

Jan CHOJCAN, Lucjan KARWAN, Jacek IZYDORCZYK

Instytut Elektroniki  
Politechniki Śląskiej

## ZASTOSOWANIE STATYSTYK POZYCYJNYCH DO STATYSTYCZNEJ ANALIZY UKŁADÓW ELEKTRONICZNYCH

**Streszczenie.** W pracy podano podstawowe definicje i własności statystyk pozycyjnych oraz przedstawiono możliwości ich wykorzystania w statystycznej analizie układów w celu zwiększenia efektywności metody Monte Carlo. Rozważania ilustrowane są przykładami.

## QUANTILE ARITHMETIC APPROACH TO THE STATISTICAL ANALYSIS OF NETWORKS

**Summary.** In the paper definitions and properties of the quantile arithmetic are developed. Application of the method to the statistical analysis of networks is considered. The method provide to efficiency improvement of Monte Carlo method. The article includes a lot of examples.

### 1. Wstęp

Dokonując analizy statystycznej układów często zachodzi potrzeba oceny uzysku produkcyjnego lub wartości parametrów pozycyjnych rozkładu funkcji układowej  $T$ , czyli określenia kwantyli rzędu  $p$ ; ( $0 < p < 1$ ) dla różnych wartości  $p$ . Analizę tę przeprowadza się najczęściej za pomocą metody Monte Carlo (MC) lub jej modyfikacji [2,3].

W pracy podjęto próbę wykorzystania uporządkowanych, w kolejności niemalejącej, n-ciągów wyników uzyskanych z metody MC do oszacowania przy zadanej wielkości dopuszczalnego błędu — długości n-ciągów oraz oceny efektywności takiego postępowania.

Dział rachunku prawdopodobieństwa i statystyki matematycznej poświęcony analizie rozkładów wektora losowego kwantyli z próby nazywany jest rozkładami statystyk pozycyjnych [1].

## 2. Statystyki pozycyjne [1]

W próbach o licznosci  $n$  elementów będą rozpatrywane zmienne losowe, które są odpowiednikami parametrów pozycyjnych. Takie zmienne losowe będą nazywane **statystykami pozycyjnymi**.

Niech  $(X_1, X_2, \dots, X_n)$  będzie  $n$ -wymiarowym wektorem losowym. Rozpatrzmy zmienną losową  $z_k^{(n)}$  określoną w następujący sposób: porządkujemy każdy zespół wartości  $(x_1, x_2, \dots, x_n)$  wektora losowego  $(X_1, X_2, \dots, X_n)$  od najmniejszego do największego, tj. tak, by uzyskać ciąg  $(x_{r1}, x_{r2}, \dots, x_{rm})$  spełniający nierówności  $x_{r1} \leq \dots \leq x_{rm}$ .

**Statystyką pozycyjną**  $z_k^{(n)}$  nazywamy zmienną losową, będącą funkcją wektora losowego  $(X_1, X_2, \dots, X_n)$ , która przybiera  $k$ -tą co do wielkości wartość, tj. wielkość  $x_{rk}$  każdego z możliwych zespołów wartości  $(x_1, x_2, \dots, x_n)$ . Liczbę  $k$  nazywamy rangą statystyki pozycyjnej  $z_k^{(n)}$ , a stosunek

$\frac{k}{n}$  - **rangą względną**.

Jeśli  $\lim_{n \rightarrow \infty} \frac{k}{n} = \lambda$ ; ( $0 < \lambda < 1$ ), to ciąg  $z_k^{(n)}$  nazywamy **ciągami centralnym** statystyk.

Gdy zachodzi  $k = [n \cdot \lambda] + 1$ , gdzie  $0 < \lambda < 1$ , wtedy statystykę pozycyjną  $z_k^{(n)}$  nazywamy **kwantylem z próby**. Ciąg kwantyli z próby jest ciągiem centralnym statystyk.

Funkcje statystyk pozycyjnych  $z_k^{(n)}$  będziemy także nazywali statystykami pozycyjnymi.

Funkcja zmiennej  $x$ ; ( $-\infty < x < \infty$ ) równa zero dla  $x \leq x_{r1}$  i równa  $\frac{m}{n}$ ; ( $m = 1, 2, \dots, n$ ) dla  $x > x_{r1}$ , gdzie  $m$  jest największym wskaźnikiem, dla którego zachodzi nierówność  $x_{rm} < x$ , nosi nazwę **dystrybuanty empirycznej**  $S_n(x)$ .  $S_n(x)$  jest dla każdej wartości argumentów  $x$  zmienną losową.

Jeżeli zmienne losowe  $X_r$ ; ( $r = 1, 2, \dots, n$ ) są niezależne i mają jednakową dystrybuantę ciągłą  $F(x)$ , a  $x_{r1}, x_{r2}, \dots, x_{rm}$  oznacza uporządkowany według wielkości zespół wyników obserwacji zmien-

nych losowych  $X_1, X_2, \dots, X_n$ , wówczas  $x_{r1} \leq x_{r2} \leq \dots \leq x_{rn}$ , a dystrybucję empiryczną  $S_n(x)$  można przedstawić w następującej postaci:

$$S_n(x) = \begin{cases} 0 & \text{gdyn } x \leq x_{r1} \\ \frac{k}{n} & \text{gdyn } x_{rk} < x < x_{r(k+1)} \\ 1 & \text{gdyn } x > x_{rn} \end{cases} \quad (k = 1, 2, \dots, n-1) \quad (1)$$

Z przyjętych założeń wynika, że dla każdego ustalonego  $x$  zachodzi relacja  $P(X_r < x) = F(x) = p = \text{const}$  ( $k = 1, 2, \dots, n$ ). Tak więc  $S_n(x)$  jest dla stałej wartości  $x$  częstością w schemacie Bernoulliego, jest więc częstością realizacji zdarzenia losowego w  $n$  niezależnych doświadczeniach, gdy prawdopodobieństwo realizacji zdarzenia w poszczególnym doświadczeniu wynosi  $p = F(x)$ . Stąd otrzymujemy:

$$P \left[ S_n(x) = \frac{m}{n} \right] = \frac{n!}{m!(n-m)!} [F(x)]^m \cdot [1-F(x)]^{n-m} \quad (2)$$

Oznaczamy dystrybucję zmiennej losowej  $z_k^{(n)}$  przez  $F_{kn}(x)$ , wówczas

$F_{kn}(x) = P(z_k^{(n)} < x)$ . Zauważmy, że zdarzenie  $z_k^{(n)}$  oznacza, iż w zespole  $n$  obserwacji  $k$ -ta co do wielkości obserwacja jest mniejsza od  $x$ , a więc jest co najmniej  $k$  obserwacji mniejszych od  $x$ . Otrzymujemy stąd:

$$\begin{aligned} F_{nk}(x) &= P(z_k^{(n)} < x) = P(S_n(x) \geq \frac{k}{n}) = \sum_{m=k}^n P(S_n(x) = \frac{m}{n}) = \sum_{m=k}^n \binom{n}{m} F(x)^m \cdot [1-F(x)]^{n-m} = \\ &= \frac{n!}{(k-1)!(n-k)!} \int_0^{F(x)} t^{k-1} (1-t)^{n-k} dt \end{aligned} \quad (3)$$

Ponieważ założyliśmy istnienie gęstości  $f(x) = \frac{dF(x)}{dx}$  zmiennej  $X$ , przeto istnieje również

gęstość  $f_{kn}(x)$  zmiennej losowej  $z_k^{(n)}$  w postaci:

$$f_{nk}(x) = \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} [1-F(x)]^{n-k} \cdot f(x) \quad (4)$$

**Przykład 1**

Należy określić rozkład statystyki  $z$  dla prostych prób liczących  $n$  elementów wybranych z populacji o rozkładzie prostokątnym, określonym gęstością:

$$f(x) = \begin{cases} \text{lgdy } 0 \leq x \leq 1 \\ 0 \text{ gdy } x < 0; \vee x > 1 \end{cases}$$

Ze wzoru (4) gęstość  $f(x)$  jest równa:

$$f_{nk}(x) = \begin{cases} \frac{n!}{(k-1)!(n-k)!} \left[ \int_0^x t^{k-1} dt \right] \cdot \left[ \int_x^1 dt \right]^{n-k} = \frac{n!}{(k-1)!(n-k)!} \cdot x^k \cdot (1-x)^{n-k} & \text{gdy } 0 \leq x \leq 1 \\ 0 & \text{gdy } x < 0; \vee x > 1 \end{cases}$$

Niech np.:  $n = 5$ ,  $k = 1, 2, 3$ , wówczas:  $f_{1,5}(x) = 5 \cdot (1-x)^4$ ,  $f_{2,5}(x) = 20 \cdot x \cdot (1-x)^3$ ,

$f_{3,5}(x) = 30 \cdot x^2 \cdot (1-x)^2$ . Prawdopodobieństwo, że wartość zmiennej  $z \binom{5}{}$  będzie zawarta między  $a$  i  $b$ , wynosi:

$$P\left(a \leq z \binom{5}{1} \leq b\right) = \int_a^b f_{1,5}(x) dx = \left[ -1 \cdot (1-x)^5 \right]_a^b$$

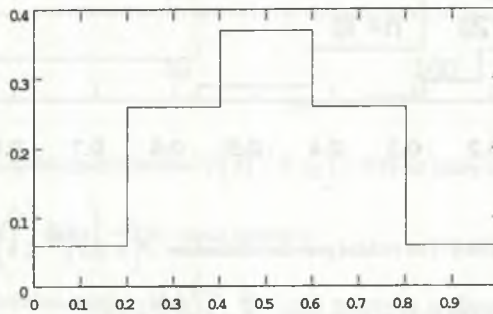
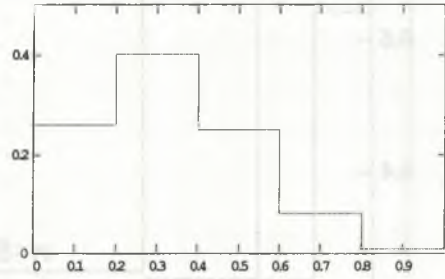
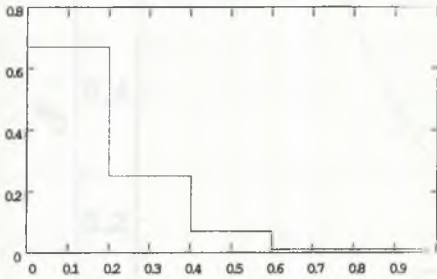
analogicznie  $P\left(a \leq z \binom{5}{2} \leq b\right) = \int_a^b f_{2,5}(x) dx = \left[ -1 \cdot (1-x)^5 - 5 \cdot x \cdot (1-x)^4 \right]_a^b$  oraz

$$P\left(a \leq z \binom{5}{3} \leq b\right) = \int_a^b f_{3,5}(x) dx = \left[ -1 \cdot (1-x)^5 - 5 \cdot x \cdot (1-x)^4 - 10 \cdot x^2 \cdot (1-x)^3 \right]_a^b$$

Na rysunku 1a,b,c przedstawiono wykresy  $P\left(a \leq z \binom{5}{k} \leq b\right)$ , gdy  $a = 1 \times 0.2$ ,  $b = (1+1) \times 0.2$  ( $1=0,1,2,3$ ) dla  $k=1$  (rys.1a),  $k=2$  (rys.1b) oraz  $k=3$  (rys.1c).

Wpływ wielkości  $n$  (dla  $k=1$ ) na rozkład prawdopodobieństw  $P\left(a \leq z \binom{5}{1} \leq b\right)$  można prześledzić na rys. 2. Zauważmy, jeśli dla  $n = 5$  zmienia losowa  $z \binom{5}{1}$  przyjmuje z prawdopodobieństwem 0.99 wartości nie większe od  $x=0.6$ , to dla  $n=20$  zmienna  $z \binom{20}{1}$  przyjmuje z tym samym, prawdopodobieństwem wartości nie większe od  $x = 0.2$ .





Rys.1.a. Wykres prawdopodobieństwa  $P\left(a \leq z_1^{(5)} \leq b\right)$  dla  $a=0.2x1, b=0.2x(1+1), l=0,1,2,3$

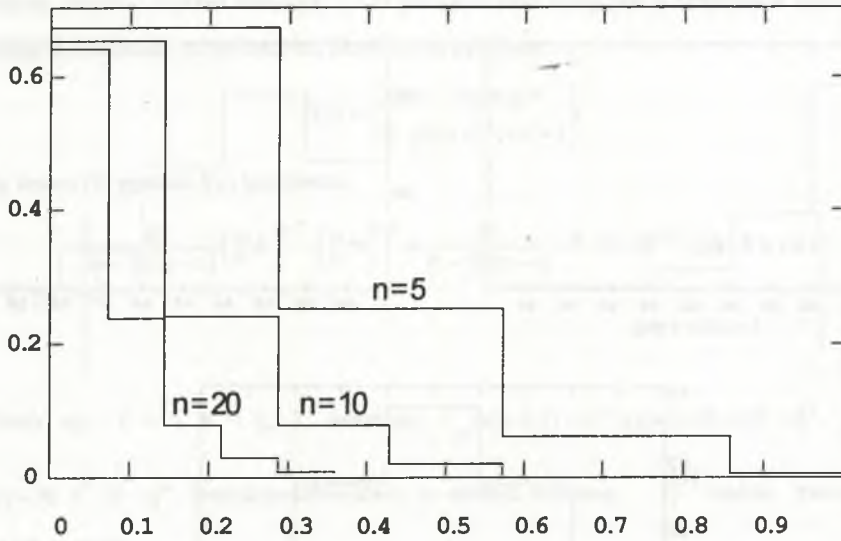
Rys.1. b. Wykres prawdopodobieństwa  $P\left(a \leq z_2^{(5)} \leq b\right)$  dla  $a=0.2x1, b=0.2x(1+1), l=0,1,2,3$

Rys.1.c. Wykres prawdopodobieństwa  $P\left(a \leq z_3^{(5)} \leq b\right)$  dla  $a=0.2x1, b=0.2x(1+1), l=0,1,2,3$

Fig.1.a. Probability  $P\left(a \leq z_1^{(5)} \leq b\right)$  for  $a=0.2x1, b=0.2x(1+1), l=0,1,2,3$

Fig.1.b. Probability  $P\left(a \leq z_2^{(5)} \leq b\right)$  for  $a=0.2x1, b=0.2x(1+1), l=0,1,2,3$

Fig.1.c. Probability  $P\left(a \leq z_3^{(5)} \leq b\right)$  for  $a=0.2x1, b=0.2x(1+1), l=0,1,2,3$



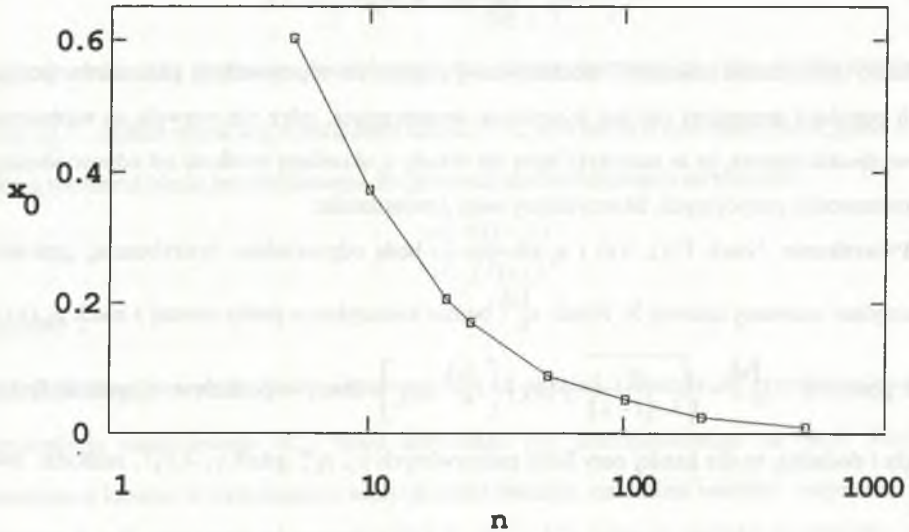
Rys.2. Wpływ wielkości n (dla  $k=1$ ) na rozkład prawdopodobieństw  $P(a \leq z_3^{(n)} \leq b)$

Fig.2. Influence of number n (for  $k=1$ ) on probability  $P(a \leq z_3^{(n)} \leq b)$

W tabeli 1 (i na rys.3) przedstawiono wartości  $x_0$  dla różnych wartości n (dla  $k=1$  oraz  $P(z^{(n)} \leq x_0) \approx 0.99$ ).

Tabela 1

n=	5	10	20	25	50	100	200	500
$x_0$	0.602	0.369	0.206	0.168	0.088	0.05	0.025	0.01



Rys.3. Zależność wartości  $x_0$  spełniającej równanie  $P\left(z_k^{(n)} \leq x_0\right) = 0.99$  od liczby  $n$

Fig.3. Root of equation  $P\left(z_k^{(n)} \leq x_0\right) = 0.99$  versus number  $n$

Wynika z nich, że rozkład statystyki  $z_k^{(n)}$  dla prób prostych  $n$ -elementowych, wybranych z populacji o rozkładzie prostokątnym, jest ze wzrostem  $n$  coraz bardziej skupiony. Czy jest to ogólna prawidłowość?

Aby odpowiedzieć na to pytanie rozpatrzmy ciąg kwantyli z próby  $\left\{z_k^{(n)}\right\}$ , a więc  $k = [n \cdot \lambda] + 1$ , gdzie  $0 < \lambda < 1$ . Dla każdej wartości  $\lambda$  z przedziału  $0 < \lambda < 1$  istnieje co najmniej jedna taka wartość  $a_\lambda$ , że:

$$1 - P(X = a_\lambda) \leq F(a_\lambda) \leq \lambda \tag{5}$$

Wartość  $a_\lambda$  jest kwantylem rozkładu zmiennej losowej  $X$ . Można udowodnić [1], że jeżeli relację (5) spełnia tylko jedno  $a_\lambda$ , to ciąg kwantyli z próby  $\left\{z_k^{(n)}\right\}$  jest stochastycznie zbieżny do odpowiadającego mu kwantyla  $a_\lambda$  z populacji generalnej, czyli:

$$\lim_{c>0} \lim_{n \rightarrow \infty} P\left(z_k^{(n)} - a_\lambda \geq c\right) = 0 \quad (6)$$

Samo stwierdzenie zbieżności stochastycznej ciągów do odpowiednich parametrów pozycyjnych populacji generalnej nie jest w praktyce wystarczające, gdyż nie pozwala na wyznaczenie prawdopodobieństwa, że te statystyki będą się różniły o określoną wielkość od odpowiadających im parametrów pozycyjnych. Skorzystajmy więc z twierdzenia:

**Twierdzenie:** Niech  $F(x)$ ,  $f(x)$  i  $a_\lambda$  ( $0 < \lambda < 1$ ) będą odpowiednio dystrybuantą, gęstością i kwantylem zmiennej losowej  $X$ . Niech  $z_k^{(n)}$  będzie kwantylem z próby prostej i niech  $g_{kn}(y)$  będzie gęstością  $Y_k^{(n)} = \sqrt{\frac{n}{\lambda \cdot (1-\lambda)}} \cdot f(a_\lambda) \cdot [z_k^{(n)} - a_\lambda]$ . Jeżeli w punkcie  $x=a_\lambda$  gęstość  $f(x)$  jest ciągła i dodatnia, to dla każdej pary liczb rzeczywistych  $y_1, y_2$ , gdzie  $y_1 < y_2$ , zachodzi związek:

$$\lim_{n \rightarrow \infty} P\left(y_1 < Y_k^{(n)} < y_2\right) = \lim_{n \rightarrow \infty} \int_{y_1}^{y_2} g_{kn}(y) dy = \frac{1}{\sqrt{2\pi}} \int_{y_1}^{y_2} \exp\left(-\frac{y^2}{2}\right) dy \quad (7)$$

Z twierdzenia tego wynika, że jeżeli w rozważanej populacji generalnej zmienna losowa  $X$  jest typu ciągłego o gęstości  $f(x)$  i jeżeli ta gęstość jest ciągła i dodatnia w punkcie  $a_\lambda$ , gdzie  $a_\lambda$  jest kwantylem populacji generalnej, to kwantyl z prostej  $z_k^{(n)}$  ma rozkład asymptotycznie normalny:

$$N(m_I \sigma_I) = N\left(a_\lambda \cdot \frac{1}{f(a_\lambda)} \cdot \sqrt{\frac{\lambda \cdot (1-\lambda)}{n}}\right) \quad (8)$$

Z zależności (8) wynika, że kwantyl  $z_k^{(n)}$  mieści się z poziomem ufności  $1-\alpha$   $0 < \alpha < 1$  w przedziale ufności  $[c,d]$  określonym wzorem:

$$P\left(c < z_k^{(n)} < d\right) \approx P\left(a_\lambda - a \cdot \sigma_I < z_k^{(n)} < a_\lambda + a \cdot \sigma_I\right) = 1 - \alpha = \gamma \quad (9)$$

gdzie:  $a_\lambda = a \cdot (1 - \lambda)$ , natomiast błąd bezwzględny



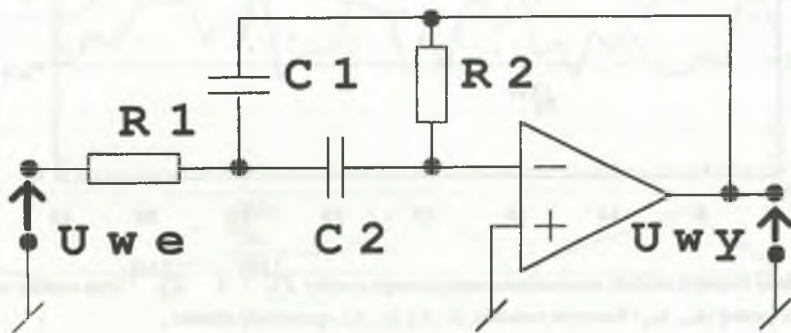
$$e = \frac{d-c}{2} = a \cdot \sigma_I = \frac{a}{f(a_\lambda)} \cdot \sqrt{\frac{\lambda \cdot (1-\lambda)}{n}} \quad (10)$$

Z zależności (10) można wyznaczyć niezbędną liczbę obserwacji  $n$ , przy której kwantyl  $z_k^{(n)}$  próby  $z_k^{(n)}$  będzie różnił się, z poziomem ufności  $1-\lambda$ , od kwantyla  $a_\lambda$  o mniej niż  $e$ , gdzie  $e$  jest żadaną wartością błędu bezwzględnego. Po prostych przekształceniach otrzymamy:

$$n = \frac{a^2 \cdot \lambda(1-\lambda)}{e^2 \cdot f^2(a_\lambda)} \quad (11)$$

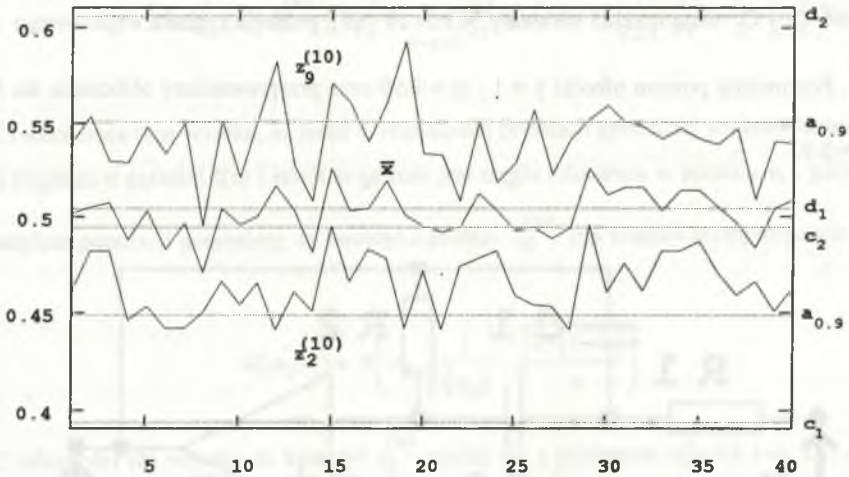
### Przykład 2

Należy określić wpływ liczby obserwacji  $n$  na rozkład kwantyla  $z_k^{(n)}$  próby z modułu wzmocnienia napięciowego  $|K_u|$  filtru aktywnego [4], przedstawionego na rys.4. Przyjmając: wzmacniacz idealny o wzmocnieniu nieskończenie wielkim, natomiast wartości rezystancji rezystorów  $R_1$  i  $R_2$  mają rozkład normalny  $N(1.5 \cdot 10^2)$  [ $\Omega$ ], również wartości pojemności kondensatorów  $C_1$  i  $C_2$  mają rozkład normalny  $N(1.5 \cdot 10^2)$  [F], pulsacja sygnału wejściowego  $\omega = 1 \left[ \frac{\text{rad}}{\text{s}} \right]$ . Przyjmijmy poziom ufności  $\gamma = 1 - \alpha = 0.99$  oraz przeprowadźmy obliczenia dla  $\lambda_1 = 0.1$  i  $\lambda_2 = 0.9$ .



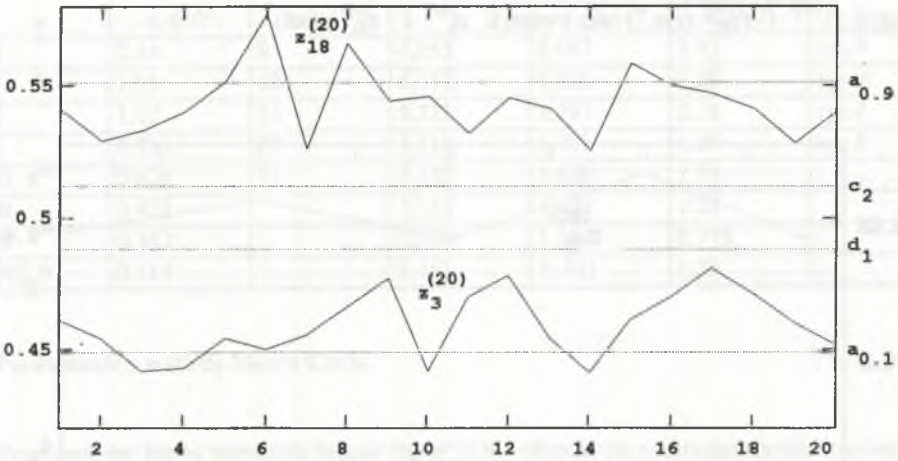
Rys.4. Filtr aktywny  
Fig. 4. An active filter

Wygenerowano metodą Monte Carlo  $N=1000$  wyników i otrzymano  $|K_u|=0.50056$  oraz  $\sigma = 3.9304 \cdot 10^{-2}$ . Wyniki te różnią się nieznacznie od wartości uzyskanych metodą analityczną, ponieważ  $|K_u| = 0.5$ , natomiast  $\sigma = 3.95 \cdot 10^{-2}$ . Łatwo sprawdzić w tablicach rozkładu normalnego, że  $a_{0,9} = |K_u| + z_{0,9} \cdot \sigma = 0.5506$ , natomiast  $a_{0,1} = |K_u| - z_{0,1} \cdot \sigma = 0.4494$ . Wartości odchyłeń standardowych  $\sigma_1$ , przedziałów ufności i błędów bezwzględnych (dla poziomu ufności  $1-\alpha=0.99$ ,  $a=2.576$ ) oraz  $k$  dla różnych wartości  $n$  przedstawiono w tabeli 2. Wielkość  $k$  obliczono z zależności  $k=[n \cdot \lambda] + 1$ . Zauważmy, że dla małych  $n$  zależność ta jest dokładna. Dla  $n=10$  i  $\lambda=0.1$   $k=2$ , a dla  $\lambda=0.9$   $n-k+1=9$ . Rozkład kwantyli z próby  $z_2^{(10)}$  i  $z_9^{(10)}$  modułu wzmocnienia przedstawiono na rysunku 5. Na rysunku tym przedstawiono też rozkład wartości średniej arytmetycznej  $\bar{x}$  dla  $t=40$  realizacji obserwacji  $n=10$ -elementowych. Zaznaczono również granice przedziałów ufności i wartości (graniczne) błędów bezwzględnych.

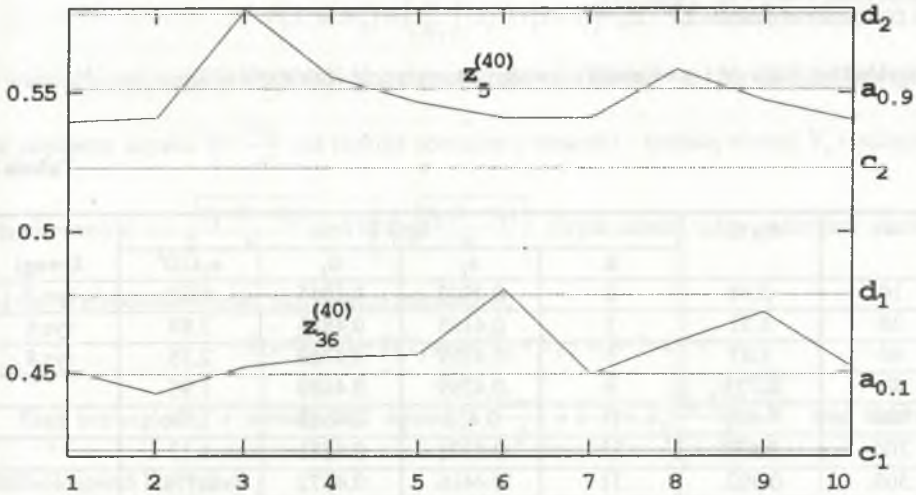


Rys.5. Rozkłady kwantyli modułu wzmocnienia napięciowego z próby  $Z_2^{(10)}$  i  $Z_9^{(10)}$  oraz rozkład średniej arytmetycznej ( $a_{0,1}$ ,  $a_{0,9}$  - kwantyle rozkładu;  $[c_1, d_1]$ ,  $[c_2, d_2]$  - przedziały ufności ,

Fig. 5. Distribution of quantiles of voltage amplification for  $Z_2^{(10)}$  i  $Z_9^{(10)}$  and distribution arithmetic mean ( $a_{0,1}$ ,  $a_{0,9}$  - quantiles of distribution;  $[c_1, d_1]$ ,  $[c_2, d_2]$  - confidence intervals)

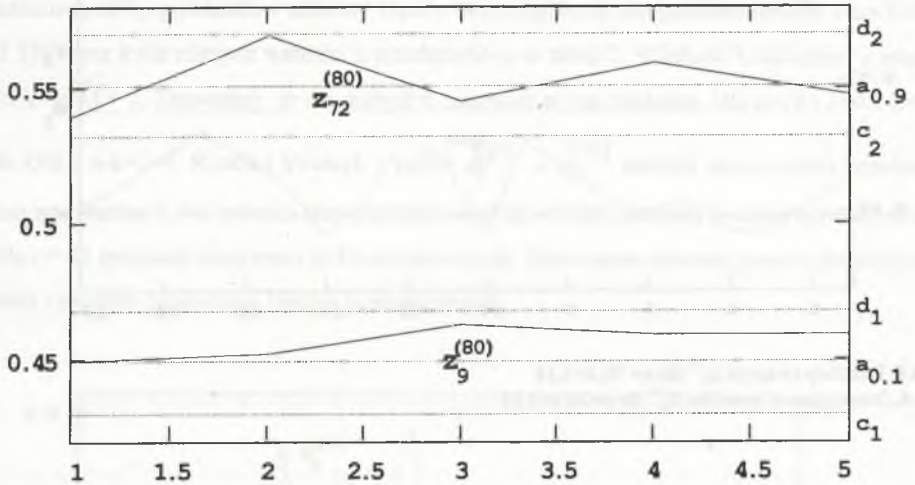


Rys.6. Rozkłady kwantyli  $Z_k^{(n)}$  dla  $n=20, k=3,18$   
 Fig.6. Distribution of quantiles  $Z_k^{(n)}$  for  $n=20, k=3,18$



Rys.7. Rozkład kwantyli  $Z_5^{(40)}, Z_{36}^{(40)}$   
 Fig.7. Distribution of quantiles  $Z_5^{(40)}, Z_{36}^{(40)}$

Rozkłady kwantyli  $z_3^{(20)}$  i  $z_{18}^{(20)}$  przedstawiono na rysunku 6. Przedstawiono również rozkłady kwantyli  $z_5^{(40)}$  i  $z_{36}^{(40)}$  (rys.7) oraz kwantyli  $z_9^{(80)}$  i  $z_{72}^{(80)}$  (rys.8).



Rys.8. Rozkłady kwantyli  $Z_9^{(80)}$ ,  $Z_{72}^{(80)}$   
Fig.8. Distribution of quantiles  $Z_9^{(80)}$ ,  $Z_{72}^{(80)}$

Przykład ten ilustruje i potwierdza rozważania teoretyczne (zależność (10)).

Tabela 2

n	$\sigma_1 \times 10^2$	$\lambda=0.1$				Uwagi
		k	$c_1$	$d_1$	$e_1 \times 10^2$	
10	2.14	2	0.3943	0.5045	5.51	rys. 5
20	1.51	3	0.4105	0.4883	3.89	rys.6
40	1.07	5	0.4219	0.4769	2.75	rys.8
80	0.755	9	0.4299	0.4689	1.95	
100	0.676	11	0.432	0.4668	1.74	
200	0.478	21	0.4371	0.4671	1.23	
500	0.302.	51	0.4416	0.4572	0.778	
1,000	0.214	101	0.4439	0.4549	0.55	



n	$\sigma_1 \times 10^2$	$\lambda=0.9$				Uwagi
		n-k+1	$e_2$	$d_2$	$e_2 \times 10^2$	
10	2.14	9	4,955	6,057	5.51	rys.5
20	1.51	18	5,117	5,895	3.89	rys. 6
40	1.07	36	5,231	5,781	2.75	rys.7
80	0.755	72	5,311	5,701	1.95	rys.8
100	0.676	91	5,332	5,680	1.74	
200	0.478		5,383	5,629	1.23	
500	0.302		5,428	5,584	0.778	
1,000	0.214		5,451	5, 561	0.55	

### 3. Porównanie z metodą Monte Carlo

Przypnijmy jak w metodzie Monte Carlo [2,3] określa się niezbędną liczbę losowań zapewniającą błąd mniejszy od zadanego. Rozpatrzmy proces losowania MC mający na celu określenie uzysku produkcyjnego  $Y_s$ . Jest to schemat Bernoulliego, w którym  $Y_s$  jest prawdopodobieństwem sukcesu, a  $Y_F=1-Y_s$  prawdopodobieństwem niepowodzenia. Prawdopodobieństwo, że w  $N$  losowaniach MC uzyskamy dokładnie  $N$  sukcesów, wynosi:

$$P(X = N_s) = \binom{N}{N_s} \cdot Y_s^{N_s} \cdot (1 - Y_s)^{(N - N_s)},$$

a wartość oczekiwana  $m = E(X) = N \cdot Y_s$  oraz wariancja  $\text{var}(X) = \sigma^2 = n \cdot Y_s \cdot (1 - Y_s)$ . Stąd wynika,

że estymator uzysku  $Y = \frac{N_s}{N}$  ma rozkład normalny o wartości średniej równej  $Y_s$  i odchyleniu

standardowym  $\sigma = \sqrt{\frac{Y_s \cdot (1 - Y_s)}{N}}$ , czyli  $N \left( Y_s \sqrt{\frac{Y_s \cdot (1 - Y_s)}{N}} \right)$ . Uzysk mieści się z poziomem ufności  $\gamma$

$= 1 - \alpha$  w przedziale ufności określonym zależnością:

$$P(c < Y_s < d) = P(Y_s - a \cdot \sigma < Y_s < Y_s + a \cdot \sigma) = 1 - \alpha = \gamma$$

Błąd bezwzględny  $e$  metody MC wynosi  $e = \frac{d - c}{2} = a \cdot \sigma = a \cdot \sqrt{\frac{Y_s \cdot (1 - Y_s)}{N}}$ , stąd niezbędna liczba losowań  $N$  wynosi:

$$N = \frac{Y_s \cdot (1 - Y_s) \cdot a^2}{e^2} \quad (12)$$

Jeśli zdefiniujemy błąd względny  $e_s$  jako stosunek błędu bezwzględnego do uzysku, czyli  $e = \frac{e}{Y_s}$ , wówczas

$$N = \left( \frac{a}{e_s} \right)^2 \cdot \frac{1 - Y_s}{Y_s} \quad (13)$$

Ponieważ metoda MC zastosowana jest do określenia uzysku  $Y_s$ , a metoda wykorzystująca do uporządkowania wylosowanych wyników statystyki pozycyjne (SP) przeznaczona jest do określenia kwantyli, niemożliwe jest w ogólnym przypadku ich porównanie. Załóżmy zatem, że  $\lambda = \frac{1 - Y_s}{2}$ , co umożliwi porównanie.

### Przykład 3

Należy porównać niezbędną liczbę losowań  $N$  w metodzie MC dla różnych wartości uzysku  $Y_s$ , z liczbą losowań  $n$  w metodzie SP dla  $\lambda = \frac{1 - Y_s}{2}$ , jeśli próby będą wybierane z populacji o rozkładzie prostokątnym określonym w przykładzie 1, błąd bezwzględny  $e = 0.01$ , a poziom  $\gamma = \alpha = 0.99$  ( $a = 2.576$ ). Korzystając z zależności (11) i (12) sporządzono wykresy zależności  $n = n(\lambda)$  i  $N(Y_s)$  przedstawione na rys.9. Z rysunku tego wynika, że dla  $Y_s = \frac{1}{3}$  ( $\lambda < \frac{1}{3}$ )  $N > n$ , co oznacza, że liczba losowań w metodzie MC jest większa.

Wykres względnej procentowej różnicy liczby losowań zdefiniowanej zależnością :

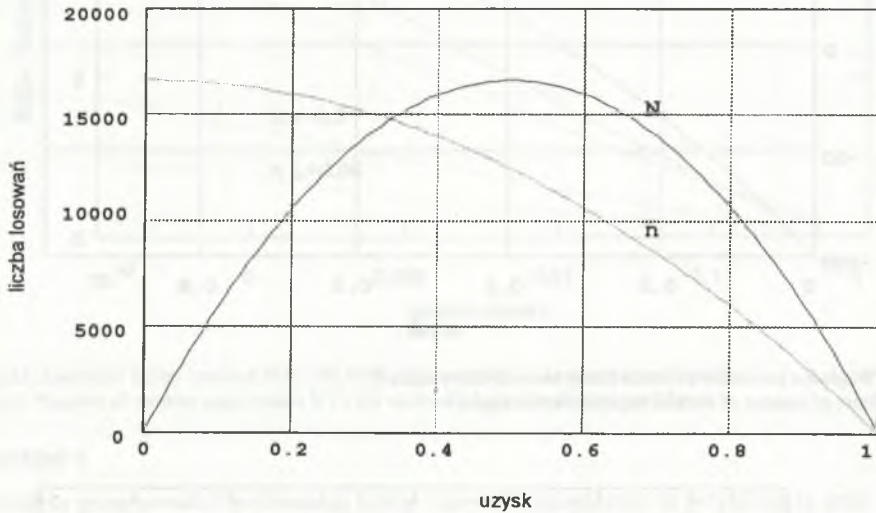
$$\sigma\% = \frac{N - n}{n} \cdot 100\% = \frac{\frac{a^2 \cdot Y_s(1 - Y_s)}{e^2} - \frac{a^2 \cdot \lambda(1 - \lambda)}{e^2 \cdot f^2(a_\lambda)}}{\frac{a^2 \cdot \lambda(1 - \lambda)}{e^2 \cdot f^2(a_\lambda)}} \quad (14)$$

przedstawiono na rys.10. Oczywiście  $f(a_\lambda) = 1$ , gdy  $0 < a_\lambda < 1$ , ponieważ populacja ma rozkład prostokątny. Względna procentowa różnica jest największa dla dużych uzysków.

Do potwierdzenia wniosków wynikających z wykresów przedstawionych na rysunku 9 i 10 można dojść rozpatrując wykresy zależności liczby losowań przy stałym błędzie względnym  $e_s$ .

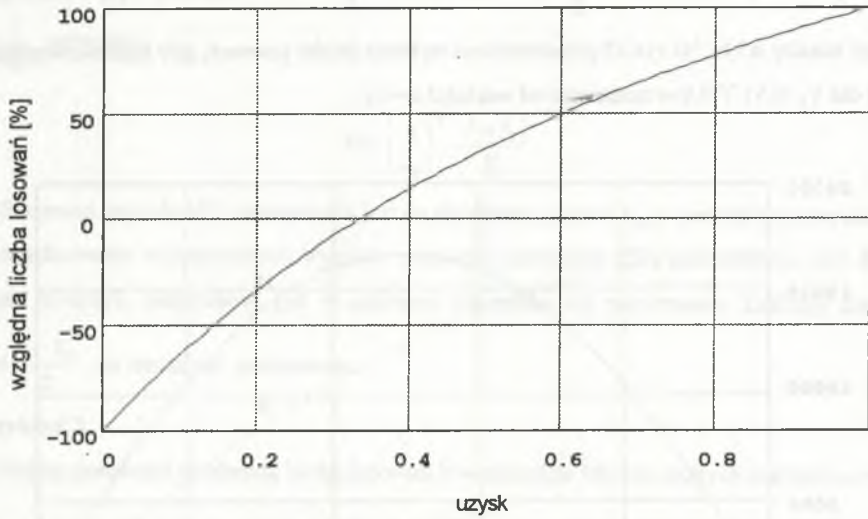
Na rysunku 11 przedstawiono wykresy liczby losowań  $n$  i  $N$  dla  $e_s = \frac{e}{Y_s}$  przy niezmięnionej

wartości poziomu ufności i  $\lambda = \frac{1 - Y_s}{2}$ . Również zmiana poziomu ufności nie wpływa na zmianę relacji między  $n$  i  $N$ . Na rys.12 przedstawiono wykresy liczby losowań, gdy błąd bezwzględny  $e_s = 0.01$  dla  $Y_s = 0.5$  i  $Y = 0.9$ , w zależności od wartości  $\alpha = 1 - y$ .

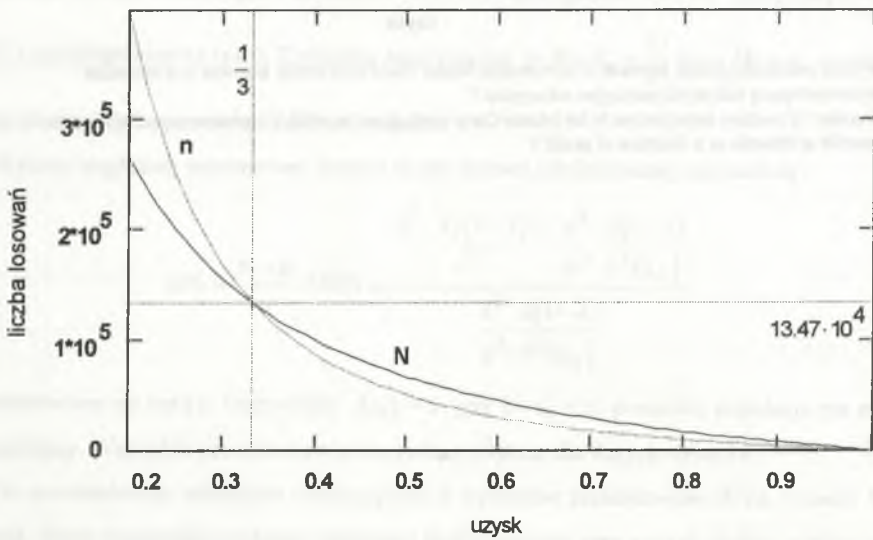


Rys.9. Wykres zależności liczby losowań  $N$  w metodzie Monte Carlo oraz liczby losowań  $n$  w metodzie wykorzystującej statystyki pozycyjne od uzysku  $Y$

Fig. 9. Number of random experiments  $N$  for Monte Carlo method and number of random experiments  $n$  for quantile arithmetic as a function of profit  $Y$

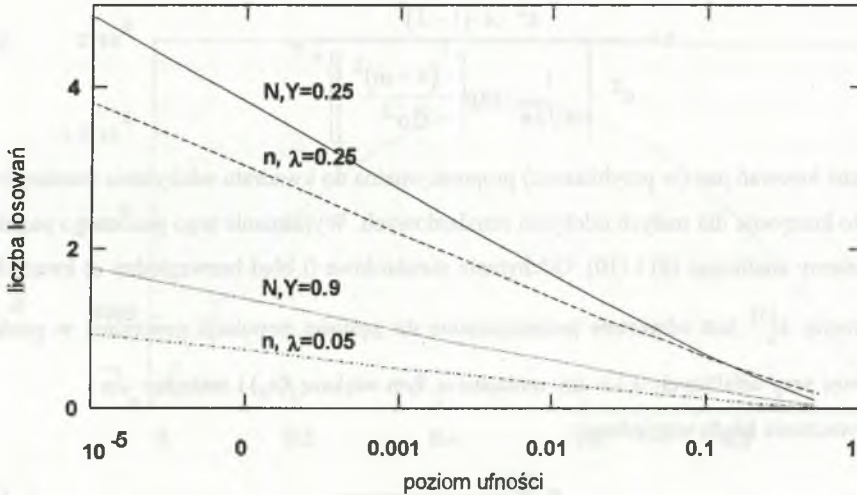


Rys.10. Względna procentowa różnica liczby losowań dla rysunku 9  
 Fig.10. Ratio of number of random experiments for figure 9



Rys.11. Wykres N i n jak na rysunku 9) dla stałego błędzi  $e/Y=0.01$   
 Fig. 11. Figures of N and n (as on figure 9) for constant error  $e/Y=0.01$





Rys.12. Zależność liczby losowań  $N$  i  $n$  dla wybranych wartości  $Y, \lambda$  od poziomu ufności  
 Fig.12. Number of random experiments  $N$  i  $n$  for various  $Y, \lambda$  as a function of confidence

#### Przykład 4

Należy przeprowadzić porównanie liczby losowań (analogicznie do przykładu 3), jeśli próby będą wybierane z populacji o rozkładzie normalnym określonym gęstością:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp \left[ \frac{-(x - m)^2}{2\sigma^2} \right] \quad (15)$$

Porównanie przeprowadzić dla  $\sigma=0.1$ ,  $e=0.01$ ,  $\gamma=1-\alpha=0.99$  ( $a=2.576$ ). Korzystając z zależności (11) na liczbę losowań  $n$  (metoda SP) oraz zależności (13) na liczbę losowań  $N$  (metoda MC) sporządzono wykresy  $n=n(\lambda)$  dla  $\lambda = \frac{1 - Y_s}{2}$  oraz  $N=N(Y_s)$  przedstawione na rys.13. Z wykresów tych wynika, że dla  $Y_s < 0.935n < N$ .

Stosowanie przy porównywaniu liczby losowań błędu bezwzględnego  $e$ , gdy próby są wybierane z populacji o rozkładzie różnym od prostokątnego, powoduje uzależnienie liczby losowań od parametrów losowań. Prześledźmy to na przykładzie rozkładu normalnego.

Podstawiając (15) dla  $x=a_\lambda$  do(11) otrzymujemy:

$$n = \frac{a^2 \cdot \lambda \cdot (1 - \lambda)}{e^2 \cdot \left\{ \frac{1}{\sigma \sqrt{2\pi}} \cdot \exp \left[ \frac{-(x - m)^2}{2\sigma^2} \right] \right\}^2}, \quad (16)$$

czyli liczba losowań jest (w przybliżeniu) proporcjonalna do kwadratu odchylenia standardowego. Jest to korzystne dla małych odchyżeń standardowych. Wyjaśnienie tego pozornego paradoksu znajdziemy analizując (8) i (10). Odchylenie standardowe (i błąd bezwzględny  $e$ ) kwantyla z próby prostej  $z_k^{(n)}$  jest odwrotnie proporcjonalne do gęstości populacji generalnej w punkcie  $x=a_\lambda$ , a więc przy ustalonych  $\lambda$  i  $e$ -im mniejsze  $\sigma$ , tym większe  $f(a_\lambda)$  i mniejsze  $\sqrt{n}$ .

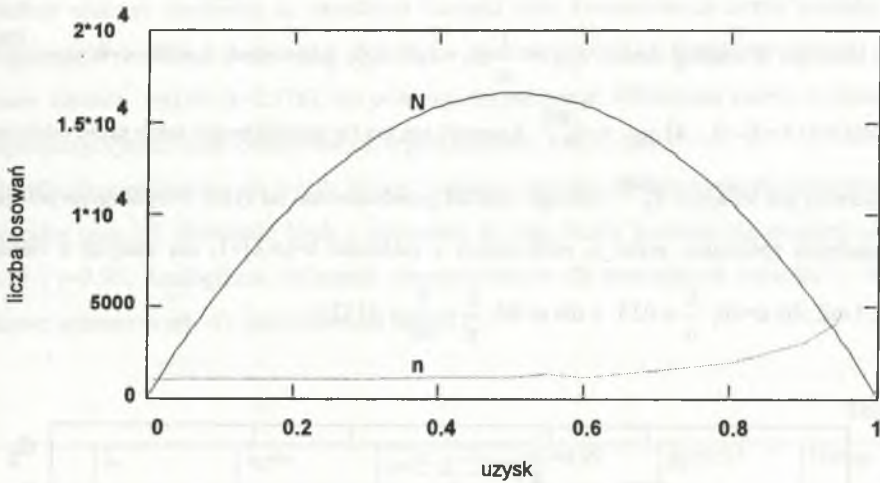
Wprowadzenie błędu względnego:

$$e_s = \frac{e}{f(a_\lambda) = e \cdot f(a_\lambda)} \quad (17)$$

uniezależnia liczbę losowań od rozkładu gęstości populacji generalnej. Otrzymujemy wówczas, wstawiając (17) do (11), wzór na liczbę losowań  $n$  w zależności od przyjętego poziomu ufności  $\gamma = 1 - \alpha$  (stąd  $a$ ), wielkości  $\lambda (0 < \lambda < 1)$  i dopuszczalnego błędu względnego  $e_s$ .

$$n = \frac{a^2 \cdot \lambda \cdot (1 - \lambda)}{e_s^2} \quad (18)$$

Wykresy  $n=n(\lambda)$  będą identyczne z charakterystykami uzyskanymi wówczas, gdy próby były wybierane z populacji o rozkładzie prostokątnym (rys.9 i 10), ponieważ  $f(a_\lambda) = 1$ , a  $e_s = e \cdot f(a_\lambda) = e$ .



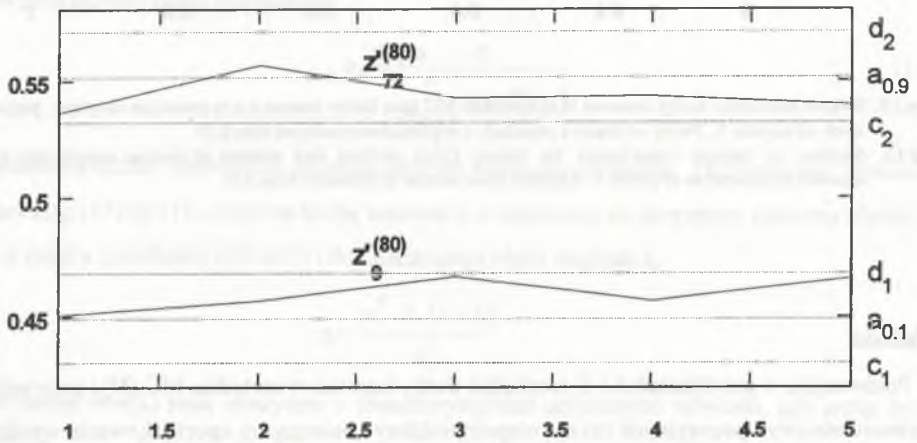
Rys.13. Wykres zależności liczby losowań  $N$  w metodzie MC oraz liczby losowa  $n$  w metodzie statystyk pozycyjnych od uzysku  $Y$ . Próby wybrano z populacji o rozkładzie normalnym  $N(m,0.1)$

Fig.13. Number of random experiments for Monte Carlo method and number of random experiments  $n$  for quantile arithmetic as of profit  $Y$ . Samples have normal distribution  $N(m,0.1)$

### Wnioski

Porównując, w przykładach 3 i 4, niezbędne liczby losowań w metodzie MC ( $N$ ) i przy wykorzystaniu statystyk pozycyjnych ( $n$ ) nie rozpatrywaliśmy konieczności uporządkowania wyników obserwacji (rezultatów analizy na m.c) w drugiej metodzie. Poświęćmy teraz temu zagadnieniu trochę uwagi. Uporządkowanie, w kolejności niemalejącej, nawet dużej liczby obserwacji  $n$  nie jest dla m.c. w porównaniu z czasem analizy dużego obwodu, zajęciem czasochłonnym. Można ten czas wydatnie skrócić, jeśli zamiast kwantyla  $z_k^{(n)}$ , gdzie  $n=n_1 \cdot m$  obserwacji, weźmy średnią arytmetyczną  $m$  kwantyli  $z_{k_1}^{(n_1)}$  (oczywiście  $k=[n \cdot \lambda]+1$ , a  $k_1=[n_1 \cdot \lambda]+1$ ), czyli zamiast porządkować  $n=n_1 \cdot m$  obserwacji dzielimy je na  $m$   $n_1$ -elementowych (uporządkowanych) prób. Wówczas kwantyl  $z_k^{(n)}$  ma rozkład  $N(a, \sigma_1)$ .

Na rys.14 przedstawiono rozkład kwantyla  $z_9^{(80)}$  (i kwantyla  $z_{72}^{(80)}$ ) otrzymanego z uśredniania kwantyli z według wzoru:  $Y_{j+1} = \frac{1}{m}(x_1 + x_2 + \dots + x_m)$ , gdzie  $m=4$ ,  $n_1=20$ ,  $n=80$ ,  $x_i = z_3^{(20)}$  z rys.6, dla  $t=i+4 \times j$  ( $= 0, \dots, 4$ ),  $y_{j+1} = z_9^{(80)}$ . Kwantyl ten ma (w przybliżeniu) takie samo odchylenie standardowej jak kwantyl  $z_9^{(80)}$ , którego rozkład przedstawiono na rys.8. Przybliżenie polega na niedokładnym spełnianiu przez  $k$ , obliczonym z zależności  $k=[n\lambda]+1$ , dla małych  $n$  warunku  $\frac{k}{n} = \lambda$  i tak, dla  $n=20$ ,  $\frac{k}{n} = 0.15$ , a dla  $n=80$ ,  $\frac{k}{n} = \frac{9}{80} \doteq 0.1125$ .



Rys.14. Rozkład kwantyli  $z_9^{(80)}$  i  $z_{72}^{(80)}$  uzyskanych przez uśrednienie kwantyli  $z_3^{(20)}$  i  $z_{18}^{(20)}$

Fig.14. Distribution of quantiles  $z_9^{(80)}$  i  $z_{72}^{(80)}$  obtained as a mean of  $z_3^{(20)}$  i  $z_{18}^{(20)}$

Na rys.14 zaznaczono dodatkowo kwantyle  $a_{0.15}$  i  $a_{0.20}$  oraz odpowiadający temu ostatniemu przedział ufności  $[c', d']$ . Wydaje się, że metoda statystyk pozycyjnych może być z powodzeniem wykorzystana do określania liczb kwantylowych[15]. Zilustrujmy to przykładem.



**Przykład 5**

Należy obliczyć niezbędną do określenia wartości liczb kwantylowych liczbę losowań, jeśli próby będą wybierane z populacji o rozkładzie normalnym  $(N(0,1))$ . Przedziały ufności, z poziomem ufności  $\gamma=0.99$  ( $a=2.576$ ), nie powinny się pokrywać. Obliczenia należy wykonać dla przystopozycyjnych liczb kwantylowych o pozycyjności  $\lambda=2,4,\dots,20$ .

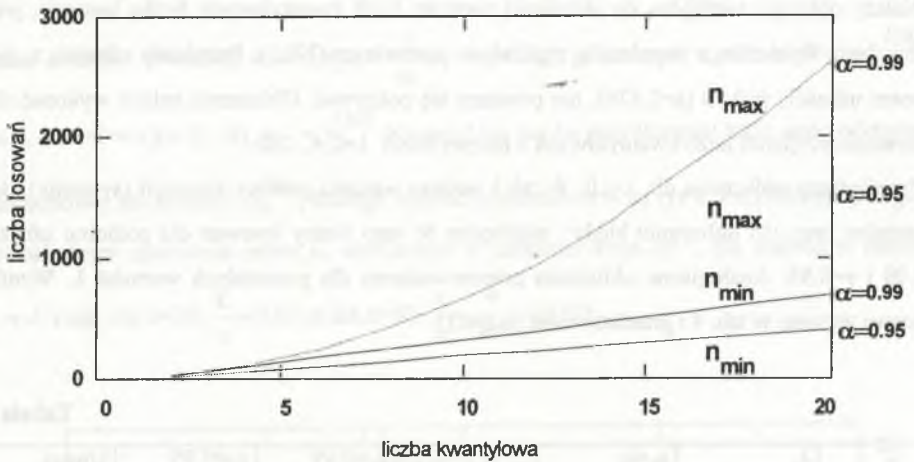
Prześledźmy obliczenia dla  $\lambda=10$ . W tab.3 podano wartości połowy kwantyli (symetria), dopuszczalne przy ich obliczaniu błędy i niezbędne do tego liczby losowań dla poziomu ufności  $\gamma=0.99$  i  $\gamma=0.95$ . Analogiczne obliczenia przeprowadzono dla pozostałych wartości  $\lambda$ . Wyniki końcowe zebrano w tab. 4 i przedstawiono na rys.15.

**Tabela 3**

i	$\lambda_i$	$a_\lambda = a_i$	$e = \frac{a_{i-1} - a_i}{2}$	$g_1=0.99$ $n_i$	$g_2=0.95$ $n_i$	Uwagi
6	0.55	0.126	0.126	671	391	$n_{\max}$
7	0.65	0.3854	0.1297	654	381	
8	0.75	0.674	0.1443	591	344	
9	0.85	1.04	0.183	468	272	
10	0.95	1.645	0.3025	324	189	$n_{\min}$

**Tabela 4**

l		2	4	6	8	10	12	14	16	18	20
$\gamma_1 =$ $=0.99$	$n_{\min}$	27	99	175	246	324	395	469	547	623	698
	$n_{\max}$	27	106	240	427	671	949	1,291	1,717	2,131	2,630
$\gamma =$ $=0.95$	$n_{\min}$	16	58	102	142	189	229	272	317	361	404
	$n_{\max}$	16	68	139	247	391	550	747	994	1,234	1,523



Rys.15. Zależność liczby losowań od pozycyjności liczby kwantylowej dla dwóch poziomów ufności, gdzie  $n_{max}$  to liczba losowań niezbędna do określenia kwantyli środkowych

Fig. 15. Number of random experiments as a function of for two confidence levels

## LITERATURA

1. Fisz M.: Rachunek prawdopodobieństwa i statystyka matematyczna. PWN, Warszawa 1969
2. Zieliński R.: Metody Monte Carlo. WNT Warszawa 1970
3. Stybliński M.: Metody analizy i optymalizacji statystycznej układów elektronicznych. Warszawa 1980
4. Chojcan J.L., Lasek.: Metody analizy wrażliwościowej układów elektronicznych. Skrypty uczelniane Politechniki Śląskiej nr 937, Gliwice 1980
5. Chojcan J., Karwan L.: Optymalny dobór tolerancji elementów oraz zastosowanie liczb kwantylowych (ss. 105-145) w monografii „Metody statystycznej i wrażliwościowej analizy i optymalizacji układów”, IPE Politechniki Warszawskiej, Warszawa 1981

Recenzent: Prof.dr hab. inż. Stanisław Kuta

Wpłynęło do Redakcji 28.09.1994 r.

**Abstract**

In the paper definitions and properties of the quantile arithmetic are developed. There are definition of positions statistics, range of position statistics, quantile from sample and empirical distribution. An example of position statistics for uniformly distributed random variable is discussed. Position statistics can be used to interval estimation of some parameters of the experimental distribution. In the article application of the method to the statistical analysis of an active pass-band filter is given. The method provides an improvement in number of necessary observations compared to the Monte Carlo method. The most improvement is obtained when the mean value of the experimental quantiles is computed. The statistical properties of the method is discussed in detail.