METODICKÉ VARIÁCIE NA TÉMU PRAVDEPODOBNOSŤ III

BELOSLAV RIEČAN, Bratislava

Pojem náhodnej premennej (= náhodnej veličiny) je u nás v školskej praxi menej populárny ako pojem náhodného javu (= udalosti), hoci je rovnako názorný. Veď pri hode kockou nás zaujíma počet bodiek na padnutej stene, v lotérii výška výhry, vo výrobe počet bezchybných výrobkov a pod. A funkcia, ktorá každému výsledku náhodného pokusu priraďuje reálne číslo — to je náhodná premenná.

Pritom však náhodná premenná akoby patrila do zakázanej trinástej komnaty. Pozrime sa napr. na pravdepodobnostnú časť dosiaľ platnej učebnice pre gymnáziá (napísanej v čase, keď sa tieto školy ešte nazývali SVŠ). Nepochádza od nikoho menšieho ako od Jaroslava Hájka. Obsahuje síce aj rozsiahlu časť venovanú náhodným veličinám, ale tá je vyznačená zreteľne na okraji s tým, že ju možno vynechať. A netreba pochybovať o jej vynechávaní pri vyučovaní už aj tak dosť nabitom látkou*. Podobne to dopadlo s náhodnými veličinami aj v nedávno vyjdenej učebnici pre gymnáziá s rozšíreným vyučovaním matematiky. Autor učebnice V. Dupač koncipoval príslušnú partiu veľmi umiernene a vyznačil, čo možno vynechať celkom. (Odporúčali by sme čitateľovi pre potešenie toto nápadité a poučné Dupačovo dielko, keby sme mu vedeli poradiť, ako by ho mohol zohnať.) Ani pripravovaná experimentálna učebnica pre gymnáziá nebude obsahovať časť venovanú náhodnej premennej.

Zavedenie pojmu náhodnej premennej by malo veľký význam najmä pre aplikácie v matematickej štatistike. Na druhej strane je tu významný dôvod proti: jej zavedenie si vyžaduje vybudovanie osobitného a osobitého matematického aparátu. Na to niet času a najmä tradície. A lepšie vynechať ako prebrať povrchne.

---

* Možno, že by sa situácia trochu zlepšila, keby spomínaná časť bola označená ako nevhodná pre mládež.

V nasledujúcich riadkoch sa pokúsime o elementárny výklad základov pravdepodobnosti, založený od samého začiatku na pojme náhodnej premennej. Nepôjdeme hlbšie do aplikácií, ale uspokojíme sa s Bernoulliho schémou, opisujúcou nezávislé opakovanie pokusov. Však ten, kto našiel v pravdepodobnosti záľubu vie, aká je to partia dôležitá pre aplikácie.

Náhodná premenná a jej stredná hodnota

Náš svet bude ohraničený konečným pravdepodobnostným priestorom, t. j. množinou $\Omega = \{\omega_1, \ldots, \omega_n\}$ spolu s usporiadanou n-ticou $(p_1, \ldots, p_n)$ nezáporných čísel, takou, že $\sum_{i=1}^{n} p_i = 1$. Náhodnou premennou rozumieme ľubovoľnú funkciu $\xi\colon \Omega \to R$. Jej strednou hodnotou je číslo

$$E(\xi) = \sum_{i=1}^{n} \xi(\omega_i) p_i$$

Pravdepodobnosť množiny $A \subset \Omega$ definujeme pomocou rovnosti

$$P(A) = E(\chi_A)$$

Príklad 1. Nech $p_1 = p_2 = \ldots = p_n$. Potom $np_i = 1$, teda $p_i = \frac{1}{n}$ pre všetky i. Pre ľubovoľnú náhodnú premennú ξ platí

$$E(\xi) = \sum_{i=1}^{n} \xi(\omega_i) \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{n} \xi(\omega_i)$$

teda stredná hodnota je aritmetickým priemerom hodnôt funkcie ξ.

Čo je pravdepodobnosť množiny A? Je to číslo

$$P(A) = \frac{1}{n} \sum_{i=1}^{n} \chi_A(\omega_i)$$

V súčte sú len nuly a jednotky, pričom jednotiek je práve toľko, koľko je tých ω_i, pre ktoré $\omega_i \in A$. Preto

$$P(A) = \frac{m}{n}$$

kde m je počet prvkov množiny A.

Príklad 2. Ako sa vypočíta $P(A)$ vo všeobecnom prípade? Platí

$$P(A) = \sum_{i=1}^{n} \chi_A(\omega_i) p_i$$

Do súčtu vstúpia niektoré p_i. Totiž práve tie, pre ktoré $\omega_i \in A$. Ak $\omega_i \notin A$, tak $\chi_A(\omega_i) = 0$. Preto

$$P(A) = \sum_{\omega_i \in A} p_i$$

Príklad 3. Hrajme túto hru: Hádžem kockou. Ak mi padne šestka, zaplatí mi súper 5 korún, ak padne 4 alebo 5, súper mi dá korunu, v ostatných prípadoch naopak ja platím 2 koruny. Je táto hra pre mňa výhodná?

Máme $\Omega = \{1, 2, \ldots, 6\}$, $p_1 = p_2 = \ldots = p_6 = \frac{1}{6}$, $\xi(6) = 5$, $\xi(4) = \xi(5) = 1$, $\xi(1) = \xi(2) = \xi(3) = -2$. Číslo $\xi(\omega)$ je môj zisk (strata) pri výsledku ω. Preto

$$E(\xi) = \frac{1}{6}(5 + 1 + 1 - 2 - 2 - 2) = \frac{1}{6}$$

Dosiahnutý výsledok môžeme interpretovať takto: Môj zisk je $\frac{1}{6}$ koruny v priemere na jednu partiu.

Všimnime si, že nie je dôležité koľko má priestor Ω prvkov, ale aké sú pravdepodobnosti jednotlivých hodnôt. V našom prípade máme 3 hodnoty: $\alpha_1 = 5$, $\alpha_2 = 1$, $\alpha_3 = -2$. Ďalej

$$\{\omega; \xi(\omega) = 5\} = \{6\}, \quad \{\omega; \xi(\omega) = 1\} = \{4,5\},$$
$$\{\omega; \xi(\omega) = -2\} = \{1, 2, 3\}$$

teda (zapísané skrátene)

$$P(\xi = 5) = \frac{1}{6}, \quad P(\xi = 1) = \frac{2}{6} = \frac{1}{3}, \quad P(\xi = -2) = \frac{3}{6} = \frac{1}{2}$$

$$E(\xi) = 5 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + (-2) \cdot \frac{1}{6} + (-2) \cdot \frac{1}{6} + (-2) \cdot \frac{1}{6} =$$

$$= 5 \cdot \frac{1}{6} + 1 \cdot \left(\frac{1}{6} + \frac{1}{6}\right) + (-2)\left(\frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right) =$$
$$= 5 \cdot P(\xi = 5) + 1 \cdot P(\xi = 1) + (-2)P(\xi = -2)$$

Vieme, ako túto vlastnosť sformulovať vo všeobecnom prípade, preto sa nebudeme formuláciou zbytočne zaoberať.

Nezávislé opakovanie pokusov

Ide o schému riešiacu úlohy takého typu: Aká je pravdepodobnosť toho, že pri stonásobnom bode kockou šestka padne aspoň dvadsaťkrát? Aká je pravdepodobnosť toho, že medzi 100 výrobkami bude aspoň 80 bezchybných, ak výrobca pracuje so spoľahlivosťou 95 %?

Na tieto otázky sa možno pozerať z rôznych hľadísk. My budeme skúmať problém z hľadiska náhodnej premennej — počtu ξ úspešných pokusov. Pri pevnom počte n všetkých pokusov, ξ môže nadobudnúť hodnoty 0, 1, ..., n. Zdá sa, že túto skutočnosť by mohla opísať funkcia

$$\xi = \sum_{i=1}^{n} \chi_{A_i}$$

kde pod A_i si predstavujeme úspech pri n-tom opakovaní pokusu. Veličina $\xi(\omega)$ nadobudne totiž hodnotu k práve vtedy, keď práve k sčítancov $\chi_{A_i}^{(\omega)}$ má hodnotu 1, a to je práve vtedy, keď ω patrí do práve k množín spomedzi $A_1, \ldots, A_n$. (Ak chceme byť celkom akurátni, môžeme za priestor Ω vziať množinu všetkých usporiadaných n-tíc $\omega = (\omega_1, \ldots, \omega_n)$ a položiť

$$A_i = \{(\omega_1, \ldots, \omega_n); \omega_i \in A\}$$

kde A je udalosť (padnutie šestky, vybratie bezchybného výrobku a pod.), ktorej výskyt sledujeme.

Zostáva nám opísať dve skutočnosti: 1. opakovanie, 2. nezávislosť. Opakovanie spočíva v tom, že pravdepodobnosti všetkých udalostí A_i sú rovnaké:

$$P(A_i) = p, \qquad i = 1, \ldots, n \tag{1}$$

O niečo komplikovanejšie je to s opísaním nezávislosti opakovania. Skúsime to najprv na príklade. Hádžme napr. dvakrát kockou, A spočíva v padnutí šestky. Udalosť $A_1 \cap A_2$ spočíva v tom, že šestka padne prvý aj druhý raz. Preto

$$P(A_1 \cap A_2) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(A_1)\, P(A_2)$$

(Výsledok druhého bodu nezávisí od prvého.) Podobne, ak označíme A_1' (resp. A_2') komplement množiny A_1 (resp. A_2), teda udalosť spočívajúcu v tom, že A_1 (resp. A_2) nenastane, t. j. šestka pri prvom (resp. druhom) hode nepadne, dostaneme

$$P(A_1 \cap A_2') = \frac{5}{36} = P(A_1)\, P(A_2'),$$

$$P(A_1' \cap A_2) = \frac{5}{36} = P(A_1')\, P(A_2),$$

$$P(A_1' \cap A_2') = \frac{25}{36} = P(A_1')\, P(A_2')$$

Nezávislosť opakovania opíšeme pomocou uvedených rovností, pravda, pri n súčiniteľoch, teda napr. (pri $n = 3$), platí aj rovnosť

$$P(A_1 \cap A_2' \cap A_3) = P(A_1)\, P(A_2')\, P(A_3)$$

aj rovnosť

$$P(A_1 \cap A_2) = P(A_1)\, P(A_2)$$

Pre ľubovoľné n to vyzerá o niečo komplikovanejšie:

$$P\left(\bigcap_{i \in \mu} A_i \cap \bigcap_{j \in \nu} A_j'\right) = \prod_{i \in \mu} P(A_i) \prod_{j \in \nu} P(A_j')$$

$$\nu \cap \mu = \emptyset, \quad \nu, \mu \subset \{1, 2, \ldots, n\}$$

Príklad 4. Aká je pravdepodobnosť toho, že z piatich nezávislých opakovaní daného pokusu nastane sledovaná udalosť (ktorej pravdepodobnosť je p) práve dvakrát?

Ak zostaneme pri predchádzajúcich označeniach, v ktorých

$$\xi = \sum_{i=1}^{5} \chi_{A_i}$$

našou úlohou je vypočítať pravdepodobnosť $P(\xi = 2)$. Dva úspechy môžu nastať prvé dva razy, posledné dva razy, prvý a posledný raz a pod.; celkove je tu $\binom{5}{2} = \frac{5 \cdot 4}{2 \cdot 1} = 10$ možností. Nebudeme ich všetky vypisovať:

$$P(\xi = 2) = P(A_1 \cap A_2 \cap A_3' \cap A_4' \cap A_5') +$$
$$+ P(A_1 \cap A_2' \cap A_3 \cap A_4' \cap A_5') + \ldots + P(A_1' \cap A_2' \cap A_3' \cap A_4 \cap A_5) =$$
$$= P(A_1)P(A_2)P(A_3')P(A_4')P(A_5') + P(A_1)P(A_2')P(A_3)P(A_4')P(A_5') + \ldots$$
$$\ldots + P(A_1')P(A_2')P(A_3')P(A_4)P(A_5) =$$
$$= p^2(1-p)^3 + p^2(1-p)^3 + \ldots + p^2(1-p)^3 = \binom{5}{2} p^2(1-p)^3$$

Predchádzajúci príklad nám naznačuje všeobecný vzťah v Bernouilliho schéme

$$P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Tento výsledok možno sformulovať a dokázať úplne podrobne:

Veta 1. Nech $A_1, \ldots, A_n$ sú podmnožiny konečného pravdepodobnostného priestoru vyhovujúce podmienkam (1) a (2). Nech ξ je náhodná premenná definovaná rovnosťou

$$\xi = \sum_{i=1}^{n} \chi_{A_i}$$

Potom pre ľubovoľné celé k, $0 \leqq k \leqq n$ platí

$$P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Dôkaz. Fixujme k a označme znakom C množinu všetkých k-prvkových podmnožín množiny $\{1, \ldots, n\}$. Potom

$$\{\omega ; \xi(\omega) = k\} = \bigcup_{\mu \in C} \left(\bigcap_{i \in \mu} A_i \cap \bigcap_{j \notin \mu} A_j' \right)$$

teda

$$P(\xi = k) = \sum_{\mu \in C} P\left(\bigcap_{i \in \mu} A_i \cap \bigcap_{j \notin \mu} A_j' \right) = \sum_{\mu \in C} \prod_{i \in \mu} P(A_i) \prod_{j \notin \mu} P(A_j') =$$
$$= \sum_{\mu \in C} p^k (1-p)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}$$

Príklad 5. Aká je pravdepodobnosť toho, že pri stonásobnom hode kockou šestka padne aspoň dvadsaťkrát?

V tomto prípade $p=\frac{1}{6}$, $n=100$. Hľadaná pravdepodobnosť bude súčtom

$$\sum_{k=20}^{100} P(\xi=k)=\sum_{k=20}^{100}\binom{100}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{100-k}\doteq 0,18$$

Príklad 6. Pravdepodobnosť toho, že jednotlivý výrobok má prvotriednu kvalitu je 0,95. Aká je pravdepodobnosť toho, že spomedzi 100 náhodne vybratých výrobkov bude aspoň 80 bezchybných?

V tomto príklade $p=0,95$, $n=100$, teda hľadaná pravdepodobnosť je

$$\sum_{k=80}^{100}\binom{100}{k}0,95^k 0,5^{100-k}\doteq 0,98.$$

Stredná hodnota binomického rozdelenia

Hovoríme, že náhodná premenná ξ má binomické rozdelenie s parametrami $n\in N$ a $p\in(0, 1)$, ak

$$P(\xi=k)=\binom{n}{k}p^k(1-p)^{n-k}, \qquad k=0, 1, \ldots, n$$

Videli sme, že binomické rozdelenie má náhodná premenná — počet úspechov pri n-násobnom nezávislom opakovaní nejakého pokusu.

Nazdávame sa, že výsledky príkladov 5 a 6 (zaoberajúcich sa binomickým rozdelením) nie sú uspokojivé, nedávajú nám dostatočnú číselnú predstavu. Číselne ich vyjadrovať by bolo veľmi prácne. Možno ich, pravda, odhadnúť, ale aj na tento odhad potrebujeme poznať strednú hodnotu*, nehovoriac už o tom, že stredná hodnota dáva sama osebe veľmi cennú informáciu o skúmanej náhodnej premennej — v našom prípade počte pokusov v sérii nezávislých opakovaní daného pokusu.

---

* Spomínanými odhadmi pravdepodobnosti z príkladov 5 a 6 sa tu nebudeme zaoberať. Táto úloha je riešená napr. v knihe L. Kosmák, Úvod do kombinatorickej teórie pravdepodobnosti.

Ostaneme pri doterajšom prístupe k problematike, v ktorom sme pracovali so sumou

$$\xi = \sum_{i=1}^{n} \chi_{A_i}$$

kde A_i predstavoval úspech pri i-tom nezávislom opakovaní, teda platili rovnosti (1) a (2). Kedykoľvek budeme hovoriť o náhodnej premennej s binomickým rozdelením (s parametrami n, p), budeme mať pred očami práve uvedené vyjadrenie. Ľahko vidieť, že kľúčom na riešenie danej úlohy (výpočet strednej hodnoty binomického rozdelenia) je nasledujúce tvrdenie.

Veta 2. Ak ξ, η sú dve náhodné premenné definované na tom istom konečnom pravdepodobnostnom priestore, tak

$$E(\xi + \eta) = E(\xi) + E(\eta)$$

Dôkaz. Podľa definície

$$E(\xi + \eta) = \sum_{i=1}^{n} (\xi + \eta)(\omega_i) p_i =$$

$$= \sum_{i=1}^{n} \xi(\omega_i) p_i + \sum_{i=1}^{n} \eta(\omega_i) p_i = E(\xi) + E(\eta)$$

Veta 3. Nech ξ je náhodná premenná s binomickým rozdelením s parametrami n, p. Potom

$$E(\xi) = np$$

Dôkaz. Videli sme, že stredná hodnota súčtu je súčet stredných hodnôt. Náhodná premenná χ_{A_i} nadobúda hodnoty 1, resp. 0 s pravdepodobnosťou p, resp. $1 - p$. Preto

$$E(\chi_{A_i}) = 1 \cdot p + 0 \cdot (1 - p) = p$$

Podľa vety 2 teda

$$E(\xi) = E\left(\sum_{i=1}^{n} \chi_{A_i}\right) = \sum_{i=1}^{n} E(\chi_{A_i}) = \sum_{i=1}^{n} p = np$$

V príklade 5 bolo $n = 100$, $p = \frac{1}{6}$, teda $E(\xi) = np = 100/6 = 16{,}\bar{6}$, teda počty úspechov (šestiek) pri 100 bodoch sa budú „grupovať" okolo sedemnástky. V príklade 6 bolo $n=100$, $p=0{,}95$, teda $E(\xi)=np=95$. Ale ako ďaleko sa budú od tejto strednej hodnoty odchyľovať, rozptyľovať skutočné hodnoty? To je otázka, na ktorú sa práve chystáme dať odpoveď.

Disperzia

Za mieru rozptylu hodnôt náhodnej premennej ξ (okolo jej strednej hodnoty) sa zvyčajne berie číslo

$$D(\xi)=\sum_{i=1}^{n}(\xi(\omega_i)-E(\xi))^2 p_i$$

prípadne druhá odmocnina z neho. (Toto číslo sa volá disperzia, alebo rozptyl, alebo variancia, odmocnina z neho má názov stredná kvadratická odchýlka; ale to všetko nie je podstatné.) Disperzia je tým väčšia, čím sú väčšie vzdialenosti $|\xi(\omega_i)-E(\xi)|$ čísel $\xi(\omega_i)$, $E(\xi)$, pravdaže, pravdepodobnejšie hodnoty sú preferované. Presne o týchto veciach hovorí Čebyševova nerovnosť:

Veta 4. Pre ľubovoľné $\varepsilon>0$ platí

$$P(\{\omega\,;\,|\xi(\omega)-E(\xi)|\geqq\varepsilon\})\leqq\frac{1}{\varepsilon^2}D(\xi)$$

Dôkaz. Položme $B=\{\omega\,;\,|\xi(\omega)-E(\xi)|\geqq\varepsilon\}$, $\alpha=\{i\,;\,|\xi(\omega_i)-E(\xi)|\geqq\varepsilon\}$. Potom

$$P(B)=\sum_{i\in\alpha}p_i=\frac{1}{\varepsilon^2}\sum_{i\in\alpha}\varepsilon^2 p_i\leqq$$

$$\leqq\frac{1}{\varepsilon^2}\sum_{i\in\alpha}(\xi(\omega_i)-E(\xi))^2 p_i\leqq$$

$$\leqq\frac{1}{\varepsilon^2}\sum_{i=1}^{n}(\xi(\omega_i)-E(\xi))^2 p_i=\frac{1}{\varepsilon^2}D(\xi)$$

Položme špeciálne $\varepsilon=3\sqrt{D(\xi)}$. Potom podľa vety 4

$$P(|\xi-E(\xi)|\geqq\varepsilon)\leqq\frac{1}{9D(\xi)}D(\xi)=\frac{1}{9}$$

teda je málo pravdepodobné, že by sme vypadli z intervalu

$$(E(\xi)-\varepsilon,\ E(\xi)+\varepsilon)=(E(\xi)-3\sqrt{D(\xi)},\ E(\xi)+3\sqrt{D(\xi)})$$

Tento interval je tým užší (a teda hodnoty funkce ξ sa viac primykajú k strednej hodnote), čím je menšie $D(\xi)$.

Na výpočet disperzie stačí (podobne ako v prípade strednej hodnoty) poznať hodnoty náhodnej premennej ξ a ich pravdepodobnosti.

Veta 5. Nech $\{\alpha_1, \dots, \alpha_k\}$ je množina hodnôt náhodnej premennej ξ, $B_i=\{\omega\,;\ \xi(\omega)=\alpha_i\}$. Potom

$$D(\xi)=\sum_{i=1}^{k}(\alpha_i-E(\xi))^2P(B_i)$$

Dôkaz. Stačí vyňať pred zátvorku:

$$D(\xi)=\sum_{i=1}^{n}(\xi(\omega_i)-E(\xi))^2p_i=\sum_{j=1}^{k}(\alpha_j-E(\xi))^2\sum_{\xi(\omega_i)=\alpha_j}p_i=$$

$$=\sum_{j=1}^{k}(\alpha_j-E(\xi))^2\sum_{\omega_i\in B_j}p_i=\sum_{j=1}^{k}(\alpha_j-E(\xi))^2P(B_j)$$

Príklad 7. Vypočítajme disperziu náhodnej premennej ξ z príkladu 3. Máme $\alpha_1=5$, $\alpha_2=1$, $\alpha_3=-2$, $P(B_1)=1/6$, $P(B_2)=1/3$, $P(B_3)=1/2$. Vypočítali sme už $E(\xi)=1/6$. Preto

$$D(\xi)=\left(5-\frac{1}{6}\right)^2\frac{1}{6}+\left(1-\frac{1}{6}\right)^2\frac{1}{3}+\left(-2-\frac{1}{6}\right)^2\frac{1}{2}=6{,}47$$

Za vetou 4 sme hovorili o čísle $3\sqrt{D(\xi)}=7.62$.

Príklad 8. Nech $\xi=\chi_A$, $P(A)=p$. Podľa definície $E(\xi)=p$. Ďalej $\alpha_1=1$, $\alpha_2=0$, $P(\xi=1)=p$, $P(\xi=0)=1-p$. Preto

$$D(\xi)=(1-p)^2p+(0-p)^2(1-p)=(1-p)p$$

Príklad 9. Nech ξ je konštantná funkcia, $\xi=c$. Potom

$$E(\xi)=\sum_{i=1}^{n}\xi(\omega_i)p_i=\sum_{i=1}^{n}cp_i=c\sum_{i=1}^{n}p_i=c$$

$$D(\xi)=\sum_{i=1}^{n}(\xi(\omega_i)-c)^2p_i=\sum_{i=1}^{n}(c-c)^2p_i=0$$

Disperzia binomického rozdelenia

Ľahko nám je teraz vyskakovať, keď máme k dispozícii také kvalitné nástroje. Náhodnú premennú ξ vyjadríme v tvare

$$\xi = \sum_{i=1}^{n} \chi_{A_i}$$

kde $A_1, \ldots, A_n$ sú množiny vyhovujúce podmienkam (1) a (2). Už vieme, že

$$D(\chi_{A_i}) = p(1-p)$$

(pozri príklad 8). Keby sme vedeli, že disperzia súčtu je súčtom disperzií, všetko by bolo hotové, lebo potom by

$$D(\xi) = \sum_{i=1}^{n} D(\chi_{A_i}) = np(1-p)$$

Vo všeobecnosti sa takéto fígle nedajú robiť. V našom prípade nám to umožňuje „nezávislosť" množín $A_1, \ldots, A_n$, t. j. rovnosť (2). Ani tú celkom nevyužijeme. Vystačíme s rovnosťou

$$P(A_i \cap A_j) = P(A_i)P(A_j), \qquad i \neq j \tag{3}$$

Podrobnosti technického rázu sú obsiahnuté v nasledujúcich štyroch lemách.

Lema 1. Nech $A_1, \ldots, A_n$ sú udalosti vyhovujúce podmienke (3). Potom

$$E(\chi_{A_i}\chi_{A_j}) = E(\chi_{A_i})E(\chi_{A_j}), \qquad i \neq j$$

Dôkaz. Treba len rozšifrovať komplikované zápisy. Predovšetkým $E(\chi_{A_i}) = P(A_i)$. Potom $\chi_{A_i}\chi_{A_j} = \chi_{A_i \cap A_j}$. Preto

$$E(\chi_{A_i}\chi_{A_j}) = E(\chi_{A_i \cap A_j}) = P(A_i \cap A_j) =$$
$$= P(A_i)P(A_j) = E(\chi_{A_i})E(\chi_{A_j})$$

Lema 2. Pre každú náhodnú premennú ξ a každé reálne číslo α platí $E(\alpha\xi) = \alpha E(\xi)$.

Dôkaz.

$$E(\alpha\xi) = \sum_{i=1}^{n} \alpha\xi(\omega_i)p_i = \alpha \sum_{i=1}^{n} \xi(\omega_i)p_i = \alpha E(\xi)$$

Lema 3. Ak ξ, η sú také ľubovoľné náhodné premenné, že $E(\xi\eta) = E(\xi)E(\eta)$, tak $E((\xi - E(\xi))(\eta - E(\eta))) = 0$.

Dôkaz. Dvojčlen dvojčlenom násobiť vieme:

$$(\xi - E(\xi))(\eta - E(\eta)) = \xi\eta - E(\xi)\eta - \xi E(\eta) + E(\xi)E(\eta)$$

Potom použijeme vetu 2, lemu 2 a naostatok príklad 9:

$$\begin{aligned} &E((\xi - E(\xi))(\eta - E(\eta)) = \\ &= E(\xi\eta) - E(E(\xi)\eta) - E(E(\eta)\xi) + E(E(\xi)E(\eta)) = \\ &= E(\xi\eta) - E(\xi)E(\eta) - E(\eta)E(\xi) + E(\xi)E(\eta) = 0 \end{aligned}$$

Lema 4. Pre každú náhodnú premennú ξ platí

$$D(\xi) = \sum_{i=1}^{n} E((\xi - E(\xi))^2)$$

Dôkaz.

$$\begin{aligned} D(\xi) &= \sum_{i=1}^{n} (\xi(\omega_i) - E(\xi))^2 p_i = \sum_{i=1}^{n} (\xi - E(\xi))^2(\omega_i) p_i = \\ &= E((\xi - E(\xi))^2) \end{aligned}$$

Lema 5. Ak $E(\xi_i\xi_j) = E(\xi_i)E(\xi_j)$ pre všeky $i \neq j$, tak

$$D\left(\sum_{i=1}^{n} \xi_i\right) = \sum_{i=1}^{n} D(\xi_i)$$

Dôkaz. Platí

$$\begin{aligned} D\left(\sum_{i=1}^{n} \xi_i\right) &= E\left(\left(\sum_{i=1}^{n} \xi_i - E\left(\sum_{i=1}^{n} \xi_i\right)\right)^2\right) = \\ &= E\left(\left(\sum_{i=1}^{n} (\xi_i - E(\xi_i))\right)^2\right) = \\ &= E\left(\sum_{i=1}^{n} (\xi_i - E(\xi_i))^2\right) + 2\sum_{i<j} E((\xi_i - E(\xi_i))(\xi_j - E(\xi_j))) = \\ &= \sum_{i=1}^{n} E(\xi_i - E(\xi_i))^2 = \sum_{i=1}^{n} D(\xi_i) \end{aligned}$$

Veta 6. Nech ξ je náhodná premenná s binomickým rozdelením s parametrami n, p. Potom

$$D(\xi)=np(1-p)$$

Dôkaz. Podľa lemy 1 platí

$$E(\chi_{A_i}\chi_{A_j})=E(\chi_{A_i})E(\chi_{A_j}), \qquad i\neq j$$

Preto podľa lemy 5 platí

$$D(\xi)=D\left(\sum_{i=1}^{n}\chi_{A_i}\right)=\sum_{i=1}^{n}D(\chi_{A_i})=$$

$$=\sum_{i=1}^{n}p(1-p)=np(1-p)$$

Teraz sa môžeme vrátiť k príkladu 5 a 6. V prvom z nich platí

$$D(\xi)=100\cdot\frac{1}{6}\cdot\frac{5}{6}=\frac{500}{36}, \qquad \sqrt{D(\xi)}=3{,}727$$

Keďže $E(\xi)=np = 100/6 = 16{,}\bar{6}$ a $E(\xi) - 3\sqrt{D(\xi)} = 5{,}48$, $E(\xi) + 3\sqrt{D(\xi)} = 27{,}84$, môžeme očakávať, že počet padnutých šestiek bude v rozmedzí od 5 do 28. Väčší počet úspechov je zriedkavý.

V príklade 6 máme

$$D(\xi)=100\cdot 0{,}95\cdot 0{,}05=4{,}75, \qquad \sqrt{D(\xi)}=2{,}179$$

$$E(\xi)-3\sqrt{D(\xi)}=88{,}46, \qquad E(\xi)+3\sqrt{D(\xi)}=101{,}54$$

teda počet bezchybných výrobkov by sa mal nachádzať medzi 88 a 100. (Viac ako 100 ich v nijakom prípade nemôže byť.)

Zákon veľkých čísel

Možno má už čitateľ (a v každom prípade potenciálny žiak) toho dosť, ale v terajšej situácii, keď už máme prebraté, čo sme prebrali, bolo by hriechom neprebrať aj zákon veľkých čísel. To by bolo podobné, ako byť na Krivánskom sedle a nevystúpiť na Kriváň.

O čo ide v zákone veľkých čísel? Urobíme n nezávislých pokusov a zistíme počet ξ_n tých pokusov, v ktorých nastal sledovaný jav (ktorého pravdepodobnosť je p). Podiel ξ_n/n je relatívna početnosť, ktorá sa podľa zákona veľkých čísel pri „dosť veľkom n príliš nelíši od pravdepodobnosti p“. Lenže ξ_n je vlastne počet úspechov pri n nezávislých pokusoch, teda náhodná premenná s binomickým rozdelením pravdepodobnosti s parametrami n, p. A o tejto premennej vieme všetko. A keď aj nie všetko, tak dosť — určite toľko, koľko potrebujeme.

Veta 7. Nech ξ_n je náhodná premenná s binomickým rozdelením s parametrami n, p (t. j. počet úspechov pri n nezávislých pokusoch). Potom pre každé $\varepsilon > 0$ platí

$$P\left(\left\{\omega; \left|\frac{\xi_n(\omega)}{n} - p\right| \geqq \varepsilon\right\}\right) \leqq \frac{1}{4n\varepsilon^2}$$

Dôkaz. Podľa viet 4, 3 a 6 platí

$$P\left(\left|\frac{\xi_n}{n} - p\right| \geqq \varepsilon\right) = P(|\xi_n - np| \geqq n\varepsilon) =$$

$$= P(|\xi_n - E(\xi_n)| \geqq n\varepsilon) \leqq \frac{1}{n^2\varepsilon^2} D(\xi_n) = \frac{np(1-p)}{n^2\varepsilon^2} \leqq \frac{1}{4n\varepsilon^2}$$

Veta 8. (zákon veľkých čísel). Nech $\frac{\xi_n}{n}$ $(n = 1, 2, \ldots)$ je relatívna početnosť udalosti A (ktorej pravdepodobnosť je p) v postupnosti n nezávislých pokusov (t. j. ξ_n je náhodná premenná s binomickým rozdelením s parametrami n a p). Potom pre každé $\varepsilon > 0$ platí

$$\lim_{n\to\infty} P\left(\left\{\omega; \left|\frac{\xi_n(\omega)}{n} - p\right| \geqq \varepsilon\right\}\right) = 0.$$