Počet pravděpodobnosti

6. Markovův jednoduchý řetěz o dvou eventualitách

In: Bohuslav Hostinský (author): Počet pravděpodobnosti. Druhá část. (Czech). Praha: Jednota československých matematiků a fysiků, 1950. pp. 22–48.

Persistent URL: http://dml.cz/dmlcz/403305

KAPITOLA ŠESTÁ

MARKOVŮV JEDNODUCHÝ ŘETĚZ O DVOU EVENTUALITÁCH

57. Pojem Markovova řetězu. V druhé kapitole bylo pojednáno o pravděpodobnostech vztahujících se k opětovaným pokusům. Při tom se předpokládalo, že je určitá konstantní pravděpodobnost p, že se pokus zdaří, stejná pro všechny pokusy a nezávislá na tom, zda se pokusy, provedené před pokusem právě uvažovaném, zdařily či nezdařily. Nyní vezmeme v úvahu pravděpodobnosti vztahující se k opětovaným pokusům předpokládajíce, že pravděpodobnost, se kterou se vyskytne nějaký zjev jako výsledek určitého pokusu, závisí na výsledcích pokusů předcházejících.

Budiž $E^{(1)}, E^{(2)}, E^{(3)}, \ldots, E^{(n)}, \ldots$ řada zjevů, z nichž první je výsledek prvního pokusu, druhý je výsledek druhého pokusu atd. *Pravíme, že pravděpodobnosti, se kterými se vyskytují zjevy $E^{(n)}$, jsou spojeny v Markovův řetěz, závisí-li pravděpodobnost, že n-tý pokus dá výsledek $E^{(n)}$ na tom, jaké výsledky daly pokusy předcházející.*

Závislosti mezi pravděpodobnostmi výsledků, které může míti n-tý pokus, a mezi výsledky, které daly pokusy předcházející, mohou býti různého druhu. V odst. 58 podáme přehled rozmanitých úloh o řetězech, kterými se budeme zabývati v kap. VI, VII a VIII.

58. Přehled úloh o Markovových řetězech. a) V kap. VI budeme jednati o *jednoduchém Markovově řetězu se dvěma eventualitami a s konstantními pravděpodobnostmi přechodu.* Každý z pokusů má jen dvě eventuality: buď se podaří, nebo se nepodaří. Řetěz je jednoduchý; to znamená, že, je-li znám výsledek n-tého pokusu, je tím určena jednoznačně pravděpodobnost, že se $(n + 1)$-tý pokus zdaří nezávisle na tom, jak dopadly pokusy předcházející před n-tým. Řečená pravděpodobnost je konstantní, nezávisí na pořadovém čísle pokusu.*) Veli-

*) Místo řetěz s konstantními pravděpodobnostmi přechodu se říká též *homogenní řetěz.*

činy definující takový řetěz a příklad, jak se řetěz uskuteční, jsou uvedeny v odst. 59. Pravděpodobnost, že n-tý pokus se podaří, závisí obecně na n; za určitých předpokladů má limitu, roste-li n do nekonečna (odst. 60), dodatky k této větě v odst. 61 a 62. Podobně jako v případě nezávislých pokusů zavádí se zde pro řadu n pokusů střední hodnota počtu m zdařených pokusů (63); střední hodnota čtverce úchylky, dělená počtem pokusů, má za určitých předpokladů limitu, roste-li n do nekonečna (64). Ve zvláštním případě může býti pravděpodobnost, že se n-tý pokus zdaří, nezávislá na n (65). Srovnání výsledků s obdobnými platnými pro nezávislé pokusy (66). Odst. 67—69 jednají o rozmanitých aplikacích. V odst. 70 se jedná o t. zv. charakteristické rovnici přiřaděné danému řetězu.

b) Kap. VII začíná definicí jednoduchého Markovova řetězu o libovolném (konečném) počtu eventualit (odst. 71). K rozboru některých úloh hodí se zavésti proměnnou veličinu závislou na výsledku pokusu (72). Geometrické znázornění řetězu (73). Na základě vlastností středních hodnot veličin vztahujících se k řetězu (74) dokazuje se věta o limitě pravděpodobnosti (75) s dodatky (76, 77). Kořeny charakteristické rovnice (78) slouží v některých případech k výhodnému vyjádření pravděpodobností přechodu (79). Po přehledu method k výpočtu disperse (80) je určena disperse pro libovolný počet n pokusů (81). Odst. 82 jedná o stacionárním řetězu.

c) Užití řetězů je objasněno v kap. VIII, kde jsou rozebírány úlohy o míchání karet (odst. 83, 84) a úloha o tazích ze dvou osudí se záměnou koulí (86—87); v této poslední úloze vyskytuje se případ zjevů $F_1, F_2, F_3, \ldots$ jež *netvoří* řetěz, jichž pravděpodobnosti jsou po řadě závislé na zjevech $E_1, E_2, E_3, \ldots$, které *tvoří* řetěz (86). Zákon velkých čísel platí též v případě řetězů (89). V odst. 41 jsme pojednali o regularisaci v případě geometrických pravděpodobností. V případě řetězů nastává regularisace, t. j. určitá pravděpodobnost se blíží limitě, vzrůstá-li počet pokusů do nekonečna (90, 91).

59. Jednoduchý řetěz se dvěma eventualitami a konstantními pravděpodobnostmi přechodu. a) Jsou dána dvě osudí, jedno bílé, druhé černé. V každém z nich jsou bílé a černé koule. Je tedy určitá pravděpodobnost p_{11}, že z bílého osudí vytáhneme kouli bílou a pravděpodobnost p_{12}, že z něho vytáhneme černou; platí

$$p_{11} + p_{12} = 1. \tag{1}$$

Podobně pravděpodobnosti p_{21} a p_{22}, že z černého osudí vytáhneme bílou resp. černou kouli, vyhovují rovnici

$$p_{21} + p_{22} = 1. \tag{1a}$$

Předpokládáme, že konáme postupně tahy za těchto podmínek: vytáhneme-li při n-tém tahu kouli bílou, koná se $(n + 1)$-tý tah z osudí bílého; vytáhneme-li při n-tém tahu kouli černou, koná se $(n + 1)$-tý tah z osudí černého.*) Pravděpodobnosti, že při 1., 2., 3., ... tahu vyjde na př. bílá koule, jsou spojeny v *jednoduchý Markovův řetěz*; p_{ik} jsou *pravděpodobnosti přechodu*, $(i, k = 1, 2)$. Aby byl řetěz úplně určen, musíme nějak určiti, ze kterého osudí se koná první tah. Toto určení se může provésti dvojím způsobem:

Buď je přímo dáno, ze kterého osudí se koná první tah, nebo jsou dány jen pravděpodobnosti $p_1^{(0)}$ a $p_2^{(0)}$, že první tah se koná z osudí bílého resp. černého; rozhodnouti o tom, ze kterého osudí se skutečně táhne, závisí na výsledku *předběžného pokusu*; uvedené pravděpodobnosti splňují podmínku

$$p_1^{(0)} + p_2^{(0)} = 1. \tag{2}$$

Není-li tedy dáno, ze kterého osudí se koná první tah, je pravděpodobnost $p_1^{(1)}$, že při prvním tahu vyjde bílá koule, rovna úhrnné pravděpodobnosti složené ze dvou sčítanců: buď se koná tah z bílého a vyjde bílá, nebo se koná z černého a vyjde bílá, tedy

$$p_1^{(1)} = p_1^{(0)} \cdot p_{11} + p_2^{(0)} \cdot p_{21}. \tag{3}$$

b) Za předpokladu, že pravděpodobnosti $p_1^{(0)}$ a $p_2^{(0)}$ jsou dány, budiž $p_1^{(n)}$ *prostá pravděpodobnost, že při n-tém tahu vyjde bílá koule a* $p_2^{(n)}$, že při n-tém tahu vyjde černá. Srovnáváme-li $p_k^{(n+1)}$ s $p_k^{(n)}$, $(k = 1, 2)$, dostaneme podle věty o úhrnné pravděpodobnosti

$$p_k^{(n+1)} = p_1^{(n)} \cdot p_{1k} + p_2^{(n)} \cdot p_{2k}, \; k = 1, 2. \tag{3a}$$

*) Zde i v dalším stále předpokládáme, že vytažená koule se vždy vloží zpět do osudí, z něhož byla vytažena; počet bílých koulí i počet černých zůstává v jednom i ve druhém osudí beze změny.

Napíšeme-li tuto rovnici postupně*) pro $n = 0, 1, 2, \dots$, vypočteme po řadě veličiny $p_k^{(n)}$ $(n = 1, 2, 3, \dots)$. Vychází

$$p_k^{(1)} = p_1^{(0)} \cdot p_{1k} + p_2^{(0)} p_{2k}, \quad k = 1, 2,$$

což je zobecněná rovnice (3), a dále

$$p_k^{(2)} = p_1^{(0)}(p_{11}p_{1k} + p_{12}p_{2k}) + p_2^{(0)}(p_{21}p_{1k} + p_{22}p_{2k}), \quad \dots$$

$$p_k^{(n)} = p_1^{(0)} \sum_{\alpha=1}^{2} \sum_{\beta=1}^{2} \dots \sum_{\omega=1}^{2} p_{1\alpha} p_{\alpha\beta} p_{\beta\gamma} \dots p_{\omega k} +$$

$$+ p_2^{(0)} \sum_{\alpha=1}^{2} \sum_{\beta=1}^{2} \dots \sum_{\omega=1}^{2} p_{2\alpha} p_{\alpha\beta} p_{\beta\gamma} \dots p_{\omega k}, \quad k = 1, 2. \tag{4}$$

Každý z obou součtů je $(n - 1)$-násobný; sčítá se podle $n - 1$ indexů $\alpha, \beta, \gamma, \dots, \omega$.

c) Koeficienty při $p_1^{(0)}$ a $p_2^{(0)}$ v rovnici (4) jsou

$$P_{ik}^{(n)} = \sum_{\alpha=1}^{2} \sum_{\beta=1}^{2} \dots \sum_{\omega=1}^{2} p_{i\alpha} p_{\alpha\beta} \dots p_{\omega k}, \quad P_{ik}^{(1)} = p_{ik}, \tag{5}$$

$$i, k = 1, 2; \quad n = 1, 2, 3, \dots$$

$P_{ik}^{(n)}$ je pravděpodobnost, že, konal-li se první tah z i-tého osudí ($i = 1$ odpovídá bílému, $i = 2$ černému), vyšla při n-tém tahu koule barvy k ($k = 1$ odpovídá bílé, $k = 2$ černé). Význam rovnice (5) pro $P_{ik}^{(n)}$ objasní se takto: $p_{i\alpha} p_{\alpha\beta} \dots p_{\omega k}$ je složená pravděpodobnost, že při prvním, druhém, ..., n-tém tahu byly vytaženy postupně koule barev $\alpha, \beta, \dots, \omega, k$; sečteme-li tyto složené pravděpodobnosti pro všechny možné hodnoty indexů $\alpha, \beta, \dots, \omega$, dostaneme úhrnnou pravděpodobnost $P_{ik}^{(n)}$, že při n-tém tahu vyjde koule barvy k.

Obecněji ukazuje rovnice (5), že $P_{ik}^{(n)}$ má tento význam: je to *pravděpodobnost, že při $(m + n)$-tém tahu vyjde koule barvy k*, vyšla-li při m-tém tahu koule barvy i.

Místo rovnice (4) pišme stručněji

*) Index 0 platí pro *předběžný pokus.*

$$p_k^{(n)} = p_1^{(0)} P_{1k}^{(n)} + p_2^{(0)} P_{2k}^{(n)}. \tag{6}$$

Z definice veličin $P_{ik}^{(n)}$ plyne přímo, že

$$P_{ik}^{(m+n)} = \sum_{j=1}^{2} P_{ij}^{(m)} P_{jk}^{(n)}, \quad i, k = 1, 2, \tag{7}$$

kde m a n jsou libovolná kladná celá čísla; sečítajíce v rovnici (5) postupně podle indexů $k, \omega, \ldots, \beta, \alpha$ vždy od 1 do 2, najdeme s ohledem na (1) a (1a), že

$$\sum_{k=1}^{2} P_{ik}^{(n)} = 1, \quad n = 1, 2, 3, \ldots \tag{8}$$

Z rovnic (2), (6) a (8) plyne vztah

$$p_1^{(n)} + p_2^{(n)} = 1, \quad n = 0, 1, 2, \ldots, \tag{9}$$

který ostatně je přímým důsledkem definice veličin $p_i^{(n)}$. Neboť rovnice (9) vyjadřuje větu, že při n-tém tahu je jisto, že bude tažena buď bílá nebo černá koule.

d) Obecný pojem jednoduchého řetězu se dvěma eventualitami a s konstantními pravděpodobnostmi přechodu p_{ik} stanoví se takto: Konáme postupně pokusy takové, že každý z nich dá za výsledek jeden ze dvou zjevů E_1, E_2. Je-li známo, že n-tý pokus dal výsledek E_i, je p_{ik} pravděpodobnost, že $(n + 1)$-tý pokus dá výsledek E_k, $(k = 1, 2)$; pravděpodobnosti p_{ik} nejsou závislé na tom, jak dopadly pokusy $(n—1)$-tý, $(n — 2)$-tý, ... Pravděpodobnosti vyhovují rovnicím (1) a (1a), jež shrneme v jedinou

$$p_{i1} + p_{i2} = 1, \quad i = 1, 2.$$

Rovnice (5) určuje pravděpodobnost $P_{ik}^{(n)}$, že zjev E_k se vyskytne jakožto výsledek n-tého pokusu, dal-li „předběžný pokus“, který si myslíme vykonaný bezprostředně před prvním, zjev E_i. Veličiny $P_{ik}^{(n)}$ vyhovují rovnicím (7) a (8). Jsou-li $p_1^{(0)}$ a $p_2^{(0)}$ pravděpodobnosti, že E_1 resp. E_2 se vyskytne jakožto výsledek předběžného pokusu, je pravděpodobnost $p_k^{(n)}$, že n-tý pokus dá za výsledek E_k, určena rovnicí (6).

60. Markovova věta o limitě pravděpodobnosti. $P_{ik}^{(n)}$. a) Předpokládáme, že všechny pravděpodobnosti p_{ik} jsou kladné:

$$0 < p_{ik} < 1, \; i, k = 1, 2 \tag{1}$$

a že jsou splněny rovnice (1) a (1a), odst. 60, že tedy

$$p_{i1} + p_{i2} = 1, \; i = 1, 2. \tag{2}$$

Položme

$$\delta = p_{11} - p_{21}.$$

Z rovnice (3a) a (9), odst. 59 plyne pro $k = 1$, že

$$p_1^{(n+1)} = p_{21} + \delta \, . \, p_1^{(n)},$$

anebo, upravíme-li,

$$p_1^{(n+1)} - \frac{p_{21}}{1 - \delta} = \delta \left(p_1^{(n)} - \frac{p_{21}}{1 - \delta} \right).$$

Dosazujme do této rovnice na místo n postupně 0, 1, 2, ..., $n - 1$ a znásobme všechny tak utvořené rovnice. Vychází

$$p_1^{(n)} - \frac{p_{21}}{1 - \delta} = \delta^n \left(p_1^{(0)} - \frac{p_{21}}{1 - \delta} \right). \tag{3}$$

Poněvadž vzhledem k (1) je $|\delta| < 1$, máme pro $n \to \infty$

$$\lim_{n \to \infty} p_1^{(n)} = \frac{p_{21}}{1 - \delta}. \tag{3a}$$

Stejným postupem dostaneme z rovnic (3a) a (9), odst. 60 pro $k = 2$

$$p_2^{(n+1)} = p_{12} + \delta \, . \, p_2^{(n)}, \; \delta = p_{22} - p_{12} = p_{11} - p_{21},$$

$$p_2^{(n)} - \frac{p_{12}}{1 - \delta} = \delta^n \, . \left(p_2^{(0)} - \frac{p_{12}}{1 - \delta} \right)$$

a tedy

$$\lim_{n\to\infty} p_2^{(n)} = \frac{p_{12}}{1-\delta}. \tag{3b}$$

Limitní hodnoty (3a) a (3b) označíme P_1, P_2. Je tedy

$$P_1 = \frac{p_{21}}{1-\delta}, \; P_2 = \frac{p_{12}}{1-\delta}, \; P_1 + P_2 = \frac{p_{21} + 1 - p_{11}}{1 - p_{11} + p_{21}} = 1. \tag{4}$$

Rovnice (3a) a (3b) vyjadřují *Markovovu větu o limitě pravděpodobnosti*, kterou vyslovíme, majíce na mysli tahy ze dvou osudí podle předpokladů odst. 59a, takto:

Prostá pravděpodobnost, že při n-tém pokuse bude vytažena koule bílá (černá), má, roste-li n do nekonečna, limitu P_1 (P_2) určenou rovnicí (4) *a nezávislou na počátečních pravděpodobnostech $p_1^{(0)}$, $p_2^{(0)}$* (jinými slovy: nezávislou na tom, konal-li se první tah z osudí bílého nebo z černého).

Kdybychom se drželi abstraktního pojetí řetězu podle odst. 59d, zněla by věta takto:

Prostá pravděpodobnost, že n-tý pokus má za výsledek zjev E_k, ($k = 1, 2$), *má pro $n \to \infty$ limitu P_k.*

b) Vyšetřme nyní limitní hodnoty veličin $P_{ik}^{(n)}$ pro $n \to \infty$. Dosaďme do rovnice (6), odst. 59

$$p_1^{(0)} = 1, \; p_2^{(0)} = 0.$$

Vychází

$$p_k^{(n)} = P_{1k}^{(n)}$$

a tedy podle (3), pro $k = 1$

$$P_{11}^{(n)} = \frac{p_{21}}{1-\delta} + \delta^n \left(1 - \frac{p_{21}}{1-\delta}\right) = \delta^n + P_1(1-\delta^n). \tag{5}$$

Dosaďme nyní do (6), odst. 59

$$p_1^{(0)} = 0, \; p_2^{(0)} = 1.$$

Pak bude

$$p_k^{(n)} = P_{2k}^{(n)}$$

a rovnice (3) dá pro $k = 1$

$$P_{21}^{(n)} = \frac{p_{21}}{1-\delta} - \delta^n \frac{p_{21}}{1-\delta} = P_1(1-\delta^n). \tag{5a}$$

Z rovnic (5) a (5a) plyne, že

$$P_{11}^{(n)} - P_{21}^{(n)} = \delta^n \tag{6}$$

a dále

$$\lim_{n\to\infty} P_{11}^{(n)} = \lim_{n\to\infty} P_{21}^{(n)} = P_1. \tag{7}$$

Poněvadž pak podle (8), odst. 59 a podle (4), odst. 60 je

$$P_{11}^{(n)} + P_{12}^{(n)} = 1,\ P_{21}^{(n)} + P_{22}^{(n)} = 1,\ P_1 + P_2 = 1, \tag{5b}$$

platí též

$$\lim_{n\to\infty} P_{12}^{(n)} = \lim_{n\to\infty} P_{22}^{(n)} = P_2. \tag{7a}$$

Považujeme rovnice (7) a (7a) za jiný výraz shora uvedené Markovovy věty. Jejich smysl je: *Pravděpodobnost $P_{ik}^{(n)}$, že n-tý pokus dá výsledek E_k, dal-li předběžný (nultý) pokus výsledek E_i, má za limitu P_k, ($k = 1, 2$) nezávislou na i.*

Limity P_1, P_2 jsou určeny rovnicemi (4).

61. Dodatek k větě o limitě pravděpodobnosti. Vyhovují-li pravděpodobnosti p_{ik} nejen podmínkám (1) a (2), odst. 60, nýbrž také podmínkám

$$p_{1k} + p_{2k} = 1,\ k = 1, 2, \tag{1}$$

je

$$1 = p_{11} + p_{12} = p_{11} + p_{21} = p_{21} + p_{22} = p_{21} + p_{11}$$

a tedy

$$p_{12} = p_{21},\ p_{11} = p_{22}. \tag{2}$$

Z toho plyne podle (4), odst. 60, že

$$P_1 = \frac{p_{21}}{1-\delta} = \frac{1-p_{11}}{1-p_{11}+p_{21}} = \frac{1-p_{11}}{2(1-p_{11})} = \tfrac{1}{2} = P_2. \quad (3)$$

Naopak plyne z rovnosti $P_2 = P_1$, že $p_{12} = p_{21}$ (srv. rovnice (4), odst. 60), z čehož následuje platnost rovnic (1).

Jsou-li splněny rovnice (1) *a* (2), *odst. 60, vyjadřují rovnice* (1) *nutnou a postačující podmínku pro to, aby limity* P_1 *a* P_2 *byly stejné.*

V případě dvou osudí (odst. 59a) mají podmínky (1) nebo z nich plynoucí důsledky tento význam: v bílém osudí je poměr počtu bílých koulí k počtu černých zrovna takový, jako je v černém osudí poměr černých k bílým. Po velmi velkém počtu tahů je pravděpodobnost vytáhnouti bílou kouli stejně veliká jako černou.

62. Zvláštní případ, kdy podmínky věty o limitě nejsou splněny. Předpoklad vyjádřený nerovnostmi (1), odst. 60 je podstatný pro platnost Markovovy věty o limitě. Kdyby na př.

$$p_{11} = 0,\ p_{22} = 0,\ p_{12} = p_{21} = 1, \quad (1)$$

zjevy E_1 a E_2 by se pravidelně střídaly; kdyby n-tý pokus vedl k E_1 (E_2), vedl by $(n+1)$-tý pokus k E_2 (E_1). Pravděpodobnosti $p_k^{(n)}$ a $P_{ik}^{(n)}$ by neměly limitu, rovnice (3a), (7) a (7a), odst. 60 by neplatily.

Vrátíme-li se k tahům ze dvou osudí (odst. 59a), mají podmínky (1) tento význam: V bílém osudí jsou jen černé koule, v černém jen bílé. Konáme-li n-tý tah z bílého (černého) osudí, koná se $(n+1)$-tý z černého (bílého). Kdyby se první tah konal z bílého, byla by pravděpodobnost $P_{11}^{(n)}$, že při n-tém tahu vyjde bílá koule, rovna nule pro liché n a rovna jedné pro sudé n.

63. Střední hodnota počtu zdařených pokusů. Vykonejme celkem n pokusů za předpokladů odst. 59d. Vyskytne-li se zjev E_1, pravíme, že se pokus zdařil; vyskytne-li se E_2, pravíme, že se pokus nezdařil. r-tému pokusu přiřadíme veličinu $x^{(r)}$;

zdaří-li se pokus, budiž $x^{(r)} = 1$, nezdaří-li se, budiž $x^{(r)} = 0$. Je-li v řadě n pokusů m pokusů zdařilých, je (jako v odst. 14b)

$$m = x^{(1)} + x^{(2)} + \dots + x^{(n)}.$$

Budiž jako v odst. 59c $P_{ik}^{(s)}$ pravděpodobnost, že s-tý pokus bude míti výsledek E_k, dal-li předběžný (nultý) pokus výsledek E_i.

Pravděpodobnost, že se s-tý pokus zdaří, je tedy $P_{i1}^{(s)}$; střední hodnota veličiny $x^{(s)}$ je

$$E(x^{(s)}) = 1 \,.\, P_{i1}^{(s)} + 0 \,.\, (1 - P_{i1}^{(s)}) = P_{i1}^{(s)} \tag{2}$$

a střední hodnota počtu m zdařených pokusů je (pro $i = 1, 2$)

$$E(m) = E(x^{(1)}) + E(x^{(2)}) + \dots + E(x^{(n)}) =$$
$$= P_{i1}^{(1)} + P_{i1}^{(2)} + \dots + P_{i1}^{(n)}. \tag{3}$$

Předpokládáme, že jsou splněny podmínky (1), odst. 60. Podle (5) a (5a), odst. 60 je

$$P_{11}^{(s)} = \frac{p_{21}}{1 - \delta} + \frac{1 - p_{11}}{1 - \delta} \,.\, \delta^s,$$

$$P_{21}^{(s)} = \frac{p_{21}}{1 - \delta} - \frac{p_{21}}{1 - \delta} \,.\, \delta^s.$$

Hodnota $E(m)$ závisí tedy na indexu i, ale

$$\lim_{n \to \infty} \frac{E(m)}{n} = \lim_{n \to \infty} \frac{P_{i1}^{(1)} + P_{i1}^{(2)} + \dots + P_{i1}^{(n)}}{n} = P_1, \tag{4}$$

neboť podle (4) a (7), odst. 60 má $P_{i1}^{(n)}$ limitu P_1 pro $n \to \infty$. *Střední hodnota počtu podařených pokusů dělená celkovým počtem pokusů má pro $n \to \infty$ limitu P_1 nezávislou na výsledku předběžného pokusu.*

64. Markovova věta o limitě disperse. V případě nezávislých pokusů jsme dokázali — viz rovnici (1), odst. 15 — že střední hodnota druhé mocniny úchylky, totiž $(m - np)^2$, dělená počtem pokusů n, rovná se pro každou hodnotu čísla n výrazu

$p(1-p)$. V případě pokusů spojených v řetěz je výraz pro $E(m)$ (viz rovnici (3), odst. 63) složitější; vezmeme-li na místo p pravděpodobnost P_1, že n-tý pokus při nekonečně velikém n se podaří, vyjádří se střední hodnota druhé mocniny úchylky takto:

$$E(m-nP_1)^2 = E[\sum_{s=1}^{n}(x^{(s)}-P_1)]^2.$$

Veličina $x^{(s)}$ je zde definována jako v odst 62. Z toho plyne pro střední hodnotu výrazu $(m-nP_1)^2/n$, který se nazývá *disperse*, rovnice

$$E\,\frac{(m-nP_1)^2}{n} = \frac{\sum_{s=1}^{n}E(x^{(s)}-P_1)^2}{n} + $$

$$+\,2\,\frac{\sum_{r=1}^{n-1}\sum_{s=1}^{n-r}E[(x^{(r)}-P_1)(x^{(r+s)}-P_1)]}{n}. \tag{5}$$

První člen napravo v této rovnici se snadno vypočte. Uvažme, že $x^{(s)}$ má hodnotu rovnou jedné nebo nule dle toho, vyjde-li při s-tém tahu koule bílá nebo černá; dále je (za předpokladu, že první tah se děje z i-tého osudí)

$P_{i1}^{(s)}$ pravděpodobnost rovnosti $x^{(s)}=1$,

$P_{i2}^{(s)}$ pravděpodobnost rovnosti $x^{(s)}=0$.

Z toho plyne, že

$$E(x^{(s)}-P_1)^2 = P_{i1}^{(s)}(1-P_1)^2 + P_{i2}^{(s)}P_1^2. \tag{6}$$

K výpočtu druhého členu na pravé straně rovnice (5) poznamenejme (viz odst. 59c):

$P_{i1}^{(r)}P_{11}^{(s)}$ je pravděpodobnost, že platí obě rovnice

$$x^{(r)}=1,\ x^{(r+s)}=1,$$

$P_{i1}^{(r)}P_{12}^{(s)}$ je pravděpodobnost, že platí obě rovnice

$$x^{(r)}=1,\ x^{(r+s)}=0,$$

$P_{i2}^{(r)}P_{21}^{(s)}$ je pravděpodobnost, že platí obě rovnice

$$x^{(r)} = 0,\ x^{(r+s)} = 1,$$

$P_{i2}^{(r)}P_{22}^{(s)}$ je pravděpodobnost, že platí obě rovnice

$$x^{(r)} = 0,\ x^{(r+s)} = 0.$$

Je tedy

$$E[(x^{(r)} - P_1)(x^{(r+s)} - P_1)] =$$
$$= P_{i1}^{(r)}P_{11}^{(s)}(1 - P_1)^2 + P_{i1}^{(r)}P_{12}^{(s)}(1 - P_1)\,.\,(- P_1) +$$
$$+ P_{i2}^{(r)}P_{21}^{(s)}\,.\,(- P_1)(1 - P_1) + P_{i2}^{(r)}P_{22}^{(s)}\,.\,P_1^2. \tag{7}$$

Dosaďme podle (6) a (7) příslušné výrazy do pravé strany rovnice (5), a na místo $P_{ik}^{(n)}$ dosaďme podle (5) a (5a), odst. 60 přihlížejíce k prvním dvěma rovnicím (5b), odst. 60. Pravá strana rovnice (5) objeví se nám jako výraz složený z několika geometrických řad; vychází

$$E\,\frac{(m - nP_1)^2}{n} = P_1(1 - P_1)\left[\frac{1 + \delta}{1 - \delta} - \frac{2}{n}\,\frac{\delta(1 - \delta^n)}{(1 - \delta)^2}\right] +$$
$$+ \frac{(1 - 2P_1)(2 - i - P_1)}{n(\delta - 1)}\left[(2n - 1)\delta^{n+1} -\right.$$
$$\left. - \delta - 2\,\frac{\delta^2 - \delta^{n+1}}{1 - \delta}\right]; \tag{8}$$
$$i = 1, 2.$$

To je *obecná formule pro střední hodnotu čtverce úchylky dělenou počtem n pokusů* neboli pro dispersi. Předpokládá se, že p_{jk}, $(j, k = 1, 2)$ jsou kladné a dále

$$\sum_{k=1}^{2} p_{jk} = 1,\ j = 1, 2;\ P_1 = \frac{p_{21}}{1 - \delta},\ \ \delta = p_{11} - p_{21}.$$

Index i se vztahuje k výsledkům předběžného pokusu. V případě, že pokusy nejsou spojeny v řetěz, nýbrž že jsou nezávislé, platí

$$p_{11} = p_{21} = P_1 = p,\ \delta = 0;$$

rovnice (8) přechází v dříve odvozený vzorec (1), odst. 15. Uveďme tyto speciální případy formule (8):

a) Roste-li n do nekonečna, je $\lim \delta^n = 0$ poněvadž $|\delta| < 1$, a formule (8) přechází v

$$\lim_{n\to\infty} E \frac{(m - nP_1)^2}{n} = P_1(1 - P_1) \frac{1 + \delta}{1 - \delta}, \tag{9}$$

což je *Markovova věta o limitě disperse.*

b) Je-li

$$p_{12} = p_{21}, \; p_{11} = p_{22}$$

a tedy

$$\delta = 2p_{11} - 1, \; 1 + \delta = 2p_{11}, \; 1 - \delta = 2(1 - p_{11}), \\ P_1 = \tfrac{1}{2},$$

dá rovnice (8) pro libovolné celé n

$$E \frac{(m - \frac{1}{2}n)^2}{n} = \frac{1}{4} \frac{p_{11}}{1 - p_{11}} + \frac{1 - 2p_{11}}{8n} \cdot \frac{1 - (2p_{11} - 1)^n}{(1 - p_{11})^2}. \tag{10}$$

Poznamenejme, že ani ve formuli (9) ani v (10) se neprojevuje výsledek předběžného pokusu; pravé strany obou těchto rovnic jsou nezávislé na indexu i.

65. Stacionární řetěz. a) Podle 59b prostá pravděpodobnost $p_1^{(n)}$, že se n-tý pokus podaří (neboli, podle odst. 59a, pravděpodobnost, že při n-tém tahu vytáhneme bílou kouli), je funkcí indexu n; je obecně pro každý pokus jiná. Řetěz se nazývá *stacionární*, je-li $p_1^{(n)}$ nezávislá na n. Podmínka, která musí býti splněna, aby řetěz byl stacionární, vyplývá přímo z rovnice (3), odst. 60; $p_1^{(n)}$ bude veličina nezávislá na n, bude-li koeficient při δ^n na pravé straně oné rovnice roven nule. Podmínka, aby řetěz byl stacionární, zní tedy

$$p_1^{(0)} = \frac{p_{21}}{1 - p_{11} + p_{21}}. \tag{1}$$

Splněním této podmínky je zaručeno, že řetěz je stacionární; naopak, je-li řetěz stacionární, musí býti nutně splněna tato podmínka. Srovnáním rovnic (3) a (4), odst. 60 plynou z podmínky stacionárnosti důsledky:

$$p_1^{(0)} = p_1^{(1)} = p_1^{(2)} = \ldots = p_1^{(n)} = \ldots = P_1 = \frac{p_{21}}{1 - p_{11} + p_{21}}; \tag{2}$$

rovnice (5) a (5a), odst. 60 dají pro stacionární řetěz

$$P_{11}^{(n)} = \delta^n + p_{21}\frac{1-\delta^n}{1-\delta}, \quad P_{21}^{(n)} = p_{21}\frac{1-\delta^n}{1-\delta}; \quad \delta = p_{11} - p_{21}. \tag{3}$$

Rovnice (1) *vyjadřuje nutnou a postačující podmínku, aby řetěz byl stacionární*; pravděpodobnost, že se zdaří předběžný pokus, rovná se, je-li podmínka (1) splněna, pravděpodobnosti, že se zdaří n-tý pokus a limitě P_1 pravděpodobnosti přechodu $P_{i1}^{(n)}$ $(i = 1, 2)$ pro $n \to \infty$. Srv. odst. 82c.

b) Ve smyslu odst. 59a uskutečníme stacionární řetěz takto: první, druhý, třetí, ... tah konají se buď z bílého nebo z černého osudí; pravděpodobnost vytáhnouti bílou kouli z bílého je p_{11}, pravděpodobnost vytáhnouti bílou z černého je p_{21}. Předběžný (nultý) tah koná se z pomocného osudí, ve kterém jsou též bílé a černé koule; pravděpodobnost vytáhnouti bílou je $p_1^{(0)}$, určená rovnicí (1). Vytáhneme-li při předběžném tahu bílou (černou) kouli, koná se první tah z bílého (černého) osudí. n-tý tah koná se z osudí té barvy, kterou má koule vytažená při $(n - 1)$-tém tahu. Při kterémkoli tahu má pravděpodobnost vytáhnouti bílou kouli konstantní hodnotu $p_1^{(0)}$ danou rovnicí (1).

66. Srovnání s případem nezávislých pokusů. Srovnejme případ stacionárního řetězu, kdy konáme tahy ze dvou osudí, jak bylo vyloženo v odst. 65, s případem nezávislých tahů z jednoho osudí.

Uvedeme číselný *příklad.*

a) Máme dvě osudí, jedno bílé, druhé černé; pravděpodobnosti p_{11} a p_{21} vytáhnouti bílou kouli z prvního resp. ze druhého jsou

$$p_{11} = \tfrac{2}{3}, \; p_{21} = \tfrac{1}{3};$$

předběžný tah nechť se koná z pomocného osudí, kde je jedna bílá koule a jedna černá; je tedy $p_1^{(0)} = \frac{1}{2}$. Podmínka stacionárnosti, vyjádřená rovnicí (1), odst. 65, je splněna a platí (viz odst. 63)

$$p_1^{(0)} = p_1^{(1)} = \ldots = P_{11}^{(n)} = P_{21}^{(n)} = P_1 = P_2 = \tfrac{1}{2},$$
$$n = 1, 2, 3, \ldots$$

Podle (3), odst. 63 je střední hodnota počtu m tahů, při nichž vyjde bílá koule, je-li n celkový počet tahů, roven

$$E(m) = \tfrac{1}{2}n. \tag{1}$$

Podle rovnice (9), odst. 64 je limita střední hodnoty čtverce úchylky

$$\lim_{n\to\infty} E\,\frac{(m - \frac{1}{2}n)^2}{n} = \tfrac{1}{2}. \tag{2}$$

b) V případě nezávislých pokusů, kdy konáme tahy z jediného osudí, ve kterém je tolik černých koulí kolik bílých, je pravděpodobnost, že vytáhneme bílou kouli, $p = \frac{1}{2}$. Poněvadž $p = p_{11} = p_{21} = \frac{1}{2}$, platí rovnice (1) jako v případě stacionárního řetězu, ale (viz odst. 15.)

$$E\,\frac{(m - \frac{1}{2}n)^2}{n} = \tfrac{1}{4}. \tag{2a}$$

Vykonáme-li tedy velký počet n tahů, je poměr m/n počtu tahů, při kterých byla tažena bílá koule, k celkovému počtu tahů přibližně roven $\frac{1}{2}$ v obou případech a) i b).

Ale *střední hodnota disperse,**) t. j. poměru $(m - \frac{1}{2}n)^2 : n$ je

---

*) Statisticky určíme střední hodnotu disperse z velikého počtu k serií tahů; každá serie má n tahů, v nichž celkem vyjde m-krát bílá koule. Aritmetický střed ze všech k hodnot disperse dává hledanou její střední hodnotu.

v případě a) stacionárního řetězu rovna $\frac{1}{2}$, kdežto v případě b) nezávislých tahů z jediného osudí je rovna $\frac{1}{4}$.

67. Výskyt samohlásek a souhlásek v souvislém textu. Markov vybral z Puškinova „*Oněgina*" část textu o 20 000 hláskách a rozdělil ji, bez porušení daného pořadí hlásek, na 200 skupin po 100 hláskách. V každé skupině čítal, kolik je v ní samohlásek; výsledek je vyjádřen statistickou tabulkou:

Počet samohlásek ve skupině	37	38	39	40	41	42	43	44	45	46	47	48	49
Počet skupin	3	1	6	18	12	31	43	29	25	17	12	2	1

Nejčetnější jsou skupiny, jež obsahují střední počet 43 samohlásek; je jich 43. Čím více se počet samohlásek ve skupině uchyluje od 43, tím méně takových skupin se vyskytuje. Největší úchylky od středního počtu 43 jsou —6 a +6. Pravděpodobnost, že hláska zvolená namátkou ve skupině, která obsahuje 37 (nebo 38, 39, ...) samohlásek, je samohláska, rovná se 0,37 (nebo 0,38, 0,39, ...). Průměrná pravděpodobnost P_1, že hláska zvolená v některé skupině, je samohláskou, je dána rovnicí

$$P_1 = \tfrac{1}{200}(3 \,.\, 0{,}37 + 1 \,.\, 0{,}38 + \ldots + 1 \,.\, 0{,}49) = 0{,}432. \quad (1)$$

V jedné skupině je průměrně $100 \,.\, P_1 = 43{,}2$ samohlásek.

Jde nyní o to srovnati úchylky, o které se liší skutečný počet samohlásek vyskytujících se v té které ze 200 skupin od středního jich počtu 43,2. Za tím účelem vypočteme střední hodnotu čtverce úchylky (dispersi) ze dvou různých předpokladů: a) výskyt samohlásek je obdobný výskytu bílých koulí, konáme-li tahy z jednoho osudí za předpokladů odst. 13 a násl. b) výskyt samohlásek je obdobný výskytu bílých koulí, konáme-li tahy ze dvou osudí v případě Markovova řetězu uvažovaného v odst. 59a a násl.

a) Očíslujme 200 skupin hlásek čísly 1, 2, ..., 200 a budiž m_i počet samohlásek obsažených v i-té skupině. Pak bude

$(m_i - 100P_1)^2$ čtverec úchylky pro i-tou skupinu; P_1 je pravděpodobnost, že zvolená hláska je samohláskou, daná rovnicí (1). Přirovnáme-li výskyt samohlásek k výskytu bílých koulí při tazích z jednoho osudí, vzájemně nezávislých, má býti podle odst. 15 a 22 rovnost mezi theoretickým výrazem $np(1 - p)$ a empirickou střední hodnotou

$$\frac{1}{s}\left[(m_1 - np)^2 + (m_2 - np)^2 + \dots + (m_s - np)^2\right].$$

Dosadíme-li sem $s = 200$ (počet serií), $n = 100$ (počet tahů v jedné serii), $p = P_1$ docházíme k závěru, že by měla platiti (přibližně) rovnost

$$\frac{1}{20000}\sum_{i=1}^{200}(m_i - 100P_1)^2 \doteq P_1(1 - P_1). \qquad (2)$$

Provedeme-li výpočet se shora uvedenými hodnotami pro m_i a pro P_1, shledáme, že levá strana v rovnici (2) má hodnotu 0,051, kdežto pravá hodnotu 0,245. Rovnost (2) není tedy v našem případě správná, výskyt samohlásek nelze přirovnati k vzájemně nezávislým tahům z osudí.

b) Skupiny po 100 hláskách jsou vzaty z textu, ve kterém pořadí hlásek je dáno smyslem textu. Hlásky nejsou zde vybrány jako koule vytahované z osudí.

Podle Markova jsou pravděpodobnosti, se kterými se vyskytne samohláska na prvém, druhém, třetím, ... místě textu spojeny v jednoduchý řetěz. Příslušné pravděpodobnosti stanovíme takto: Index 1 nechť odpovídá samohlásce, index 2 souhlásce; p_{11} bude pravděpodobnost, že po samohlásce bezprostředně následuje samohláska; p_{12} pravděpodobnost, že po samohlásce bezprostředně následuje souhláska atd. Platí rovnosti

$$p_{11} + p_{12} = 1,\ p_{21} + p_{22} = 1.$$

Veličiny p_{ik} se určí z textu statisticky: spočítáme, kolikrát po samohlásce bezprostředně následuje samohláska, dělíme číslo tak nalezené celkovým počtem samohlásek, čímž nalezneme p_{11} atd. Tak nalézá Markov hodnoty

$$p_{11} = \tfrac{1104}{8638} = 0{,}128,\ p_{21} = 0{,}663,$$

$$P_1(1 - P_1) \cdot \frac{1 + p_{11} - p_{21}}{1 - p_{11} + p_{21}} = 0{,}06. \tag{3}$$

Podle theorie o dispersi — viz rovnici (9), odst. 64 — je nutno v případě řetězu nahraditi pravou stranu rovnice (2), odst. 67 výrazem (3). Levá strana má, jak víme, hodnotu 0,051, takže souhlas je nyní lepší než v případě a).

Výsledek úvah shrneme takto: Rovnice (2) by platila, kdyby 200 skupin po 100 hláskách bylo vybráno z textu namátkou beze vztahu k souvislosti textu (kdybychom na př. každou hlásku napsali na zvláštní lístek, vložili lístky do osudí a pak vytahali a libovolně uspořádali do 200 skupin po 100 hláskách); jsou-li však hlásky ponechány ve skupinách tak, jak následují jedna za druhou v daném textu, přicházejí k platnosti pravděpodobnosti p_{ik}, máme případ řetězu a nutno pravou stranu rovnice (2) nahraditi výrazem (3).

68. Brownův pohyb po přímce. a) Malé hmotné částice (na př. zrnka pylového prášku) suspendované ve vodě jsou v ustavičném pohybu, jenž je způsobován nárazy molekul vody na částice. Pohyb toho druhu nazývá se *Brownův pohyb* podle botanika Browna, jenž jej pozoroval r. 1828. Předpokládáme, že molekuly vody se pohybují nepravidelně ve všech směrech a že je vhodné počítati se středními hodnotami pošinutí, která jsou vyvolána nárazy molekul. Zjednodušme úlohu co nejvíce, předpokládajíce, že pohyb nějaké částice, která je vystavena náhodným nárazům, děje se podél přímky. Pak jsme vedeni k úloze, kterou se zabýval Rayleigh r. 1880: Někdo kráčí po rovné cestě. Každý krok má stejnou délku a, buď vpřed nebo vzad; oba případy mají pro každý krok stejnou pravděpodobnost $\frac{1}{2}$ a to nezávisle na tom, jak dopadly kroky ostatní. Kam se chodec dostane po n krocích? Úsudek, že chodec se mnoho od původního svého místa nevzdálí, poněvadž učiní asi tolik kroků vpřed jako vzad, není správný. Počítejme každý krok dopředu za kladnou délku a, krok dozadu za zápornou délku $-a$; algebraický součet všech n délek budiž ma, dráha chodcem skutečně uražená. m je celé číslo buď kladné nebo záporné nebo rovné nule. Platí rovnice

$$E(ma) = a \,.\, E(m) = 0, \tag{1}$$

ale veličina

$$E[(ma)^2] = a^2 \,.\, E(m^2)$$

nerovná se nule.

Abychom určili $E(m^2)$, užijeme postupu obdobného tomu, kterého jsme užili v odst. 14b a 15b. Přiřadíme k-tému kroku veličinu $x^{(k)}$ rovnou $+1$ nebo -1 dle toho, je-li to krok vpřed nebo vzad. Pak bude

$$m = x^{(1)} + x^{(2)} + \ldots + x^{(n)},$$
$$E(x^{(k)}) = 0,\ E(x^{(k)})^2 = 1,\ E[x^{(k)} \,.\, x^{(l)}] = 0 \text{ pro } k \neq l;$$

správnost poslední rovnice vyplývá z nezávislosti k-tého a l-tého kroku. Z napsaných rovnic plyne, že

$$E(m) = E(x^{(1)}) + E(x^{(2)}) + \ldots + E(x^{(n)}) = 0$$

ve shodě s rovnicí (1) a

$$E(m^2) = E(x^{(1)})^2 + E(x^{(2)})^2 + \ldots + E(x^{(n)})^2 + \\ + 2E[x^{(1)} \,.\, x^{(2)}] + \ldots = n,$$

a tedy

$$E[(ma)^2] = na^2. \tag{2}$$

Je-li pravděpodobnost kroku vpřed stejně veliká jako kroku vzad a rovna $\frac{1}{2}$, a délka jednoho kroku rovna a, je střední hodnota druhé mocniny dráhy uražené po n krocích rovna na^2.

Kdyby všechny kroky byly vpřed (nebo všechny vzad), měla by úhrnná dráha délku na, a tedy její druhá mocnina by byla n^2a^2. Avšak dráha uražená po n krocích se zkracuje tím, že nejsou všechny kroky v témže směru; střední hodnota čtverce dráhy je podle (2) na^2 a nikoli n^2a^2.

Výsledek vyjádřený rovnicí (2) je matematicky skoro totožný s větou o střední hodnotě druhé mocniny chyby, která vzniká, měříme-li danou délku rozdělenou na n částí tak, že měříme každou část zvláště a výsledky sečteme (odst. 46a).

b) Odůvodnění rovnice (2) předpokládá, že pravděpodobnosti jednotlivých kroků vpřed nebo vzad jsou vzájemně nezávislé. Připusťme nyní, že tyto pravděpodobnosti jsou spojeny v jednoduchý Markovův řetěz. To znamená: byl-li některý krok vpřed, existují pravděpodobnosti

p_{11}, že bezprostředně následující krok bude také vpřed,
p_{12}, že bezprostředně následující krok bude vzad;

a dále: by-li některý krok vzad, jsou pravděpodobnosti

p_{21}, že bezprostředně následující krok bude vpřed,
p_{22}, že bezprostředně následující krok bude vzad.

Mimo to předpokládáme, že

$$p_{11} = p_{22} = p, \; p_{12} = p_{21} = 1 - p; \tag{3}$$

z toho plyne (viz odst. 61), že

$$P_1 = P_2 = \tfrac{1}{2}.$$

Po velmi velkém počtu kroků je tedy pravděpodobnost kroku vpřed rovna pravděpodobnosti kroku vzad.

Označme písmenem n úhrnný počet kroků; pro přehlednost budiž n sudé číslo. Délka kroku budiž a a celkový počet kroků vpřed budiž m'. Přiřaďme k-tému kroku číslo $x^{(k)}$, které je rovno 1, je-li krok vpřed, a které se rovná 0, je-li krok vzad; pak je

$$m' \doteq x^{(1)} + x^{(2)} + \ldots + x^{(n)}.$$

Podle (4), odst. 63 je pro veliké n střední hodnota čísla m' rovna $E(m') = nP_1 = \frac{1}{2}n$. Poněvadž z celkového počtu n kroků je m' vpřed a $(n - m')$ vzad, je uražená dráha rovna $(2m' - n)\,a$. Zaveďme úchylku

$$h = m' - \tfrac{1}{2}n = m' - nP_1,$$

tedy $m' = h + \frac{1}{2}n$; uražená dráha bude rovna $2ha$. Z toho plyne střední hodnota druhé mocniny uražené dráhy $4a^2 E(h^2)$. Hodnota $E(h^2) : n$ se počítá podle rovnice (10), odst. 64.

Píšeme-li p na místo p_{11}, platí

$$E(h^2) = E(m' - \tfrac{1}{2}n)^2 =$$

$$\tfrac{1}{4}n \frac{p}{1-p} + \frac{1-2p}{8} \cdot \frac{1-(2p-1)^n}{(1-p)^2}.$$

Střední hodnota druhé mocniny dráhy uražené po n krocích je tedy

$$4a^2 E(h^2) = \left[\frac{np}{1-p} + \frac{1-2p}{2} \cdot \frac{1-(2p-1)^n}{(1-p)^2}\right] a^2. \qquad (4)$$

Pro velmi velké hodnoty čísla n přechází (4) v přibližný vzorec

$$4a^2 E(h^2) = \frac{npa^2}{1-p}. \qquad (5)$$

Rovnice (4) zobecňuje rovnici (2) pro libovolnou hodnotu čísla. Kdyby pravděpodobnosti vztahující se k jednotlivým krokům byly nezávislé, bylo by vzhledem k (3)

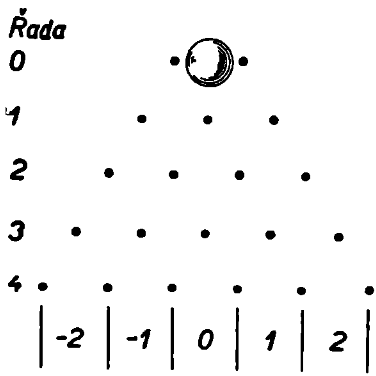
$$p_{11} = p_{21} = p,\ p_{12} = p_{22} = p = \tfrac{1}{2},$$

pravá strana rovnice (4) by se ztotožnila s pravou stranou rovnice (2).

V případě, že pravděpodobnost $p_{11} = p_{22} = p$ je veliká proti $p_{12} = p_{21} = 1 - p$, projevuje se závislost pravděpodobností tím, že chodec má jakousi „setrvačnost"; učiní-li k-tý krok vpřed (vzad), je pravděpodobnější, že $(k + 1)$-tý krok bude zase vpřed (vzad) než opačným směrem.

69. Theorie Galtonova přístroje. Galtonův přístroj je prkno, do kterého jsou zaraženy hřebíky ve vrcholech pravidelné sítě složené z rovnostranných trojúhelníků. Prkno se nakloní tak, aby hřebíky tvořily horizontální řady. Kulička, jejíž průměr je o něco menší (asi o čtvrtinu) než mezera mezi dvěma sousedními hřebíky, vloží se mezi dva hřebíky, které tvoří horní řadu. Kulička padá po nakloněné rovině, naráží na hřebíky a uchyluje se po každém nárazu buď napravo nebo nalevo; když prošla všemi vodorovnými řadami hřebíků, spadne do jedné z přihrádek umístěných pod spodní řadou hřebíků. Řadu, kde kuličku vkládáme do přístroje, označíme indexem 0, další řady pak 1, 2, 3, ... Přihrádky značíme indexy —1, —2, ... nebo 1, 2, ... dle toho, jsou-li nalevo nebo napravo od středu. Je-li n řad (nultou nepočítaje) a n sudé číslo, je $n + 1$ přihrádek s indexy $0, \pm 1, \pm 2, \ldots, \pm \frac{1}{2}n$. Obrazec č. 23 odpovídá případu $n = 4$.

Otázka zní: *Jak veliká je pravděpodobnost a_h, že kulička dopadne do přihrádky, označené číslem h, a jak veliká je střední*



Obr. 23.

hodnota $E(h^2)$? Budeme řešiti úlohu dvojím způsobem (pro n sudé).

a) Theorie s předpokladem o nezávislých pravděpodobnostech vychází z těchto podmínek:

I. Kulička vložená do horní mezery naráží na hřebík v řadě 1, který je pod onou mezerou, a uchýlí se buď napravo nebo nalevo. Pak narazí shora na nejbližší hřebík v řadě 2, zase se uchýlí napravo nebo nalevo, pak narazí na hřebík v řadě 3 atd. V každé horizontální řadě narazí na jediný hřebík; po n nárazech spadne do jedné z přihrádek.

II. Pravděpodobnost, že kulička při jednom (kterémkoli) nárazu se uchýlí napravo, rovná se pravděpodobnosti ($= \frac{1}{2}$), že se uchýlí nalevo.

III. Pravděpodobnost, že se kulička po nárazu na určitý hřebík uchýlí v určitém smyslu, nezávisí na tom, jak se uchýlila po nárazech v předcházejících řadách.*)

Kulička narazí n-krát; uchýlí-li se $\frac{1}{2}n$-krát napravo a tedy $\frac{1}{2}n$-krát nalevo, dopadne do přihrádky označené 0. Uchýlí-li se $(\frac{1}{2}n + h)$-krát napravo a tedy $(\frac{1}{2}n - h)$-krát nalevo, padne do přihrádky h ($h = 0, \pm 1, \pm 2, \ldots, \pm \frac{1}{2}n$). Všech možných případů (po každém z n nárazů napravo nebo nalevo) je 2^n; příznivých (t. j. těch, kdy mezi n nárazy je $(\frac{1}{2}n + h)$ takových, že po nich následuje úchylka napravo) je $(n)_{\frac{1}{2}n+h}$, takže

$$a_h = (n)_{\frac{1}{2}n+h} \cdot \frac{1}{2^n}. \tag{1}$$

Dále je střední hodnota čtverce úchylky:

$$E(h^2) = \sum_{h=-\frac{1}{2}n}^{+\frac{1}{2}n} a_h \, . \, h^2. \tag{2}$$

(1) je Newtonova formule (viz (1) v odst. 13, kde je dosaditi $p = \frac{1}{2}$, $m = \frac{1}{2}n + h$); podle vzorce (1), odst. 15 (zase $p = \frac{1}{2}$, $m = \frac{1}{2}n + h$) je

$$\frac{E(h^2)}{n} = \frac{1}{4}. \tag{3}$$

b) Theorie s předpokladem o závislých pravděpodobno-

---

*) Podmínka III nevyplývá z podmínky II; viz srovnání dvou úloh o tazích z osudí v odst. 66.

stech. Vzorec (3) dá se kontrolovati pokusy; v poslední době několik autorů se zabývalo theorií Galtonova přístroje a výsledky theorie byly srovnávány s pokusy.*) Ukázalo se, že vzorec (3) nevyhovuje a že lepší výsledek dává theorie, která podržuje podmínky I a II shora uvedené, která však nahrazuje podmínku III touto:

III. Pravděpodobnost p_{11}, že kulička se uchýlí napravo po nárazu na hřebík v některé řadě, uchýlila-li se po bezprostředně předcházejícím nárazu napravo, liší se obecně od pravděpodobnosti p_{21}, že se uchýlí napravo, uchýlila-li se po bezprostředně předcházejícím nárazu nalevo. Dále máme obecně různé pravděpodobnosti p_{12}, že kulička, která se uchýlila napravo, uchýlí se po následujícím náraze nalevo, a p_{22}, že po náraze nalevo se uchýlí zase nalevo.

Poněvadž celý přístroj je souměrný, předpokládáme

$$p_{11} = p_{22},\ p_{12} = p_{21}.$$

Padá-li kulička, jsou úchylky napravo či nalevo zjevy spojené v Markovův řetěz. Podmínky (1) a (2), odst. 60 a (2), odst. 61 jsou splněny, takže $P_1 = \frac{1}{2}$ podle (3), odst. 61. Pravděpodobnost $p_1^{(k)}$, že kulička po nárazu na hřebík v k-té řadě se uchýlí napravo, je konstantní:

$$p_1^{(1)} = p_1^{(2)} = \ldots = p_1^{(k)} = \ldots = p_1^{(n)} = \tfrac{1}{2}.$$

za předpokladu, že $p_1^{(1)} = \frac{1}{2}$. Řetěz je stacionární (odst. 65). Máme zde $h = m - \frac{1}{2}n$ a podle (9), odst. 64

$$\lim_{n\to\infty} E\left(\frac{h^2}{n}\right) = \frac{1}{4}\,\frac{1 + p_{11} - p_{21}}{1 - p_{11} + p_{21}} = \frac{1}{4}\,\frac{p_{11}}{1 - p_{11}}. \tag{4}$$

*) *B. Schulz*: Zur Theorie des Galtonschen Brettes (Zeitschr. f. Physik 92, 747—754; 1934). — *H. Münzner*: Über eine spezielle Markoffsche Kette am Galtonbrett (Zeitschr. f. angew. Mathematik und Mechanik 14, 343—346; 1934). — *W. Seitz-K. Hamacher-Odenhausen*: Untersuchungen über das Galton'sche Brett (Phys. Zeitschrift 35, 530—532; 1934). — *B. Hostinský*: Sur les probabilités relatives aux variables aléatoires liées entre elles (Annales de l'Institut Poincaré t. VII, fasc. II, p. 89—119; 1937).

Úloha o střední hodnotě h^2 je vlastně totožná s úlohou o lineárním Brownově pohybu (odst. 68).

Pravá strana (4) se může značně lišiti od pravé strany (3); je-li na př. $p_{11} = \frac{3}{4}$, $p_{21} = \frac{1}{4}$, je

$$\frac{1 + p_{11} - p_{21}}{1 - p_{11} + p_{21}} = \frac{1 + \frac{1}{2}}{1 - \frac{1}{2}} = 3.$$

Pro veliké n je tedy $E\left(\frac{h^2}{n}\right)$ třikrát větší podle vzorce (4) než podle (3).

c) Vzorec (4) lze kontrolovati pokusně; předpokládáme, že n je velké (stačí n rovné asi 20). Vložme do přístroje velký počet N kuliček a budiž N_h počet těch, které spadnou do přihrádky h. Pak je

$$E(h^2) = \frac{1}{N} \sum_{h=-\frac{1}{2}n}^{+\frac{1}{2}n} h^2 N_h; \quad \sum_{-\frac{1}{2}n}^{+\frac{1}{2}n} N_h = N.$$

Budiž a_{11} (a_{22}) celkový počet případů, ve kterých po úchylce kuličky napravo (nalevo) následuje bezprostředně úchylka napravo (nalevo), a a_{12} (a_{21}) počet těch, kdy po úchylce napravo (nalevo) následuje úchylka nalevo (napravo). Přibližné hodnoty pravděpodobností p_{11} a p_{21} jsou

$$p'_{11} = \frac{a_{11}}{a_{11} + a_{12}}, \quad p'_{21} = \frac{a_{21}}{a_{21} + a_{22}}.$$

Pokusy ukázaly, že v některých případech (při vhodném sklonu prkna) je $p'_{11} = \frac{3}{4}$, $p'_{21} = \frac{1}{4}$, takže theoretický vzorec (4) se potvrzuje. Okolnost, že p'_{11} je značně větší než p'_{21}, ukazuje k tomu, že kulička má jakousi *setrvačnost*: uchýlí-li se napravo, je pravděpodobnost, že po následujícím nárazu na hřebík se uchýlí napravo, značně větší než pravděpodobnost úchylky nalevo. Sklon prkna musí býti malý. Při větším sklonu se stává, že kulička proběhne uličkou mezi hřebíky a protne několik vodorovných řad, aniž by narazila na hřebík; podmínky I, II nejsou pak splněny.

Poznamenejme, že vzorec (10), odst. 64 dává přesnou hodnotu disperse pro jakýkoli počet n řad hřebíků.

70. Charakteristická rovnice. a) Rovnice druhého stupně v λ

$$\begin{vmatrix} \lambda p_{11} - 1, & \lambda p_{21} \\ \lambda p_{12}, & \lambda p_{22} - 1 \end{vmatrix} = 0, \tag{1}$$

neboli

$$\lambda^2(p_{11}p_{22} - p_{12}p_{21}) - \lambda(p_{11} + p_{22}) + 1 = 0 \tag{2}$$

nazývá se *charakteristickou rovnicí* příslušnou řetězu o pravděpodobnostech p_{ik}. Jeden její kořen je vždy $\lambda = 1$, neboť, dosadíme-li do determinantu (1) jednotku na místo λ, a přičteme-li pak prvky druhého řádku k příslušným prvkům prvého řádku, změní se determinant ve tvar

$$\begin{vmatrix} p_{11} + p_{12} - 1, & p_{21} + p_{22} - 1 \\ p_{12}, & p_{22} \end{vmatrix};$$

tento determinant se rovná nule vzhledem k rovnicím (1) a (1a), odst. 59. Z těchto dvou rovnic plyne, že

$$p_{11}p_{22} - p_{12}p_{21} = p_{11}(1 - p_{21}) - (1 - p_{11})\, p_{21} =$$
$$= p_{11} - p_{21};$$

rovnici (2) lze tedy psáti ve tvaru

$$\lambda^2 - \lambda \frac{p_{11} + p_{22}}{p_{11} - p_{21}} + \frac{1}{p_{11} - p_{21}} = 0;$$

z toho plyne, že druhý kořen rovnice (1) je

$$\lambda = \frac{1}{p_{11} - p_{21}} = \frac{1}{\delta},$$

užijeme-li označení zavedeného v odst. 60a.

b) Položme v rovnicích (7), odst. 59 $n = 1$; dostaneme pro $k = 1, 2$

$$P_{i1}^{(m+1)} = P_{i1}^{(m)} p_{11} + P_{i2}^{(m)} p_{21},$$
$$P_{i2}^{(m+1)} = P_{i1}^{(m)} p_{12} + P_{i2}^{(m)} p_{22}.$$

Za předpokladu, že všechny pravděpodobnosti p_{ik} jsou kladné, mají $P_{i1}^{(m)}$ a $P_{i2}^{(m)}$ pro $m \to \infty$ limitní hodnoty P_1 resp. P_2 (viz odst. 60); je tedy

$$\left.\begin{aligned} P_1 &= P_1 p_{11} + P_2 p_{21}, \\ P_2 &= P_1 p_{12} + P_2 p_{22}. \end{aligned}\right\} \tag{3}$$

Veličiny P_1 a P_2 se jeví jakožto řešení dvou lineárních homogenních rovnic (3), k nimž připojíme podle rovnice (5b), odst. 60 podmínku

$$P_1 + P_2 = 1. \tag{4}$$

Determinant rovnic (3) vzhledem k P_1 a k P_2 musí býti roven nule, tedy

$$\begin{vmatrix} p_{11} - 1, & p_{21} \\ p_{12}, & p_{22} - 1 \end{vmatrix} = 0.$$

Determinant je roven determinantu v rovnici (1) pro $\lambda = 1$, rovná se tedy skutečně nule. Veličiny P_1 a P_2 jsou jednoznačně určeny rovnicemi (3) a (4); jejich hodnoty se shodují s formulemi (4), odst. 60.