

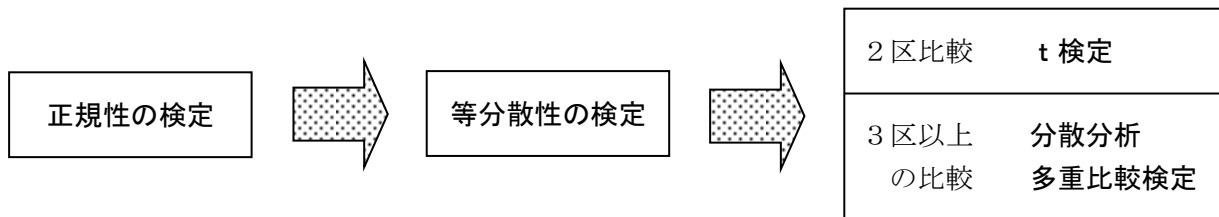
VI. 統計処理

受精率やふ化率は、卵重や体重などのように正規分布しないことが多い。このような場合の検定法としては、①正規分布するように数値変換を行ってパラメトリック検定で行う方法、②ノンパラメトリック検定で行う方法、③分割表などによって比率の検定を行う方法がある。①と②の場合は、1区1回のふ化率を入卵個数に関係なく1個のデータとして扱うため、反復試験の必要がある。③の場合は、受精卵と無精子の個数によって検定するため反復試験の必要は無い。①の場合の数値変換には、アークサイン変換が良く用いられる。その他に、受精卵には1、無精子には0を個数だけ入力する1-0分散分析法もあるが、データ数が多い場合はわずかな差でも有意になってしまうため、鶏の場合では差がないことを強調する検定以外は通常行わない。

1. 数値変換を行ってパラメトリック検定で行う方法

(1) 検定の順序

差の検定を行う前に、正規性の検定を行い正規分布に従わない場合は数値変換を行い正規分布に従う形に変換する。次に、等分散性の検定を行い、分散が等しいか否かで差の検定方法が変わる。差の検定は、試験区が2つであればt検定、3つ以上であれば分散分析と多重比較検定で行う。



(2) 正規性と等分散性の検定

t検定や分散分析の前に行わなければならないのが正規性の検定で、コルモゴロフ・スミルノフの検定、シャピロ・ウィルクスの検定、リリフォースの検定などの検定法がある。P>0.001であれば正規分布に従っていると見る。これらの検定をしなくても正規確率プロットを作成して、一直線に並んでいれば正規分布していると見る方法もある。

t検定や多重比較検定は、分散が等しいかどうかで差の検定方法が変わる。等分散性の検定は、試験区が2つの場合は、F検定で行い、3つ以上であれば、ハートレイの検定、コ克蘭の検定、バートレットの検定、ルビーン（レーベン）の検定などがあり、P>0.05であれば分散は等しいと見る。

分散分析でも最小二乗分散分析で行う場合は、正規性の前提条件が外れるため、正規性の検定を飛ばして等分散性の検定から行う。

(3) アークサイン変換 ($\sin^{-1}\sqrt{X}$)

二項分布のような変数を正規分布に近づける数値変換法。数値が 80%以上や 20%以下に多くのデータがある場合に有効。鶏の受精率では、対数変換よりもアークサイン変換が良く用いられている。

計算順序は、

- ① パーセントの数値を小数点の数値に戻す。 85% → 0.85
- ② ①の値の平方根を計算する。 0.85 → 0.9219544...
- ③ ②の値のアークサインを計算する。 0.9219544... → 1.1730969...

EXCEL 関数では、平方根は SQRT、アークサインは ASIN。平均値を出す場合は、アークサイン変換後の平均値を逆に計算してパーセントに戻す。(サインを求めてから 2 乗し 100 倍する)

数値変換各種

データ変換には、アークサイン変換の他にも次のようなものがあります。

$y = \log x$	...	対数変換
$y = \sqrt{x}$	...	ルート変換
$y = x^a$	...	べき乗変換
$y = e^{ax}$	...	指数変換
$y = 1/2 \cdot \log(x / (1 - x))$	...	ロジスティック変換
$y = 1/2 \cdot \log((1 + x) / (1 - x))$	...	フィッシャーの z 変換
$x^{(\lambda)} = \begin{cases} (x^\lambda - 1) / \lambda & \lambda \neq 0 \\ \log x & \lambda = 0 \end{cases}$	...	ボックス・コックス変換

(4) t 検定

t 検定には 3 種類ある。分散が等しい場合は、等分散を仮定した t 検定を行い、分散が等しくない場合は、分散が等しくないと仮定した t 検定（ウェルチの検定）を行う。同じ個体を用いて 1 回目と 2 回目とか、使用前と使用後の様な検定をする場合は、対応のある場合（従属 2 標本、一対の標本）の t 検定を行う。試験区が 3 つ以上で t 検定を繰り返して検定した場合は、有意確率が上昇してしまうので、行った場合は P 値を修正する必要がある。

(5) 分散分析

試験区が 3 つ以上の場合に検定する方法で、因子の効果を判定するもの。ある実験で、貯卵温度を 5℃、10℃、15℃で調査したとする。この場合、貯卵温度が因子であり、5℃、10℃、15℃が水準となる。因子が 1 つであれば一元配置分散分析、因子が 2 つであれば二元配置分散分析、3 つであれば三元配置分散分析となる。温度と湿度条件を変えたとなると、因子が 2 つなので二元配置となる。同じ個体で 1 分後、2 分後、3 分後と言うように 3 個以上ある場合は、それを因子と捉えて二元配置分散分析で検定ができる。

単に試験区ごとに差を見たい場合は、試験区を因子と捉えて一元配置分散分析を行う。t検定と一元配置分散分析では、反復データが必須であるが、二元配置以上の場合は実験の繰り返しが無くても分析できる。分散分析では、因子ごとの効果と交互作用の効果が判るが、因子の数と繰り返しの有無によって判定できる項目が変わる。交互作用が有意と判断された場合は、因子内における水準の差の有意性は無意味となり、どの区が優れているかは一元配置の形で多重比較検定を行う。

分散分析によって判定できるもの

因子の数	実験の繰り返し	有意性を判定できる項目
1	なし	—
	あり	S_A
2	なし	S_A S_B
	あり	S_A S_B $S_{A \times B}$
3	なし	S_A S_B S_C $S_{A \times B}$ $S_{A \times C}$ $S_{B \times C}$
	あり	S_A S_B S_C $S_{A \times B}$ $S_{A \times C}$ $S_{B \times C}$ $S_{A \times B \times C}$

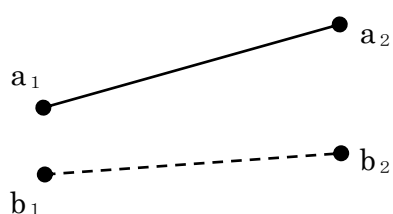
S_A は因子 A の効果、 $S_{A \times B}$ は因子 A と因子 B の交互作用効果

因子

分散分析でいきなり「因子」という語句が出てきますが要因と同じ意味で、英語ではどちらも **factor** です。ただし、因子は統計用語なので、要因と言わずに因子と言いましょう。

交互作用

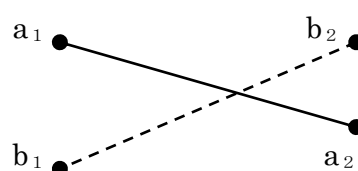
交互作用とは、因子と因子が互いに影響していることを指しており、ある条件の結果と別の条件での結果が異なる場合には交互作用があると言います。交互作用には、A の条件の時に a は b よりも成績が良いが、B の条件になるとさらに差が開くといった相乗効果と、A の条件の時には a は b よりも成績が良いが、B の条件になると成績が逆転する場合の相殺効果と呼ばれるものがあります。



条件 A 条件 B

相乗効果

条件が変わると差が広がる



条件 A 条件 B

相殺効果

条件が変わると逆転する

(6) 多重比較検定

分散分析で判るのは、因子の効果と交互作用であって、どの区とどの区に差があるかは判らない。一元配置分散分析の形で、これを調べるのが多重比較検定と呼ばれるもので、LSD 検定 (最小有意差検定)、対比 (ついで) による検定、ダンカンの多重範囲検定、テューキーの HSD 検定、シェフェの検定等がある。検定法により検出力の強さ (有意と出やすいか) に差があり、LSD 検定 = 対比による検定 > ダンカンの検定 > テューキーの検定 > シェフェの検定の順となる。LSD 検定やダンカンの検定は、多重比較として適切でないとされているが、学会誌等の報告ではどれも頻繁に用いられている。t 検定など 2 標本の検定を繰り返して検定した後、P 値を修正する方法 (ボンフェローニ検定など) もある。多重比較検定は、一元配置分散分析のあとで行うのが基本だが、必ずしも分散分析を先に行う必要はなく、多重比較検定のみを行っても問題はない。どの検定を用いるかは、結論の持って行き方によって変えても良いが、テューキーの HSD 検定が無難。二元配置以上の場合、因子ごとの水準を 2 つにしておくとも多重比較を行う必要が無くなる。もし水準が 3 つ以上の場合、因子ごとに一元配置の形に集計して多重比較を行う。

ボンフェローニ検定

ボンフェローニの不等式を利用した多重比較で、t 検定など 2 標本の検定を繰り返した後で有意確率を修正するもの。実際には、算出された有意確率に繰り返し行った検定数 ${}_nC_2$ をかけるだけのこと。試験区が 3 つであれば検定数は、1 と 2 区、2 と 3 区、1 と 3 区の 3 回比較を行うので (${}_3C_2 = 3$)、算出された有意確率 (P 値) を 3 倍します。したがって、P 値が 0.016 だと 0.048 と変換され 5% 水準で有意となりますが、0.017 では 0.051 に変換されて有意差無しとなります。試験区が 4 つの場合は ${}_4C_2 = 6$ なので、P 値を 6 倍します。試験区が 5 つの場合では、 ${}_5C_2 = 10$ となり、P 値を 10 倍することになります。この方法は、2 標本の比較が出来れば用いることが出来るので、ノンパラメトリック検定での多重比較検定をしたい場合などでも行えます。ただし、試験区が多くなるほど差が出にくくなるので、試験区は 5 区程度までに抑えることが望ましいとされています。

ボンフェローニの不等式

ある A といった事象に対する確立を $P(A)$ としたとき、次の不等式が成り立つと言うもの。

$$P(\cup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$$

5 つのグループを t 検定で調べようとする、 ${}_5C_2 = 10$ で 10 組の t 検定を行うことになるので、

$$P(A_1 \cup A_2 \cup A_3 \cdots \cup A_{10}) \leq P(A_1) + P(A_2) + P(A_3) \cdots P(A_{10})$$

となり、有意水準を 5% とすると、

$$P(A_1 \cup A_2 \cup A_3 \cdots \cup A_{10}) \leq 10 \times 0.05 = 0.5$$

となるので有意水準は最大で 10 倍に増える可能性があるというもの。

多重比較検定の符号の記入方法

多重比較検定を行ったら、どの区とどの区が有意で、どの区とどの区は差がないかを表にまとめます。通常は、データの右肩に小文字のアルファベットを記入し、異符号間で有意と言った書き方をします。この符号の付け方は、次の通りです。

試験区の数4区であって、次のような結果であったとします。

	I	II	III	IV
I	—	P<0.05	NS	NS
II	P<0.05	—	P<0.05	NS
III	NS	P<0.05	—	NS
IV	NS	NS	NS	—

NS：有意差無し

- ① まず、試験区を数値の大きい順に並べます。ここでは、I区>III区>IV区>II区 であつたとします。
- ② 次に、下の図のように差の無い区同士を線で結びます。IはIIIとIVと差が無いのでこの3つを線で結び、IIはIVと差がないのでIVとIIを線で結びます。III区はI区と全く同じになるので線を引きません。IV区はどの区とも差がないので、この場合も線を引きません。この場合2本の線が引けたので、上から順に a、b という様に小文字のアルファベットを付けます。
- ③ ②で作成した表を試験区ごとに縦に見て、どの線が引かれているかを集計します。IとIIIはaの線1本なのでa、II区はbとなり、IVの場合はaの線とbの線の両方あるのでabとなります。

区	I	III	IV	II
I	————— ← a の線			
II			————— ← b の線	
	=====			
	a	a	ab	b

- ④ 最後に、試験区の平均値にの右肩に③のアルファベットを記入して表を作成します。

表	×××
I	○○○ a
II	△△△ b
III	□□□ a
IV	▽▽▽ ab

a—b 異符号間で有意 (P<0.05)

統計ソフトでの多重比較検定

統計ソフトも色々ありますが、統計ソフトによって使用できる検定に制限がかかります。例えば、STATISTICA というソフトでは LSD 検定、ダンカンの検定、テューキーの検定、シェフェの検定などがありますが、これらは全て等分散が成り立つ場合の検定方法しか出来ません。SPSS では、さらに多くの多重比較検定が用意されており、等分散が成り立たない場合の検定も 4 種類行えます。LSMLMW では最小二乗分散分析のため、正規性の検定は不要ですが、対比による検定しか出来ません。SAS は、最小二乗分散分析が行えて、全ての検定が可能なので、論文では SAS が最も頻繁に使われおり、さすがの SAS などと呼ばれています。

(7) 統計モデル（数学モデル、線型モデル）

一般線型モデルのことを、general linear model の略で GLM と呼ぶ。

一元配置の分散分析では、

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

Y_{ij} は測定値、 μ は全平均、 α_i は水準 A_i の主効果、 ε_{ij} は誤差。

二元配置の分散分析では、

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

β_j は水準 B_j の主効果、 $(\alpha\beta)_{ij}$ は交互作用。

これらの統計モデルは、103 ページの有意性を判定できる項目に全平均 (μ) と誤差 (ε) を加えたものと一致する。三元配置以上の場合で、特に交互作用を考えなくて良い場合などは交互作用を併合して分析する場合もある。ただし、交互作用を併合すると有意と出やすくなるので注意が必要。論文発表の際には、基本モデルの場合は記載の必要がないが、誤差を併合して分析した場合は用いた統計モデルを記載する。

試験区を減らす方法

多元配置分散分析で、因子内の水準が全て 2 つだとすると、試験区は二元配置では 4 つ、三元配置では 8 つ、四元配置では 16 必要です。しかし、ラテン方格法と直行法の利用により、三元配置では 4 つ、四元配置では 8 つの試験区に減少させることが出来ます。ただし、この場合の交互作用は算出されません。

四元配置のデータを 8 つに減少させた場合の試験区（黄色部分だけ行う）

		D1		D2	
		B1	B2	B1	B2
C1	A1	X1111	X1211	X1112	X1212
	A2	X2111	X2211	X2112	X2212
C2	A1	X1121	X1221	X1122	X1222
	A2	X2121	X2221	X2122	X2222

ABCD は因子を示し数字は水準。X の横の数字は、ABCD のそれぞれの水準を示す。

2. ノンパラメトリック検定

(1) ノンパラメトリック検定の種類

計測値を用いて検定するのがパラメトリック検定であったのに対し、計測値を順位に置き換えて検定するのがノンパラメトリック検定。受精率の検定にはあまり見ないが、正規分布に従わない計測値の検定に用いられる方法として知られている。ノンパラメトリック検定にも多くの検定があるが、通常のパラメトリック検定と対応させて表にまとめた。ただし、反復数が通常5個以上必要なので反復数が少ない場合は数値変換してパラメトリック検定を行う。

パラメトリック検定と対応するノンパラメトリック検定

目 的	パラメトリック検定	ノンパラメトリック検定
2つの標本の比較	t 検定 ウェルチの検定	マン・ホイットニーの U 検定 ウィルコクソンの順位和検定
対応のある2つの標本の比較	従属2標本の t 検定	ウィルコクソンの符号付順位検定
3つ以上の標本の比較 (因子1つ或いは因子の効果は無視)	一元配置分散分析	クラスカル・ウォリスの検定
3つ以上の標本の比較(因子2つ) (対応のある3つ以上の標本の比較)	二元配置分散分析	フリードマンの検定
共変量がある場合	共分散分析	
多重比較検定	最小有意差検定 対比による検定 ダンカンの多重範囲検定 テューキーの HSD 検定 シェフェの検定	
2変数の相関	相関係数	ケンドールの順位相関係数 τ (タウ) スピアマンの順位相関係数 R
ある変数から別の変数を推定	回帰分析 (1個の変数から推定) 重回帰分析 (複数の変数から推定)	

(2) ノンパラメトリック検定での多重比較検定

STATISTICA や SPSS では、ノンパラメトリック検定での多重比較検定が無いので、2区間の比較を行うマン・ホイットニーの U 検定（ウィルコクソンの順位和検定と結果は同じ）を繰り返した後で、ボンフェローニの方法によって修正を行う。

3. 分割表などによる比率の検定（カイ 2 乗検定）

（1）2 × 2 の分割表による独立性の検定

通常はカイ 2 乗検定を指すが、例数が 5 以下の例数を含む場合はフィッシャーの直接確率計算法を行う。

例：	試験区	入卵個数	無精卵	受精卵	受精率
	1	100	10	90	90.0%
	2	125	25	100	80.0

分散分析や t 検定などでは受精率の数値を使って検定を行うが、分割表による独立性の検定は無精卵個数と受精卵個数を用いる。分割表は、

区	無精卵	受精卵	合計
1	a	b	a+b
2	c	d	c+d
合計	a+c	b+d	n

$$\chi^2 = n(|ad-bc| - n/2)^2 / (a+b)(c+d)(a+c)(b+d)$$

$$\text{自由度(df)} = (\text{行} - 1)(\text{列} - 1)$$

であり、例では a、b、c、d にそれぞれ、10、90、25、100 を当てはめて計算する。

（2）2 × m、m × n の分割表による検定

試験区が多くなった場合などもカイ 2 乗検定で行えるが、一元配置の分散分析と同様でどの区とどの区が有意かは解らない。統計ソフトでも SPSS は出来るが STATISTICA では出来ない。ただし、手計算でも行える。どの区とどの区が有意かを知るには、2 × 2 の形を繰り返すしかないので、この場合は算出された有意確率をボンフェローニの方法で修正する。

（3）適合度の検定

試験区の数に関係なく、期待度数と観測度数をカイ 2 乗検定によって検定する。この場合の期待度数は、全区の平均から算出するが、観測度数の合計（無精卵と受精卵の合計など）と期待度数の合計を等しくして検定する。

通常と比較では、数値の右肩に数値の高いものから順に a b c を書き、異符号間で有意といった書き方がされるが、この場合は、全平均から算出された期待度数との検定となるので、試験区ごとに P<0.05 などを書く方法がとられる。

試験区	受精卵	無精卵	P 値
1	○○	◇◇	NS
2	××	▽▽	NS
3	□□	☆☆	P<0.05
4	△△	◎◎	NS
合計	●●	◆◆	