KAPITOLA DRUHÁ

OPĚTOVANÉ VZÁJEMNĚ NEZÁVISLÉ POKUSY

13. Pravděpodobnost různých výsledků v řadě opakovaných vzájemně nezávislých pokusů. a) Budiž p pravděpodobnost, že zjev Z se dostaví jakožto výsledek nějakého pokusu; $1 - p$ pak je pravděpodobnost, že Z se nedostaví. Pokus, jehož výsledkem je Z, nazveme zkrátka zdařeným, v opačném případě nezdařeným.

Vykonáme postupně n vzájemně nezávislých pokusů; pro každý z nich je pravděpodobnost p, že se zdaří, stejná. Klademe si otázku: jak veliká je pravděpodobnost P_m, že v řadě n pokusů bude právě m zdařených a tedy $n - m$ nezdařených $(0 \leqq m \leqq n)$?

Očíslujme n pokusů pořadovými čísly $1, 2, \ldots, n$ a vyvolme nejprve určitých m z těchto čísel jakožto pořadová čísla pokusů, které se mají zdařiti. Pravděpodobnost, že v řadě n pokusů budou zdařené právě na těchto m předepsaných místech, má hodnotu

$$p^m(1 - p)^{n-m}.$$

Ale v otázce shora položené nepřihlížíme k pořadí, ve kterém se vyskytují pokusy zdařené nebo nezdařené, nýbrž jen k celkovému počtu m zdařených. Proto musíme násobiti hodnotu právě nalezenou počtem kombinací bez opakování m-té třídy z n prvků (viz odst. 3c). Tak dostaneme vzorec pro hledanou pravděpodobnost P_m:

$$P_m = \frac{n!}{m!(n - m)!}\, p^m (1 - p)^{n-m}, \qquad (1)$$

který se někdy nazývá vzorcem Newtonovým. Pro $m = 0, 1, 2, \ldots n$ dává vzorec (1) $n + 1$ čísel $P_0, P_1, P_2, \ldots P_n$, která udávají pravděpodobnosti, že v řadě n pokusů se nezdaří ani jeden, resp. že se jich zdaří 1, 2, ..., n. Poznamenejme, že, pokud $0 \leq p \leq 1$ a pokud m a n jsou celá kladná čísla, je hodnota pravé strany (1) menší než jedna nebo nejvýše rovna jedné.

b) Poněvadž

$$\frac{n!}{n!(n-m)!} = (n)_m,$$

máme podle binomické formule

$$1 = [p + (1-p)]^n =$$
$$= p^n + (n)_1 p^{n-1}(1-p) + (n)_2 p^{n-2}(1-p)^2 + \ldots +$$
$$+ (n)_k p^{n-k}(1-p)^k + \ldots + (1-p)^n.$$

Jednotlivé členy tohoto výrazu představují pravděpodobnosti, že v řadě n pokusů bude $n, n-1, \ldots, n-k, \ldots, 0$ zdařených; součet všech těchto pravděpodobností se rovná jedné,

$$P_0 + P_1 + P_2 + \ldots + P_n = 1.$$

c) Položme si další otázku: Je-li n dáno, jak voliti m, aby pravděpodobnost P_m určená vzorcem (1) měla co největší hodnotu?

Vypočteme, neměníce čísel p a n, veličiny $P_0, P_1, \ldots, P_n$ a utvořme podíl u_m m-té veličiny k veličině $(m+1)$-té:

$$u_m = \frac{P_m}{P_{m+1}} = \frac{1-p}{p}\,\frac{m+1}{n-m}.$$

Hledaná hodnota indexu m, pro který P_m dosahuje největší hodnoty, má tu vlastnost, že $u_m < 1$ a že $u_{m-1} < 1$.

Z toho plynou dvě nerovnosti:

$$\frac{1-p}{p}\cdot\frac{m+1}{n-m} > 1, \frac{1-p}{p}\,\frac{m}{n-m+1} < 1,$$

které upraveny dají

$$(m + 1)(1 - p) > (n - m)p, \quad m(1 - p) < (n - m + 1)p,$$

aneb

$$np + p - 1 < m < np + p. \tag{2}$$

Číslo m je takto uzavřeno mezi dvěma mezemi, jichž rozdíl se rovná jedné. Není-li právě jedna z mezí rovna celému číslu (druhá mez je pak také celé číslo), je celé číslo m jednoznačně určeno nerovnostmi (2). Je-li n dosti veliké, takže lze čísla p a $p - 1$ vynechati vedle n, vychází z (2) np jakožto přibližná hodnota pro m (s chybou menší než 1). Je tedy, za předpokladu, že P_m má co největší hodnotu, přibližně np zdařených a přibližně $n(1 - p)$ nezdařených pokusů. Výsledek výpočtu vyjádříme takto:

Ze všech případů,) které se mohou vyskytnouti, vykonáme-li pokus n-krát, má největší pravděpodobnost ten, ve kterém počet zdařených pokusů se má k počtu nezdařených jako p k* $(1 - p)$.

14. Střední hodnota počtu zdařených pokusů. a) Počet m zdařených pokusů může míti hodnoty $m = 0, 1, 2, \dots, n$; příslušné pravděpodobnosti jsou $P_0, P_1, P_2, \dots, P_n$. Počet pokusů n považujeme za dané neproměnné číslo. Podle definice střední hodnoty (odst. 8) bude

$$\text{s. h.}\ (m) = \sum_{m=0}^{n} m \,.\, P_m.$$

Abychom ustanovili hodnotu tohoto součtu, položme $q = 1 - p$, $p + q = 1$, takže

$$\begin{aligned} 1 &= (p + q)^n = \\ &= p^n + (n)_1 p^{n-1} q + (n)_2 p^{n-2} q^2 + \dots + npq^{n-1} + q^n. \end{aligned}$$

Derivujme tuto rovnici dle p a násobme pak veličinou p; vychází

$$\begin{aligned} n(p + q)^{n-1} p = np^n &+ (n - 1)(n)_1\, p^{n-1} q + \\ &+ (n - 2)\,(n)_2 p^{n-2} q + \dots + npq^{n-1}, \end{aligned} \tag{1}$$

*) Při daném n je celkem $n + 1$ různých případů ($m = 0, 1, 2, \dots, n$).

aneb, zavedeme-li

$$(p+q)=1,\ p^n=P_n,\ (n)_1p^{n-1}q=P_{n-1},\ \dots,\ npq^{n-1}=P_1,$$

$$np=nP_n+(n-1)P_{n-1}+\dots+P_1=\sum_{m=0}^{n}m\,.\,P_m.$$

Je tedy

$$\text{s. h.}\ (m)=np. \tag{2}$$

b) Uvedeme ještě jiný důkaz rovnice (2) užívajíce methody, která se hodí i k jiným výpočtům.

Budiž $x^{(i)}$ hodnota přiřaděná i-tému pokusu takto:

$$x^{(i)}=1,\ \text{zdaří-li se } i\text{-tý pokus},$$

$$x^{(i)}=0,\ \text{nezdaří-li se } i\text{-tý pokus}.$$

Podle toho je

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

neboť v součtu na pravé straně je tolik sčítanců rovných 1, kolik pokusů se podaří (t. j. m); ostatní sčítanci jsou rovni nule.

Podle věty dokázané v odst. 10a je

$$\text{s. h.}\ (m)=\text{s. h.}\ (x^{(1)}+x^{(2)}+\dots+x^{(n)})=\sum_{i=1}^{n}\text{s. h.}\ (x^{(i)});$$

poněvadž pak

$$\text{s. h.}\ (x^{(i)})=1\,.\,p+0\,.\,(1-p)=p,$$

je

$$\text{s. h.}\ (m)=np. \tag{2}$$

Pro každý jednotlivý pokus má příslušná veličina $x^{(i)}$ hodnotu buď 1 nebo 0; její střední hodnota je p; součet všech n veličin $x^{(i)}$ má střední hodnotu n-krát větší, totiž np. Výsledek (2) vyjádříme větou:

V řadě n pokusů postupně provedených je střední hodnota počtu zdařených pokusů rovna np. (Srv. odst. 22.)

15. Střední hodnota druhé mocniny úchylky. Veličina $h = m - np$, která se rovná rozdílu mezi počtem m skutečně zdařených pokusů a mezi střední hodnotou np čísla m, nazývá se *úchylka*.

Střední hodnota úchylky se rovná nule, neboť

$$\text{s. h. } (m - np) = \text{s. h. } (m) - \text{s. h. } (np) = np - np = 0.$$

Počítejme střední hodnotu čtverce úchylky; podle definice je

$$\text{s. h. } (m - np)^2 = \sum_{m=0}^{n} P_m (m - np)^2.$$

Výpočet pravé strany provedeme dvojím způsobem:

a) Derivujme rovnici (1) odst. 14 dle p a násobme pak veličinou p. Vychází

$$p[n(n-1)(p+q)^{n-2}p + n(p+q)^{n-1}] =$$
$$= n^2p^n + (n-1)^2\,(n)_1 p^{n-1}q + \ldots .$$

Dosadíme-li sem

$$p + q = 1, \; p^n = P_n, \; (n)_1 p^{n-1}q = P_{n-1}, \ldots,$$

bude

$$p[n(n-1)p + n] = \sum_{m=0}^{n} m^2 P_m = \text{s. h. } (m^2).$$

K této rovnici připojíme další dvě (užíváme vztahu s. h. $(m) = np$)

$$-2n^2p^2 = -2 \text{ s. h. } (mnp),$$
$$n^2p^2 = \text{ s. h. } (n^2p^2).$$

Sečtením všech tří rovnic vychází

$$\text{s. h. } (m - np)^2 = np(1 - p). \tag{1}$$

b) Druhý důkaz vzorce (1) provedeme užívajíce methody vyložené v odst. 14b. Budiž zase $x^{(i)}$ veličina přiřaděná i-tému pokusu; $x^{(i)} = 1$ nebo 0 podle toho, zdařil-li se i-tý pokus nebo ne. Máme

$$
\begin{aligned}
&\text{s. h. } (m - np)^2 = \\
&= \text{s. h. } (x^{(1)} + x^{(2)} + \ldots + x^{(n)} - np)^2 = \\
&= \text{s. h. } [(x^{(1)} - p) + (x^{(2)} - p) + \ldots + (x^{(n)} - p)]^2 = \\
&= \text{s. h. } [\sum_{i=1}^{n}(x^{(i)} - p)^2 + 2\sum_{i<k}\sum(x^{(i)} - p)(x^{(k)} - p)] = \\
&= \sum_{i=1}^{n} \text{s. h. } (x^{(i)} - p)^2 + 2\sum_{i<k}\sum \text{s. h. } (x^{(i)} - p)(x^{(k)} - p); \qquad (2)
\end{aligned}
$$

dvojnásobný součet se vztahuje ke všem dvojicím indexů i a k $(i < k)$ utvořeným z čísel od 1 do n.

Poněvadž $x^{(i)}$ může nabýti jen hodnot 1 a 0 s pravděpodobnostmi p resp. $(1 - p)$, je

$$\text{s. h. } (x^{(i)} - p)^2 = (1 - p)^2 \,.\, p + (-p)^2(1 - p) = p(1-p)$$

a tedy

$$\sum_{i=1}^{n} \text{s. h. } (x^{(i)} - p)^2 = np(1 - p). \qquad (3)$$

Ještě je třeba vypočísti s. h. $(x^{(i)} - p)(x^{(k)} - p)$ pro $i \neq k$. Výpočet provedeme dvojím způsobem. *Předně* uvážíme, že součin $(x^{(i)} - p)(x^{(k)} - p)$ má (pro $x^{(i)} = 1, 0$; $x^{(k)} = 1, 0$) možné hodnoty

$$(1 - p)^2, \; (1 - p) \,.\, -p, \; -p \,.\, (1 - p), \; -p \,.\, -p$$

a že příslušné pravděpodobnosti jsou

$$p^2, \; p(1 - p), \; (1 - p)p, \; (1 - p)^2.$$

Násobíce každou z uvedených hodnot příslušnou pravděpodobností a sečtouce čtyři součiny tak utvořené dostáváme

$$\text{s. h. } (x^{(i)} - p)(x^{(k)} - p) = 0.$$

Druhý způsob výpočtu se zakládá na větě dokázané v odst. 10b: poněvadž výsledek i-tého pokusu nemá vlivu na podmínky k-tého pokusu, je

$$\text{s. h. } [(x^{(i)} - p)(x^{(k)} - p)] = \text{s. h. } (x^{(i)} - p) \cdot \text{s. h. } (x^{(k)} - p),$$

a ježto

$$\text{s. h. } (x^{(i)} - p) = (1 - p)p + (-p)(1 - p) = 0,$$

máme zase

$$\text{s. h. } (x^{(i)} - p)(x^{(k)} - p) = 0.$$

Každý člen dvojitého součtu v (2) je tedy roven nule a vzhledem k (3) obdržíme zase

$$\text{s. h. } (m - np)^2 = np(1 - p); \tag{1}$$

tento vzorec udává *střední hodnotu čtverce úchylky*. Konáme-li jen jeden pokus ($n = 1$), je s. h. $(x^{(1)} - p)^2 = p(1 - p)$; konáme-li n pokusů, jest s. h. druhé mocniny úchylky n-krát větší.

c) Někdy se zavádí do počtu t. zv. *relativní úchylka*, t. j. úchylka $m - np$ dělená počtem pokusů: $(m : n) - p$. Patrně je

$$\text{s. h. } \left(\frac{m}{n} - p\right)^2 = \frac{p(1 - p)}{n}. \tag{4}$$

16. Bernoulliova věta. Podle Čebyševovy věty (nerovnosti (3) odst. 12) je při $\varepsilon > 0$

$$P\left(-\varepsilon < \frac{m}{n} - p < \varepsilon\right) \geqq 1 - \frac{\text{s. h. } \left(\frac{m}{n} - p\right)^2}{\varepsilon^2}$$

a tedy vzhledem ke vzorci (4) odst. 15.

$$P\left(-\varepsilon < \frac{m}{n} - p < \varepsilon\right) \geqq 1 - \frac{p(1 - p)}{n\varepsilon^2}.$$

Z toho plyne dále, že

$$\lim_{n=\infty} P\left(-\varepsilon < \frac{m}{n} - p < \varepsilon\right) = 1,$$

což je právě Bernoulliova věta:

Pravděpodobnost, že relativní úchylka $(m : n) - p$ *nebude číselně větší než dané, jakkoli malé, číslo* ε, *blíží se jistotě, když počet pokusů* n *roste do nekonečna.*

Tato věta dokázaná ve spise Ars conjectandi Jakuba Bernoulliho vyšlém r. 1713 (osm let po autorově smrti), je jedním z hlavních výsledků počtu pravděpodobnosti; v odst. 24 pojednáme o tom, jak byla později zobecněna.

Připomeňme, že v uvedeném důkaze Bernoulliovy věty a vůbec ve výpočtech odst. 13—16 užíváme předpokladu o vzájemné nezávislosti jednotlivých pokusů; při každém pokuse je konstantní pravděpodobnost p, že se pokus zdaří, nezávislá na tom, jak dopadly pokusy ostatní.

17. Wallisova formule. V odst. 17—19 uvedeme důkazy některých pomocných vzorců, kterých se užívá v různých výpočtech pravděpodobností.

Budiž m celé kladné číslo a počítejme hodnotu integrálu

$$A_m = \int_0^{\frac{1}{2}\pi} \sin^m x \, dx.$$

Integrujíce po částech obdržíme

$$A_m = \left[-\cos x \sin^{m-1} x\right]_0^{\frac{1}{2}\pi} + (m-1)\int_0^{\frac{1}{2}\pi} \sin^{m-2} x \,.\, \cos^2 x \, dx;$$

výraz $[\ldots]_0^{\frac{1}{2}\pi}$ je roven nule. Píšeme-li $1 - \sin^2 x$ místo $\cos^2 x$, je

$$A_m = (m-1)\int_0^{\frac{1}{2}\pi} \sin^{m-2} x \, dx - (m-1)\int_0^{\frac{1}{2}\pi} \sin^m x \, dx =$$

$$= (m-1)\, A_{m-2} - (m-1) A_m$$

a tedy

$$A_m = \frac{m-1}{m} A_{m-2},$$

$$A_{m-2} = \frac{m-3}{m-2} A_{m-4},$$

. . . .

Řada rovnic, které takto dostaneme snižujíce index m postupně o dvě jednotky, končí, je-li $m = 2p$ sudé číslo, rovnicí

$$A_0 = \int_0^{\frac{1}{2}\pi} \mathrm{d}x = \tfrac{1}{2}\pi;$$

je-li $m = 2p + 1$ liché číslo, končí řada rovnicí

$$A_1 = \int_0^{\frac{1}{2}\pi} \sin x \, \mathrm{d}x = 1.$$

Znásobme všechny rovnice; v případě sudého m dostaneme

$$A_{2p} = \frac{2p-1}{2p} \cdot \frac{2p-3}{2p-2} \cdots \tfrac{1}{2} \cdot \tfrac{1}{2}\pi$$

a v případě lichého m

$$A_{2p+1} = \frac{2p}{2p+1} \cdot \frac{2p-2}{2p-1} \cdots \tfrac{2}{3}.$$

Poněvadž uvnitř integračního intervalu $(0, \frac{1}{2}\pi)$ platí

$$\sin x < 1, \ \sin^{m+1} x < \sin^m x,$$

zmenšuje se A_m s rostoucím m, takže

$$A_{2p+1} < A_{2p} < A_{2p-1}$$

čili

$$\frac{2p}{2p+1} \frac{2p-2}{2p-1} \cdots \tfrac{2}{3} < \frac{2p-1}{2p} \frac{2p-3}{2p-2} \cdots \tfrac{1}{2} \tfrac{1}{2}\pi <$$

$$< \frac{2p-2}{2p-1} \cdot \frac{2p-4}{2p-3} \cdots \tfrac{2}{3}.$$

Po úpravě dostaneme

$$\frac{2^2 \cdot 4^2 \ldots (2p-2)^2 (2p)^2}{3^2 \cdot 5^2 \ldots (2p-1)^2 (2p+1)} < \tfrac{1}{2}\pi < \frac{2^2 \cdot 4^2 \ldots (2p-2)^2 \cdot 2p}{3^2 \cdot 5^2 \ldots (2p-1)^2}.$$

Položme

$$F(p) = \frac{2 \,.\, 2 \,.\, 4 \,.\, 4 \quad \ldots 2p \,.\, 2p}{1 \,.\, 3 \,.\, 3 \,.\, 5 \,.\, 5 \ldots (2p - 1)(2p + 1)};$$

předešlé nerovnosti dají se napsati takto:

$$F(p) < \tfrac{1}{2}\pi < F(p) \,.\, \frac{2p + 1}{2p}. \tag{1}$$

Funkce $F(p)$ roste, roste-li p, neboť

$$\frac{F(p + 1)}{F(p)} = \frac{(2p + 2)\,(2p + 2)}{(2p + 1)\,(2p + 3)} = \frac{4p^2 + 8p + 4}{4p^2 + 8p + 3} > 1;$$

poněvadž je stále menší než $\frac{1}{2}\pi$, má limitu pro $\lim p = \infty$. Vzhledem k (1) je $\lim\limits_{p=\infty} F(p) = \frac{1}{2}\pi$

nebo

$$\lim_{p=\infty} \left[\frac{2 \,.\, 2 \,.\, 4 \,.\, 4 \ldots 2p \,.\, 2p}{1 \,.\, 3 \,.\, 3 \,.\, 5 \ldots (2p - 1) \,.\, (2p + 1)} \right] = \tfrac{1}{2}\pi. \tag{2}$$

To je *Wallisova formule* z r. 1655.

18. Stirlingova formule. Budiž n celé kladné číslo. Hodnota výrazu $1 \,.\, 2 \,.\, 3 \,.\, \ldots \,.\, n = n!$ dá se vyjádřiti pro případ, že n je veliké číslo, přibližnou formulí, kterou máme odvoditi. Položme

$$\varphi(n) = \frac{n!}{\sqrt{2\pi} \,.\, e^{-n}\, n^{n+\frac{1}{2}}}.$$

Ze vzorců

$$[\varphi(n)]^4 = \frac{(n!)^4}{4\pi^2\, e^{-4n}\, n^{4n+2}},\quad [\varphi(2n)]^2 = \frac{[(2n)!]^2}{2\pi\, e^{-4n}\, n^{4n+1}\, 2^{4n+1}}$$

plyne, že

$$\frac{[\varphi(n)]^4}{[\varphi(2n)]^2} = \frac{(2n + 1) \,.\, 2 \,.\, 2 \,.\, 4 \,.\, 4 \ldots (2n) \,.\, (2n)}{\pi \,.\, n \,.\, 1 \,.\, 3 \,.\, 3 \,.\, 5 \,.\, 5 \ldots (2n - 1) \,.\, (2n + 1)}.$$

Podle Wallisovy formule (2) odst. 17. má pravá strana této rovnice za limitu 1, roste-li n do nekonečna. Je tedy

$$\lim_{n=\infty} \frac{[\varphi(n)]^4}{[\varphi(2n)]^2} = 1 \text{ a z toho } \lim_{n=\infty} \frac{[\varphi(n)]^2}{\varphi(2n)} = 1. \qquad (1)$$

Dále je

$$\frac{\varphi(n)}{\varphi(n+1)} = \frac{n!}{(n+1)!} \frac{e^{-n-1}(n+1)^{n+\frac{3}{2}}}{e^{-n} n^{n+\frac{1}{2}}} = e^{-1}\left(1+\frac{1}{n}\right)^{n+\frac{1}{2}},$$

$$\lg \frac{\varphi(n)}{\varphi(n+1)} = -1 + (n+\tfrac{1}{2}) \lg\left(1+\frac{1}{n}\right).$$

Rozviňme $\lg\left(1+\frac{1}{n}\right)$ v Maclaurinovu řadu postupující podle mocnin proměnné $\frac{1}{n}$. Vychází

$$(n+\tfrac{1}{2}) \lg\left(1+\frac{1}{n}\right) = \sum \frac{(-1)^k}{(k+1)\, n^k} + \sum \frac{(-1)^{k+1}}{2k\, n^k},$$

$$\lg \frac{\varphi(n)}{\varphi(n+1)} = \sum_{k=2}^{\infty} \frac{(-1)^k}{n^k} \left[\frac{1}{k+1} - \frac{1}{2k}\right] =$$

$$= \sum_{k=2}^{\infty} \frac{(-1)^k (k-1)}{n^k\, 2k\, (k+1)} = \frac{1}{12n^2} - \frac{1}{12n^3} + \dots$$

Členy této řady se zmenšují co do absolutní hodnoty, roste-li index k, a mají střídavá znamení. Proto platí

$$0 < \lg \frac{\varphi(n)}{\varphi(n+1)} < \frac{1}{12n^2},\ 1 < \frac{\varphi(n)}{\varphi(n+1)} < e^{\frac{1}{12n^2}}.$$

Připojme k poslední nerovnosti dalších $(n-1)$, které dostaneme dosazujíce postupně $(n+1)$, $(n+2)$, ..., $(2n-1)$ na místo n:

$$1 < \frac{\varphi(n+1)}{\varphi(n+2)} < e^{\frac{1}{12(n+1)^2}}$$

$$1 < \frac{\varphi(2n+1)}{\varphi(2n)} \quad e^{\frac{1}{12(2n-1)^2}}.$$

Znásobme všech těchto n nerovností; mocnitel čísla e napravo bude

$$\frac{1}{12}\left[\frac{1}{n^2} + \frac{1}{(n+1)^2} + \ldots + \frac{1}{(2n-1)^2}\right] < n\frac{1}{12n^2} = \frac{1}{12n},$$

takže
$$1 < \frac{\varphi(n)}{\varphi(2n)} < e^{\frac{1}{12n}}, \quad \lim_{n=\infty} \frac{\varphi(n)}{\varphi(2n)} = 1.$$

Dělme výraz, jenž stojí za znamením lim v poslední rovnici, výrazem, jenž se vyskytuje v rovnici (1). Pro $\lim n = \infty$ vychází $\lim\limits_{n=\infty}\varphi(n) = 1$ aneb, vypíšeme-li $\varphi(n)$,

$$\lim_{n=\infty} \frac{n!}{\sqrt{2\pi}\, e^{-n}\, n^{n+\frac{1}{2}}} = 1. \tag{2}$$

Je-li tedy n veliké číslo, můžeme nahraditi $n!$ asymptotickým výrazem:

$$n! \sim \sqrt{2\pi}\, e^{-n}\, n^{n+\frac{1}{2}}.$$

Formule (2) pochází od *Stirlinga* (1730). Důkaz na základě Wallisovy formule zde uvedený pochází od *J. A. Serreta.*

19. Laplaceův integrál a jiné pomocné vzorce. a) Abychom určili hodnotu L Laplaceova integrálu

$$L = \int_{-\infty}^{+\infty} e^{-x^2}\, dx,$$

zavedeme v rovině Oxy polární souřadnice. Druhou mocninu integrálu L považujeme za dvojnásobný integrál:

$$L^2 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-x^2-y^2}\, dx\, dy$$

vztažený k celé rovině Oxy. Polární souřadnice: průvodič r a polární úhel φ souvisí s x, y podle rovnic

$$x = r\cos\varphi,\ y = r\sin\varphi;\ \frac{D(x, y)}{D(r, \varphi)} = r;$$

element plošného obsahu v polárních souřadnicích je roven $r\,dr\,d\varphi$ (je to obsah čtyrúhelníka omezeného jednak dvěma průvodiči příslušnými polárním úhlům φ a $\varphi + d\varphi$, jednak dvěma kruhovými oblouky o společném středu 0 a o poloměrech r a $r + dr$). Je tedy

$$L^2 = \int_0^{2\pi} \int_0^{\infty} e^{-r^2}\, r\, dr\, d\varphi,$$

kterýžto integrál je roven součinu dvou integrálů

$$\int_0^{2\pi} d\varphi = 2\pi \quad \text{a} \quad \int_0^{\infty} e^{-r^2}\, r\, dr = \tfrac{1}{2};$$

je tedy $L^2 = \pi$ a z toho plyne hledaný vzorec pro L:

$$L = \int_{-\infty}^{+\infty} e^{-x^2}\, dx = \sqrt{\pi}.$$

Poněvadž e^{-x^2} je sudá funkce, je

$$\int_{-\infty}^{0} e^{-x^2}\, dx = \int_0^{\infty} e^{-x^2}\, dx = \tfrac{1}{2}\int_{-\infty}^{+\infty} e^{-x^2}\, dx = \tfrac{1}{2}\sqrt{\pi}. \qquad (1)$$

b) Abychom ustanovili hodnotu integrálu

$$I_r = \int_0^{\infty} e^{-x^2}\, x^r\, dx,$$

kde r je libovolné celé kladné číslo, vyjděme z rovnice platné pro každé $r > 2$:

$$I_r = \left[-\tfrac{1}{2}\, e^{-x^2}\, x^{r-1}\right]_0^{\infty} + \frac{r-1}{2} \int_0^{\infty} e^{-x^2}\, x^{r-2}\, dx$$

aneb

$$I_r = \frac{r-1}{2}\, I_{r-2}.$$

Je-li r sudé, $r = 2m$, je

$$I_{2m} = \frac{2m-1}{2} \cdot \frac{2m-3}{2} \dots \tfrac{1}{2} \cdot I_0;$$

I_0 jest integrál (1), tedy

$$I_{2m} = \frac{1 \cdot 3 \cdot 5 \dots (2m-1)}{2^{m+1}} \sqrt{\pi}. \tag{2}$$

Pro liché r, $r = 2m + 1$, je

$$I_{2m+1} = \frac{2m}{2} \cdot \frac{2m-2}{2} \dots \tfrac{2}{2} \cdot I_1$$

a poněvadž

$$I_1 = \int_0^\infty e^{-x^2} x \, dx = \left[-\frac{e^{-x^2}}{2} \right]_0^\infty = \tfrac{1}{2},$$

vychází

$$I_{2m+1} = \frac{2 \cdot 4 \cdot 6 \dots (2m)}{2^{m+1}} = \frac{m!}{2}. \tag{3}$$

20. Přibližný vzorec pro P_m. Zavedení spojité proměnné. a) Podle Newtonova vzorce (viz (1) v odst. 13.) je

$$P_m = \frac{n!}{m!\,(n-m)!} p^m (1-p)^{n-m}$$

pravděpodobnost, že v řadě n pokusů bude m zdařených. Zavedeme-li úchylku h rovnicí

$$m = np + h,$$

nabude hořejší vzorec tvaru

$$P_m = \frac{n!}{(np+h)!\,(n-np-h)!} p^{np+h} (1-p)^{n-np-h}.$$

Předpokládejme, že n je tak veliké číslo, že lze faktoriály nahraditi přibližnými výrazy podle Stirlingovy formule (odst. 18.). Po úpravě dostaneme tuto přibližnou hodnotu pro P_m:

$$P_m \doteq \frac{\left(1+\frac{h}{np}\right)^{-np-h-\frac{1}{2}}\left(1-\frac{h}{n-np}\right)^{-n+np+h-\frac{1}{2}}}{\sqrt{2\pi\, np\,(1-p)}}.$$

Položme

$$A=\left(1+\frac{h}{np}\right)^{-np-h-\frac{1}{2}},\; B=\left(1-\frac{h}{n-np}\right)^{-n+np+h-\frac{1}{2}}$$

a rozviňme $\lg A$ a $\lg B$ v Maclaurinovy řady postupující podle mocnin veličiny $\frac{h}{n}$, při čemž budeme předpokládati, že $\frac{h}{\sqrt{n}}$ *je menší než určité konečné číslo* a že tedy veličiny

$$\frac{h}{n}=\frac{h}{\sqrt{n}}\cdot\frac{1}{\sqrt{n}},\; \frac{h^3}{n^2}=\left(\frac{h}{\sqrt{n}}\right)^3\frac{1}{\sqrt{n}}$$

jsou libovolně malé, roste-li n neomezeně. Tak dostaneme

$$\lg A = -\,(np+h+\tfrac{1}{2})\lg\left(1+\frac{h}{np}\right)=$$

$$=(-np-h-\tfrac{1}{2})\left(\frac{h}{np}-\tfrac{1}{2}\frac{h^2}{n^2p^2}+\tfrac{1}{3}\frac{h^3}{n^3p^3}-\ldots\right)=$$

$$=-h-\tfrac{1}{2}\frac{h^2}{np}+\ldots$$

$$\lg B=(-n+np+h-\tfrac{1}{2})\lg\left(1-\frac{h}{n-np}\right)=$$

$$=(-n+np+h-\tfrac{1}{2})\left(-\frac{h}{n-np}-\tfrac{1}{2}\frac{h^2}{(n-np)^2}-\ldots\right)=$$

$$=h-\frac{h^2}{2n(1-p)}+\ldots$$

V obou řadách jsme podrželi (vedle členu h nekonečné velkého) jen konečné veličiny; vynechané členy jsou nekonečně malé. Máme tedy

$$\lg(AB) = -\frac{h^2}{2np(1-p)},\quad AB = e^{-\frac{h^2}{2np(1-p)}}$$

a

$$P_m = \frac{e^{-\frac{h^2}{2np(1-p)}}}{\sqrt{2\pi\, np(1-p)}}. \tag{1}$$

Z tohoto vzorce plyne, že pravděpodobnost P_m dosahuje maximální hodnoty v případě, že $h = 0$, t. j. když $m = np$; v tom případě je počet zdařených pokusů ($= np$) v poměru $p : (1 - p)$ k počtu nezdařených (srv. odst. 13c) a máme

$$P_{m_{max}} = \frac{1}{\sqrt{2\pi\, np(1-p)}} \tag{2}$$

Roste-li úhrnný počet pokusů n do nekonečna, konverguje i tato maximální hodnota pravděpodobnosti P_m k nule.

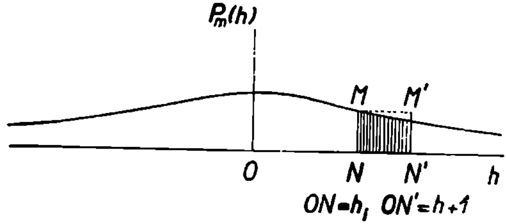
b) Ve vzorci $h = m - np$ může m nabývati hodnot $0, 1, 2, \ldots n$ a tedy h jen hodnot $-np, -np + 1, \ldots, -np + n$. Ale k některým výpočtům se nám hodí považovati ve vzorci (1) h za spojitě proměnnou veličinu.

Pro veliké n je nejen P_m, vyjádřená přibližně pravou stranou vzorce (1), malá, nýbrž také derivace podle h je malá, neboť

$$\frac{dP_m}{dh} = -\frac{h}{\sqrt{2\pi}\,[np(1-p)]^{\frac{3}{2}}}\, e^{-\frac{h^2}{2np(1-p)}}.$$

Tato okolnost dovoluje vyjádřiti P_m jakožto plochu. V diagramu (obr. 1), jenž udává P_m jakožto funkci proměnné h, je, poněvadž tečna křivky má přibližně nulový sklon k ose Oh, (vyčárkovaná) plocha omezená dvěma pořadnicemi, příslušnými úsečkám h a $h + 1$, přibližně rovna obdélníku

$NN'M'M$ o základně $= 1$ a o výšce $NM = P_m(h)$. P_m je číselně rovna ploše obdélníka a tedy také velmi přibližně oné vyčárkované ploše. Pravděpodobnost, že úchylka jest rovna buď h_1, nebo $h_1 + 1$, nebo $h_1 + 2, \ldots$ nebo h_2 (jinými slovy: že jest obsažena v mezích h_1 až h_2), rovná se součtu $P_m(h_1) +$



Obr. 1.

$+ P_m(h_1 + 1) + \ldots + P_m(h_2)$. Tento součet se dá nahraditi součtem obdélníků takových jako je $NN'M'M$ anebo součtem vyčárkovaných ploch, který se rovná integrálu ($h_1 <$ $< h_2$)

$$\int_{h_1}^{h_2+1} \frac{e^{-\frac{h^2}{2np(1-p)}}}{\sqrt{2\pi\, np\,(1-p)}}\, dh.$$

Pro případ, že n a h_2 jsou velká čísla, můžeme psáti h_2 na místo $h_2 + 1$ a máme vzorec

$$P(h_1 < m - np < h_2) = \int_{h_1}^{h_2} \frac{e^{-\frac{h^2}{2np(1-p)}}}{\sqrt{2\pi\, np\,(1-p)}}\, dh. \qquad (3)$$

Výsledek výpočtů shrneme touto větou:

Budiž p konstantní pravděpodobnost, že nějaký pokus se zdaří, n počet vzájemně nezávislých pokusů, m počet zdařených mezi nimi a tedy np střední hodnota počtu zdařených pokusů; pravděpodobnost, že úchylka $m - np$ jest obsažena v mezích

h_1 a h_2 ($h_1 < h_2$), jest udána přibližně vzorcem (3) a to tím přesněji, čím je n větší.

Připomeňme ještě předpoklady, za kterých byla odvozena formule (3); n je tak veliké, že faktoriály v původní formuli pro P_m (viz (1), odst. 13.) se dají nahraditi přibližnými výrazy podle Stirlingovy formule; $\frac{h}{\sqrt{n}}$ zůstává menší než určité konečné číslo; h_2 je tak veliké, že v horní mezi integrálu (3) můžeme psáti h_2 místo $h_2 + 1$.

Počítáme-li podle (3) pravděpodobnost, že úchylka, $m - np$ jest obsažena v mezích $-\infty \ldots +\infty$, vychází

$$P(-\infty < m - np < +\infty) = \int_{-\infty}^{+\infty} \frac{e^{-\frac{h^2}{2np(1-p)}}}{\sqrt{2\pi\, np\, (1-p)}}\, dh$$

a zavedeme-li integrační proměnnou u rovnicí

$$u\sqrt{2np\,(1-p)} = h, \quad \sqrt{2np\,(1-p)}\, du = dh,$$

je (viz odst. 19a)

$$P(-\infty < m - np < +\infty) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-u^2}\, du = 1.$$

Tento výpočet však není přesný, neboť vzorec (3) platí jen pro velké n; h nemůže býti větší než n, zde se však integruje podle h, při konstantním n, od $-\infty$ do $+\infty$. K přesnějším výpočtům se doporučuje užívati původní Newtonovy formule (1) odst. 13. pro P_m.

c) Jakožto příklad uveďme pokusy s mincí. Padne-li na líc, považujeme pokus za zdařený, padne-li na rub, za nezdařený. Poněvadž je zde

$$p = 1 - p = \tfrac{1}{2},$$

platí

$$np = \frac{n}{2},\ h = m - \frac{n}{2},\ P_m = \frac{n!}{(\frac{1}{2}n + h)!\,(\frac{1}{2}n - h)!} \cdot \frac{1}{2^n}.$$

Následující tabulky udávají hodnotu P_m jakožto funkci veličiny h a to pro $n = 2, 4, 6, 8$:

$n = 2$

h	-1	0	1
P_m	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$n = 4$

h	-2	-1	0	1	2
P_m	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

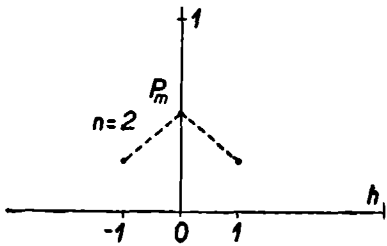
$n = 6$

h	-3	-2	-1	0	1	2	3
P_m	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$

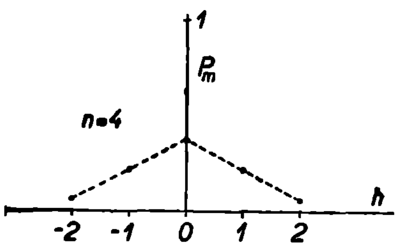
$n = 8$

h	-4	-3	-2	-1	0	1	2	3	4
P_m	$\frac{1}{256}$	$\frac{8}{256}$	$\frac{28}{256}$	$\frac{56}{256}$	$\frac{70}{256}$	$\frac{56}{256}$	$\frac{28}{256}$	$\frac{8}{256}$	$\frac{1}{256}$

Příslušné čtyři diagramy (obr. 2. až 5.)

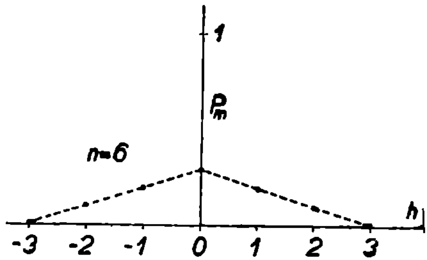


Obr. 2.

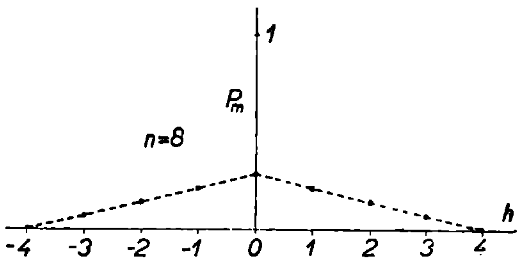


Obr. 3.

ukazují, jak se lomená čára, spojující po dvou sousední body diagramu, blíží — roste-li n — ke křivce, která probíhá velmi blízko osy Oh a která má také všude velmi malý sklon (srv. hořejší graf v odst. b), obr. 1.).



Obr. 4.



Obr. 5.

21. Laplaceova věta. Číselné příklady. a) Rovnice (3) odst. 20. vyjadřuje t. zv. Laplaceovu větu (kterou však znal již dříve Bayes). K odhadu pravděpodobností, že úchylka je v daných mezích, upravíme onu rovnici na jednodušší tvar. Za tím účelem zavedeme pomocnou funkci proměnné t:

$$\Theta(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-x^2}\, dx,$$

jejíž hodnoty najdeme v tabulkách. Uveďme výtah z tabulek:

t	$\Theta(t)$	t	$\Theta(t)$
0,00	0,0000000	1,20	0,9103140
0,20	0,2227025	1,50	0,9661052
0,40	0,4283922	2,00	0,9953223
0,50	0,5204999	3,00	0,9999779
1,00	0,8427008	4,00	0,9999999847

$\Theta(t)$ se tedy velmi rychle blíží jedné, roste-li t.

Zaveďme do rovnice (3) odst. 20. integrační proměnnou x rovnicí

$$x\sqrt{2np\,(1-p)} = h, \quad \sqrt{2np\,(1-p)}\,.\,\mathrm{d}x = \mathrm{d}h;$$

pro $h_1 < h_2$ a pro veliké n dostáváme

$$P(h_1 < m - np < h_2) = \\ = \tfrac{1}{2}\Theta\left(\frac{h_2}{\sqrt{2np\,(1-p)}}\right) - \tfrac{1}{2}\Theta\left(\frac{h_1}{\sqrt{2np\,(1-p)}}\right). \qquad (1)$$

Klademe-li $h_1 = 0$, $h_2 = h > 0$, je

$$P\,(0 < m - np < h) = \tfrac{1}{2}\Theta\left(\frac{h}{\sqrt{2np\,(1-p)}}\right);$$

klademe-li $-h_1 = h_2 = h > 0$, je

$$P\,(-h < m - np < h) = \Theta\left(\frac{h}{\sqrt{2np\,(1-p)}}\right)$$

nebo

$$P\,(|m - np| < h) = \Theta\left(\frac{h}{\sqrt{2np\,(1-p)}}\right).$$

b) Píšeme-li

$$t_1 = \frac{h_1}{\sqrt{2np(1-p)}}, \quad t_2 = \frac{h_2}{\sqrt{2np(1-p)}},$$

vychází Laplaceova věta (1) v této úpravě (pro $t_1 < t_2$):

$$P[np + t_1\sqrt{2np(1-p)} < m < np + t_2\sqrt{2np(1-p)}] =$$
$$= \int_{t_1}^{t_2} \frac{e^{-x^2}}{\sqrt{\pi}}\,dx. \tag{2}$$

Věta (2) platí přibližně pro velké hodnoty čísla n; přesně vzato je

$$\lim_{n=\infty} P[np + t_1\sqrt{2np(1-p)} < m < np + t_2\sqrt{2np(1-p)}] =$$
$$= \int_{t_1}^{t_2} \frac{e^{-x^2}}{\sqrt{\pi}}\,dx. \tag{2a}$$

Pravděpodobnost, že v řadě n pokusů bude počet zdařených obsažen v mezích

$$np - t\sqrt{2np(1-p)} \quad \text{a} \quad np + t\sqrt{2np(1-p)}$$

rovná se $\Theta(t)$ pro velké n.

Tato pravděpodobnost závisí tedy jen na čísle t a nikoli na p ani na n.

Volme za t postupně čísla 0,476936, 1, 2, 3, 4. Z tabulek funkce $\Theta(t)$ najdeme tyto pravděpodobnosti:

Následující tabulka obsahuje v levém sloupci meze pro úchylku h a v pravém sloupci příslušné hodnoty pravděpodobnosti, že h je v těch mezích:

$$\left.\begin{array}{r|l} np \pm 0{,}4769 \,.\, \sqrt{2np(1-p)} & 0{,}5 \\ np \pm \sqrt{2np(1-p)} & 0{,}8427 \\ np \pm 2\sqrt{2np(1-p)} & 0{,}9953 \\ np \pm 3\sqrt{2np(1-p)} & 0{,}99998 \end{array}\right\} \tag{3}$$

Příklad 1.: Házíme penízem; pravděpodobnost, že padne líc, je $p = \frac{1}{2}$. Dosazujeme-li do předešlých čtyř řádků za n postupně

$$n = 20\,000;\ 2\,000\,000;\ 200\,000\,000,$$

takže

$$\sqrt{2np\,(1-p)} = 100;\ 1000;\ 10\,000,$$

dostaneme tuto tabulku:

Meze, v nichž má býti obsažen počet m zdařených pokusů			Příslušná pravděpodobnost
pro $n = 2.10^4$	pro $n = 2.10^6$	pro $n = 2.10^8$	
$10^4 \pm 48$	$10^6 \pm 476$	$10^8 \pm 4\,769$	0,5
$10^4 \pm 100$	$10^6 \pm 1000$	$10^8 \pm 10\,000$	0,8427 ..
$10^4 \pm 200$	$10^6 \pm 2000$	$10^8 \pm 20\,000$	0,9953 ..
$10^4 \pm 300$	$10^6 \pm 3000$	$10^8 \pm 30\,000$	0,99998 ..

Příklad 2.: Házíme kostkou; pravděpodobnost, že padne jedno oko, je $p = \frac{1}{6}$. Zase dosazujeme postupně $n = 18\,000$; $1\,800\,000$; $180\,000\,000$, takže $\sqrt{2np\,(1-p)} \doteq 71,\ 707,\ 7071$ (poněvadž $\sqrt{50} = 7{,}071, \ldots$) a dostaneme tyto výsledky:

Meze, v nichž má býti obsažen počet m zdařených pokusů			Příslušná pravděpodobnost
pro $n = 18.10^3$	pro $n = 18.10^5$	pro $n = 18.10^7$	
$3.10^3 \pm 34$	$3.10^5 \pm 337$	$3.10^7 \pm 3\,372$	0,5
$3.10^3 \pm 71$	$3.10^5 \pm 107$	$3.10^7 \pm 7\,071$	0,8427
$3.10^3 \pm 141$	$3.10^5 \pm 1\,414$	$3.10^7 \pm 14\,142$	0,9953
$3.10^3 \pm 212$	$3.10^5 \pm 2\,121$	$3.10^7 \pm 21\,213$	0,99998

22. Srovnání theoretických vzorců s výsledky pokusů. Vzorce uvedené v předešlém odstavci dají se kontrolovati, srovnáme-li je s výsledky skutečně provedených pokusů. Rozdělme všechny pokusy v s serií; provedeme nejprve n pokusů prvé serie, mezi kterými bude m_1 zdařených, pak n pokusů druhé serie, mezi kterými bude m_2 zdařených atd. Předpokládáme, že pravděpodobnost p, že se jeden pokus zdaří, je konstantní, že jsou pokusy nezávislé jeden na druhém a že daná čísla n (počet pokusů v jedné serii) a s (počet serií) jsou veliká. Srovnání theorie s pokusy je zajímavé hlavně v těchto věcech:

a) Empirická (t. j. odvozená ze statistiky pokusů) střední hodnota počtu zdařených pokusů v jedné serii je arithmetický střed čísel m_i:

$$\text{s. h}'.\,(m) = \frac{m_1 + m_2 + \ldots + m_s}{s}; \tag{1}$$

s. h′. značí empirickou střední hodnotu. Toto číslo se má přibližně shodovati s theoretickou střední hodnotou s. h. (m), která je rovna np (odst. 14a).

b) Empirická střední hodnota čtverce $(m - np)^2$ úchylky pro jednu serii je

$$\text{s. h}'.\,(m - np)^2 = \frac{(m_1 - np)^2 + (m_2 - np)^2 + \ldots + (m_s - np)^2}{s}. \tag{2}$$

Toto číslo má se shodovati s theoretickou střední hodnotou čtverce úchylky, která je podle rovnice (1) odst. 15. s. h. $(m - np)^2 = np\,(1 - p)$. Střední hodnota čtverce relativní úchylky (viz odst. 15c) je

$$\text{s. h}'.\left(\frac{m}{n} - p\right)^2 = \frac{\left(\frac{m_1}{n} - p\right)^2 + \left(\frac{m_2}{n} - p\right)^2 + \ldots + \left(\frac{m_s}{n} - p\right)^2}{s}$$

theoretická hodnota je s. h. $\left(\frac{m}{n} - p\right)^2 = \frac{p(1-p)}{n}$.

c) Napišme úchylky pro každou z s serií:

$$m_1 - np,\ m_2 - np,\ \ldots,\ m_s - np$$

a spočítejme, kolik z nich jest obsaženo v mezích h_1 až h_2; toto číslo dělené počtem s serií, udává empirickou hodnotu pravděpodobnosti, že úchylka je v oněch mezích a má se shodovati s theoretickou hodnotou té pravděpodobnosti (vzorec (1) odst. 21.).

Jinak vyjádřeno: theoretický počet serií, ve kterých jest odchylka obsažena v mezích h_1 až h_2, je podle citovaného vzorce

$$\frac{s}{2}\left[\Theta\left(\frac{h_2}{\sqrt{2np(1-p)}}\right) - \Theta\left(\frac{h_1}{\sqrt{2np(1-p)}}\right)\right]. \tag{3}$$

Pro $h_1 = -h_2 = h > 0$ bude

$$s \,.\, \Theta\left(\frac{h}{\sqrt{2np(1-p)}}\right)$$

theoretický počet těch serií, u nichž absolutní hodnota úchylky je menší než h.

d) Očekáváme, že čísla $m_1, m_2, \ldots, m_s$ budou se lišiti od np nejvýše asi o $\pm 3\sqrt{2np(1-p)}$, neboť podle čtvrtého vzorce (3) odst. 21 je jen nepatrná pravděpodobnost, že m by se lišilo od np více než o $\pm 3\sqrt{2np(1-p)}$.

Statistika o výsledcích velkého počtu pokusů může sloužiti, podle toho, co bylo uvedeno v tomto odstavci, ke kontrole theoretických vzorců o pravděpodobnostech a středních hodnotách.

23. Zobecnění Laplaceovy věty. Vyjdeme z Laplaceovy věty vyjádřené rovnicí (2a) odst. 21. (pro $t_1 < t_2$):

$$\lim_{n=\infty} P[np + t_1 \sqrt{2np(1-p)} < m < np + t_2 \sqrt{2np(1-p)}] =$$

$$= \int_{t_1}^{t_2} \frac{1}{\sqrt{\pi}} e^{-x^2} dx. \qquad (1)$$

Veličina m (počet zdařených pokusů) dá se podle odst. 14b pojímati jakožto součet n veličin: $x^{(1)} + x^{(2)} + \ldots + x^{(n)}$; $x^{(i)}$ závisí na výsledku i-tého pokusu.

Laplaceova věta nazývá se někdy také věta o limitě pravděpodobnosti, poněvadž se vztahuje k limitě pravděpodobnosti (pro $n = \infty$), že tento součet jest obsažen v mezích závislých určitým způsobem na n. S tohoto stanoviska lze Laplaceovu větu zobecniti jak následuje.

Předpokládejme, že konáme řadu pokusů vzájemně nezávislých, takže výsledek některého z nich nemá vlivu na pravděpodobnosti, se kterými se očekávají výsledky jiných. i-tému pokusu přiřadíme veličinu $x^{(i)}$, která může nabývati různých hodnot, podle toho k jakému zjevu vedl onen pokus. Nechť jsou $E_1^{(i)}$, $E_2^{(i)}$, ... zjevy, které se mohou vyskytnouti jakožto výsledek i-tého pokusu a nechť nabude $x^{(i)}$ hodnoty $\alpha_k^{(i)}$, vede-li i-tý pokus ke zjevu $E_k^{(i)}$. Pravděpodobnost, že se $E_k^{(i)}$ vyskytne, budiž $p_k^{(i)}$; platí, že $p_k^{(i)} = P(x^{(i)} = \alpha_k^{(i)})$, a máme

$$a^{(i)} = \text{s. h. } (x^{(i)}) = \sum_k p_k^{(i)} \alpha_k^{(i)}, \quad \sum_k p_k^{(i)} = 1;$$

součty se vztahují ke všem možným eventualitám i-tého pokusu (součet má tolik členů, kolik má i-tý pokus různých možných výsledků).

Zaveďme do počtu *absolutní moment* $d_i^{(\delta)}$ stupně δ ($\delta > 0$) veličiny $x^{(i)}$:

$$d_i^{(\delta)} = \text{s. h. } |x^{(i)} - a^{(i)}|^\delta = \sum_k p_k^{(i)} |\alpha_k^{(i)} - a^{(i)}|^\delta.$$

Absolutní moment $d_i^{(2)}$ druhého stupně je totožný se střední hodnotou čtverce úchylky.

$$d_i^{(2)} = \text{s. h. } (x^{(i)} - a^{(i)})^2 = \sum_k p_k^{(i)} (\alpha_k^{(i)} - a^{(i)})^2.$$

Úchylka se zde počítá tak, že se každá veličina $x^{(i)}$ odečítá od své střední hodnoty $a^{(i)}$, která sama závisí na i (v jednoduchém případě 15b) byla s. h. všech veličin $x^{(i)}$ stejná, rovná p).

Zobecniti Laplaceovu větu znamená nalézti podmínky, za kterých platí (pro $t_1 < t_2$)

$$\lim_{n=\infty} P\left[t_1 \sqrt{2 \sum_{i=1}^{n} d_i^{(2)}} < \sum_{i=1}^{n}(x^{(i)} - a^{(i)}) < t_2 \sqrt{2 \sum_{i=1}^{n} d_i^{(2)}}\right] = \int_{t_1}^{t_2} \frac{e^{-x^2}}{\sqrt{\pi}}\, dx. \qquad (2)$$

Rovnice (1) je speciálním případem rovnice (2). Připustíme-li totiž, že $x^{(i)}$ může nabýti (jako v odst. 14 a 15) jen hodnot 1 až 0 a to s konstantními pravděpodobnostmi p resp. $(1 - p)$, bude

$$a^{(i)} = p, \sum_{i=1}^{n} x^{(i)} = m, \; d_i^{(2)} = \text{s. h. } (x^{(i)} - p)^2 = p(1 - p),$$

$$\sum_{i=1}^{n}(x^{(i)} - a^{(i)}) = m - np, \sum_{i=1}^{n} d_i^{(2)} = np\,(1 - p) \text{ (podle odst. 15);}$$

rovnice (2) pak přejde v (1).

Laplace se zabýval myšlenkou odvoditi obecný zákon pro pravděpodobnost, že součet velikého počtu n náhodných veličin jest obsažen v určitých mezích (závislých na n). Dokázal větu ve zvláštním případě (1). Obtíž zobecnění je v tom, že třeba voliti pravděpodobnosti $p^{(i)}$ i hodnoty $\alpha_k^{(i)}$ tak, aby bylo vyhověno rovnici (2). Čebyšev učinil v tomto směru další kroky, a ačkoli nejsou jeho výsledky úplné, ukázaly se jeho methody cennými. Čebyševovy důkazy zdokonalil Markov. Později Ljapunov dokázal Laplaceovu větu v obecnějším znění než Markov. Nové výsledky a nové methody v tomto oboru jsou uvedeny v knihách: *A. Khintchine* (Chinčin): Asymptotische Gesetze der Wahrscheinlichkeitsrechnung (Berlin 1933); *S. Bernštejn*: Těorija věrojatnostěj (Moskva 1946, 4. vyd.).

24. Zákon velkých čísel. Markovova věta. a) Konáme neomezenou řadu pokusů; výsledek i-tého pokusu určuje hodnotu veličiny $x^{(i)}$. Možné hodnoty veličiny $x^{(i)}$ nechť jsou $\alpha_1^{(i)}$, $\alpha_2^{(i)}, \ldots$ s příslušnými pravděpodobnostmi $p_1^{(i)}$, $p_2^{(i)}, \ldots$. Budiž $a^{(i)}$ jako v odst. 23., s. h. veličiny $x^{(i)}$, tedy

$$\text{s. h. } (x^{(i)}) = a^{(i)} = \sum_k p_k^{(i)} \alpha_k^{(i)}, \quad i = 1, 2, 3, \ldots .$$

Užívajíce názvosloví obvyklého u ruských matematiků pravíme, že *veličiny* $x^{(1)}, x^{(2)}, \ldots$ *splňují zákon velkých čísel*, je-li

$$\lim_{n=\infty} P\left(\left|\frac{x^{(1)} + x^{(2)} + \ldots + x^{(n)}}{n} - \right.\right.$$
$$\left.\left. - \frac{a^{(1)} + a^{(2)} + \ldots + a^{(n)}}{n}\right| \leqq \varepsilon\right) = 1, \qquad (1)$$

kde ε je libovolně volená kladná konstanta. Smysl rovnice (1) je ten: pravděpodobnost, že aritmetický střed veličin $x^{(1)}$, $\ldots, x^{(n)}$ se liší od aritmetického středu jejich středních hodnot o méně než ε, blíží se jistotě, roste-li n do nekonečna.

Kdybychom volili hodnoty $\alpha_k^{(i)}$, kterých mohou nabývati veličiny $x^{(i)}$ jakož i příslušné pravděpodobnosti $p_k^{(i)}$ docela libovolně, neplatil by zákon (1) obecně. Aby platil, je třeba omeziti nějakým způsobem tyto veličiny. Uvádím zde *Markovovu větu*, která stanoví postačující podmínky k platnosti zákona velikých čísel a která je pozoruhodná tím, že platí také pro veličiny $x^{(i)}$ vzájemně závislé: *Budiž*

$$B_n = \text{s. h. } [x^{(1)} + x^{(2)} + \ldots + x^{(n)} - (a^{(1)} + a^{(2)} +$$
$$+ \ldots + a^{(n)})]^2;$$

veličiny $x^{(i)}$ splňují zákon velkých čísel, je-li

$$\lim_{n=\infty} \frac{B_n}{n^2} = 0. \tag{2}$$

Markovovův důkaz užívá Čebyševovy nerovnosti (viz odst. 12.). Položme pro stručnost

$$y_n = x^{(1)} + x^{(2)} + \dots + x^{(n)} - (a^{(1)} + a^{(2)} \dots + a^{(n)});$$

$$B_n = \text{s. h. } y_n^2.$$

Podle druhé nerovnosti (4) odst. 12. je

$$P(|y_n| < t\sqrt{B_n}) \geqq 1 - \frac{1}{t^2}$$

aneb

$$P\left(\frac{|y_n|}{n} < t\sqrt{\frac{B_n}{n^2}}\right) \geqq 1 - \frac{1}{t^2}.$$

Kladné číslo t může být libovolně veliké, takže $1 - \frac{1}{t^2}$ se liší libovolně málo od jednotky; vzhledem k předpokladu (2) můžeme voliti pak n tak veliké, že $t\sqrt{\frac{B_n}{n^2}} = \varepsilon$, kde ε je libovolně dané kladné číslo. Je tedy

$$P\left(\left|\frac{x^{(1)} + x^{(2)} + \dots + x^{(n)}}{n} - \right.\right.$$

$$\left.\left. - \frac{a^{(1)} + a^{(2)} + \dots + a^{(n)}}{n}\right| < \varepsilon\right) \geqq 1 - \frac{1}{t^2},$$

z čehož plyne platnost limitního vztahu (1), t. j. zákona velkých čísel.

b) Uveďme některé speciální případy, ve kterých je splněna podmínka (1). *Zákon velkých čísel platí, jsou-li veličiny $x^{(i)}$ vzájemně nezávislé a je-li*

$$\text{s. h. } (x^{(i)} - a^{(i)})^2 = \sum_k p_k^{(i)} (\alpha_k^{(i)} - a^{(i)})^2 < C, \qquad (3)$$

kde C je konstanta (Čebyšev). Neboť v tomto případě je

$$B_n = \sum_{i=1}^{n} \text{s. h. } (x^{(i)} - a^{(i)})^2 +$$
$$+ 2 \sum \sum_{1 \leq i < k \leq n} \text{s. h. } [(x^{(i)} - a^{(i)}) (x^{(k)} - a^{(k)})].$$

Poněvadž první součet má n členů, z nichž každý je menší než C a poněvadž

$$\text{s. h. } (x^{(i)} - a^{(i)}) = a^{(i)} - a^{(i)} = 0,$$

a vzhledem k nezávislosti veličin $x^{(i)}$ (odst. 10b)

$$\text{s. h. } [x^{(i)} - a^{(i)}) (x^{(k)} - a^{(k)})] =$$
$$= \text{s. h. } (x^{(i)} - a^{(i)}) \,.\, \text{s. h. } (x^{(k)} - a^{(k)}) = 0,$$

je $B_n < nC$; rovnice (2) platí.

Poissonova věta: Je-li $p^{(i)}$ pravděpodobnost, že v neomezené řadě navzájem nezávislých pokusů i-tý pokus se zdaří a je-li mezi prvními n pokusy m zdařených, je

$$\lim_{n=\infty} P\left(\left|\frac{m}{n} - \frac{1}{n}\sum_{i=1}^{n} p^{(i)}\right| \leq \varepsilon\right) = 1. \qquad (4)$$

Neboť zde jsou pro každý pokus dva možné výsledky, a

$$\alpha_1^{(i)} = 1, \ \alpha_2^{(i)} = 0, \ p_1^{(i)} = p^{(i)}, \ p_2^{(i)} = 1 - p^{(i)},$$
$$\text{s. h. } x^{(i)} = a^{(i)} = p^{(i)}, \ \text{s. h. } (x^{(i)} - a^{(i)})^2 =$$
$$= p^{(i)} (1 - p^{(i)}) < 1,$$

takže je splněna podmínka (3) pro $C = 1$. Ježto $x^{(1)} + x^{(2)} + \ldots + x^{(n)} = m$, přechází rovnice (1) ve (4).

Bernoulliova věta (odst. 16.) je speciální případ Poissonovy věty (4); obdržíme ji ze vzorce (4) předpokládajíce, že každý pokus má konstantní na i nezávislou pravděpodobnost $p^{(i)} = p$.

25. Náhodné rozdělování předmětů do přihrádek. a) N předmětů se rozdělí do ν přihrádek; každý předmět má stejnou pravděpodobnost, že přijde do jedné určité přihrádky jako do jakékoliv jiné (pravděpodobnost dostati se do určité přihrádky je $p = 1 : \nu$). Jak velká je pravděpodobnost, že v dané přihrádce bude právě n předmětů ($n \leq N$)? Za předpokladu nezávislosti (dostane-li se několik předmětů do jedné přihrádky, nemá to vliv na pravděpodobnost, se kterou se tam může dostati další předmět) je počet všech možných případů ν^N, t. j. počtu variací N-té třídy z ν prvků (odst. 3b). Příznivý případ je ten, že do zvolené přihrádky se dostane n předmětů a že zároveň všechny ostatní předměty budou jakkoli rozděleny do ostatních $\nu - 1$ přihrádek. Počet příznivých případů je tedy roven počtu kombinací n-té třídy z N prvků, což je $(N)_n$ podle odst. 3c, násobenému počtem variací $(N - n)$-té třídy z $\nu - 1$ prvků. Hledaná pravděpodobnost P je tedy dána vzorcem

$$P = \frac{N(N-1)\dots(N-n+1)}{1.2.\dots.n} \cdot \frac{(\nu-1)^{N-n}}{\nu^N}. \qquad (1)$$

Vzorec (1) je v podstatě totožný s Newtonovým vzorcem (1) odst. 13. Neboť položíme-li $p = 1/\nu$ a násobíme-li čitatele i jmenovatele v (1) číslem ν^{N-n}, vychází $P = (N)_n p^n (1-p)^{N-n}$, což se shoduje až na označení s vzorcem (1) odst. 13. Srovnejme úlohu o zařaďování předmětů s úlohou o opětovaných pokusech (odst. 13): Dvěma možnostem: zařaditi předmět do určité přihrádky A či nezařaditi, odpovídají dvě možnosti: pokus se buď zdaří nebo nezdaří; N (počet předmětů) odpovídá úhrnnému počtu pokusů; ν (počet přihrádek) jest obráceně úměrný pravděpodobnosti p, že se pokus zdaří ($p = \nu^{-1}$)); n (počet předmětů v přihrádce A) odpovídá počtu zdařených pokusů.

b) Ve vzorci (1) udává poměr N/ν kolik předmětů je průměrně v jedné přihrádce. Předpokládejme že N i ν rostou do nekonečna, ale tak, že je vždy průměrně k předmětů v jedné přihrádce; k je dané číslo. Bude tedy

$$\lim_{\nu=\infty} \frac{N}{\nu} = k,$$

$$\lim_{\nu=\infty} P = \lim_{\nu=\infty} \frac{\frac{N}{\nu} \cdot \left(\frac{N}{\nu} - \frac{1}{\nu}\right) \dots \left(\frac{N}{\nu} - \frac{n-1}{\nu}\right)}{1 \,.\, 2 \,.\, 3 \dots . \, n} \left(1 - \frac{1}{\nu}\right)^{\frac{N}{\nu} \nu - n}$$

nebo

$$\lim_{\nu=\infty} P = \frac{k^n}{n!} e^{-k}. \tag{2}$$

Je-li tedy velmi velký počet předmětů rozdělen do velmi velkého počtu přihrádek a to tak, že na jednu přihrádku připadne průměrně k předmětů, udává pravá strana rovnice (2) pravděpodobnost, že v dané přihrádce je přesně m předmětů. (2) se nazývá *Poissonovou formulí.*

Úhrnná pravděpodobnost, že v dané přihrádce buď není žádný předmět ($n = 0$), nebo jen jeden ($n = 1$), nebo jen dva ($n = 2$) atd. je

$$\left(1 + \frac{k}{1!} + \frac{k^2}{2!} + \frac{k^3}{3!} + \dots\right) . \, e^{-k} = e^k \,.\, e^{-k} = 1,$$

rovná se tedy jistotě.

c) Poissonovu formuli lze vyložiti geometricky takto: Na neomezené přímce q jsou rozsety body tak, že na jednotku délky připadne průměrně k bodů; pravděpodobnost, že jich bude přesně n na zvolené úsečce o délce 1 cm, rovná se pravé straně rovnice (2).

Pravděpodobnost, že na úsečce o délce x, zvolené na přímce q, bude přesně n bodů, je podle (2) rovna

$$\frac{(kx)^n}{n!} e^{-kx},$$

neboť, volíme-li úsečku o délce x za jednotku délky, bude na této nově zvolené jednotce průměrně kx bodů.*)

---

*) Pro čtenáře, kteří se zajímají o počet pravděpodobnosti, uvádím názvy některých učebnic určených pro začátečníky:

Fréchet-Halbwachs: Le Calcul des probabilités à la portée de tous (Paris, 1924).

Borel-Deltheil: Probabilités, erreurs (Paris, 1923).

Coolidge: An Introduction to Mathematical Probability (Oxford, 1925).

Vyšlo též německy:

Coolidge-Urban: Einführung in die Wahrscheinlichkeitsrechnung (Leipzig, 1927).

Castelnuovo: Calcolo di Probabilità, seconde ediz., ve 2 svazcích (Bologna, 1925—28).

Uspensky: Introduction to Mathematical Probability (New York, 1937).

Borel: Éléments de la Théorie des probabilités, 3ième édition (Paris, 1924).

Czuber: Die Wahrscheinlichkeitsrechnung und ihre Anwendungen auf Fehlerausgleichung, Statistik und Lebensversicherung, 3. Aufl. (Leipzig, I. 1914, II. 1921).

O počtu pravděpodobnosti v souvislosti s jeho užitím ve fysice, v biologii a psychologii a s otázkami filosofickými jedná spis:
Borel: Le hasard, Paris 1920.

Konečně upozorňuji na encyklopedické dílo, které dává přehled o otázkách počtu pravděpodobnosti a o jeho aplikacích v různých oborech:

Traité du Calcul des Probabilités et de ses Applications publié par *E. Borel.* Vyšlo ve čtyřech svazcích v Paříži 1925—39.

Jako doplněk uvádí redaktor knihu:

V. J. Glivěnko: Těorija věrojatnostěj (Moskva, 1939).

Spis obsahuje axiomaticky budované základy počtu pravděpodobnosti s obsahem menším, než tato knížka. Český překlad spisu od prof. Dra K. Rychlíka je v tisku.