2013年度 森林統計学

第13回 7月9日 平均値の差の検定
講義資料

高知大学農学部　森林科学コース
鈴木保志

復習

「検定」の復習

- 帰無仮説と対立仮説
  - H_0とH_1
- 2種類の過誤
  - αとβ
- 平均値の検定
  - 検定のやり方

復習

帰無仮説と対立仮説

- 帰無仮説 H_0
  - 検定する仮説。結果が「有意」でなければ採択される。
  - 例)目だった差がない、薬品などの効果がない
- 対立仮説 H_1
  - 結果が「有意」のとき、採択される。
  - 例) 目だった差がある、薬品などの効果がある

※ H は hypothesis (仮説) の頭文字

復習

2種類の過誤

	検定の結果 有意でない	検定の結果 有意
真実(誰にもわからない)	H_0を採択	H_1を採択
H_0が真	正しい (確率：$1-\alpha$)	✕ 第1種の過誤 (確率：α=有意水準)
H_1が真	✕ 第2種の過誤 (確率：β)	正しい (確率：$1-\beta$=検出力)

(教科書p.159, 表1 に加筆)

- 第1種の過誤をしてしまう確率はαでコントロールできる
  - α: 有意確率 (“$1-\alpha$” は「推定」での信頼確率と同じ)
- 第2種の過誤の確率は神のみぞ知る
  - αを小さくしすぎなければ、βも大きくならないことがわかっている

2種類の過誤: 教科書p.160～の頭蓋骨の例

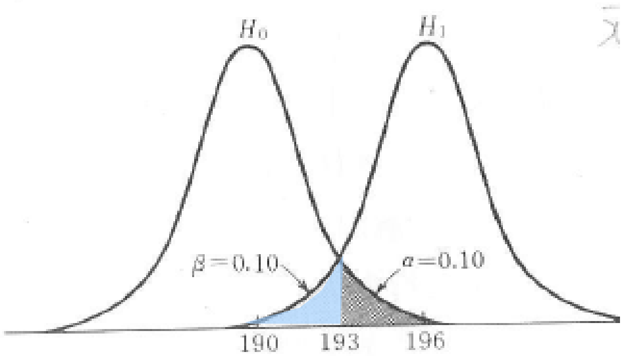


図 1　H_0 と H_1 のもとでの $\bar{x}$ の分布

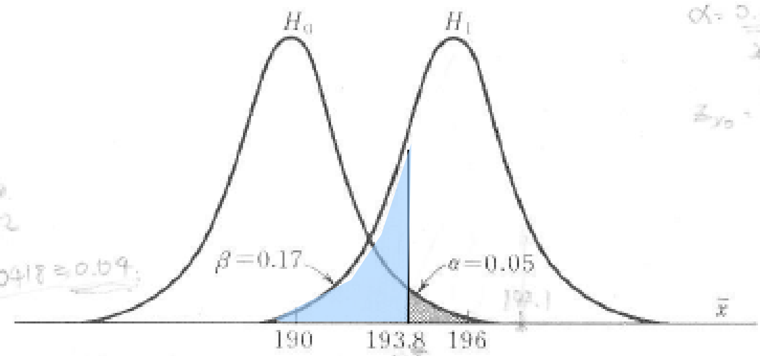


図 2　選ばれた棄却域をもつ，H_0 と H_1 のもとでの $\bar{x}$ の分布

- 左: α=0.10 としたとき、この例では β=0.10
- 右: αをより小さく、0.05 とすると、βは0.17 と大きくなる
- αとβはトレードオフの関係

復習

平均値(1つの平均)の検定

- 教科書 例2.(p.166)の適性検査の例で復習
- 新入生の平均点118 (n = 50) は、これまでの学生の平均点μ=115 (σ=20) と同じくらいといえるか?
  - $H_0 : \mu$=115
  - $H_1 : \mu$≠115 [両側検定]
  - $\bar{x}$ =118, $\sigma_{\bar{x}}$ =20 /√50 = 2.8
- 方法①: 信頼区間内に $\bar{x}$ があるか?
  - α=0.05 とすると z =1.960（表IVから）

正規分布表の見方

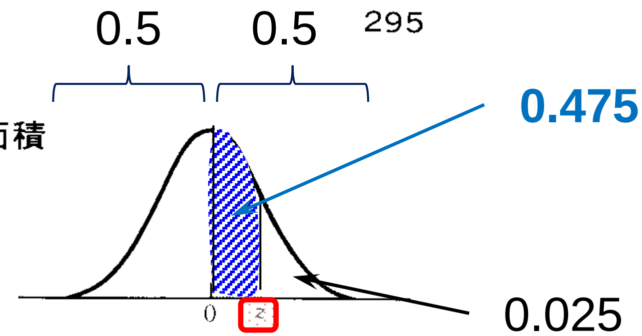
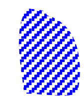


表 IV　標準正規分布の面積

表の中の数字は $z=0$ から z の正値までの曲線下の部分の面積である．z の負値に対する面積は対称性を利用して求めればよい．

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0										
0.1										
0.2										
0.3										
0.4										
0.5										
0.6										
0.7										
0.8										
0.9										
1.0										
1.1										
1.2										
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857

- 両側の棄却域が5%となるzは?
- 片側は2.5% (0.025)
- 正規分布は左右対称
-  の部分は 0.5 − 0.025 = 0.475
- z は 1.96

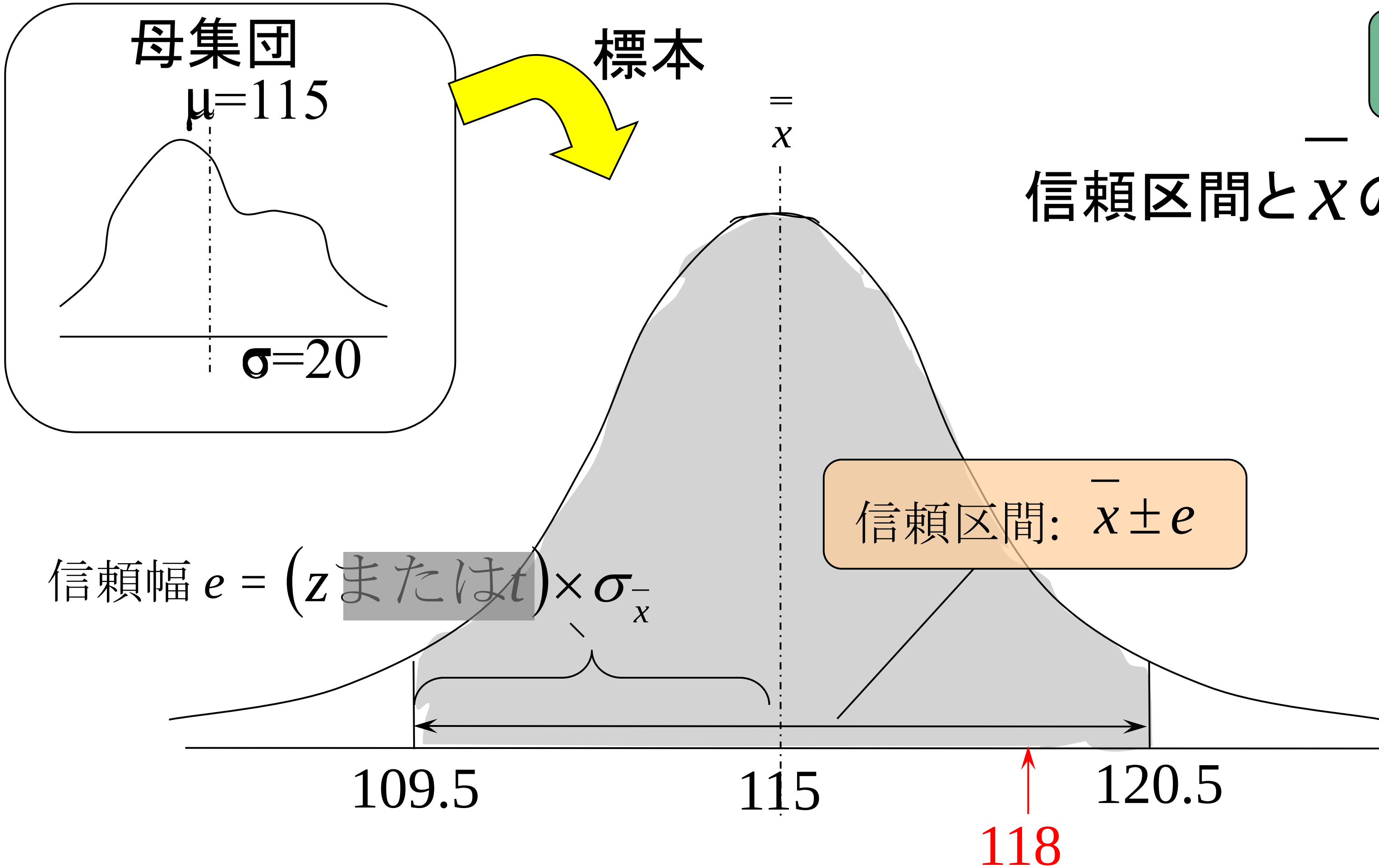
復習

平均値(1つの平均)の検定

- 教科書 例2.(p.166)の適性検査の例で復習
- 新入生の平均点118 (n = 50) は、これまでの学生の平均点μ=115 (σ=20) と同じくらいといえるか?
  - $H_0 : \mu$=115
  - $H_1 : \mu$≠115 [両側検定]
  - $\bar{x}$ =118, $\sigma_{\bar{x}}$ =20 /√50 = 2.8
- 方法①: 信頼区間内に $\bar{x}$ があるか?
  - α=0.05 とすると z =1.960（表IVから）
  - 信頼区間 115±1.960×2.8 = 115±5.5=109.5～120.5

復習

信頼区間と $\bar{X}$ の分布



- 標本(n=50)がμ=115, σ=20の母集団から採られたら、その標本平均 $\bar{x}$ は95%の確率で109.5～120.5の間の値をとる。
- 実際の $\bar{x}$ は118 で信頼区間内

→ 有意でない (H_0は棄却されない)

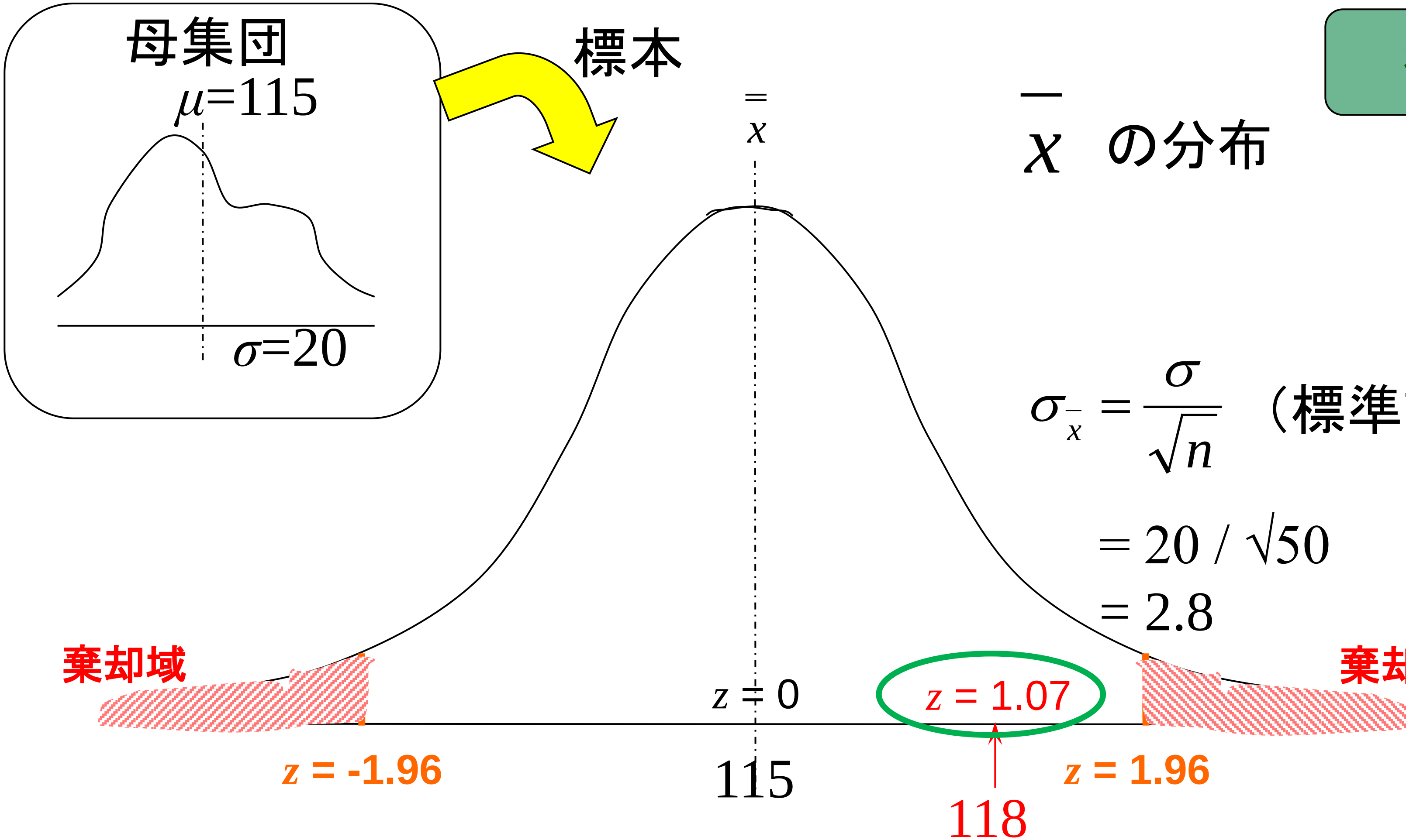
復習

平均値(1つの平均)の検定

- 教科書 例2.(p.166)の適性検査の例
- 新入生の平均点118 (n = 50) は、これまでの学生の平均点μ=115 (σ=20) と同じくらいといえるか?
  - $H_0: \mu$=115
  - $H_1: \mu$≠115 [両側検定]
  - $\bar{x}$=118, $\sigma_{\bar{x}}$ =20 /√50 = 2.8
- 方法②: $\bar{x}$ の z は閾値のz (1.96) より大きいか?
  - 118を標準化すると z'= (118 – 115) / 2.8 = 1.07

復習

$\bar{x}$ の分布



$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$ （標準誤差）

$= 20 / \sqrt{50}$

$= 2.8$

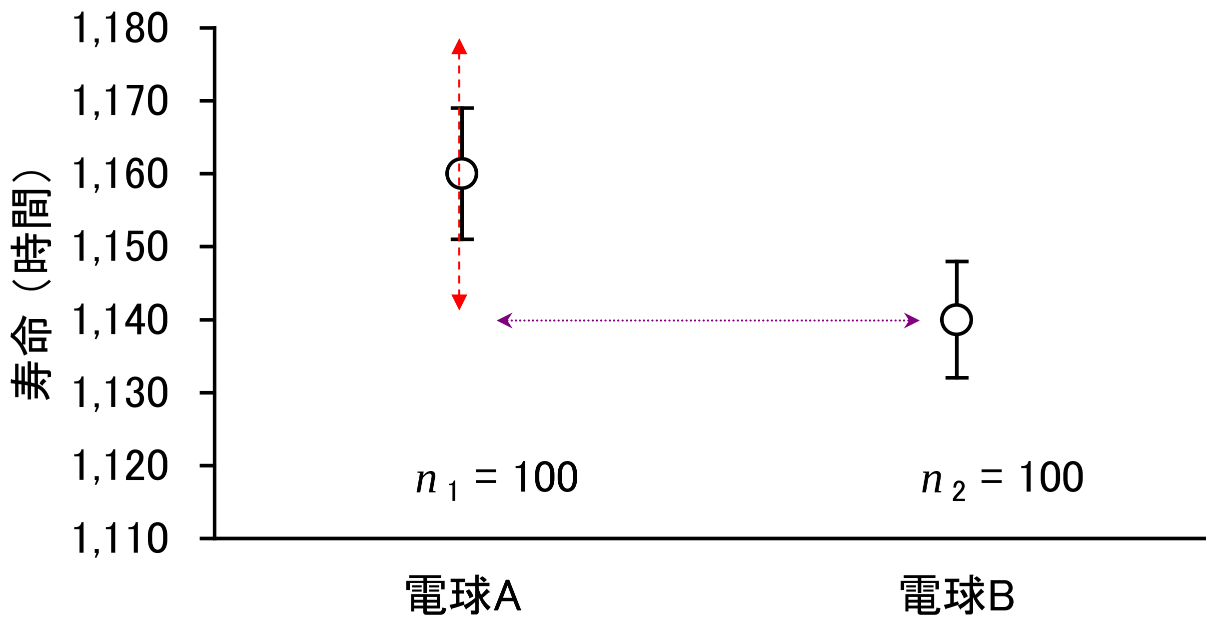
- 118 は 平均値の115から標準偏差の1.07倍離れている
- α=0.05 とすると z =1.96（表IVから）　＞ 1.07
- 118 はα=0.05 の危険率による閾値内
  - 差は<u>有意でない</u> (H_0は棄却されない)

ここからが今日の本題

2つの平均値の差の検定

- 教科書p.172～の例
- 電球Aと, 電球Bの寿命には差があるか?
  - 電球A: n_1=100, s_1=90, $\overline{x}_1$=1160
  - 電球B: n_2=100, s_2=80, $\overline{x}_2$=1140
- まず推定(信頼区間)の考え方で考察してみると…
  - 電球Aの標準誤差: $s_{\overline{x}1} = s_1 / \sqrt{n_1} = 90 / \sqrt{100} = 9.0$
  - 電球B: $s_{\overline{x}2} = s_2 / \sqrt{n_2} = 80 / \sqrt{100} = 8.0$

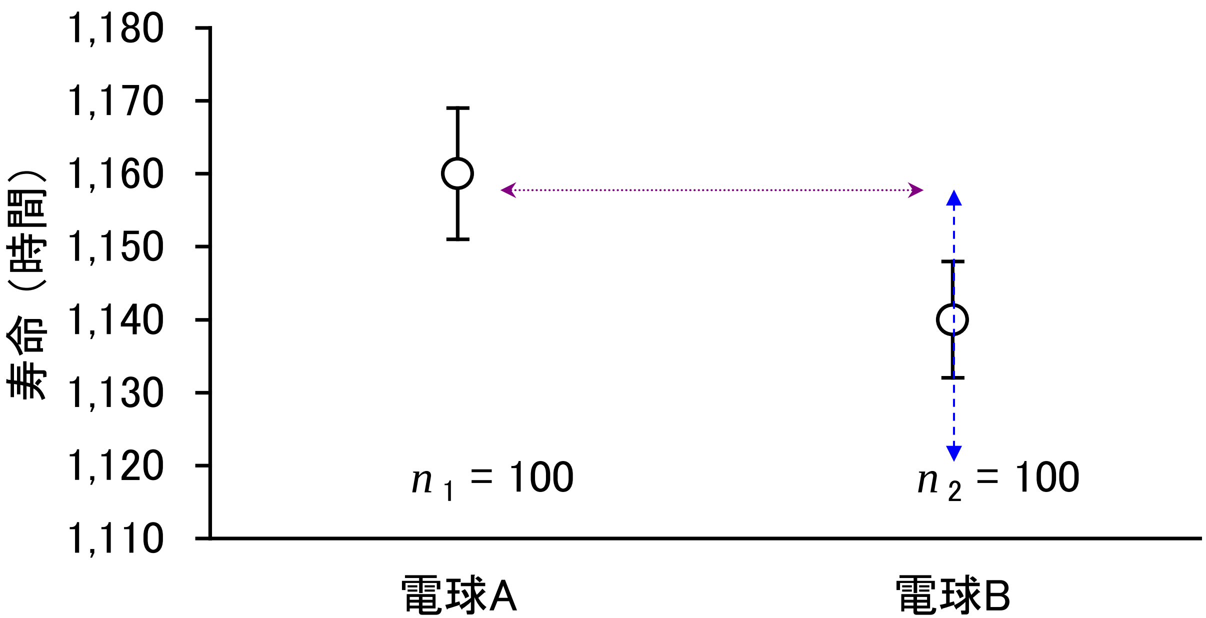
平均値±標準誤差の図



図a. 教科書p.172の例の平均値と標準誤差

- 95%信頼区間はおおむね標準誤差の2倍なので、その範囲に他方の平均値があるかどうかで有意差の有無の見当がつく

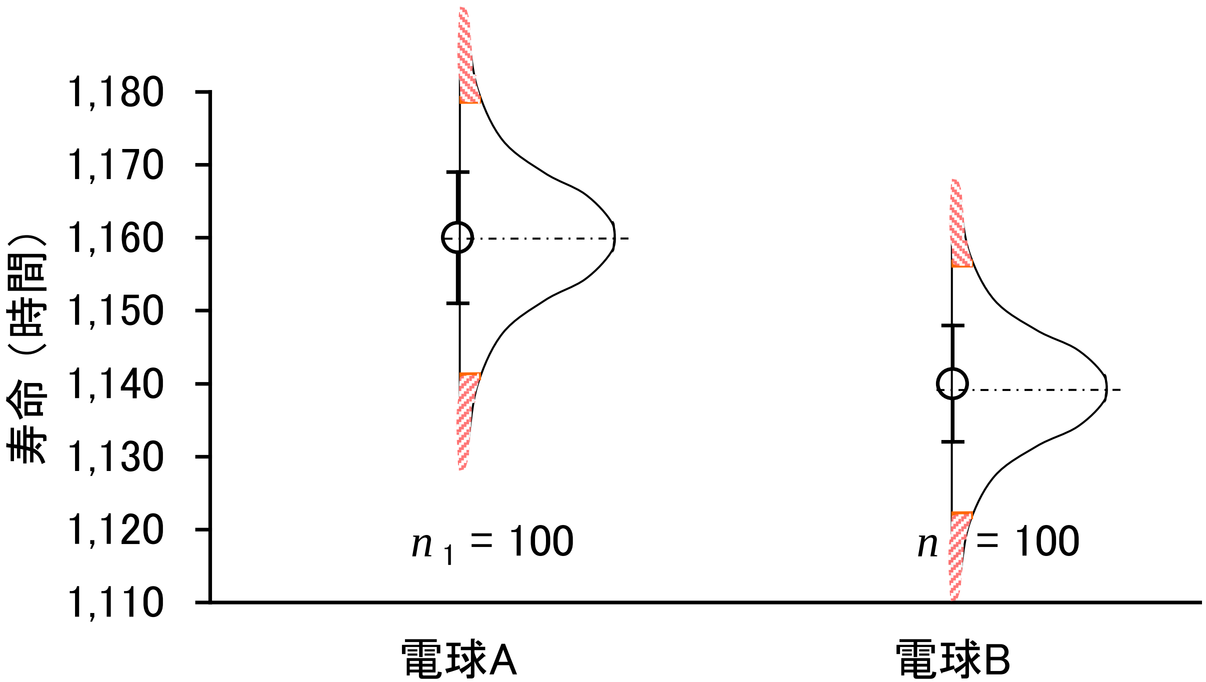
平均値±標準誤差の図



図a. 教科書p.172の例の平均値と標準誤差

- 平均値±標準誤差の図は、平均値の差の検定前の基礎情報として重要。
- ただし正確な結果のためには、検定をする必要がある。

2つの信頼区間で差を考えるイメージ



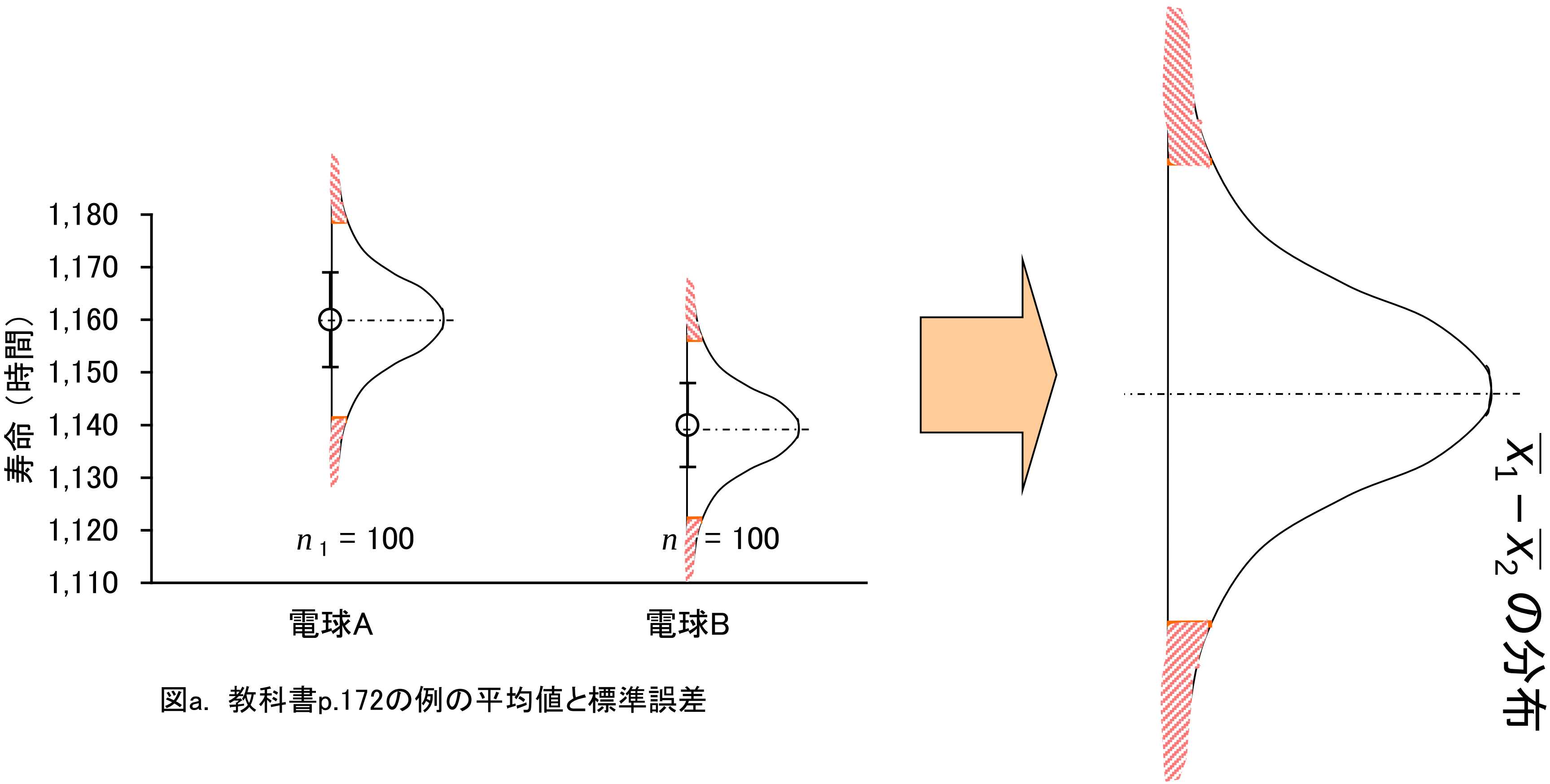
図a. 教科書p.172の例の平均値と標準誤差

- 1回の推定は0.95の信頼性
- 2つの信頼区間で考えると、実際には $0.95^2 = 0.90$ の信頼性での検定になってしまう

2つの平均値の差の検定
(教科書p.172～の例)

- 1回の推定は0.95の信頼性
- 2つの信頼区間で考えると、実際には $0.95^2 = 0.90$ の信頼性での検定になってしまう
- <u>「平均値の差」という新しい分布を作る</u>と、1変数で検定できる。

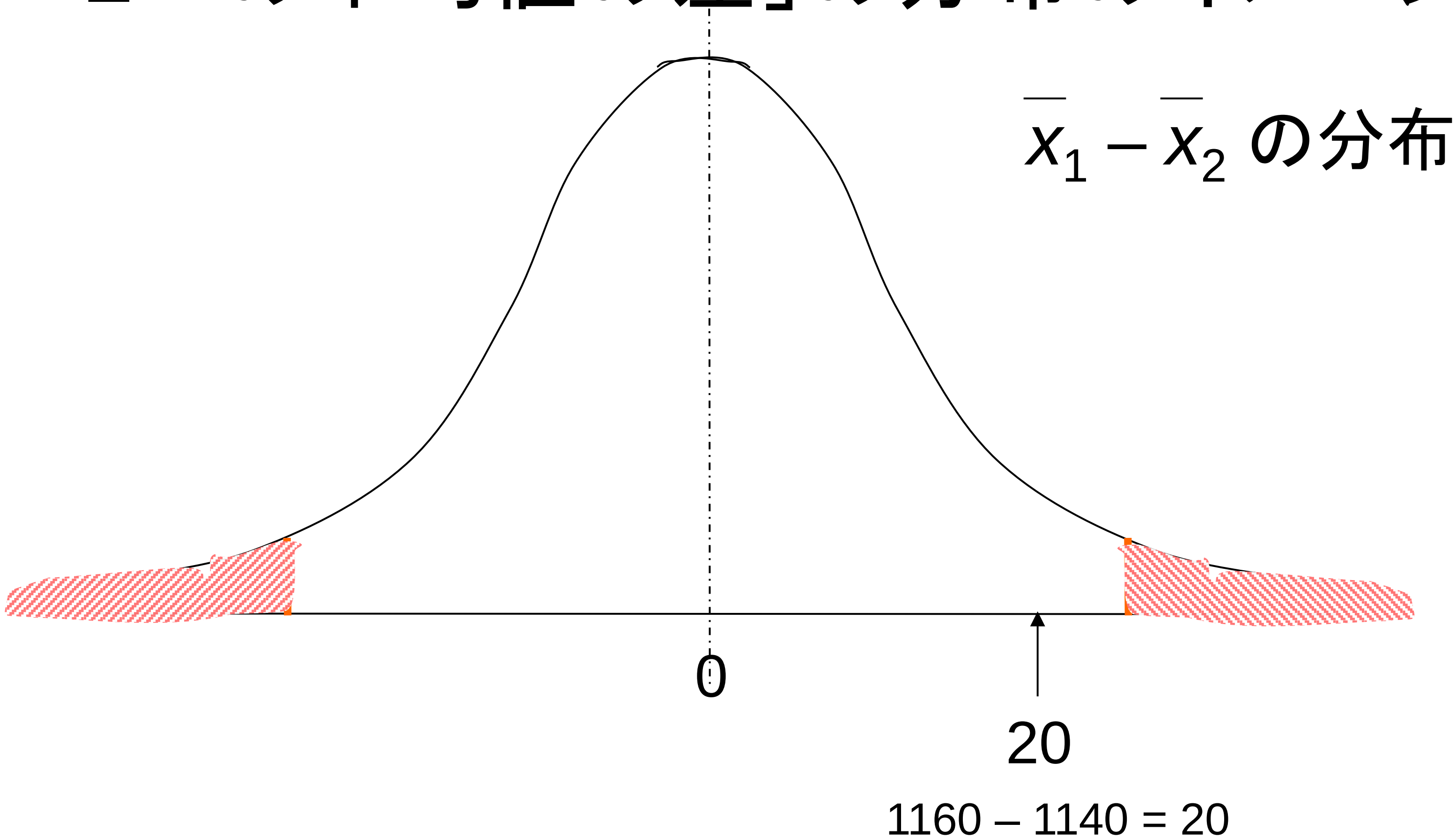
「2つの平均値の差の分布」で考える



図a. 教科書p.172の例の平均値と標準誤差

- ひとつの分布になるので、これまでの1変数の検定の考え方が使えるようになる

「2つの平均値の差」の分布のイメージ



- 「2つの平均値の差」という新しい分布を考える
  - 差がなければ その平均は 0
  - 実際の平均値の差が、0に対して大きいか小さいか、を検定する

2つの平均値の差の検定
(教科書p.172～の例)

- 二つの標本平均の差$(\bar{x}_1 - \bar{x}_2)$という新しい変数を作ると、その標準偏差$\sigma_{\bar{x}_1-\bar{x}_2}$は

  $\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\dfrac{\sigma_1{}^2}{n_1} + \dfrac{\sigma_2{}^2}{n_2}}$ に従うことがわかっている（教科書 p.174；実際にはσはｓで置きかえる）。これから z を計算して判定する。

- 教科書p.173, (4) の定理

2つの平均値の差の検定
(教科書p.172～の例)

- 標準化の公式で考える。区別するためにアスタリスクをつける。（式番号は資料”Statistics'13_13a_docx.pdf”のもの）

$$z^{*} = \frac{x^{*} - \mu^{*}}{\sigma^{*}} \tag{1}$$

x*とμ *は平均値の差の検定の場合、「2 つの平均値の差」という分布のある値と母平均になるので

$$x^{*} = \bar{x}_1 - \bar{x}_2 \tag{2}$$

$$\mu^{*} = \mu_1 - \mu_2 = 0 \tag{3}$$

σ *は「2 つの平均値の差」という分布の標準偏差なので、σ を s で置きかえて

$$\sigma^{*} = \sqrt{{s_{\bar{x}.1}}^2 + {s_{\bar{x}.2}}^2} = \sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}} \tag{4}$$

2つの平均値の差の検定
(教科書p.172～の例)

まとめると

$$z^* = \frac{x^* - \mu^*}{\sigma^*} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_{\bar{x}.1}^{\ 2} + s_{\bar{x}.2}^{\ 2}}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^{\ 2}}{n_1} + \frac{s_2^{\ 2}}{n_2}}} \tag{1a}$$

このz*を計算して、閾値と比較する

計算結果

正規検定	
$\bar{x}.1 - \bar{x}.2$	20.0
$\sigma_{\bar{x}.1-\bar{x}.2}$	12.0
z	1.661
p値(両側)	9.67%

< 1.96

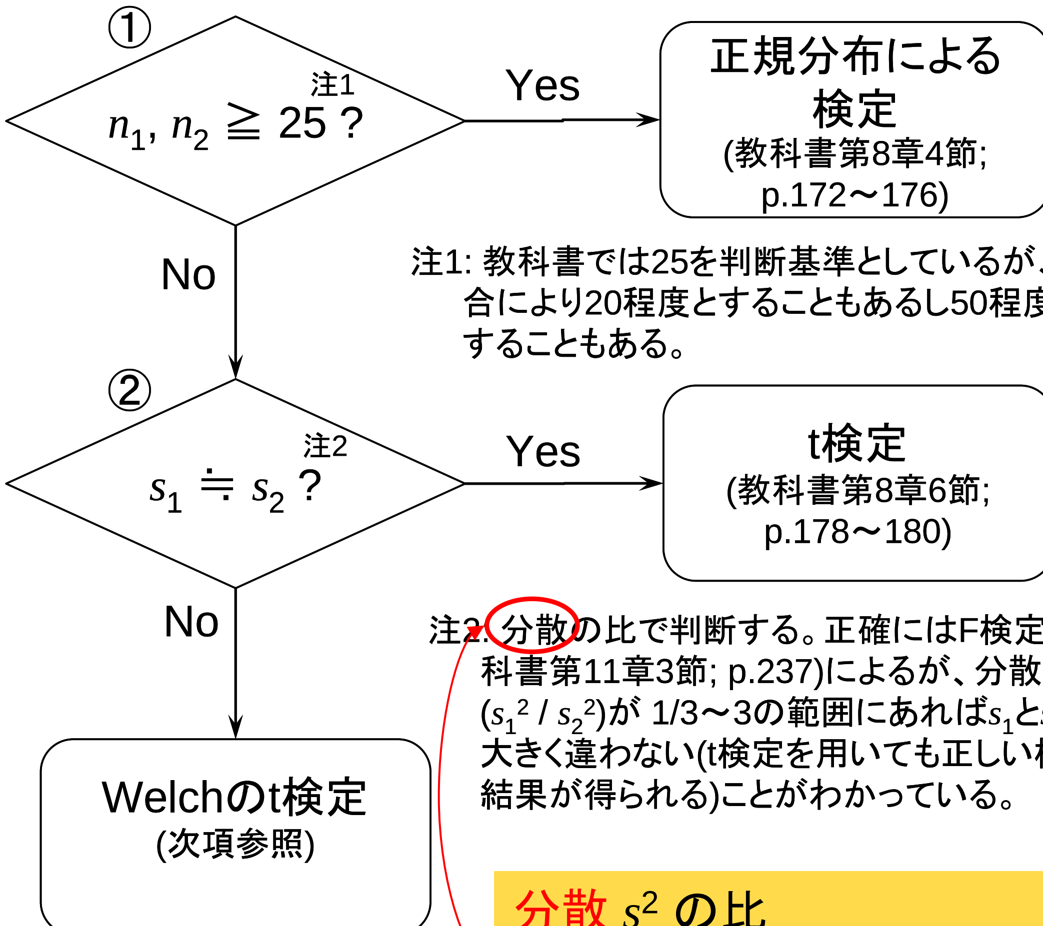
α =0.05 (両側) のzの閾値

- 検定結果は有意でない。(電球 A の寿命と電球 B の寿命に統計的に有意な差があるとは言えない。)

3つの検定方法と検定方法の選択

- 資料” Statistics'13_13a_docx.pdf”で説明
- [1] 正規分布による方法：nが十分大きいとき
- [2] t検定：nが大きくないが、2つの母標準偏差が同じとみなせるとき
- [3] Welchのt検定：nが大きくなく、2つの母標準偏差も同じとみなせないとき
- [1] > [2] > [3] の順に検出力は低くなるので、条件が許す限り[1]あるいは[2]を用いる方がよい

3つの検定方法と検定方法の選択



注1: 教科書では25を判断基準としているが、場合により20程度とすることもあるし50程度とすることもある。

注2: 分散の比で判断する。正確にはF検定(教科書第11章3節; p.237)によるが、分散比(s_1^2 / s_2^2)が 1/3～3の範囲にあればs_1とs_2は大きく違わない(t検定を用いても正しい検定結果が得られる)ことがわかっている。

分散 s^2 の比
(標準偏差 s の比ではない)
ことに注意

2つの平均値の差の検定

3. 検定

「2 つの平均値の差の検定」では基本的な手順は［1］・［2］・［3］どの方法でも同じである。標本平均の差が仮説で定められた母平均の差とどのくらい異なるか、を示す統計量 z^* を標準化の公式（教科書 p.105 ; (4)式）に従って算出する。z^*が定めた α（危険率・第 1 種の過誤の確率）に相当する閾値（［1］の場合正規分布表の z ;［2］・［3］の場合 t 分布表の t）より大きければ「有意差あり」、同じか小さければ「有意差なし」と判定する。

$$z^* = \frac{x^* - \mu^*}{\sigma^*} \tag{1}$$

ここで x^* と μ^* は［1］・［2］・［3］どの方法でも共通である。x^* は 2 つの標本平均の差

$$x^* = \bar{x}_1 - \bar{x}_2 \tag{2}$$

である。μ^* は 2 つの母平均の差であるが、「2 つの平均値の差の検定」では帰無仮説としてほとんどの場合「2 つの母平均は等しい」（$\mu_1 = \mu_2$）としているので、

$$\mu^* = \mu_1 - \mu_2 = 0 \tag{3}$$

となる。σ^* は「2 つの標本平均の差」の分布の標準偏差のことで、これは［1］・［2］・［3］それぞれの方法により異なる。

2つの平均値の差の検定

- [1] 正規分布による場合: p.172〜の例(先ほど説明)
- [2] t分布による場合: σ^* の計算方法が異なる

t 検定

教科書 p.179, (6)式で正規分布による検定の場合の z に相当する t 値を計算、表 V の t 値と比較。より大きければ仮説を棄却（有意）。

※ (6)の式は一見複雑だが、まず二つの標本分散をあわせた s_p^2

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

を計算して、これを正規分布の場合のように合成する

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

と考えると覚えやすい。

2つの平均値の差の検定

- [1] 正規分布による場合: p.172～の例(先ほど説明)
- [2] t分布による場合: σ^* の計算方法が異なる
- [3] Welchのt検定: 自由度ν の計算方法が異なる (σ^*の計算方法は[1]正規分布による場合　と同じ)

2つの平均値の差の検定の例

- 教科書 p.179, 例 2 のデータで説明。

n_1	11	n_2	39
$\bar{x}_1$	2.3	$\bar{x}_2$	5.2
s_1	1.0	s_2	2.7
$s_{\bar{x}1}$	0.3	$s_{\bar{x}2}$	0.4

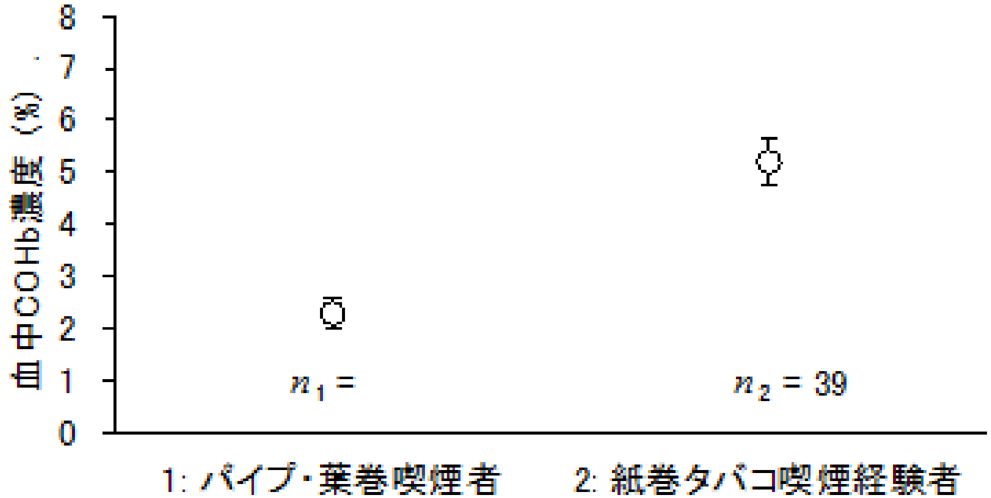


図-3a. 教科書p.179, 例2のデータの平均値と標準誤差

平均値±標準誤差 の図からは、有意差があることが伺える。

検定の手順

1. 平均値、標準偏差、標準誤差を算出
2. 平均値±標準誤差の図を作成
   - 検定に先立ち、視覚的に有意差の有無を考察する

ここまでやってから検定に入る

3. n と <u>分散比</u> から適切な検定手法を選択

再掲

3つの検定方法と検定方法の選択

①

$n_1, n_2 \geqq 25$?[注1]

Yes → 正規分布による検定
(教科書第8章4節; p.172～176)

No ↓

注1: 教科書では25を判断基準としているが、場合により20程度とすることもあるし50程度とすることもある。

②

$s_1 \fallingdotseq s_2$?[注2]

Yes → t検定
(教科書第8章6節; p.178～180)

No ↓

Welchのt検定
(次項参照)

注2: 分散の比で判断する。正確にはF検定(教科書第11章3節; p.237)によるが、分散比(s_1^2 / s_2^2)が 1/3～3の範囲にあればs_1とs_2は大きく違わない(t検定を用いても正しい検定結果が得られる)ことがわかっている。

分散 s^2 の比
(標準偏差 s の比ではない)
ことに注意

2つの平均値の差の検定の例

計算結果

	$\overline{x}.1 - \overline{x}.2$	−2.9
	$\sigma_{\overline{x}.1-\overline{x}.2}$	0.527

t*は、差を標準偏差で割って…

t*	−5.502

- 自由度 ν = 44.2
- 表V では、切捨て側で最も近い値をみる (この場合 40)

表Vを読む

α = 0.05 (両側)

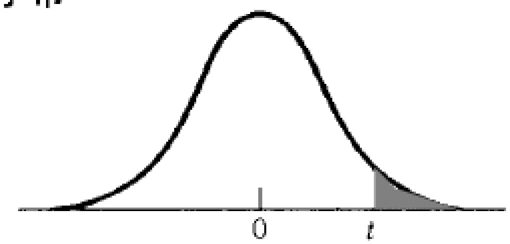
ν = 40

- n_2 = 39 で 25より大きいが、n_1 = 11 で 25より小さい
  - 正規分布による検定は不適
- $s_1 ≒ s_2$ か?
  - 分散比 (s_1^2とs_2^2の比)で判断
  - s_1^2 / s_2^2 = 1.02 / 2.72 = 0.137
  - 1/3 〜 3 の範囲にない (分散比が互いに3倍以上離れている)
  - t検定も不適
- Welchのt検定が適した方法

危険率 $P = \alpha / 2$
(両側検定)

表 V スチューデントの t 分布

表側の数字は自由度 (ν) を表わす．表頭の数字は t が表の中の数値を超える確率 (P) を表わす．負の t の値に対しては分布の対称性を利用すればよい．



自由度
ν(ニュー)
は 40

ν \ P	.10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
⋮	⋮	⋮	⋮	⋮	⋮
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

2つの平均値の差の検定の例

計算結果

	$\bar{x}.1 - \bar{x}.2$	−2.9
	$\sigma_{\bar{x}.1-\bar{x}.2}$	0.527

t*は、差を標準偏差で割って…

t*	−5.502

- 自由度ν = 44.2
- 表V では、<u>切捨て側で</u>最も近い値をみる (この場合 40)

- 閾値のt値は **2.021**
- t* = -5.502 で |t*| > t なので… **<u>有意差あり</u>** と判定

- n_2 = 39 で 25より大きいが、n_1 = 11 で 25より小さい
  - 正規分布による検定は不適
- $s_1 \fallingdotseq s_2$ か?
  - 分散比 (s_1^2とs_2^2の比)で判断
  - s_1^2 / s_2^2 = 1.02 / 2.72 = 0.137
  - 1/3 〜 3 の範囲にない (分散比が互いに3倍以上離れている)
  - *t*検定も不適
- Welchの*t*検定が適した方法

2つの平均値の差の検定の例

正規検定	
$\bar{x}_{.1} - \bar{x}_{.2}$	−2.9
$\sigma_{\bar{x}.1-\bar{x}.2}$	0.5
z	−5.502
p値(両側)	0.00%
t検定	
ν	48
s1^2/s2^2	0.137
sp	2.445
$\sigma_{\bar{x}.1-\bar{x}.2}$	0.835
t	−3.474
p値(両側)	0.11%
Welchのt検定	
ν	44.2
t	−5.502
p値(両側)	0.00%
ν	p値(両側)
44.0	0.00%
45.0	0.00%

p値は低いほど有意性が高い (5%以下で有意)

- 教科書p.179の例2では、適した検定方法は Welchのt検定
- 正規検定、t検定は不適だが当てはめてみると有意性は一般に

  正規検定 < t検定 < Welchのt検定

  の順となる
- つまり、この順に第1の過誤の危険性は小さくなるが、第2の過誤の危険性は高くなる
- <u>適切な検定方法を使うことが大切</u>

o 標本分散比が 1/3 以下なので、t 検定は不適。

o ν は $n_1 + n_2 - 2 = 48$ より少し小さくなっている。

課題

- 課題資料を参照。