2013年後期
応用確率統計

⑪ 仮説検定

河野 行雄
kawano@pe.titech.ac.jp
2014年1月9日

統計的検定とは？

- 統計的推測

  母数推定（標本平均、最尤推定）

  仮説検定、信頼区間、回帰曲線

- 仮説検定

  観測されたデータと想定する仮定が一致するかどうかを確認する方法

- 例

  A県で生産されたリンゴの糖度： $X_1,\cdots,X_n$

  B県で生産されたリンゴの糖度： $Y_1,\cdots,Y_m$

  ⇩

  標本平均 $\overline{X},\overline{Y}$ を比べることで平均の糖度が等しいといえるか？

復習

- 尤度関数

$$L(\theta)=\prod_{i=1}^{n} f(X_i,\theta)$$

- 最尤推定量

$$\hat{\theta}^*=\arg\max_\theta L(\theta)=\arg\max_\theta \log L(\theta)$$

← 漸近不偏、漸近有効（クラメールラオの下界を達成）

- クラメール・ラオの不等式

$$V\left(\hat{\theta}\mid\theta_0\right)\geq\frac{1}{nI}$$

$$I=E\left(\left(\frac{\partial}{\partial\theta}\log f(X,\theta_0)\right)^2\right)=-E\left(\frac{\partial^2}{\partial\theta^2}\log f(X,\theta_0)\right)$$

仮説検定の枠組み

- 観測データ

$$X_i=\theta_1+\varepsilon_{Xi}\quad(i=1,\cdots,n)$$

$$Y_j=\theta_2+\varepsilon_{Yj}\quad(j=1,\cdots,m)$$

  $\varepsilon_{Xi},\varepsilon_{Yi}$ は i.i.d. で平均 0 分散 σ^2 の正規分布 $N\left(0,\sigma^2\right)$ に従う

- 仮説

  仮説 H_0 ： 「 $\theta_1=\theta_2$ 」 （帰無仮説、統計的仮説）

  仮説 H_1 ： 「 $\theta_1\neq\theta_2$ 」 （対立仮説）

  ⇩

  有限の標本数 n,m で判定する枠組み → 仮説検定

検定統計量

■ 母平均の推定量

$$\hat{\theta}_1 = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \hat{\theta}_2 = \bar{Y} = \frac{1}{m}\sum_{j=1}^{m} Y_j$$

■ 検定統計量

仮説 H_0 が真 → $\bar{Y} - \bar{X}$ は $N\left(0, \left(\frac{1}{n} + \frac{1}{m}\right)\sigma^2\right)$ に従う

検定統計量: $T \equiv \dfrac{\bar{Y} - \bar{X}}{\sqrt{\frac{1}{n} + \frac{1}{m}}\,\sigma} = \sqrt{\dfrac{nm}{n+m}}\dfrac{\bar{Y} - \bar{X}}{\sigma}$

検定統計量は平均 0 分散 1 の正規分布 $N(0,1)$ に従う

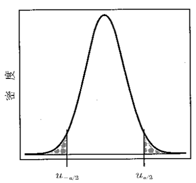
仮説の受容と棄却

■ 有意水準

判定誤り(過誤)の確率を α とする判定域 μ

$P(|T| \geq \mu) = \alpha$

判定域(例:1.96)　有意水準(例:0.05)



■ 仮説の受容

$|T| < \mu$ ⟶ 仮説 H_0 を受容 「 $\theta_1 = \theta_2$ 」

■ 仮説の棄却

$|T| \geq \mu$ ⟶ 仮説 H_0 を棄却 「 $\theta_1 \neq \theta_2$ 」

(確率 α の過誤を含む)

例12.1

■ 抜き取り検査

あるエ場で製造される機械の寿命は平均 θ_0 時間

抜き取られた n 個の機械の寿命は $X_1, \cdots, X_n$

仮説 H_0 : 「機械の寿命は θ_0」 は正しいか?

■ 仮説検定

観測データ: $X_i = \theta_0 + \varepsilon_i \quad (i = 1, \cdots, n) \qquad \varepsilon \to N(0, \sigma^2)$

推定量: $\hat{\theta} = \bar{X}$

検定統計量: $T = \dfrac{\sqrt{n}\left(\bar{X} - \theta_0\right)}{\sigma}$

判定: $|T| < \mu = 1.96$ ⟶ 仮説を受容 「 $\theta = \theta_0$ 」

$|T| \geq \mu = 1.96$ ⟶ 仮説を棄却 「 $\theta \neq \theta_0$ 」

分散の検定

■ 観測データ

$X_i = \theta_0 + \varepsilon_i \quad (i = 1, \cdots, n) \qquad \varepsilon \to N(0, \sigma^2)$

■ 推定量

$\hat{\theta} = \bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ (標本平均)

$\hat{\sigma}^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \hat{\theta}\right)^2$ (不偏分散)

■ 仮説

仮説 H_0 : 「 $\sigma^2 = \sigma_0^2$ 」 (帰無仮説)

仮説 H_1 : 「 $\sigma^2 \neq \sigma_0^2$ 」 (対立仮説)

カイ二乗分布

■ n 個の標準正規分布に従う確率変数の二乗和

$$Y = X_1^2 + X_2^2 + \cdots + X_n^2 = Z_1 + Z_2 + \cdots + Z_n$$

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$

$$p(z) = \frac{1}{\sqrt{2\pi z}} \exp\left(-\frac{z}{2}\right) = \frac{1}{2^{1/2}\Gamma(1/2)} z^{-1/2} e^{-z/2}$$

$$\varphi_Z(t) = \frac{1}{2^{1/2}\Gamma(1/2)} \int_0^\infty z^{-1/2} e^{-(1/2 - jt)z} dz = \frac{1}{\sqrt{1-2jt}}$$

$$\varphi_Y(t) = \frac{1}{(1-2jt)^{n/2}}$$

$$p(y) = \frac{1}{2^{n/2}\Gamma(n/2)} z^{n/2-1} e^{-z/2}$$

← 自由度 n のカイ二乗分布

ガンマ関数

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$$

$$\Gamma(1/2) = \sqrt{\pi},\ \Gamma(1) = 1$$

例12.4

■ ばらつきの年度比較

昨年調べたリンゴの糖度のばらつきは σ_0^2

今年売られているリンゴの糖度は $X_1, \cdots, X_n$

仮説 H_0： 「糖度のばらつき σ^2 は昨年と同じ」 は正しいか？

■ 仮説検定

観測データ： $X_i = \theta_0 + \varepsilon_i \quad (i = 1, \cdots, n) \qquad \varepsilon \to N(0, \sigma^2)$

推定量： $\hat{\theta} = \bar{X}$

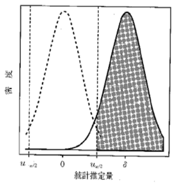
検定統計量： $T = \frac{1}{\sigma_0^2} \sum_{i=1}^n (X_i - \hat{\theta})^2$ ← 自由度 $n-1$ の カイ二乗分布

有意水準： $\int_0^{\mu_1} P(T) dT = \frac{\alpha}{2} \qquad \int_{\mu_2}^\infty P(T) dT = \frac{\alpha}{2}$

判定： $\mu_1 < T < \mu_2$ ⟶ 仮説を受容

$T \le \mu_1$ または $T \ge \mu_2$ ⟶ 仮説を棄却

過誤と検出力



■ 仮説

帰無仮説 H_0： 「 $\theta_1 = \theta_2$ 」

対立仮説 H_1： 「 $\theta_1 \ne \theta_2$ 」

■ 過誤

第1種の過誤： 帰無仮説が正しいのに対立仮説 → 有意水準 α

第2種の過誤： 対立仮説が正しいのに帰無仮説 → 検出力 β

■ 検出力

検出力： 対立仮説が正しいときに帰無仮説を棄却する確率

検定統計量： $T \equiv \sqrt{\frac{nm}{n+m}} \frac{\hat{\theta}_2 - \hat{\theta}_1}{\sigma}$

$\theta_1 \ne \theta_2$ のとき T は $N(\delta, 1)$ に従う

検出力： $P(|T| \ge \mu \mid \delta) \qquad \delta = \sqrt{\frac{nm}{n+m}} \frac{\theta_2 - \theta_1}{\sigma}$

ネイマン・ピアソンの補題

■ 観測データ

$X_1, \cdots, X_n$ は i.i.d. で確率密度関数 $f(X, \theta)$ に従う

■ 仮説

帰無仮説 H_0： 「 $\theta = \theta_0$ 」

対立仮説 H_1： 「 $\theta = \theta_1$ 」

■ 最適な検定統計量（ネイマン・ピアソンの補題）

検定統計量： $T \equiv \prod_{i=1}^n \frac{f(X_i, \theta_1)}{f(X_i, \theta_0)}$ ← 尤度比検定 検出力を最大化

有意水準： $P(T \ge \mu \mid \theta_0) = \alpha$ ⟶ $T < \mu$ or $T \ge \mu$ で判定

まとめ

■ 仮説

帰無仮説 H_0 : 「 $\theta_1 = \theta_2$ 」「 $\theta = \theta_0$ 」

対立仮説 H_1 : 「 $\theta_1 \neq \theta_2$ 」「 $\theta = \theta_1$ 」

■ 検定統計量

$$T \equiv \sqrt{\frac{nm}{n+m}} \frac{\overline{Y} - \overline{X}}{\sigma}$$ ← 分散1の正規分布

■ 有意水準

$$P(|T| \geq \mu) = \alpha$$ ← 第一種の過誤

■ 仮説の受容と棄却

$|T| < \mu$ ⟶ 帰無仮説を受容 「 $\theta_1 = \theta_2$ 」

$|T| \geq \mu$ ⟶ 帰無仮説を棄却 「 $\theta_1 \neq \theta_2$ 」