生命科学の統計学

第３部

Katsumi Wakabayashi, Ph. D.
Prof. Emer. Gunma University
Technical consultant, Shibayagi Co. Ltd.

分散分析

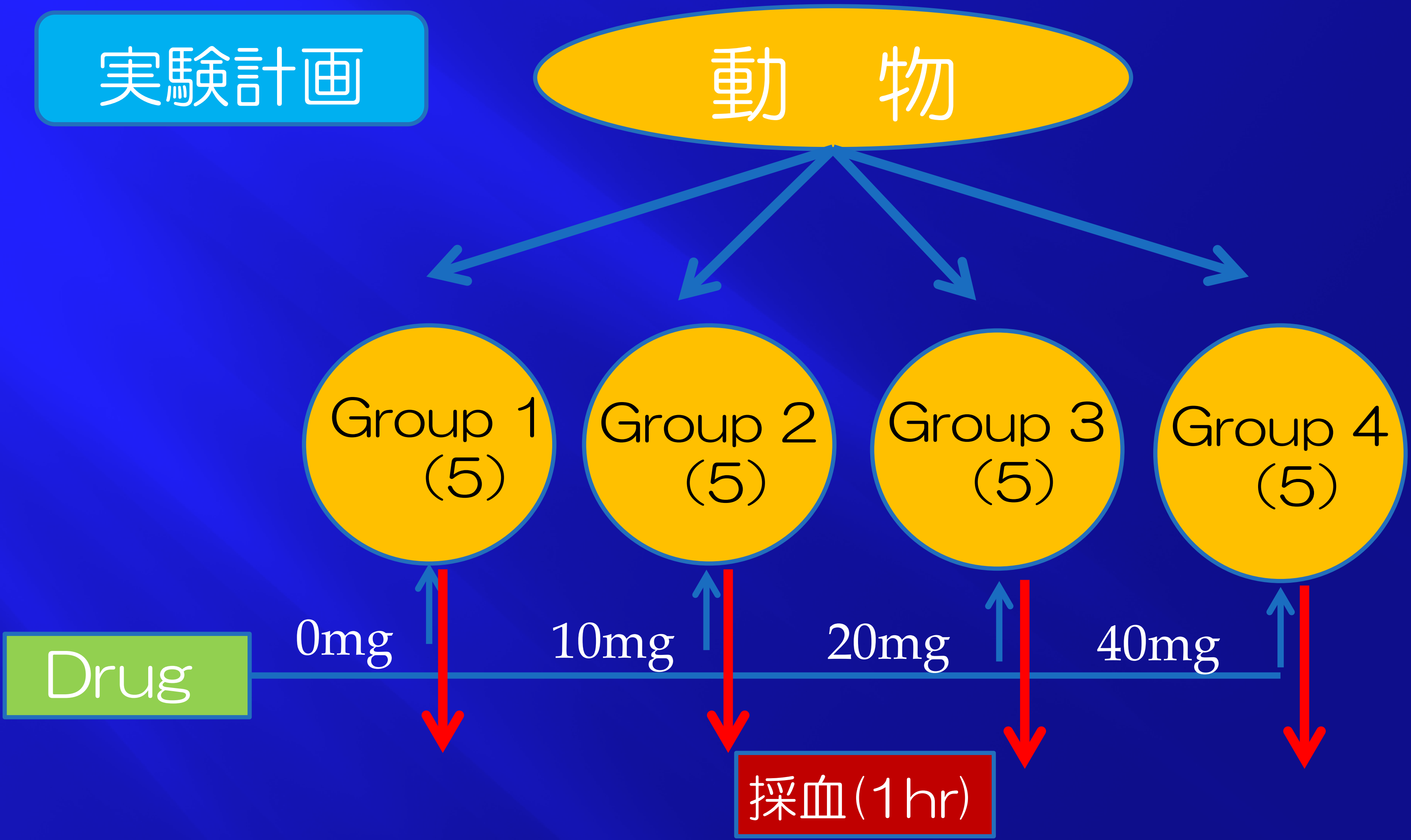
一元配置分散分析

：一つの要因の影響の有無を調べる

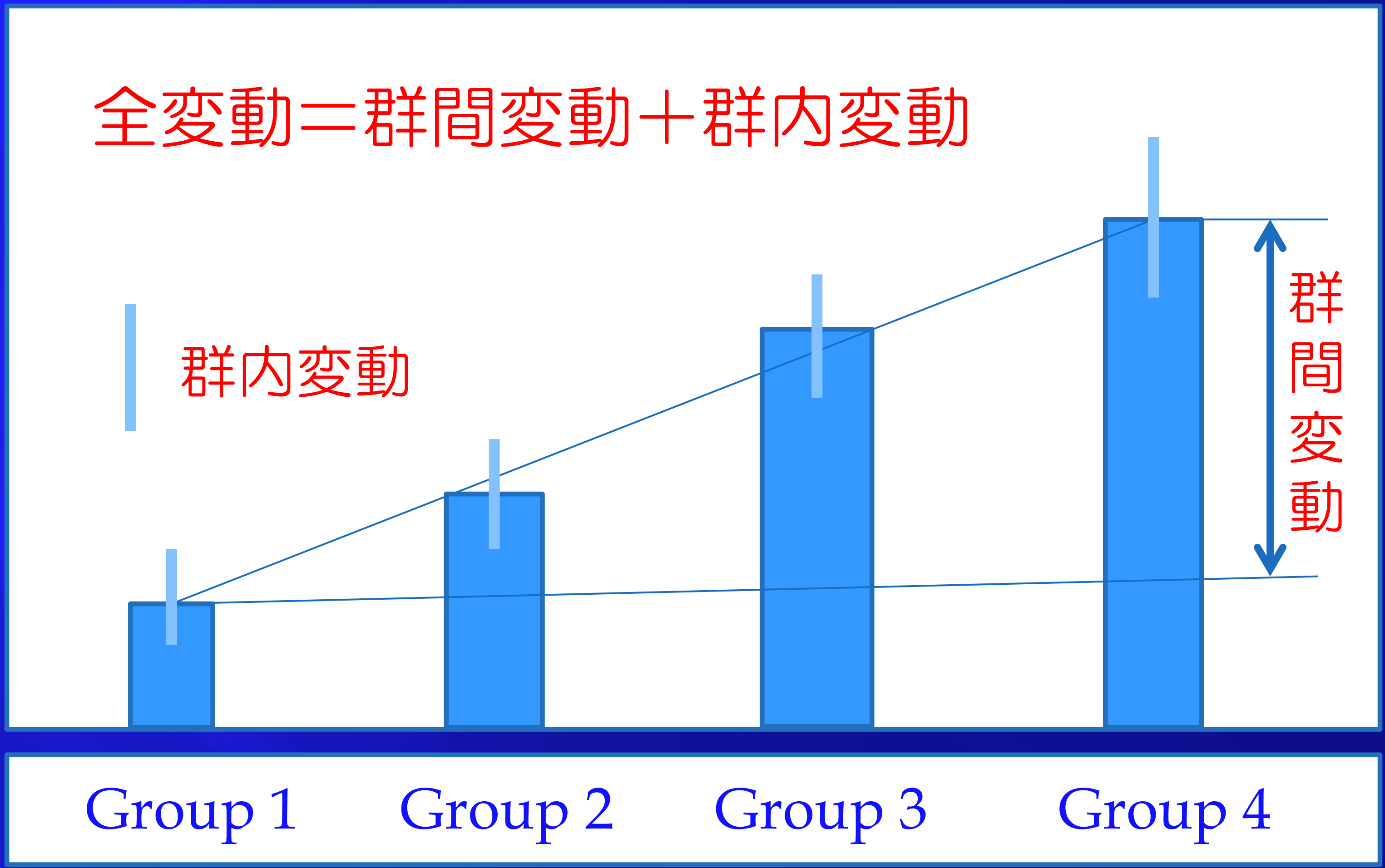
二元配置分散分析

：二つの要因の，単独での作用と相互作用を調べる

一元配置分散分析の考え方



一元配置分散分析の考え方



群間変動は群内変動と比べて有意に大きいか？
を F 分布から判定する

実験結果

Group 1	Group 2	Group 3	Group 4	Sum
X_{11} X_{12} X_{13} X_{14} X_{15}	X_{21} X_{22} X_{23} X_{24} X_{25}	X_{31} X_{32} X_{33} X_{34} X_{35}	X_{41} X_{42} X_{43} X_{44} X_{45}	i=1~4 j=1~5
T1 $\bar{X}1 = T1/r$	T2 $\bar{X}2 = T2/r$	T3 $\bar{X}3 = T3/r$	T4 $\bar{X}4 = T4/r$	$SS_A = r\sum(\overline{Xi} - \bar{\bar{X}})^2$ (群間変動) $SS_{AR} = \sum(xij - \bar{\bar{X}})^2$ (全変動)
SS1	SS2	SS3	SS4	$SS_{R(A)} = \sum SSi$ 群内変動=誤差変動

Tは“total”即ち総和のことである

具体例

Rat No.	Group 1 (0mg)	Group 2 (10mg)	Group 3 (20mg)	Group 4 (40mg)	変動Total
1	11	16	19	23	
2	13	14	21	28	
3	12	13	23	29	
4	14	15	18	27	
5	11	18	21	25	全変動 642.95
平均値	12.2	15.2	20.4	26.4	
群内変動	6.8	14.8	15.20	23.04	郡内変動 59.84

全変動＝$\sum\sum X_{ij}{}^{2} - (\sum\sum X_{ij})^{2}/N$＝642.95
（平均値を用いず上のような計算式が便利である）
群間変動＝全変動－群内変動＝642.95－59.84＝583.11

分散分析表

変動元	自由度	平方和	分散	F　値
処理 （群間変動）	群数－１	$SS_A = r\sum(\overline{Xi} - \bar{\bar{X}})^2$	SSA/3	
群内変動 （エラー）	全動物数－群数	$SS_{R(A)} = \sum SSi$	SSR(A)/16	
全変動	全動物数－１	$SS_{AR} = \sum(xij - \bar{\bar{X}})^2$		

変動元	自由度	平方和	分散	F　値
処理 （群間変動）	3	583.11	194.37	51.97
誤差変動 （群内変動）	16	59.84	3.740	
全変動	19	642.95		

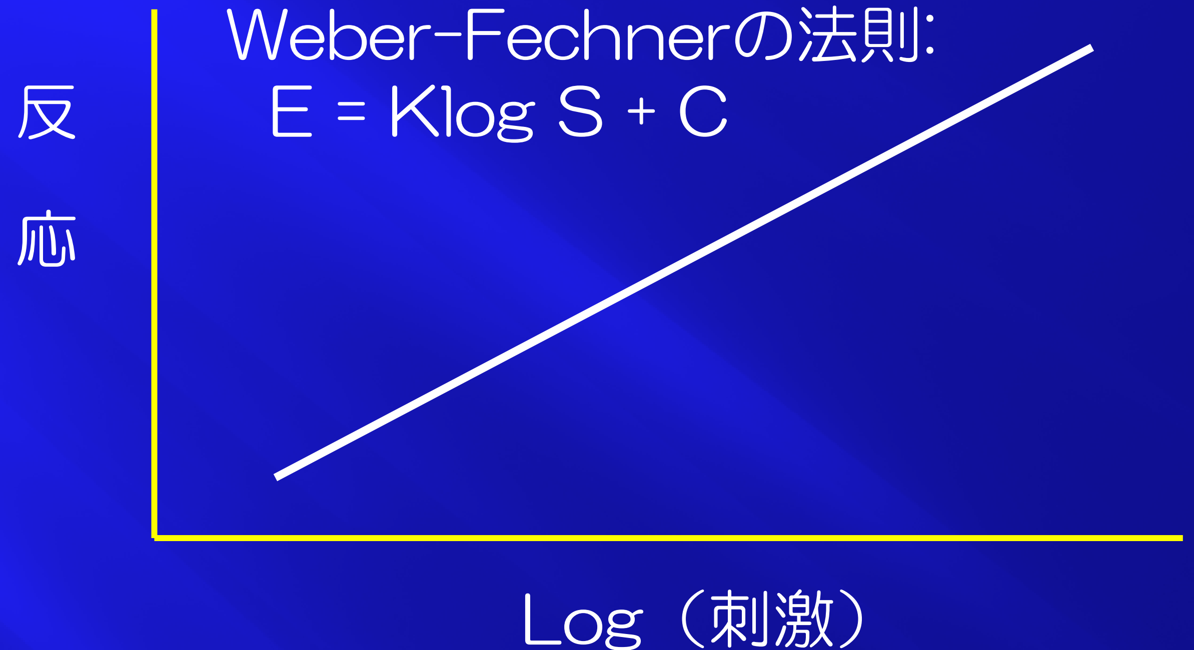
F (0.01)＝5.29　　第１自由度３，第２自由度１６

刺激　⇒　反応（Response）

Graded response　⇒　正規分布を基にした統計処理

Weber-Fechnerの法則:

$E = K\log S + C$



All or nothing response（全か無か反応、悉無率）

刺激　⇒　反応（有あor無し、１or ０）

反応の例：死、排卵、有効無効、行動、事象出現率を基に判定

出現率の検定（All or nothing response）（１）

	死亡	生存	合計
薬物投与無し（対照群）	6	4	10
薬物投与あり（処置群）	2	8	10
合計	8	12	20

このように例数の少ない場合には（ひと枡５以下がある）
Fisherの直接確率検定法が使われる
表の組み合わせ及びそれより極端な組み合わせを考え、
それらの起こる確率を合計する．その合計が0.05よりも
小さい時は薬物投与の効果があったと判定する

6	4
2	8

7	3
1	9

8	2
0	10

0.0268 + 0.0095 + 0.0004 = 0.0367

出現率の検定（All or nothing response）（2）
例数がひと枡５以上の場合にはカイ自乗検定が用いられる

	死亡（期待値）	生存（期待値）	合計
薬物投与無し	23（97）	159（84）	182
薬物投与あり	182（107）	18（92）	200
合計	205	177	382

括弧内は薬物投与の影響がないと考えた時の出現率
例えば 205×182/382＝97、205×200/382＝107
(観察値ー期待値)2／期待値　の合計がカイ自乗(計算値）でこの計算値を自由度１のカイ自乗表から得られた数値と比較する．計算値の方が大きければ薬物投与と生存、死亡の間には関係があると判定する
上の場合、 χ2cal = 235.318、　χ2(0.001)=10.83 なので危険率0.1％以下で薬物投与の生存に対する効果は有意であると判定される

m×n contigency table（m×n 分割表）

2×2分割表をさらに拡大して精度を改善することもできる

支持政党	20～29	30～39	40～49	50～59	60～69	70～
政党 A	13	32	36	45	42	33
政党 B	21	12	11	15	18	25
政党 C	15	10	12	6	10	11
政党 D	12	8	6	4	5	3
政党 E	5	8	12	7	4	3
無し	34	30	23	23	21	25

$\chi^2 cal = 52.982$

$X^2(0.001) = 52.62$

年齢層が異なると支持政党は異なる　　$p < 0.001$

さらに例えば２０代と３０代で差があるかの検定もできる

統計入門編　終わり