

## 第 3 章 統計学の復習

James H. Stock 著・Mark W. Watson 著・宮尾 龍蔵 訳  
『入門計量経済学』（共立出版、2016）

<https://www.kyoritsu-pub.co.jp/book/b10003746.html>

2025-08-25

## 統計学の基本的な役割

- ▶ 現実世界の問題に答えるためには、**母集団 (population)** の分布の性質を知る必要がある。
  - ▶ 例: 大学卒業生の平均賃金はいくらなのか、男性と女性で異なるのか、その違いはどれほどなのか、といった問題である。
- ▶ しかし、母集団全体を調査するには **膨大なコストと時間**がかかる。
  - ▶ 例: 2000 年の米国国勢調査では 100 億ドルもの費用がかかり、10 年の期間を要した。
- ▶ 統計学の基本的な役割は、母集団から**無作為に (randomly) 標本 (sample)** を選ぶことで、その母集団の分布の性質を見出すことである。
  - ▶ 限られた標本から、**統計的な推論 (statistical inference)** に基づく結論を導き出すことが可能となる。

## 計量経済学で利用する3つの統計的手法

計量経済学では、主に以下の3つのタイプの統計的手法を利用する。

### 1. 推定 (estimation) :

- ▶ 母集団の分布の特性を表す値 (例えば平均) について「**最善の推測 (best guess)**」を求める作業である。

### 2. 仮説検定 (hypothesis testing) :

- ▶ 母集団に関して特定の仮説を設定し、そして標本を使ってその仮説が正しいかどうかを判断することである。

### 3. 信頼区間 (confidence interval) :

- ▶ データを使って未知の母集団の性質に関して、区間や範囲を推定することである。

これらの手法は、未知の母集団の平均に関する推測、仮説検定、信頼区間の構築、さらには2つ以上の変数間の関係や異なる母集団の比較といった経済学の問題で頻繁に議論される。

## 3.1 母集団の平均の推定 (Estimation of Population Mean)

### 推定量と推定値

- ▶ **推定量 (estimator) :**
  - ▶ 母集団からランダムに抽出される標本データの関数として表現される。
  - ▶ 無作為抽出という標本の性質から **確率変数 (random variable)** となる。
  - ▶ 例: 標本平均  $\bar{Y}$  や、最初の観測値  $Y_1$  は母集団の平均  $\mu_Y$  の推定量である。
- ▶ **推定値 (estimate) :**
  - ▶ 特定の標本から実際に計算された推定量の値である。
  - ▶ 非確率な値となる。

## 推定量の望ましい性質 (Desirable Properties of Estimators)

複数の推定量が考えられる場合、ある推定量が別の推定量と比べて「より良い」といえるために、以下の3つの望ましい性質が挙げられる。

### 1. 不偏性 (unbiasedness) :

- ▶ 推定量が不偏であるとは、無作為な標本抽出を何度も繰り返したとき、平均的には正しい答えが得られることである。
- ▶ 数学的には、 $E(\hat{\rho}_Y) = \mu_Y$  のとき、推定量  $\hat{\rho}_Y$  は不偏 (unbiased) であると呼ばれる。
- ▶ そうでなければ、偏り (bias) を持つ (biased) といわれる。

# 一貫性

## 2. 一貫性 (consistency) :

- ▶ 標本数が大きいとき、推定量のランダムな変動に関する不確実性が非常に小さいことである。
- ▶ より正確には、推定量  $\hat{\rho}_Y$  が真の値  $\mu_Y$  を中心とした小さな範囲に収まる確率が、**標本数が大きくなるにつれて1に近づく**性質である。
- ▶ 数学的には、 $\hat{\rho}_Y \xrightarrow{p} \mu_Y$  のとき、 $\hat{\rho}_Y$  は一貫性 (consistency) を満たすといわれる。

## 効率性

### 3. 効率性 (efficiency) :

- ▶ 複数の不偏推定量がある場合、**標本分布の分散 (variance) が最も小さい推定量**が「より良い」とされる。
- ▶ 分散が小さい推定量は、データに含まれる情報をより効率的に使うという特徴を持つ。
- ▶  $\text{var}(\hat{\rho}_Y) < \text{var}(\tilde{\rho}_Y)$  であれば、 $\hat{\rho}_Y$  は  $\tilde{\rho}_Y$  よりも効率的 (efficient) と呼ばれる。

## 標本平均の性質 (Properties of the Sample Mean)

標本平均  $\bar{Y}$  は、母集団の平均  $\mu_Y$  を推定する自然な方法である。

▶ **不偏性 (unbiasedness) :**

▶  $E(\bar{Y}) = \mu_Y$  なので、 $\bar{Y}$  は  $\mu_Y$  の不偏推定量である。

▶ **一貫性 (consistency) :**

▶ 大数の法則 (Law of Large Numbers) より、 $\bar{Y} \xrightarrow{p} \mu_Y$  なので、 $\bar{Y}$  は一貫性を満たす。

▶ **効率性 (efficiency) :**

▶ 観測値  $Y_1, \dots, Y_n$  の加重平均に基づくすべての不偏推定量のなかで、 $\bar{Y}$  は最も効率的である。

▶ すなわち、 $\bar{Y}$  は **最良な線形不偏推定量 (Best Linear Unbiased Estimator, BLUE)** である。

▶  $n \geq 2$  の場合、 $\text{var}(\bar{Y}) = \sigma_Y^2/n$  は  $Y_1$  の分散  $\sigma_Y^2$  よりも小さい。

## 最小二乗推定量としての標本平均 (Sample Mean as a Least Squares Estimator)

標本平均  $\bar{Y}$  は、観測値との差の二乗の平均がすべての推定量のなかで最小となり、その意味でデータに最も良くフィットする推定量である。

- ▶ 以下の二乗和を最小にする推定量  $\hat{\mu}_Y$  は、**最小二乗推定量 (least squares estimator)** と呼ばれる。

$$\sum_{i=1}^n (Y_i - \hat{\mu}_Y)^2$$

- ▶ この二乗和を最小化する  $\hat{\mu}_Y$  は、 $\bar{Y}$  であることが代数的に示される。
  - ▶ 予測誤差の二乗和を最小化するために微分し、それをゼロとすることで  $\hat{\mu}_Y = \bar{Y}$  となる。

## 無作為な標本抽出の重要性 (Importance of Random Sampling)

- ▶ 標本がランダムに抽出されたものでなければ、標本平均  $\bar{Y}$  は **バイアス (bias)** を持ってしまう。
  - ▶ 例: 公園のベンチに座っている勤労年齢の大人をインタビューして失業率を推計すると、失業者に偏った標本となり、推計値にバイアスが生じる。
  - ▶ 歴史的例: 1936 年の米国大統領選挙での「ガゼット紙」の予測ミスは、電話帳と自動車保有者登録ファイルから抽出したサンプルが、**富裕層 (共和党支持者)** に偏っていたためである。この偏ったサンプルが推定量にバイアスをもたらし、大きな誤りを犯した。
- ▶ 標本の選択方法を考える際、このようなバイアスをできるだけ小さくすることが非常に大切である。

## 3.2 母集団の平均に関する仮説検定

### 仮説検定の基本

- ▶ 仮説検定 (hypothesis test) は、母集団の分布に関する特定の仮説について、標本に基づく証拠から答えを出す手法である。
- ▶ 統計学における仮説検定の出発点は、テストされる仮説をまず特定化することである。

## 帰無仮説 (Null Hypothesis)

- ▶ **帰無仮説 (null hypothesis,  $H_0$ ) :**
  - ▶ 母集団の平均  $E(Y)$  がある特定の値  $\mu_{Y,0}$  と等しい、と設定される仮説である。
  - ▶ 例:  $H_0 : E(Y) = 20$  (大卒者の賃金は母集団の平均で見て時間当たり 20 ドル)。
  - ▶ 数学的表現:  $H_0 : E(Y) = \mu_{Y,0}$ 。

## 対立仮説 (Alternative Hypothesis)

- ▶ **対立仮説 (alternative hypothesis,  $H_1$ )** :
  - ▶ 帰無仮説が誤りの場合に成立する仮説である。
  - ▶ 最も一般的なものは、**両側の対立仮説 (two-sided alternative hypothesis)** である。
    - ▶  $H_1 : E(Y) \neq \mu_{Y,0}$  ( $E(Y)$  は  $\mu_{Y,0}$  より大きい値と小さい値の両方を含む)。
  - ▶ **片側の対立仮説 (one-sided alternative hypothesis)** を設定することもできる。
    - ▶ 例:  $H_1 : E(Y) > \mu_{Y,0}$  (大卒者の方が賃金が高い)。

## 帰無仮説の棄却

- ▶ 統計分析者にとっての問題は、無作為な標本を用いて、 $H_0$  を採択 (accept) するのか、あるいはそれを却下 (reject) して  $H_1$  を採択するのか、判断を下すことである。
  - ▶ 帰無仮説を「採択する」というよりも、「**棄却できない (fail to reject)**」と理解されることが多い。

## p 値 (p-value)

▶ **p 値 (p-value)**、あるいは **有意確率 (significance probability)** とは:

- ▶ 帰無仮説が正しいという仮定の下で、統計量と帰無仮説の値との距離が、手元の標本から測った距離よりも離れてしまう（つまり帰無仮説の成立がより難しくなる）そのような統計量を引き出す確率のことである。
- ▶ 数学的表現:

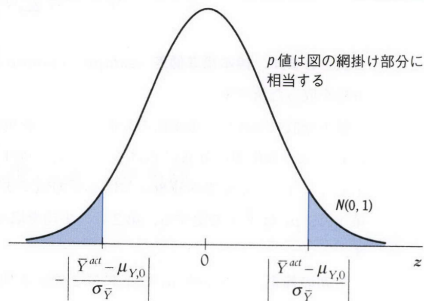
$$p \text{ 値} = \Pr_{H_0} [ |\bar{Y} - \mu_{Y,0}| > |\bar{Y}^{\text{act}} - \mu_{Y,0}| ]$$

ここで  $\bar{Y}^{\text{act}}$  は手元の標本から得られた標本平均である。

## 図 3.1

図 3.1  $p$  値の求め方

$p$  値は、次のような  $\bar{Y}$  を引き出す確率のことである。すなわち、 $\bar{Y}$  と真の平均  $\mu_{Y,0}$  との差が、 $\bar{Y}^{act}$  以上となる、そのような  $\bar{Y}$  の値を抽出する確率を表す。大標本の場合には、 $\bar{Y}$  は帰無仮説の下で、 $N(\mu_{Y,0}, \sigma_{\bar{Y}}^2)$  に従い、 $(\bar{Y} - \mu_{Y,0})/\sigma_{\bar{Y}}$  は  $N(0, 1)$  に従う。したがって  $p$  値は、図でいえば標準正規分布の両端で、 $\pm|\bar{Y}^{act} - \mu_{Y,0}|/\sigma_{\bar{Y}}$  より外側の領域（網掛け部分）に当たる。



## p 値の解釈

- ▶ p 値が小さい（例: 0.5%）場合: 帰無仮説の下でこのようなサンプルを得る可能性は非常に低い。したがって、**帰無仮説は正しくない**と推論してよいことになる。
- ▶ p 値が大きい（例: 40%）場合: 帰無仮説を否定する証拠としては弱い。無作為なサンプリングがもたらすばらつきにより真の平均とは異なる標本が抽出されたと推論され、**帰無仮説は棄却できない**と判断される。

## p 値の計算に必要な情報

- ▶ p 値を計算するには、帰無仮説の下での標本平均  $\bar{Y}$  の標本分布を求める必要がある。
  - ▶ 標本数が多い場合、**中心極限定理 (Central Limit Theorem)** により  $\bar{Y}$  の標本分布は正規分布 (Normal distribution) に近似できる。
    - ▶  $H_0$  の下で  $\bar{Y}$  は  $N(\mu_{Y,0}, \sigma_{\bar{Y}}^2)$  に従う ( $\sigma_{\bar{Y}}^2 = \sigma_Y^2/n$ )。

## $\sigma_Y$ が既知のときの p 値の計算

- ▶ 標本数が大きい場合、 $H_0$  の下で  $\bar{Y}$  の標本分布は  $N(\mu_{Y,0}, \sigma_{\bar{Y}}^2)$  に従う。
- ▶ したがって、基準化された  $\bar{Y}$ 、すなわち  $(\bar{Y} - \mu_{Y,0})/\sigma_{\bar{Y}}$  は標準正規分布 (**standard normal distribution**) に従う。
- ▶ p 値は、標準正規分布において  $\pm(\bar{Y}^{\text{act}} - \mu_{Y,0})/\sigma_{\bar{Y}}$  を超える両端の領域に相当する。

$$\text{p 値} = \Pr_{H_0} \left[ \left| \frac{\bar{Y} - \mu_{Y,0}}{\sigma_{\bar{Y}}} \right| > \left| \frac{\bar{Y}^{\text{act}} - \mu_{Y,0}}{\sigma_{\bar{Y}}} \right| \right] = 2\Phi \left( - \left| \frac{\bar{Y}^{\text{act}} - \mu_{Y,0}}{\sigma_{\bar{Y}}} \right| \right)$$

ここで  $\Phi$  は標準正規分布の累積密度関数 (cumulative density function) である。

## 標本分散、標本標準偏差、標準誤差

- ▶ 母集団の分散  $\sigma_Y^2$  は通常未知である。
- ▶ **標本分散 (sample variance)**  $s_Y^2$ : 母集団の分散  $\sigma_Y^2$  の推定量である。

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

- ▶  $n-1$  で割るのは、 $\mu_Y$  を  $\bar{Y}$  で推計することで生じる微小なバイアスを修正し、不偏性を得るためである (**自由度 (degrees of freedom) の修正**)。
- ▶ 標本分散は **母集団の分散の一致推定量 (consistent estimator)** である:  $s_Y^2 \xrightarrow{p} \sigma_Y^2$ 。

## 標本標準偏差と標準誤差

- ▶ **標本標準偏差 (sample standard deviation)**  $s_Y$ : 標本分散に平方根を取った値である。
- ▶ **標本平均  $\bar{Y}$  の標準誤差 (standard error,  $SE(\bar{Y})$ )**:  $\bar{Y}$  の標本分布の標準偏差  $\sigma_{\bar{Y}} = \sigma_Y/\sqrt{n}$  の推定量である。
  - ▶  $SE(\bar{Y}) = \hat{\sigma}_{\bar{Y}} = s_Y/\sqrt{n}$  と表記される。

## $\sigma_Y$ が未知のときの p 値の計算

- ▶  $s_Y^2$  は  $\sigma_Y^2$  の一致推定量なので、p 値は (3.6) 式で  $\sigma_Y$  を標準誤差  $SE(\bar{Y})$  に置き換えれば計算できる。

$$\text{p 値} = 2\Phi\left(-\left|\frac{\bar{Y}^{\text{act}} - \mu_{Y,0}}{SE(\bar{Y})}\right|\right)$$

- ▶ この標準化された標本平均  $(\bar{Y} - \mu_{Y,0})/SE(\bar{Y})$  は、**t 統計量 (t-statistic)** または **t 比率 (t-ratio)** と呼ばれる。

$$t = \frac{\bar{Y} - \mu_{Y,0}}{SE(\bar{Y})}$$

## p 値の計算における t 分布の役割

- ▶ 標本数  $n$  が大きい場合、t 統計量は近似的に標準正規分布  $N(0, 1)$  に従う。
  - ▶ p 値の計算は  $2\Phi(-|t^{\text{act}}|)$  となる。
  - ▶ 例:  $n = 200$ ,  $\bar{Y}^{\text{act}} = \$22.64$ ,  $s_Y = \$18.14$  の場合、 $\bar{Y}$  の標準誤差は  $s_Y/\sqrt{n} = 18.14/\sqrt{200} \approx 1.28$  となる。t 統計量の値は  $t^{\text{act}} = (22.64 - 20)/1.28 = 2.06$  である。
  - ▶ このとき、p 値は  $2\Phi(-2.06) \approx 0.039$  (3.9%) となる。

## 特定の有意水準を使った仮説検定

- ▶ **有意水準 (significance level)  $\alpha$** : 帰無仮説が実際には正しいのに、誤って棄却してしまう確率 (**タイプIエラー (type I error)**) を犯す確率) である。
  - ▶ 一般的に5%が用いられることが多い。
- ▶ **p値**が有意水準を下回るとき、帰無仮説を棄却する。
  - ▶ 例: 5%水準で棄却するシンプルなルールは、「 $|t^{\text{act}}| > 1.96$  ならば  $H_0$  を棄却する」である。
    - ▶ これは、標準正規分布の両端の領域が  $\pm 1.96$  の外側で合計5%となるためである。
- ▶ **t統計量の絶対値が臨界値 (critical value)** を上回れば帰無仮説を棄却することになる。

## 帰無仮説の棄却と採択

- ▶ 帰無仮説を棄却する統計量の範囲を **棄却域 (rejection region)**、棄却しない範囲を **採択域 (acceptance region)** と呼ぶ。
- ▶ **タイプ II エラー (type II error)** : 帰無仮説が実際には誤りであるのに、棄却すべきなのに棄却しない誤りである。
- ▶ **テストのサイズ (size of a test)** : 帰無仮説が正しいときに実際に誤って棄却してしまう確率である。

## 有意水準の選択に関する考慮事項

- ▶ **テストのパワー (power of a test)** : 対立仮説が正しい (帰無仮説が誤りである) ときに、実際正しく帰無仮説を棄却する確率である。
- ▶  $p$  値は、帰無仮説を棄却できる最も小さな有意水準に相当する。
- ▶ 有意水準の設定はトレードオフを伴う。
  - ▶ 低い有意水準 (例: 1%や 0.1%) はタイプ I エラーを避けるが、誤った帰無仮説を棄却するパワーが低下してしまう。
  - ▶ 経済学や政策問題に応用する際には、多くの場合、法廷のケースほど控えめな有意水準は必要とされず、5%水準が「ほどよい水準」とみなされることが多い。

## 片側の対立仮説 (One-Sided Alternative Hypothesis)

- ▶ 対立仮説が特定の方向に限定される場合 (例:  
 $H_1 : E(Y) > \mu_{Y,0}$ )。
- ▶ p 値の求め方や仮説検定の方法は両側テストと基本的に同じだが、t 統計量の絶対値ではなく、**特定の方向の値が大きいときにのみ帰無仮説を棄却する**点が異なる。
- ▶ p 値は、標準正規分布において、求められた t 統計量より右側の領域に当たる。

$$p \text{ 値} = \Pr_{H_0}(Z > t^{\text{act}}) = 1 - \Phi(t^{\text{act}})$$

- ▶ 片側 5%有意水準の臨界値は 1.645 である。

## 3.3 母集団の平均に関する信頼区間

### 信頼集合と信頼区間

- ▶ ランダム・サンプリングに伴うばらつきのため、一つのサンプル情報から真の平均の値を正確に知ることは不可能である。
- ▶ **信頼集合 (confidence set)** : 無作為に選ばれた標本から、真の平均  $\mu_Y$  をある一定の確率で含む範囲 (集合) を求めることである。
- ▶ **信頼水準 (confidence level)** : その集合内に  $\mu_Y$  が含まれる確率のことである。
- ▶ **信頼区間 (confidence interval)** : 母集団の平均に対する信頼集合は、平均に関するある下限から上限までに取りうるすべての値に相当するため、区間となる。

## 信頼区間の導出方法

- ▶ 母集団の平均  $\mu_Y$  が取りうるすべての値について、 $H_0 : E(Y) = \mu_{Y,0}$  をテストし、 $H_1 : E(Y) \neq \mu_{Y,0}$  に対して棄却されない  $\mu_{Y,0}$  の値をリストアップすることで信頼集合が得られる。
  - ▶ このリストに母集団の平均の真の値が含まれる確率は、設定した信頼水準（例: 95%）と等しくなる。
- ▶ より簡単な方法として、t 統計量の公式を利用する。
  - ▶  $H_0 : E(Y) = \mu_{Y,0}$  が 5% 水準で棄却されないのは、
$$\left| \frac{\bar{Y} - \mu_{Y,0}}{SE(\bar{Y})} \right| \leq 1.96$$
 のときである。
  - ▶ これを変形すると、
$$\bar{Y} - 1.96SE(\bar{Y}) \leq \mu_Y \leq \bar{Y} + 1.96SE(\bar{Y})$$
 となる。

## 信頼区間の公式

標本数が大きいとき、母集団の平均  $\mu_Y$  に対する信頼区間は以下のようになる。

- ▶ **95%信頼区間:**  $\bar{Y} \pm 1.96SE(\bar{Y})$
- ▶ **90%信頼区間:**  $\bar{Y} \pm 1.64SE(\bar{Y})$
- ▶ **99%信頼区間:**  $\bar{Y} \pm 2.58SE(\bar{Y})$

例:  $n = 200$ ,  $\bar{Y} = \$22.64$ ,  $SE(\bar{Y}) = 1.28$  の場合、95%信頼区間は  $(22.64 \pm 1.96 \times 1.28) = (\$20.13, \$25.15)$  となる。

## カバー確率

- ▶ 母集団の平均に関する信頼区間の **カバー確率 (coverage probability)** とは、標本抽出が繰り返される下で計算される、真の平均を含む確率のことである。
- ▶ 95%信頼区間は、ランダムに抽出されたすべてのサンプルの95%で、 $\mu_Y$  の真の値を含む区間を表す。

## 3.4 異なる母集団の平均の比較

### 2つの平均の差に関する仮説検定

- ▶ 2つの異なる母集団の平均（例: 大卒男性と女性の平均賃金  $\mu_M, \mu_W$ ）の差を比較する必要がある。
- ▶ 帰無仮説と両側の対立仮説は以下のように表される。
  - ▶  $H_0 : \mu_M - \mu_W = d_0$
  - ▶  $H_1 : \mu_M - \mu_W \neq d_0$
  - ▶ 男性と女性で母集団の平均賃金が同じであるという仮説は、 $d_0 = 0$  と置けばよいことがわかる。
- ▶  $\mu_M - \mu_W$  の推定量は、それぞれの標本平均の差  $\bar{Y}_M - \bar{Y}_W$  である。
- ▶ 中心極限定理と独立な標本抽出の仮定から、 $\bar{Y}_M - \bar{Y}_W$  は近似的に正規分布に従う。

## 標準誤差と t 統計量

- ▶  $\bar{Y}_M - \bar{Y}_W$  の **標準誤差 (standard error)** は以下で表される。

$$SE(\bar{Y}_M - \bar{Y}_W) = \sqrt{\frac{s_M^2}{n_M} + \frac{s_W^2}{n_W}}$$

ここで  $s_M^2, s_W^2$  は各グループの標本分散、 $n_M, n_W$  は各グループの標本数である。

- ▶ 帰無仮説をテストする **t 統計量 (t-statistic)** は、推定量から帰無仮説で設定されている差  $d_0$  を差し引き、その標準誤差で割ることで求められる。

$$t = \frac{(\bar{Y}_M - \bar{Y}_W) - d_0}{SE(\bar{Y}_M - \bar{Y}_W)}$$

- ▶  $n_M, n_W$  がともに大きければ、この t 統計量は標準正規分布に従う。p 値は両側テストの場合、(3.14) 式で求められる。

## 2つの平均の差に関する信頼区間

- ▶ 平均の差  $\mu_M - \mu_W$  に対する 95%両側の信頼区間は、 $(\bar{Y}_M - \bar{Y}_W)$  の  $\pm 1.96$  標準誤差の範囲に入ることになる。

$$(\bar{Y}_M - \bar{Y}_W) \pm 1.96SE(\bar{Y}_M - \bar{Y}_W)$$

例: 2004年の米国における大卒者の賃金男女格差について、95%信頼区間は (\$2.91, \$4.12) と推定された。

## 3.5 実験データに基づく因果関係の効果の推定

### 実験

- ▶ ランダムにコントロールされた実験 (**randomly controlled experiment**) では、対象者がトリートメント・グループ (treatment group) かコントロール・グループ (control group) に無作為に割り振られる。
- ▶ **因果関係の効果 (causal effect)** あるいは**処置の効果 (treatment effect)** は、2つの条件付期待値の差で表現される。
  - ▶  $E(Y|X = 1) - E(Y|X = 0)$  (トリートメントを受けたグループと受けなかったグループの平均の差)。

## 因果関係の効果

- ▶ この因果関係の効果は、トリートメント・グループとコントロール・グループの**標本平均の差によって推定できる**。
- ▶ 「その処置に効果はない」という仮説は、2つの平均が等しいという仮説と同一であり、前述の2つの平均を比較する  $t$  統計量 (3.20) 式を使ってテストできる。
- ▶ 2グループの平均の差に関する 95%信頼区間は、因果関係の効果に関する 95%信頼区間に等しい。

## 自然実験と準実験

- ▶ **自然実験 (natural experiments) または準実験 (quasi-experiments)** は、経済学で用いられることが多い。これは、あたかもランダムにコントロールされた実験であるかのような状況を分析するものである。
  - ▶ 例: 401(k) プランのデフォルト選択肢を「非加入」から「加入」へと変更した企業を分析し、自動加入が加入率を大幅に向上させるという因果関係の効果を推定した研究。

## 3.6 標本数が小さい場合の t 統計量

### 小標本とスチューデント t 分布

- ▶ 3.2 節から 3.5 節まででは、**中心極限定理**に基づき、t 統計量の分布が標準正規分布に近似されることを前提としてきた。
- ▶ しかし、**標本数が小さい (small sample size)** 場合には、標準正規分布では t 統計量の分布をうまく近似できない。
- ▶ もし **母集団の分布自体が正規分布 (normal distribution)** に従うならば、1 つの母集団の平均をテストする t 統計量の正確な分布は、**自由度 (degrees of freedom)  $n - 1$  のスチューデント t 分布 (Student's t-distribution)** となる。

## t 統計量

- ▶ t 統計量は、以下のように定義される。

$$t = \frac{\bar{Y} - \mu_{Y,0}}{s_Y/\sqrt{n}}$$

- ▶ スチューデント t 分布は  $Z/\sqrt{W/(n-1)}$  の分布で定義され、 $Z$  は標準正規分布、 $W$  は自由度  $n-1$  のカイ二乗分布に従う確率変数である。
- ▶ スチューデント t 分布からの臨界値は、標準正規分布の臨界値よりも少し幅が広がる。

## 平均の差をテストする t 統計量の修正版

- ▶ 2つの平均の差をテストする t 統計量 (3.20) 式は、たとえ母集団の分布が正規分布であったとしても、スチューデント t 分布には従わない。
- ▶ 「プールされた」標準誤差 (**pooled standard error**) をベースにした修正版の t 統計量は、2つのグループが同じ分散を持つ（あるいは各グループの観測値の数が同じ）という特別なケースに限定される場合にスチューデント t 分布に従う。

## プールされた分散の推定量

- ▶ プールされた分散の推定量

$$s_{\text{pooled}}^2 = \frac{1}{n_M + n_W - 2} \left( \sum_{i=1}^{n_M} (Y_{Mi} - \bar{Y}_M)^2 + \sum_{j=1}^{n_W} (Y_{Wj} - \bar{Y}_W)^2 \right)$$

- ▶ プールされた標準誤差

$$SE_{\text{pooled}}(\bar{Y}_M - \bar{Y}_W) = s_{\text{pooled}} \times \sqrt{\frac{1}{n_M} + \frac{1}{n_W}}$$

- ▶ しかし、2つの母集団の分散が異なる場合、プールされた分散推定量にはバイアスが生まれ、一致性が保証されない。

## スチューデント t 分布の実際の利用 (1)

- ▶ 経済変数 (economic variables) が正規分布に従うことは例外的である (例: 2004 年米国における賃金の分布、ウォール街の最悪の日)。
- ▶ 標本数が大きい場合、t 統計量の分布を正規分布で近似することは妥当である。
- ▶ 2つの平均を比較する際、2つのグループで平均が異なることに経済学的な理由がある場合には、2つのグループで分散が異なることも十分に考えられる。そのときは、プールされた標準誤差の式は不適切であり、正しい標準誤差はグループごとに異なる分散を容認する (3.19) 式のような表現となる。
- ▶ したがって、経済学の実証分析では、**大標本の標準正規分布への近似に依拠することが多い。**

## スチューデント t 分布の実際の利用 (2)

- ▶ 実際、標本数が  $n > 15$  のとき、スチューデント t 分布と標準正規分布で計算される p 値の違いは無視できるほどわずかである ( $> 0.01$  を超えない)。現代のほとんどの実証研究において標本数は数百もしくは数千と十分大きく、スチューデント t 分布と標準正規分布の違いは無視できる。

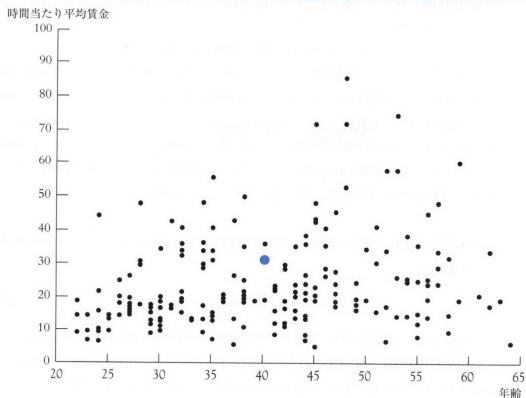
## 3.7 散布図、標本共分散、標本相関

### 散布図

- ▶ **散布図 (scatterplot)** は、2つの変数  $X$  と  $Y$  の  $n$  個の観測値  $(X_i, Y_i)$  を  $X$  軸、 $Y$  軸平面上の点としてプロットした図である。
- ▶ 変数間の関係を視覚的に捉えることができる。
  - ▶ 例: 情報産業に勤務する技術者の年齢と時間当たり賃金の散布図から、年長の技術者の方が若い技術者よりも賃金が多いという正の関係が見て取れる。

## 図 3-2

図 3.2 平均賃金と年齢の散布図



図の各点は、標本に含まれる 200 人の就業者それぞれの年齢と平均賃金を表す。青色の点は、年齢 40 歳で時間当たり賃金が 31.25 ドルの就業者に対応する。データは情報産業に勤務する技術者に関するもので、2005 年 3 月の現代人口調査より取得。

## 標本共分散と標本相関

母集団の共分散や相関は未知だが、標本から推定できる。

- ▶ **標本共分散 (sample covariance)**  $s_{XY}$ :  $X$  と  $Y$  の線形関係の方向を示す。

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

- ▶ 標本分散と同様に  $n-1$  で割る。
- ▶ 標本共分散は母集団の共分散の一致推定量 (**consistent estimator**) である:  $s_{XY} \xrightarrow{p} \sigma_{XY}$

## 標本相関係数

- ▶ **標本相関係数 (sample correlation coefficient)** または **標本相関 (sample correlation)**  $r_{XY}$ :  $X$  と  $Y$  の線形関係の強さを示す。

$$r_{XY} = \frac{s_{XY}}{s_X s_Y}$$

ここで  $s_X, s_Y$  はそれぞれの標本標準偏差である。

- ▶ 単位がなく、 $-1$  から  $1$  の間の値をとる:  $|r_{XY}| \leq 1$ 。
- ▶  $r_{XY} = 1$  は強い正の線形関係、 $r_{XY} = -1$  は強い負の線形関係、 $r_{XY} = 0$  は線形関係がないことを示す。
- ▶ 標本相関係数も**一貫性 (consistency)**を持つ。

## 相関と線形関係 (Correlation and Linear Relationship)

- ▶ 標本相関係数は **線形関係 (linear relationship)** の指標である。
- ▶ 相関がゼロであっても、非線形な関係が存在する場合がある。
  - ▶ 例: Xが増加するにつれてYが当初増加し、その後減少するような関係の場合、XとYの関係は顕著だが、相関はゼロになりうる。
  - ▶ このような場合は散布図を見ることで関係性を把握できる。

## まとめ

1. **標本平均  $\bar{Y}$** : 母集団の平均  $\mu_Y$  の不偏、一致、効率的な推定量であり、大標本では正規分布に近似される。
2. **t 統計量**: 母集団の平均が特定の値を取るという帰無仮説をテストする際に用いられ、大標本では帰無仮説が正しい下で標準正規分布に従う。
3. **p 値**: 帰無仮説が正しいという仮定の下で、観察された統計量よりも極端な値が得られる確率であり、帰無仮説を棄却できる最小の有意水準に相当する。
4. **信頼区間**: 繰り返し抽出される標本の 95% (例えば) で、真の母集団平均を含む区間となるよう構築される。
5. **2つの母集団の平均の比較**: 1つの母集団の平均に関する仮説検定と信頼区間と概念的に同様である。
6. **標本相関係数**: 2変数間の線形関係の強さを測定する母集団相関係数の推定量である。