Listy Biometryczne
Vol. XXII(1985) No.1
Biometrical Letters

ESTYMACJA PROBITÓW

TADEUSZ BEDNARSKI, STANISŁAW GNOT

Instytut Matematyczny PAN, Wrocław
Instytut Immunologii i Terapii Doświadczalnej PAN, Wrocław

Praca wpłynęła 20 marca 1984; w wersji ostatecznej 5 listopada 1984

Bednarski T., Gnot S., 1985. Probit analysis. Listy Biometryczne XXII, z.1. Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu (Adam Mickiewicz University Press) pp. 3-12. PL ISSN 0458-0036.

In the paper the basic methods of one-dimensional probit analysis are presented. Also an analysis of real and simulated examples is given which enables the reader to test the applicability of the methods in a particular real situation.

1. WSTĘP

Jednym z częstych tematów statystycznej praktyki konsultacyjnej jest opracowanie danych, pochodzących z tzw. prób biologicznych. Są to eksperymenty, których celem jest określenie stopnia działania pewnych czynników na podstawie reakcji zaobserwowanych na żywych organizmach. Czynnikami tymi mogą być na przykład: witaminy, hormony, preparaty toksyczne, lekarstwa, itp. Podawane są one na różnych poziomach, określonych stężeniem lub czasem oddziaływania. Reakcje badanych organizmów na dany czynnik mogą być różne. Dla przykładu - reakcją na podanie witaminy może być zmiana ciężaru osobnika, odpowiedzią na podanie insuliny może być spadek poziomu cukru, a podanie preparatu toksycznego może spowodować śmierć osobnika. Podstawowym problemem w eksperymentach tego typu jest oszacowanie poziomu czynnika, przy którym zareaguje z góry zadany procent obiektów badanej populacji. Analiza statystyczna prowadząca do rozwiązania tego problemu nosi nazwę analizy probitów. W pracy przedstawimy podstawowe metody i modele tej analizy.

W zagadnieniu analizy probitów stosuje się w zasadzie dwie metody estymacji: metodę najmniejszych kwadratów, której pochodną jest tzw. metoda graficzna, oraz metodę największej wiarogodności. Praktyczne zastosowanie tych metod wymaga dokładnego opisu modelu, tzn. przyjęcia założeń o rozkładach prawdopodobieństwa zmiennych losowych. Najczęściej przyjmuje się model probitowy, czyli rozkład normalny lub model logitowy, wynikający z przyjęcia rozkładu logitowego. W pracy przedstawimy sposób stosowania powyższych metod w obu modelach. Obecna powszechność użycia techniki komputerowej umożliwia zastosowanie metody największej wiarogodności do bardziej złożonych modeli. Model taki wprowadzony przez Prentice'a (1976) będzie szczegółowo przedstawiony w naszej pracy. Dla modeli najczęściej stosowanych podajemy dokładną postać estymatora probitów lub algorytm, pozwalający wyznaczyć wartość estymatora. Podajemy również sposób szacowania wariancji wariancji estymatorów. Daje to możliwość konstrukcji przedziałów ufności dla probitów. Przedstawione przez nas metody i modele nie są nowe, choć naszym zdaniem niewystarczająco znane. W sposobie ich prezentacji staraliśmy się przede wszystkim uwzględnić zainteresowanie praktyka, któremu korzystanie z prac oryginalnych może sprawić trudność. Metody przedstawione w pracy ilustrowane są przykładami rzeczywistymi i symulacyjnymi. Pozwolą one czytelnikowi ocenić efektywność metod, wagę założeń modelowych oraz wpływ "wielkości eksperymentu" na dokładność otrzymanych wyników.

2. MODEL STATYSTYCZNY

Analiza probitów poprzedzona jest doświadczeniem, którego niezmienny schemat ilustruje następujący przykład (Bliss 1935). Badano w nim oddziaływanie toksyny na pewne insekty zbożowe. W tabeli poniżej zawarte są wyniki tego doświadczenia.

Logarytm dawki	1.69	1.72	1.76	1.78	1.81	1.84	1.86	1.88
Liczba insektów w doświadczeniu	59	60	62	56	63	59	62	60
Liczba zgonów	6	13	18	28	52	53	61	60

Czynnikiem jest tutaj toksyna. Dla niej ustalono najpierw wielkość dawek i dla tych dawek badano reakcję zwierząt w poszczególnych grupach. Celem analizy jest określenie dawki toksyny, która powoduje zgon z góry zadanej frakcji osobników. Przyjmuje się, że każdemu osobnikowi badanej populacji odpowiada pewna minimalna dawka d toksyny powodująca zgon. Zmienność osobnicza w populacji sprawia, że wygodnie jest traktować wartość d jako wielkość losową. Zwykle zakłada się, że rozkład zmiennej d w populacji zadany jest dystrybuantą F_θ, która należy do z góry zadanej rodziny rozkładów, indeksowanej przez nieznany parametr (wektor parametrów) θ.

Wtedy dla ustalonej dawki d_o prawdopodobieństwo tego, że losowo wybrany osobnik posiada wartość $d<d_o$ wyraża się wzorem

$$\Pr\left\{d<d_o\right\}=F_\Theta(d_o).$$

Wielkość d_o będąca kwantylem rzędu $p=F_\Theta(d_o)$ rozkładu F_Θ nazywa się w teorii analizy probitów probitem rzędu p. Powyższa definicja jest powszechnie przyjętym uogólnieniem pierwotnego pojęcia probitu wprowadzonego przez Blissa dla rozkładów normalnych. Podstawowym celem analizy probitów jest podanie metod statystycznych, pozwalających oszacować d przy zadanym p. Metody te sprowadzają się w zasadzie do problemu estymacji parametru Θ przy następującym ogólnym schemacie doświadczenia. Dla $i=1,2,\ldots,k$ niech w i-tej grupie eksperymentalnej: n_i będzie liczebnością grupy, d_i wielkością dawki, natomiast x_i liczbą reakcji w grupie. Wówczas wielkości n_i, d_i są powiązane ze zmiennymi losowymi x_i przez następujące zależności stochastyczne:

(i) zmienne losowe x_i posiadają rozkłady dwumianowe $b(p_i,n_i)$ z prawdopodobieństwami sukcesu $p_i=F_\Theta(d_i)$ i są niezależne,

(ii) frakcje grupowe x_i/n_i są "dobrym" oszacowaniem prawdopodobieństw p_i.

3. METODA NAJMNIEJSZYCH KWADRATÓW

Wybór metody statystycznej prowadzącej do estymacji probitów zależy w istotny sposób od wyboru modelu statystycznego lub mówiąc bardziej precyzyjnie od wyboru rodziny dystrybuant F_Θ. W przypadku gdy $\Theta=(\mu,\sigma)$, gdzie μ jest parametrem przesunięcia, a σ jest parametrem skali dystrybuanty F, tzn. gdy $F_\Theta(d)=F(\frac{d-\mu}{\sigma})$, przybliżone równości $x_i/n_i \approx p_i$ implikują przybliżoną równość

$$F^{-1}(x_i/n_i)\approx d_i/\sigma - \mu/\sigma\ , \quad i=1,2,\ldots,k,$$

która wiąże liniowo zmienne losowe $y_i=F^{-1}(x_i/n_i)$ z wielkościami dawek d_i i wektorem parametrów Θ. Estymację probitów można w tym przypadku przeprowadzić w sposób prosty, choć przybliżony, stosując odpowiednio przeskalowany papier. Jeżeli F jest dystrybuantą rozkładu normalnego, to przekształcenie $F^{-1}(x_i/n_i)$ nosi nazwę przekształcenia probitowego, gdy F jest dystrybuantą rozkładu logitowego, to znaczy $F(d)=e^d/(1+e^d)$, przekształcenie to nazywa się ligitowym. Metody graficzne można w sposób bezpośredni zastąpić metodami numerycznymi przyjmując dla y_i model regresji liniowej

$$y_i = ad_i + b + e_i, \tag{3.1}$$

gdzie a,b są nieznanymi parametrami, a zmienne e_i są nieskorelowanymi błędami losowymi. Rozwiązując równanie $F^{-1}(p) = \hat{a}d_p+\hat{b}$, otrzymujemy estymator probitu rzędu p

$$\hat{d}_p = [F^{-1}(p) - \hat{b}]/\hat{a}.$$

Zmienne losowe a i b są otrzymane przez minimizację sumy kwadratów $\sum_{i=1}^{k}(y_i-ad_i-b)^2$. Oszacowanie wariancji $\hat{d}_p$ można otrzymać stosując następujące rozumowanie.

Z rozważań asymptotycznych wynika, że zmienna losowa $F^{-1}(x_i/n_i)$ posiada rozkład bliski rozkładowi normalnemu

$$N\{F^{-1}(p_i),\ p_i(1-p_i)/(f^2[F^{-1}(p_i)]n_i)\},$$

gdzie f jest gęstością rozkładu F. Można zatem przyjąć, że błędy losowe występujące w modelu (3.1) mają rozkłady normalne o wartościach oczekiwanych zero i wariancjach

$$\sigma_i^2 = p_i(1-p_i)/(f^2[F^{-1}(p_i)]n_i. \tag{3.2}$$

Ponieważ nie znamy dokładnych wartości p_i, nieznane są również wariancje σ_i^2. Można je jednak oszacować podstawiając do wzoru (3.2) w miejsce p_i frakcje x_i/n_i. Korzystając z klasycznych metod szacowania wariancji zadanych funkcji zmiennych losowych (por. Green, Margerison 1978) otrzymujemy:

$$\operatorname{var}\hat{d}_p = [\hat{S}_a^2 + 2(\hat{d}_p-\bar{d})\hat{S}_{ab} + (\hat{d}_p-\bar{d})^2\hat{S}_b^2]/\hat{a}^2,$$

gdzie

$$\hat{S}_a^2 = \sum_{i=1}^{k}\hat{\sigma}_i^2/k,\quad \hat{S}_{ab} = \sum_{i=1}^{k}\hat{\sigma}_i^2(d_i-\bar{d})/\sum_{i=1}^{k}(d_i-\bar{d})^2,$$

$$\hat{S}_b^2 = \sum_{i=1}^{k}\hat{\sigma}_i^2(d_i-\bar{d})^2/\left[\sum_{i=1}^{k}(d_i-\bar{d})^2\right]^2,$$

dla

$$\bar{d} = \sum_{i=1}^{k}d_i/k,\quad \hat{\sigma}_i^2 = x_i(n_i-x_i)/(n_i^3 f^2(d_i)),$$

$$a = \sum_{i=1}^{k}y_i(d_i-\bar{d}) / \sum_{i=1}^{k}(d_i-\bar{d})^2.$$

Z praktycznych zastosowań metody regresji liniowej wynika, że daje ona zadowalające wyniki dla estymacji probitów, których rząd jest bliski 0.5. Potwierdza to również przykład symulacyjny, opisany w końcowej części. Zaletą metody regresji jest wynikająca z niej względna prostota obliczeń. Istotną wadą jest niemożność jej zastosowania w sytuacji, gdy przy pewnych dawkach wszystkie bądź żaden z badanych obiektów nie zareagował. Wówczas wielkość $F^{-1}(x_i/n_i)$ jest nieokreślona i obserwacja x_i musi być pominięta. Bazująca również na rozważaniach asymptotycznych metoda największej wiarogodności, którą przedstawimy w następnej części pracy, omija tę trudność.

ność. Obliczenia są jednak bardziej złożone, a zastosowanie tej metody wymaga użycia komputera.

4. METODA NAJWIĘKSZEJ WIAROGODNOŚCI

Dodatkową wadą metody regresji liniowej w zastosowaniu do estymacji probitów jest konieczność ograniczenia się do modelu z parametrem przesunięcia i skali, czyli do szczególnej, dwuparametrowej rodziny rozkładów. Może to mieć istotny wpływ na dokładność estymacji probitów, których rząd jest bliski jedności lub zeru. W wielu zagadnieniach biologicznych dużą wagę przykłada się do szacowania dawek dających więcej niż 90% reakcji. Badania praktyczne i teoretyczne wykazują, że stosowanie w takich sytuacjach modeli z większą liczbą parametrów zwiększa dokładność estymacji. Omówimy dalej pewien model czteroparametrowy, którego przydatność w teorii estymacji probitów uzasadnił Prentice (1976), estymując jego parametry metodą największej wiarogodności. Definiuje on następującą rodzinę rozkładów prawdopodobieństwa:

$$P_{\mu,\sigma,m_1,m_2}(d) = \int_{-\infty}^{(d-\mu)/\sigma} f_{m_1,m_2}(\omega)d\omega, \tag{4.1}$$

gdzie $f_{m_1,m_2}(\omega) = e^{\omega m_1}/[\beta(m_1,m_2)(1+e^{\omega})^{m_1+m_2}]$, natomiast $\beta(m_1,m_2)$ jest znaną funkcją beta. W tym modelu μ i σ są parametrami przesunięcia i skali, natomiast m_1 i m_2 są parametrami mającymi wpływ na wielkość "ogonów" i symetrię rozkładów. Model ten jako szczególne przypadki zawiera niemal wszystkie stosowane dotąd w analizie probitów rozkłady. Ilustruje to następująca tabela.

Parametry	Typ rozkładu
$m_1=m_2=1$	logitowy
$m_1,m_2 \to \infty$	normalny
$m_1=1, m_2\to\infty$; $m_1\to\infty, m_2=1$	rozkład statystyk ekstremalnych
$m_1,m_2 \to 0$	podwójny wykładniczy
$m_1\neq 0,\ m_2 \to 0$	wykładniczy
$m_1 \to 0,\ m_2\neq 0$	odbity wykładniczy (reflected exponential)
$m_1=m_2$	symetryczny
$m_1 < m_2$	ujemnie skośny (negatively skewed)
$m_1 > m_2$	dodatnio skośny (positively skewed)

Logarytm prawdopodobieństwa zaobserwowania $x_1,x_2,\ldots,x_k$ przy dawkach $d_1,d_2,\ldots,d_k$ i liczebnościach grupowych $n_1,n_2,\ldots,n_k$ jest, z dokładnością do składnika niezależnego od parametru, równy

$$L(\theta) = \sum_{i=1}^{k}\left[x_i \log P(d_i) + (n_i-x_i)\log(1-P(d_i))\right],$$

gdzie

$$P(d_i) = \int_{-\infty}^{(d_i-\mu)/\sigma} f_{m_1,m_2}(\omega)d\omega.$$

Wyznaczenie estymatora probitu poprzedza estymacja wektora $\theta = (\theta_1,\theta_2,\theta_3,\theta_4)$ dla $\theta_1=\mu$, $\theta_2=\sigma$, $\theta_3=m_1$ i $\theta_4=m_2$, metodą największej wiarogodności. Przypomnijmy, że $\hat{\theta}$ jest estymatorem największej wiarogodności, jeżeli $L(\hat{\theta}) = \max_{\theta} L(\theta)$. Aby znaleźć wartość tego estymatora wygodnie jest posłużyć się metodą gradientów lub metodą Rapsona-Newtona (patrz Dyer i McReynolds 1970). Przykładowo podajemy tutaj prostszy algorytm, wykorzystujący metodę gradientów. Niech $L_i'(\theta) = \partial L(\theta)/\partial\theta_i$ oraz $L'(\theta) = (L_1'(\theta),\ldots,L_4'(\theta))$:

(i) wybrać wstępne oszacowanie θ^j prawdziwej wartości parametru θ i obliczyć $L(\theta^j)$,

(ii) dla pewnego $\varepsilon = (\varepsilon_1,\ldots,\varepsilon_4)$ obliczyć $\theta^{j+1} = \theta^j + \varepsilon L'(\theta^j)$ oraz $L(\theta^{j+1})$, jeśli $L(\theta^{j+1}) > L(\theta^j)$, to za j wstawiamy $j+1$ i wracamy do (ii), w przeciwnym przypadku za ε wstawiamy ε/β, gdzie β jest stałą większą od jedności (zmniejszamy krok) i wracamy do (ii).

Cykl ten powtarzamy do momentu otrzymania zadowalającej dokładności. Pochodne cząstkowe logarytmu ilorazu wiarogodności mają dla modelu (4.1) następującą postać:

$$\partial L(\theta)/\partial\theta_j = \sum_{i=1}^{k}\left[\frac{x_i}{P(d_i)} - \frac{n_i-x_i}{1-P(d_i)}\right](\partial P(d_i)/\partial\theta_j), \quad j=1,2,3,4,$$

gdzie

$$\partial P(d_i)/\partial\mu = -f(z_i)/\sigma, \quad \partial P(d_i)/\partial\sigma = -z_i f(z_i)/\sigma, \quad z_i(d_i-\mu)/\sigma,$$

oraz

$$\partial P(d_i)/\partial m_l = \int_{-\infty}^{z_i}\left[\partial \log f_{m_1,m_2}(\omega)/\partial m_l\right] f_{m_1,m_2}(\omega)\, d\omega, \quad l=1,2.$$

Wiadome jest, że w sensie asymptotycznej wariancji estymatory największej wiarogodności są najbardziej efektywne spośród estymatorów asymptotycznie normalnych. Dla określenia asymptotycznej wariancji estymatora θ, a dalej estymatora probitu, należy obliczyć tzw. macierz informacji Σ, której elementami są drugie pochodne cząstkowe logarytmu ilorazu wiarogodności, tzn. element (j,h) macierzy Σ jest równy

$$\sum_{i=1}^{k} \frac{n_i}{P(d_i)(1-P(d_i))}(\partial P(d_i)/\partial\theta_j)(\partial P(d_i)/\partial\theta_h).$$

Podstawowe twierdzenie z teorii estymacji metodą największej wiarogodności mówi, że asymptotycznym rozkładem wektora $\sqrt{n}(\hat{\Theta}-\Theta)$ jest wielowymiarowy rozkład normalny o zerowym wektorze wartości oczekiwanej i macierzy kowariancji Σ^{-1}, $(n=n_1+n_2+\ldots+n_k)$.

Analizując dane Blissa, Prentice stosuje trzyparametrowy model (4.1) z $m_2=1$. W tym szczególnym przypadku

$$P(d) = e^{zm_1}(1+e^{z})^{-m_1},$$

gdzie $z=(d-\mu)/\sigma$. Estymator probitu rzędu p wyraża się wtedy wzorem

$$\hat{d}_p=\hat{\mu} - \hat{\sigma}\log(p^{-1/\hat{m}_1}-1),$$

gdzie $\hat{\mu}$, $\hat{\sigma}$ i $\hat{m}_1$ są składowymi wektora będącego estymatorem największej wiarogodności w modelu trzyparametrowym. Oszacowaniem wariancji estymatora $\hat{d}_p$ jest $c'\Sigma^{-1}c$ dla

$$c' = (1,\ -\log(p^{-1/\hat{m}_1}-1),\ -\hat{\sigma}\hat{m}_1^{-2}\log[p/(1-p^{1/\hat{m}_1})]).$$

W tabeli poniżej podane są liczebności x_i z przykładu Blissa oraz liczebności teoretyczne, wynikające z zastosowania metody największej wiarogodności i przyjęcia kolejno modelu normalnego, logitowego i modelu (4.1) z $m_2=1$. W ostatnich dwóch wierszach tabeli podajemy liczebności teoretyczne, wynikające z zastosowania metody regresji. W ostatnich kolumnach zamieszczone są wartości statystyki χ^2 i odpowiadające im wartości P. Wielkości te są obliczone w celu zbadania zgodności danych z kolejnymi modelami teoretycznymi. Wartość statystyki χ^2 obliczona jest ze wzoru

$$\chi^2 = \sum_{i=1}^{k} (x_i-x_i')^2/x_i,$$

gdzie x_i' są liczebnościami teoretycznymi. Liczba stopni swobody χ^2 jest równa $k-r$, gdzie r jest liczbą estymowanych parametrów.

	Liczebności x_i z przykładu Blissa									
	6	13	18	28	52	53	61	60		
Model	liczebności teoretyczne x_i' metoda największej wiarogodności									
Normalny	3.3	10.9	23.6	33.9	49.6	53.3	59.6	59.2	5.22	.50
	$\hat{\mu}$ = 1.77, $\hat{\sigma}$ = .051									
Logitowy	3.4	9.8	22.5	33.9	50.1	53.3	59.2	58.7	4.95	.55
	$\hat{\mu}$ = 1.77, $\hat{\sigma}$ = .029									
(4.1), m_2=1	6.1	11.2	20.1	29.7	48.6	54.8	60.9	59.8	.90	.98
	$\hat{\mu}$ = 1.82, $\hat{\sigma}$ = .016, $\hat{m}_1$ = .279									

c.d. tabeli

	metoda regresji liniowej									
Normalny	3.9	12.1	25.2	35.7	50.9	54.2	60.1	-	4.85	.45
			$\hat{\mu}$ = 1.77,		$\hat{\sigma}$ = .051					
Logitowy	3.8	11.1	25.0	36.5	52.3	54.5	59.9	-	5.55	.35

5. NIEPARAMETRYCZNA METODA ESTYMACJI LD_{50}

W zagadnieniach medycznych często jesteśmy zainteresowani estymacją tzw. medialnej dawki śmiertelnej - LD_{50} (lethal dose 50%), tj. probitu rzędu 0.5. Możemy wówczas skorzystać z metody estymacji, której jedynym ograniczeniem jest założenie o symetrii rozkładu prawdopodobieństwa dawki d w populacji. Metoda ta nosi nazwę metody Barentsa-Karbera, a jej szczegółowy opis można znaleźć w książce Perkala (1958). Korzystając z wcześniej wprowadzonych oznaczeń podajemy wzór na estymator LD_{50}:

$$LD_{50} = (1/2)\left[d_k+d_{k-1} - \sum_{i=1}^{k-1}(d_{i+1}-d_{i-1})x_i/n_i\right].$$

Można wykazać, że jest to estymator zgodny o rozkładzie asymptotycznie normalnym. Wariancję estymatora otrzymujemy z następującego wzoru:

$$\mathrm{Var}\ LD_{50} = (1/4)\sum_{i=1}^{k-1}\left[(d_{i+1}-d_{i-1})^2 x_i(n_i-x_i)/(n_i^2(n_i-1))\right].$$

6. PRZYKŁAD SYMULACYJNY

Opisane w pracy metody estymacji probitów bazują na faktach z teorii metod asymptotycznych. W praktyce oznacza to, że wiarogodne zastosowanie tych metod wymaga dużej liczby badanych obiektów. Często przy estymacji probitów liczba danych w grupach eksperymentalnych nie przekracza 20. Mając na względzie użyteczność przedstawionych metod podajemy niewielki przykład symulacyjny, który pozwoli czytelnikowi skonfrontować teorię z praktyką analizy danych prób biologicznych. W przykładzie przyjęliśmy, że rozkład zmiennej d jest normalny N(5,1), tj. μ=5, σ=1. Dla każdego z poziomów dawki d_1=2, d_2=3,...,d_7=8 generowano 20 liczb losowych o rozkładzie zero-jedynkowym z prawdopodobieństwem sukcesu $P(d_i)= \Phi^{-1}(d_i-5)$, gdzie Φ jest dystrybuantą standardowego rozkładu normalnego. Tak otrzymane dane odpowiadają doświadczeniu probitowemu, gdy rozkład d jest N(5,1), a eksperymentowi podlega każdorazowo 20 obiektów przy ustalonym poziomie dawki równym odpowiednio 2,3...,8. Eksperyment losowy powtórzono 30-krotnie, estymując metodą regresji i metodą największej wiarogodności probity rzędu 0.5 i 0.1. Probit rzędu 0.5 oszacowano także metodą nieparametryczną. Wyniki symulacji przedstawia tabela poniżej, w której wprowadzono następujące oznaczenia: d_p - teoretyczna wartość probitu, $\bar{d}_p$ - średnia arytmetyczna 30 wartości estymatorów $\hat{d}_p$, s_1 - odchylenie standardowe dla 30 wartości estymatorów $\hat{d}_p$, s_2 - odchylenie standardowe estymatora probitu $\hat{d}_p$, wyliczone na podstawie wzorów przedstawionych w pracy, przy dokładnych wartościach p_i.

	Metody				
	nieparametryczna	regresji liniowej		największej wiarogodności	
p	0.5	0.5	0.1	0.5	0.1
d_p	5	5	3.72	5	3.72
$\bar{d}_p$	5.02	5.02	3.60	5.02	3.76
s_1	0.26	0.23	0.36	0.15	0.27
s_2	0.36	0.19	0.30	0.18	0.29

7. UWAGI KOŃCOWE

W rozdziale tym przedstawiamy uwagi i komentarze natury praktycznej, które mogą czytelnikowi ułatwić wybór metody estymacji probitu oraz pomogą zaplanować eksperyment, umożliwiający wiarogodne opracowanie wyników.

Metodą, która najdokładniej uwzględnia strukturę probabilistyczną modeli parametrycznych opisanych w pracy, jest metoda największej wiarogodności. Daje ona także estymatory probitów o asymptotycznie minimalnej wariancji. Przykład Blissa oraz przykład symulacyjny potwierdzają to również w sytuacji, gdy liczba danych mieści się w zakresie nawet skromnych eksperymentów probitowych. Istotną zaletą tej metody jest możliwość jej zastosowania do różnych parametrycznych rodzin rozkładów dla estymacji probitów, praktycznie dowolnego rzędu. Przy jej zastosowaniu niezbędne jest jednak użycie komputera (mini-komputera). W sytuacji gdy dysponujemy jedynie kalkulatorem, możemy posłużyć się metodą regresji. Przykłady podane w pracy wskazują na względnie dużą efektywność tej metody. Należy jednak pamiętać, że metoda regresji znajduje zastosowanie jedynie dla modeli z parametrami przesunięcia i skali, a w zasadzie stosuje się ją do modelu normalnego i logitowego. Metoda ta nie powinna być stosowana dla estymacji probitu rzędu wyższego niż 0.8, a niższego niż 0.2. Estymatory otrzymane metodą regresji można wykorzystać jako wartości początkowe w algorytmie obliczającym wartość estymatorów największej wiarogodności. Wadą metody regresji jest to, że niejednokrotnie zmuszeni jesteśmy pominąć informację pochodzącą z grup eksperymentów, w których reagują wszystkie obiekty lub nie reaguje żaden. Zaprezentowaliśmy w pracy również metodę nieparametryczną. Uczyniliśmy to jednak głównie ze względu na to, że jest ona obliczeniowo prosta i tradycyjnie stosowana. Należy pamiętać o tym, że dokładność tej metody jest niewielka i zależy w dużej mierze od zaplanowania eksperymentu.

Chcielibyśmy zwrócić uwagę na to, że ten fragment teorii analizy probitu, który przedstawiliśmy, dotyczy estymacji, w której badamy wpływ jednego czynnika. Istnieje już bogata literatura na temat wieloczynnikowej analizy probitów. Zainteresowanych czytelników odsyłamy do książki Finneya (1952) oraz do pracy Hasselblada i in. (1980).

LITERATURA

Bliss, C.I., (1935). The calculation of the dosage-mortality curve. Annals of Applied Biology, 22, 134-167.

Finney, D.J., (1952). Probit analysis, a statistical treatment of the sigmoid response curve. Wyd. 2. Cambridge University Press.

Green, J.R., Margerison, D.(1978). Statistical treatment of statistical data. Elsevier Scientific Publishing Company, Amsterdam-Oxford-New York.

Hasselblad, V., Stead, A.G., Creason, J.P. (1980). Multiple probit analysis with a nonzero background. Biometrics 36, 659-664.

Dyer, P., McReynolds, S.R. (1970). The computation and theory of optimal control. Academic Press New York and London.

Perkal, J. (1958). Matematyka dla rolników. PWN, Warszawa.

Prentice, R.L., (1976). A generalization of the probit methods for dose response curves. Biometrics 32, 761-768.