第 20 回　区間推定（11.5）

村澤 康友

2022 年 12 月 13 日

今日のポイント

1. 母数を含む領域を定める推定を区間推定という．ある確率で母数を含む確率的な領域を，その母数の信頼域という．信頼域が母数を含む確率を，その信頼域の信頼係数という．
2. 正規母集団の母平均・母分散の信頼区間は標本平均・標本分散の標本分布から求まる．2 標本問題も同様.
3. ベルヌーイ母集団の母比率（＝母平均）の信頼区間は標本比率（＝標本平均）の漸近分布から求まる.

目次

1　信頼域（p. 225）

定義 1. 母数を含む領域を定める推定を**区間推定**という.

定義 2. ある確率で母数を含む確率的な領域を，その母数の**信頼域**という.

定義 3. 1 次元の信頼域を**信頼区間**という.

定義 4. 信頼域が母数を含む確率を，その信頼域の**信頼係数**という.

注 1. 「信頼係数 α の信頼域」を「100α %信頼域」と略すことが多い.

2　正規母集団

2.1　母平均の信頼区間（p. 226）

2.1.1　母分散が既知

母集団分布を $\mathrm{N}\left(\mu,\sigma^2\right)$ とする．μ の 95 %信頼区間を求める．ただし σ^2 は既知とする．大きさ n の無作為標本の標本平均を $\bar{X}$ とすると

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

標準化すると

$$\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} \sim \mathrm{N}(0,1)$$

標準正規分布表より

$$\Pr\left[-1.96 \le \frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} \le 1.96\right] = .95$$

ここで

$$
\begin{aligned}
&-1.96 \le \frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} \le 1.96 \\
&\iff -1.96\sqrt{\frac{\sigma^2}{n}} \le \bar{X}-\mu \le 1.96\sqrt{\frac{\sigma^2}{n}} \\
&\iff \bar{X}-1.96\sqrt{\frac{\sigma^2}{n}} \le \mu \le \bar{X}+1.96\sqrt{\frac{\sigma^2}{n}}
\end{aligned}
$$

したがって μ の 95 %信頼区間は

$$\left[\bar{X}-1.96\sqrt{\frac{\sigma^2}{n}}, \bar{X}+1.96\sqrt{\frac{\sigma^2}{n}}\right]$$

2.1.2　母分散が未知

標本分散を s^2 とする．σ^2 を s^2 に置き換えると

$$\frac{\bar{X}-\mu}{\sqrt{s^2/n}} \sim \mathrm{t}(n-1)$$

t 分布表より，例えば $n=10$ なら

$$\Pr\left[-2.262 \le \frac{\bar{X}-\mu}{\sqrt{s^2/10}} \le 2.262\right] = .95$$

ここで

$$
\begin{aligned}
&-2.262 \le \frac{\bar{X}-\mu}{\sqrt{s^2/10}} \le 2.262 \\
&\iff -2.262\sqrt{\frac{s^2}{10}} \le \bar{X}-\mu \le 2.262\sqrt{\frac{s^2}{10}} \\
&\iff \bar{X}-2.262\sqrt{\frac{s^2}{10}} \le \mu \le \bar{X}+2.262\sqrt{\frac{s^2}{10}}
\end{aligned}
$$

したがって $n=10$ なら μ の 95 %信頼区間は

$$\left[\bar{X}-2.262\sqrt{\frac{s^2}{10}}, \bar{X}+2.262\sqrt{\frac{s^2}{10}}\right]$$

例 1 (p. 226)**.** $n=100$, $\bar{X}=2.346$, $s=.047$ のとき，μ の 90 %信頼区間を求める．正規母集団なら

$$\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim \mathrm{t}(n-1)$$

t 分布表より，$n=100$ なら

$$\Pr\left[-1.661 \le \frac{\bar{X}-\mu}{s/10} \le 1.661\right] = .9$$

ここで

$$
\begin{aligned}
&-1.661 \le \frac{\bar{X}-\mu}{s/10} \le 1.661 \\
&\iff -.1661s \le \bar{X}-\mu \le .1661s \\
&\iff \bar{X}-.1661s \le \mu \le \bar{X}+.1661s
\end{aligned}
$$

したがって μ の 90 %信頼区間は

$$[2.346-.1661\cdot .047, 2.346+.1661\cdot .047]$$

すなわち $[2.338, 2.354]$.

2.2　母分散の信頼区間（p. 226）

2.2.1　母平均が既知

σ^2 の 95 %信頼区間を求める．ただし μ は既知とする．標本分散は

$$\hat{\sigma}^2 := \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$$

このとき

$$\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n)$$

χ^2 分布表より，例えば $n=10$ なら

$$\Pr\left[3.24697 \le \frac{10\hat{\sigma}^2}{\sigma^2} \le 20.4832\right] = .95$$

ここで

$$
\begin{aligned}
&3.24697 \le \frac{10\hat{\sigma}^2}{\sigma^2} \le 20.4832 \\
&\iff \frac{1}{20.4832} \le \frac{\sigma^2}{10\hat{\sigma}^2} \le \frac{1}{3.24697} \\
&\iff \frac{10\hat{\sigma}^2}{20.4832} \le \sigma^2 \le \frac{10\hat{\sigma}^2}{3.24697}
\end{aligned}
$$

したがって $n=10$ なら σ^2 の 95 %信頼区間は

$$\left[\frac{10}{20.4832}\hat{\sigma}^2, \frac{10}{3.24697}\hat{\sigma}^2\right]$$

2.2.2　母平均が未知

μ が未知なら標本分散は

$$s^2 := \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2$$

このとき

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

χ^2 分布表より，例えば $n = 10$ なら

$$\Pr\left[2.70039 \le \frac{9s^2}{\sigma^2} \le 19.0228\right] = .95$$

ここで

$$\begin{aligned}
&2.70039 \le \frac{9s^2}{\sigma^2} \le 19.0228 \\
\iff &\frac{1}{19.0228} \le \frac{\sigma^2}{9s^2} \le \frac{1}{2.70039} \\
\iff &\frac{9s^2}{19.0228} \le \sigma^2 \le \frac{9s^2}{2.70039}
\end{aligned}$$

したがって $n = 10$ なら σ^2 の 95 %信頼区間は

$$\left[\frac{9}{19.0228}s^2, \frac{9}{2.70039}s^2\right]$$

2.3　母平均の差の信頼区間（p. 227）

2.3.1　母分散が既知

母集団分布を $\mathrm{N}\left(\mu_X, \sigma_X^2\right)$，$\mathrm{N}\left(\mu_Y, \sigma_Y^2\right)$ とする．$\mu_X - \mu_Y$ の 95 %信頼区間を求める．ただし σ_X^2, σ_Y^2 は既知とする．各母集団から独立に抽出した大きさ m, n の無作為標本の標本平均を $\bar{X}, \bar{Y}$ とすると

$$\begin{aligned}
\bar{X} &\sim \mathrm{N}\left(\mu_X, \frac{\sigma_X^2}{m}\right) \\
\bar{Y} &\sim \mathrm{N}\left(\mu_Y, \frac{\sigma_Y^2}{n}\right)
\end{aligned}$$

両者は独立だから

$$\bar{X} - \bar{Y} \sim \mathrm{N}\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}\right)$$

標準化すると

$$\frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\sigma_X^2/m + \sigma_Y^2/n}} \sim \mathrm{N}(0, 1)$$

標準正規分布表より

$$\Pr\left[-1.96 \le \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\sigma_X^2/m + \sigma_Y^2/n}} \le 1.96\right] = .95$$

ここで

$$\begin{aligned}
&-1.96 \le \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\sigma_X^2/m + \sigma_Y^2/n}} \le 1.96 \\
&\iff -1.96\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}} \le \bar{X} - \bar{Y} - (\mu_X - \mu_Y) \\
&\quad \le 1.96\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}} \\
&\iff \bar{X} - \bar{Y} - 1.96\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}} \le \mu_X - \mu_Y \\
&\quad \le \bar{X} - \bar{Y} + 1.96\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}
\end{aligned}$$

したがって $\mu_X - \mu_Y$ の 95 %信頼区間は

$$\left[\bar{X} - \bar{Y} - 1.96\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}, \right.$$
$$\left. \bar{X} - \bar{Y} + 1.96\sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}\right]$$

2.3.2　母分散が未知

$\sigma_X^2 = \sigma_Y^2 = \sigma^2$ なら

$$\frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\sigma^2(1/m + 1/n)}} \sim \mathrm{N}(0, 1)$$

プールした標本分散を s^2 とする．σ^2 を s^2 に置き換えると

$$\frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{s^2(1/m + 1/n)}} \sim \mathrm{t}(m + n - 2)$$

t 分布表より，例えば $m = 4$，$n = 6$ なら

$$\Pr\left[-2.306 \le \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{s^2(1/4 + 1/6)}} \le 2.306\right] = .95$$

ここで

$$\begin{aligned}
&-2.306 \le \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{s^2(1/4 + 1/6)}} \le 2.306 \\
&\iff -2.306\sqrt{s^2\left(\frac{1}{4} + \frac{1}{6}\right)} \le \bar{X} - \bar{Y} - (\mu_X - \mu_Y) \\
&\quad \le 2.306\sqrt{s^2\left(\frac{1}{4} + \frac{1}{6}\right)} \\
&\iff \bar{X} - \bar{Y} - 2.306\sqrt{s^2\left(\frac{1}{4} + \frac{1}{6}\right)} \le \mu_X - \mu_Y \\
&\quad \le \bar{X} - \bar{Y} + 2.306\sqrt{s^2\left(\frac{1}{4} + \frac{1}{6}\right)}
\end{aligned}$$

したがって $m=4$, $n=6$ なら $\mu_X-\mu_Y$ の 95 % 信頼区間は

$$\left[\bar{X}-\bar{Y}-2.306\sqrt{s^2\left(\frac{1}{4}+\frac{1}{6}\right)},\right.$$
$$\left.\bar{X}-\bar{Y}+2.306\sqrt{s^2\left(\frac{1}{4}+\frac{1}{6}\right)}\right]$$

注 2. $\sigma_X^2\neq\sigma_Y^2$ なら近似的な信頼区間を用いる．

2.4 母分散の比の信頼区間（p. 229）

2.4.1 F 分布の上側確率と下側確率

補題 1.

$$X\sim\mathrm{F}(m,n)\Longrightarrow\frac{1}{X}\sim\mathrm{F}(n,m)$$

証明. F 分布の定義より

$$X=\frac{U/m}{V/n}$$

ただし $U\sim\chi^2(m)$ と $V\sim\chi^2(n)$ は独立．逆数は

$$\frac{1}{X}=\frac{V/n}{U/m}$$

したがって $1/X\sim\mathrm{F}(n,m)$. □

注 3. $\mathrm{F}(m,n)$ の下側確率は

$$\Pr[X\leq x]=\Pr\left[\frac{1}{X}\geq\frac{1}{x}\right]$$

より $\mathrm{F}(n,m)$ の上側確率から求まる（図 1）.

2.4.2 母平均が既知

σ_X^2/σ_Y^2 の 95 % 信頼区間を求める．ただし μ_X,μ_Y は既知とする．標本分散を $\hat{\sigma}_X^2,\hat{\sigma}_Y^2$ とすると

$$\frac{m\hat{\sigma}_X^2}{\sigma_X^2}\sim\chi^2(m)$$
$$\frac{n\hat{\sigma}_Y^2}{\sigma_Y^2}\sim\chi^2(n)$$

両者は独立だから

$$\frac{\hat{\sigma}_X^2/\sigma_X^2}{\hat{\sigma}_Y^2/\sigma_Y^2}\sim\mathrm{F}(m,n)$$

すなわち

$$\frac{\hat{\sigma}_X^2/\hat{\sigma}_Y^2}{\sigma_X^2/\sigma_Y^2}\sim\mathrm{F}(m,n)$$

F 分布表より，例えば $m=4$, $n=6$ なら

$$\Pr\left[\frac{1}{9.197}\leq\frac{\hat{\sigma}_X^2/\hat{\sigma}_Y^2}{\sigma_X^2/\sigma_Y^2}\leq 6.227\right]=.95$$

ここで

$$\begin{aligned}&\frac{1}{9.197}\leq\frac{\hat{\sigma}_X^2/\hat{\sigma}_Y^2}{\sigma_X^2/\sigma_Y^2}\leq 6.227\\ \Longleftrightarrow\ &\frac{1}{6.227}\leq\frac{\sigma_X^2/\sigma_Y^2}{\hat{\sigma}_X^2/\hat{\sigma}_Y^2}\leq 9.197\\ \Longleftrightarrow\ &\frac{1}{6.227}\frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2}\leq\frac{\sigma_X^2}{\sigma_Y^2}\leq 9.197\frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2}\end{aligned}$$

したがって $m=4$, $n=6$ なら σ_X^2/σ_Y^2 の 95 %信頼区間は

$$\left[\frac{1}{6.227}\frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2},9.197\frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2}\right]$$

2.4.3 母平均が未知

標本分散を s_X^2,s_Y^2 とすると

$$\frac{(m-1)s_X^2}{\sigma_X^2}\sim\chi^2(m-1)$$
$$\frac{(n-1)s_Y^2}{\sigma_Y^2}\sim\chi^2(n-1)$$

両者は独立だから

$$\frac{s_X^2/\sigma_X^2}{s_Y^2/\sigma_Y^2}\sim\mathrm{F}(m-1,n-1)$$

すなわち

$$\frac{s_X^2/s_Y^2}{\sigma_X^2/\sigma_Y^2}\sim\mathrm{F}(m-1,n-1)$$

F 分布表より，例えば $m=4$, $n=6$ なら

$$\Pr\left[\frac{1}{14.885}\leq\frac{s_X^2/s_Y^2}{\sigma_X^2/\sigma_Y^2}\leq 7.764\right]=.95$$

ここで

$$\begin{aligned}&\frac{1}{14.885}\leq\frac{s_X^2/s_Y^2}{\sigma_X^2/\sigma_Y^2}\leq 7.764\\ \Longleftrightarrow\ &\frac{1}{7.764}\leq\frac{\sigma_X^2/\sigma_Y^2}{s_X^2/s_Y^2}\leq 14.885\\ \Longleftrightarrow\ &\frac{1}{7.764}\frac{s_X^2}{s_Y^2}\leq\frac{\sigma_X^2}{\sigma_Y^2}\leq 14.885\frac{s_X^2}{s_Y^2}\end{aligned}$$

したがって $m=4$, $n=6$ なら σ_X^2/σ_Y^2 の 95 %信頼区間は

$$\left[\frac{1}{7.764}\frac{s_X^2}{s_Y^2},14.885\frac{s_X^2}{s_Y^2}\right]$$

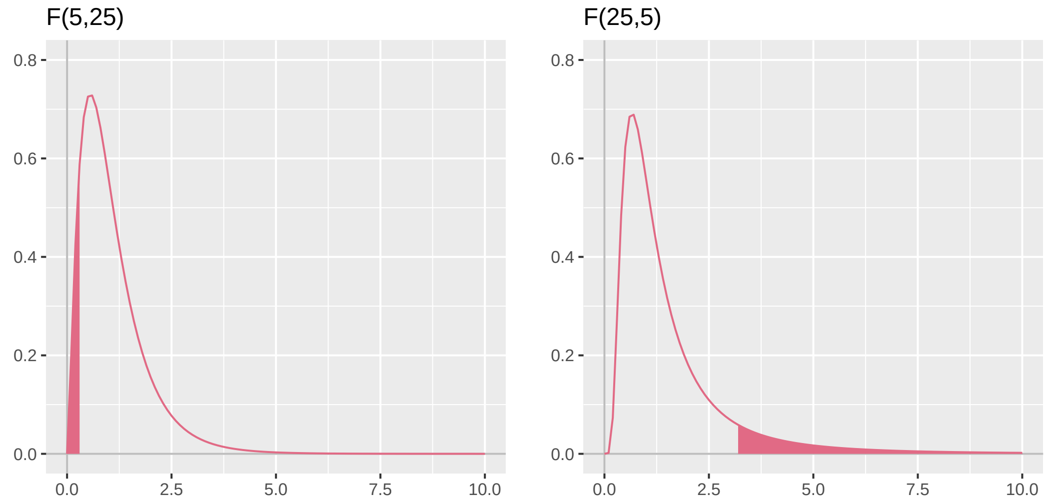

図 1　F$(5,25)$ の 10 %下側確率と F$(25,5)$ の 10 %上側確率

3　ベルヌーイ母集団

3.1　母比率と標本比率（p. 250）

母集団分布を $\mathrm{Bin}(1,p)$ とする．

定義 5. ベルヌーイ母集団における成功（$=1$）の比率を**母比率**という．

注 4. $\mathrm{Bin}(1,p)$ の平均は p なので母比率＝母平均．

定義 6. ベルヌーイ母集団からの標本における成功（$=1$）の比率を**標本比率**という．

注 5. 値が 1 か 0 なので標本比率＝標本平均．

3.2　母比率の信頼区間（p. 229）

p の 95 %信頼区間を近似的に求める．$\mathrm{Bin}(1,p)$ の平均は p，分散は $p(1-p)$．したがって母平均の信頼区間を近似的に求めればよい．大きさ n の無作為標本の標本平均を $\hat{p}$ とすると，中心極限定理より

$$\hat{p} \overset{a}{\sim} \mathrm{N}\left(p, \frac{p(1-p)}{n}\right)$$

標準化すると

$$\frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \overset{a}{\sim} \mathrm{N}(0,1)$$

分母の p を $\hat{p}$ に置き換えると

$$\frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \overset{a}{\sim} \mathrm{N}(0,1)$$

標準正規分布表より

$$\Pr\left[-1.96 \le \frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \le 1.96\right] \approx .95$$

ここで

$$\begin{aligned}
&-1.96 \le \frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \le 1.96 \\
&\iff -1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le \hat{p}-p \le 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\
&\iff \hat{p}-1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p}+1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}
\end{aligned}$$

したがって p の 95 %信頼区間は近似的に

$$\left[\hat{p}-1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p}+1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

例 2. 100 世帯を対象にある番組の視聴率を調査したら 10 %の視聴率であった．真の視聴率 p の 95 %信頼区間は $n=100$，$\hat{p}=.1$ を代入すると

$$\left[.1-1.96\frac{\sqrt{.1(1-.1)}}{10}, .1+1.96\frac{\sqrt{.1(1-.1)}}{10}\right]$$

計算すると $[.0412, .1588]$，すなわち 4.12〜15.88 % となる．

4　今日のキーワード

区間推定，信頼域，信頼区間，信頼係数，母平均の信頼区間，母分散の信頼区間，母平均の差の信頼区間，母分散の比の信頼区間，母比率，標本比率，母比率の信頼区間

5　次回までの準備

提出　復習テスト 14–20

復習　教科書第 11 章 5 節，復習テスト 20

試験　(1) 教科書を読む (2) 用語の定義を覚える (3) 復習テストを自力で解く (4) 過去問に挑戦