１１．検定力

内　容

１．帰無仮説検定の問題点

２．効果量

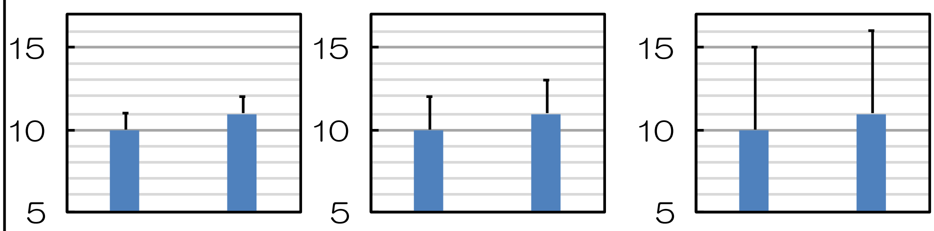
３．検定力分析

3

1．帰無仮説検定の問題点

4

帰無仮説検定の問題点　*t* 検定



n	*t*	*p*
5	1.581	0.153
10	2.236	0.038

n	*t*	*p*
5	1.118	0.296
10	1.581	0.131
15	1.937	0.063
20	2.236	0.031

n	*t*	*p*
5	0.707	0.500
10	1.000	0.331
15	1.225	0.231
20	1.414	0.165
30	1.732	0.089
40	2.000	0.049

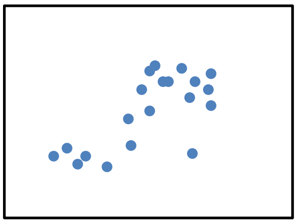
n は各グループの人数

誤差線は標準偏差

5

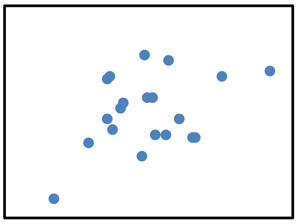
帰無仮説検定の問題点　無相関検定

$r = 0.7$



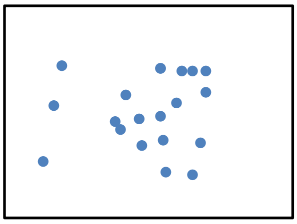
n	t	p
5	1.700	0.188
10	2.772	0.024

$r = 0.4$



n	t	p
5	0.756	0.505
10	1.234	0.252
15	1.574	0.140
20	1.852	0.081
30	2.309	0.029

$r = 0.1$



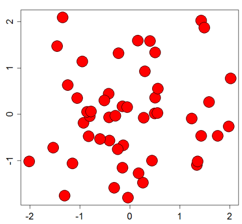
n	t	p
5	0.174	0.873
50	0.696	0.490
100	0.995	0.322
200	1.414	0.159
300	1.735	0.083
400	2.005	0.046

6

帰無仮説検定の問題点　無相関検定

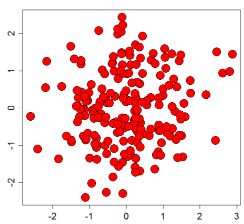
$r = 0.1$

$n = 50$



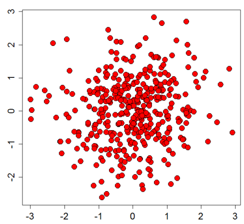
t	p
0.696	0.490

$n = 200$



t	p
1.414	0.159

$n = 400$



t	p
2.005	0.046

7

標本数の等しい対応のない t 検定

$\overline{X_1}, \overline{X_2}$ ：各群の標本平均
u_1^2, u_2^2 ：各群の不偏分散
n ：各群のデータ数

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{u_1^2 + u_2^2}{n}}} = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{u_1^2 + u_2^2}} \times \sqrt{n}$$

8

p 値の落とし穴

p 値は標本数の影響を大きく受ける。

p 値は効果の大きさを表す指標ではない。

9

帰無仮説検定の流れ

A群	a_1	a_2	•••	a_n
B群	b_1	b_2	•••	b_n

1. 仮説の設定
   帰無仮説（H_0）A = B
   対立仮説（H_1）A ≠ B
2. 検定統計量の計算
3. 検定統計量が臨界値より
   大きい→H_0棄却→ AとBに有意差あり
   小さい→H_0棄却しない→有意差なし

9

10

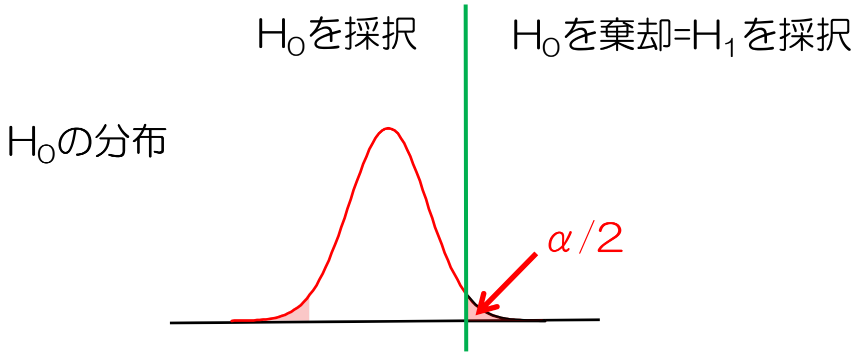
p 値の本当の意味

p 値

帰無仮説を棄却した場合に、
本当は帰無仮説が真であったという
間違いを犯す確率

	H_0を採択する	H_0を棄却する
本当はH_0が真	○	α（危険率）

11



12

2．効果量

帰無仮説の正しくなさを表す量的な指標

標本数に影響されない効果（差）の指標

検定	効果量の指標		効果量の目安		
	母集団	標本	小	中	大
t	$\delta = \dfrac{\lvert\mu_A - \mu_B\rvert}{\sigma}$	$d = \dfrac{\overline{X_1} - \overline{X_2}}{S_p}$	0.2	0.5	0.8
相関	$\lvert\rho\rvert$	$\lvert r\rvert$	0.1	0.3	0.5

13

t 検定の効果量

Cohen の d $$d=\frac{\overline{X_1}-\overline{X_2}}{S_p}$$

２群の標本数が異なる場合

$$S_p=\sqrt{\frac{(n_1-1)u_1^2+(n_2-1)u_2^2}{n_1+n_2}}$$

２群の標本数が等しい場合

$$S_p=\sqrt{\frac{u_1^2+u_2^2}{2}}$$

14

標本数の等しい対応のない t 検定

$$t=\frac{\overline{X_1}-\overline{X_2}}{\sqrt{\frac{u_1^2+u_2^2}{n}}}=\frac{\overline{X_1}-\overline{X_2}}{\sqrt{u_1^2+u_2^2}}\times\sqrt{n}=\underbrace{\boxed{\frac{\overline{X_1}-\overline{X_2}}{\sqrt{\frac{u_1^2+u_2^2}{2}}}}}_{\text{効果量}}\times\sqrt{\frac{n}{2}}$$

平均の差をデータの散らばりで割って、標準化

検定統計量＝標本効果量×f(サンプルサイズ)

15

t 検定の効果量（効果の大きさ）

$$d = \frac{\overline{X_1} - \overline{X_2}}{S}$$

コーエンの d

$$S = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2}}$$

効果大	.8
効果中	.5
効果小	.2

$S_1 : n_1 - 1$で割った標準偏差
$s_1 : n_1$で割った標準偏差

16

対応がある場合のt検定の効果量

相関で調整をする場合

$$d_{paired-t} = \frac{d}{\sqrt{2(1-r)}}$$

調整しない場合

17

効果量に関する Cohen　（1982）　の基準

	効果量	大	中	小
t 検定 t	d	0.8	0.5	0.2
無相関検定 r	r	0.5	0.3	0.1
カイ二乗検定 chisq	Φ係数（2×2） CramerのV	0.5	0.3	0.1
一元配置分散分析 anov	f, f^2	0.4	0.25	0.1
二元配置分散分析	η, η^2	0.14	0.06	0.01

```
install.packages("pwr")
library(pwr)
cohen.ES(test="t", size="small")
```

18

分散分析の効果量

f^2	要因の平方和／誤差の平方和 Ｆ値＊（df1/df2）
f	
η^2	要因の平方和／全体の平方和
η	相関比

19

	平方和	自由度	平均平方	分散比
要因	要因平方和	要因自由度	要因平均平方	分散比
誤差	誤差平方和	誤差自由度	誤差平均平方	
全体	全体平方和			

$$\eta^2 = \frac{\text{要因平方和}}{\text{全体平方和}}$$

$$\eta_p^2 = \frac{\text{要因平方和}}{\text{誤差平方和}+\text{要因平方和}}$$

$$\omega^2 = \frac{\text{要因自由度}\times(\text{要因平均平方}-\text{誤差平均平方})}{\text{全体平方和}+\text{誤差平均平方}}$$

20

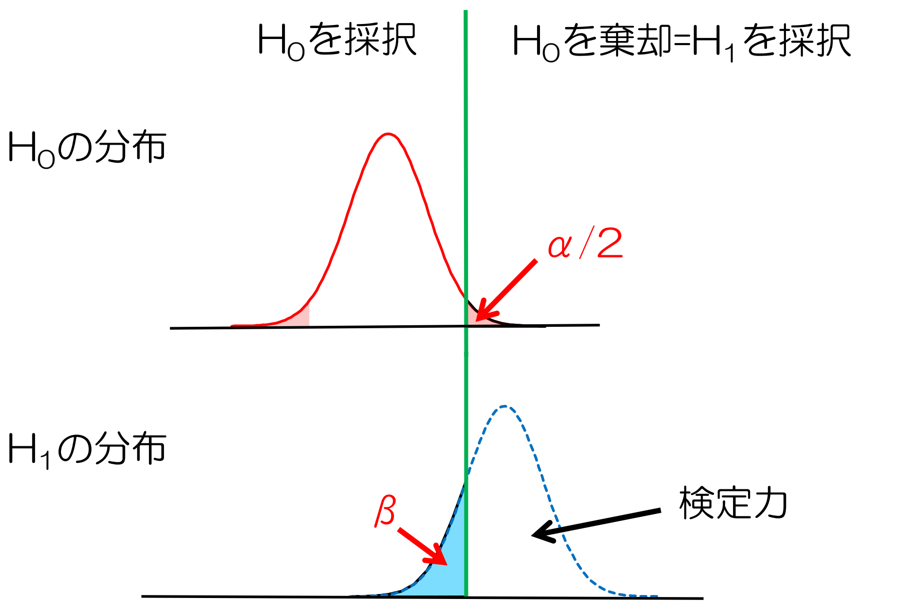
3．検定力分析

第一種の誤りと第二種の誤り

p 値

帰無仮説を棄却した場合に、
本当は帰無仮説が真であったという
間違いを犯す確率

	H_0を採択する	H_0を棄却する
本当はH_0が真	○	α（危険率） 第一種の誤り
本当はH_1が真	β 第二種の誤り	○ 検定力



検定力

検定力（1－β）

本当は対立仮説が真であった場合に、
対立仮説を正しく採択できる確率

本当は対立仮説が偽であった場合に、
対立仮説を採択してしまわない確率

検定力が 0.8 が望ましいとされる

検定力が 0.5 を切るような場合は問題

検定力分析の使い方

事前分析

- 予想される効果量（差）から、標本数を決めたい

  有意になる標本数という意味ではない！

- 標本数から、検出が期待できる効果量（差）を求めたい

事後分析

- 検定結果をどの程度主張していいのかを知りたい

25

検定力分析の使い方

有意でないのは、

- 標本数が少ないせい？
- 効果量が小さいせい？

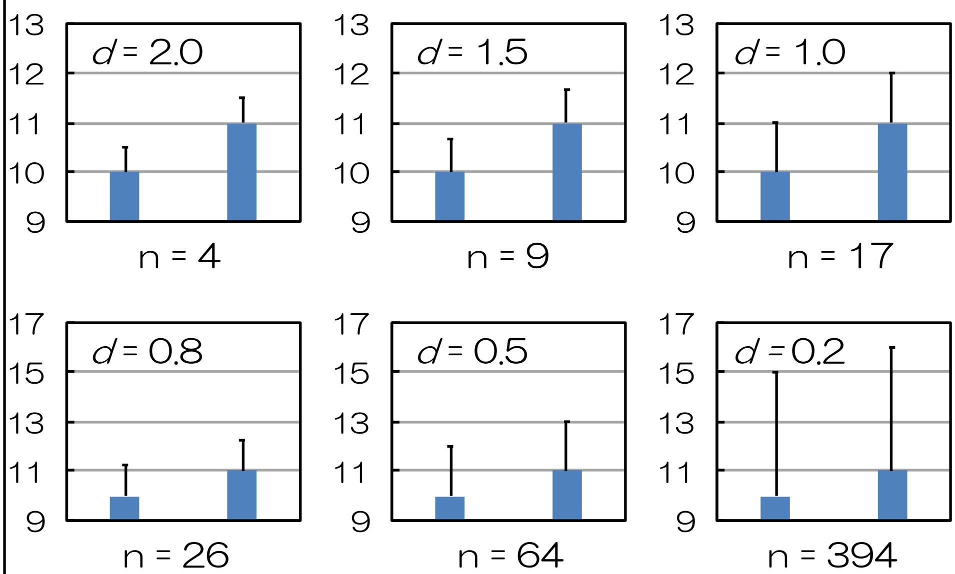
有意だけど本当？

- 検定力があればOKだが、なければ怪しい

検定力＝帰無仮説検定の結果が得られる確率

26

事前分析　*t* 検定：標本数を決めたい



誤差線は標準偏差

27

事前分析　t 検定：標本数を求めたい

検定力0.8

危険率5%

予想されるd	必要なn／グループ
2	4
1.5	9
1.0	17
0.8	26
0.5	64
0.2	394

28

power.t.test(sig.level=0.05, <u>power=0.8</u>, <u>delta=0.5</u>)

↑ 検定力　↑ 効果量

```
Two-sample t test power calculation
      n = 63.76576
      delta = 0.5
      sd = 1
      sig.level = 0.05
      power = 0.8
      alternative = two.sided
NOTE: n is number in *each* group
```

29

事前分析 無相関検定：標本数を求めたい

検定力0.8

危険率5%

予想されるr	必要なn
0.9	6
0.8	9
0.7	13
0.6	19
0.5	29
0.4	46
0.3	84
0.2	193
0.1	782

30

無相関検定でのサンプルサイズ

```
pwr.r.test(r=0.4, power=0.8)
```

効果量＝母相関係数

```
approximate correlation power
calculation (arctangh transformation)

      n = 46.5707
      r = 0.4
      sig.level = 0.05
      power = 0.8
      alternative = two.sided
```

31

カイ二乗検定でのサンプルサイズ

pwr.chisq.test(w=0.3, df=1, power=0.8)

効果量

```
Chi squared power calculation

      w = 0.3
      N = 87.20955
      df = 1
      sig.level = 0.05
      power = 0.8

NOTE: N is the number of observations
```

32

一元配置分散分析でのサンプルサイズ

pwr.anova.test(k=3, f=0.25, power=0.8)

効果量

```
Balanced one-way analysis of variance
power calculation

      k = 3
      n = 52.3966
      f = 0.25
      sig.level = 0.05
       power = 0.8

NOTE: n is number in each group
```

33

事前分析 t 検定：効果量を求めたい

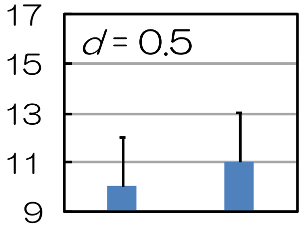
検定力0.8

危険率5%

n／グループ	検出可能なd
10	1.325
20	0.909
30	0.736
40	0.634
50	0.566
100	0.398
250	0.251
500	0.177

34

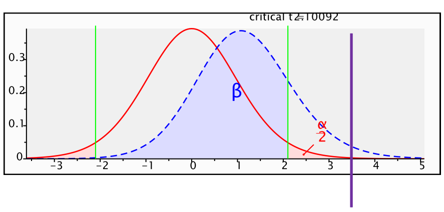
事後分析 t 検定：検定力を確認したい



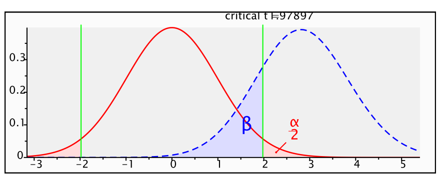
差＝1、SD = 2

n／グループ＝10

t (18)=1.581, p =0.131



検定力＝0.185



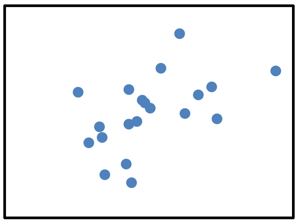
検定力＝0.801

n／グループ＝64

35

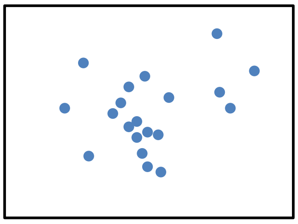
事前分析 無相関検定：標本数を決めたい

r = 0.5



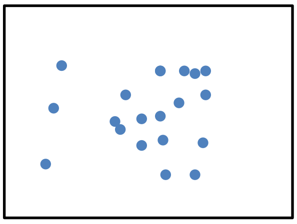
n=29

r = 0.3



n=84

r = 0.1



n=782

36

事前分析 無相関検定：効果量を求めたい

検定力0.8

危険率5%

n	検出可能なr
10	0.761
20	0.579
30	0.485
40	0.426
50	0.384
100	0.276
250	0.176
500	0.125

例題） 効果量と検定力を求める（t 検定）

37

A社	6	4	4	4	6	5	4	6	4	3
B社	5	4	3	3	5	3	5	5	4	3

```
m1 <- mean(a)
m2 <- mean(b)
n1 <- length(a)
n2 <- length(b)
s1 <- sd(a)
s2 <- sd(b)
sp <- sqrt(((n1-1)*s1^2+(n2-1)*s2^2)/(n1+n2))
d <- abs(m1-m2)/sp
pwr.t.test(n=length(a),d=d)
pwr.t2n.test(n1=length(a),n2=length(b),d=d)
```

38

A社	6	4	4	4	6	5	4	6	4	3
B社	5	4	3	3	5	3	5	5	4	3

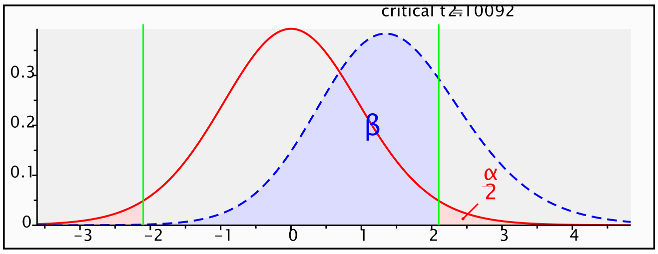
```
Two-sample t test power calculation

      n = 10
      d = 0.6255432
      sig.level = 0.05
      power = 0.2631322
      alternative = two.sided

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

39

G*Power



40

例題）　効果量と検定力を求める（対応のある t 検定）

~~A社~~	6	4	4	4	6	5	4	6	4	3
~~B社~~	5	4	3	3	5	3	5	5	4	3

$$d = \frac{\overline{X_1} - \overline{X_2}}{S}$$

S : 差の標準偏差（$n-1$で割った）

41

~~A社~~	6	4	4	4	6	5	4	6	4	3
~~B社~~	5	4	3	3	5	3	5	5	4	3

```
a <- c(6,4,4,4,6,5,4,6,4,3)
b <- c(5,4,3,3,5,3,5,5,4,3)
m <- a-b
d <- abs(mean(m))/sd(m)

pwr.t.test(n=length(a),type="paired",d=d)
```

42

~~A社~~	6	4	4	4	6	5	4	6	4	3
~~B社~~	5	4	3	3	5	3	5	5	4	3

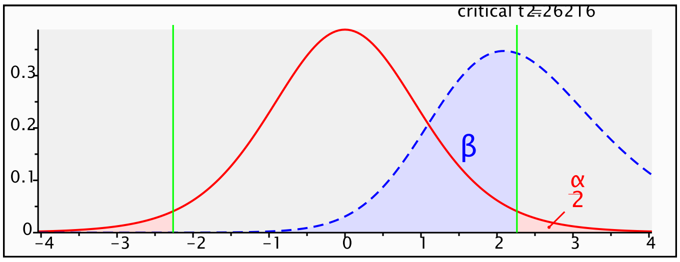
```
Paired t test power calculation

          n = 10
          d = 0.7115125
          sig.level = 0.05
          power = 0.5192634
          alternative = two.sided

 NOTE: n is number of *pairs*
```

43

G*Power



44

例題）　効果量と検定力を求める（無相関検定）

x	6	4	4	2	6	5	4	7	4	3
y	5	4	3	3	6	3	5	5	6	2

```
x <- c(6, 4, 4, 2, 6, 5, 4, 7, 4, 3)
y <- c(5, 4, 3, 3, 6, 3, 5, 5, 6, 2)
r <- cor(x, y)
pwr.r.test(n=length(x), r=r)
```

```
x <- c(6,4,4,2,6,5,4,7,4,3)
y <- c(5,4,3,3,6,3,5,5,6,2)
r <- cor(x,y)
pwr.r.test(n=length(x), r=r)
```

```
approximate correlation power
calculation (arctangh transformation)

        n = 10
        r = 0.5791078
        sig.level = 0.05
        power = 0.4096608
        alternative = two.sided
```


例題）　効果量と検定力を求める（χ^2 検定）

	はい	いいえ
男子	23	26
女子	12	19

```
dat <- matrix(c(23,26,12,19),ncol=2, byrow=T)
x <- chisq.test(dat,correct=F)
(a <- sqrt(x$statistic/sum(dat)))

pwr.chisq.test(N=sum(dat), df=1, w=a)
```

```
dat <- matrix(c(23,26,12,19),ncol=2, byrow=T)
x <- chisq.test(dat,correct=F)
(a <- sqrt(x$statistic/sum(dat)))
pwr.chisq.test(N=sum(dat), df=1, w=a)
```

```
Chi squared power calculation

     w = 0.08081479
     N = 80
     df = 1
     sig.level = 0.05
     power = 0.1116691

NOTE: N is the number of observations
```


G*Power