Kurs: Statystyczne sterowanie procesami produkcji (SPC)

Wykład 2 (8.10.2022)

Temat. Elementy teorii prawdopodobieństwa i statystyki matemat. w zastosowaniach SPC.

1. Wprowadzenie.

Na W1 ustaliliśmy, że metody wspomagające procesy wytwarzania to SPC. Jest ono niezbędne, aby w wyniku: monitoringu procesu wytwarzania zauważyć obecność zakłóceń tych procesów, określić ich identyfikację i w efekcie końcowym skorygować taki proces. Dalej będziemy zakładali, że ograniczymy się do tzw. zakłóceń losowych. To, jak zobaczymy dalej, wymagać będzie:

a) wprowadzenie pojęcia ZJAWISKA LOSOWEGO i jego modelu

b) opisu ilościowego tego zjawiska

c) podania metod badania tego zjawiska.

Z matematycznego punktu widzenia oznacza to, że musimy posłużyć się TP i SM.

2. Zjawisko losowe i jego model probabilistyczny.

Wszystko co może być przedmiotem obserwacji nazywamy zjawiskiem.

Oznacza to, iż z każdym zjawiskiem stowarzyszony jest co najmniej jeden jego obserwator

Z relacji zjawisko – obserwator powstaje wynik obserwacji.

Def 1 (zjawiska losowe)

Powiemy, iż Zjawisko jest LOSOWE (RANDOM), jeśli
a) obserwator jest w stanie ustalić całe spektrum wyników obserwacji
b) w chwili obserwowania zjawiska nie może z całą pewnością ustalić, z którym wynikiem ma do czynienia.

Zadanie 1. Zapoznać się z materiałem z [RR] str. 32 – 38.

FAKT 1 (model prob. ZL)

Każde ZL ma swój co najmniej jeden opis teoretyczny – tzw. model probabilistyczny Kołmogorowa (MPK) w postaci

(Ω, Σ, P), gdzie

(i) Ω jest zbiorem, którego elementami $\omega \in \Omega$ są wszystkie wyniki obserwacji ZL (jego spektrum)

Ω nazywamy przestrzenią próbkową lub przestrzenią probabilistyczną, ω – zdarzeniem elementarnym lub punktem próbkowym.

(ii) Σ jest zbiorem złożonym z pewnych podzbiorów Ω gdzie: a) $\emptyset, \Omega \in \Sigma$

b) $A \in \Sigma \Rightarrow A^c \in \Sigma$

c) $A, B \in \Sigma \Rightarrow A \cup B \in \Sigma$

Wtedy A nazywamy zdarzeniem. Dalej Σ nazywamy ciałem zdarzeń

(i) oraz (ii) stanowią opis jakościowy ZL.

Do pełnego opisu potrzebujemy kategorii realizującej opis ilościowy, czyli MIARY.

Tę rolę pełni P, czyli funkcja

$$P : \Sigma \longrightarrow [0,1], \text{ gdzie}$$

$$\Sigma \ni A \longrightarrow P(A) \in [0,1]$$

P nazywamy <u>funkcją prawdopodobieństwa</u>, liczbę $P(A)$ – prawdopodobieństwem zdarzenia A, gdy P ma własności

(i) $P(\emptyset) = 0$, $P(\Omega) = 1$

(ii) $P(A) + P(A^c) = 1$ dla k. $A \in \Sigma$

(iii) $P(A \cup B) = P(A) + P(B)$ dla $A, B \in \Sigma$ i $A \cap B = \emptyset$.

<u>Zad 2</u>. Zapoznać się z materiałem [RR] str. 13 – 37.

[Dalej będziemy zakładali, że ilekroć mowa będzie o ZL, to będziemy mieli MPK (Ω, Σ, P) i na odwrót, każdy MPK jest opisem teoretycznym co najmniej jednego ZL]

3. Pojęcie rozkładu prawdopodobieństwa i zmiennej losowej.

Opis ZL za pomocą MPK jest niezbędny, ale niewygodny z punktu widzenia przetwarzania (np. na maszynie).

Powody tego są następujące:

1°. $\omega \in \Omega$ na ogół nie jest liczbą!

2°. mamy niejednoznaczność odpowiedniości

$$ZL \longrightarrow MPK.$$

Aby wyeliminować powyższe, dla danego ZL i jego MPK (Ω, Σ, P) wprowadza się mechanizm „tłumaczenia" X każdej $\omega \in \Omega$ na liczby rzeczywiste ($\mathbb{R}$), czyli

$$\Omega \ni \omega \longrightarrow X(\omega) \in \mathbb{R},$$

gdzie zachodzi warunek

$$\forall_{t \in \mathbb{R}} \; \{\omega \in \Omega : X(\omega) < t\} \in \Sigma.$$

Wtedy X nazywamy ZMIENNĄ LOSOWĄ.

ZL $\quad \Omega \ni \omega \longrightarrow X(\omega) \in \mathbb{R}$ – liczbowy wynik obserwacji

obserwator

W T.P. rozważa się dwie sytuacje:

(i) $X(\Omega)$ (zbiór wszystkich liczbowych obserwacji 'X') jest skończony, a jeśli nie to równoliczny ze zbiorem l. naturalnych $\mathbb{N}$ (mówimy wtedy przeliczalny)

(ii) $X(\Omega)$ jest przedziałem liczbowym, być może nieograniczonym.

W sytuacji (i) mówimy, że X jest typu dyskretnego, w sytuacji (ii) typu ciągłego.

Podstawowym opisem ilościowym X jest jej funkcja dystrybuanty F, gdzie

$$\mathbb{R} \ni t \longrightarrow F(t) \stackrel{df}{=} P(\{\omega \in \Omega : X(\omega) < t\}),$$

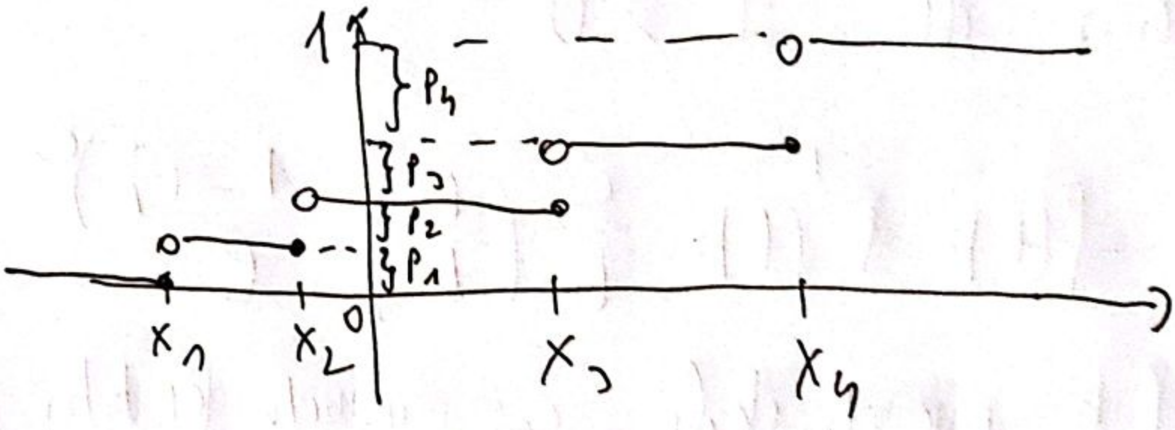
która definiuje rozkład prawdopodobieństwa X.

W sytuacji (i) mamy tzw. rozkład dyskretny, w (ii) — rozkład ciągły.

3a. Charakterystyka rozkładu dyskretnego:

Niech X będzie dyskretna, czyli dla (Ω, Σ, P) mamy $X(\Omega) = \{x_1, x_2 \dots, x_n\}$, gdzie zakładamy, że $x_1 < x_2 < \dots < x_n$.

Wtedy F ma postać (np. dla $n=4$)



Mówimy wtedy, że F jest „schodkowa".

Wtedy
$$p_1 = P(\{\omega \in \Omega : X(\omega) = x_1\})$$
$$p_2 = P(\{\omega \in \Omega : X(\omega) = x_2\})$$
$$p_3 = P(\{\omega \in \Omega : X(\omega) = x_3\})$$
$$p_4 = P(\{\omega \in \Omega : X(\omega) = x_4\}),$$
co można przedstawić tabelką

X:

x_1	x_2	x_3	x_4
p_1	p_2	p_3	p_4

I na odwrót, mając tabelę postaci

v_1	v_2	$\dots$	v_n
q_1	q_2	$\dots$	q_n

i gdy

$0 < q_j < 1$ i $q_1 + q_2 + \dots + q_n = 1$

umiemy to, i.e.

(i) istnieje ZL i jego MPK (Ω, Σ, P)

(ii) istnieje zm. losowa X, i.e.

$$\forall_{1 \le j \le n} \quad q_j = P(\{\omega \in \Omega : X(\omega) = v_j\}),$$

czyli v_j – liczbowe wyniki obserwacji ZL,

p_j – ich prawdop.

3b 3 ważne przykłady r. dyskretnych

(I) standardowy 2-punktowy

X: $\begin{array}{c|c} 0 & 1 \\ \hline q & p \end{array}$ $\quad q = 1 - p$

II $B(n,p)$ – dwumianowy Bernoulliego,

$(n \geq 2,\ p \in (0,1)$ – ustalone$)$,

gdy $X \in B(n,p)$ oznacza, iż $X(\Omega) = \{0,1,2,\dots,n\}$,

$$\forall_{0 \leq k \leq n} \quad P(\{\omega \in \Omega : X(\omega) = k\}) = \binom{n}{k} p^k (1-p)^{n-k}$$

Wtedy zdarzenie $\{\omega \in \Omega : X(\omega) = k\} = S_k$

„liczba sukcesów".

Dlatego $X \in B(n,p)$ opisuje ZL, gdy

(i) powtarzamy n-krotnie pewne doświadczenie, tak aby wyniki kolejnych obserwacji nie zależały od siebie

(ii) ~~dos~~ obserwacja pojedynczego doświadczenia daje wynik TAK = 1 = „sukces"
NIE = 0 = „porażka"

(iii) prawdopodobieństwo sukcesu $= p$

Dalej zdarzenie S_k oznacza, że mamy ~~[illegible]~~ SERIĘ

$$(b_1, b_2 \dots b_n), \text{ gdzie } b_j \in \{0,1\}$$

i liczba $\#\{j : b_j = 1\} = k$

Mając rozkład $B(n,p)$ mogę ustalić prawdop. S_k.

Zad 3. zapoznać się z materiałem [RR] ~~[illegible]~~.

iii $P(\lambda)$ – rozkład POISSONA ($\lambda > 0$ ustalone), gdy $X \in P(\lambda)$ oznacza, że $X(\Omega) = \{0,1,2,\dots\}$

oraz $\forall_{0 \le k} \; P(\{\omega \in \Omega : X(\omega) = k\}) = e^{-\lambda} \frac{\lambda^k}{k!}$

Analiza procesu za pomocą $X \in P(\lambda)$ pozwala odpowiedzieć na pytanie, czy określona liczba niezgodności na jednostkę została przekroczona i z jakim prawdop. (Np. liczba wad izolacji, liczba zarysowań karoserii samochodu itp.).

Zad 4. Analiza listy zadań #1 (strona www wykładowcy)

3c) 2 ważne rozkłady ciągłe.

I Rozkład gaussowski (normalny, centralny)

$X \in N(m, \sigma)$, gdy $m \in \mathbb{R}$, $\sigma > 0$,

oznacza, że

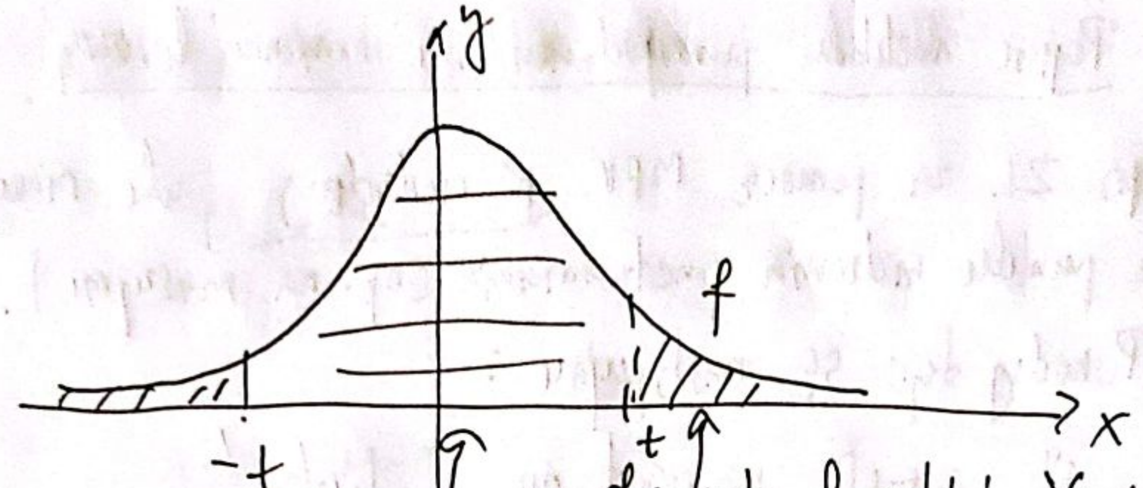
$$f(t) = F'(t) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(t-m)^2}{2\sigma^2}}, \quad t \in \mathbb{R}$$

Wtedy $f = F'$ – funkcja gęstości rozkładu X

Jeśli $m = 0$, $\sigma = 1$, to mówimy, że X ma standardowy r. normalny i $F = \Phi$.

Wiadomo, że $X \in N(m, \sigma)$, to $\frac{X-m}{\sigma} \in N(0,1)$ (procedura standaryzacji), czyli każdy $X \in N(m,\sigma)$ można sprowadzić do $N(0,1)$.

Niech dalej $X \in N(0,1)$. Wtedy mamy:



f: wykres f. gęstości $X \sim N(0,1)$

tzw. „krzywa dzwonowa"

Wtedy obszar ograniczony f i OX ma pole = 1,)

$$\int_{-\infty}^{+\infty} f = 1$$

Dla $t>0$, obszar

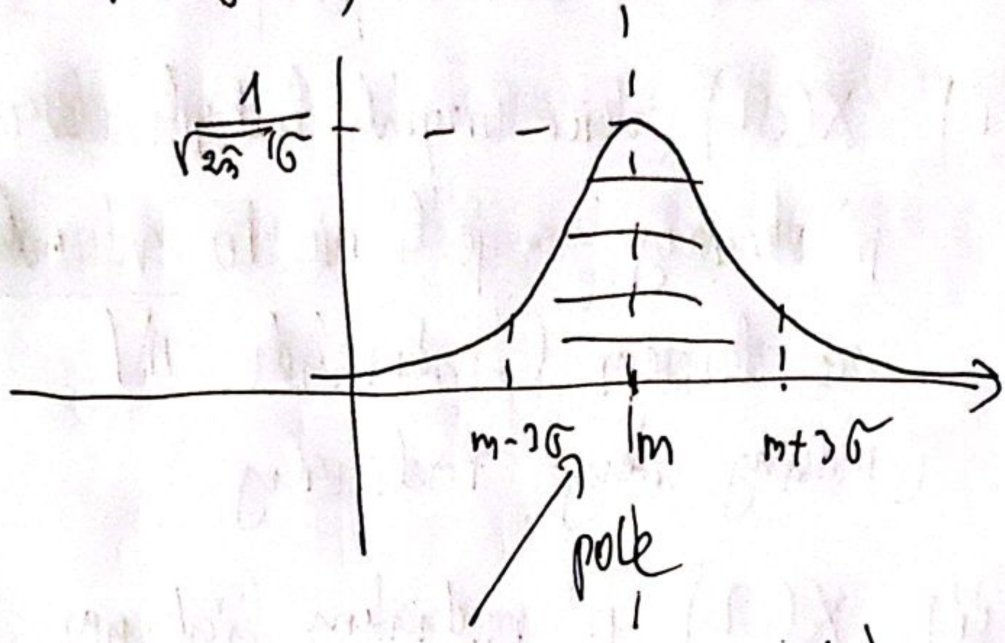
$$1-\Phi(t) = P(\{\omega\in\Omega : X(\omega)\geq t\}) = P(\{\omega\in\Omega : X(\omega)< -t\})$$
$$= \Phi(-t)$$

Dla obszaru

$$P(\{\omega\in\Omega : -t\leq X(\omega)< t\})$$
$$= \Phi(t) - \Phi(-t) = 2\Phi(t)-1$$

Dla $t=3$, pole tego obszaru $\approx$ 99,7% całości.

W sytuacji ogólnej, $X \in N(m, \sigma)$ mamy



$$P(\{\omega \in \Omega : m-3\sigma \leq X(\omega) < m+3\sigma\}) \approx 0{,}9973$$

(tzw. reguła „3σ").

zadkie: Jeśli $\omega_0 \in \{\omega \in \Omega : X(\omega) \geq m+3\sigma\}$, to mówimy, że „skrajnie duże", $\omega_0 \in \{\omega \in \Omega : X(\omega) < m-3\sigma\}$ „skrajnie małe".

<u>Problem</u> Co oznacza m i σ?

II Rozkład Weibulla.

$X \in W(b,c)$, $b>0$, $c>0$ dane, jeśli

$X(\Omega) = \mathbb{R}_+$ (wyniki liczbowe obserwacji >0)

~~oraz~~ czyli $F(t)=0$, $t \le 0$,

oraz dla $t>0$

$$f(t) = F'(t) = \frac{c}{b}\left(\frac{t}{b}\right)^{c-1} e^{-\left(\frac{t}{b}\right)^c}$$

(b - parametr skali, c - parametr kształtu)

X opisuje czas poprawnej pracy obiektu (systemu, procesu) do chwili wystąpienia wady (defektu).

Uwaga.

Rozkłady ciągłe są stablicowane. Patrz zadanie 1 na www wykładowcy.