講義にて配布した資料の抜粋・縮小版です



平成22年度
統計工学　講義資料(8)

～統計的推論：仮説検定Ⅱ～

(C) Sadami SUZUKI



## 2つの母集団比較：t検定(対応なし)

### Step1：仮説の設定

- 問題意識の明確化
- 考察の対象となる母集団特性値に関する仮説の設定

店舗Xの店長は最近出店してきた競合店Yと比較して主婦層のウケが悪いように感じている・・・

**「ウチの顧客満足度は店舗Yに比べてどうだろうか？」**

検定仮説

(C) Sadami SUZUKI



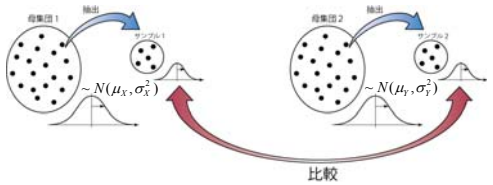
## 2つの母集団比較：t検定(対応なし)

### Step2：データ収集および基本統計量の計算

- 仮説に応じたデータの収集(母集団からの標本抽出)
- 基本統計量の計算

検定仮説

$$H_0 : \mu_X = \mu_Y$$




**対象となる母集団と収集すべきデータは？**

(C) Sadami SUZUKI



## 2つの母集団比較：t検定(対応なし)

### Step3：検定に用いる統計量の選択とその分布に関する検討

検定仮説

$$H_0 : \mu_X = \mu_Y$$

- 考察の対象：**2つの母集団平均値の差**

$$\hat{\mu}_X - \hat{\mu}_Y = \bar{X} - \bar{Y} = \frac{\sum X_i}{n} - \frac{\sum Y_i}{m}$$

- 2つの正規母集団平均値の差の分布：

平均値

分散

(C) Sadami SUZUKI



## 2つの母集団比較：t検定(対応なし)

### Step3：検定に用いる統計量の選択とその分布に関する検討

検定仮説

$$H_0 : \mu_X = \mu_Y$$

- 考察の対象：**2つの母集団平均値の差**

$$\hat{\mu}_X - \hat{\mu}_Y = \bar{X} - \bar{Y} = \frac{\sum X_i}{n} - \frac{\sum Y_i}{m}$$

- 2つの正規母集団平均値の差の分布：

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

標準正規分布に従う

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

$$\hat{\sigma}_Y^2 = \frac{1}{m-1}\sum_{j=1}^{m}(Y_j - \bar{Y})^2$$

$$T' = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\hat{\sigma}_X^2}{n} + \frac{\hat{\sigma}_Y^2}{m}}}$$

・・・2つの不偏分散に影響された分布

(C) Sadami SUZUKI



## 2つの母集団比較：t検定(対応なし)

### Step3：検定に用いる統計量の選択とその分布に関する検討

検定仮説

$$H_0 : \mu_X = \mu_Y$$

- **2つの正規母集団平均値の差**の分布：

**2つの母集団の等分散性**を仮定する

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

**2つの標本集団をプールして分散を推定すると・・・**

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{j=1}^{m}(Y_j - \bar{Y})^2}{(n-1)+(m-1)} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{j=1}^{m}(Y_j - \bar{Y})^2}{n+m-2}$$

・・・プールされた(不偏)分散

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\hat{\sigma}^2}}$$

自由度 $n+m-2$ の $t$ 分布に従う

(C) Sadami SUZUKI

講義にて配布した資料の抜粋・縮小版です



## 2つの母集団比較：t検定（対応なし）

**検定仮説**

$$H_0 : \mu_X = \mu_Y$$

### Step4：検定量を用いて仮説が正しいか否かを判定する

- 帰無仮説が成り立つとき

$$T_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\hat{\sigma}^2}}$$

自由度 $n+m-2$ の $t$ 分布に従う

- 得られた標本を基に検定量の計算 $t_0 = \frac{\bar{x} - \bar{y}}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right)\hat{\sigma}^2}}$

$$|t_0| > t_{n+m-2}(\alpha)$$ となる確率が$\alpha$

**有意水準$\alpha$で仮説は棄却される**

(C) Sadami SUZUKI



## 分散の検定：$\chi^2$検定

### Step1：仮説の設定

- 平均値の検定のときと同様に考えると・・・

**帰無仮説**

$$H_0 : \sigma^2 = \sigma_0^2$$

**対立仮説**

$H_1 : \sigma^2 < \sigma_0^2$
$H_1 : \sigma^2 > \sigma_0^2$ ・・・片側検定

$H_1 : \sigma^2 \neq \sigma_0^2$ ・・・両側検定

### Step2：データ収集および基本統計量の計算

(C) Sadami SUZUKI



## 分散の検定：$\chi^2$検定

**検定仮説**

$$H_0 : \sigma^2 = \sigma_0^2$$

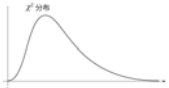
### Step3：検定に用いる統計量の選択とその分布に関する検討

- 考察の対象：**母分散**
- 検定量：母分散の**不偏推定量**としての**不偏分散**をベースに考える

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

- 分布：標本平均からの偏差平方和と母分散との比が **$\chi^2$分布**

$$\chi^2 = \frac{(n-1)\hat{\sigma}^2}{\sigma^2} = \frac{\sum(X_i - \bar{X})^2}{\sigma^2} \equiv \frac{S_X}{\sigma^2}$$




自由度 $n$-1 の $\chi^2$ 分布に従う

(C) Sadami SUZUKI



## 分散の検定：$\chi^2$検定

**検定仮説**

$$H_0 : \sigma^2 = \sigma_0^2$$

### Step4：検定量を用いて仮説が正しいか否かを判定する

- 帰無仮説が成り立つとき

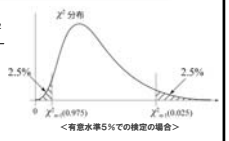
$$\chi_0^2 = \frac{(n-1)\hat{\sigma}^2}{\sigma_0^2} = \frac{\sum(X_i - \bar{X})^2}{\sigma_0^2} \equiv \frac{S_X}{\sigma_0^2}$$

自由度 $n$-1 の $\chi^2$ 分布に従う

- 得られた標本を基に検定量の計算

$$\chi_0^2 = \frac{(n-1)\hat{\sigma}^2}{\sigma_0^2} = \frac{\sum(x_i - \bar{x})^2}{\sigma_0^2}$$

$$\chi_0^2 > \chi_{n-1}^2\left(\frac{\alpha}{2}\right) \text{ or } \chi_0^2 < \chi_{n-1}^2\left(1 - \frac{\alpha}{2}\right)$$ となる確率が$\alpha$




＜有意水準5%での検定の場合＞

**有意水準$\alpha$で仮説は棄却される**

(C) Sadami SUZUKI



## 等分散性の検定：$F$検定

**検定仮説**

$$H_0 : \sigma_X^2 = \sigma_Y^2$$

### Step3：検定に用いる統計量の選択とその分布に関する検討

- 考察の対象：**2つの母分散の比**
- 検定量：母分散の**不偏推定量**としての**不偏分散**をベースに考える

$$\hat{\sigma}_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \qquad \hat{\sigma}_X^2 = \frac{1}{m-1}\sum_{i=1}^{m}(Y_i - \bar{Y})^2$$

$$\chi_X^2 = \frac{(n-1)\hat{\sigma}_X^2}{\sigma_X^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma_X^2} = \frac{S_X}{\sigma_X^2} \qquad \chi_Y^2 = \frac{(m-1)\hat{\sigma}_Y^2}{\sigma_Y^2} = \frac{\sum_{i=1}^{m}(Y_i - \bar{Y})^2}{\sigma_Y^2} = \frac{S_Y}{\sigma_Y^2}$$

自由度 $n$-1 の $\chi^2$ 分布に従う　　自由度 $m$-1 の $\chi^2$ 分布に従う

- 分布：それぞれの不偏分散の比が**$F$分布**に従う

$$F = \frac{\chi_Y^2/(m-1)}{\chi_X^2/(n-1)} = \frac{\hat{\sigma}_Y^2/\sigma_Y^2}{\hat{\sigma}_X^2/\sigma_X^2} = \frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2} \cdot \frac{\sigma_X^2}{\sigma_Y^2}$$

自由度 $m$-1, $n$-1 の $F$ 分布に従う

(C) Sadami SUZUKI



## 等分散性の検定：$F$検定

**検定仮説**

$$H_0 : \sigma_X^2 = \sigma_Y^2$$

### Step4：検定量を用いて仮説が正しいか否かを判定する

- 帰無仮説が成り立つとき

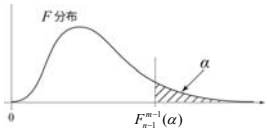
$$F_0 = \frac{\hat{\sigma}_Y^2}{\hat{\sigma}_X^2} = \frac{S_Y/(m-1)}{S_X/(n-1)}$$

自由度 $m$-1, $n$-1 の $F$ 分布に従う

- 得られた標本を基に検定量の計算

$\hat{\sigma}_Y^2 > \hat{\sigma}_X^2$としてF値を求め，片側検定を行うものとすると…

$$F_0 > F_{n-1}^{m-1}(\alpha)$$ となる確率が$\alpha$




**有意水準$\alpha$で仮説は棄却される**

(C) Sadami SUZUKI