ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XI, 4 (1970)

S. ZUBRZYCKI (Wrocław)

WIELKOPRÓBKOWA ANALIZA PEWNEGO ZADANIA EKOLOGICZNEGO, CZYLI POGOŃ ZA ŚLIMAKAMI

Przedmiotem tej noty jest dyskusja zadania statystycznego podpowiedzianego przez badania populacji ślimaków (Łomnicki [2]). Chodzi o szacowanie średniej liczby zwierząt pewnego gatunku na jednostce powierzchni siedliska na podstawie łowienia, znakowania i powtórnego odłowu na ograniczonym poletku, gdy musimy się liczyć zarówno z ruchliwością zwierząt, jak i z możliwością ukrycia się ich w podłożu. Uwagi nasze odnoszą się w krytycznym punkcie rozważań do wielkich próbek, gdy możemy uważać średnie próbkowe za dostatecznie dokładne przybliżenie średnich w populacji macierzystej. Podajemy tylko wstępne uwagi o tym, jak należałoby uwzględnić wahania losowe przy małej liczbie obserwacji.

1. Zadanie przyrodnicze. W siedlisku zamieszkałym przez ślimaki wyróżnia się kwadratowe poletko, aby na podstawie obserwacji na tym poletku ocenić średnią liczbę ślimaków na jednostce powierzchni. Będziemy tę średnią oznaczali przez λ i nazywali *gęstością*.

Zadanie byłoby stosunkowo proste, gdyby należało się liczyć tylko z ruchliwością ślimaków (por. § 2). Nie byłoby również kłopotów z oceną, gdyby nie było ruchliwości ślimaków, a tylko możliwość ukrycia się ich w podłożu. Porównując wówczas liczbę ślimaków powtórnie odłowionych z liczbą ślimaków oznaczonych podczas pierwszej inspekcji, łatwo byłoby oszacować prawdopodobieństwo pozostania ślimaka na powierzchni i odpowiednio skorygować zaobserwowaną pozorną gęstość ślimaków w siedlisku. Trudność wynika z konieczności jednoczesnego uwzględnienia ruchliwości i możliwości ukrycia się. To z tego powodu, gdy w czasie powtórnej inspekcji nie znajdziemy poprzednio oznaczonego ślimaka, nie mamy pewności, czy ukrył się on, czy też opuścił poletko. Obserwacje przeprowadza się więc w ten sposób, że poletko ogląda się wielokrotnie w pewnych odstępach czasu, numeruje się wszystkie znalezione ślimaki dla późniejszej identyfikacji, notuje ich pozycje i zwraca na to samo miejsce. Podstawowym zagadnieniem jest rozważenie, jak można wnioskować o gęstości λ z dwukrotnej obserwacji poletka. W przypadku kilku inspekcji, każda z nich stanowi punkt wyjściowy dla następnej, a to pozwala w oczywisty sposób wykorzystać w tej bardziej złożonej sytuacji wnioski odnoszące się do dwukrotnej inspekcji.

2. Ocena punktowa i przedziałowa, gdy ślimaki się nie kryją. Gdy ślimaki się nie kryją, możemy podczas inspekcji zaobserwować i zliczyć wszystkie ślimaki obecne na poletku. Biorąc za podstawę rozważań model losowego rozmieszczenia ślimaków w siedlisku, sprowadzamy zadanie do szacowania parametru rozkładu Poissona: jeśli w siedlisku przypada średnio λ ślimaków na jednostkę powierzchni, to liczba x ślimaków, znajdujących się w danym momencie na poletku o powierzchni p, jest zmienną losową o rozkładzie Poissona z wartością oczekiwaną λp. A zatem, *przy losowym rozmieszczeniu w siedlisku nie kryjących się ślimaków, iloraz x/p, gdzie x jest zaobserwowaną w czasie inspekcji liczbą ślimaków na poletku o powierzchni p, jest nieobciążonym punktowym oszacowaniem gęstości λ.*

Dla dużych x można zastosować przybliżenie normalne do rozkładu Poissona i twierdzić, że x/p ma w przybliżeniu rozkład normalny z wartością oczekiwaną λ i wariancją λ/p. Wówczas w przybliżeniu z prawdopodobieństwem $1-\alpha$ zaobserwowana liczba ślimaków x znajdzie się w przedziale

$$p\lambda - c_\alpha\sqrt{p\lambda} < x < p\lambda + c_\alpha\sqrt{p\lambda}, \tag{2.1}$$

gdzie c_α jest odczytaną z tablic wartością krytyczną taką, że dla zmiennej losowej ξ o rozkładzie normalnym ze średnią 0 i wariancją 1 mamy

$$\Pr\{|\xi| > c_\alpha\} = \alpha.$$

Przekształcając nierówności (2.1), otrzymujemy

$$x - c_\alpha\sqrt{p\lambda} < p\lambda < x + c_\alpha\sqrt{p\lambda}$$

i następnie

$$\frac{x}{p} - c_\alpha\sqrt{\frac{\lambda}{p}} < \lambda < \frac{x}{p} + c_\alpha\sqrt{\frac{\lambda}{p}}.$$

Wreszcie, zastępując w skrajnych wyrazach nieznaną gęstość λ przez jej oszacowanie x/p, widzimy, że *wzór*

$$\frac{x}{p} - \frac{c_\alpha}{p}\sqrt{x} < \lambda < \frac{x}{p} + \frac{c_\alpha}{p}\sqrt{x} \tag{2.2}$$

określa (przybliżone) przedziały ufności dla λ na poziomie ufności $1-\alpha$.

Moglibyśmy wprost odwrócić nierówności (2.1) dla oszacowania λ. Zamiast (2.1) możemy mianowicie napisać

$$-c_\alpha\sqrt{p\lambda} < x - p\lambda < c_\alpha\sqrt{p\lambda},$$

a to jest równoważne z tym, że $(x-p\lambda)^2 < c_\alpha^2 p\lambda$, czyli że

$$p^2\lambda^2 - (2x + c_\alpha^2)p\lambda + x^2 < 0,$$

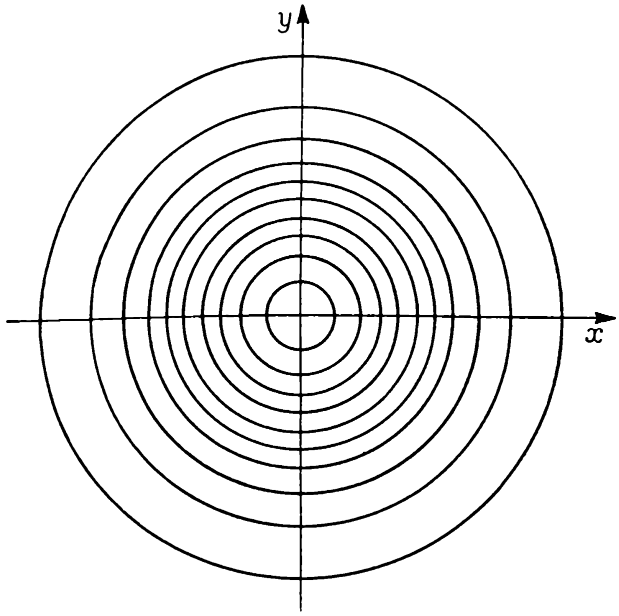
a więc λ ma być zawarte między pierwiastkami trójmianu kwadratowego po lewej stronie ostatniej nierówności. Tak otrzymujemy inny *przybliżony wzór*

$$\frac{x}{p}+\frac{c_a^2}{2p}-\frac{c_a}{p}\sqrt{x+\left(\frac{c_a}{2}\right)^2}<\lambda<\frac{x}{p}+\frac{c_a^2}{2p}-\frac{c_a}{p}\sqrt{x+\left(\frac{c_a}{2}\right)^2} \tag{2.3}$$

na przedziały ufności dla gęstości λ *na poziomie ufności* $1-\alpha$.

PRZYKŁAD. Dla $p = 25$, $x = 300$ i $\alpha = 0{,}05$ mamy $c_a = 1{,}96$. Nierówności (2.2) dają wówczas $10{,}642 < \lambda < 13{,}358$, a nierówności (2.3) dają $10{,}679 < \lambda < 13{,}399$.

3. Ocena ruchliwości. Przez *ruchliwość* ślimaków będziemy tu rozumieli rozkład prawdopodobieństwa wektora łączącego pozycję wyjściową ślimaka w chwili pierwszej inspekcji z jego pozycją końcową w chwili



Rys. 1

powtórnej inspekcji. O rozkładzie tym będziemy zakładali, że jest centralnie symetrycznym rozkładem normalnym ze środkiem w początku układu (zob. rys. 1 pokazujący kilka warstwic gęstości takiego rozkładu). Innymi słowy, zakładać będziemy, że wektor ten ma rozkład prawdopodobieństwa o gęstości

$$f(x, y \mid \sigma) = \frac{1}{2\pi\sigma^2}\exp\{-(x^2+y^2)/2\sigma^2\}. \tag{3.1}$$

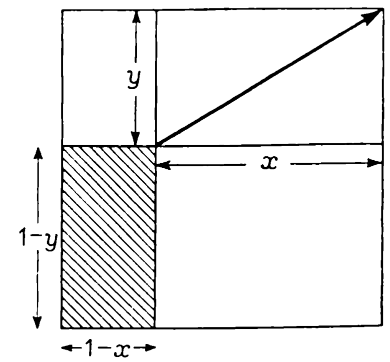
Rozkład taki jest scharakteryzowany przez jeden parametr, a mianowicie przez odchylenie standardowe σ swych składowych. Wobec tego ocenę ruchliwości będziemy uważali za jednoznaczną z oceną odchylenia standardowego σ.

Ocenę tą mamy robić na podstawie obserwacji ograniczonych do ustalonego połetka. Dlatego nie będziemy mogli obserwować pozycji końcowej, gdy będzie ona poza tym poletkiem. A więc będziemy mieli do dyspozycji jako podstawę oceny pewną ilość pozycji wyjściowych i pozycji końcowych wewnątrz poletka.

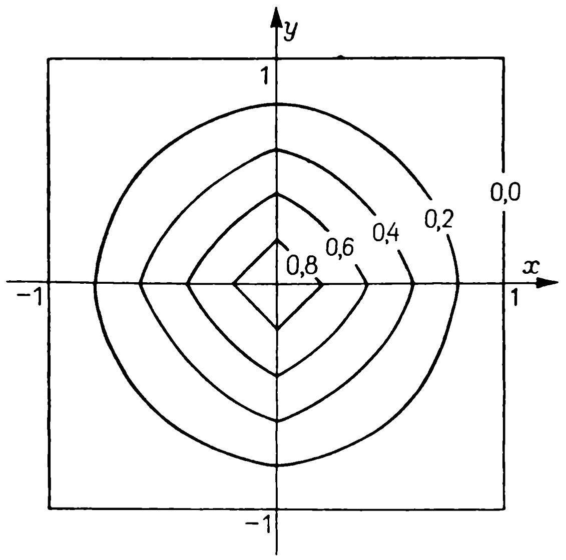
Gdyby wykluczyć możliwość krycia się ślimaków, każdy brak pozycji końcowej należałoby tłumaczyć opuszczeniem poletka przez ślimaka. Ponieważ jednak dopuszczamy krycie się ślimaków, brak pozycji końcowej ślimaka można również przypisywać temu, że ślimak się ukrył. Jak wobec tego oceniać ruchliwość? Otóż dla dalszej dyskusji będziemy zakładali, że krycie się ślimaków jest statystycznie niezależne od ich ruchliwości. Innymi słowy, przyjmować będziemy, że ukrycie się ślimaka jest niezależne od tego, czy oddalił się on od swej pozycji wyjściowej dużo czy mało. Będziemy także przyjmowali, że poszczególne ślimaki kryją się niezależnie jeden od drugiego. Wówczas te ślimaki, u których zaobserwowaliśmy zarówno pozycję wyjściową, jak i pozycję końcową możemy uważać za próbkę wylosowaną spośród tych wszystkich ślimaków, u których zaobserwowaliśmy pozycję wyjściową i które nie były ukryte w chwili dokonywania powtórnej inspekcji. Są to mianowicie ślimaki znalezione podczas pierwszej inspekcji i przebywające na powierzchni podczas drugiej inspekcji, które ponadto znajdują się podczas drugiej inspekcji na badanym poletku. Przy poczynionych założeniach o niezależności możemy z tak otrzymanej próbki oceniać ruchliwość dokładnie tak samo, jak robilibyśmy to, gdyby nie było krycia się ślimaków.

Ponieważ obserwacje ograniczone są do kwadratowego poletka, obserwowane wektory łączące pozycje wyjściowe z końcowymi nie będą miały rozkładu prawdopodobieństwa o gęstości (3.1), lecz będzie to rozkład powstały z (3.1) przez rodzaj obcięcia. Przyjmijmy dla dalszych rozważań, że bok kwadratowego poletka ma długość 1, a osie układu współrzędnych (x, y) obierzmy tak, aby były równoległe do krawędzi poletka. Wówczas na to, aby wektor (x, y), łączący pozycję wyjściową z pozycją końcową, mógł być zaobserwowany, jego składowe muszą być co do modułu mniejsze niż 1, a ponadto pozycja wyjściowa musi leżeć wewnątrz pewnego zawartego w poletku prostokąta o bokach $1-|x|$ i $1-|y|$ (zob. rys. 2, który przedstawia sytuację dla $0 < x < 1$ i $0 < y < 1$). Tak więc, prawdopodobieństwo tego, że wektor, łączący pozycję wyjściową z pozycją końcową, zostanie zaobserwowany, gdy jest on równy (x, y), a pozycja wyjściowa jest losowana z jednostajnym rozkładem prawdopodobieństwa z badanego poletka, jest dane przez zależność

$$P(x, y) = \begin{cases} (1-|x|)(1-|y|) & \text{dla } |x| < 1 \text{ i } |y| < 1, \\ 0 & \text{poza tym.} \end{cases} \tag{3.2}$$



Rys. 2



Rys. 3

Kilka warstwic funkcji $P(x, y)$, z których każda jest złożona z czterech łuków hiperbol, przedstawia rys. 3.

Dalej, przy losowym rozmieszczeniu ślimaków w siedlisku, pozycje wyjściowe zachowują się łącznie tak, jakby pochodziły z niezależnego losowania punktów z jednostajnym rozkładem prawdopodobieństwa na badanym poletku. Dlatego obserwowany rozkład prawdopodobieństwa wektora, łączącego pozycję wyjściową z pozycją końcową, ma gęstość prawdopodobieństwa

$$\varphi(x, y \mid \sigma) = \frac{f(x, y \mid \sigma) P(x, y)}{\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f(x, y \mid \sigma) P(x, y)\, dx\, dy}. \tag{3.3}$$

Wedle poczynionych założeń ruchliwość ślimaków jest izotropowa, to jest niezależna od kierunku. Dlatego chcielibyśmy ją oceniać na podstawie zaobserwowanych odległości między pozycją wyjściową i końcową. Zanotujmy, że odległości te mają rozkład prawdopodobieństwa o gęstości

$$r(x \mid \sigma) = \begin{cases} \dfrac{1}{\sigma^2} x \exp\{-x^2/2\sigma^2\} & \text{dla } x > 0, \\ 0 & \text{poza tym.} \end{cases} \tag{3.4}$$

Jest to tzw. *rozkład Rayleigh'a* (zob. np. Helstrom [1], rozdz. 5, § 3). Jednak obserwowany rozkład odległości między pozycją wyjściową i końcową będzie inny, a mianowicie taki, jak rozkład odległości od początku układu współrzędnych punktu losowanego według rozkładu o gęstości (3.3). Łatwy rachunek pokazuje, że jest to rozkład o gęstości

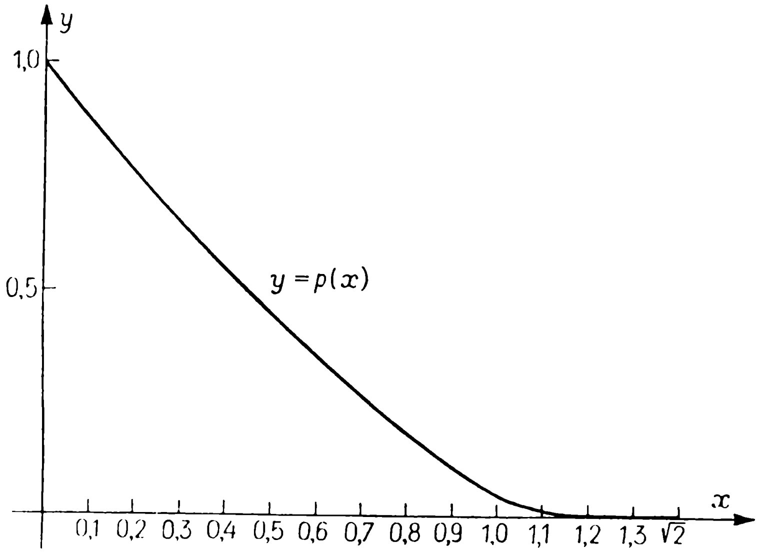
$$\varrho(x \mid \sigma) = r(x \mid \sigma) p(x) \Big/ \int\limits_{0}^{\infty} r(x \mid \sigma) p(x)\, dx, \tag{3.5}$$

gdzie $p(x)$ jest średnią wartością funkcji $P(u, v)$ na kole $u^2+v^2 = x^2$.

Funkcję $p(x)$ można wyrazić w postaci jawnej przez całkowanie. Mamy mianowicie

$$p(x) = \begin{cases} \frac{4}{\pi}\left\{\frac{\pi}{4} - x + \frac{1}{4}x^2\right\} & \text{dla } 0 < x \leqslant 1, \\ \frac{4}{\pi}\left\{\frac{\pi}{4} - \left(\arccos\frac{1}{x}\right) + \sqrt{x^2-1} - \frac{1}{4}x^2 - \frac{1}{2}\right\} & \text{dla } 1 \leqslant x \leqslant \sqrt{2}, \\ 0 & \text{poza tym.} \end{cases}$$

Niektóre wartości funkcji $p(x)$ podano w tablicy 1, a jej wykres na rys. 4.



Rys. 4

TABLICA 1

x	$p(x)$	x	$p(x)$	x	$p(x)$
0	1	0,50	0,4430	1,00	0,0451
0,05	0,9371	0,55	0,3960	1,05	0,0256
0,10	0,8759	0,60	0,3506	1,10	0,0146
0,15	0,8162	0,65	0,3069	1,15	0,0079
0,20	0,7581	0,70	0,2647	1,20	0,0039
0,25	0,7016	0,75	0,2241	1,25	0,0016
0,30	0,6467	0,80	0,1851	1,30	0,0005
0,35	0,5934	0,85	0,1477	1,35	0,0001
0,40	0,5416	0,90	0,1119	1,40	0,0000
0,45	0,4915	0,95	0,0777	$\sqrt{2}$	0

4. Jakiej statystyki użyć do estymowania ruchliwości? Gęstość prawdopodobieństwa n niezależnych obserwacji o gęstości danej wzorem (3.3) można zapisać w postaci

$$(4.1)\quad \varphi(x_1, y_1|\sigma)\cdot\ldots\cdot\varphi(x_n, y_n|\sigma) = h(\sigma)\cdot k(x_1, y_1, \ldots, x_n, y_n)\cdot e^{-TU/2},$$

gdzie

$$h(\sigma) = \Big\{(2\pi\sigma^2)^n\Big[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y|\sigma)P(x, y)\,dxdy\Big]^n\Big\}^{-1},$$

$$k(x_1, y_1, \ldots, x_n, y_n) = P(x_1, y_1)\cdot\ldots\cdot P(x_n, y_n),$$

$$T = x_1^2+y_1^2+\ldots+x_n^2+y_n^2, \qquad U = 1/\sigma^2.$$

Wobec twierdzenia o faktoryzacji (por. [3], § 47) dowodzi to, że *suma kwadratów zaobserwowanych odległości między pozycją wyjściową i końcową jest statystyką dostateczną dla wnioskowania o odchyleniu standardowym* σ, a zatem ta suma kwadratów powinna być użyta jako podstawa estymacji odchylenia σ.

5. Wielkopróbkowa ocena odchylenia standardowego σ. Niech K oznacza oczekiwaną wartość kwadratu długości obserwowanego wektora łączącego pozycję wyjściową i końcową. K zależy od ruchliwości ślimaków, mianowicie $K = k_1(\sigma)$, gdzie

$$(5.1)\qquad k_1(\sigma) = \int_0^{\infty} x^2\varrho(x|\sigma)\,dx.$$

Niech $k_1^{-1}(u)$ oznacza funkcję odwrotną do funkcji $k_1(\sigma)$. Otóż za ocenę odchylenia standardowego σ możemy przyjąć $\hat{\sigma} = k_1^{-1}(\hat{K})$, gdzie $\hat{K}$ jest średnią kwadratów zaobserwowanych odległości między wyjściowymi i końcowymi pozycjami ślimaków.

Zgodnie z teorią asymptotyczną (zob. np. [3], § 44), otrzymamy w ten sposób estymator odchylenia σ, który ma asymptotycznie rozkład normalny z wartością oczekiwaną σ i wariancją

$$(5.2)\qquad \left(\frac{1}{[dk_1(t)/dt]_{t=\sigma}}\right)^2\cdot\frac{w}{n},$$

gdzie σ jest prawdziwą wartością odchylenia standardowego charakteryzującego ruchliwość, w jest wariancją kwadratu obserwowanej odległości między pozycją wyjściową i końcową, czyli

$$w = \int_0^{\infty} x^4\varrho(x|\sigma)\,dx-\big(k_1(\sigma)\big)^2,$$

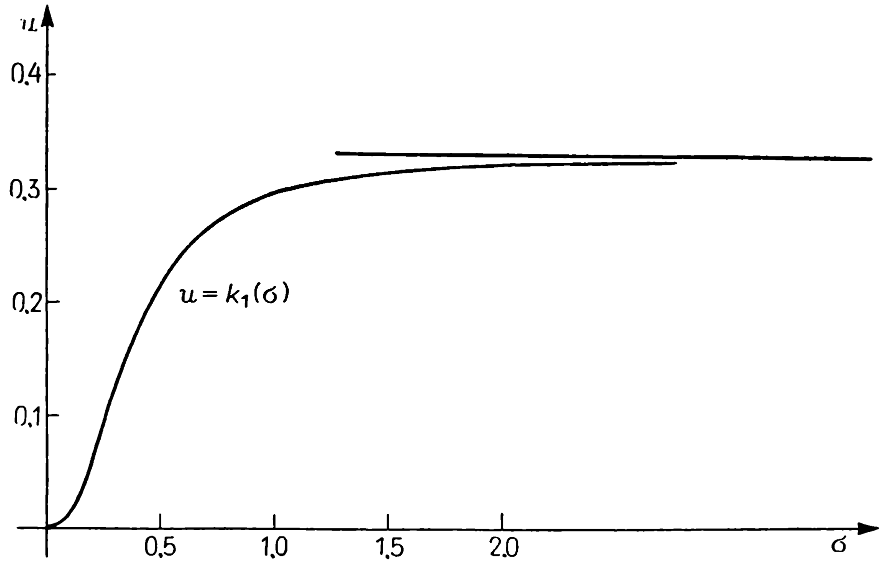
a n jest liczbą obserwacji dotyczących odległości. Możemy uzyskać informację o dokładności oceny, podstawiając w wyrażeniu (5.2) znalezioną przez nas ocenę $\hat{\sigma}$.

Niektóre wartości funkcji $k_1(\sigma)$, uzyskane przez całkowanie numeryczne, podajemy w tablicy 2, a jej wykres na rys. 5.

Opisane wielkopróbkowe postępowanie ma jedną wadę. Mianowicie przy $\sigma \to \infty$ funkcja $k_1(\sigma)$ dąży do skończonej granicy

$$\int_0^\infty x^2 p(x)\,dx \Big/ \int_0^\infty p(x)\,dx \approx 0{,}3331,$$

podczas gdy zaobserwowane średnie kwadraty odległości między pozycją wyjściową i końcową mogą przyjąć dowolną wartość między 0 i 2. Dla-



Rys. 5

tego wprawdzie z małym, i tym mniejszym im liczniejsza próbka i im mniejsza ruchliwość ślimaków, ale jednak z dodatnim prawdopodobieństwem proponowana wyżej ocena odchylenia standardowego σ nie będzie określona. Jest to niepożądana konsekwencja obcinania obserwacji. Dodajmy, że metoda największej wiarogodności prowadzi do równania $k_1(\sigma) = \hat{K}$, a zatem do oceny $\hat{\sigma} = k_1^{-1}(\hat{K})$.

Skonstruowanie racjonalnej oceny odchylenia σ bez omawianej wyżej przykrej własności pozostaje sprawą otwartą.

6. Wielkopróbkowa ocena gęstości λ. Ocena odchylenia standardowego σ interesuje nas przede wszystkim jako środek do oceny gęstości λ. Otóż, obserwowana gęstość ślimaków jest równa

$$g = \lambda \cdot Q,$$

gdzie Q jest prawdopodobieństwem pozostania ślimaka na powierzchni. Stąd

$$\lambda = \frac{g}{Q}. \tag{6.1}$$

Gęstość g możemy łatwo ocenić stosując postępowanie z § 2 do ślimaków zauważonych na powierzchni. Chodzi o ocenę Q.

Otóż, porównując liczbę ślimaków, u których zaobserwowaliśmy pozycję wyjściową, z liczbą ślimaków, u których zaobserwowaliśmy pozycję końcową, możemy ocenić prawdopodobieństwo q ponownego zaobserwowania ślimaka, tzn. prawdopodobieństwo jednoczesnego zajścia zdarzenia, że ślimak zaobserwowany podczas pierwszej inspekcji będzie w czasie drugiej inspekcji na powierzchni, i zdarzenia, że znajdować się on będzie na poletku. Prawdopodobieństwo pierwszego z tych zdarzeń oznaczyliśmy przez Q; oznaczmy prawdopodobieństwo drugiego zdarzenia przez R. Tak więc

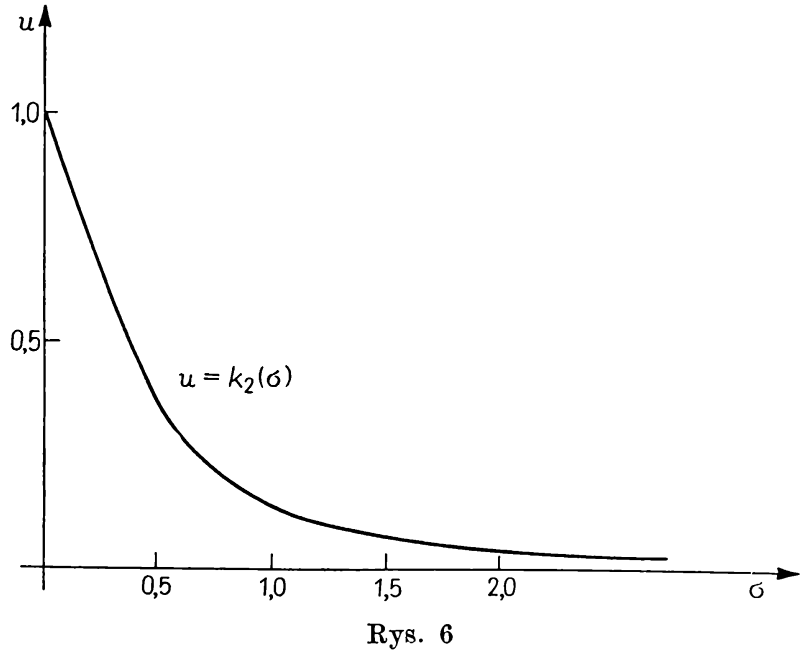
$$q = Q \cdot R. \tag{6.2}$$

W związku tym q możemy ocenić bezpośrednio z obserwacji, a R zależy od ruchliwości ślimaków, mianowicie $R = k_2(\sigma)$, gdzie

$$k_2(\sigma) = \int_0^\infty r(x|\sigma)p(x)dx \tag{6.3}$$

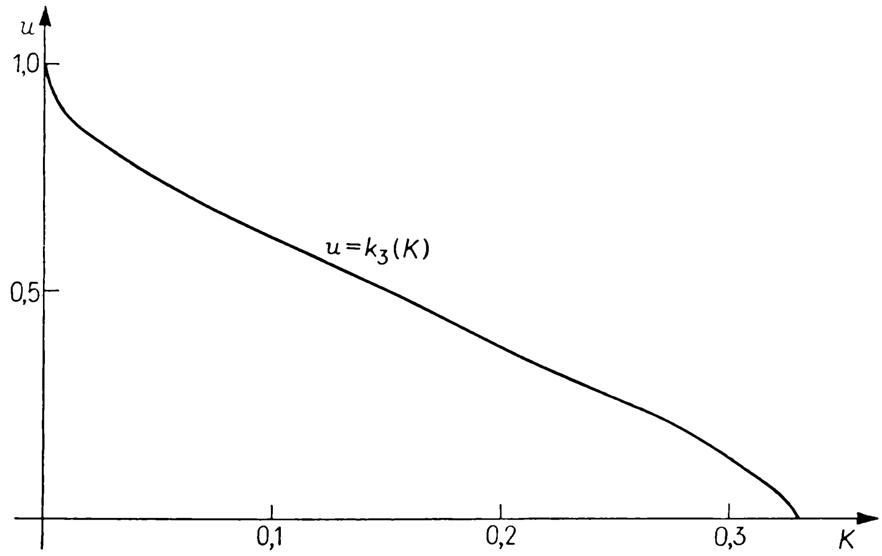
(por. wzór (3.5)).

Funkcję $k_2(\sigma)$ możemy stablicować. Kilka wartości otrzymanych przez całkowanie numeryczne podajemy w tablicy 2, a jej wykres na rys. 6.



Rys. 6

Znając funkcje $K = k_1(\sigma)$ i $R = k_2(\sigma)$, możemy także wyznaczyć zależność R od K. Niech $R = k_3(K)$. Zależność tę ilustruje wykres na rys. 7 otrzymany z danych tablicy 2.



Rys. 7

Ostatecznie

$$\lambda = \frac{g}{Q} = \frac{g \cdot R}{q} \tag{6.4}$$

i na tej podstawie można zaproponować zależność

$$\hat{\lambda} = \frac{\hat{g} \cdot k_3(\hat{K})}{\hat{q}} \tag{6.5}$$

gdzie $\hat{g}$ jest zaobserwowaną gęstością ślimaków przebywających na powierzchni, $\hat{q}$ jest zaobserwowaną frakcją ślimaków powtórnie złowionych, $\hat{K}$ jest średnią kwadratów zaobserwowanych odległości między pozycją wyjściową i końcową ślimaków, a $k_3(\hat{K})$ jest oceną nieznanego prawdopodobieństwa R, jako wielkopróbkową ocenę gęstości λ.

Zwróćmy uwagę, że zgodnie z (6.2) musi być $q \leqslant R$, a przypadek $q = R$ odpowiada sytuacji, kiedy ślimaki się nie kryją. Może się jednak z dodatnim prawdopodobieństwem zdarzyć, że $\hat{q} > \hat{R} = k_3(\hat{K})$. Jest to niepożądana własność estymatora (6.5). Można ją formalnie wykluczyć zmieniając definicję estymatora na następującą:

$$\hat{\lambda} = \begin{cases} \dfrac{\hat{g} \cdot k_3(\hat{K})}{\hat{q}}, & \text{gdy } \hat{q} \leqslant k_3(\hat{K}), \\ \hat{g}, & \text{gdy } \hat{q} > k_3(K). \end{cases} \tag{6.6}$$

Znów jednak sprawa racjonalnie uzasadnionej postaci estymatora gęstości λ i prawdopodobieństwa Q pozostaje otwarta.

TABLICA 2 (1)

σ	$k_1(\sigma)$	$k_2(\sigma)$	σ	$k_1(\sigma)$	$k_2(\sigma)$
0	0	1	2,20	0,3254	0,0318
0,05	0,0048	0,9218	2,40	0,3272	0,0269
0,10	0,0202	0,8468	2,60	0,3276	0,0230
0,20	0,0648	0,7063	2,80	0,3284	0,0199
0,30	0,1236	0,5787	3,00	0,3290	0,0174
0,40	0,1769	0,4657	3,20	0,3296	0,0153
0,50	0,2170	0,3715	3,40	0,3300	0,0136
0,60	0,2453	0,2971	3,60	0,3304	0,0121
0,70	0,2651	0,2398	3,80	0,3307	0,0109
0,80	0,2792	0,1959	4,00	0,3309	0,0098
0,90	0,2895	0,1622	4,20	0,3311	0,0089
1,00	0,2972	0,1360	4,40	0,3313	0,0082
1,10	0,3031	0,1153	4,60	0,3315	0,0075
1,20	0,3077	0,0989	4,80	0,3317	0,0067
1,30	0,3113	0,0856	5,00	0,3318	0,0063
1,40	0,3142	0,0748	...	...	...
1,50	0,3166	0,0658	∞	0,3331	0
1,60	0,3186	0,0583			
1,70	0,3202	0,0520			
1,80	0,3216	0,0467			
1,90	0,3228	0,0421			
2,00	0,3238	0,0382			

7. Przykład. Na koniec chcę opisać następujący eksperyment wykonany przy użyciu tablic liczb losowych. Przyjęto

$$\lambda = 50, \quad Q = 0{,}7, \quad \sigma = 0{,}5.$$

Wielkościom tym odpowiada

$$K = 0{,}2170, \quad g = 35, \quad R = 0{,}3715, \quad q = 0{,}2701.$$

Wylosowano według przybliżenia normalnego do rozkładu Poissona, że w chwili pierwszej inspekcji będzie na poletku 47 ślimaków. Wylosowano z kolei, że 35 z nich pozostanie na powierzchni, a 12 ukryje się w podłożu. Następnie 35 ślimakom ujawnionym w czasie pierwszej inspekcji wylosowano wektory ruchliwości i stwierdzono, że 23 z nich w czasie drugiej inspekcji znajdą się poza poletkiem, a tylko 12 pozostanie na poletku. Z nich wylosowano 8 jako te, które się nie ukryją. Na podstawie tych ośmiu ślimaków znaleziono

$$\hat{K} = 0{,}2109, \quad \hat{\sigma} = 0{,}4848,$$

(1) Wdzięczny jestem p. Marii Kusiakowej za wstępne obliczenia i p. Halinie Pielat za obliczenia na maszynie cyfrowej funkcyj rozważanych w tej pracy.

a następnie

$$\hat{g} = 35, \quad \hat{q} = 0{,}2286, \quad k_3(\hat{K}) = 0{,}3858.$$

Stąd ostatecznie otrzymujemy estymator gęstości ślimaków

$$\hat{\lambda} = \frac{35 \cdot 0{,}3858}{0{,}2286} = 59{,}07$$

dość bliski założonej wartości $\lambda = 50$.

Prace cytowane

[1] C. W. Helstrom, *Statistical theory of signal detection,* Oxford — London — New York — Paris 1960.

[2] A. Łomnicki, *Modyfikacja metody znakowania i powtórnego odłowu do oceny liczebności zwierząt przy uwzględnieniu ich migracji poza badaną powierzchnię,* Listy Biometryczne 16-18, Wrocław 1967, str. 51-52.

[3] S. Zubrzycki, *Wykłady z rachunku prawdopodobieństwa i statystyki matematycznej,* Warszawa 1966.

Praca wpłynęła 1. 10. 1968

С. ЗУБЖИЦКИ (Вроцлав)

СТАТИСТИЧЕСКИЙ АНАЛИЗ ОДНОЙ ЭКОЛОГИЧЕСКОЙ ЗАДАЧИ В УСЛОВИЯХ БОЛЬШОЙ ВЫБОРКИ ИЛИ ПРЕСЛЕДОВАНИЕ УЛИТОК

РЕЗЮМЕ

В работе рассматривается статистическая задача, возникающая при исследовании популяции улиток (А. Ломницки [2]). Нужно оценить среднее количество животных определенного вида на единицу поверхности среды на основании ловли, маркировки и вторичной ловли на ограниченной площадке, учитывая как подвижность так и способность животных прятаться в почве. Автор дает решение в случае большой выборки, это значит, когда выборочные средние можно принять как удовлетворительно блиские приближения средних в исходной популяции.

В конце работы дается пример оценки на основании искуственного эксперимента, составленного посредством случайных чисел.

S. ZUBRZYCKI (Wrocław)

LARGE SAMPLE ANALYSIS OF A PROBLEM IN ECOLOGY OR THE HUNTING FOR SNAILS

SUMMARY

The paper deals with a statistical problem discussed by A. Łomnicki [2] in a research concerning a population of snails. One has to estimate the average number

of animals of a given species for a unit of the area of habitat applying the well known technique of catching, marking, and catching anew. The research is carried on in a limited area and one has to take into account both the mobility of animals and their ability of taking cover in the litter. The author presents a solution in the case of a large sample, i.e. in the case where the sample means may be accepted for sufficiently precise approximations of the unknown averages in the population.

In the last section there is qiven an example of the estimation based on the data from a fictitious experiment performed by means of random numbers.