Ch.7 統計的仮説検定

- 帰無仮説 H_0, 対立仮説 H_1
- 母数空間 $\Theta = \Theta_0 \sqcup \Theta_1$.

$H_0 : \theta \in \Theta_0$, $H_1 : \theta \in \Theta_1 = \Theta_0^c$.

Def. (検定)

仮説検定方式とは，標本空間 $\mathcal{X}$ を

$R = \{x \in \mathcal{X} \mid H_0 \text{を棄却}\}$ と $A = \{x \in \mathcal{X} \mid H_0 \text{を受容}\}$

に分割するルールのことである:

$\mathcal{X} = R \sqcup A$.

R を H_0 の棄却域，A を H_0 の受容域という.

$T(\boldsymbol{X}) = T(X_1, \cdots, X_n)$ により R, A が決まるとき，$T(\boldsymbol{X})$ を

検定統計量という. □

- 統計的仮説検定では，帰無仮説を棄却するのに高い信頼性を与えている.

Def. (有意水準)

$\alpha \in (0,1)$ が $\sup_{\theta \in \Theta_0} P_\theta(\boldsymbol{X} \in R) \leq \alpha$ をみたすとき，

α を有意水準という. □

§7.2 正規母集団に関する検定.

$X_1, \cdots, X_m \overset{i.i.d.}{\sim} N(\mu_1, \sigma_0^2)$

$Y_1, \cdots, Y_n \overset{i.i.d.}{\sim} N(\mu_2, \sigma_0^2)$

σ_0^2 : known.

(A) $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$. 両側検定

(B) $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 > \mu_2$. 片側検定

(A) $T(\bar{X}, \bar{Y}) = \dfrac{|\bar{X} - \bar{Y}|}{\sigma_0}$

$\dfrac{|\bar{X} - \bar{Y}|}{\sigma_0} > C$ で H_0 を棄却.

$\mu = \mu_1 = \mu_2$ とする.

$$(\bar{X} - \mu) + (\bar{Y} - \mu) \sim N\left(0, \left(\frac{1}{n} + \frac{1}{m}\right)\sigma_0^2\right)$$

$Z := \dfrac{\bar{X} - \bar{Y}}{\sigma_0 \sqrt{\frac{1}{n} + \frac{1}{m}}}$ とすると $Z \sim N(0,1)$

$$P_{H_0}\left(|Z| > C\sqrt{\frac{mn}{m+n}}\right) = \alpha$$

をみたすように C を決定 : $C = \sqrt{\dfrac{m+n}{mn}}\, z_{\frac{\alpha}{2}}$.

(B) $P\left(Z > C\sqrt{\dfrac{mn}{m+n}}\right) = \alpha$

をみたすように C を決定 : $C = \sqrt{\dfrac{m+n}{mn}}\, z_{\alpha}$.

7.2.2 t-検定

$$X_1, \cdots, X_m \overset{i.i.d.}{\sim} N(\mu_1, \sigma^2)$$

$$Y_1, \cdots, Y_n \overset{i.i.d.}{\sim} N(\mu_2, \sigma^2)$$

σ^2 : unknown.

σ^2の推定量:

$$\hat{\sigma}^2 = \frac{1}{m+n-2}\left(\sum_{i=1}^{m}(X_i - \overline{X})^2 + \sum_{i=1}^{n}(Y_i - \overline{Y})^2\right)$$

検定: $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$.

$$T(\overline{X}, \overline{Y}) = \frac{|\overline{X} - \overline{Y}|}{\hat{\sigma}}$$

$(\overline{X}, \overline{Y}) \perp\!\!\!\perp \hat{\sigma}^2$.

$$U := (m+n-2)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{m+n-2},$$

$$Z := \frac{\overline{X} - \overline{Y}}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}}$$ とすると $Z \sim N(0, 1)$

$$\therefore T := \frac{Z}{\sqrt{\frac{U}{m+n-2}}} \sim t_{m+n-2}.$$

7.2.3 F-検定.

$$X_1, \cdots, X_m \overset{i.i.d.}{\sim} N(\mu_1, \sigma_1^2)$$

$$Y_1, \cdots, Y_n \overset{i.i.d.}{\sim} N(\mu_2, \sigma_2^2)$$

検定： $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \neq \sigma_2^2$.

σ_1^2, σ_2^2の推定量.

$$\hat{\sigma}_1^2 = \frac{1}{m-1}\sum_{i=1}^{m}(X_i - \bar{X})^2, \quad \hat{\sigma}_2^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2.$$

検定統計量： $T = \dfrac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$

H_0のもとでは, $\sigma^2 := \sigma_1^2 = \sigma_2^2$.

$$F := \frac{\frac{\hat{\sigma}_1^2}{\sigma^2}}{\frac{\hat{\sigma}_2^2}{\sigma^2}} \sim F_{m-1, n-1}.$$

§7.3 検定統計量の導出方法

7.3.1. 尤度比検定

$X_1, \cdots, X_n \overset{i.i.d.}{\sim} f(x|\theta)$.

<u>Def.</u> (尤度比検定統計量)

$H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1 = \Theta_0^c$ を検定するための

尤度比検定統計量は、

$$\lambda(\mathbb{X}) := \frac{\sup_{\theta \in \Theta_0} L(\theta|\mathbb{X})}{\sup_{\theta \in \Theta} L(\theta|\mathbb{X})}$$

で定められる。このとき、$C > 0$ に対して H_0 の棄却域が

$$R = \{ x \in \mathcal{X} \mid \lambda(x) \leq C \}$$

で与えられる検定を、尤度比検定という。 $\square$

$\hat{\theta}$: Θ での MLE, $\hat{\theta}_0$: Θ_0 での MLE とすると、

$$\lambda(\mathbb{X}) = \frac{L(\hat{\theta}_0, \mathbb{X})}{L(\hat{\theta}, \mathbb{X})}.$$

<u>Prop.</u> θ : 1-dim. 適当な正則条件のもと、

$H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1 = \Theta_0^c$ における

尤度比検定統計量は、H_0 のもとで $-2\log\lambda(\mathbb{X}) \xrightarrow{d} \chi_1^2$. $\square$

Prop. $X_1, \cdots, X_n \overset{i.i.d.}{\sim} f(x|\theta)$, $\theta \in \Theta$: k-dim

$\dim \Theta_0 = r\ (<k)$.

$H_0 : \theta \in \Theta_0$ vs. $H_1 : \theta \in \Theta_1 = \Theta_0^c$ における

尤度比検定統計量は, H_0のもとで

$$-2\log\lambda(X) \xrightarrow{d} \chi^2_{k-r}$$ □

7.3.2 Wald検定とスコア検定

Def. (Wald検定)

W_n : θの推定量. S_n^2 : $V[W_n]$の推定量.

$$\frac{W_n - \theta}{S_n} \xrightarrow{d} N(0,1)$$

が成立しているとする.

$H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$ における H_0の棄却域を

$$R = \left\{ x \in \mathcal{X} \;\middle|\; \frac{|W_n - \theta_0|}{S_n} \geq z_{\frac{\alpha}{2}} \right\}$$

とする検定方式を**Wald検定**という. □

ex $W_n = \hat{\theta}_n$: θのMLE.

漸近正規性より $\sqrt{nI_1(\theta)}\,(\hat{\theta}_n - \theta) \xrightarrow{d} N(0,1)$.

Slutskyの定理より $\sqrt{nI_1(\hat{\theta}_n)}\,(\hat{\theta}_n - \theta) \xrightarrow{d} N(0,1)$.

$\therefore$ Wald検定の棄却域は、

$$R = \left\{ x \in \mathcal{X} \;\middle|\; \sqrt{n I_1(\hat{\theta}_n)}\, |\hat{\theta}_n - \theta_0| \geq z_{\frac{\alpha}{2}} \right\}. \qquad \square$$

Def. (スコア検定)

$H_0 : \theta = \theta_0 \quad vs. \quad H_1 : \theta \neq \theta_0$ における H_0 の棄却域を

$$R = \left\{ x \in \mathcal{X} \;\middle|\; \frac{|S_n(\theta_0, x)|}{\sqrt{n I_1(\theta_0)}} \geq z_{\frac{\alpha}{2}} \right\}$$

とする検定方式をスコア検定という. $\square$

ex $X_1, \cdots, X_n \overset{i.i.d.}{\sim} B(1, p)$

$H_0 : p = p_0 \quad vs. \quad H_1 : p \neq p_0.$

$$\begin{aligned}
I_1(p) &= -\mathbb{E}\left[\frac{d^2}{dp^2} \log f(X|p) \right] \\
&= \mathbb{E}\left[\frac{X}{p^2} + \frac{1-X}{(1-p)^2} \right] \\
&= \frac{p}{p^2} + \frac{1-p}{(1-p)^2} \\
&= \frac{1}{p} + \frac{1}{1-p} \\
&= \frac{1}{p(1-p)}.
\end{aligned}$$

Wald検定の棄却域:

$$R = \left\{ x \in \mathcal{X} \;\middle|\; \sqrt{n}\, |\hat{p} - p_0| \geq z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})} \right\}$$

また，

$$L(p, x) := \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$ なので．

$$\begin{aligned}
S_n(p, \mathbb{X}) &= \frac{d}{dp}\log L(p, \mathbb{X}) \\
&= \frac{d}{dp}\left(\sum_{i=1}^{n}\left(X_i \log p + (1-X_i)\log(1-p)\right)\right) \\
&= \sum_{i=1}^{n}\left(\frac{X_i}{p} - \frac{1-X_i}{1-p}\right) \\
&= \frac{n\hat{p}}{p} - \frac{n-n\hat{p}}{1-p} \qquad (\bar{X} = \hat{p}) \\
&= \frac{n(\hat{p}-p)}{p(1-p)}
\end{aligned}$$

スコア検定の棄却域は，

$$R = \left\{ x \in \mathcal{X} \;\middle|\; \frac{\sqrt{n}\,|\hat{p}-p_0|}{\sqrt{p_0(1-p_0)}} \geq z_{\frac{\alpha}{2}} \right\}$$

$\square$

§7.4 適合度検定

7.4.1 カイ2乗適合度検定

nコのデータが $C_1, \cdots, C_K$ の Kコのカテゴリーに分類．

それぞれ $X_1, \cdots, X_K$ コ観測．

$\rightarrow \sum_{k=1}^{K} X_k = n$．

各カテゴリーに入る確率：p_k $(k=1,\cdots,K)$

p_kの推定量：$\frac{X_k}{n}$．

p_kの理論値：π_k．

$H_0 : \forall k,\ p_k = \pi_k$ vs. $H_1 : \exists k,\ p_k \neq \pi_k$．

→ カテゴリーに関するカイ2乗適合度検定

・H_0が正しいとき，C_kに入るデータ数の理論値：$n\pi_k$

<u>Def.</u> (Pearsonのχ^2検定統計量)

$X = (X_1, \cdots, X_K)$，$\pi = (\pi_1, \cdots, \pi_K)$ に対し，

Pearsonのχ^2検定統計量を

$$Q(X, \pi) := \sum_{k=1}^{K} \frac{(X_k - n\pi_k)^2}{n\pi_k}$$

と定める． □

Prop. 帰無仮説 H_0 のもとで

$$Q(X, \pi) \xrightarrow{d} \chi^2_{K-1}$$

$\square$

・棄却域を

$$R = \{ x \in \mathcal{X} \mid Q(x, \pi) > \chi^2_{K-1, \alpha} \}$$

とすればよい（上側検定）

7.4.2 クロス表における独立性検定

・ A の事象 $\quad A = \bigsqcup_{i=1}^{r} A_i$,

B の事象 $\quad B = \bigsqcup_{i=1}^{c} B_i$.

n 個のデータ.

A_i かつ B_j の観測数 : X_{ij}

$P(A_i \cap B_j) =: p_{ij}$: 真の確率.

$$X_{i\cdot} := \sum_{j=1}^{c} X_{ij}, \quad X_{\cdot j} := \sum_{i=1}^{r} X_{ij}.$$

$$p_{i\cdot} := \sum_{j=1}^{c} p_{ij}, \quad p_{\cdot j} := \sum_{i=1}^{r} p_{ij}.$$

パラメータ数
$p_{i\cdot}$ $(i=1, \cdots, r-1)$
$p_{\cdot j}$ $(j=1, \cdots, c-1)$
$r+c-2$ コ

・ $H_0 : \forall (i,j), \; p_{ij} = p_{i\cdot} \times p_{\cdot j}$. ←

vs.

$H_1 : \exists (i,j), \; p_{ij} \neq p_{i\cdot} \times p_{\cdot j}$. ←

パラメータ数
$rc-1$ コ.

$$Q(\mathbb{X}) := \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{\left(X_{ij} - \frac{X_{i\cdot} \cdot X_{\cdot j}}{n}\right)^2}{\frac{X_{i\cdot} \cdot X_{\cdot j}}{n}}$$

$Q(\mathbb{X})$ は，H_0 のもとで，

$$Q(\mathbb{X}) \xrightarrow{d} \chi^2_{(r-1)(c-1)} .$$

§7.5 検定方式の評価.

7.5.1 検定のサイズと検出力.

- 第1種の誤り : H_0 が正しいのに棄却する : $P_{H_0}(X \in R)$.
- 第2種の誤り : H_0 が誤っているのに受容する : $P_{H_1}(X \notin R)$.

Def. (検出力函数)

- $\beta(\theta) := P_\theta(X \in R)$ を検出力函数という.

  $\theta \in \Theta_0$ に対しては, $\beta(\theta)$ は type I error を表し,

  $\theta \in \Theta_1$ に対しては, $1-\beta(\theta)$ は type II error を表す.

- 有意水準 $\alpha \in (0,1)$ に対し. $\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha$ のとき,

  サイズ α の検定といい, $\sup_{\theta \in \Theta_0} \beta(\theta) \leq \alpha$ のとき

  レベル α の検定という. $\square$

Def. (強力)

二つの検定手法 T_1, T_2 について, それぞれの検出力函数を

$\beta_1(\theta), \beta_2(\theta)$ とする. 次をみたすとき, T_1 は T_2 より強力であるという

(1) $\forall \theta \in \Theta_0$, $\beta_1(\theta) \leq \alpha$, $\beta_2(\theta) \leq \alpha$ T_1, T_2 ともレベル α の検定

(2) $\forall \theta \in \Theta_0^c$, $\beta_1(\theta) \geq \beta_2(\theta)$, 少なくとも1点で strict に成立. $\square$

type II error は T_1 の方が小さい

Def. (最強力検定)

C_α : レベル α の検定全体.

検定 $T \in C_\alpha$ が (一様)最強力 とは,

$\forall T' \in C_\alpha$ に対し, T は T' より強力であることをいう. $\square$

7.5.2 Neyman-Pearson の補題

$H_0 : \theta = \theta_0$ vs. $H_1 : \theta = \theta_1 \ (\neq \theta_0)$

$\mathbb{X} = (X_1, \cdots, X_n)$: random sample

$f_n(x|\theta)$: $\mathbb{X}$ の joint pdf.

$$\to \lambda(\mathbb{X}) = \frac{f_n(\mathbb{X}|\theta_0)}{f_n(\mathbb{X}|\theta_1)}.$$

尤度比検定の棄却域は.

$$R = \{x \in \mathcal{X} \mid \lambda(\mathbb{X}) < c\}$$

$$= \{x \in \mathcal{X} \mid f_n(\mathbb{X}|\theta_1) > k f_n(\mathbb{X}|\theta_0)\} \quad \left(k = \frac{1}{c}\right)$$

Th. (Neyman-Pearson の補題)

棄却域が $\{x \in \mathcal{X} \mid f_n(\mathbb{X}|\theta_1) > k f_n(\mathbb{X}|\theta_0)\}$ となる

尤度比検定は, 最強力. $\square$