Materiał do wykładu PMRiSt, 26.11.2020

Część 2

Temat. Znaczenie wartości oczekiwanej i wariancji w T.P.

Punktem wyjścia do rozumowania wyjaśniającego znaczenie EX i $var(X)$ będzie udowodniona n. Czebyszewa (⊛⊛),

czyli: Niech $X \neq const.$ z 2-m momentem.

Wtedy $\forall t > 0$

$$\boxed{P(\{\omega \in \Omega : |X(\omega) - m| \geq t\sigma\}) \leq \frac{1}{t^2}, \quad m = EX, \quad \sigma^2 = var(X), \quad \sigma = \sqrt{\sigma^2}}$$

Zanim to zrobimy weźmy przykład.

P1. Niech $X \in N(m, \sigma^2)$. Z zasady standaryzacji wiemy, że $\mathcal{N} = \frac{X - m}{\sigma} \in N(0, 1)$, dlatego zastosowanie N. Czebyszewa daje

$$P(\{\omega \in \Omega : |\mathcal{N}(\omega)| \geq t\}) \leq \frac{1}{t^2}$$

Biorąc $t = 3$ mamy:

$$P(\{\omega \in \Omega : |X(\omega) - m| \geq 3\sigma\}) = P(\{\omega \in \Omega : |\mathcal{N}(\omega)| \geq 3\}) \leq \frac{1}{9} \approx 0{,}111$$

Ale można zrobić to dokładniej:

$$P(\{v\in\Omega: |N(v)|\geq 3\}) = 1-P(\{v\in\Omega: |N(v)|<3\})$$

$$= 1-(\Phi(3)-\Phi(-3)) = 1-(\Phi(3)-(1-\Phi(3)))$$

$$= 1-(2\Phi(3)-1) = 2(1-\Phi(3))$$

Z tablicy $N(0,1)$ mamy:

$$\Phi(3) > 0{,}9986, \quad \text{dlatego}$$

$$P(\{v\in\Omega: |X(v)-m|\geq 3\sigma\}) = 2(1-\Phi(3))$$

$$< 2(1-0{,}9986) = 0{,}0028$$

Wynik uzyskany nazywamy ZASADA 3σ

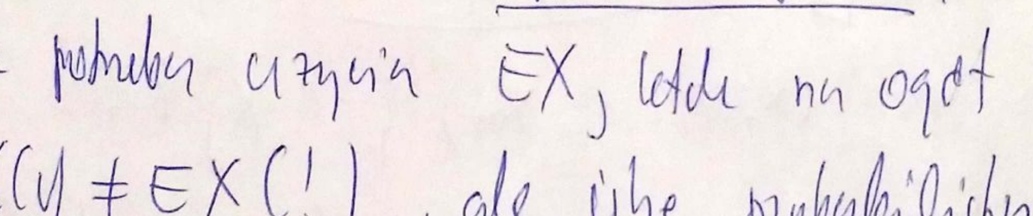
$$\boxed{\forall_{X\in N(m,\sigma^2)} \; P(\{v\in\Omega: |X(v)-m|\geq 3\sigma\}) < 0{,}002}$$

Wracamy do głównego wątku.

W cz. 1 sformułowały tzw. Stochastyczną ZASADĘ
(S.Z.N) NIEOZNACZONOŚCI (*), która mówi, że dla zm. losowej $X \neq const$ o 2-m momencie

$$\forall_{\varepsilon>0} \exists_{\delta>0} \; P(\{\omega\in\Omega: |X(\omega) - EX| < \delta\}) \geq 1-\varepsilon,$$

która uzasadnia, że w T.P. $X(\omega)$ istnieje tylko teoretycznie, w praktyce „jest nieosiągalne”, dlatego istnieje potrzeba użycia EX, która na ogół $X(\omega) \neq EX$ (!), ale istnieje probabilistyczna kontrola rozproszenia $X(\omega)$ względem EX



Według S.Z.N.

Zauważmy, że S.Z.N. jest konsekwencją N.Cz. (***)

Najpierw zauważmy, że (***) $\Leftrightarrow$

(****) $P(\{\omega \in \Omega : |X(\omega) - m| < t\sigma\}) \geq 1 - \frac{1}{t^2}$

Wystarczy przyjąć, że w (****)

$$\varepsilon = \frac{1}{t^2} \ (t>0), \quad \delta = t\sigma$$

Wtedy prawdą jest, że dla dowolnego $\varepsilon > 0$ (bo $t>0$) istnieje $\delta = t\sigma$ ($\sigma = \sqrt{\sigma^2}$), że mamy (***), a więc S.Z.N.

Z tego powodu $\sigma = \sqrt{var(X)}$ nazywamy rozproszeniem rozkładu

Podsumowanie.

Wynikiem obserwacji Z.L. (a S.Z.N. należy oczekiwać, iż istotny nadmiar i spełnione założenia n. Czebyszewa) o zmiennej X, jest wartość średnia

$m = EX$. Jej znaczenie w sytuacji kiedy teoretyczny wynik obserwacji $X(\omega_0) \neq EX$

polega na tym, iż z dowolnie bliskim liczbie 1 prawdopodobieństwem (bo co najmniej $= 1-\varepsilon$, $\varepsilon > 0$) możemy kontrolować rozproszenie $X(\omega_0)$ względem EX, czyli $|X(\omega_0) - EX|$, bowiem

z prawd. $> 1-\varepsilon$, $X(\omega_0) \in (EX - t\sigma, EX + t\sigma)$,

gdzie $\sigma = \sigma_X = \sqrt{\sigma_X^2} = \sqrt{var(X)}$,

$var(X) = E(X-m)^2$ „średnie kwadratowe odchylenie”

Na koniec wstępu jeszcze jeden przykład

P.2. Niech X oznacza liczbę sukcesów w n niezal. próbach z prawd. sukcesu w pojedynczej próbie p.

Dla $n=100$, $p=0{,}2$ obliczamy

$$pr = \mathbb{P}(\underbrace{\{\omega\in\Omega : 8 < X(\omega) < 32\}}_{A}).$$

Uwaga.

pr można obliczyć wzorem (dlaczego?)

$$pr = \sum_{k=9}^{31} \binom{100}{k} 0{,}2^{k}\, 0{,}8^{31-k}.$$

Zad. Napisać procedurę pozwalającą obliczyć pr.

Oszacowanie polega na zastosowaniu nier. Czebyszewa.

Z założenia $X \in B(100, 0{,}2)$.

Stąd $EX = n\cdot p = 20$,

$$\sigma^2 = npq = 100\cdot 0{,}2\cdot 0{,}8 = 16$$

$$\sigma = 4$$

Chcemy zastosować nier. Czeb.:

$t > 0$

$$\underbrace{P(\{\omega \in \Omega : |X(\omega) - m| \geq t\sigma\}) < 1/t^2}$$

Zauważmy, że dla zdarzenia A mamy:

$$A = \{\omega \in \Omega : 8-20 < X(\omega) - m < 32-20\}$$

$$= \{\omega \in \Omega : -12 < X(\omega) - m < 12\}$$

$$= \{\omega \in \Omega : |X(\omega) - m| < 12\}$$

Dalej, ponieważ $12 = 3 \cdot 4 = 3 \cdot \sigma$ $(t = 3)$

oraz

$$P(A) = 1 - P(A^c) = 1 - P(\{\omega \in \Omega : |X(\omega) - m| \geq 12\})$$

$$= 1 - P(\{\omega \in \Omega : |X(\omega) - m| \geq 3\sigma\}) \mid c'$$

z N. Cz. $P(\{\omega \in \Omega : |X(\omega) - m| \geq 3\sigma\}) < 1/3^2$,

mamy $P(A) \geq 1 - 1/9 = 8/9 \approx 0{,}9$.

Rozważania dotyczące roli EX w T.P. zakończmy pewną uwagą, która w przyszłości będzie przedmiotem analizy.

Problem „NIEPEWNOŚCI POMIARU"

Ilekroć podejmujemy „coś" zmierzyć, szczególnie wtedy kiedy robimy to po raz pierwszy, „aby upewnić się" pomiar powtarzamy kilkakrotnie.

Dostajemy tzw. serię pomiarów:

$(x_1, x_2, \ldots, x_n)$, $x_j \in \mathbb{R}$.

Wtedy za wynik pomiaru bierzemy średnią z takiej serii.

No tak, dlaczego jaką? (są różne!) ora

Czy ma to uzasadnienie?

Odpowiedzi (rozwiązaniem) na to pytanie są T.P. i treść dzisiejszego wykładu!

Np. biorąc średnią arytmetyczną

$\frac{1}{n}(X_1 + \ldots + X_n)$ bierzemy EX, dla

zm. losowej symetrycznej.

Czasami (KIEDY?) istnieje potrzeba użycia

innej średniej — np. WAŻONEJ, gdy

$$\sum_{j=1}^{n} p_j X_j \quad \text{dla tzw}$$

$$d_X \quad \begin{array}{c|c|c|c} x_1 & x_2 & \ldots & x_n \\ \hline p_1 & p_2 & & p_n \end{array}.$$

<u>Zad</u>.

Napisz ~~~~ S.Z N dla tego

przykładu.