# Wykład 4
# Karty kontrolne sum skumulowanych CUSUM

## Test sum skumulowanych CUSUM

Dla modelu z losowym dryfem wartości oczekiwanej proces „wędruje” w czasie z obszaru o zadowalającym działaniu do obszaru o działaniu niezadowalającycm. Do tego celu dobrym narzędziem będzie sekwencyjny test oparty na ilorazie wiarogodności SPRT. Jest to nico inna filozofia niż przy klasycznych kartach kontrolnych dla średniej. Test ten jest mniej wrażliwy na pojedyncze sygnały rozregulowania, świetnie natomiast się zachowuje w przypadku trwałego przesunięcia się wartości średniej.

### Model

Załóżmy, że obserwowane wielkości mają rozkład normalny z wartością oczekiwaną $\mu$ i wariancją $\sigma^2$. Zakładamy, że $\sigma^2$ jest znana i nie zmienia się w czasie. Chcemy jak najszybciej wykryć możliwą zmianę wartości średniej z $\mu_0$ na $\mu_1$.

Wykrycie tego typu rozregulowania przy pomocy tradycyjnych kart $Xbar$ oraz $S$ szczególnie, gdy przesunięcie średniej jest nieznaczne, jest zwykle trudne i może pozostać niezauważone. Załóżmy, że w chwili $N \geqslant 20$ mamy $N$ niezależnych próbek o liczności $n$. Utwórzmy logarytm ilorazu wiarogodności

$$\begin{aligned}
\mathcal{R}_1 &= \ln\left(\frac{\prod_{j=1}^{N} \frac{1}{\sqrt{2\pi}\frac{\sigma}{\sqrt{n}}} \exp\left[-\frac{(\overline{x}_j-\mu_1)^2}{2\sigma^2/n}\right]}{\prod_{j=1}^{N} \frac{1}{\sqrt{2\pi}\frac{\sigma}{\sqrt{n}}} \exp\left[-\frac{(\overline{x}_j-\mu_0)^2}{2\sigma^2/n}\right]}\right) = \\
&= \frac{1}{2\sigma^2/n}\sum_{i=1}^{N}[(\overline{x}_j-\mu_0)^2-(\overline{x}_j-\mu_1)^2] = \\
&= \frac{\mu_0-\mu_1}{2\sigma^2/n}[N(\mu_0+\mu_1)-2N\bar{\bar{x}}] = \\
&= N\frac{\mu_1-\mu_0}{\sigma^2/n}\left[\bar{\bar{x}}_N-\frac{\mu_0+\mu_1}{2}\right]
\end{aligned} \tag{1}$$

W przypadku, gdy $\mu_1$ nie jest wyspecyfikowane, to można je zastąpić wartością estymatora $\mu$, tzn. $\bar{\bar{x}}$. Wówczas statystyka testowa przyjmuje postać

$$\mathcal{R}_1 = N\frac{(\bar{\bar{x}}_N-\mu_0)^2}{2\sigma^2/n} \tag{2}$$

Postępowanie to jest podobne do kart kontrolnych dla wartości średniej, z tym, że w $j$-tym kroku przeprowadzamy weryfikację w oparciu o $\bar{\bar{x}}_j$, a nie $\bar{x}_j$. Jasne jest zatem, że test ten nie będzie wykrywał sporadycznych sygnałów rozregulowania a tylko istotną i trwałą zmianę wartości średniej.

## Wykorzystanie testu CUSUM do SSP

Test CUSUM można wykorzystać do statystycznego sterowania procesem (SSP). W tym celu przyjmijmy, że

- $\mu_0$ - dolna dopuszczalna granica dla wartości średniej procesu,
- $\mu_1$ - górna dopuszczalna granica dla wartości średniej procesu,
- $\mu^* = \frac{\mu_0+\mu_1}{2}$ - środek pasa tolerancji.

Zwykle, ale nie zawsze, chcemy utrzymać proces w środku pasa tolerancji. Odpowiednikiem obszaru dopuszczalnego dla testu sekwencyjnego jest obszar kontynuacji.

$$\ln B = \ln(k_0) < N\frac{\mu_1-\mu_0}{\sigma^2/n}\left[\bar{\bar{x}}_N - \frac{\mu_0+\mu_1}{2}\right] < \ln(k_1) = \ln A, \tag{3}$$

gdzie $k_0 = B$ a $k_1 = A$ (w SSP ograniczenia SPRT oznacza się jako $k_0$ i $k_1$).

Aby zilustrować działanie testu sekwencyjnego rozważmy następujący przykład. Załóżmy, że $N$ pierwszych próbek jest idealnie w środku pasa ktolerancji, a $N+1$ różni się od pozostałych, tzn, $\bar{x}_1 = \ldots = \bar{x}_N = \mu^*$ oraz $\bar{x}_{N+1} = \mu^* + \delta$. Wówczas

$$\mathcal{R}_1^N = N\frac{\mu_1-\mu_0}{\sigma^2/n}\left[\frac{\mu_0+\mu_1}{2} - \frac{\mu_0+\mu_1}{2}\right] = 0$$

Tymczasem

$$\mathcal{R}_1^{N+1} = (N+1)\frac{\mu_1-\mu_0}{\sigma^2/n}\left[\frac{\mu_0+\mu_1}{2} + \frac{\delta}{N+1} - \frac{\mu_0+\mu_1}{2}\right] = \frac{\mu_1-\mu_0}{\sigma^2/n}\cdot\delta$$

i nie zależy od $N$. Zatem $N$ pierwszych próbek „nominalnych" nie uryje $N+1$ próbki odbiegającej od nominału. Z drugiej strony, gdyby $\bar{x}_{N+2} = \mu^* - \delta$, to $\mathcal{R}_1^{N+2} = 0$ i szybko po sobie następujące sygnały rozregulowania o przeciwnych znakach mogą się zniwelować.

Rozważmy następujący przykład. Załóżmy, że sterujemy procesem, dla którego $\mu = 75$, $\sigma = 5$, $USL = 95$ oraz $LSL = 55$. Linie specyfikacji nie mogą być przekroczone. Każdy produkt poza liniami specyfikacji jest niezgodny z wymaganiami. Przy założeniu próbek czteroelementowych ($n = 4$) mamy $GLK = 82,5$ a $DLK = 67,5$ (zakładamy stabilność procesu $\bar{\bar{x}} = \mu$, $\hat{\sigma} = \sigma$). Załóżmy, że nastąpiła zmiana średniej do wartości 80 (o $2\sigma_{\bar{x}}$). Pytanie: w jaki sposób to wpłynie na zwiększenie liczby wyrobów niezgodnych oraz jak długo trzeba czekać na wykrycie tej zmiany przy pomocy karty kontrolnej $Xbar$?

Nietrudno obliczyć, że dla procesu stabilnego, sznsa na przekroczenie dolnej lub górnej linii specyfikacji jest jak 3 : 100 000. Zmiana średniej zmieni te szanse na 3 : 10 000 000, dla dolnej, oraz 1 : 1 000, dla górnej linii specyfikacji. Tym samym było 6 niezgodności na 100 000 wyprodukowanych, a po zmianie mamy 100 na 100 000. Aby odpowiedzieć na pytanie jak długo trzeba czekać na wykrycie zmiany przez kartę $Xbar$, wystarczy zauważyć, że po zmianie średniej, średnie próbkowe $\bar{x}$ mają rozkłady normalne $N(80, (2,5)^2)$. Tym samym $P(\bar{x} > GLK) = 0,1587$ oraz $P(\bar{x} < DLK) = 0$.

**Średnią długością serii - ARL** (Average Run Length) nazywamy średni czas oczekiwania na zdarzenie o prawdopodobieństwie $p$. Nietrudno zauważyć, że $ARL = \frac{1}{p}$, gdyż czas oczekiwania na pierwszy sukces w schemacie Bernoulli'ego ma rozkład geometryczny.

Zatem w naszym przykładzie dla procesu stabilnego mamy

$$ARL = \frac{1}{0,0027} = 370,37$$

a dla procesu z aburzoną średnią

$$ARL = \frac{1}{0,1587} = 6,30,$$

tzn. jeżeli pobieramy próbki co pół godziny, to zmianę średniej zarejestrujemy średnio po 3,5 godz. co przy produkcji 100 000 detali podczas jednej zmiany, daje nam średnio 60 niezgodności do wykrycia zmiany średniej.

Okazuje się, że dla takich przesunięć ($2\sigma_{\bar{x}}$) lepsze są testy CUSUM. Dla dużych przesunięć ($3\sigma_{\bar{x}}$ i $4\sigma_{\bar{x}}$) lepsza jest karta $Xbar$. Dla klasycznego testu SPRT o sile $(0.01, 0.01)$ $\ln(k_0) = \ln \frac{\beta}{1-\alpha} = -4,595$, oraz $\ln(k_1) = 4,595$. W praktyce zwykle $4 \leqslant \ln(k_1) \leqslant 5$, a jako wartości różnicy $\mu_1 - \mu_0$ przyjmuje się zwykle $0.5\sigma$ lub $\sigma$. W praktycznych zastosowaniach tej karty stosuje się tzw. $V$-maski, które ułatwiają wykrycie sygnału rozregulowania.

## Karta kontrolna CUSUM Shewharta

Często jako alternatywę dla karty sum skumulowanych CUSUM używa się tzw. karty kontronej CUSUM Shewharta. Weryfikacja przebiega w oparciu o bieżącą wartość średnią próbek połączonych

$$\bar{\bar{x}}_N = \frac{1}{N}\sum_{j=1}^{N} \bar{x}_j, \tag{4}$$

gdzie $N$ odgrywa rolę indeksu bieżącego. Zgodnie z przyjętymi założeniami wszystkie średnie próbkowe są niezależne, zatem

$$E(\bar{\bar{x}}_N) = \mu$$

oraz

$$V(\bar{\bar{x}}_N) = \frac{\sigma_0^2}{Nn}.$$

Do wyznaczenia obszaru uregulowania używamy statystyki

$$\mathcal{R}_2^N = \frac{\bar{\bar{x}}_N - \mu_0}{\frac{\sigma_0}{\sqrt{nN}}}, \tag{5}$$

która ma w przybliżeniu standardowy rozkład normalny. Zatem

$$GLK = 3 \text{ a } DLK = -3.$$

Niestety karta ta posiada dość istotną wadę, gdyż np. dla $\bar{x}_1 = \ldots = \bar{x}_N = \mu_0$ oraz $\bar{x}_{N+1} = \mu_0 + \delta$, gdzie $\mu_0$ oraz $\sigma_0$ są parametrami procesu uregulowanego (nominalnego), to pomimo, że $\mathcal{R}_2^N = 0$ mamy

$$\mathcal{R}_2^{N+1} = \frac{\bar{\bar{x}}_{N+1} - \mu_0}{\frac{\sigma_0}{\sqrt{n(N+1)}}} = \frac{\frac{\delta}{\sqrt{N+1}}}{\frac{\sigma_0}{\sqrt{n}}}$$

tzn. wartość statystyki $\mathcal{R}_2$ zależy od tego jak długo proces był uregulowany. Wynika stąd, że im większe $N$ tym mniejsza szansa wykrycia przesunięcia średniej. Karta ta, choć ze względu na swoją prostotę, często stosowana, ma sporą inercję i rzadko wykrywa pojedyncze sygnały rozregulowania, a zmianę średniej zbyt późno. Nie zaleca się stosowanie jej bez jednoczesnego stosowania klasycznej karty średniej $Xbar$.

## Karta kontrolna CUSUM Page'a

Jest to najczęściej stosowana karta kontrolna sum skumulowanych ze względu na to, że bardzo dobrze i szybko wykrywa trwałą zmianę średniej.

Zakładamy, że proces jest uregulowany z parametrami $\mu_0$ oraz $\sigma$. Możliwa jest zmiana średniej do wartości $\mu_1$, ale wariancja procesu pozostaje nie zmieniona. Tak jak poprzednio $N$ oznacza indeks bieżący, a pobierane w regularnych odstępach czasu próbki mają liczność $n$. Niech

$$z_N = \frac{\bar{\bar{x}}_N - \mu_0}{\sigma/n} \tag{6}$$

będzie standaryzowaną wartością średnią $N$-tej próbki. Dla każdej chwili $N$ wyznaczamy wartości liczbowe dwóch statystyk - górnej i dolnej sumy skumulowanej:

$$SH_N = \max[0, (z_N - k) + SH_{N-1}] \tag{7}$$

oraz

$$SL_N = \max[0, (-z_N - k) + SL_{N-1}], \tag{8}$$

gdzie $SH_0 = SL_0 = 0$ a $k$ jest pewną stałą. Zauważmy, że górna suma skumulowana (7), jest zawsze nieujemna i wzrasta jeżeli $z_N > k$. Podobnie dolna suma skumulowana (8) jest nieujemna i wzrasta jeżeli $z_N < -k$. Wielkość $k$ jest zwykle równa połowie założonej wartości bezwzględnej przesunięcia wartości średniej procesu, wyrażonej w jednostkach standaryzowanych, tzn.

$$k = \sqrt{n}|\mu_1 - \mu_0|/(2\sigma). \tag{9}$$

Sygnałem rozregulowania jest przekroczenie przez jedną ze statystyk (7) lub (8) zadanej z góry wielkości $h$. Wartość $h$ zwykle jest równa 4, rzadziej 5. Przy czym $SH_N > h$ oznacza przesunięcie średniej w górę, a $SL_N > h$ oznacza przesunięcie średniej w dół. Wybór wartości $k$ oraz $h$ przeprowadzono w oparciu o gruntowną analizę średniej długości serii $ARL(D)$, gdzie $D$ oznacza wielkość przesunięcia średniej. Według tych badań wynika, że przesunięcie (9) prowadzi zwykle do procedury o najkrótszej długości serii $ARL(k)$ przy ustalonym $ARL(0)$. Ponadto wybrane wartości $h$ zapewniają największą średnią długość serii w przypadku, gdy proces jest uregulowany przy jednocześnie małej średniej $ARL$ dla procesu rozregulowanego.

W tabeli poniżej podano porównanie średniej długości serii $ARL$ dla klasycznej katy kontrolnej $Xbar$ oraz karty CUSUM Page'a dla $k = 0,5$ oraz $h = 4$ i $h = 5$.

| $ARL(D)$ | | | |
|---|---|---|---|
| $D$ (w $\sigma$) | $Xbar$ | $h = 4$ | $h = 5$ |
| 0 | 370,37 | 168 | 465 |
| 1 | 43,956 | 8,38 | 10,4 |
| 2 | 6,301 | 3,34 | 4,01 |
| 3 | 2,00 | 2,19 | 2,57 |
| 4 | 1,89 | 1,71 | 2,01 |
| 5 | 1,02 | 1,31 | 1,69 |

Mamy zatem trzy procedury oparte na sumach skumulowanych:

1. test sekwencyjny SPRT - najlepszy do wyboru jednej z dwóch możliwych wartości parametrów, jednakże przesunięcie średniej jest czymś innym (daje gorzsze wyniki);

2. karta kontrona CUSUM Shewharta;

3. karta kontrolna Page'a.

Związek między testem sekwencyjnym CUSUM bezpośrednio opartym na sekwencyjnym teście ilorazowym i testem Page'a zbadał Johnson w JASA - 1961. Załóżmy, że $\delta = (\mu_1 - \mu_0)/(\sigma/\sqrt{n})$ oznacza przesunięcie średniej procesu mierzone w jednostkach standaryzowanych, wówczas nierówność (3) może być zapisana

$$\ln{(k_0)} < \delta(\sum_{i=1}^{N} z_i - \frac{\delta}{2}N) < \ln{(k_1)}.$$

Dla uproszczenia rozważań, załóżmy, ze chcemy testować czy średnia nie uległa przesunięciu w górę ($\mu_1 > \mu_0$), analogiczne rozważania można przeprowadzić dla przesunięcia w dół. Korzystając z przybliżonej wartości górnego ograniczenia w teście SPRT ($k_1 = \frac{1-\beta}{\alpha}$), górna granica obszaru kontynuacji jest równa

$$\sum_{i=1}^{N} z_i = \frac{1}{\delta}\ln\left(\frac{1-\beta}{\alpha}\right) + \frac{\delta}{2}N.$$

Otrzymana zależność może być zapisana inaczej: mamy sygnał rozregulowanie jeżeli $\widetilde{SH}_N$ przekroczy wartość $h$, gdzie

$$\widetilde{SH}_N = (z_N - k) + \widetilde{SH}_{N-1},$$

gdzie $k = \frac{\delta}{2}$ oraz $h = \frac{1}{\delta} \ln \left( \frac{1-\beta}{\alpha} \right)$. Związek ten ilustruje podobieństwo i różnicę między tymi dwoma testami. Jak widać test Page'a jest pewną odmianą testu SPRT z nieco lepiej dopasowanymi do testowanej sytuacji parametrami.

Z symulacji wynika, że wartości początkowe $SH_0$ oraz $SL_0$ różne od zera nie zmieniają średniej długości serii $ARL$ dla procesu uregulowanego, jednocześnie znacząco ją skracając dla procesu rozregulowanego. Zaleca się stosowanie wartości początkowych równych $h/2$. Taki test nosi nazwę FIR CUSUM (Fast Initial Response). Należy jednak pamiętać, że testy te służą głównie wykryciu przesunięcia średniej i nigdy nie powinny być używane bez jednoczesnego stosowania klasycznej karty $Xbar$.

Często stosuje się jednocześnie karty Page'a i Shewharta. Jednakże dla takich połączonych kart zwykle przyjmuje się dla kart Shewharta $|z| \leqslant 3,5$ zamiast $|z| \leqslant 3$. takie połączenie zwiększa prawdopodobieństwo fałszywego sygnału dla procesu niezaburzonego, ale jednocześnie znacznie zwiększa prawdopodobieństwo wykrycia sygnału przy dużych odchyleniach od średniej. Najlepiej ilustruje to porównanie średniej długości serii $ARL$ zamieszczone w tabeli poniżej.

| $ARL(D)$ | | | | | |
|---|---|---|---|---|---|
| Przesunięcie średniej $D$ (w $\sigma$) | $Xbar$ | CUSUM Page'a $h=5$ | Shewhart CUSUM $z=3$ | FIR CUSUM | Shewhart FIR CUSUM $z=3,5$ |
| 0 | 370,37 | 465 | 391 | 430 | 359,7 |
| 0,25 | 281,14 | 139 | 130,9 | 122 | 113,9 |
| 0,5 | 179,26 | 38 | 37,15 | 28,7 | 28,9 |
| 0,75 | 81,22 | 17 | 16,80 | 11,2 | 11,15 |
| 1 | 43,956 | 10,4 | 10,21 | 6,35 | 6,32 |
| 2 | 6,301 | 4,01 | 3,77 | 2,36 | 2,36 |
| 3 | 2,00 | 2,57 | 2,10 | 1,54 | 1,54 |
| 4 | 1,89 | 2,01 | 1,34 | 1,16 | 1,16 |
| 5 | 1,02 | 1,69 | 1,07 | 1,02 | 1,02 |

## Test sum skumulowanych CUSUM do sterowania zmiennością procesu

W zasadzie testy (karty kontrolne) CUSUM zwykle stosuje się jedynie do sterowania średnią procesu. Na ogól nie przywiązuje się wagi do sterowania tego typu kartami zmiennościa procesu. Problemami tymi zajmowali się Johnson i Leone - *Statistics and Experimental Design in Engineering and Physical Sciences*, Vol. 1, Wiley, (1977), którzy stworzyli procedury dla $R$ i $S^2$.

Z nowszych rozwiązań, w roku 1981, Hawkins zaproponował procedurę CUSUM wykrywającą zmiany w zmienności procesu w oparciu o pojedyncze obserwacje.

Normy brytyjskie BS (British Standards) przewidują podejście zaproponowane przez Johnsona i Leone, które pokrótce omówimy.

Analogicznie jak dla procedury CUSUM dla wartości średniej obliczamy

$$S_i = \max\{0, SD_i - k \cdot \overline{SD} + S_{i-1}\}, \tag{10}$$

gdzie $SD_i$ oznacza $i$-te odchylenie standardowe, $\overline{SD}$ wartość nominalną odchylenia standardowego (zwykle jest to średnia z odpowiednio dużej liczby poprzedzających odchyleń standardowych). Sygnał rozregulowania otrzymujemy, gdy $S_i > h$. Dobór $h$ i $k$ jest zupełnie inny niż dla kart dla średniej. Tutaj odchylenie standardowe nie jest standaryzowane. Wybór stałych $k$ i $h$ jest szczegółowo omówiony w normie brytyjskiej **BS 5703** oraz w książce Johnsona i Leone.

## Karty EWMA (Exponentially Weighted Moving Averages)

Bardzo popularną formą kart CUSUM są karty dla średniej ruchomej z wagami wykładniczymi EWMA.

Zdefiniujmy $N$-tą średnią ruchomą z wagami wykładniczymi jako

$$\hat{\overline{x}}_N = r\overline{x}_N + (1-r)\hat{\overline{x}}_{N-1}, \tag{11}$$

gdzie $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_N$ są średnimi próbkowymi, $r \in (0, 1]$ a $\hat{\overline{x}}_0 = \overline{\overline{x}}$ lub znannej wartości średniej próbkowej $\mu_0$. Zauważmy, że dla $r = 1$ mamy $\hat{\overline{x}}_N = \overline{x}_N$, tzn. nie ma ważenia średniej. W przeciwnym przypadku, wcześniejsze średnie $\overline{x}_{N-i}$ wchodzą do średniej ruchomej z wagami $r(1-r)^i$. Tzn. dla $0 < r < 1$ mamy

$$\hat{\overline{x}}_N = \sum_{i=0}^{N-1} r(1-r)^i \cdot \overline{x}_{N-i} + \underbrace{(1-r)^N \cdot \hat{\overline{x}}_0}_{const} \tag{12}$$

Przyjmując, że wszystkie próbki zawierają po $n$ elementów z tej zależności można wyznaczyć wariancję średniej ruchomej z wagami wykładniczymi, mianowicie

$$V(\hat{\bar{x}}_N) = \frac{\sigma^2}{n}\left(\sum_{i=0}^{N-1} r^2(1-r)^{2i}\right) = \frac{\sigma^2}{n}\cdot\frac{r[1-(1-r)^{2N}]}{2-r}. \tag{13}$$

Zatem

$$\sigma_{\hat{\bar{x}}_N} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{r[1-(1-r)^{2N}]}{2-r}} \approx \frac{\sigma}{\sqrt{n}}\sqrt{\frac{r}{2-r}}. \tag{14}$$

Tę ostatnią, przybliżoną postać odchylenia standardowego średniej ruchomej z wagami wykładniczymi, wykorzystuje się do konstrukcji linii kontrolnych karty EWMA.

$$GLK = \overline{\overline{x}} + 3\frac{a(n)\overline{s}}{\sqrt{n}}\sqrt{\frac{r}{2-r}} \tag{15}$$

oraz

$$DLK = \overline{\overline{x}} - 3\frac{a(n)\overline{s}}{\sqrt{n}}\sqrt{\frac{r}{2-r}}. \tag{16}$$

Jest oczywiste, że średnia ważona jest wrażliwa na przesunięcie średniej procesu. Dokładna analiza pokazuje, ze w celu umożliwienia szybkiego wykrywania małych przesunięć wartości średniej, czynnik $r$ powinien leżeć gdzieś pomiędzy 0,2 i 0,5. Zwykle przyjmuje się dwie wartości: $r = \frac{1}{3}$ lub $r = \frac{1}{4}$.

## **Karta średniej ruchomej MA** (Moving Average)

Inną kartą, która pozwala na wygładzenie danych i zauważenie trendu, jest karta średniej ruchomej MA.

Załóżmy, że mamy pewien proces $x(t)$, średnią ruchomą rzędu $q$ dla procesu $x(t)$ nazywamy

$$z(t) = \begin{cases} \frac{x(t)+x(t-1)+\ldots+x(t-q+1)}{q} & \text{dla} \quad t \geqslant q \\ \\ \frac{x(t)+x(t-1)+\ldots+x(1)}{t} & \text{dla} \quad t < q \end{cases}$$

Wprowadźmy pomocniczą wielkość

$$q(t) = \begin{cases} q & \text{dla} \quad t \geqslant q \\ \\ t & \text{dla} \quad t < q \end{cases}$$

Wówczas przyjmując $x(t) = \bar{x}_t$, gdzie $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_N$ są średnimi próbkowymi możemy badać proces średnich ruchomych. Wariancja średniej ruchomej w chwili $t$ wynosi $\frac{\sigma^2}{nq(t)}$, gdzie $n$ jest licznością próbek. Wówczas linie kontrolne mają postać

$$\mu \pm 3 \cdot \frac{\sigma}{\sqrt{n \cdot q(t)}}.$$

Oczywiście, w zależności od potrzeb, parametry procesu uregulowanego $\mu$ i $\sigma$ możemy zastąpić ich estymatorami nieobciążonymi.

## Test CUSUM dla pojedynczych obserwacji

### Karta kontrolna średniej

W przypadku, gdy zamiast próbek mamy do dyspozycji jedynie pojedyncze obserwacje $x_1, x_2, \ldots, x_N$ o rozkładzie normalnym, to dokonujemy ich standaryzacji

$$z_i = \frac{x_i - \bar{x}}{\hat{\sigma}_x},$$

gdzie $\bar{x}$ jest średnią z całości, a estymator wodchylenia standardowego

$$\hat{\sigma}_x = \begin{cases} s \cdot a(N) & \text{dla odchylenia z próby} \\ \\ R \cdot b(N) & \text{dla rozstępu z próby} \end{cases}$$

Karty sum skumulowanych tworzymy dokładnie tak samo jak dla testu Page'a. Oczywiście za oszacowania parametrów można przyjąć wartości pochodzące ze wcześniejszych obserwacji.

### Karta kontrolna dla zmienności

Hawkins (1981) skonstruował kartę kontrolną dla zmienności w przypadku pojedynczych obserwacji. Metoda oparta jest na następującym spostrzeżeniu. Jeżeli $X \sim N(0, \sigma^2)$, to

$$E\left(\sqrt{\frac{|X_i|}{\sigma}}\right) = 0,82218 \quad \text{oraz} \quad V\left(\sqrt{\frac{|X_i|}{\sigma}}\right) = (0,34914)^2.$$

Dokonujemy wówczas standaryzacji

$$z_i = \frac{\sqrt{\frac{|X_i|}{\sigma}} - 0,82218}{0,34914}. \tag{17}$$

W przypadku, gdy $E(X_i) \neq 0$, tworzymy zmienne pomocnicze $\Delta_i = X_i - X_{i-1}$ ($E(\Delta_i) = 0$). Dla $z_i$ tworzymy kartę Page'a.

## Wykorzystanie testów CUSUM do wykrywania losowego dodatniego dryfu wariancji

Testy oparte na sumach skumulowanych są bardzo dobrym narzędziem do wykrywania rozregulowań powstałych na skutek „zmęczenia” badanego procesu, co odpowiada modelowi dodatniego, losowego dryfu wariancji.

Jeżeli mamy $j = 1, 2, \ldots, N$, $n$ elementowych próbek $x_{j,1}, x_{j,2}, \ldots, x_{j,n}$ pobranych z populacji o rozkładzie normalnym, to

$$s_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_{j,i} - \bar{x}_j)^2$$

jest nieobciążonym estymatorem wariancji populacyjnej $\sigma^2$. Ponadto

$$z_j = \frac{n-1}{\sigma^2}s_j^2 \sim \chi^2[n-1]$$

ma rozkład chi-kwadrat z $(n-1)$ stopniami swobody. Korzystając z faktu, iż dla $z.0$ gęstość rozkładu $\chi^2[n-1]$ ma postać

$$f(z) = \frac{1}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{n-1}{2}}}z^{\frac{n-1}{2}-1}e^{\frac{z}{2}}$$

otrzymujemy postać logarytmu funkcji wiarogodności

$$\mathcal{R}_3 = \ln\left(\frac{(\frac{1}{\sigma_1^2})^{N\frac{n-3}{2}}\exp\left(-\frac{n-1}{2\sigma_1^2}\sum_{j=1}^{N}s_j^2\right)}{(\frac{1}{\sigma_0^2})^{N\frac{n-3}{2}}\exp\left(-\frac{n-1}{2\sigma_0^2}\sum_{j=1}^{N}s_j^2\right)}\right).$$

Zatem po zaobserwowani $N$-tej próbki, obszar kontynuacji sekwencyjnego testu ilorazowego, przyjmuje postać

$$\ln(k_0) < \mathcal{R}_3 < \ln(k_1).$$

Bez większych problemów można także skonstruować zwykłą kartę CUSUM Shewharta. Zauważmy, że dla danych normalnych

$$V(s_j^2) \;=\; E(s_j^2 - \sigma^2)^2 = \frac{2\sigma^4}{n-1},$$

ponieważ zmienna losowa o rozkładzie $\chi^2[n-1]$ ma wariancję równą $2(n-1)$. Tym samym otrzymujemy postać statystyki Shewharta opartej na sumach skumulowanych

$$\mathcal{R}_4 = \frac{\sum_{j=1}^{N}s_j^2 - N\sigma_0^2}{\frac{\sigma_0^2}{\sqrt{N}}\sqrt{\frac{2}{n-1}}}.$$

## Test CUSUM oparty na ocenach alternatywnych

Użycie sekwencyjnego testu w przypadku ocen alternatywnych jest jak najbardziej naturalne. Załóżmy, że frakcja jednostek niezgodnych dla gotowego produktu wynosi $p$. Producent twierdzi, że wartość $p = p_0$. Załóżmy, że monitorujemy dostawy i kiedy frakcja jednostek niezgodnych rośnie do $p_1 > p_0$, to następuje interwencja. Załóżmy, że $j$-ta próbka ma liczność $n_j$. Wówczas statystyka testowa dla testu opartego na ilorazie wiarygodności jest równa:

$$\begin{aligned} \mathcal{R}_5 &= \ln\left(\frac{\prod_{j=1}^{N}\frac{n_j!}{x_j!(n_j-x_j)!}p_1^{x_j}(1-p_1)^{n_j-x_j}}{\prod_{j=1}^{N}\frac{n_j!}{x_j!(n_j-x_j)!}p_0^{x_j}(1-p_0)^{n_j-x_j}}\right) = \\ &= \ln\left(\frac{p_1}{p_0}\right)\sum_{j=1}^{N}x_j + \ln\left(\frac{1-p_1}{1-p_0}\right)\sum_{j=1}^{N}(n_j - x_j) \end{aligned} \tag{18}$$

Daje to nam przedział akceptacji postaci

$$\ln(k_0) < \mathcal{R}_5 < \ln(k_1). \tag{19}$$

Podobnej analizy można dokonać również przy pomocy karty kontrolnej CUSUM Shewharta. Aby ją skonstruować, zauważmy, że jeżeli w $j$-tej próbce liczba niezgodności wynosi $x_j$ przy frakcji jednostek niezgodnych $p_0$, to

$$E\Big(\sum_{j=1}^{N} x_j\Big) = p_0 \sum_{j=1}^{N} n_j.$$

Zatem statystyka testowa dla karty CUSUM Shewharta jest równa

$$\mathcal{R}_6 = \frac{\sum_{j=1}^{N} x_j - p_0 \sum_{j=1}^{N} n_j}{\sqrt{\sum_{j=1}^{N} n_j p_0 (1-p_0)}}, \tag{20}$$

a przedział akceptacji ma postać:

$$-3 < \mathcal{R}_6 < 3. \tag{21}$$