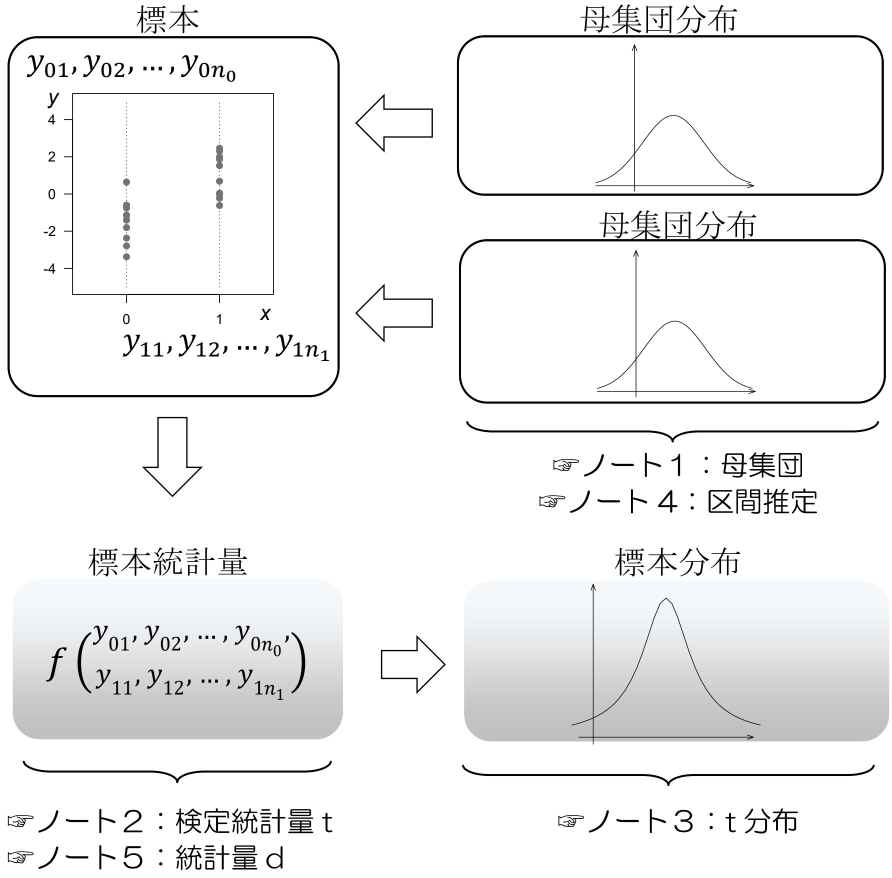
第3講　t 検定（対応なし）

学習の目標

- □ 二群の平均値の差の検定である「**対応のない t 検定**」のロジックがわかり、論文で出会った際に適切に解釈できる。
- □ t 検定を行う際に母集団に想定される仮定を理解できる。
- □ 仮説検定における二つの仮説（**帰無仮説**と**対立仮説**）が理解でき、適切に想定できる。
- □ 標本分布の標準偏差を**標本誤差**と呼び、**標本平均値差**が従う標本分布の標準誤差を$\sigma_{\bar{y}_1-\bar{y}_0}$と書くことが分かる。
- □ t 値という統計量は、標本平均値差とその標本誤差の推定量の**比の値**であることが分かる。
- □ 帰無仮説の下で t 値が従う標本分布を **t 分布**と言い、自由度が大きくなると標本正規分布に近づくことが分かる。
- □ t 検定では、データから得られた t 値が、**限界値**を超えて t 分布の**棄却域**に入るかどうかを検定する。棄却域に入った場合、帰無仮説という仮定が誤っていたと**棄却**する。
- □ t 値が棄却域に落ちる確率を**有意水準** α と呼び、帰無仮説が真の時、誤って棄却する**第一種の過誤**を表すことが分かる。
- □ 対立仮説が正しい時に t 値が従う分布を**非心 t 分布**と呼び、これに基づいて対立仮説が正しいのに帰無仮説を保持してしまう誤りをβ（**第二種の過誤**）と呼ぶことが分かる。
- □ **信頼区間**の作り方がわかり、95%信頼区間とは「標本から計算された区間が母集団のパラメータ値を含む確率」が95%であることを意味するということが分かる。
- □ **サンプルサイズ**をあまりに大きくすると、細かい差まで際限なく検出できるようになり、この結果実質的には小さな差しかないのに有意になってしまうことが分かる。
- □ 平均値の差が標準偏差が何個分なのかを表す統計量として**コーエンの d** という効果量があることが分かる。
- □ 二種類の過誤と母集団効果量から**サンプルサイズの決定**に利用できる。

本講から数回にわたり、t 検定と呼ばれる統計手法を学びます。これから学んでいく統計手法の中ではもっとも単純なモデルの一つで、論文などでも目にすることの多い伝統的な統計手法です。しかし「単純」なモデルではありますが、より発展的な手法の基礎となるモデルでもあります。しっかりと身に付けてください。

見取り図



データの形式

ID	グループ（説明変数）	観測値（応答変数）
1	0	2.1
2	0	3.2
⋮	⋮	⋮
n_0	0	3.1
n_0+1	1	2.3
n_0+2	1	1.2
⋮	⋮	⋮
n_0+n_1	1	1.5

（１）目的と考え方

① 目的（リサーチクエスチョン）

二つのグループからの**小標本**に差があるか調べる。

② 考え方

各グループの標本の**平均**を計算し全体のばらつきを考えた上でそれらの差がどのくらいぶれるかを判断する。

（２）仮説検定

① 仮説の設定

H0：二つの母集団平均は完全に同じだ

H1 :「二つの母集団平均は完全に同じだ」は偽だ

② 検定統計量の算出

「二つのグループの平均値の差：そのばらつき」という比を考え、その比の値を t とおく。

$$t = \frac{\text{「二つのグループの平均値の差」}}{\text{「二つのグループの平均値の差」のばらつきの推定量}}$$

③ 有意水準の設定

H0 のとき t がどういう値を取るかを示した分布。

④ 仮説の採択

③の基準に照らし①の値が大きいなら、H0 を保持することを諦め、代わりに、対立仮説である H1 を採択する。

（３）区間推定

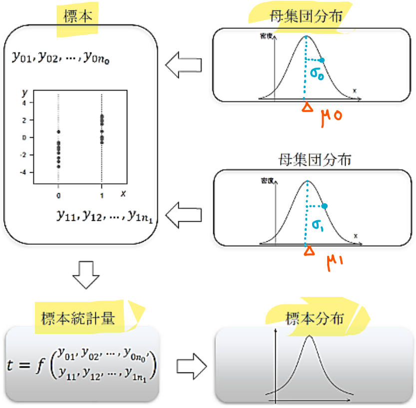
x%信頼区間という標本から作った区間は 100 回に x 回母集団のパラメータの値を含むようなように作られている。

（４）効果量（コーエンの d）

「二つのグループの平均値の差：全体のばらつき」という比を考え、その比の値（d）から実質的な差を考える。

$$d = \frac{\text{二つのグループの平均値の差}}{\text{全体のばらつきの推定値}}$$

📖 ノート１　母集団に対する仮定：帰無仮説と対立仮説



グループ０　構文Aの容認度　A組のテストの点　…

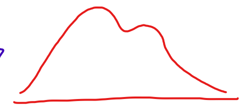
グループ１　構文Bの容認度　B組のテストの点　…

Ⓟ t検定の目標

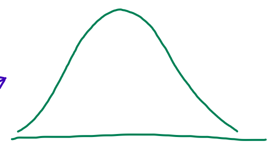
母集団における二群の平均値差の有無を検討すること

Ⓟ 母集団分布を仮定するということ

① 全知全能の視点



② 人間の視点



わたしたちは、何らかの分布を母集団に想定せざるを得ず、正規分布をはじめとする分布を仮定するのだが、それが真の分布と一致している保証はない。

全知全能（神）の視点に立つと母集団分布の形が１つ明確に見える

（１）母集団の実存

頻度主義統計学では、真の母集団分布の存在を仮定する。
パラメータには、真の値があると仮定される。

（２）頻度主義統計学の立場

人間は、真の母集団分布も、ましてや、パラメータの真の値も知りえない。

そこで、仮説を立て、その仮説が現実に得られたデータとどのくらい整合性があるかで、仮説の是非を判断する。

（３）標本分布を同定するために必要な仮定

【標本の抽出の仕方に関する仮定】

① **独立性の仮定**：　各要素は互いに独立
independently

② **同一分布の仮定**：　標本は同じ分布から抽出
identical

③ **無作為性の仮定**：　標本はランダムに抽出

$$y_1, y_2, \cdots, y_n \overset{i.i.d}{\sim} N(\mu, \sigma^2)$$

Ⓟ 独立していないとき

ID		y_i
1	山田	30
2	山田	40
3	中山	10
⋮		⋮
n	川田	50

30に近い値が予測

同じ人物が反復されて観測されたことで、ID = 1のデータとID = 2のデータに相関が生じてしまった（= 独立ではない！）

【母集団の姿に関する仮定】

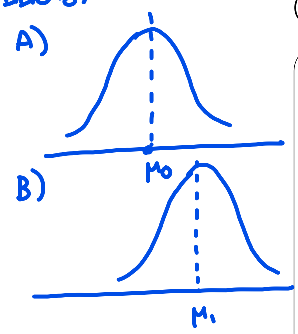
① 正規性の仮定：　母集団の分布は正規分布

② 分散の等質性の仮定：　二つの母集団の分散は同じ

→ ③ 平均値の差に対する仮説：二つの母集団の平均は同じ

話の都合上「差はない」という仮定の下、推論を進めるが、むしろ、データとの整合性を踏まえると、棄却して、退けたい仮説
⇒ 帰無仮説と呼ぶ。

例 英語の教授法



$H0: \mu_0 = \mu_1$
（2つの教授法に差がない）

$H1: \mu_0 \neq \mu_1$
（2つの教授法に差がある）

こっちが採択されてほしい!!

（４）検証する仮説の設定

☞ 帰無仮説と対立仮説

① **帰無仮説** Null Hypothesis (H0)
それを棄却する目的で設定される仮説のこと。

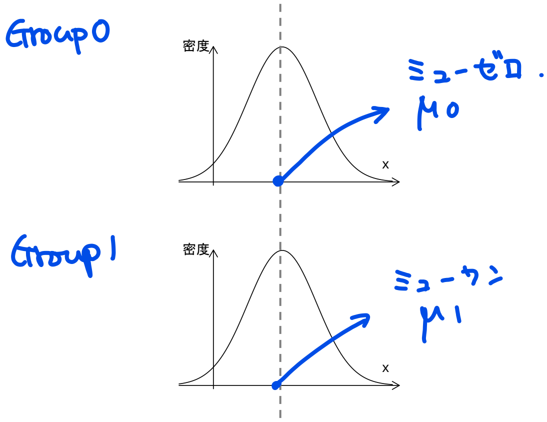
研究者のもくろみとして、まちがっている（無）だろうと、最終的に結論づけたい（帰）仮説のこと。

② **対立仮説** Alternative Hypothesis (H1)
帰無仮説が、棄却された場合に採択する仮説のこと。

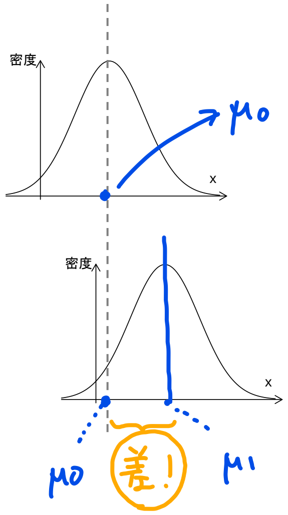
帰無仮説の代わりに採択する、という意味

今回の t 検定では、それぞれ次のようになる。

① **帰無仮説** Null Hypothesis (H0)
H0：二つの母集団平均は完全に同じだ



② **対立仮説** Alternative Hypothesis (H1)
H1：「二つの母集団平均は完全に同じだ」は偽だ



📖 ノート2　標本分布：t 分布

（1）標準化 Standardization

① 例題

関東の予備校の英語の模試で 60 点を取った A 君と関西の予備校の英語の模試で 60 点を取った B 君はどちらが英語の成績が良いと言えるのだろうか？

② 偏差値

異なる標本間での点数の比較のために各データが「平均から標準偏差何個分」の位置にあるか計算したもの。

$$z_i = 50 + 10 \times \frac{y_i - \bar{y}}{s_y}$$

※ただし、もとのテストの点数と似たような値に収まるよう 10 をかけて 50 を足している。

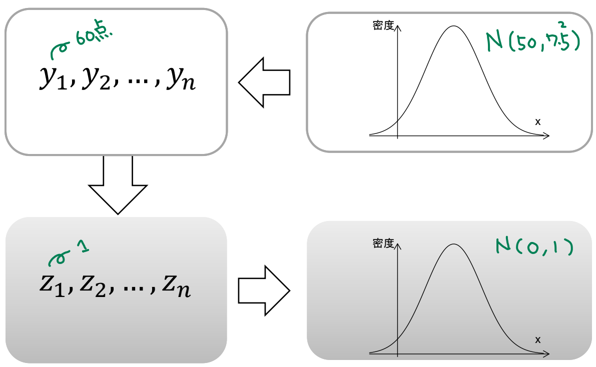
③ 標準化

異なる標本間での点数の比較のために各データが「平均から標準偏差何個分」の位置にあるか計算したもの。

$$z_i = \frac{y_i - \bar{y}}{s_y}$$

④ 標準正規分布

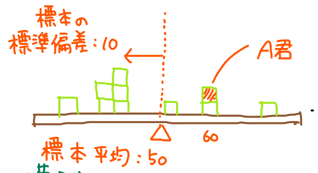
これは、平均が 0、分散が 1 の正規分布$N(0,1)$。なお、変数y_iが正規分布$N(\mu,\sigma^2)$に従うとき、標準化という操作で作られた新しい統計量z_iは標準正規分布に従う。



Ⓟ 偏差値

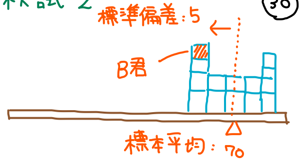
模試 1

60



模試 2

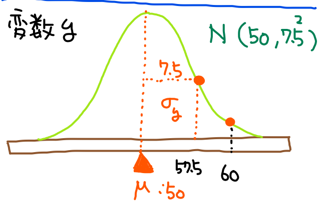
30



$$\underbrace{50 + \left(\frac{y_B - \bar{y}}{s_y}\right) \times 10}_{\text{偏差値}}$$

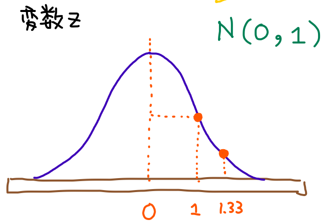
$\underbrace{\frac{y_B - \bar{y}}{s_y}}_{\text{標準化得点}}$

Ⓟ 正規分布と標準正規分布



$$z_B = \frac{y_B - \mu}{\sigma_y}$$

変数の標準化



Ⓟ 中心極限定理のバージョンアップ

（第2講）

平均がμで，分散がσ^2の
任意の母集団分布から
$y_1, y_2, \cdots, y_n$を取りだし，
その標本平均$\bar{y}$を計算すると，
サンプルサイズnが大きいとき

$\bar{y} \sim N(\mu, \frac{\sigma^2}{n})$.

（今回の話）

平均がμで，分散がσ^2の
正規分布$N(\mu, \sigma^2)$の母集団から
$y_1, y_2, \cdots, y_n$を取りだし，
その標本平均$\bar{y}$を計算すると，
サンプルサイズnが大きくなくても

$\bar{y} \sim N(\mu, \frac{\sigma^2}{n})$.

Ⓟ 人間の限界

この z の値をもとに，議論をすすめたいのだが，μとσ^2は母集団に属するので人間にはわからない。

① μについて：
「母集団のμがゼロである」という仮定（帰無仮説）の下で話を進めることにする

② σについて：
得られた標本からσの値を推定し，それを$\hat{\sigma}$とおいて代入する．

サンプルを取るたびにぶれるけど今回のデータをふまえると
$\hat{\sigma}$ = 　σは10なんじゃないかな
σは10.1じゃ

Ⓟ 記法：^ （ハット）

σ：真の値
$\hat{\sigma}$：人間が推定したもの

Ⓟ z値とt値の比較

帰無仮説下では，$\mu_1 - \mu_0 = 0$なので

$$z = \frac{\bar{y}}{\sigma/\sqrt{n}} \qquad t = \frac{\bar{y}}{\hat{\sigma}/\sqrt{n}}$$

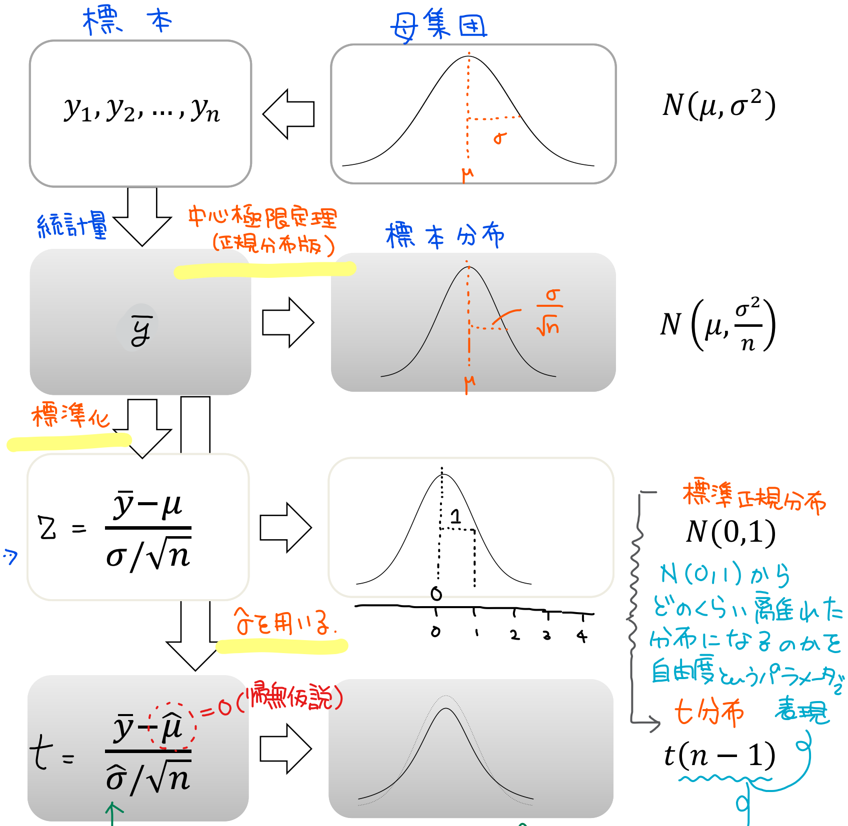
神の視点ではこの母集団の値σは定数

サンプルをとる度にかわる統計量
t値とは2つの統計量の比！

（2）t分布 t-distribution

"ばらつきのソースが増えたことによる"

これは、統計量の算出の際σではなく$\hat{\sigma}$を用いたことによる補正のせいで標準正規分布へのなりかけとなった分布。



ばらつきのソース：$\bar{y}$に加え，$\hat{\sigma}$もばらつく．このため，標準正規分布よりばらついた分布に。

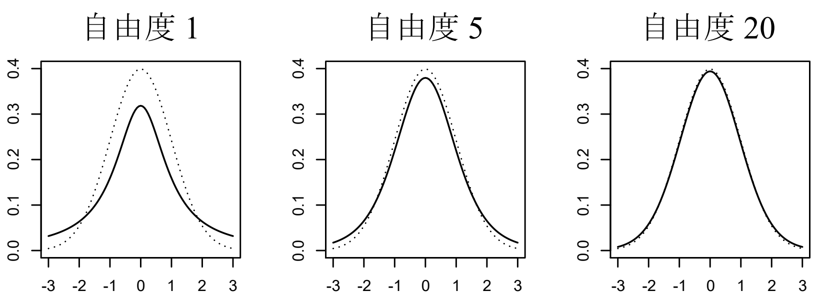
① 形状

t分布の形状は、サンプルサイズで決まる。
※対応のない二群のt検定では$t(n_0 + n_1 - 2)$を用いる。

② 標準正規分布との関係

$\hat{\sigma}$がばらつかなくなっていくので，ばらつきの源は$\bar{y}$だけに限られてゆくから。

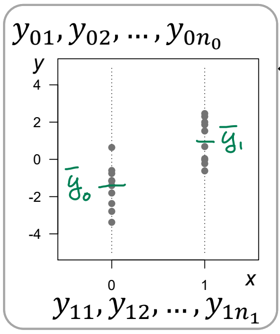
t 分布は、その自由度をどんどん大きくさせていくと、標準正規分布に近づいていく。



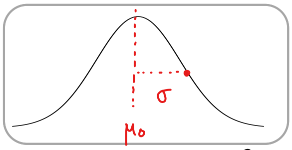
📖 ノート３　検定統計量ｔと仮説検定

（１）二群の平均値差

母集団

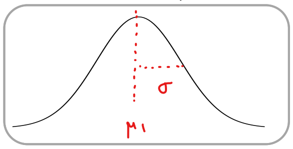


グループ0



$N(\mu_0, \sigma^2)$

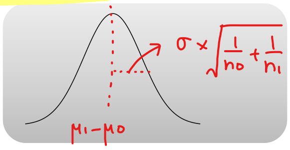
グループ1



$N(\mu_1, \sigma^2)$

中心極限定理 ＋ 再生性　A

$\bar{y}_1 - \bar{y}_0$



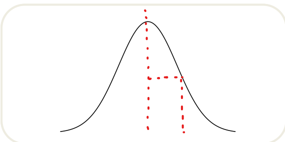
$N\left(\mu_1 - \mu_0, \frac{\sigma^2}{n_0} + \frac{\sigma^2}{n_1}\right)$

B　標準化

0（帰無仮説）

$$z = \frac{(\bar{y}_1 - \bar{y}_0) - (\mu_1 - \mu_0)}{\sigma\sqrt{\frac{1}{n_0}+\frac{1}{n_1}}}$$

人間が推定

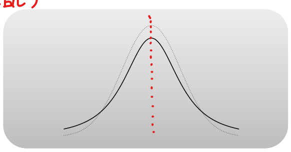


$N(0,1)$

$\hat{\sigma}$を用いる　C

0（帰無仮説）

$$t = \frac{(\bar{y}_1 - \bar{y}_0) - (\mu_1 - \mu_0)}{\hat{\sigma}\sqrt{\frac{1}{n_0}+\frac{1}{n_1}}}$$



$t(n_0+n_1-2)$

Ⓟ 記法

グループ0の　y_{01}
何番のデータか

n_0：グループ0のサンプルサイズ
n_1：〃　1　〃

$\bar{y}_0$：グループ0の標本平均
$\bar{y}_1$：〃　1　〃

Ⓟ 正規分布の再生性

これは，正規分布にしたがう
uとvが存在しているとき，
そのuとvの和/差も正規分布
になるという性質のこと

$u \sim N(\mu_0, \sigma_0^2)$
$v \sim N(\mu_1, \sigma_1^2)$
$v \pm u \sim N(\mu_1 \pm \mu_0, \sigma_0^2 + \sigma_1^2)$

Ⓟ 二群の平均値差　$\bar{y}_1 - \bar{y}_0$

（ステップ1）中心極限定理から
$\bar{y}_0 \sim N\left(\mu_0, \frac{\sigma^2}{n_0}\right)$
$\bar{y}_1 \sim N\left(\mu_1, \frac{\sigma^2}{n_1}\right)$
（ステップ2）正規分布の再生性から
$\bar{y}_1 - \bar{y}_0 \sim N\left(\mu_1 - \mu_0, \sigma^2\left(\frac{1}{n_0}+\frac{1}{n_1}\right)\right)$

Ⓟ 標準誤差 standard error

$$t = \frac{\bar{y}_1 - \bar{y}_0}{\hat{\sigma}_{\bar{y}_1-\bar{y}_0}}$$

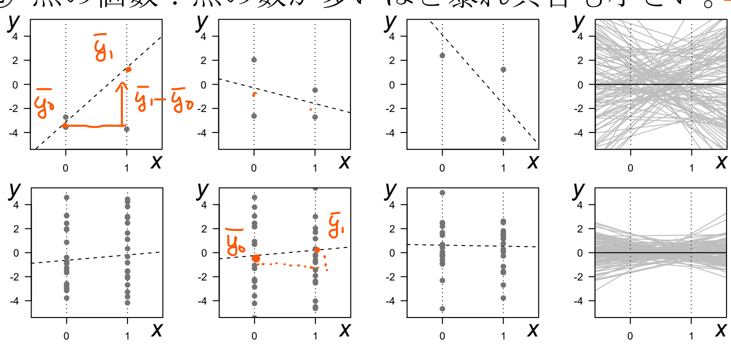
「統計量である$\bar{y}_1 - \bar{y}_0$が従う
標本分布の標準誤差
$\sigma_{\bar{y}_1-\bar{y}_0}$を推定したもの」

※ 標準誤差に影響を与える要素

標本分布の標準偏差を標準誤差と言い、平均値差$y_1 - y_0$の従う標本分布の標準誤差を$\sigma_{\bar{y}_1 - \bar{y}_0}$と表す。

① もともとのばらつきの推定値：$\hat{\sigma}_y$

② 点の個数：点の数が多いほど暴れ具合も小さい。



$$\hat{\sigma}_{\bar{y}_1-\bar{y}_0} = \hat{\sigma} \times \sqrt{\frac{1}{n_0}+\frac{1}{n_1}}$$

$$t = \frac{\bar{y}_1 - \bar{y}_0}{\hat{\sigma}} \times \frac{1}{\sqrt{\frac{1}{n_0}+\frac{1}{n_1}}}$$

効果の大きさ　サンプルサイズ

Ⓟ 背理法

これは，ある仮定を導入し，議論を進め，もし，その仮定の下で，矛盾(整合性がとれない状況)が生じたら，翻って，その仮定を退ける推論。

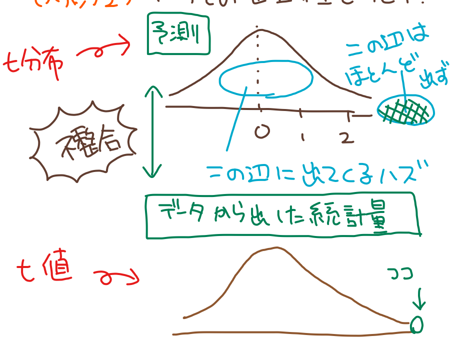
(ステップ1) 動物X = 哺乳類
(ステップ2) Xは エラ呼吸する
(ステップ3)「動物X = 哺」は誤り.

Ⓟ 検定統計量

これは，仮説検定に用いる統計量のこと。二群の平均値差の検定(t検定)では，左ページのt値を用いる.

Ⓟ 標本分布を用いた背理法

(ステップ1)「H0: $\mu_0=\mu_1$」とする.
(ステップ2) データとの整合性を判断.



(ステップ3)「H0: $\mu_0=\mu_1$」は偽

（２）仮説検定：t 検定

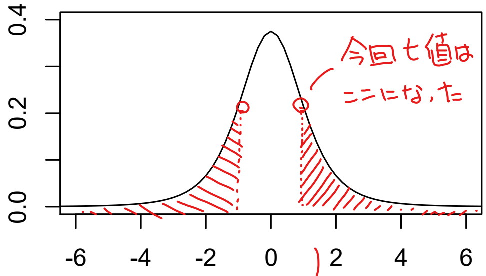
$$\begin{cases} H0: \mu_0 = \mu_1 \quad (\mu_1 - \mu_0 = 0) \\ H1:「\mu_0 = \mu_1」は偽 \end{cases}$$

手順 1　仮説の設定

母集団に関する帰無仮説と対立仮説を設定する。

手順 2　検定統計量の計算（データが決定するもの）

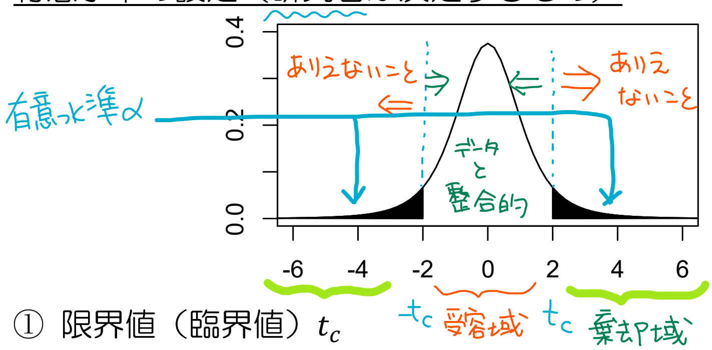
データから t 値を計算し、その p 値を求める。



※ p 値

帰無仮説が正しい時に、今回の検定統計量の実現値が、今以上に極端になる確率のこと。

手順 3　有意水準の設定（研究者が決定するもの）



① 限界値（臨界値）t_c

よくあることと滅多にないことを分ける基準。

② 棄却域と受容域

棄却域とは、限界値より極端な値を示した領域。
棄却域ではない領域が受容域。

③ 有意水準α

t 値が棄却域に入る確率のこと。

手順 4　仮説の選択

① 検定統計量の実現値が棄却域に入った場合
帰無仮説のもとではありえなかったことが起きたと考え、帰無仮説を棄却する。

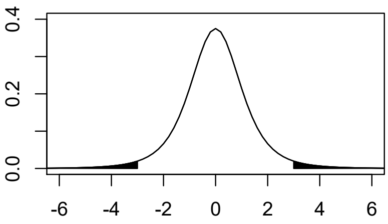
② 検定統計量の実現値が受容域に入った場合
帰無仮説と得られたデータは、整合的だったと考え、帰無仮説は棄却しない。

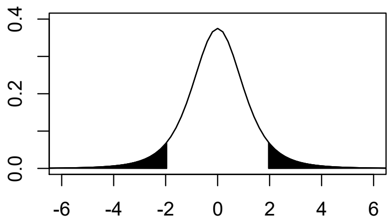
H0が正しいことが証明される訳ではない，ということに注意！

📖 補足：二種類の過誤

（1）研究者が判断基準を用意するということ

データから$t = 2$だった！





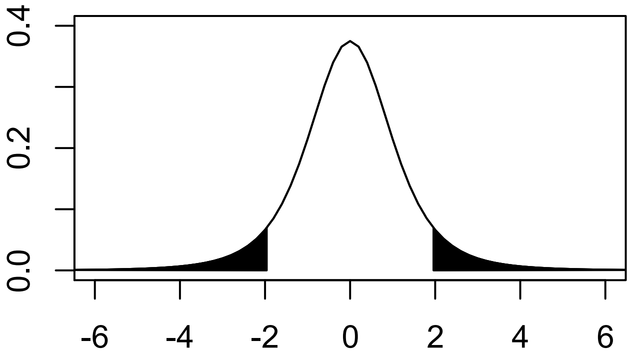
（2）二種類の過誤

真実/判断	H0	H1
H0	正しい判断 $1-\alpha$	第一種の過誤 α
H1	第二種の過誤 β	正しい判断 $1-\beta$

① **第一種の過誤** α

「H0 が真」であるときの条件付き確率

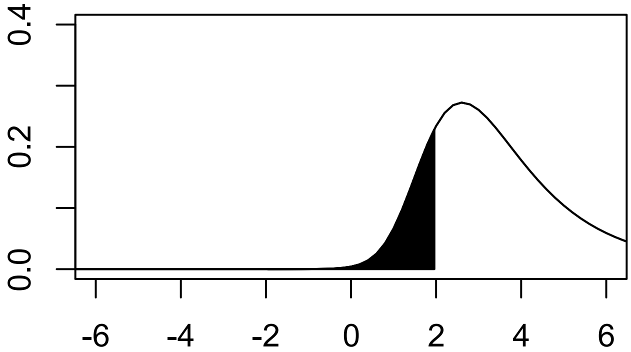
＝　帰無仮説の方が正しいのに棄却してしまうリスク



② **第二種の過誤** β

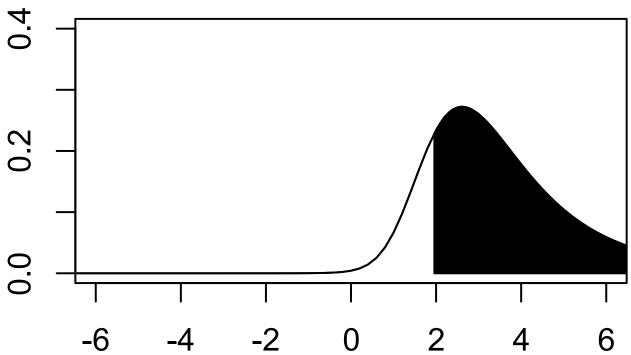
「H1 が真」であるときの条件付き確率

＝　H1 が正しいのに H0 を棄却できないリスク。



③ **検定力**　$1-\beta$　「H1 が真」であるときの条件付き確率

H1 の方が正しいときにちゃんと H1 を採用できる確率。



（3）片側検定（非推奨）

何らかの理由で t 値の符号が必ず正/負であることが明白なとき、一方の棄却域を消し反対側の棄却域を倍にする検定。

ネイマン流の区間推定
信頼区間の作り方を学ぶ

（1）目的

母集団の$\mu_1-\mu_0$に対する幅を持たせた推論をしたい。

① 検定

$\mu_1-\mu_0=0$ (H0) がデータと整合的か考える。

② 区間推定

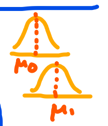
データと整合的な$\mu_1-\mu_0$の値の集合を考える。

$\mu_1-\mu_0=2$ データと整合 Y/N
$\mu_1-\mu_0=1$ データと整合 Y/N
$\mu_1-\mu_0=0$ データと整合 Y/N
$\mu_1-\mu_0=-1$ データと整合 Y/N
$\mu_1-\mu_0=-2$ データと整合 Y/N

母集団の平均値差

Ⓟ t値の拡張

標 ← 母
↓
統 → 標分



①帰無仮説が正しいとき
$\mu_0=\mu_1$なので、
$\bar{y}_1-\bar{y}_0$もきっと0のまわりに分布するハズ。
→ 標本分布は0が中心。

② $\mu_1-\mu_0=a$ $(a\neq 0)$のとき
$\mu_1-\mu_0$がaなら、
$\bar{y}_1-\bar{y}_0$もaに近い値に偏るはず。
→ 偏りをのぞくために、
$\bar{y}_1-\bar{y}_0-a$のように
aを引いてやると、
帰無分布と同様、
中心が0のt分布に従う。

$$t=\frac{(\bar{y}_1-\bar{y}_0)-(\mu_1-\mu_0)}{\hat{\sigma}_{\bar{y}_1-\bar{y}_0}}$$

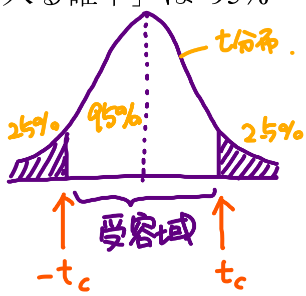
これがt分布に従う。

（2）考え方

① 「 t値 が 受容域 に 入る確率」は 95%

$$-t_c \leq t \leq t_c$$

$$t=\frac{(\bar{y}_1-\bar{y}_0)-(\mu_1-\mu_0)}{\hat{\sigma}_{\bar{y}_1-\bar{y}_0}}$$



議論に沿って各値を代入する！

② 「$-t_c \leq \dfrac{(\bar{y}_1-\bar{y}_0)-(\mu_1-\mu_0)}{\hat{\sigma}_{\bar{y}_1-\bar{y}_0}} \leq t_c$ の確率」は95%

㊟ 式が複雑に見えるかもしれませんが、①を表しただけで、これがらないで!!

わたしたちの最終目的は、$\mu_1-\mu_0$の範囲を考えることだったので、②の不等式を変形して、
△△ $\leq \mu_1-\mu_0 \leq$ □□ という形を作りましょう！

③ 「$\bar{y}_1-\bar{y}_0-t_c\times\hat{\sigma}_{\bar{y}_1-\bar{y}_0} \leq \mu_1-\mu_0 \leq \bar{y}_1-\bar{y}_0+t_c\times\hat{\sigma}_{\bar{y}_1-\bar{y}_0}$ の確率」は95%

緑 $\leq$ 赤 $\leq$ 緑

緑：標本から計算されたもの　赤：母集団の値　緑：標本から計算されたもの

④ 「標本から計算された区間が母集団のパラメータの値を含む確率」は95%

標本から計算された区間：必ず標本抽出のせいでばらつく。

母集団のパラメータの値：頻度主義統計学では、固定的で、ばらつくことはない。

Ⓟ信頼区間の確率の(イ)

この確率がでてくる原因は、緑がばらつくことにある！

㊟ 不等式の解き方："$\leq$"が2つあるので1つ1つばらして解く！

(i) $-t_c \leq \dfrac{(\bar{y}_1-\bar{y}_0)-(\mu_1-\mu_0)}{\hat{\sigma}_{\bar{y}_1-\bar{y}_0}}$

↓ 分母を両辺にかける。

$-t_c\times\hat{\sigma}_{\bar{y}_1-\bar{y}_0} \leq (\bar{y}_1-\bar{y}_0)-(\mu_1-\mu_0)$

↓ 移項（緑と赤を別の辺に！）。

$\mu_1-\mu_0 \leq \bar{y}_1-\bar{y}_0+t_c\times\hat{\sigma}_{\bar{y}_1-\bar{y}_0}$

(ii) $\dfrac{(\bar{y}_1-\bar{y}_0)-(\mu_1-\mu_0)}{\hat{\sigma}_{\bar{y}_1-\bar{y}_0}} \leq t_c$

↓ 分母を両辺にかける。

$(\bar{y}_1-\bar{y}_0)-(\mu_1-\mu_0) \leq t_c\times\hat{\sigma}_{\bar{y}_1-\bar{y}_0}$

↓ 移項

$\bar{y}_1-\bar{y}_0-t_c\times\hat{\sigma}_{\bar{y}_1-\bar{y}_0} \leq \mu_1-\mu_0$

$$\bar{y}_1-\bar{y}_0-t_c\times\hat{\sigma}_{\bar{y}_1-\bar{y}_0} \leq \mu_1-\mu_0 \leq \bar{y}_1-\bar{y}_0+t_c\times\hat{\sigma}_{\bar{y}_1-\bar{y}_0}$$

言語学に携わる研究者としては、この読み方さえまちがわなければ、細かい点は、統計ソフトが行うので、覚えておかなくてもOK.

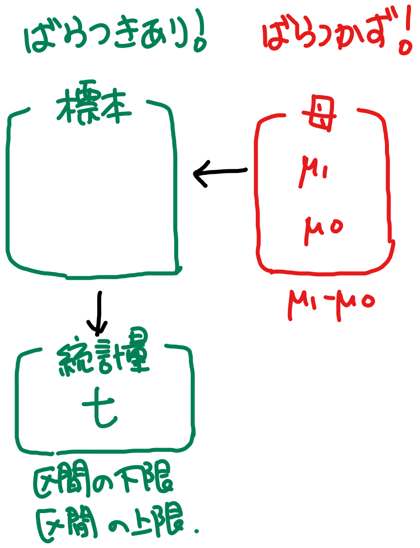
（3）読み方

頻度主義とよばれるゆえん.

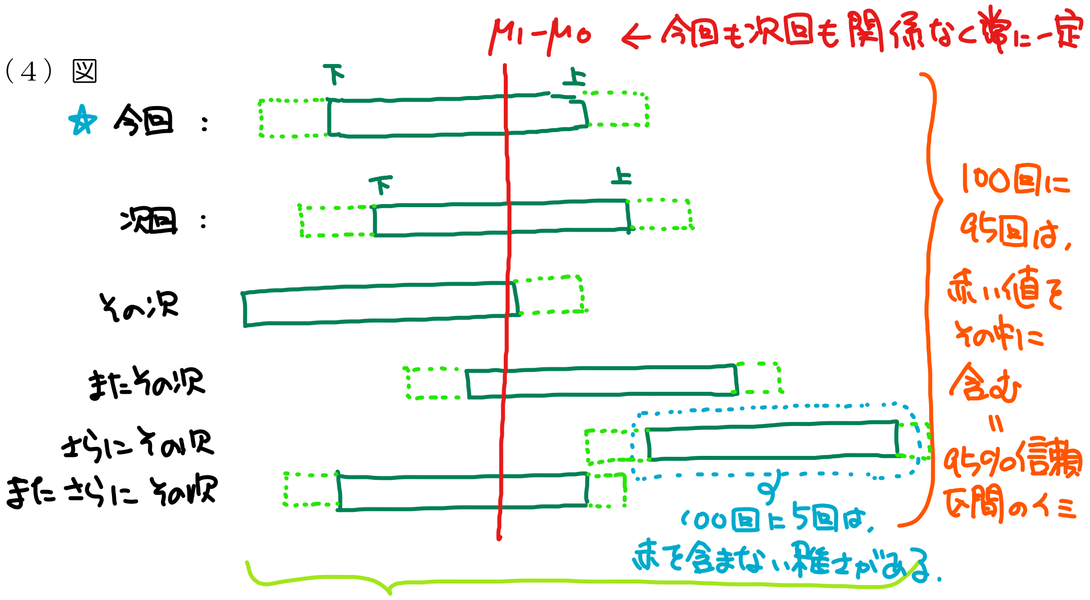
何度も何度も、同じようにサンプルを取ったとしたら
この区間が母集団の値($\mu_1-\mu_0$)を含む 相対頻度が 100回に95回あたる.

95%信頼区間 ＝ [−4.28,　　0.28]

※よくある間違い
95%信頼区間とは、
○)「この区間が真の$\mu_1 - \mu_0$を含む相対頻度が 95%であるということ」
であり、
この書き方をすると、あたかも$\mu_1-\mu_0$にばらつきがあると読めてしまう。
×)「母集団の$\mu_1 - \mu_0$がこの間の値を取る確率が 95%である」
という意味ではない！



（4）図



信頼区間もまた標本から抽出された統計量(緑)なのでばらつく!

確認問題：サンプルサイズとｐ値の落とし穴

英語の教授法 A と B の効果を調べるために、調査を行った。集めた実験協力者を無作為に半分に分け、一方に教授法 A、残り半分に教授法 B を採用した授業を受けてもらった。この実験協力者に共通のテスト（100 点満点）を受けてもらった。ここから教授法 A を受けた人と教授法 B を受けた人という二つの母集団の英語成績の母集団平均値に差があるのかどうかを検証するために、対応のない t 検定を行った（両側検定）。

（１）実験協力者が６人のとき、次の結果を得た。

① t 値：0.41

② p 値：0.70

③ 95%信頼区間：CI (-17.3, 23.3)

問１　α 水準が 0.01%であるとき帰無仮説を棄却できるか。

問２　実質的に二つのグループには差があると言えるか。

$\hat{\sigma}_{\bar{y}_1-\bar{y}_0}$ がそこそこ大きい

（２）実験協力者が 6,000 人のとき、次の結果を得た。

① t 値：5.09

② p 値：0.0000003717

③ 95%信頼区間：CI (0.79, 1.79)

問１　α 水準が 0.01%であるとき帰無仮説を棄却できるか。

問２　実質的に二つのグループには差があると言えるか。

問３　問１と問２に一見食い違いが見られるのはなぜか。

$\hat{\sigma}_{\bar{y}_1-\bar{y}_0}$ はかなーり小さい．ので，とても小さい
$\bar{y}_1-\bar{y}_0$ の値でも，有意に見えてしまう！
⇒ t という値のみで判断するのは危険

Ⓟ 検定とサンプルサイズ

母集団 → 小さいサイズ 標本：t 値 ㊇
母集団 → 大きいサイズ 標本：t 値 ㊅

$$t = \frac{\bar{y}_1 - \bar{y}_0}{\hat{\sigma}_{\bar{y}_1-\bar{y}_0}} = \frac{\bar{y}_1 - \bar{y}_0}{\hat{\sigma}_y \times \sqrt{\frac{1}{n_0}+\frac{1}{n_1}}}$$

サンプルサイズ

ここが大きくなると結果として t 値も大きくなる！

分母とは，$\bar{y}_1-\bar{y}_0$ の標準誤差（ばらつき）
→ サンプルサイズが大きいと小さくなる！

Ⓟ t 値，P 値だけに注目すべからず

サンプルサイズによって，どのくらい細かい差が検出できるのか変わってしまうから！

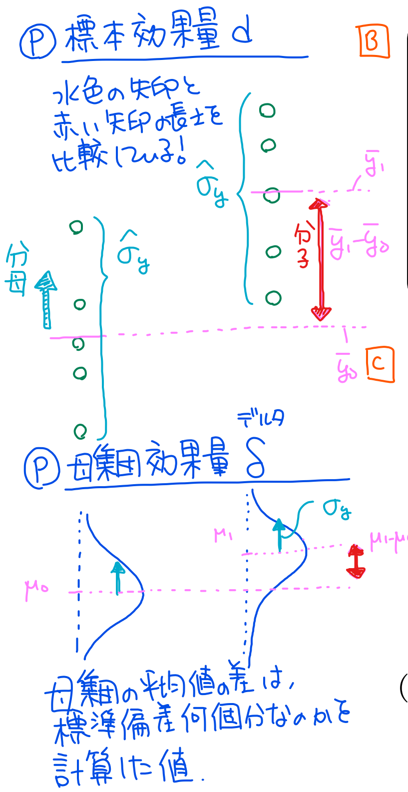
↓

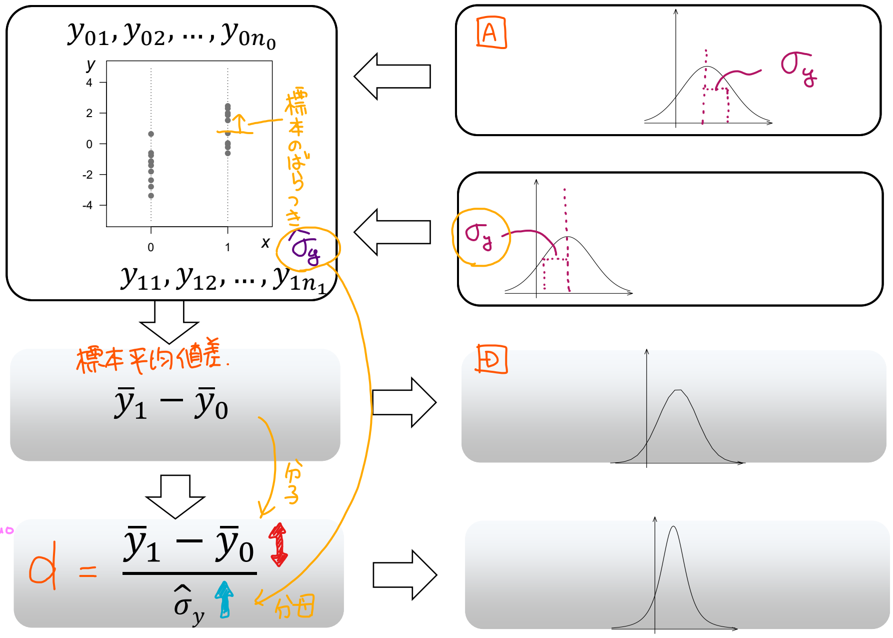
対策

① 信頼区間を見る．

② 効果量を見る．

③ 適切なサンプルサイズを，実験を行う前に判断する．





（１）母集団の効果量：母集団標準化平均値差 δ

$$\delta = \frac{\mu_1 - \mu_0}{\sigma}$$

（２）母集団効果量の点推定

母集団の標準化平均値差の**点推定値**として標本標準化平均値差 d が用いられる。

$\hat{\delta}$

$$\hat{\delta} = d = \frac{\bar{y}_1 - \bar{y}_0}{\hat{\sigma}}$$

（３）母集団効果量の区間推定

サンプルサイズが大きい時 d の標本分布が正規分布となることを利用すると、δの 95%信頼区間は次のようになる。

$1.96 \times \hat{\sigma}_d$ 統計量 d の標準誤差

$\hat{\delta}$

$1.96 \times \hat{\sigma}_d$ 統計量 d の標準誤差

$$\hat{\sigma}_d = \sqrt{\frac{1}{n_0} + \frac{1}{n_1} + \frac{d^2}{2(n_0 + n_1 - 2)}}$$

📖 補足：効果量を用いたサンプルサイズの決定

（1）サンプルサイズの決定

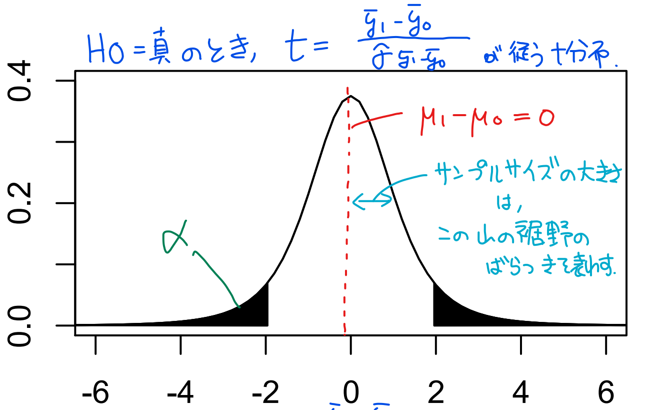
次の四つは、三つが決まると残り一つが決まる関係にある。

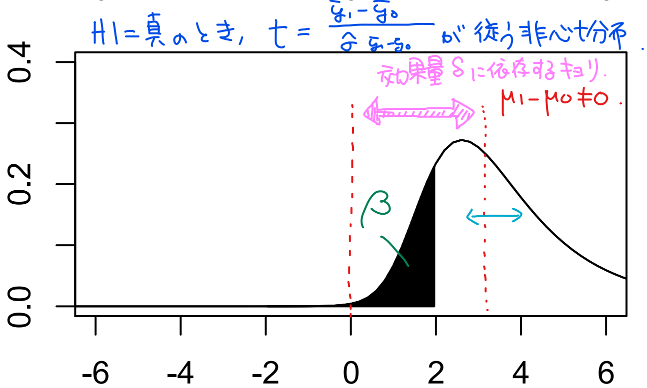
①第一種の過誤の確率（α）、H0=真のとき, H0を棄却!(tが棄に入る確率).

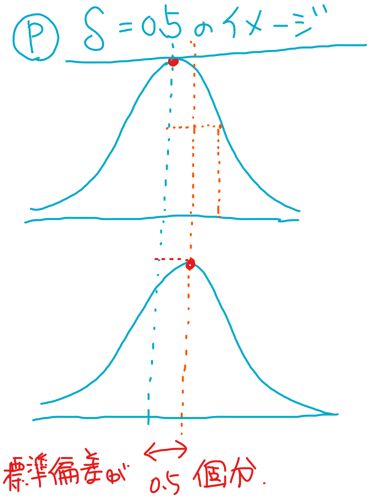
②第二種の過誤の確率（β）、H1=真のとき, H0を保持!(tが受 〃).

③想定される母集団効果量（δ）母集団の平均値差が標準偏差何個分?

④サンプルサイズ（n） 標本の大きさ.







（２）実践上の手続き

① **表**：標本数を決める表は数多く発表されている。

例：$\alpha = 0.05$、検定力$1-\beta = 0.8$の平均値の差の検定

	$\delta = 0.2$	$\delta = 0.5$	$\delta = 0.8$
標本数	394	64	26

$\mu_1 - \mu_0$ が 標準偏差 σ_y の 0.5個分.

② **統計ソフト**：

・R のような統計ソフトでは、簡単に計算してくれる（pwr ライブラリの power.t.test）。

・検定力関係では、G*Power も有名。

想定する標準化平均値差 が 0.5くらい

有意水準 0.05 (5%)

```
> library(pwr)
> pwr.t.test(d = 0.5, sig.level = 0.05, power = 0.8)
     Two-sample t test power calculation

              n = 63.76561
              d = 0.5
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

検定力 0.8

```
> pwr.t.test(n = 3000, sig.level = 0.05, power = 0.8)
     Two-sample t test power calculation

              n = 3000
              d = 0.07236001
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

標準偏差が 0.07個分の差があったとしても見つけられる.