似することもできる（分母の p **には，予めその推定値を代入する）。**

その場合，上の例題では，$\frac{\overline{x} - p_0}{\sqrt{\overline{x}(1-\overline{x})/n}} = \frac{0.126 - 0.109}{\sqrt{0.126(1-0.126)/1350}} = 1.88$ となり，結果的には同じ結論となる。