

統計学II

早稲田大学政治経済学術院

西郷 浩

本日の目標

- 対標本における母平均の差の検定
- 母分散に関する検定
 - 母分散の値に関する検定
 - 母分散の比に関する検定

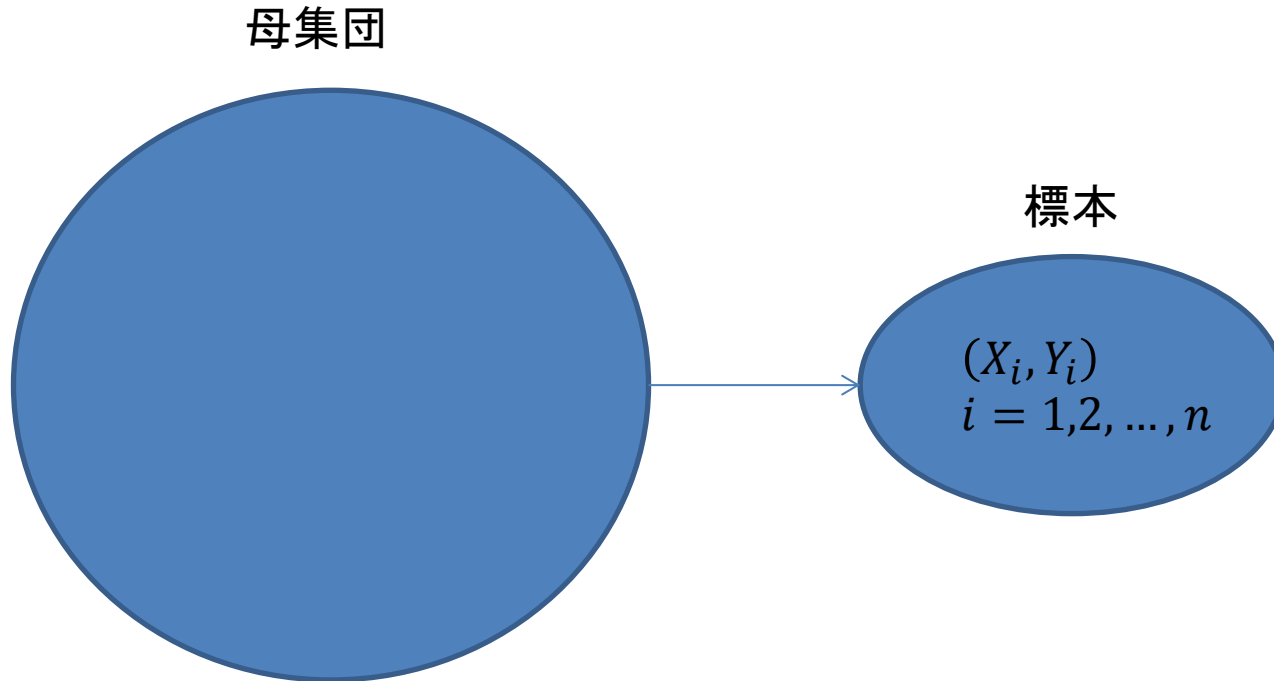
対標本における検定(1)

- 問題

- Pearson の親子の身長データから、20組の親子の身長を無作為抽出する。
 - (67, 71), (70, 73), (69, 71), (71, 72), (68, 67),
(67, 67), (66, 71), (70, 65), (65, 64), (66, 71),
(66, 70), (68, 71), (66, 69), (69, 71), (65, 62),
(70, 68), (66, 67), (67, 70), (64, 67), (68, 67).
 - 教科書の数値例と異なることに注意する。
- 父親の平均身長と息子の平均身長とに有意な差があるか？

対標本における検定(2)

図1: 対標本の例



対標本における検定(3)

- 標本平均

- $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$

- 標本平均の差の性質

- $E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_X - \mu_Y$

- $V(\bar{X} - \bar{Y}) = V(\bar{X}) + V(\bar{Y}) - 2Cov(\bar{X}, \bar{Y})$

- X_i と Y_i が独立でない。

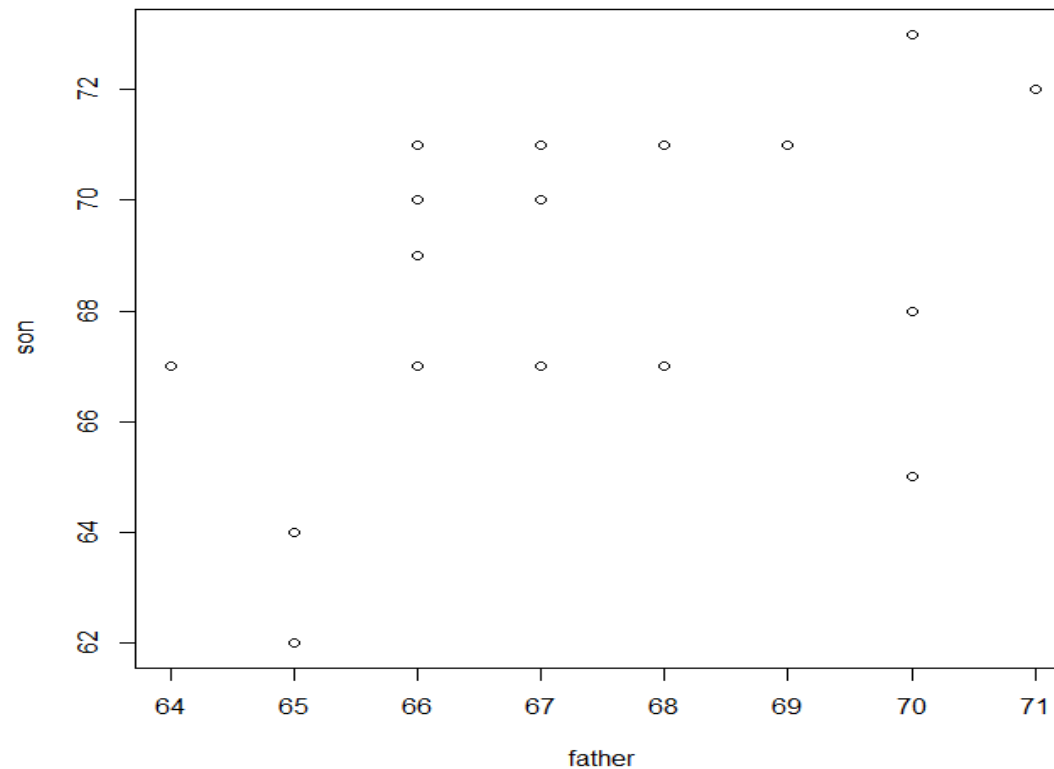
- 父親が長身だと、息子も長身になりやすい。

- したがって、 $Cov(\bar{X}, \bar{Y}) > 0$.

- » 注: X_i と Y_i が独立である。→ $Cov(\bar{X}, \bar{Y}) = 0$.

対標本における検定(4)

図2: 父親の身長と息子の身長散布図(対標本の場合)



対標本における検定(5)

- 対標本の検定方法(その1)

- 標本

- $(X_i, Y_i) \sim iid\{(\mu_X, \mu_Y), (\sigma_X^2, \sigma_Y^2, \sigma_{XY})\}$

- 仮説

- $\begin{cases} H_0: \mu_X - \mu_Y = 0 \\ H_1: \mu_X - \mu_Y \neq 0 \end{cases}$

- 検定統計量

- $|T| = \frac{|\bar{X} - \bar{Y}|}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{n} - 2\frac{S_{XY}}{n}}}$

対標本における検定(6)

– ただし、

$$\gg S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$\gg S_Y^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2$$

$$\gg S_{XY} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

– 両側検定(有意水準0.05)

- $|T| > 1.96$ であれば、 H_0 を棄却する。
 - 正規母集団を想定できるのであれば、1.96の代わりに $t_{0.025}(n-1)$ を用いる。

対標本における検定(7)

- 対標本の検定方法(その2)
 - 結果はその1と全く同じになる。
 - 各々の観察値における、 X_i と Y_i との差
 - $D_i = X_i - Y_i$
 - D_i の標本平均
 - $\bar{D} = \frac{1}{n} \sum_{i=1}^n D_i = \bar{X} - \bar{Y}$
 - \bar{D} の性質
 - $\mu_D = E(\bar{D}) = E(\bar{X} - \bar{Y}) = \mu_X - \mu_Y$
 - $V(\bar{D}) = \frac{\sigma_D^2}{n}$, ただし、 $\sigma_D^2 = V(D_i) = \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}$

対標本における検定(8)

– 仮説

$$\bullet \begin{cases} H_0: \mu_X - \mu_Y = 0 \\ H_1: \mu_X - \mu_Y \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0: \mu_D = 0 \\ H_1: \mu_D \neq 0 \end{cases}$$

– 検定統計量

$$\bullet |T| = \frac{\bar{D}}{\sqrt{\frac{S_D^2}{n}}}$$

$$- S_D^2 = \frac{1}{n-1} \sum_{i=1}^n (D_i - \bar{D})^2 = S_X^2 + S_Y^2 - 2S_{XY}$$

– 両側検定

• $|T| > 1.96$ のとき、帰無仮説 H_0 を棄却する。

– 正規母集団を仮定できるときは、1.96 を $t_{0.025}(n-1)$ に代える。

対標本における検定(9)

– Pearson の親子の身長データ(続き)

- $\bar{D}_{obs} = -1.3, S_{D_{obs}}^2 = 7.59$
- $|T|_{obs} = 2.11$
- 正規母集団を仮定すると、自由度19のt分布の上側0.025点は2.09
- H_0 は棄却される。
 - 父親の平均身長と息子の平均身長には差がある。

母分散の検定(1)

- 正規母集団

- $X_i \sim_{iid} N(\mu, \sigma^2) \quad i = 1, 2, \dots, n.$

- 標本分散

- $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

- 標本分散の性質

- $E(S^2) = \sigma^2$

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

母分散の検定(2)

- 母分散の値に関する検定

- 仮説

- $\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 \neq \sigma_0^2 \end{cases}$ ただし、 σ_0^2 は具体的な正の値。

- 検定統計量

- $$U = \frac{(n-1)S^2}{\sigma_0^2}$$

- 両側検定(有意水準0.05)

- $U < \chi_{0.025}^2(n-1)$ または $U > \chi_{0.975}^2(n-1)$ なら H_0 を棄却する。

母分散の検定(3)

－ 例：

- Pearson データから、父親の身長に関する大きさ10の無作為標本を得る。

－ 65, 67, 70, 64, 68, 66, 67, 66, 62, 64

» 教科書の数値例と異なることに注意する。

- 仮説

$$- \begin{cases} H_0: \sigma^2 = 7.5 \\ H_1: \sigma^2 \neq 7.5 \end{cases}$$

- 臨界点

$$- \chi_{0.025}^2(10 - 1) = 2.7004, \chi_{0.975}^2(10 - 1) = 19.0228$$

母分散の検定(4)

- 検定統計量の観察値

$$- U_{obs} = \frac{(10-1) \times 5.21}{7.5} = 6.25$$

- 両側検定(有意水準0.05)
 - H_0 は棄却されない。

母分散の検定(5)

- 分散比の検定

- 2標本問題

- 標本A $X_i \sim iid N(\mu_A, \sigma_A^2) \quad i = 1, 2, \dots, n.$
 - 標本B $Y_j \sim iid N(\mu_B, \sigma_B^2) \quad j = 1, 2, \dots, m.$
 - 2つの標本は独立とする。

- 2つの標本分散の性質

- $\frac{(n-1)S_A^2}{\sigma_A^2} \sim \chi^2(n-1)$ ただし、 $S_A^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$
 - $\frac{(m-1)S_B^2}{\sigma_B^2} \sim \chi^2(m-1)$ ただし、 $S_B^2 = \frac{1}{m-1} \sum_{i=1}^m (Y_i - \bar{Y})^2$

母分散の検定(6)

- 2つの標本が独立なので

$$- F = \frac{S_A^2/\sigma_A^2}{S_B^2/\sigma_B^2} \sim F(n-1, m-1)$$

– 仮説

$$- \begin{cases} H_0: \sigma_A^2 = \sigma_B^2 \\ H_1: \sigma_A^2 \neq \sigma_B^2 \end{cases} \Leftrightarrow \begin{cases} H_0: \sigma_A^2 / \sigma_B^2 = 1 \\ H_1: \sigma_A^2 / \sigma_B^2 \neq 1 \end{cases}$$

– 検定統計量

$$\bullet F = \frac{S_A^2}{S_B^2}$$

母分散の検定(7)

－ 両側検定 (有意水準0.05)

- $F < F_{0.025}(n-1, m-1)$ または
 $F > F_{0.975}(n-1, m-1)$ のとき H_0 を棄却する。

母分散の検定(7)

－ 例：

- Pearson データから
 - － 父親の身長が無作為標本(大きさ20)
 - » 64, 65, 69, 68, 69, 60, 71, 67, 72, 69, 70, 66, 69, 68, 71, 66, 73, 67, 68, 72.
 - － 息子の身長が無作為標本(大きさ20)
 - » 68, 69, 73, 70, 70, 66, 69, 72, 71, 65, 70, 72, 73, 68, 70, 68, 69, 70, 63, 72.
 - － 両者は独立に標本抽出されている。
 - » 教科書の数値例と異なることに注意。
 - － 臨界点
 - » $F < 0.40$ または $F > 2.53$.

母分散の検定(7)

- 検定統計量

- $F_{obs} = \frac{S_{A_{obs}}^2}{S_{B_{obs}}^2} = \frac{9.54}{6.78} = 1.41$

- 両側検定(有意水準0.05)

- H_0 は棄却されない。