

2° z całkowalności funkcji  $g_k(x)$  wynika na mocy własności (c29) całkowalność funkcji  $|x|^j$ , gdyż

$$|x|^j \leq g_k(x)$$

Wobec powyższego z istnienia momentu skończonego  $m_{k+1}$  wynika istnienie skończonych momentów  $m_1, \dots, m_k$  i

$$(221) \quad h_k(t) = \frac{d^k h_0(t)}{dt^k} = \frac{d^k \varphi(t)}{dt^k} = \int_{-\infty}^{+\infty} (ix)^k e^{itx} dP(x)$$

Ponieważ z definicji

$$h_k(0) = i^k m_k$$

wynika stąd żądana własność.

### § 238. Informacja i entropia

Teoria informacji wywodzi się z zagadnień praktycznych telekomunikacji. Jest to obecnie dział matematyki zyskujący stale na znaczeniu i mający duże powiązania z rachunkiem prawdopodobieństwa i statystyką matematyczną. Wprowadzenie teorii informacji do probabilistyki przyniosło nie tylko nowe spojrzenie na znane zagadnienia, ale spowodowało wykrycie nowych praw i konstrukcje nowych metod. Z tego względu jest pożądaną, by współczesny wykład probabilistyki obejmował również elementy teorii informacji.

W teorii informacji za jednostkę informacji przyjmuje się bit, czyli cyfrę binarną (skrót nazwy angielskiej "binary digit"). Jest to naturalne, ponieważ za elementarną informację można uważać odpowiedź na jedno pytanie, wyrażoną słowami "tak" lub "nie", którym to słowom można wzajemnie jednoznacznie przyporządkować cyfry binarne 0 lub 1. Jeśli przyjąć, że alfabet języka polskiego zawiera 32 znaki, to każdy z tych znaków może być wzajemnie jednoznacznie przyporządkowany jednej z liczb od 0 do 31 zapisanych w systemie binarnym 5 bitami od 00000 do 11111. W ten sposób informacja przesyłana jednym znakiem alfabetu może być traktowana jako informacja złożona z 5 kolejnych bitów. Za pośrednictwem znaków alfabetu każdy tekst może być traktowany jako informacja zło-



żona ze skończonego ciągu bitów. Informacja przesyłana obrazem telewizyjnym może być traktowana również jako ciąg bitów, jeśli obraz traktujemy jako układ skończonej liczby punktów, w których może występować sygnał świetlny lub nie, co znowu odpowiada cyfrom binarnym 0 lub 1.

Zauważmy, że uzyskanie określonej informacji likwiduje jakiś stan naszej niewiedzy czy niepewności. W teorii informacji wprowadza się pojęcie entropii, jako wielkości wyrażającej stopień naszej niewiedzy w danym zagadnieniu. Entropia mierzy się ilością informacji potrzebnej do pełnej likwidacji naszej niewiedzy. Uzyskana informacja jest równa ubytkowi entropii. Wynika stąd, że entropię mierzy się też w bitach. Na przykład entropia wyboru jednego spośród 8 elementów jest równa 3 bity, ponieważ wyboru można dokonać uzyskując kolejno następujące jednostkowe informacje:

1<sup>o</sup> Czy dany element należy do wybranych dowolnie 4 elementów?

Jeśli "tak", to już wiemy, że dany element należy wybrać spośród tych 4 elementów, jeśli "nie", to należy go wybrać z elementów pozostałych.

2<sup>o</sup> Czy dany element należy do 2 elementów wybranych dowolnie z wyselekcjonowanej poprzednio czwórki elementów?

Jeśli "tak", to już wiemy, że dany element należy wybrać spośród tej pary elementów, jeśli "nie", to z pary pozostałej.

3<sup>o</sup> Czy dany element jest wskazanym elementem wybranej poprzednio pary?

Jeśli "tak", to element został już znaleziony, jeśli "nie", to elementem szukanym jest pozostały element danej pary.

Jak łatwo zauważyć, entropia wyboru jednego spośród  $2^n$  elementów jest równa  $n$  bitów.

Pozornie mogłoby się wydawać, że za entropię wyboru jednego spośród  $N$  elementów w przypadku  $2^n < N < 2^{n+1}$  powinniśmy przyjąć liczbę  $n+1$  bitów, gdyż tyle pytań jednostkowych wystarcza, aby wyboru dokonać. Można jednak to zagadnienie potraktować nieco inaczej, biorąc pod uwagę, że w pewnych przypadkach wystarcza mniejsza liczba pytań jednostkowych. Na przykład, dla wybrania elementu spośród  $N = 5$  elementów, stawiamy pytanie, czy dany element znajduje się w wybranej dowolnie trójce elementów. Jeśli odpowiedź będzie "nie", to szukany element znajduje się w pozosta-



tej parze elementów i dla jego znalezienia wystarczy jeszcze tylko jedno pytanie jednostkowe. Zatem wyboru dokonujemy za pomocą tylko 2 pytań jednostkowych. Zastanówmy się zatem, ile średnio trzeba pytań jednostkowych, aby spośród 5 elementów wybrać jeden. Przyjmując, że prawdopodobieństwo znalezienia się danego elementu w wybranej trójce jest równe  $\frac{3}{5}$ , a w przypadku, gdy dany element znajdzie się w tej trójce, prawdopodobieństwo znalezienia się jego w wybranej parze elementów jest równe  $\frac{2}{3}$ , otrzymujemy następującą średnią liczbę pytań jednostkowych:

$$\frac{3}{5} \cdot \frac{2}{3} \cdot 3 + \frac{3}{5} \cdot \frac{1}{3} \cdot 2 + \frac{2}{5} \cdot 2 = \frac{12}{5}$$

Bylibyśmy zatem raczej skłonni wybrać liczbę  $\frac{12}{5}$  bitów jako entropię wyboru jednego spośród 5 elementów, ale i takie wyjście nie byłoby w pełni zadowalające. Jeśli bowiem mówi się o średniej liczbie pytań jednostkowych, to ma się na myśli jakieś serie przypadków, w których z  $N$  elementów wybieramy 1 element. Rozpatrzmy serię  $k$  takich przypadków. Jeśli za pierwszym razem szukamy elementu  $e_1$ , za drugim  $e_2, \dots$ , za  $k$ -ym elementu  $e_k$ , to zagadnienie sprowadza się do wyszukania 1 ciągu  $(e_1, \dots, e_k)$  spośród  $N^k$  możliwych ciągów. Ale do tego celu wystarczy  $n_k$  pytań jednostkowych, gdzie  $n_k$  jest taką liczbą, że

$$2^{n_k-1} < N^k \leq 2^{n_k}$$

czyli

$$k \log_2 N \leq n_k < k \log_2 N + 1$$

Zatem średnio na jeden przypadek, w którym wybieramy 1 element spośród  $N$  elementów, przypada

$$\alpha = \frac{n_k}{k}$$

pytań jednostkowych, gdzie

$$\log_2 N \leq \alpha < \log_2 N + \frac{1}{k}$$

Ze względu na dowolność liczby naturalnej  $k$  przyjmuje się, że entropia wyboru 1 elementu spośród  $N$  elementów jest równa

$$(i) \quad \alpha = \log_2 N$$

Ogólniej powiemy, że entropia układu  $N$  jednakowo prawdopodobnych możliwości jest dana wzorem (i). Wzór ten nosi nazwę wzoru Hartley'a.

Entropią zbioru będziemy nazywać entropię wyboru 1 elementu z tego zbioru.

Niech teraz będzie dany zbiór złożony z  $K$  grup po  $L$  elementów każda. Zatem dany zbiór składa się z

$$N = KL$$

elementów. Jego entropia jest równa

$$(ii) \quad \log_2 N = \log_2 K + \log_2 L$$

Widzimy, że entropia wyboru elementu z całego zbioru jest sumą entropii wyboru grupy i entropii wyboru wewnątrz grupy.

Uogólnimy teraz pojęcie entropii na przypadki, gdy grupy nie są jednakowo liczne, w ten sposób, aby utrzymać w mocy prawo (ii) w postaci

$$(iii) \quad \log_2 N = E_1 + E_2$$

gdzie  $E_1$  jest entropią wyboru między grupami, a  $E_2$  średnią entropią wyboru wewnątrz grup. Niech  $N_k$  oznacza liczebność  $k$ -tej grupy,  $k = 1, \dots, K$ ,  $N = N_1 + \dots + N_K$ . Wtedy entropia wyboru wewnątrz  $k$ -tej grupy jest równa  $\log_2 N_k$ , a średnią entropię wyboru wewnątrz grup określimy wzorem

$$(iv) \quad E_2 = \sum_{k=1}^K \frac{N_k}{N} \log_2 N_k$$

Wprowadzając prawdopodobieństwa poszczególnych grup jako

$$p_k = \frac{N_k}{N}, \quad k = 1, \dots, K$$

możemy wzór (iv) napisać w postaci

$$E_2 = \sum_{k=1}^K p_k \log_2 (N p_k) = \sum_{k=1}^K p_k \log_2 p_k + \log_2 N \cdot \sum_{k=1}^K p_k =$$



$$= \sum_{k=1}^K p_k \log_2 p_k + \log_2 N$$

Wstawiając tę wartość do wzoru (iii) otrzymujemy

$$E_1 = - \sum_{k=1}^K p_k \log_2 p_k$$

Zgodnie z tym wzorem definiujemy entropię zbioru co najwyżej przeliczalnego  $\{e_1, e_2, \dots\}$ , dla którego  $p_k$  oznacza prawdopodobieństwo elementu  $e_k$ ,  $k = 1, 2, \dots$  i  $p_1 + p_2 + \dots = 1$ , jako

$$(222) \quad \mathcal{E} \stackrel{\text{def}}{=} - \sum_{k=1}^n p_k \log_2 p_k, \quad n \in \mathbb{N} \vee n = \infty$$

Jest to tzw. wzór Shannona.

W przypadku, gdy zbiór składa się z elementów  $e_{kl}$ , gdzie  $k$  jest numerem grupy, a  $l$  numerem elementów w grupie,  $k = 1, \dots, K$ ,  $K \in \mathbb{N} \vee K = \infty$ ,  $l = 1, \dots, n_k$ ,  $n_k \in \mathbb{N} \vee n_k = \infty$ , a prawdopodobieństwo elementu  $e_{kl}$  jest - zgodnie ze wzorem (88) z § 142 - równe

$$p_{kl} = p_k \pi_{kl}$$

gdzie  $p_k$  oznacza prawdopodobieństwo  $k$ -ej grupy, a  $\pi_{kl}$  prawdopodobieństwo  $l$ -go elementu wewnątrz  $k$ -ej grupy, mamy na mocy (222)

$$\begin{aligned} \mathcal{E} &= - \sum_{k=1}^K \sum_{l=1}^{n_k} p_{kl} \log_2 p_{kl} = - \sum_{k=1}^K \sum_{l=1}^{n_k} p_k \pi_{kl} \log_2 (p_k \pi_{kl}) = \\ &= - \sum_{k=1}^K \sum_{l=1}^{n_k} p_k \pi_{kl} \log_2 p_k - \sum_{k=1}^K \sum_{l=1}^{n_k} p_k \pi_{kl} \log_2 \pi_{kl} = \\ &= - \sum_{k=1}^K p_k \log_2 p_k \sum_{l=1}^{n_k} \pi_{kl} - \sum_{k=1}^K p_k \left( \sum_{l=1}^{n_k} \pi_{kl} \log_2 \pi_{kl} \right) = \end{aligned}$$

$$= - \sum_{k=1}^K p_k \log_2 p_k + \sum_{k=1}^K p_k \left( - \sum_{l=1}^{n_k} \pi_{kl} \log_2 \pi_{kl} \right)$$

Otrzymaliśmy zatem wzór

$$(223) \quad \mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2$$

gdzie

$$\mathcal{E} = - \sum_{k=1}^K \sum_{l=1}^{n_k} p_{kl} \log_2 p_{kl}$$

jest entropią całego zbioru,

$$\mathcal{E}_1 = - \sum_{k=1}^K p_k \log_2 p_k$$

entropią zbioru grup, czyli entropią międzygrupową, a

$$\mathcal{E}_2 = \sum_{k=1}^K p_k E_k$$

średnią entropią wewnątrzgrupową, gdzie

$$E_k = - \sum_{l=1}^{n_k} \pi_{kl} \log_2 \pi_{kl}$$

oznacza entropię  $k$ -ej grupy.

Wzór (223) jest uogólnieniem wzorów (ii) i (iii) i wyraża prawo dodawania informacji.

Jak wynika ze wzoru (222), entropia tym wzorem określona zawsze istnieje i jest nieujemna, ponieważ

$$\bigwedge_{k \in \mathfrak{K}} (0 \leq p_k \leq 1 \Rightarrow -p_k \log_2 p_k \geq 0)$$

Zgodnie z umową z § 8 jest tu w przypadku  $p_k = 0$

$$0 \cdot \log_2 0 = 0 \cdot (-\infty) = 0$$



## Twierdzenie (A)

Entropia (222) jest równa zeru wtedy i tylko wtedy, gdy jeden z elementów danego zbioru ma prawdopodobieństwo równe 1, a pozostałe mają prawdopodobieństwa równe 0.

D Ponieważ

$$\bigwedge_{k \in \mathfrak{N}} p_k \log_2 p_k \geq 0$$

więc entropia (222) jest zerem wtedy i tylko wtedy, gdy

$$\bigwedge_{k \in \mathfrak{N}} p_k \log_2 p_k = 0$$

czyli

$$\bigwedge_{k \in \mathfrak{N}} p_k = 0 \vee p_k = 1$$

Ponieważ

$$p_1 + p_2 + \dots = 1$$

otrzymujemy stąd tezę twierdzenia.

Twierdzenie (A) jest zgodne z intuicją, ponieważ w przypadku, gdy jeden z elementów ma prawdopodobieństwo 1 a pozostałe 0, uważamy, że informacja jest pełna, czyli niepewność jest równa 0.

## Twierdzenie (B)

Entropia (222) osiąga przy ustalonym  $n \in \mathfrak{N}$  wartość największą wtedy i tylko wtedy, gdy

$$p_1 = \dots = p_n = \frac{1}{n}$$

i jej wartością maksymalną jest  $\log_2 n$ .

D Maksimum entropii (222) poszukamy jako maksimum funkcji  $n$  zmiennych

$$\tau = \tau(p_1, \dots, p_n) = - \sum_{k=1}^n p_k \log_2 p_k$$

z warunkiem ubocznym

$$p_1 + \dots + p_n = 1$$

stosując metodę mnożników Lagrange'a, czyli szukając maksimum funkcji

$$\vartheta = \vartheta(p_1, \dots, p_n, \lambda) = - \sum_{k=1}^n p_k \log_2 p_k + \lambda \left( \sum_{k=1}^n p_k - 1 \right)$$

Mamy

$$\bigwedge_{k \in \{1, \dots, n\}} \frac{\partial \vartheta}{\partial p_k} = -\log_2 p_k - 1 + \lambda = 0$$

$$\frac{\partial \vartheta}{\partial \lambda} = \sum_{k=1}^n p_k - 1 = 0$$

skąd

$$\bigwedge_{k \in \{1, \dots, n\}} \log_2 p_k = \lambda - 1$$

i w konsekwencji

$$(v) \quad p_1 = \dots = p_n = \frac{1}{n}$$

Ze względu na to, że funkcja  $\tilde{\epsilon}$  jest regularna, nieujemna i na brzegu obszaru swej określoności osiąga na mocy twierdzenia (A) wartość 0, więc musi mieć maksimum, a ponieważ może je mieć tylko w punkcie (v), więc ma je w tym punkcie.

Ze wzorów (v) i (222) wynika, że wartością maksymalną entropii jest

$$(vi) \quad \tilde{\epsilon} = - \sum_{k=1}^n \frac{1}{n} \log_2 \frac{1}{n} = \log_2 n \sum_{k=1}^n \frac{1}{n} = \log_2 n$$

zgodnie ze wzorem (i).

Dyskretny rozkład prawdopodobieństwa można traktować jako co najwyżej przeliczalny zbiór wartości  $\{x_1, x_2, \dots\}$ , którym są przyporządkowane prawdopodobieństwa odpowiednio  $p_1, p_2, \dots, p_1 + p_2 + \dots = 1$ . Wobec tego wzór (222) będziemy również uważać za definicję entropii dyskretnego rozkładu prawdopodobieństwa. Na mocy wzoru (180) możemy entropię (222) uważać za wartość średnią funkcji



$$(224) \quad h(x) = -\log_2 p(x)$$

gdzie  $p(x)$  oznacza prawdopodobieństwo elementu  $x$  i

$$p_k \stackrel{\text{def}}{=} p(x_k)$$

Wykażemy jeszcze, że entropia (222) może przybierać wartość nieskończoną. W tym celu udowodnimy najpierw następujący lemat.

Lemat

$$\sum_{k=2}^{\infty} \frac{1}{k \log_2 k} = \infty \wedge \sum_{k=2}^{\infty} \frac{1}{k \log_2^2 k} < \infty$$

Dowód

$$\begin{aligned} \sum_{k=2}^M \frac{1}{k \log_2 k} &= \sum_{m=1}^M \sum_{n=10^{m-1}}^{10^m-1} \sum_{k=2^n}^{2^{n+1}-1} \frac{1}{k \log_2 k} > \\ &> \sum_{m=1}^M \sum_{n=10^{m-1}}^{10^m-1} \sum_{k=2^n}^{2^{n+1}-1} \frac{1}{2^{n+1} \cdot \log_2 2^{n+1}} = \sum_{m=1}^M \sum_{n=10^{m-1}}^{10^m-1} \frac{1}{2^{(n+1)} \log_2 2^{(n+1)}} > \\ &> \sum_{m=1}^M \sum_{n=10^{m-1}}^{10^m-1} \frac{1}{2 \cdot 10^m} = \sum_{m=1}^M \frac{9}{20} = \frac{9}{20} M \xrightarrow{M \rightarrow \infty} \infty \end{aligned}$$

Analogicznie

$$\begin{aligned} \sum_{k=2}^{\infty} \frac{1}{k \log_2 k} &= \sum_{n=1}^{\infty} \sum_{k=2^n}^{2^{n+1}-1} \frac{1}{k \log_2 k} < \sum_{n=1}^{\infty} \sum_{k=2^n}^{2^{n+1}-1} \frac{1}{2^n \log_2(2^n)} = \\ &= \sum_{n=1}^{\infty} \sum_{k=2^n}^{2^{n+1}-1} \frac{1}{n \cdot 2^n} = \sum_{n=1}^{\infty} \frac{1}{n \cdot 2^n} < 1 + \sum_{n=2}^{\infty} \frac{1}{n(n-1)} = \\ &= 1 + \sum_{n=2}^{\infty} \left( \frac{1}{n-1} - \frac{1}{n} \right) = 1 + \left( 1 - \frac{1}{2} \right) + \left( \frac{1}{2} - \frac{1}{3} \right) + \left( \frac{1}{3} - \frac{1}{4} \right) + \dots = 2 \end{aligned}$$

W ten sposób lemat został udowodniony.

Podamy teraz przykład dyskretnego rozkładu prawdopodobieństwa, dla którego entropia będzie nieskończona.

Niech na mocy udowodnionego lematu

$$(vii) \quad s = \sum_{k=2}^{\infty} \frac{1}{k \log_2^2 k}$$

i niech

$$(viii) \quad \bigwedge_{k \in \mathbb{N}} p_k \stackrel{\text{def}}{=} \frac{1}{s(k+1) \log_2^2(k+1)}$$

Ze wzoru (vii) wynika, że jest spełniony warunek

$$\sum_{k=1}^{\infty} p_k = 1$$

Obliczmy entropię, gdy dla danego rozkładu zachodzą równości (viii).

$$\begin{aligned} \mathcal{E} &= - \sum_{k=1}^{\infty} \frac{1}{s(k+1) \log_2^2(k+1)} \log_2 \frac{1}{s(k+1) \log_2^2(k+1)} = \\ &= \sum_{k=2}^{\infty} \frac{\log_2 (sk \log_2^2 k)}{sk \log_2^2 k} > \frac{1}{s} \sum_{k=2}^{\infty} \frac{\log_2 (sk)}{k \log_2^2 k} = \\ &= \frac{\log_2 s}{s} \sum_{k=2}^{\infty} \frac{1}{k \log_2^2 k} + \frac{1}{s} \sum_{k=2}^{\infty} \frac{1}{k \log_2 k} = \\ &= \log_2 s + \frac{1}{s} \sum_{k=2}^{\infty} \frac{1}{k \log_2 k} \end{aligned}$$

Ponieważ ostatnia suma na mocy lematu ma wartość nieskończoną, więc i entropia jest nieskończona.

Gdybyśmy pojęcie entropii rozkładu prawdopodobieństwa jako wartości średniej dla funkcji (224) chcieli przedłużyć na rozkłady



ciągłe, otrzymalibyśmy dla wszystkich tych rozkładów entropię nieskończoną. W tej sytuacji definiuje się entropię ciągłego rozkładu prawdopodobieństwa podobnie, ale inaczej, a mianowicie jako wartość średnią funkcji

$$(225) \quad h(x) = -\log_2 f(x)$$

gdzie  $f$  jest gęstością prawdopodobieństwa danego rozkładu, skąd

$$(226) \quad \mathcal{E} \stackrel{\text{def}}{=} - \int_{-\infty}^{+\infty} f(x) \log_2 f(x) dx$$

O ile entropia (222) dyskretnego rozkładu prawdopodobieństwa istniała zawsze i była nieujemna, choć – jak widzieliśmy – mogła być nieskończona, o tyle entropia (226) może dla niektórych rozkładów ciągłych nie istnieć, a jeśli istnieje, może przyjmować wartości skończone lub nieskończone zarówno dodatnie jak i ujemne, z uwagi na to, że

$$f(x) > 1 \Rightarrow f(x) \log_2 f(x) > 0$$

$$0 < f(x) < 1 \Rightarrow f(x) \log_2 f(x) < 0$$

Na przykład entropia rozkładu o gęstości

$$f(x) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{dla } x < 0 \\ \frac{1}{c} & \text{dla } 0 < x < c \\ 0 & \text{dla } x > c \end{cases}$$

gdzie  $c$  jest dowolną liczbą rzeczywistą dodatnią, jest równa

$$\mathcal{E} = - \int_0^c \frac{1}{c} \log_2 \frac{1}{c} dx = \frac{1}{c} \log_2 c \cdot \int_0^c dx = \log_2 c$$

i jest dodatnia dla  $c > 1$ , równa 0 dla  $c = 1$  i ujemna dla  $0 < c < 1$ .

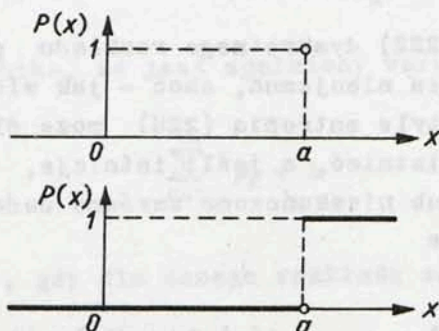
## § 239. Rozkład jednopunktowy

Rozkładem jednopunktowym  $\mathcal{P}$  nazywamy dyskretny rozkład prawdopodobieństwa na prostej, w którym istnieje taki punkt  $a \in \mathbb{R}$ , że

$$\mathcal{P}(\{a\}) = 1 \wedge \mathcal{P}(\mathbb{R} - \{a\}) = 0$$

Dystrybuanta  $F$  takiego rozkładu ma postać

$$F(x) = \begin{cases} 0 & \text{dla } x \leq a \\ 1 & \text{dla } x > a \end{cases}$$



Mamy dalej na mocy (192)

$$\bigwedge_{k \in \mathbb{N}} m_k = 1 \cdot a^k = a^k$$

W szczególności

$$\bar{x} = m_1 = a$$

Jest dalej

$$(*) \quad \bigwedge_{k \in \mathbb{N}} \mu_k = 1 \cdot (a - \bar{x})^k = 0$$

W szczególności wariancja

$$\mu_2 = 0$$

i standardowe odchylenie

$$s = \sqrt{\mu_2} = 0$$



Z uwagi na (\*) współczynniki asymetrii i spłaszczenia nie istnieją.

Funkcja charakterystyczna

$$\varphi(t) = 1 \cdot e^{iat} = e^{iat}$$

Entropia

$$\mathcal{E} = 1 \log_2 1 = 0$$

#### § 240. Rozkład równomierny

Rozkładem równomiernym będziemy nazywać rozkład dyskretny, dla którego

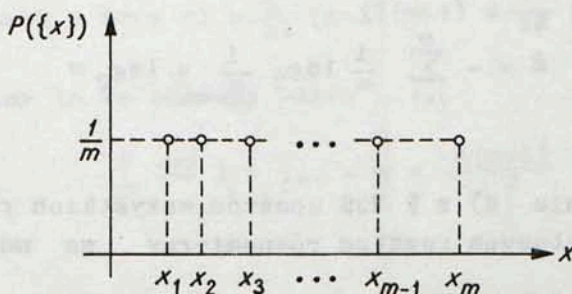
$$p_k = P(\{x_k\}) \stackrel{\text{def}}{=} \frac{1}{m}, \quad k \in \{1, \dots, m\}$$

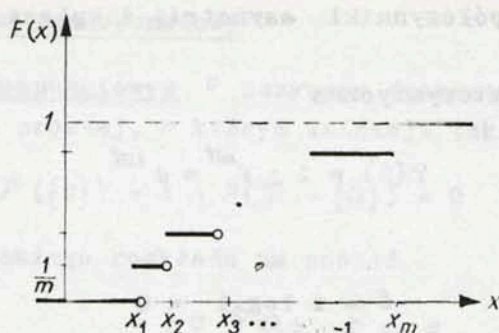
Przyjmijmy dla uproszczenia, że

$$x_1 < \dots < x_m$$

Dystrybuanta  $F$  takiego rozkładu ma postać

$$F(x) = \begin{cases} 0 & \text{dla } x \leq x_1 \\ \frac{k}{m} & \text{dla } x_k < x \leq x_{k+1}, \quad k = 1, \dots, m-1 \\ 1 & \text{dla } x > x_m \end{cases}$$





Mamy dalej

$$\bigwedge_{k \in \mathfrak{R}} m_k = \frac{1}{m} \sum_{j=1}^m x_j^k$$

W szczególności

$$\bar{x} = \frac{x_1 + \dots + x_m}{m}$$

Wariancja jest równa

$$\mu_2 = \frac{1}{m} \sum_{j=1}^m (x_j - \bar{x})^2 = \frac{x_1^2 + \dots + x_m^2}{m} - \bar{x}^2$$

Funkcja charakterystyczna

$$\varphi(t) = \frac{1}{m} \sum_{j=1}^m e^{ix_j t}$$

Entropia

$$\mathcal{E} = - \sum_{j=1}^m \frac{1}{m} \log_2 \frac{1}{m} = \log_2 m$$

Na mocy twierdzenia (B) z § 238 spośród wszystkich rozkładów dyskretnych  $m$ -punktowych rozkład równomierny ma największą entropię.



### § 241. Rozkład jednorodny

Rozkładem jednorodnym będziemy nazywać rozkład równomierny w przypadku, gdy punkty  $x_1, \dots, x_m$  idą po sobie w równych odstępach, tzn. istnieje taka liczba rzeczywista dodatnia  $h$ , że

$$x_k = x_1 + (k-1)h \quad k \in \{1, \dots, m\}$$

w tym przypadku

$$\bar{x} = \frac{x_1 + (x_1 + h) + \dots + [x_1 + (m-1)h]}{m} =$$

$$= x_1 + \frac{m-1}{2} h = \frac{x_1 + x_m}{2}$$

a wariancja jest równa

$$\mu_2 = \overline{x^2} - \bar{x}^2 =$$

$$= \frac{1}{m} [x_1^2 + (x_1 + h)^2 + \dots + (x_1 + (m-1)h)^2] - \left(x_1 + \frac{m-1}{2} h\right)^2 =$$

$$= \frac{1}{m} [mx_1^2 + 2hx_1(1 + \dots + (m-1)) + h^2(1^2 + \dots + (m-1)^2)] -$$

$$- x_1^2 - (m-1)hx_1 - \frac{(m-1)^2}{4} h^2 =$$

$$= \frac{2hx_1}{m} \cdot \frac{m(m-1)}{2} + \frac{h^2}{m} \cdot \frac{(m-1)m(2m-1)}{6} - (m-1)hx_1 - \frac{(m-1)^2}{4} h^2 =$$

$$= \frac{h^2}{12} (m-1)(4m-2 - 3m + 3) = \frac{h^2}{12} (m-1)(m+1) = \frac{h^2}{12} (m^2 - 1)$$

Skorzystaliśmy tu ze znanego wzoru

$$(227) \quad \sum_m^1 \stackrel{\text{def}}{=} 1 + \dots + m = \frac{m(m+1)}{2}$$

oraz ze wzoru

$$(228) \quad \sum_m^2 \stackrel{\text{def}}{=} 1^2 + \dots + m^2 = \frac{m(m+1)(2m+1)}{6}$$

Drugi z tych wzorów wyprowadza się łatwo metodą sumowania stronami równości

$$\begin{aligned} 2^3 &= (1 + 1)^3 = 1^3 + 3 \cdot 1^2 + 3 \cdot 1 + 1 \\ 3^3 &= (2 + 1)^3 = 2^3 + 3 \cdot 2^2 + 3 \cdot 2 + 1 \\ &\dots \dots \dots \\ (m+1)^3 &= (m + 1)^3 = m^3 + 3 \cdot m^2 + 3 \cdot m + 1 \end{aligned}$$

Po zsumowaniu mamy

$$S_m^3 - 1^3 + (m+1)^3 = S_m^3 + 3S_m^2 + 3S_m^1 + m$$

skąd

$$\begin{aligned} 3S_m^2 &= (m+1)^3 - 1 - m - 3S_m^1 = \\ &= (m+1)((m+1)^2 - 1 - \frac{3}{2}m) = (m+1)(m^2 + \frac{1}{2}m) \end{aligned}$$

a następnie

$$S_m^2 = \frac{1}{6}m(m+1)(2m+1)$$

Ogólnie, chcąc otrzymać wzór na  $S_m^k$ , mając już wzory  $S_m^1, \dots, S_m^{k-1}$ , sumujemy stronami równości

$$\begin{aligned} 2^{k+1} &= (1+1)^{k+1} = 1^{k+1} + \binom{k+1}{1}1^k \cdot 1 + \binom{k+1}{2}1^{k-1} \cdot 1 + \dots + 1 \\ 3^{k+1} &= (2+1)^{k+1} = 2^{k+1} + \binom{k+1}{1}2^k \cdot 1 + \binom{k+1}{2}2^{k-1} \cdot 1 + \dots + 1 \\ &\dots \dots \dots \\ (m+1)^{k+1} &= (m+1)^{k+1} = m^{k+1} + \binom{k+1}{1}m^k \cdot 1 + \binom{k+1}{2}m^{k-1} \cdot 1 + \dots + 1 \end{aligned}$$

i otrzymujemy

$$S_m^{k+1} - 1 + (m+1)^{k+1} = S_m^{k+1} + \binom{k+1}{1}S_m^k + \binom{k+1}{2}S_m^{k-1} + \dots + m$$

skąd

$$S_m^k = \frac{1}{k+1} \left( (m+1)^{k+1} - (m+1) - \binom{k+1}{2}S_m^{k-1} - \binom{k+1}{3}S_m^{k-2} - \dots - \binom{k+1}{k}S_m^1 \right)$$



W ten sposób otrzymujemy na przykład wzory

$$(229) \quad \zeta_m^3 = \frac{m^2(m+1)^2}{4}, \quad \zeta_m^4 = \frac{m(m+1)(2m+1)(3m^2+3m-1)}{30}$$

Postępując podobnie jak poprzednio z wariancją  $\mu_2$  obliczamy dla naszego rozkładu

$$\mu_3 = 0, \quad \mu_4 = \frac{h^4}{240} (m^2-1)(3m^2-7)$$

skąd na mocy wzorów (206) i (207) współczynniki asymetrii i spłaszczenia są równe

$$\gamma_1 = 0, \quad \gamma_2 = -\frac{6(m^2+1)}{5(m^2-1)}$$

Funkcja charakterystyczna ma postać

$$\begin{aligned} \varphi(t) &= \frac{1}{m} \left( e^{ix_1 t} + e^{i(x_1+h)t} + \dots + e^{i(x_1+(m-1)h)t} \right) = \\ &= \frac{1}{m} e^{ix_1 t} \left( 1 + e^{iht} + (e^{iht})^2 + \dots + (e^{iht})^{m-1} \right) = \\ &= \frac{1}{m} e^{ix_1 t} \frac{1 - e^{imht}}{1 - e^{iht}} \end{aligned}$$

Entropia jako entropia rozkładu równomiernego jest równa  $\log_2 m$ .

## § 242. Rozkład binomialny, czyli Bernoulliego

Rozkładem binomialnym albo rozkładem Bernoulliego, nazywamy rozkład dyskretny, dla którego

$$x_k = k \wedge p_k = P(\{x_k\}) = b(k; n, p) \stackrel{\text{def}}{=} \binom{n}{k} p^k (1-p)^{n-k}$$

gdzie  $0 < p < 1$ . (Por. schemat Bernoulliego, § 154).

Dystrybuanta takiego rozkładu jest równa

$$F(x) = \begin{cases} 0 & \text{dla } x < 0 \\ \sum_{k=1}^m \binom{n}{k} p^k (1-p)^{n-k} & \text{dla } m < x < m+1, \quad m = 0, 1, \dots, n-1 \\ 1 & \text{dla } x > n \end{cases}$$