7 2種類のアメーバによる感染の独立性

はじめに

人間の腸内に異なった2種類のアメーバがいたら，どういうことが起こるだろうか．1つの種類のアメーバは病気を引き起こすが，もう一方の種類はそうでなく，かえってこの第2の種類が同じ腸に生殖していることで，最初の種類のアメーバによって起こされる病気を防いでいるかもしれない．ここでは簡単な2×2 分割表による χ^2 検定で，第1のアメーバの感染と，第2のアメーバの感染とは互いに影響するかどうかを調べてみる．

重い腸の伝染病が，ある木工工場で働く人々の間で発生した．医者達は，その病気が，腸内で生長する，あるアメーバに原因があるとした．他のいろいろな可能性を検討したのち，保健所は，その木工工場に水を供給していた水道管に穴があいていて，しかも，その近くをこれまた穴のあいた下水管が通っていたので，アメーバが水道管にはいりこんだ，という結論を下した．その工場で働いていて，病気にかからなかった人のどれくらいが，気が付かずに，このアメーバの影響を受けていたかを調べた．保健所はその工場内で健康そうに働いている労働者138人を無作為に選び，検査した．その結果をまとめたのが表1である．70人からアメーバが検出された．疫学上このアメーバは大きな種類と小さな種類の2種類に区別されている．これによって分けると，大きな種類は35人の腸内に，小さな種類は23人に発見された．そして12人には両方とも発見されたのである．

著者は Joel E. Cohen (ハーバード大学).

表 1

		大きな種類のアメーバ		計
		検出あり	検出なし	
小さな種類のアメーバ	検出あり	12	23	35
	検出なし	35	68	103
計		47	91	138

このデータから，次のような興味ある問題が考えられる．1つの種類に感染していることは，他の種類に感染することを弱めているか，または逆に強めているか．つまり，感染している体内で，2つの種類は互いに影響し合っているのだろうか．

一般に，2つの事象 A, B が独立のとき，2つが同時に起こる確率は積 $P(A)P(B)$ で与えられた．試行を n 回繰り返すとき，この同時に起こる事象の期待度数は $nP(A)P(B)$ である．A の余事象を $\bar{A}$, B の余事象を $\bar{B}$ とするとき，事象 A, B が互いに独立ならば，事象 $\bar{A}$ と B, A と $\bar{B}$, $\bar{A}$ と $\bar{B}$ もまた互いに独立となる．これは後の練習問題6で証明してもらう．そこで，A と B が互いに独立であるとき，2×2 分割表での期待度数は次のように書ける：

	B	$\bar{B}$	計
A	$nP(A)P(B)$	$nP(A)P(\bar{B})$	$nP(A)$
$\bar{A}$	$nP(\bar{A})P(B)$	$nP(\bar{A})P(\bar{B})$	$nP(\bar{A})$
計	$nP(B)$	$nP(\bar{B})$	n

上の表の $nP(A)P(B)$ について

$$nP(A)P(B)=\frac{(\text{横の合計})\times(\text{縦の合計})}{n}$$

である．他の場合についても同じことがいえる．実際に数値を記入した 2×2 分割表で，横の合計はそれぞれ $nP(A)$, $nP(\bar{A})$ の推定値，縦の合計はそれぞれ $nP(B)$, $nP(\bar{B})$ の推定値である．

さて表1のデータにもどるとしよう．大きな種類のアメーバの検出あり，なしが小さな種類の感染に何の影響も与えないならば，大きな種類が検出された 47 人のうち，小さな種類を検出される人の割合は 35/138，つまり人数では

$47\times(35/138)=11.9$ 人いるはずである．このように感染の間には何らの相互作用がないとしてそれぞれの期待度数を求めたのが表2である．

表 2　感染の期待度数(各々の合計は前と同じにして，2種類のアメーバの間には相互作用がないとして計算した)

		大きな種類のアメーバ		計
		検出あり	検出なし	
小さな種類のアメーバ	検出あり	11.9	23.1	35.0
	検出なし	35.1	67.9	103.0
計		47.0	91.0	138.0

前の観測度数と，相互作用がないとして求めたこの期待度数を比較すれば，両者はかなり近い．したがって，一方のアメーバに感染すると他のアメーバにも感染しやすくなるという主張は否定される．

練習問題

1. すべてのマスに対する期待度数は，各マスに対する横と縦の合計を用いて(横の合計)×(縦の合計)/(全体の合計) からも求めることができる．このようにして得られた各期待度数を横に加え合わせたものは観測度数の表における横の合計に等しい．これを示しなさい．

2. 一般に 2×2 マスの中の数が次のように与えられたとして練習問題1を計算しなさい：

a	b
c	d

3. 横と縦の合計が与えられたとした表2の例で，各マスに対して適当な整数値を与えた場合，表1の数値より期待度数に近くなる数値の与え方はあるか．

4. 横と縦が独立と仮定して 2×2 分割表の χ^2 の期待値を求めなさい．

5. この例の場合について χ^2 の値を計算することは意味がないように思える．どうしてか．

6. A と B が独立ならば，$\bar{A}$ と B もそうであることを示しなさい．

7. 2つのアメーバの種類の間に独立性を発見することは重要である．それを

実際の立場から説明しなさい．強い関連があることを見いだすことはどんな重要性を持つか説明しなさい．

／解　　答／

練習問題

1.　マスの数が 23 の横の合計は 35. 同じマスの縦の合計は 91. 全体の合計は 138 だから，

$$\frac{(\text{横の合計})(\text{縦の合計})}{(\text{全体の合計})}$$

の公式を適用すれば，期待度数として $35\times91/138\fallingdotseq23.1$ を得る．同様に 12 のマスに対しては $35\times47/138=11.9$ を得る．$23.1+11.9=35.0$ となっている．

2.　表より横と縦の合計を出し，全体の合計を T とおく．

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	$a+b+c+d=T$

公式を使うと，a のマス，b のマスに対する期待度数はそれぞれ $\dfrac{(a+b)(a+c)}{T}$，$\dfrac{(a+b)(b+d)}{T}$. したがってこれを加えてみると $\dfrac{(a+b)(a+c)}{T}+\dfrac{(a+b)(b+d)}{T}=a+b$ でちょうど，前の $a+b$ と一致している．一般に，r 行 c 列の長方形の分割表に対しても同様のことが成り立つ．証明を試みるとよい．

3.　横と縦の合計が与えられたとき，この観測した結果より期待度数に近い数の組はない．

4.　χ^2 の期待値は，その自由度に等しいことが知られている．2×2 分割表の自由度は $(2-1)(2-1)=1$ だから，χ^2 の期待値も 1 である．一般の $r\times c$ 分割表では，$(r-1)(c-1)$ となる．

5.　観測した度数の組合せが期待度数に近いから，χ^2 の値はその期待値 1 よりはるかに小さくなるだろう．

6.　証明する前に，次のことを説明するとよい．今，A は大きな種類のアメーバがいること，B は小さなものがいることを表わすとする．証明すべき結果によれば，大きな種類がいることと小さなものがいることとが独立ならば（上の例で観測されたように），大きな種類がいないことは，小さなものがいることとも独立になる．

まず

$$P(B)=P(BA)+P(B\bar{A})=P(B\mid A)P(A)+P(B\mid \bar{A})P(\bar{A})$$

である．また A と B が独立であるから

$$P(B\mid A)=P(B)$$

となり，さらに

$$P(A)=1-P(\bar{A})$$

が成立する．この 2 つを最初の式に代入し，$P(\bar{A})\neq0$ に注意すると，

$$P(B)=P(B \mid \bar{A})$$

を得る．よって，B と $\bar{A}$ との独立性が示された．

7.　大小2種類のアメーバの感染は独立ではなく，病気を誘発しない小さな種類のアメーバに感染していると，大きな種類に感染することが，ほとんど起こらなくなると仮定する．このことを発見すれば，小さな種類の感染が，大きな種類の感染に対して予防効果として利用できる．同じ考えが予防注射のもととなっている．ジェンナーは，人間が一度，感染してもたいしたことのない牛痘にかかると，危険な天然痘を予防できることを発見した．今の場合アメーバの種類には独立に感染するということなので，小さな種類が病気を誘発しないといっても，大きな種類のアメーバの感染からの予防としては役立たない．

やさしい例による統計入門 下

STATISTICS BY EXAMPLE

F.モステラー　W.クラスカル　R.リンク
R.ピーターズ　G.リージング　共編
村上正康　監訳

培風館

〔共 訳 者〕

村上 正康

蔵野 正美

安田 正實

田栗 正章

中神 潤一

STATISTICS BY EXAMPLE

edited by

Frederick Mosteller, William H. Kruskal,
Richard F. Link, Richard S. Pieters,
and Gerald R. Rising