医薬安全性研究会

Bulletin of Japanese Society for Biopharmaceutical Statistics.

No.1～No.8 合併号

構　成

医薬安全性研究会 会報 合併号

総 合 目 次

安全研ニュース

No.4 1980年10月1日

医薬安全性研究会事務局発行©

101 東京都千代田区神田駿河台3-2 山崎ビル
サイエンティスト社内 〈03〉255-6847

第4回定例会

動物実験における比較のための統計量

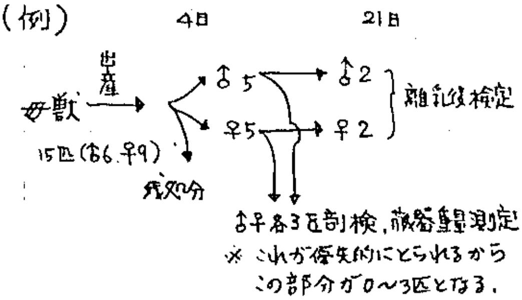
講師 吉村 功

1. 質 問

〔1〕 生殖試験における次世代（胎仔，出生仔）に関する測定値は各母体毎に算術平均された値を新たな実現値とみなし，これを再び算術平均したものを群の値として群間の差の検定を行なうか，あるいは群間で母体毎の平均値を順位和検定をすることが主流となっている．

例えば，次頁の資料は，出生仔の重量を記したものである．この場合離乳した仔のうち，離乳後検査に供する一定数の仔を除いた出生仔に対して，臓器重量が測定されるため一母体当りの検査仔数が少なく（0～3匹）かつ母体により例数も変動する．このように，サイズが非常に小さく，かつ母体毎に変動するにもかかわらず，常に母体単位で平均したものを検定するということは，統計的見地からして如何なものか？さらに最近では，異常出現率も母体毎に得られるP_iを群内で算術平均し，その際に標準偏差を求めるという指針も出されている．これは，どう考えたらよいのか？

(例)



[2] 臓器重量を比較するとき，体重を補助測定値とみなして共分散分析をするのは，どういう意味をもつのか？

以上の質問に対する吉村氏の答え

[1] の問題の定式化

データを次の記号で書く

X_{ijk}, $i=0,1$, $j=1,2,\cdots m_i$, $k=1,2\cdots r_{ij}$

ただし，$i=0$ は対照，$i=1$ は処理，m_i は母獣数，r_{ij} は各母獣における仔数．$r_{ij} \neq 0$ とする．

構造模型

$$X_{ijk} = \mu_i + U_{ij} + U_{ijk}$$

μ_i：定数：未知数

U_{ij}：母獣間の個体差　確率変数

U_{ijk}：各母獣における仔の個体差　確率変数，環境条件，測定上の変動も含む．

$$E\{U_{ij}\}=0,\quad E\{U_{ijk}\}=0$$

$$V\{U_{ij}\}=\sigma_1^2,\quad V\{U_{ijk}\}=\sigma_2^2$$

ただし，仮定が成立しないときは，別の論議になる．U_{ij}，U_{ijk} はすべて独立で，正規分布に従うものとする．

検定したい仮説

$$H_0:\ \mu_1-\mu_0=0$$

〔この問題に対する一般論は，例えば 広津千尋著「分散分析」教育出版 2800円，第8，9章に述べられている〕

$$V\{X_{ijk}\}=\sigma_1^2+\sigma_2^2$$

$$Cov\{X_{ijk},X_{i'j'k'}\}=0\quad((ij)\neq(i'j'))$$

$$Cov\{X_{ijk},X_{ijk'}\}=\sigma_1^2\quad(k\neq k')$$

$$V\{\overline{X}_{ij}\}=\sigma_1^2+\sigma_2^2/r_{ij}$$

(i) ゆえに，$\sigma_1^2 \ll \sigma_2^2$ ならば，$n_i=\sum_{j=1}^{m_i} r_{ij}$ とおき，

$$\frac{|\overline{X}_1-\overline{X}_0|}{\sqrt{\dfrac{\sum_{i=0}^{1}\sum_{j=1}^{m_i}\sum_{k=1}^{r_{ij}}(X_{ijk}-\overline{X}_i)^2}{n_1+n_0-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_0}}}$$

$$> t(n_1+n_0-2,\alpha)$$

ただし，$\overline{X}_i=\sum_{j=1}^{m_i}\sum_{k=1}^{r_{ij}} X_{ijk}/n_i$

という棄却域を用いる．

【1】 偵同の表

♂ case (g) …処理

母	仔	体重	肝重
1	1	32.7	.1288
	2	33.5	1.262
2	1	46.9	1.889
	2	46.0	2.057
3	1	39.5	1.750
	2	36.5	1.679
	3	40.5	1.846
4	1	36.6	1.676
	2	37.8	1.657
	3	39.0	1.867
5	1	40.8	1.700
	2	41.2	1.421
6	1	39.0	1.677
	2	36.6	1.723
7	1	40.5	1.789
	2	41.7	1.802
8	1	44.3	2.416
	2	53.4	1.913
9	1	37.8	1.601
10	1	46.6	1.864
	2	39.0	1.368
	3	44.1	1.648
11	1	40.8	1.574
	2	39.4	1.515
	3	41.3	1.438
12	1	45.7	2.239
	2	47.0	2.428
	3	46.2	2.200
	4	41.4	1.970
13	1	38.0	1.805
13	2	41.4	2.146
	3	40.0	2.091
14	1	48.1	2.021
	2	46.4	2.066
	3	47.9	2.006
15	1	39.7	1.839
	2	42.5	1.920
	3	36.2	1.444
16	1	36.8	1.458
	2	38.7	1.523
	3	36.4	1.584
17	1	34.2	1.454
	2	30.3	1.276
	3	31.4	1.173
18	1	42.7	1.973
	2	41.1	1.925
	3	40.3	1.778
19	1	41.8	1.936
	2	41.0	1.906
20	1	34.5	1.284
	2	35.6	1.486
	3	32.0	1.194

♂ control …対照 (g)

母	仔	体重	肝重
1	1	43.1	1.942
	2	46.3	1.884
	3	49.5	2.219
	4	51.2	2.193
2	1	48.7	2.167
	2	46.5	2.025
	3	46.0	2.027
3	1	44.5	1.836
	2	38.7	1.558
	3	44.6	1.832
4	1	42.5	2.195
	2	47.0	1.827
	3	44.0	1.961
5	1	46.2	1.769
	2	43.8	1.718
	3	42.2	1.779
6	1	41.1	1.770
	2	40.0	1.680
	3	40.5	1.786
7	1	41.5	1.633
	2	40.6	1.785
	3	29.0	1.157
8	1	41.9	1.759
	2	49.5	2.243
9	1	43.8	1.788
	2	42.9	1.824
	3	42.5	1.818
10	1	36.0	1.388
	2	35.7	1.342
	3	30.3	1.110
11	1	44.3	1.714
	2	43.2	1.760
	3	47.5	1.983
12	1	40.6	1.791
	2	43.0	1.520
	3	39.0	1.484
13	1	45.7	1.803
	2	42.0	1.592
	3	40.9	1.635
14	1	46.6	1.881
	2	42.4	1.738
	3	45.2	1.933
15	1	52.8	2.077
	2	50.4	2.243
	3	53.2	2.414
	4	49.1	2.146
16	1	36.5	1.381
	2	36.1	1.450
	3	31.9	1.310
	4	35.7	1.462
19	1	47.1	2.177
	2	44.7	2.098
	3	48.6	1.980
20	1	36.5	1.208
	2	38.2	1.409
	3	33.0	1.202

(ii) また，$\sigma_1^2 \gg \sigma_2^2$ ならば，

$$\frac{|\bar{\bar{X}}_{1..}-\bar{\bar{X}}_{0..}|}{\sqrt{\frac{\sum_{i=0}^{1}\sum_{j=1}^{m_i}(\bar{X}_{ij.}-\bar{\bar{X}}_{i..})^2}{m_1+m_0-2}}\sqrt{\frac{1}{m_1}+\frac{1}{m_0}}}$$

$$> t(m_1+m_0-2,\alpha)$$

ただし，$\bar{\bar{x}}_{i..}=\frac{1}{m_i}\sum_{j=1}^{m_i}\bar{x}_{ij}$ という棄却域を用いる．

(iii) $\sigma_1^2\sim\sigma_2^2$ の場合は問題が微妙である．この場合には，理論的に最良な方式は存在しない（Behrens-Fisher問題より複雑）．そこで，むしろ直観的にもっともらしい方式を考えて，それが，あまり不合理でないことを吟味し合格したら，それを用いることにすればよい．

$\bar{X}_{ij.}=\mu_i+U_{ij}+\bar{U}_{ij.}$,

$E\{\bar{X}_{ij.}\}=\mu_i$

$V\{\bar{X}_{ij.}\}=\sigma_1^2+\sigma_2^2/r_{ij}=\sigma^2/a_{ij}$

ただし，$\sigma^2=\sigma_1^2+\sigma_2^2$,

$$a_{ij}=r_{ij}\left(1-\frac{(r_{ij}-1)\sigma_1^2}{r_{ij}\sigma_1^2+\sigma_2^2}\right)$$

したがって，$1\leqq a_{ij} < r_{ij}$

$E\{\bar{\bar{X}}_{i..}\}=\mu_i$,

$$V\{\bar{\bar{X}}_{i..}\}=\frac{1}{m_i^2}\sum_{j=1}^{m_i}\frac{\sigma^2}{a_{ij}}$$

$$=\left(\frac{1}{m_i^2}\sum_{j=1}^{m_i}\frac{1}{a_{ij}}\right)\sigma^2$$

$$E\{(\bar{\bar{X}}_{1..}-\bar{\bar{X}}_{0..})^2\}$$

$$=(\mu_1-\mu_0)^2+\left(\frac{1}{m_1^2}\sum_{j=1}^{m_1}\frac{1}{a_{ij}}+\frac{1}{m_0^2}\sum_{j=1}^{m_0}\frac{1}{a_{0j}}\right)\sigma^2$$

$$E\left\{\sum_{j=1}^{m_i}(\bar{X}_{ij.}-\bar{\bar{X}}_{i..})^2\right\}$$

$$=\frac{m_i-1}{m_i}\sum_{j=1}^{m_i}\frac{\sigma^2}{a_{ij}}=(m_i-1)\sigma_1^2+\frac{m_i-1}{m_i}\sum_{j=1}^{m_i}\frac{\sigma_2^2}{r_{ij}}$$

ゆえに，一つの近似的な方式は，

$$\frac{|\bar{\bar{X}}_{1..}-\bar{\bar{X}}_{0..}|}{\sqrt{\frac{\sum_{j=1}^{m_1}(\bar{X}_{1j.}-\bar{\bar{X}}_{1..})^2}{m_1(m_1-1)}+\frac{\sum_{j=1}^{m_0}(\bar{X}_{0j.}-\bar{\bar{X}}_{0..})^2}{m_0(m_0-1)}}}$$

$$> t(\phi,\alpha)$$

という棄却域を用いることである．ここで，ϕは次式から求められる．

$$\phi=\frac{\left(\frac{1}{m_1^2}\sum_{j=1}^{m_1}\frac{1}{a_{1j}}+\frac{1}{m_0^2}\sum_{j=1}^{m_0}\frac{1}{a_{0j}}\right)^2}{\frac{1}{m_1^3(m_1-1)}\sum_{j=1}^{m_1}\frac{1}{a_{1j}^2}+\frac{1}{m_0^3(m_0-1)}\sum_{j=1}^{m_0}\frac{1}{a_{0j}^2}}$$

しかし，実際はこのような式で計算するのは面倒であるし，a_{ij} の値も正確に分からないので，推定値を使うため，それ程改善されない．そこで，(ii)の方式を使う方がスマートである．

■ 以上の結果を実際に検討してみると,

$$\hat{\sigma}_2^2 = \sum_i \sum_j (x_{ijk} - \bar{x}_{ij\cdot})^2 / (n_1 + n_0 - m_1 - m_0) = \frac{0.773984425 + 0.876235775}{52 + 56 - 20 - 18}$$

$$= 0.023574571$$

$$\hat{\sigma}_1^2 = \frac{1}{m_1 + m_0 - 2} \left\{ \sum_i \sum_j (\bar{x}_{ij\cdot} - \bar{\bar{x}}_{i\cdot\cdot})^2 - \sum_i \left(\frac{m_i - 1}{m_i} \sum_j \frac{1}{r_{ij}} \right) \hat{\sigma}_2^2 \right\}$$

$$= \frac{2.738475541}{20 + 18 - 2} - \left[\left(1 - \frac{1}{20}\right)\left(\frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{3}\right) + \left(1 - \frac{1}{18}\right)\left(\frac{1}{4} + \frac{1}{3} + \cdots + \frac{1}{3}\right) \right]$$

$$\times \frac{0.023574571}{36} = 0.076068765 - (8.945833 + 5.587963) \times 0.000654849$$

$$= 0.066570209$$

方式(ii)を用いると,

$$t_0 = \frac{|1.73212 - 1.77272|}{\sqrt{0.076069 \times \left(\frac{1}{20} + \frac{1}{18}\right)}} = \frac{0.0406}{0.0896} = 0.453 \quad , \quad \phi = 36$$

方式(iii)では,

$$t_0 = \frac{0.0406}{\sqrt{\frac{0.074937}{20} + \frac{0.077334}{18}}} = 0.453 \qquad \phi = 35.64$$

※ただし, σ_1^2, σ_2^2 にはその推定値を代入して計算すること

したがって, (ii), (iii) 何れもほとんど結論は同じとなる.

【2】の問題について

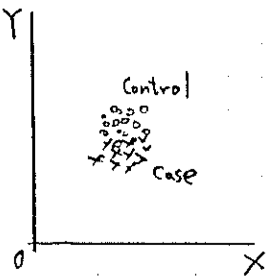
ある機器の測定値 Y_{ij}, $i = 0, 1$, $j = 1, 2, \cdots m_i$ があり, 同時に補助測定値をどう利用すればよいかは, XとYとの関係による. いいかえれば, XとYのどんな関数に対して, 線形モデルが想定できるかによる.

(i) XとYが独立の場合

$$Y_{ij} = \mu_i + U_{ij}$$

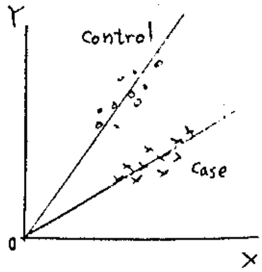
この場合は普通の t 検定を行う.

棄却域は,



$$\frac{|\bar{Y}_{1\cdot}-\bar{Y}_{0\cdot}|}{\sqrt{\frac{\Sigma(Y_{1j}-\bar{Y}_{1\cdot})^2+\Sigma(Y_{0j}-\bar{Y}_{0\cdot})^2}{m_1+m_0-2}\times\left(\frac{1}{m_1}+\frac{1}{m_0}\right)}} > t\,(m_1+m_0-2,\ \alpha)$$

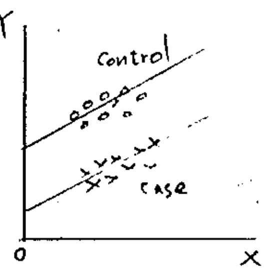
(ii) XとYが比例的で，勾配が処理によって影響を受ける場合.



$$Z_{ij} \overset{\text{def}}{=} \frac{Y_{ij}}{X_{ij}} = \mu_i + U_{ij}$$

この場合は，Zについて普通の t 検定でよいであろう. 細かいことを言えば，U_{ij} の分布にもよるが，現実には十分である.

(iii) XとYの相関が非常に大きいが，処理の結果が平行移動的な場合.



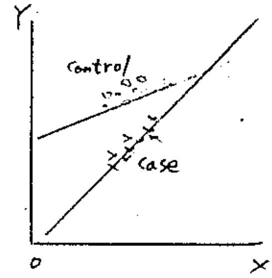
$$Y_{ij} = \mu_i + \beta X_{ij} + U_{ij}$$

この勾配はiに無関係，即ち処理と対照共通

この場合には，2水準1元配置の共分散分析がよい（スネデガー著「統計的方法」13章参照 岩波書店）.

(iv) もっと一般的に，

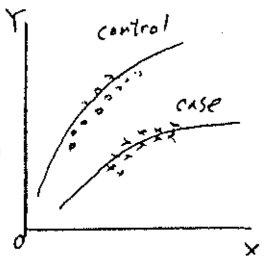
$$Y_{ij} = \alpha_i + \beta_i X_{ij} + U_{ij}$$



で勾配も，切片も効果が現われている場合には，線形仮説，$\alpha_1=\alpha_0$，$\beta_1=\beta_0$ の検定になる.（→吉村功著「数理統計学」培風館 P195～199参照）.

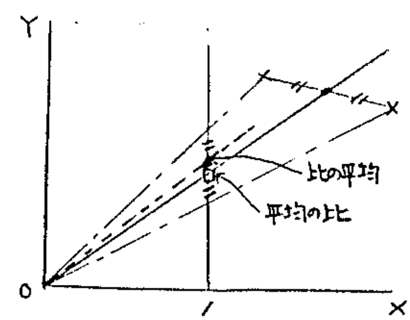
(v) さらに，別の一般化として，

$$Z_{ij} \overset{\text{def}}{=} f(Y_{ij}, X_{ij}, \theta) = \mu_i + U_{ij}$$



というモデルも考えられる. この時は，θを適当な方法で推定し，差の t 検定を行うことになる.

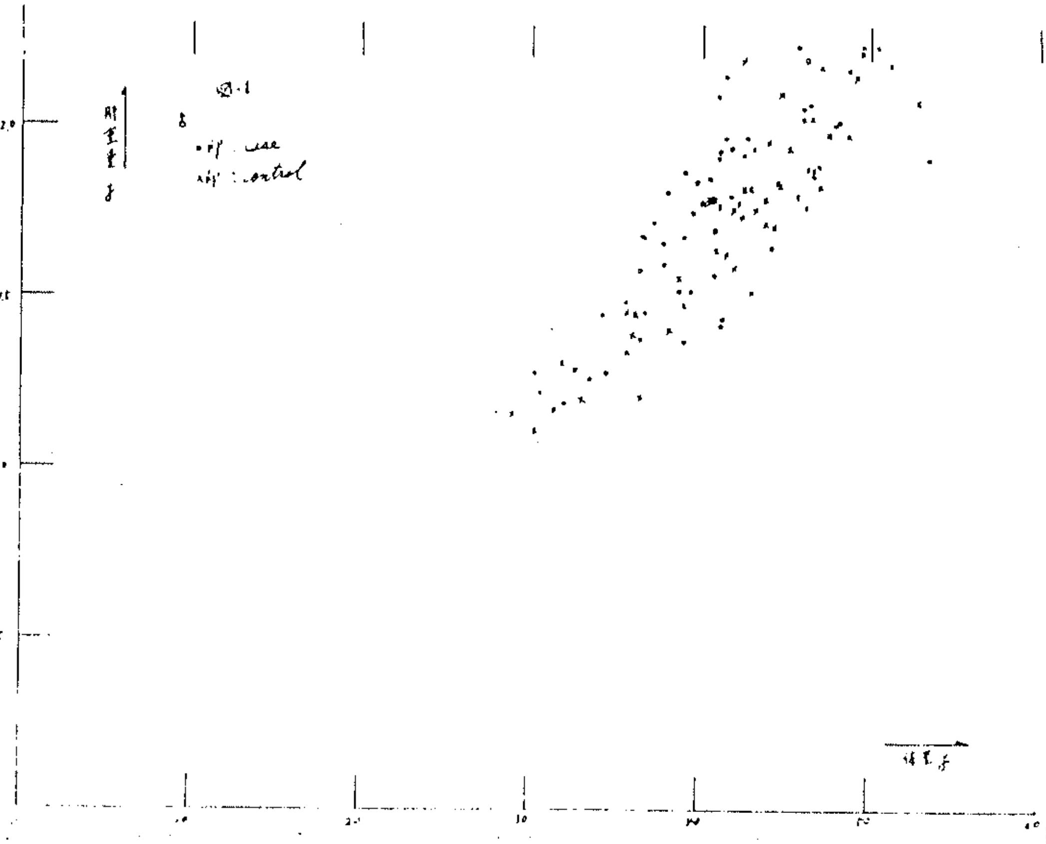
■ 比の平均と平均の比の違い



×2点を観測点とすると，□印の高さが平均の比になり，△印の高さが比の平均になる.

■【問1】の数値例とモデルについて

【1】の数値例を図示すると，下図のようになる．



これからは，モデル(i)が適合しないのは明らかである．(iv)でもなさそうである．(ii)，(iii)，(v)のどれであろうか？

定規を当ててみると，(v)において $z_{ij}=Y_{ij}/X_{ij}^{\theta}$ としたものになりそうである．これは $\theta=1$ とおけば(ii)になるから，これでいけば問題なさそうであるが，もし，本当に $\theta=1$ の場合には，θ をデータから推定した分だけ検定の精度（検出力）が落ち，理論的にあいまいな所が生じ，手順も面倒になる．この程度だったら(ii)で十分なようである．数値計算を行うと次のようになる（$\hat{\theta}\fallingdotseq 1.2$）．

Case：　$r=0.77$，勾配 $=0.049$　切片 $=-0.236$

Control： $r=0.92$， 勾配 $=0.052$， 切片 $=-0.452$

$z_{ijk}=Y_{ijk}/X_{ijk}$ について、(ii)方式の検定を行うと，

$$t_0=\frac{|0.042987-0.041331|}{\sqrt{\frac{0.0000252035+0.000111364}{20+18-2}\times\left(\frac{1}{20}+\frac{1}{18}\right)}}=1.60$$

$\phi=36$　　not significant.

第5回定例会開催要領

日時：10月25日(土) PM2:00〜4:30

講師：高橋行雄(日本ロシュ)

テーマ：一般毒性試験における相対
　　　臓器重量の意味を考える。

会場：生協会館　7F　第3会議室
　　　渋谷区千駄ケ谷4−1
　　　TEL (03) 404−3231

本来であれば、第5回定例会は9月に開催の予定でしたが、都合により延期させて頂きました。今回の講師は本会会員の高橋氏です。「応用薬理」に掲載された1年間分の報文を分析し，その現状と問題点を報告して頂き，今後の方向を考えてゆきたいと思っております。

なお、当日は会の運営についても議論したいと思っておりますので積極的な参加をお待ちしております。出欠は必ず (03−255−6847) ご連絡下さい。