SASユーザー総会 アカデミア/テクノロジー&ソリューション セッション 2009

SASによる共分散分析

浜田知久馬

東京理科大学

ANCOVA using SAS

Chikuma Hamada
Tokyo University of Science

SASによる共分散分析

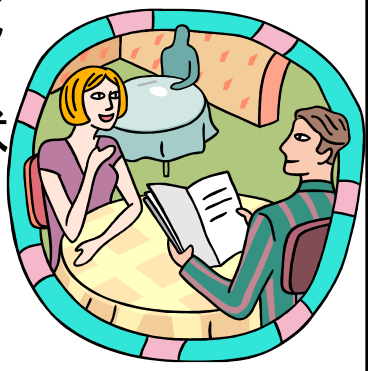
東京理科大学
浜田知久馬

SUGIJ2009

2009.7.31(金)

2

発表構成



- 医薬研究における現状
- 共変量調整の役割
- 交絡とは
- 共変量調整の原理
- 共分散分析のモデルと数理
- SASによる解析と解釈
- 共分散分析の適用例

3

医学論文における サブグループ解析・共変量調整の現状

STATISTICS IN MEDICINE 2002;21:2917-2930

Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems

Stuart J. Pocock, Susan E. Assmann, Laura E. Enos and Linda E. Kasten

4

対象ジャーナル

1群当たり50例以上の無作為化臨床試験
1997年7月～9月に公表された計50報

British Medical Journal (BMJ)5報
New England Journal of Medicine (NEJM)24報
Journal of the American Medical Association (JAMA)6報
Lancet　15報

5

Table I. Subgroup analyses in 50 clinical trial reports.

		Number of trials
Were subgroup analyses reported?	Yes	35
	No	15
Number of baseline factors included	One	17
	Two	3
	Three	3
	Four	5
	Five to six	2
	Seven or more	5
Number of outcomes for subgroup analysis	One	17
	Two	6
	Three to five	6
	Six or more	6
Total number of subgroup analyses	One	8
	Two	4
	Three to five	8
	Six to eight	9
	12 to 24	4
	Unclear	2
Statistical method used for subgroup analysis	Descriptive only	7
	Subgroup *P*-values	13
	Interaction test	15
Subgroup differences claimed	Yes	21
	No	14
If yes: Subgroup claim featured in abstract and/or conclusions	Yes	13
	No	8

Table II. Covariate adjustment in analysis of patient response by treatment.

		Number of trials
Were primary outcome analyses done using covariate adjustment?	No, unadjusted only	14
	Yes	36
Which analyses received more emphasis?	Unadjusted	38
	Covariate adjusted	11
	Equal emphasis	1
Number of covariates included	One	7
	Two	6
	Three	4
	Four	2
	Five to nine	11
	Ten or more	2
	Unclear	4
Did covariate adjusted analysis alter the trial conclusions, compared to unadjusted analyses?	No	29
	Yes	1
	Only unadjusted given	14
	Only adjusted given	6
Reasons for choice of covariates*		
	No reason given	15
	Covariates were (or expected to be) prognostic	12
	Covariates imbalanced between groups	5
	Centre or country adjusted for	4
	Baseline value of quantitative outcome	3
	Other treatment factor in a factorial trial	2
	Covariates used in stratified randomization	1

*More than one reason in some trials.

Table III. Baseline variables by treatment group in 50 clinical trial reports.

		Number of trials
Number of baseline variables compared	None	4
	One to four	1
	Five to nine	14
	Ten to 19	24
	20 to 29	5
	30 or more	2
Significance tests for baseline difference performed	Yes	24
	No	26
Baseline imbalances noted	Yes	17
	No	33

用語　分散分析と共分散分析

ANOVA:　ANalyisis Of VAriance
分散分析（群間比較）

ANCOVA: ANalyisis of COVAriance
共分散分析

反応変数に相関にある変数C(共変量）を考慮した分散分析ANOVA（群間比較）

9

分散分析

ANOVA:　ANalyisis Of VAriance

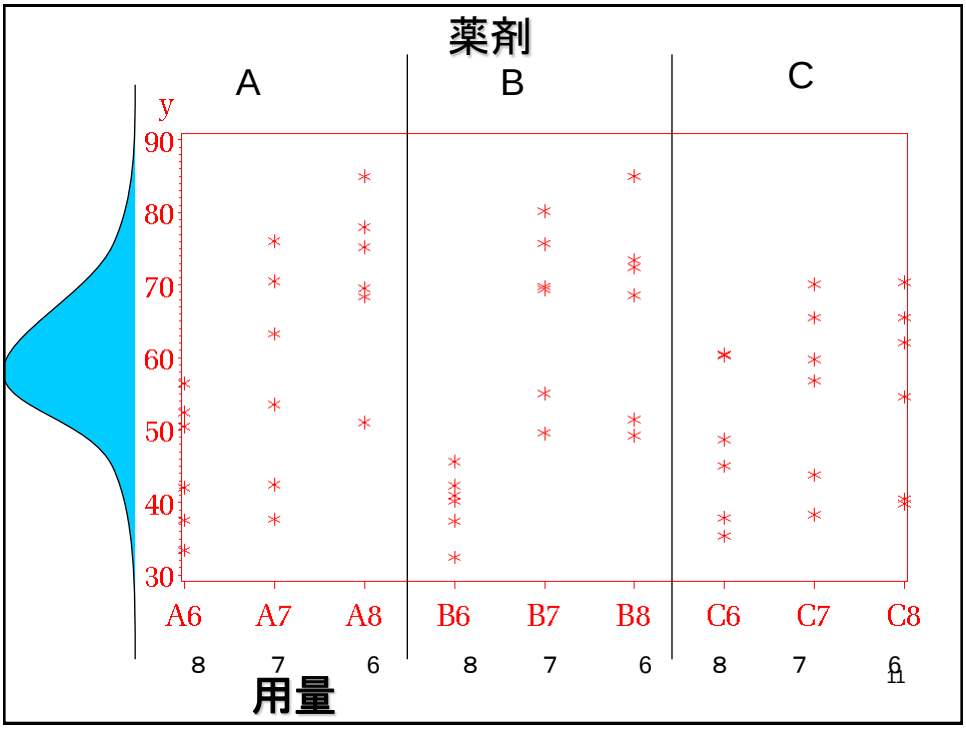
分散(variance):ばらつきを表す統計用語

分散分析:データのばらつきを構成する要因が複数存在するときに個々の要素に分解

二元配置(two-way)分散分析

要因の　例）　薬剤（A, B, C）
用量(6,7,8)

10



共変量調整の役割

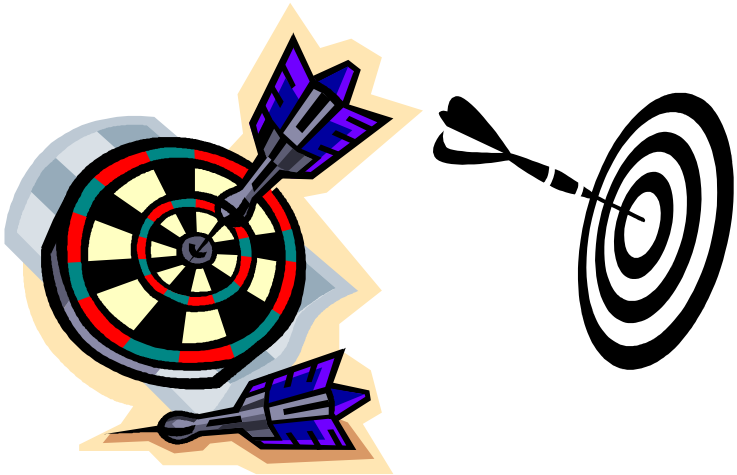
共変量: covariates　調整: adjust

1）群間の共変量の分布の偏りの除去
　交絡の調整
　群間の平均値の差を偏りなく推定

2）共変量に由来する誤差変動の除去
　解析の精度，検出力の向上，
　群間の平均値の差の分散の低下，
　信頼区間幅の減少

12

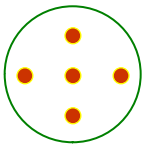
ダーツ投げの偏り（bias)と精度(precision)



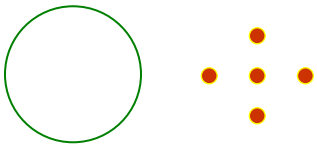
13

ダーツ投げの偏りと精度

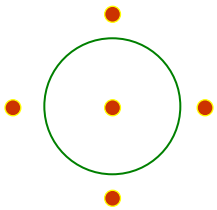
不偏で精度がよい



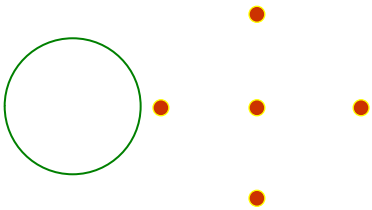
偏りあるけど精度はよい



不偏だけど精度が悪い

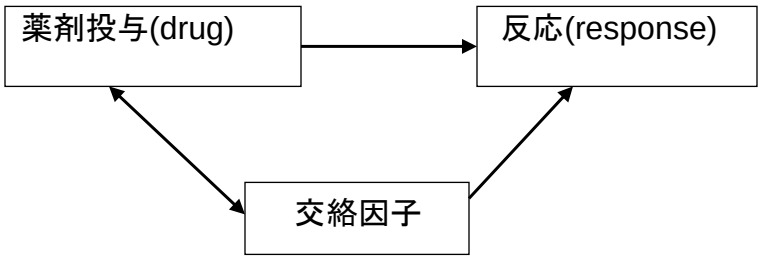


偏りありかつ精度が悪い



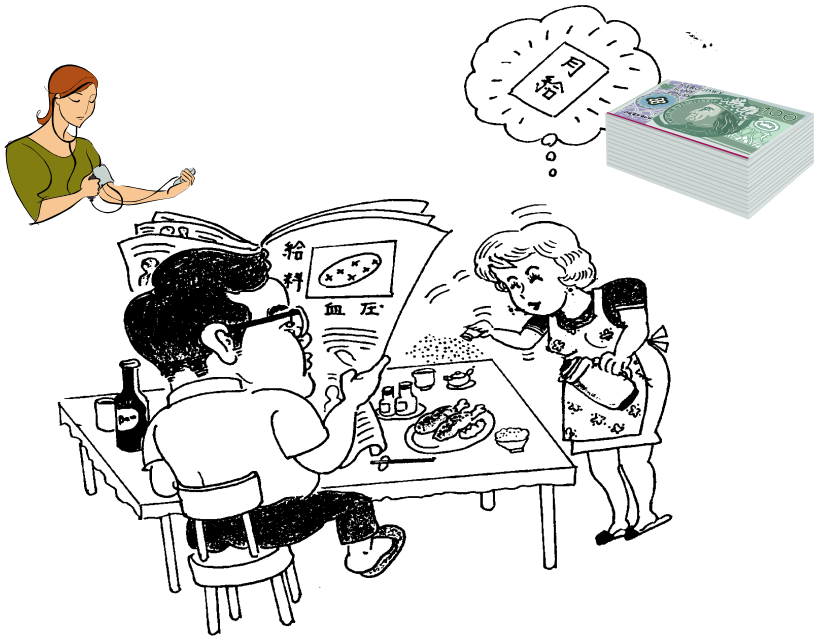
14

交絡(confounding)とは



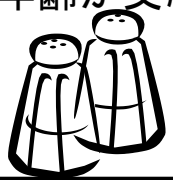
交絡因子(confounding factor)の必要条件
1)薬剤投与と関連する.
2)反応と関連する.
3)薬剤投与と反応の中間媒介変数ではない.

15



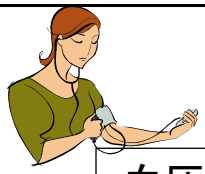
交絡(confounding)

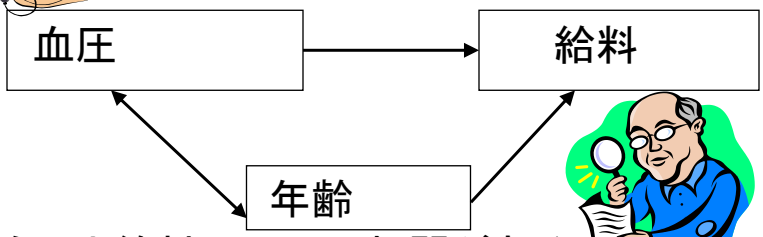
- 血圧と給料の間には正の相関がある.
- 給料を上げるためには塩辛いものを食べればよいのか？
- 年功序列的に給料は上がる（正の相関）.
- 年齢が上がれば血圧も上がる（正の相関）.
- 年齢という第3の因子が交って絡んで,
  見かけ上の相関が生じる. 年齢が交絡因子

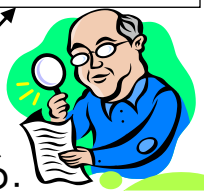


17

交絡の例







1)血圧と給料には正の相関がある.
2)年齢と血圧は正の相関がある.
3)年齢と給料には正の相関がある.
年齢が, 交絡して, 血圧と給料に見かけ上の正の相関を生じさせる .

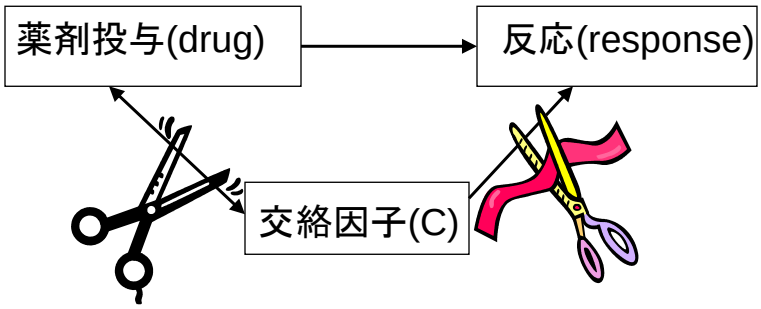
「新入社員は給料も血圧も低い. 部長は給料も血圧も高い. 」18

交絡に対する対処

1)デザインの工夫
　無作為化，層別割付，最小化割付
2)解析法の工夫
　a)共変量の群間の偏り，
　　共変量の反応変数との関連の解析
　b)デザインベースドの解析
　　層別，層を併合したMH流の調整解析
　c)モデルベースドの調整解析
　　重回帰，ロジスティック回帰，Cox回帰

19

交絡への対処



Cの群間の偏りの解析

無作為化割付

デザインベースドの解析

Cと反応変数

の関連評価

モデルによる
調整解析

20

調整した解析(adjusted analysis)

複数の交絡因子の影響を同時に考慮した解析

- 目的:交絡因子の偏りの除去，精度の増大
  調整しない解析の妥当性を保証
- 方法:1)デザインベースドの調整解析
  交絡因子の値が近い個体間で比較
  層別，層を併合した重み付き平均を算出，
  2)モデルベースドの調整解析
  共変量の影響のモデル化と除去
  重回帰，Cox回帰，ﾛｼﾞｽﾃｨｯｸ回帰

21

収縮期血圧の例(単位：mmHg)

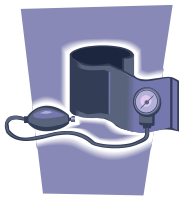
薬剤Ｐ(プラセボ群)	生データ										平均
年齢	39	41	42	43	43	44	45	46	47	49	43.9
血圧	95	99	106	111	115	116	101	117	104	119	108.3

薬剤Ａ(実薬群)	生データ										平均
年齢	36	37	38	39	39	40	41	42	44	47	40.3
血圧	88	94	96	92	98	89	103	97	110	105	97.2

血圧のt検定：p=0.0051 **

年齢のt検定：p=0.0200 *

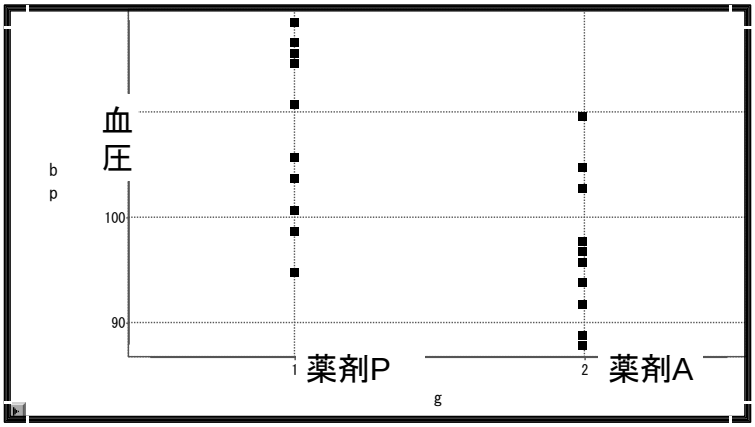
有意な結果は年齢の交絡のためか？



22

収縮期血圧のプロット

平均: 108.3　平均: 97.2



血圧のt検定:p=0.0051　**

23

GLMによる調整しない解析

```
data data;
do g=1 to 2;
   do i=1 to 10;
      input age bp @@;output;
end;end;
cards;
39 95  41  99  42 106  43 111  43 115
44 116 45 101  46 117  47 104  49 119
36 88  37  94  38  96  39  92  39  98
40 89  41 103  42  97  44 110  47 105
;
proc glm;
class g;
model bp=g;
lsmeans g/pdiff tdiff;
```

24

調整前　分散分析(t検定)の結果

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	616.050000	616.050000	10.16	0.0051
Error	18	1091.700000	60.650000		
Corrected Total	19	1707.750000	σ^2		

g	bp LSMEAN	95% Confidence Limits	
1	108.300000	103.126013	113.473987
2	97.200000	92.026013	102.373987

最小2乗平均

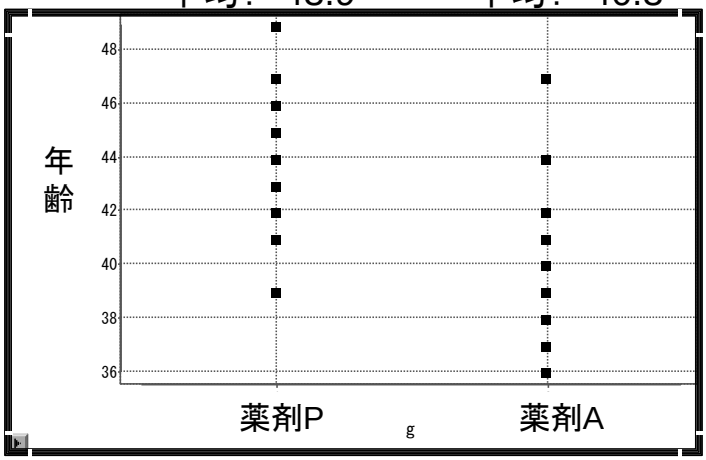
Least Squares Means for Effect g　平均値の差と信頼区間

i	j	Difference Between Means	95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	11.100000	3.782877	18.417123

25

年齢のプロット

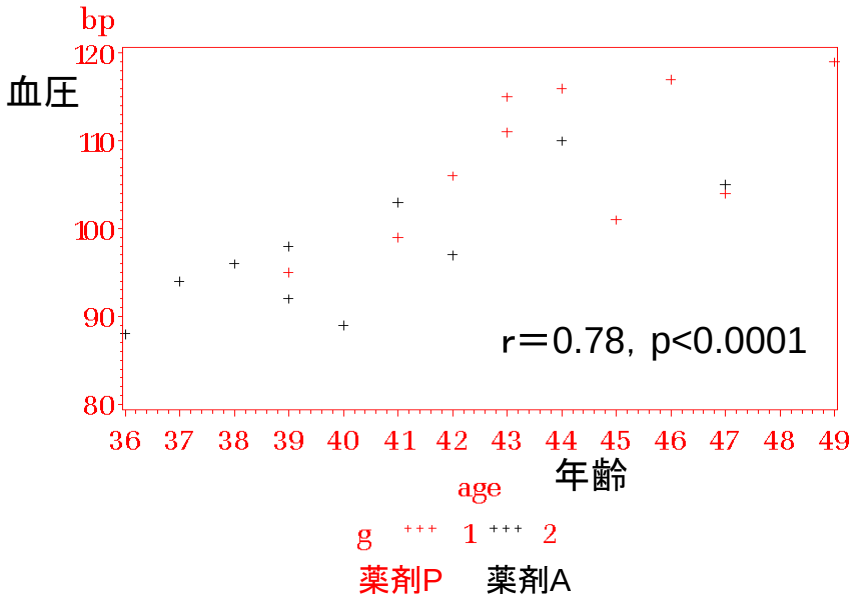
平均：43.9　　平均：40.3



年齢のt検定：p=0.0200　＊

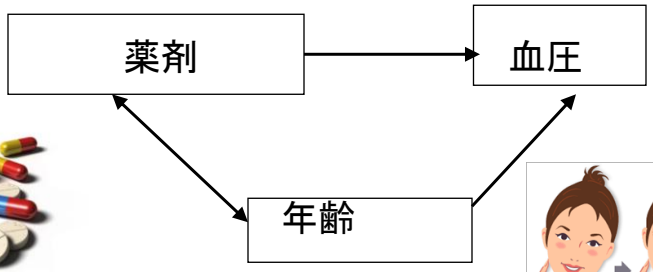
26

年齢と血圧の散布図



27

薬剤と血圧と年齢



1)薬剤群には年齢が若い人が多い.

2)年齢と血圧は正の相関がある.

3)年齢が，交絡して，薬剤と血圧に見かけ上の関連を生じさせる.

28

共変量調整の原理

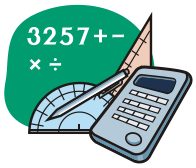
1）共変量の値が似ている個体同士で比較

マッチング

層別解析

2）数学モデルによる共変量の影響の除去

共分散分析



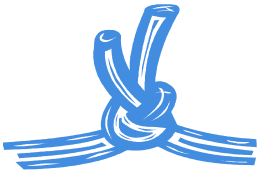
29

年齢の影響の調整　1)マッチング

薬剤 P	生データ										平均
年齢	39	41	42	43	43	44	45	46	47	49	43.9
血圧	95	99	106	111	115	116	101	117	104	119	108.3

薬剤 A	生データ										平均
年齢	36	37	38	39	39	40	41	42	44	47	40.3
血圧	88	94	96	92	98	89	103	97	110	105	97.2

- 年齢が同じ個体を集めて比較.
- 年齢分布が強制的にそろえられる.
- マッチしない個体が多くなる.



30

年齢の影響の調整　2)層別

薬剤Ｐ　　生データ　　平均

年齢	39	41	42	43	43	44	45	46	47	49	43.9
血圧	95	99	106	111	115	116	101	117	104	119	108.3
	95		108.0				113.3				

薬剤Ａ　　生データ　　平均

年齢	36	37	38	39	39	40	41	42	44	47	40.3
血圧	88	94	96	92	98	89	103	97	110	105	97.2
		93.6				99.75		105			

年齢の層ごとに平均値の差

→ Ｎで重みを付けた併合平均値の差

$$V[\hat{\mu}_{i2}-\hat{\mu}_{i1}]=\frac{\sigma^2}{n_{i2}}+\frac{\sigma^2}{n_{i1}},\ w_i=\frac{n_{i1}n_{i2}}{n_{i1}+n_{i2}}$$

31

年齢の影響の調整　2)層別

```
data data;set data;
    if age <= 39     then agec=1;
    if 40<=age <=44 then agec=2;
    if 45<=age       then agec=3;
proc glm;
class agec g;
model bp=agec g ;
lsmeans g/pdiff tdiff;
```

32

年齢の影響の調整　2)層別

年齢層	薬剤Ｐ (N)	薬剤Ａ (N)	差	重みw
30代	95.0 (1)	93.6 (5)	-1.4	$\frac{1\cdot 5}{1+5}=0.833$
40代前	108.0 (6)	99.75 (4)	-8.25	$\frac{6\cdot 4}{6+4}=2.4$
40代後	113.3 (3)	105 (1)	-8.3	$\frac{3\cdot 1}{3+1}=0.75$
併合	108.3 (10)	97.2 (10)	-6.95 p=0.069	

33

年齢の影響の調整　3)共分散分析

薬剤Ｐ　　生データ　　平均

年齢　39　41　42　43　43　44　45　46　47　49　43.9

血圧　95　99　106　111　115　116　101　117　104　119　108.3

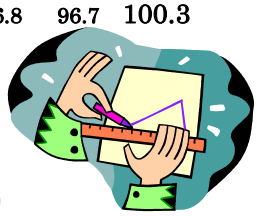
修正血圧　100.3　100.9　106.2　109.5　113.5　112.8　96.0　110.4　95.7　107.3　105.2

薬剤Ａ　　生データ　　平均

年齢　36　37　38　39　39　40　41　42　44　47　40.3

血圧　88　94　96　92　98　89　103　97　110　105　97.2

修正血圧　98.4　102.7　103.0　97.3　103.3　92.6　104.9　97.2　106.8　96.7　100.3



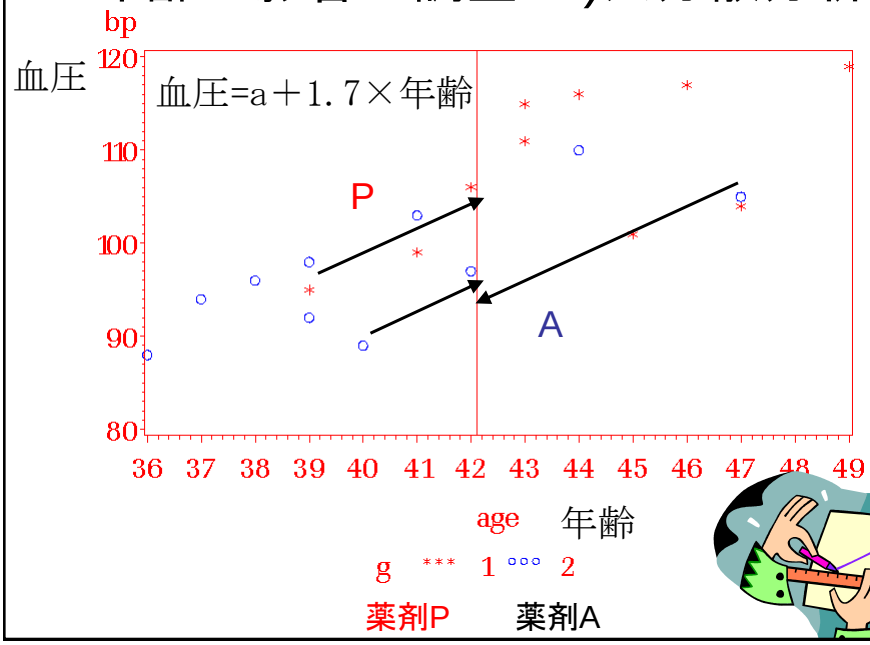
血圧=a＋b×年齢＋誤差　b=1.7

年齢を平均42.1歳に調整

修正平均の差　105.2−100.3=4.9 (p=0.1192)

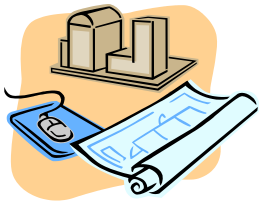
34

年齢の影響の調整　3)共分散分析



年齢の影響の調整　3)共分散分析

```
proc glm;
class g;
model bp=age g;
lsmeans g/pdiff tdiff;
```



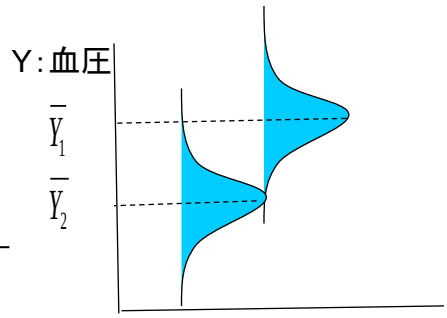
36

調整しない平均値の差と分散と信頼区間CL(Confidence Limit)

$$\hat{\mu}_1 = \overline{Y_1} \quad \hat{\mu}_2 = \overline{Y_2}$$

$$\hat{\mu}_2 - \hat{\mu}_1 = \overline{Y_2} - \overline{Y_1}$$

$$V[\hat{\mu}_2 - \hat{\mu}_1] = \frac{\sigma^2}{n_2} + \frac{\sigma^2}{n_1}$$



$$CL = \hat{\mu}_2 - \hat{\mu}_1 \pm t_{DF}\sqrt{\frac{\sigma^2}{n_2} + \frac{\sigma^2}{n_1}}$$

Z:年齢

37

調整前の平均値の差

$$\hat{\mu}_2 - \hat{\mu}_1 = \overline{Y_2} - \overline{Y_1} = 97.2 - 108.3 = -11.1$$

$$V[\hat{\mu}_2 - \hat{\mu}_1] = \frac{\sigma^2}{n_2} + \frac{\sigma^2}{n_1} = \frac{60.65}{10} + \frac{60.65}{10}$$

$$= 12.13 = 3.48^2 = SE[\hat{\mu}_2 - \hat{\mu}_1]^2$$

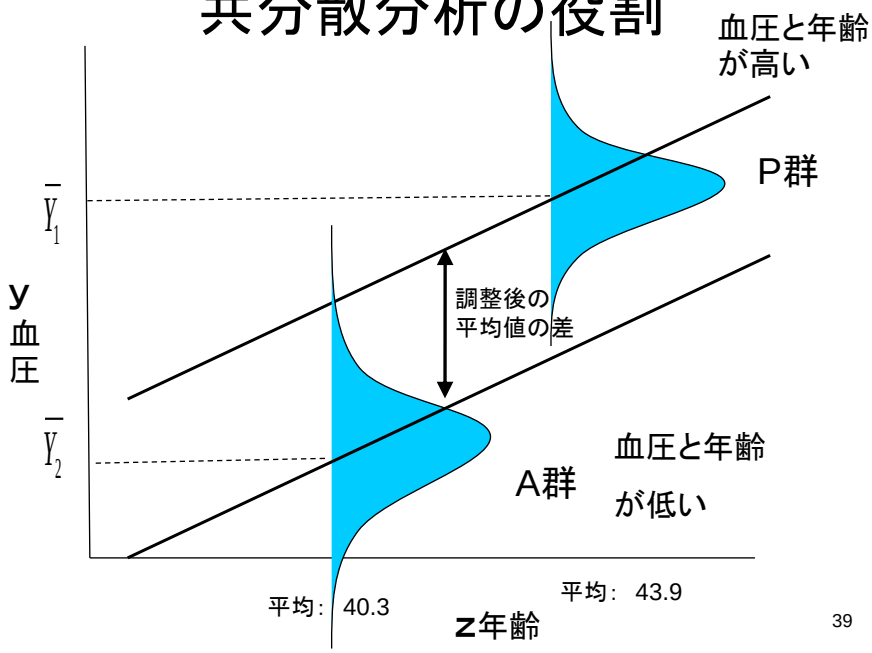
$$CL = \hat{\mu}_2 - \hat{\mu}_1 \pm t_{18}\sqrt{\frac{\sigma^2}{n_2} + \frac{\sigma^2}{n_1}}$$

$$= -11.1 \pm 2.101 \cdot 3.48 = -18.42, -3.78$$

$$p = 0.0051$$

38

共分散分析の役割

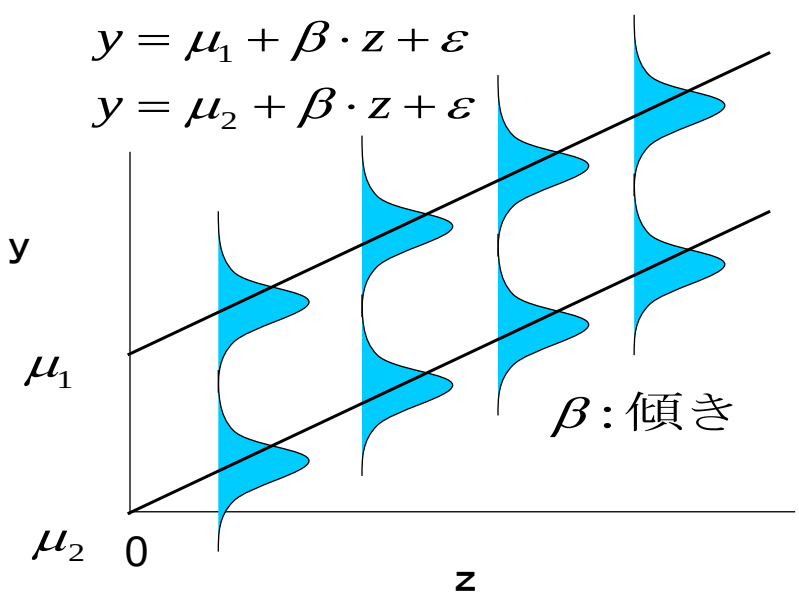


39

共分散分析のモデル

$$y = \mu_1 + \beta \cdot z + \varepsilon$$

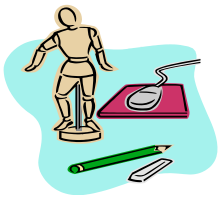
$$y = \mu_2 + \beta \cdot z + \varepsilon$$



40

共分散分析のモデル

$$Y_{ij} = \mu_i + \beta \cdot z_{ij} + \varepsilon_{ij}$$



i : 群，　j　：個体

Y : 反応変数，z : 共変量

ε_{ij}: $\mathrm{N}(0, \sigma^2)$

独立に等分散の正規分布

41

最小二乗推定

$$SS = \sum_i \sum_j (Y_{ij} - \mu_i - \beta \cdot z_{ij})^2$$

$$\frac{dSS}{d\beta} = \sum_i \sum_j z_{ij} (Y_{ij} - \mu_i - \beta \cdot z_{ij})$$

$$\frac{dSS}{d\mu_i} = -\sum_i (Y_{ij} - \mu_i - \beta \cdot z_{ij}) = 0$$

$$\hat{\mu}_i = \overline{Y}_i - \hat{\beta} \cdot \overline{z}_i$$

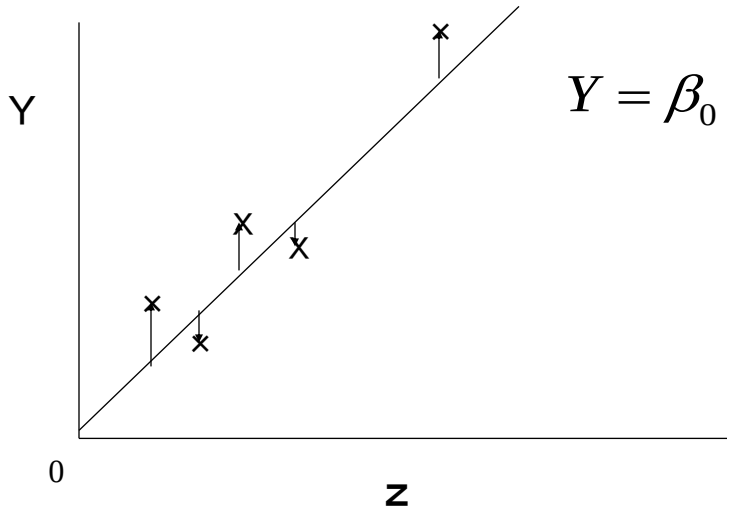
e_1^2

e_2^2

e_3^2

42

最小2乗法の模式図



$$Y = \beta_0 + \beta_1 z$$

43

共通の傾きの最小二乗推定

$$SS = \sum_i \sum_j (Y_{ij} - \mu_i - \beta \cdot z_{ij}))^2$$

$$= \sum_i \sum_j (Y_{ij} - \overline{Y}_i - \beta \cdot (z_{ij} - \overline{z}_i))^2$$

$$\frac{dSS}{d\beta} = \sum_i \sum_j -(z_{ij} - \overline{z}_i)(Y_{ij} - \overline{Y}_i - \beta \cdot (z_{ij} - \overline{z}_i)) = 0$$

$$\sum_i \sum_j -(z_{ij} - \overline{z}_i)(Y_{ij} - \overline{Y}_i) + \sum_i \sum_j \beta \cdot (z_{ij} - \overline{z}_i)^2 = 0$$

$$\hat{\beta} = \frac{\sum_i \sum_j (z_{ij} - \overline{z}_i)(Y_{ij} - \overline{Y}_i)}{\sum_i \sum_j (z_{ij} - \overline{z}_{i.})^2} = \frac{S_{zy1} + S_{zy2}}{S_{zz1} + S_{zz2}}$$

44

共通の傾きの最小二乗推定

$$\hat{\beta} = \frac{S_{zy1} + S_{zy2}}{S_{zz1} + S_{zz2}} = \frac{S_{zz1}\dfrac{S_{zy1}}{S_{zz1}} + S_{zz1}\dfrac{S_{zy1}}{S_{zz1}}}{S_{zz1} + S_{zz2}}$$

$$= \frac{\dfrac{1}{V[\hat{\beta}_1]}\hat{\beta}_1 + \dfrac{1}{V[\hat{\beta}_2]}\hat{\beta}_2}{\dfrac{1}{V[\hat{\beta}_1]} + \dfrac{1}{V[\hat{\beta}_2]}} = \frac{w_1\hat{\beta}_1 + w_2\hat{\beta}_2}{w_1 + w_2}$$

$$w_i = \frac{1}{V[\hat{\beta}_i]}$$

45

調整後の平均値の差と分散

$$\hat{\mu}_1 = \overline{Y}_1 - \hat{\beta}\cdot\overline{z}_1 \qquad \hat{\mu}_2 = \overline{Y}_2 - \hat{\beta}\cdot\overline{z}_2$$

$$\hat{\mu}_2 - \hat{\mu}_1 = \overline{Y}_2 - \overline{Y}_1 - \hat{\beta}\cdot(\overline{z}_2 - \overline{z}_1)$$

$$V[\hat{\mu}_2 - \hat{\mu}_1] = \frac{\sigma^2{}_{adj}}{n_2} + \frac{\sigma^2{}_{adj}}{n_1} + (\overline{z}_2 - \overline{z}_1)^2 V[\hat{\beta}]$$

46

調整前と調整後の比較

調整前

$$\hat{\mu}_2 - \hat{\mu}_1 = \overline{Y}_2 - \overline{Y}_1$$

$$V[\hat{\mu}_2 - \hat{\mu}_1] = \frac{\sigma^2}{n_2} + \frac{\sigma^2}{n_1}$$

調整後

$$\hat{\mu}_2 - \hat{\mu}_1 = \overline{Y}_2 - \overline{Y}_1 - \boxed{\hat{\beta}\cdot(\overline{z}_2 - \overline{z}_1)}$$

$$V[\hat{\mu}_2 - \hat{\mu}_1] = \frac{\sigma_{adj}^2}{n_2} + \frac{\sigma_{adj}^2}{n_1} + \boxed{(\overline{z}_2 - \overline{z}_1)^2 V[\hat{\beta}]}$$

47

調整前と調整後の比較

1) $\overline{z}_2 = \overline{z}_1$であれば調整前後の平均値の差は等しくなる.

無作為化で偏りの除去が期待できる.

2)zが重要な共変量であれば$\sigma^2 > \sigma_{adj}^2$

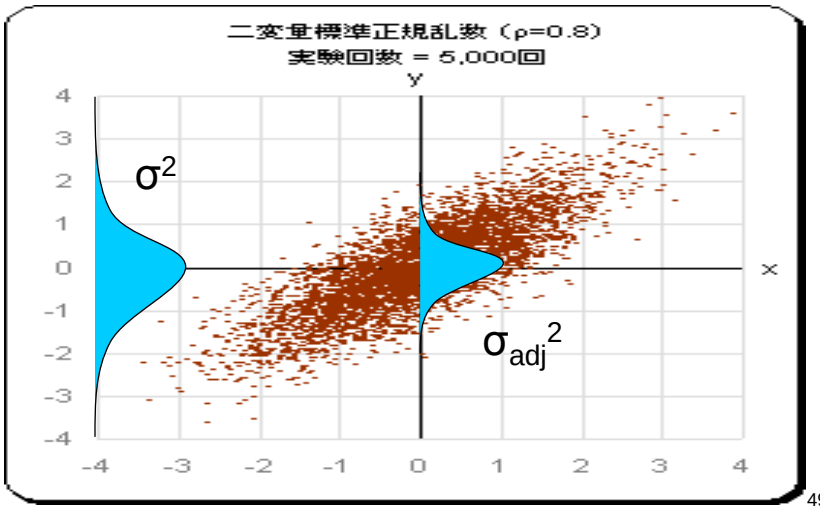
$\overline{z}_2 = \overline{z}_1$ であれば調整後の平均値の差の方が分散は小さくなる.

偏りがなくても共分散分析により，精度，

検出力が向上.

48

調整前と調整後の分散



49

調整前　分散分析(t検定)の結果

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	616.050000	616.050000	10.16	0.0051
Error	18	1091.700000	60.650000		
Corrected Total	19	1707.750000			

σ^2

g	bp LSMEAN	95% Confidence Limits	
1	108.300000	103.126013	113.473987
2	97.200000	92.026013	102.373987

最小2乗平均

Least Squares Means for Effect g　平均値の差と信頼区間

i	j	Difference Between Means	95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	11.100000	3.782877	18.417123

50

調整後　共分散分析の結果

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1134.720894	567.360447	16.83	<.0001
Error	17	573.029106	33.707594		
Corrected Total	19	1707.750000			

σ_{adj}^2

g	bp LSMEAN	95% Confidence Limits	
1	105.235978	101.026440	109.445515
2	100.264022	96.054485	104.473560

最小2乗平均

$\sigma^2 = 60.65$

Least Squares Means for Effect g　平均値の差と信頼区間				
i	j	Difference Between Means	95% Confidence Limits for LSMean(i)−LSMean(j)	
1	2	4.971955	−1.421177	11.365088

51

調整後の平均値の差

$$\hat{\beta} = 1.7,\ \bar{z}_1 = 43.9,\ \bar{z}_2 = 40.3$$

$$\hat{\mu}_2 - \hat{\mu}_1 = \bar{Y}_2 - \bar{Y}_1 - \hat{\beta}\cdot(\bar{z}_2 - \bar{z}_1)$$

$$= -11.1 - 1.7(40.3 - 43.9) = -5.0$$

$$V[\hat{\mu}_2 - \hat{\mu}_1] = \frac{\sigma_{adj}^2}{n_2} + \frac{\sigma_{adj}^2}{n_1} + (\bar{z}_2 - \bar{z}_1)^2 V[\hat{\beta}]$$

$$= \frac{33.7^2}{10} + \frac{33.7^2}{10} + (40.3 - 43.9)^2 V[\hat{\beta}]$$

$$= 9.18 = 3.03^2$$

52

調整前と調整後の比較

調整前

$$\hat{\mu}_2 - \hat{\mu}_1 = \overline{Y_2} - \overline{Y_1} = 97.2 - 108.3 = -11.1$$

$$CL = -11.1 \pm 2.101 \cdot 3.48 = -18.42, -3.78$$

$$p = 0.0051$$

調整後

$$\hat{\mu}_2 - \hat{\mu}_1 = -5.0$$

$$CL = -5.0 \pm 2.110 \cdot 3.03 = -11.37, 1.42$$

$$p = 0.1192$$

0

53

傾きの異なるモデル

$$Y_{ij} = \mu_i + \beta_i \cdot z_{ij} + \varepsilon_{ij}$$

i : 群,　j　: 個体

Y : 反応変数,　z : 共変量

ε_{ij}: $N(0, \sigma^2)$

54

最小二乗推定

$$SS = \sum_i \sum_j (Y_{ij} - \mu_i - \beta_i \cdot z_{ij}))^2$$

$$\frac{dSS}{d\mu_i} = -\sum_i (Y_{ij} - \mu_i - \beta_i \cdot z_{ij})) = 0$$

$$\hat{\mu}_i = \overline{Y_i} - \beta_i \cdot \overline{z}_i$$

55

最小二乗推定

$$SS = \sum_i \sum_j (Y_{ij} - \mu_i - \beta_i \cdot z_{ij}))^2$$

$$= \sum_i \sum_j (Y_{ij} - \overline{Y_i} - \beta_i \cdot (z_{ij} - \overline{z}_i))^2$$

$$\frac{dSS}{d\beta_i} = \sum_j -(z_{ij} - \overline{z}_i)(Y_{ij} - \overline{Y_i} - \beta \cdot (z_{ij} - \overline{z}_i)) = 0$$

$$\sum_j -(z_{ij} - \overline{z}_i)(Y_{ij} - \overline{Y_i}) + \sum_j \beta \cdot (z_{ij} - \overline{z}_i)^2 = 0$$

$$\hat{\beta}_i = \frac{\sum_j (z_{ij} - \overline{z}_i)(Y_{ij} - \overline{Y_i})}{\sum_j (z_{ij} - \overline{z}_i)^2} = \frac{S_{zyi}}{S_{zzi}}$$

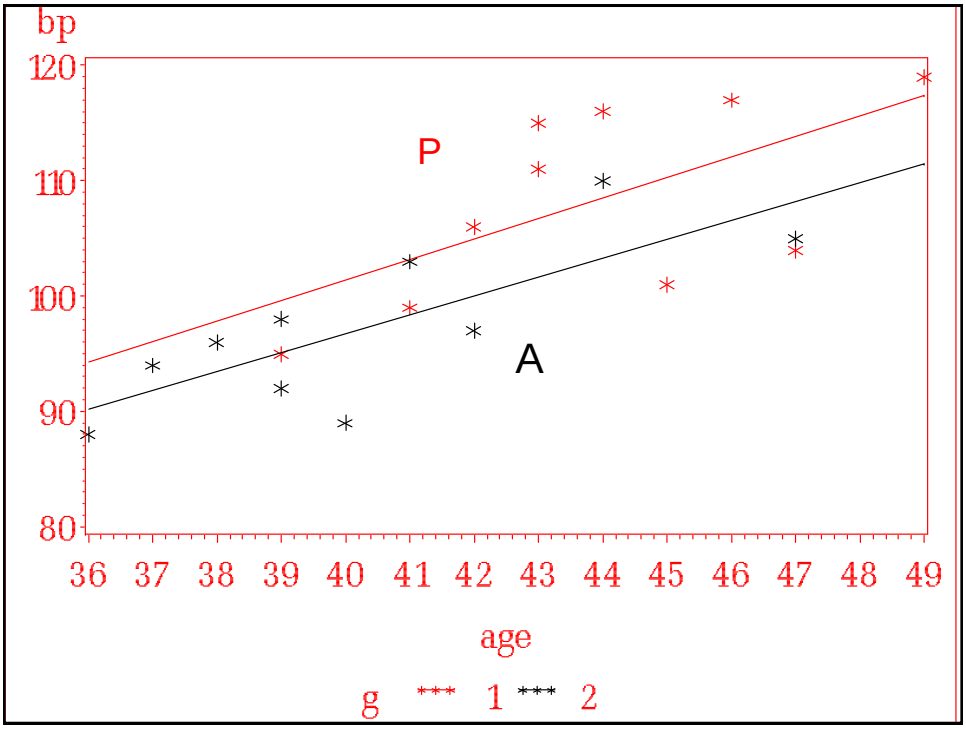
56

傾きの分散

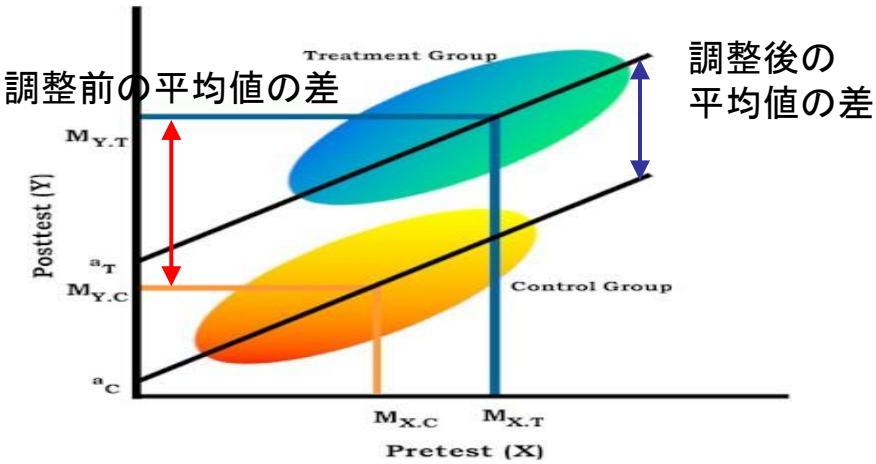
$$\hat{\beta}_i = \frac{\sum_j (z_{ij} - \bar{z}_i)(Y_{ij} - \bar{Y}_i)}{\sum_j (z_{ij} - \bar{z}_i)^2} = \frac{\sum_j (z_{ij} - \bar{z}_i)(Y_{ij})}{\sum_j (z_{ij} - \bar{z}_i)^2}$$

$$V[\hat{\beta}_i] = \frac{\sum_j (z_{ij} - \bar{z}_i)^2 \sigma^2}{\left(\sum_j (z_{ij} - \bar{z}_i)^2\right)^2} = \frac{\sigma^2}{\sum_j (z_{ij} - \bar{z}_i)^2} = \frac{\sigma^2}{S_{zzi}}$$

57

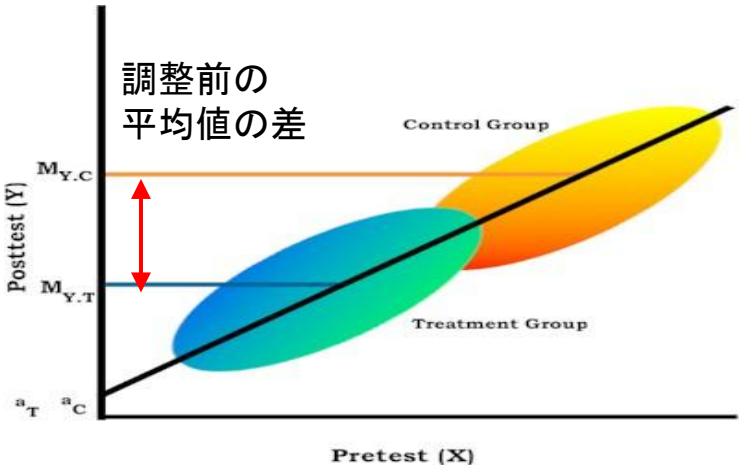


共分散分析（ANCOVA）



59

調整によって平均値の差がなくなる例



60

共分散分析の仮定

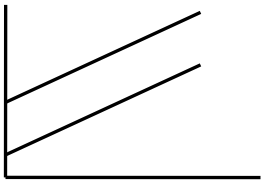
1)2つの群で共に，年齢の血圧との間には直線的な関係がある.

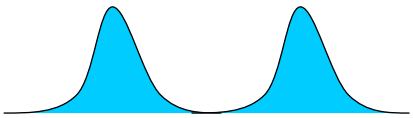
2)2つの群間で，傾きが等しい.

群×血圧の交互作用が存在しない.

3)誤差の分布に等分散性と正規性

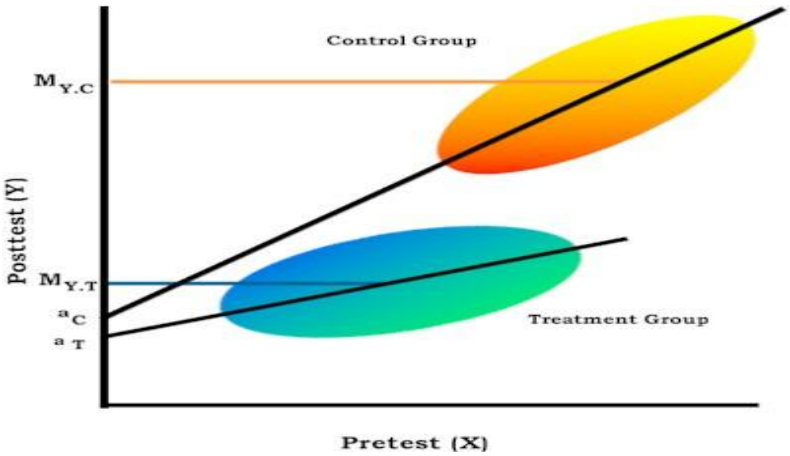
$$Y_{ij} = \mu_i + \beta \cdot z_{ij} + \varepsilon_{ij}, \quad \varepsilon_{ij}: \mathrm{N}(0, \sigma^2)$$





61

傾きが異なる例



62

共分散分析の適用例

1)経時測定データ解析における投与前値を
共変量とした共分散分析
2）臓器重量の解析における体重を
共変量とした共分散分析

63

経時データの共分散分析
投与前値Y_1の扱い

1）投与後の値(post：Y_2)と同様に単純に
モデルに含めて解析.
2）投与後の値だけで解析.
3）投与前値の差をとって解析.
4）投与前値との比をとって解析.
5）独立変数（共変量）としてモデルに含めて
解析. ANCOVA）.

64

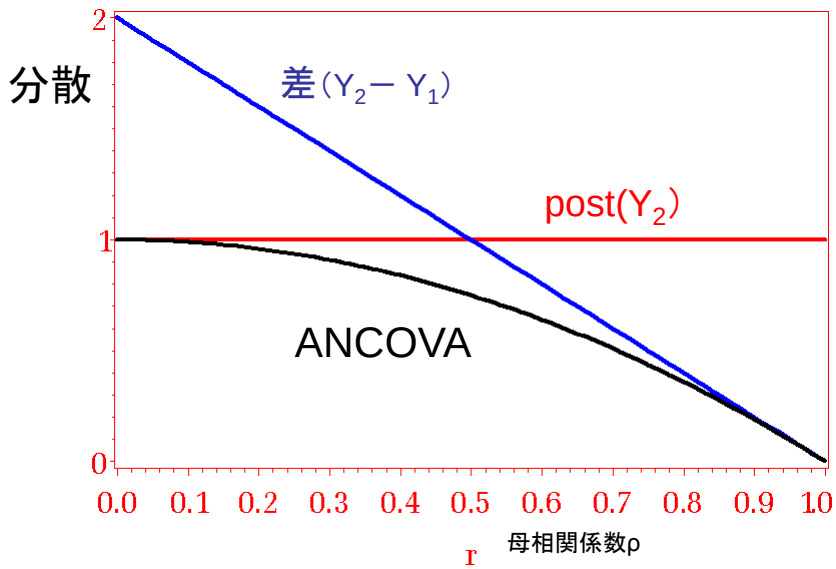
投与後値，差，共分散分析の比較

統計量	分散
post(Y_2)	σ^2
差（$Y_2 - Y_1$）	$\sigma^2+\sigma^2-2\rho\sigma^2$
ANCOVA($Y_2-\beta Y_1$)	$\sigma^2-\rho^2\sigma^2$

母相関係数ρ	ρ=0	ρ=0.5	ρ=1.0
post(Y_2)	σ^2	σ^2	σ^2
差（$Y_2 - Y_1$）	$2\sigma^2$	σ^2	0
ANCOVA	σ^2	$0.75\sigma^2$	0

65

投与後差，共分散分析の比較



66

共分散分析の利点

ANCOVA($Y_2 - \beta Y_1$)

1）$\rho = 0 \rightarrow \beta = 0$　post(Y_2)と等価

2）$\rho \fallingdotseq 1 \rightarrow 1 = 0$　差（$Y_2 - Y_1$）と等価

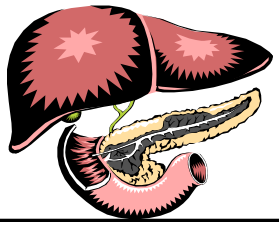
3）$0 < \rho < 1$　精度が高い

（$Y_2 - Y_1$）は$\rho > 0.5$以上ないと post(Y_2)より精度が悪い

67

ラットの体重（単位：g）と肝臓重量（単位：g）

用量 (mg)	項目										
0	体重	336.2	360.4	329.6	340.7	358.2	325.5	351	373.5	346.4	361
	肝臓	10.32	11.47	10.54	10.46	11.05	9.74	11.18	11.72	11.21	11.12
1	体重	365.4	360.5	336.1	319	321.1	351.4	343	369.8	340.6	341.2
	肝臓	12.39	10.72	10.61	9.7	9.09	10.9	10.99	11.35	9.96	10.28
3	体重	332.5	322.7	307.7	310.9	299.4	316.9	318.9	301.8	310.7	.
	肝臓	12.06	11.79	10.8	10.93	9.98	11.88	11.28	11.07	11.12	.
10	体重	261	288	287.9	303.9	285.3	269.2	266	279.2	239.7	319.8
	肝臓	10.44	11.68	12.86	12.62	11.58	11.15	11.12	11.35	9.97	13.52

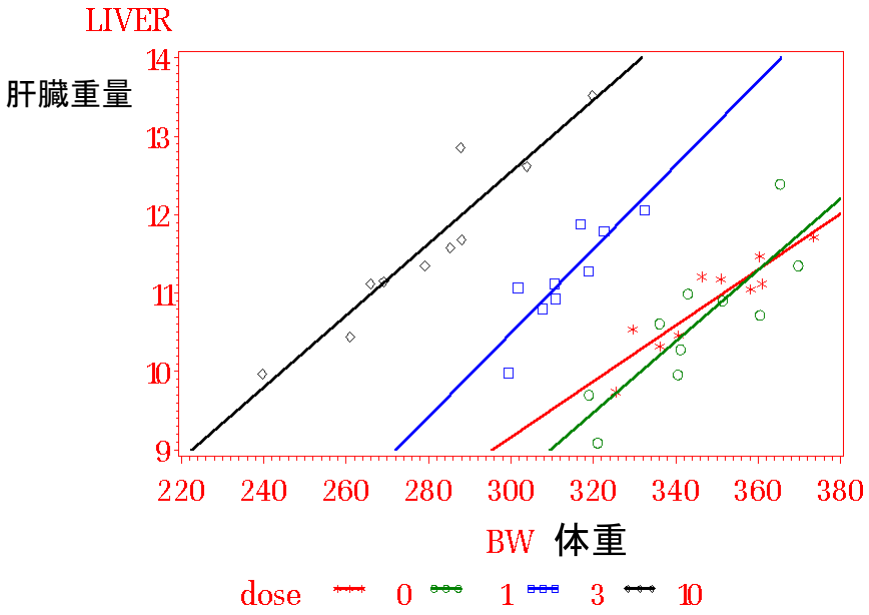


68

共分散分析の結果

	体重		肝臓 調整前		肝臓 調整後
群	平均	SD	平均	SD	平均
0	348.3	15.4	10.88	0.6	9.71
1	344.8	17.2	10.6	0.92	9.58
3	313.5**	10.4	11.21	0.64	11.58
10	280.0**	22.6	11.63	1.1	13.48

69



70

調整前後の解析プログラム

```
proc glm data=ow;
class dose;
model liver=dose/ss2;
lsmeans dose/tdiff pdiff adj=dunnett;
run;
proc glm data=ow;
class dose;
model liver=bw dose/ss2;
lsmeans dose/tdiff pdiff adj=dunnett;
run;
```

71

調整前の解析結果

Source	DF	Type II SS	Mean Square	F Value	Pr > F
dose	3	5.88060521	1.96020174	2.75	0.0575

dose	LIVER LSMEAN	H0:LSMean=Control	
		t Value	Pr > |t|
0	10.8810000		
1	10.5990000	−0.75	0.7990
3	11.2122222	0.85	0.7307
10	11.6290000	1.98	0.1371

72

調整後の解析結果

Source	DF	Type II SS	Mean Square	F Value	Pr > F
BW	1	20.08176532	20.08176532	139.39	<.0001
dose	3	25.89666101	8.63222034	59.92	<.0001

dose	LIVER LSMEAN	H0:LSMean=Control	
		t Value	Pr > |t|
0	9.7141122		
1	9.5841538	−0.76	0.7761
3	11.5812201	8.58	<.0001
10	13.4786359	12.27	<.0001

73

共変量選択方針

薬剤を投与する前に測定

1)反応変数に関連（予後因子）

2)群間で偏り（交絡可能性因子）

3)割り付け調整因子

4）施設・国

5)投与前値

薬剤を投与後に測定される結果系変数で調整する場合は結果の解釈は困難

74