TOP | 実験医学 | 単行本 | 購読案内 | 広告掲載案内 | 人材・シンポ | Podcast | Onlineコンテンツ | メールマガジン | 羊土社とは

実験医学
統計の落とし穴と蜘蛛の糸
三中信宏（農業環境技術研究所）

連載TOPへ

【書籍化決定！】本コンテンツの公開は2015年5月7日までとなります。

[SHARE] Tweet シェア

第11回 正規分布を踏まえたパラメトリック統計学の降臨

実験医学2015年3月号

はじめに

地道に計算する衆生に正規分布の王様は手を差し伸べる—前回に 続いて，今回も実験計画法の話です．典型的な実験計画法では，実際にデータが得られる前に，予想されるデータのばらつきに関する線形統計モデル（多くの場合，正規分布が前提）を仮定します．そして，首尾よく数値データが得られたならば，私たちはその次の段階に進むことができます．

データの数値のもつばらつきは私たちが手にする唯一の情報源です．したがって，仮定した統計モデルを横目にみながら，データのばらつきを 整理して数値化する必要があります．データのばらつき全体のうち，実験処理や偶然誤差といった変動要因が実際どれくらいのばらつきをデータにもたらすかは 統計量として表すことができます．今回はまずはじめにこの地道な計算について説明します．

一方，仮定した統計モデルからは，データが抽出された母集団に関して，ばらつきをあらわす統計量がどのような確率分布をするかが数学的に 導出されています．私たちがデータから地道に計算しているとき，雲の上ではパラメトリック統計学の厳密な数学理論がデータのふるまいに関する数式を操作し ているのです．そして，データから最終的に実験処理や偶然誤差のばらつきの集計が完了したとき，雲の上から正規分布の王様がおもむろに降臨し，地上の私た ちが計算してきた結果からはたして実験処理の効果があったかどうかの御神託を手渡します．データからの計算と統計理論が合体するその瞬間を私たちは体験で きます．

§ 分散は分割せよ —知りたいばらつきをあぶり出す

前回は実験計画法の総論を説明したうえで，実例としてイネの収量試験に関する具体的な数値データをお見せし，背後に仮定される統計モデルについて解説しました．この実験では，殺虫剤7水準の実験処理を4反復の完全無作為化法で実施します．前回（第10回）の図3の一部を表1として再掲します．この表1に示されているように，処理平均をみれば，水準によって収量がばらつくことがわかります．また，同じ水準のなかでも反復によって収量はばらつきます．

表1 データの全偏差を処理偏差と誤差偏差に分割する

処理	データ(kg/ha)（$=x_{ij}$）				処理和(T)	処理平均（$=\bar{x}_{i.}$）
Dol-Mix(1 kg)	2,537	2,069	2,104	1,797	8,507	2,127
Dol-Mix(2 kg)	3,366	2,591	2,211	2,544	10,712	2,678
DDT+γ-BHC	2,536	2,459	2,827	2,385	10,207	2,552
Azodrin	2,387	2,453	1,556	2,116	8,512	2,128
Dimecron-Boom	1,997	1,679	1,649	1,859	7,184	1,796
Dimecron-Knap	1,796	1,704	1,904	1,320	6,724	1,681
Control	1,401	1,516	1,270	1,077	5,264	1,316
Grand total(G)					57,110	
Grand mean						2,040 総平均（$=\bar{x}_{..}$）

誤差偏差　全偏差　処理偏差

データ　総平均　処理平均

$$x_{ij}-\bar{x}_{..}=(\bar{x}_{i.}-\bar{x}_{..})+(x_{ij}-\bar{x}_{i.})$$

全偏差
総平均に対する
各データの偏差
（総効果）

処理偏差
総平均に対する
各処理平均の偏差
（処理効果）

誤差偏差
各処理平均に対する
データの偏差
（誤差効果）

図1 偏差分割の計算式

ばらつきが2種類あるので，よくわかりません．知りたい量の見通しをよくすることはできないでしょうか？

表1は，データがもつ総平均からの偏差（全偏差）が，処理要因と誤差要因の2要因に対応して処理偏差と誤差偏差に分割されるようすを図示化しています．全偏差は総平均からの個々のデータのばらつきの偏差を数値化しますが，処理偏差は水準ごとに計算された処理平均と総平均とのばらつき，そして誤差偏差は処理平均を基準にしたときの同水準内でのデータのばらつきをそれぞれ偏差として数値化しています．表1に視覚的に示された偏差の実際の計算式は図1のようになります．

各データに対して図1に 示した偏差分割式が書けます．この例では全部で7水準4反復の計28の式になります．続いて，それらの偏差を集計してデータ全体のばらつきを数値化するた めには，左辺の偏差を平方して全水準全反復にわたって足し合わせることで平方和を計算する必要があります（この操作の意味については，連載第4回参照）．ところが，それぞれの偏差分割式の右辺については平方展開しなければならないので，途中計算はやや複雑になります．しかし，処理偏差と誤差偏差の積の総和はゼロとなり〔$\sum_i\sum_j(\bar{x}_{i.}-\bar{x}_{..})(x_{ij}-\bar{x}_{i.})=0$〕，最終的に全平方和は処理平方和と誤差平方和の和になります．すなわち，**データのばらつき全体を処理要因と誤差要因に起因する2つの部分にきれいに切り分けられるということです．**

平方和にはかならず自由度がついてまわります．全平方和は総平均に対する計28偏差から構成される統計量ですから，すべての偏差の総和がゼロ になるという制約が1つ生じます．したがって，全平方和の自由度（全自由度）は28－1＝27となります．同様に，処理平方和については，総平均からの処 理平均の偏差すべての和はゼロとなるので，その自由度は7－1＝6です．すこし複雑なのは残った誤差平方和についてです．誤差平方和を構成する誤差偏差は 全部で28個ですが，その内訳は各水準ごとに計算された処理平均に対して同一水準内の4データとの偏差を7水準にわたって平方して集計します．このとき， 同一水準の4つの偏差は和がゼロになるという制約が生じます．つまり，各水準ごとに偏差は4つありますが，このうち自由に動ける偏差は3つだけということ です．この制約がすべての水準について成立しますから，誤差平方和には全部で7つの制約が課されることになります．したがって，誤差平方和の自由度は28－7＝21となります．この値は，全平方和の自由度（27）－処理平方和の自由度（6）からも算出できます．

§ ノイズに対するシグナルの大きさ：*F* 値

平方和と自由度…これって，分散が計算できませんか？

鋭いですね．平方和を自由度で割れば分散が算出できます．実験計画法では伝統的に分散を平均平方（mean square）と呼び習わしてきたので，以下でもこの用語を使うことにします．処理要因に関する分散すなわち「処理平均平方＝処理平方和÷ 処理自由度」，および誤差要因に関する分散すなわち 「誤差平均平方＝誤差平方和÷誤差自由度」が計算されたならば，いよいよ最後のステップです．

すべての実験には目的があります．いま私たちが対象としているデータは殺虫剤という実験処理がイネの収量にどれくらい効いたかを調べるのが目 的です．このとき私たちが知りたいのは，偶然誤差によるデータのばらつきに対して，殺虫剤による実験処理がもたらすデータのばらつきがどれほど大きいかと いう相対的な比較です．それぞれの要因によるばらつきはすでに平均平方という数値として計算されています．そこで，「処理平均平方÷誤差平 均平方」という比の値を考えてみましょう．この比を*F* 値と呼びます．処理と誤差の分散比を意味する*F* 値が大きければ私たちは偶然誤差という“ノイズ”よりも実験処理の“シグナル”の方が大きいので，殺虫剤による収量のちがいは「ある」と直感的に判定できます．ところが，*F* 値が小さいと“シグナル”が“ノイズ”にかき消されてしまい，「ある」という判定は直感的に難しくなってしまいます．

では，この*F* 値はどれくらい大きければ客観的に実験要因の効果が「ある」といえるのでしょうか？

確かに，主観的な判断で*F* 値が「大きい」と言うことはできません．私たちは生データから地道に計算し続けてようやく平均平方の比*F* 値まで到達しました．しかし，数値計算で進められるのはここまでです．そのとき，次の一歩がなかなか踏み出せない私たちに，雲の上から声が聞こえてきました．

§ 正規分布の仮定から得られる御神託：仮説検定という考え方

データからの数値計算が地上で進んでいたころ，雲の上の正規分布帝国の神殿でも動きがありました．前回説明したように，実験計画の最初の段階で私たちは「データ（x_{ij}）＝平均（μ）＋処理効果（α_i）＋誤差効果（ε_{ij}）」という統計モデルを仮定し，誤差効果は平均ゼロ，分散σ^2の正規分布N（0，σ^2）に従うと仮定しました．いま，仮に処理効果がない統計モデルすなわち「データ（x_{ij}）＝平均（μ）＋誤差効果（ε_{ij}）」を考えてみましょう．ここで，処理効果をもたないこの統計モデルを帰無仮説（null hypothesis）とよびます．これに対して処理効果をもつもとのモデルを対立仮説（alternative hypothesis）と名づけます．対立仮説は処理と誤差という2つの変動要因をもつのに対して，帰無仮説は誤差が唯一の変動要因です．つまり，帰無仮 説はデータのもつばらつきはすべて偶然誤差に起因すると宣言していることになります．

データを帰無仮説「$x_{ij}=\mu+\varepsilon_{ij}$」によって説明しようとすると，$\varepsilon_{ij}$が正規分布N（0，$\sigma^2$)という仮定により，データ$x_{ij}$は平均$\mu$をもつ正規分布N（$\mu$，$\sigma^2$)に従うことになります．第9回の確率分布曼荼羅を見ると，正規分布にしたがう確率変数から計算された平方和はカイ二乗分布という確率分布にしたがい，平方和を自由度で割った平均平方の比（*F* 値）は*F*分布という別の確率分布に従うことが数学的に証明できます．この*F*分布が地上の私たちに対して雲の上から届けられた御神託なのです．

帰無仮説のもとでの*F*分布は*F* 値に関する確率分布で，分子である処理平方和の自由度（6）と分母の誤差平方和の自由度（21）の2つのパラメーターで確率分布の形が決まります（図2）．

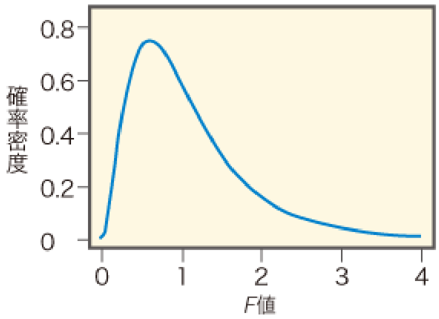

図2　F分布の確率密度関数
処理自由度（6）と誤差自由度（21）を2つのパラメーターとして描いた．

図2は，帰無仮説のもとでは*F* 値はどのような値を取りやすいかを私たちにはっきり示します．グラフの頂点が*F* 値の1あたりにあることに注意してください．これは，帰無仮説のもとでは処理平均平方と誤差平均平方との比がほぼ1であること，すなわち処理平均平方と誤 差平均平方とはほぼ同じ大きさをもつことを意味します．これは驚くようなことではなく，帰無仮説では処理効果がもともと仮定されていないわけですから，*F* 値を構成する処理平均平方はたかだか誤差平均平方程度のばらつきしか生み出さないと考えれば，直感的に納得できるでしょう．逆に言えば，殺虫剤による処理効果が大きければ大きいほど*F* 値は1よりも大きな値をとります．そこで，*F*分布の上側末端部に棄却域（critical region）を設定します．

棄却域はどのくらいにするのがよいのでしょう？

*F*分布の確率密度関数の下の部分の全面積は1ですから，棄却域はたとえば面積0.05（5%基準）あるいは0.01（1%基準）と設定するのがふつうです．そして，データから計算された*F* 値がこの棄却域に入るほど大きくなったら，そのときは「処理効果はない」と宣言する帰無仮説を捨てて「処理効果はある」とする対立仮説を採用しようという意思決定方針を立てることにします（図3）．ここで重要な点は，棄却域は数値的に決定できるので，データから得られた*F* 値が棄却域に入るかどうかは客観的に決定できるという点です．この*F*検定に基づく仮説検定は，実験計画法における分散分析法（analysis of variance）の根幹です．実際にデータから計算した結果を表2に示します．

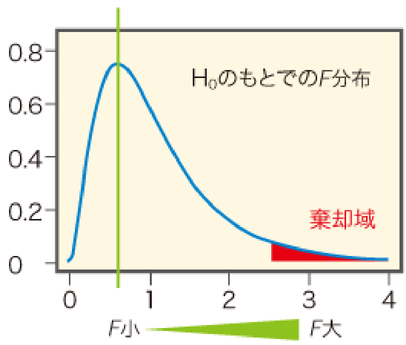

図3　F分布を用いた仮説検定の考え方
帰無仮説のもとで設定された棄却域にデータから計算されたF値が入るかどうかによって，帰無仮説を棄却するかどうかを決定する

表2　分散分析表

変動要因	自由度	平方和	平均平方	F値	$F_{0.05}(6, 21)$	$F_{0.01}(6, 21)$
全体	27	7,577,412				
処理	6	5,587,175	931,196	9.8255**	2.57	3.81
誤差	21	1,990,237	94,773			

上の結果から，Box1 の殺虫剤散布実験では，処理要因の効果は 1%レベルで有意であることが判明した．

この例では，5%棄却域は*F* 値が2.57以上，1%棄却域は3.81以上です．そして，デー

タから計算されたF値は9.82ですから，5%棄却域はもちろん1%棄却域に入るほど大きな値と判定されます．したがって，この実験では1%レベルで有意（significant）な効果を殺虫剤の処理要因は示せたという結論になります．まさにデータからの計算と統計理論が合体した瞬間です．

　このように，データからの数値計算と正規分布に基づく統計理論の両方があってはじめて，処理要因が有意であったかどうかが検定できるわけです．次回は乱塊法という実験計画法について説明したうえで，連載のしめくくりをしたいと思います．

[SHARE] Tweet シェア B!

Prev 11 Next

TOP

「第12回 データ解析と統計的推論ー連載の総括として」は，本誌2015年4月号を御覧ください