統計的推定

母集団の特徴を表す数値(母数)の近似値を標本から求めることを

統計的推定という.

点推定

例. 20歳男子の平均身長を調べるのに. 100人の標本をとり.

その平均値を近似値とする.

このように未知の母数 θ (上でいう平均身長)に対し.

ある統計量 $\Theta(X_1, \cdots, X_n)$ (上でいう標本平均)の実現値を近似値とする方法を

標本変量から計算される値.　　点推定という.

→推定量の候補は1つではない. 例えば, 上の例で.

$$\Theta(X_1, \cdots, X_{100}) = \frac{a_1X_1 + a_2X_2 + \cdots + a_{100}X_{100}}{a_1 + a_2 + \cdots + a_{100}} \quad (a_i \geq 0)$$ としてもよい.

ただし. 推定量は以下の3つをみたさないといけない.

(1) 不偏性 : $E(\Theta(X_1, \cdots, X_n)) = \theta$

(2) 一致性 : $\lim_{n\to\infty} P(|\Theta(X_1, \cdots, X_n) - \theta| > \varepsilon) = 0$

(3) 有効性 : $V(\Theta(X_1, \cdots, X_n))$ が小さい → 小さい方が優れている.

区間推定

点推定はわかりやすく計算も単純だが、本当に母数に近い値になってるか疑問.

→"母数が $X-\delta$ から $X+\delta$ の間に入る確率が $1-\alpha$"

という求め方を考える. この考え方を区間推定法という.

例4 $P(\overline{X}-1 \leqq \mu \leqq \overline{X}+1) = P(\mu-1 \leqq \overline{X} \leqq \mu+1) \geqq 0.99$ となればよい.

ここで. $\overline{X}$ は $N\left(\mu, \frac{25}{n}\right)$ に従うから.

$Z = \dfrac{\overline{X}-\mu}{\frac{5}{\sqrt{n}}} = \dfrac{\sqrt{n}}{5}(\overline{X}-\mu)$ とすると. Z は $N(0,1)$ に従い.

$P(\mu-1 \leqq \overline{X} \leqq \mu+1) = P\left(\frac{\sqrt{n}}{5}((\mu-1)-\mu) \leqq Z \leqq \frac{\sqrt{n}}{5}((\mu+1)-\mu)\right)$

$= P\left(-\frac{\sqrt{n}}{5} \leqq Z \leqq \frac{\sqrt{n}}{5}\right) = 2 \cdot P\left(0 \leqq Z \leqq \frac{\sqrt{n}}{5}\right) \geqq 0.99$ となる.

$\therefore P\left(0 \leqq Z \leqq \frac{\sqrt{n}}{5}\right) \geqq 0.495$ となるが. 正規分布表から.

$\frac{\sqrt{n}}{5} \geqq 2.5758$ となり $n \geqq 165.87$ をえる $\therefore n$ は 166 以上であればよい

問5,6.

5 $P\left(0 \leqq Z \leqq \frac{\sqrt{n}}{5}\right) \geqq 0.475$ より $\frac{\sqrt{n}}{5} \geqq 1.96$ となる

$\therefore n$ は 97 以上 であればよい.

6 $P(\overline{X}-2 \leqq \mu \leqq \overline{X}+2) = P(\mu-2 \leqq \overline{X} \leqq \mu+2) \geqq 0.9$ となればよい.

ここで $\overline{X}$ は $N\left(\mu, \frac{16}{n}\right)$ に従うから. $Z = \dfrac{\overline{X}-\mu}{\frac{4}{\sqrt{n}}} = \dfrac{\sqrt{n}}{4}(\overline{X}-\mu)$ は $N(0,1)$ に従う.

$\therefore P(\mu-2 \leqq \overline{X} \leqq \mu+2) = P\left(-\frac{\sqrt{n}}{4} \cdot 2 \leqq Z \leqq \frac{\sqrt{n}}{4} \cdot 2\right)$

$= 2 \cdot P\left(0 \leqq Z \leqq \frac{\sqrt{n}}{2}\right) \geqq 0.9.$ $\therefore P\left(0 \leqq Z \leqq \frac{\sqrt{n}}{2}\right) \geqq 0.45$

$\therefore \frac{\sqrt{n}}{2} \geqq 1.6449.$ より $n \geqq 11.$ を得る.

区間推定では 3つの数値が連動している

(1) 標本のサイズ (2) $\overline{X}-\delta < \mu < \overline{X}+\delta$ という区間 (3) $1-\alpha$ が表す信頼度.

標本の実現値 $\bar{x}$ を得たとする．

$\bar{x}-\delta \leqq \mu \leqq \bar{x}+\delta$ である確率が $1-\alpha$ であるとき，

区間 $\bar{x}-\delta \leqq \mu \leqq \bar{x}+\delta$ を母平均 μ の $1-\alpha$ 信頼区間という

例⑦ $\bar{X}$ を標本平均とし，$P(\bar{X}-\delta \leqq \mu \leqq \bar{X}+\delta)=0.99$ となる δ を求める．

$P(\bar{X}-\delta \leqq \mu \leqq \bar{X}+\delta)=P(\mu-\delta \leqq \bar{X} \leqq \mu+\delta)=0.99$ である．

ここで $\bar{X}$ は $N(\mu, \frac{1}{2})$ に従うので，$Z=\frac{\bar{X}-\mu}{\frac{1}{\sqrt{2}}}=\sqrt{2}(\bar{X}-\mu)$ は $N(0.1)$ に従う

$\therefore P(\mu-\delta \leqq \bar{X} \leqq \mu+\delta)=P(-\sqrt{2}\delta \leqq Z \leqq \sqrt{2}\delta)=2 \cdot P(0 \leqq Z \leqq \sqrt{2}\delta)=0.99$

$\therefore P(0 \leqq Z \leqq \sqrt{2}\delta)=0.495$ となり $\sqrt{2}\delta=2.5758$ $\therefore \delta=1.8$ となる．

平均身長の99%信頼区間は，$168.0 \leqq \mu \leqq 171.6$ となる．

問⑧,⑨ $\bar{X}$ は $N(\mu, \frac{1}{4})$ に従うので，$Z=\frac{\bar{X}-\mu}{\frac{1}{2}}=2(\bar{X}-\mu)$ は $N(0.1)$ に従う．

$\therefore P(\mu-\delta \leqq \bar{X} \leqq \mu+\delta)=P(-2\delta \leqq Z \leqq 2\delta)=2 \cdot P(0 \leqq Z \leqq 2\delta)=0.95$．

$\therefore P(0 \leqq Z \leqq 2\delta)=0.475$ となり $2\delta=1.96$ $\therefore \delta=0.98 \fallingdotseq 1.0$

$\therefore$ 95%信頼区間は．$169.1 \leqq \mu \leqq 171.1$ である．

⑨ $\bar{X}$ を標本平均とすると．$\bar{X}$ は $N(\mu, \frac{(60)^2}{20})=N(\mu, 180)$ に従い．

$Z=\frac{\bar{X}-\mu}{\sqrt{180}}=\frac{\bar{X}-\mu}{6\sqrt{5}}$ は $N(0.1)$ に従う．

$0.9=P(\bar{X}-\delta \leqq \mu \leqq \bar{X}+\delta)=P(\mu-\delta \leqq \bar{X} \leqq \mu+\delta)$

$=P(\frac{-\delta}{6\sqrt{5}} \leqq Z \leqq \frac{\delta}{6\sqrt{5}})=2P(0 \leqq Z \leqq \frac{\delta}{6\sqrt{5}})$

$\therefore P(0 \leqq Z \leqq \frac{\delta}{6\sqrt{5}})=0.45$．$\frac{\delta}{6\sqrt{5}}=1.6449$．$\fallingdotseq 22$

$\therefore$ 90%信頼区間は．$258 \leqq \mu \leqq 302$ である