4. ЗАКОН ВЕЛИКИХ БРОЈЕВА

Аксиоматска дефиниција вероватноће не одређује начин на који ће вероватноће случајних догађаја бити одређене у неком реалном експерименту. Зато треба наћи одговор на питање да ли се и под којим условима, на основу резултата извођења експеримента, може одредити одговарајућа вероватноћа посматраног случајног догађаја. Одговор на ово питање, важно у теорији вероватноће и статистици, дају тзв. закони великих бројева, о којима ће бити речи у поглављу које следи.

Најпре ће бити детаљније описани Бернулијева шема независних понављања експеримената и Гаусова крива, без којих развој теорије вероватноће и статистике не би био могућ.

У трећем делу поглавља ће бити изложен Бернулијев закон великих бројева и дата формулација Борел-Кантелијевог закона.

Локална и интегрална теорема Моавр-Лапласа, без доказа, су дате у четвртом делу поглавља. На основу ових теорема могуће је приближно одредити биномне вероватноће у случају када су вредности броја n велике, а Пуасоново решење за ретке догађаје је дато на крају поглавља.

4.1. БЕРНУЛИЈЕВА ШЕМА НЕЗАВИСНИХ ПОНАВЉАЊА ЕКСПЕРИМЕНАТА

Изузетно важну улогу у теорији вероватноће, њеној примени и развоју статистике, одиграла је једна класа експеримената познатих под именом ***Бернулијеви* (Bernoulli) *експерименти*** или класа независних понављања експеримента. Одређивање вероватноћа случајних догађаја везаних за Бернулијеве експерименте представља тзв. **Бернулијев проблем**.

Претпоставимо да је један експеримент могуће поновити неограничено пута. При томе се мисли да је могуће остварити неограничено пута један исти комплекс услова. Сваки пут кад се оствари тај комплекс услова, посматрани случајни догађај А се може остварити или не. Резултат експеримента, при томе, у једном понављању не утиче на резултате експеримента у осталим понављањима. Рећи ћемо да су понављања експеримента **међусобно независна**. Поред тога, предпоставља се да је у сваком понављању експеримента **вероватноћа остварења догађаја А иста**.

Наведимо неколико примера таквих експеримената.

Пример 1. Правилни новчић се баца неограничено пута и бележи се број појављивања „писма”. Може се посматрати да ли постоји гранична вредност за релативну фреквенцију појављивања „писма” кад неограничено расте број бацања новчића (понављања експеримента) и, уколико постоји, колико износи.

Пример 2. Машина аутоматски обрађује производ са вероватноћом погрешне обраде једнаком p. Ако се у великој серији од 10000 или 100000 комада појави велики број неправилно обрађених комада, да ли то значи да се поузданост машине променила?

Пример 3. Зна се да је међу свим порођајима проценат близанаца једнак p. Ако се у једној клиници у коју долазе породиље изабране на случајан начин нађе 1000 породиља, колики се може очекивати број близанаца?

Пример 4. У статистичком скупу од N елемената, $100\,p\%$ има одређено својство, док преосталих $100(1-p)\%$ нема то својство. На случајан начин се из посматраног скупа бира један елемент и бележи се да ли он има посматрано својство или не. При томе се претпоставља да се извлачење врши са враћањем, тј. сваки пут се извучени елемент враћа у посматрани скуп, па се поново из целог скупа бира следећи елемент. Може нас занимати да одредимо вероватноћу да ће међу n извучених елемената бити тачно k са посматраним својством. Може нас, на пример, интересовати и вероватноћа догађаја да се количник k/n (релативна фреквенција појављивања елемента са посматраним својством) разликује од p за не више од једног унапред датог броја. Специјално нас може интересовати чему ће тежити количник k/n када број извлачења n тежи бесконачности.

У свим наведеним примерима описани су експерименти у којима се посматра случајни догађај А, који се у једном извођењу експеримента остварује са вероватноћом p и не остварује са вероватноћом $q=1-p$. Низ независних понављања таквог експеримента представља тзв. Бернулијев експеримент.

Означимо са n број понављања посматраног експеримента. Најједноставнији и најинтересантнији проблем је одређивање вероватноће да ће се у n понављања експеримента посматрани догађај A реализовати тачно k пута. То значи да се у преосталих $(n-k)$ понављања догађај A неће реализовати. При томе је k један одређени број из скупа бројева {0, l, 2, ..., n}. Означићемо ту вероватноћу са $P_n(k)$.

Пре него пређемо на одређивање вероватноће $P_n(k)$ у општем случају, посматрајмо следећи пример:

Претпоставимо да се у једној кутији налази 5 сијалица и да је вероватноћа да је сијалица исправна једнака за све сијалице и износи 0,95. Одредимо вероватноћу да се у кутији налазе тачно три исправне и две неисправне сијалице.

Случајни догађај A који посматрамо у појединачном извођењу ,,експеримента" је догађај: ,,сијалица је исправна", а супротан догађај је $\overline{\mathrm{A}}$: ,,сијалица је неисправна". При томе је $p=0.95$, $q=0.05$, а број понављања експеримента је једнак броју сијалица, тј. $n=5$. Тражена вероватноћа је $P_5(3)$.

Ако за сваку сијалицу придружимо A или $\overline{\mathrm{A}}$, онда ће се догађај да међу пет сијалица буду три исправне реализовати ако се реализовала једна од могућности писања низа од пет словних места у којима је на 3 места слово A, а на два места слово $\overline{\mathrm{A}}$. Једна од могућности је, на пример, низ

$$\mathrm{AA}\overline{\mathrm{A}}\,\mathrm{A}\overline{\mathrm{A}}\,.$$

Вероватноћа да се оствари овај низ је једнака

$$0,95\cdot 0,95\cdot 0,05\cdot 0,95\cdot 0,05=0,95^3\cdot 0,05^2,$$

јер су догађаји у овом низу независни. Оваквих могућих низова има укупно 10. То су

1. $\mathrm{AAA}\overline{\mathrm{A}}\,\overline{\mathrm{A}}$	6. $\mathrm{A}\overline{\mathrm{A}}\,\mathrm{A}\overline{\mathrm{A}}\,\mathrm{A}$
2. $\mathrm{AA}\overline{\mathrm{A}}\,\mathrm{A}\overline{\mathrm{A}}$	7. $\overline{\mathrm{A}}\,\mathrm{AA}\overline{\mathrm{A}}\,\mathrm{A}$
3. $\mathrm{A}\overline{\mathrm{A}}\,\mathrm{AA}\overline{\mathrm{A}}$	8. $\mathrm{A}\overline{\mathrm{A}}\,\overline{\mathrm{A}}\,\mathrm{AA}$
4. $\overline{\mathrm{A}}\,\mathrm{AAA}\overline{\mathrm{A}}$	9. $\overline{\mathrm{A}}\,\mathrm{A}\overline{\mathrm{A}}\,\mathrm{AA}$
5. $\mathrm{AA}\overline{\mathrm{A}}\,\overline{\mathrm{A}}\,\mathrm{A}$	10. $\overline{\mathrm{A}}\,\overline{\mathrm{A}}\,\mathrm{AAA}$.

Сваки од ових низова има исту вероватноћу и она је једнака

$$0,95^3\cdot 0,05^2\,.$$

Наведени низови се међусобно искључују јер се у једном пакету од 5 сијалица може реализовати тачно један од њих.

Дефинисани догађај да се међу 5 изабраних сијалица налази тачно 3 неисправне, према томе, реализује се кад се реализовао један од наведених низова. Зато је његова вероватноћа једнака збиру вероватноћа сваког од њих, тј.

$$\mathrm{P}_5(3)=10\cdot 0,95^3\cdot 0,05^3\,.$$

На сличан начин се могу одредити и вероватноће

$$\mathrm{P}_5(0), \mathrm{P}_5(\mathrm{l}), \mathrm{P}_5(2), \mathrm{P}_5(4), \mathrm{P}_5(5)\,.$$

Посматрајмо, сада, општи случај. Ако се n пута понавља експеримент, одредимо вероватноћа да ће се догађај А реализовати у k понављања, а у преосталих $(n-k)$ неће.

Сваком извођењу експеримента приписаћемо слово *A* или $\overline{A}$, зависно од тога да ли се догађај *A* реализовао или није. На тај начин ћемо, низу од n понављања експеримента придружити низ од n словних места, који је састављен од слова *A* и $\overline{A}$. Посматрани догађај ће се реализовати ако се реализује низ у коме ће на k места бити слово *A*, а на преосталих $(n-k)$ места $\overline{A}$.

Вероватноћа сваког таквог низа је једнака:

$$p^k q^{n-k}$$

због независности догађаја.

Са друге стране, сви ти низови се међусобно искључују, јер једном Бернулијевом експерименту од n понављања одговара тачно један од тих могућих низова у којима се посматрани догађај *A* реализовао k пута. Таквих различитих низова има онолико колико има могућности да се од n словних места изабере k места на која се ставља слово *A*, (на преостала места се ставља $\overline{A}$), а тај број је једнак:

$$C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Низ вероватноћа:

$$P_n(0),\ P_n(1), \ldots,\ P_n(n)$$

где је:

$$P_n(k) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

је низ биномних вероватноћа *и представља вероватноће да се у Бернулијевом експерименту догађај A реализује тачно k пута, при чему је k=0,1,2,...,n.*

Биномне вероватноће $P_n(k)$ имају мале вредности за велики број n. Неке вредности вероватноће за $n = 50$ и $p = \frac{1}{3}$ су дате у Табели 4.1. Највећа вероватноћа једнака је *0,1178*. (За $n = 50$ и $p = 0{,}5$, највећа вероватноћа је мања од *0,08*.) Из Табеле 4.1. види се да биномне вероватноће расту до извесне вредности, а затим опадају.

k	$P_n(k)$	k	$P_n(k)$
5	0,0001	17	0,1178
8	0,0033	18	0,1080
10	0,0157	19	0,0910
12	0,0470	20	0,0704
14	0,0879	22	0,0332
15	0,1077	25	0,0059
16	0,1178	30	0,0001

Табела 4.1. Биномне вероватноће за $n = 50$ *и* $p = \frac{1}{3}$.

Значи да ће се за неку вредност k добити највећа вредност биномних вероватноћа. Интересује нас која је то вредност.

Биномне вероватноће за k и $(k+1)$ су:

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

$$P_n(k+1) = \frac{n!}{(k+1)!(n-k-1)!} p^{k+1} q^{n-k-1}$$

Њихов количник је број

$$\frac{P(k+1)}{P(k)} = \frac{n-k}{k+1}\frac{p}{q} \tag{4.25}$$

Ако је овај количник већи од један, значиће да са порастом k бинома вероватноћа расте. То ће бити случај кад је

$$\frac{n-k}{k+1}\frac{p}{q} \geq 1$$

односно кад је

$$k \leq np - q$$

За

$$k > np - q$$

количник (4.25) је мањи од један, што значи да ће за $k > np - q$ биномне вероватноће опадати.

Ако је $(np-q)$ цео број, количник биномних вероватноћа за $(np-q+1)$ и $(np-q)$ биће једнак

$$\frac{P_n(np-q+1)}{P_n(np-q)} = \frac{n-(np-q)}{np-q+1}\frac{p}{q}$$

Односно

$$\frac{P_n(np-q+1)}{P_n(np-q)}=1$$

На основу изложеног се може закључити:

Ако $np-q$ *није цео број, онда је број понављања у којима се реализује случајан догађај А са највећом вероватноћом, једнак:*

$$k=\left[np-q\right]+1$$

при чему је $[\ ]$ *ознака за највећи цео број који није већи од* $np-q$.

Ако је $np-q$ *цео број, онда постоје две вредности* k *које су највероватније, а то су вредности:*

$$k=np-q \qquad\qquad k+1=np-q+1$$

За те вредности, биномне вероватноће су међусобно једнаке, и веће су од свих осталих вероватноћа.

4.2. ГАУСОВА КРИВА

У примени теорије вероватноће велику улогу има функција позната под именом ***Гаусова крива (Gauss)*** *или* ***Гаус- Лапласова крива (Laplace)*** која се често зове и ***закон вероватноћа Нормалне расподеле***.

Гаусова крива је крива дата функцијом

$$\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \tag{4.2}$$

која је дефинисана за свако реално x из интервала $(-\infty, +\infty)$.

Функција $\phi(x)$ је симетрична функција која има максимум у тачки $x=0$. Максимална вредност ове функције је

$$\phi(0)=\frac{1}{\sqrt{2\pi}}\approx 0{,}3989\,.$$

Превојне тачке функције $\Phi(x)$ су тачке $x=\pm 1$ у којима функција има вредност

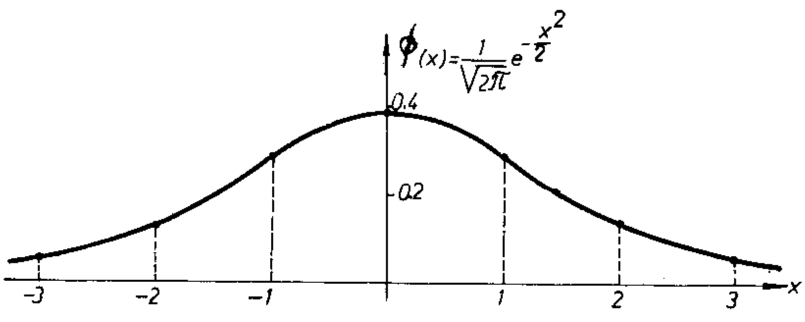
$$\phi(\pm 1)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}}\approx 0{,}2420\,.$$

Поред тога је

$$\lim_{x\to+\infty}\phi(x)=0\,,$$

$$\lim_{x\to-\infty}\phi(x)=0\,.$$

Гаусова крива је приказана на слици 4.1.



Слика 4.1. Гаусова крива $\phi(x)$

Површина између криве $\phi(x)$ и осе x је једнака интегралу

$$\int_{-\infty}^{+\infty}\phi(x)dx$$

и има вредност једнаку јединици, тј.

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}dx = 1 \qquad (4.3)$$

Вредности функције $\phi(x)$ дају се табеларно, најчешће за вредности x из интервала [0; 3,99]. За $x = 3,99$ функција има вредност *0,0001*, а за вредности x веће од *3,99* функција $\phi(x)$ има мале вредности које су практично једнаке 0.

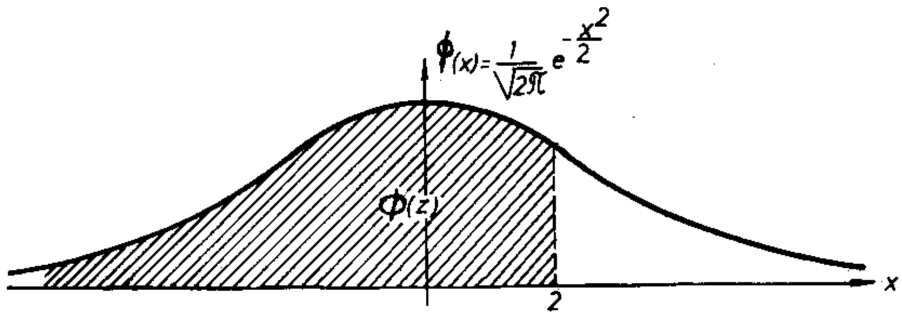
На основу Гаусове криве, дефинисана је тзв. **функција Нормалне расподеле** $\Phi(z)$ чије су вредности дате табеларно и која служи за одређивање површине између криве $\phi(x)$ и осе x у неком произвољном интервалу [a; b].

Дефиниција 4.1. **Функција Нормалне расподеле** $\Phi(z)$ *дефинисана је за свако* $z \in (-\infty, +\infty)$ *једначином*

$$\Phi(z) = \int_{-\infty}^{z} \phi(x)dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{x^2}{2}}dx$$

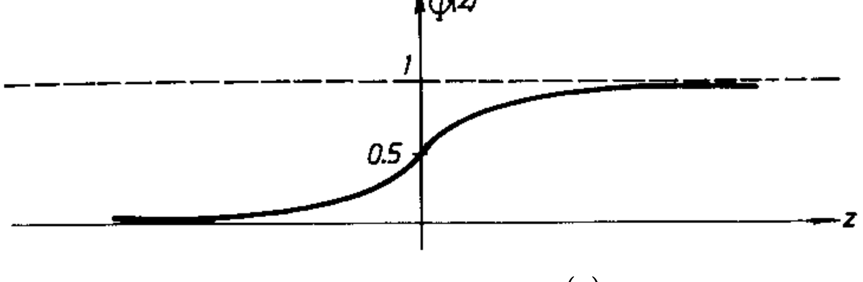
и представља површину између криве $\phi(x)$ *и осе x у интервалу* $(-\infty, z]$.

Графички приказ преко површине функције $\phi(x)$ дат је на слици 4.2.



Слика 4.2. Функција Нормалне расподеле приказана као површина

Функција Нормалне расподеле $\Phi(z)$ је монотоно растућа функција која има хоризонталне асимптоте *y=0*, када $x \to -\infty$, и *y = 1*, када $x \to \infty$. Графички је приказана на слици 4.3.



Слика 4.3. Функција $\Phi(z)$

Вредности функције нормалне расподеле $\Phi(z)$ дате су табеларно за позитивне вредности z. Због симетрије Гаусове криве, важи:

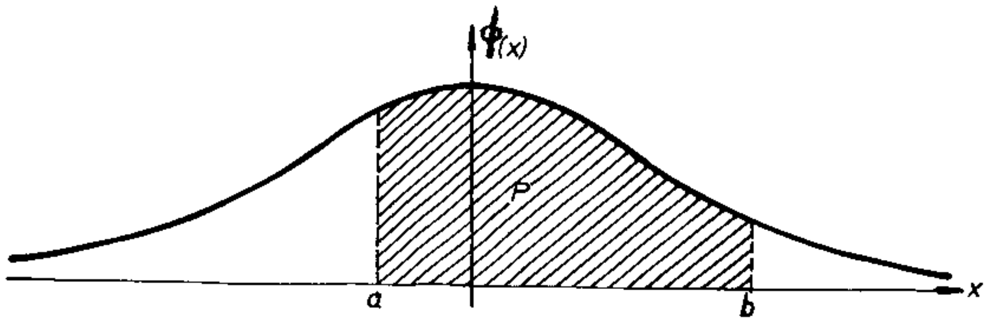
$$\Phi(-z) = 1 - \Phi(z)$$

што није тешко проверити, на основу дефиниције (4.1).

У табелама функције Нормалне расподеле најчешће су дате вредности функције за z из интервала [0; 3,49]. За вредности веће од 3,49 функција Φ практично има вредност 1.

Да би одредили ***површину између Гаусове криве и осе x*** *у произвољном интервалу* $(a, b]$ *користимо функцију* $\Phi(z)$. *Важи:*

$$P = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx = \Phi(b) - \Phi(a)$$



Слика 4.4. Приказ површине и функције $\phi(z)$

Најчешће се посматра симетрични интервал $[-z,\ z]$. Површина између функције $\phi(x)$ и осе x, у том интервалу једнака је

$$P = \frac{1}{\sqrt{2\pi}} \int_{-z}^{z} e^{-\frac{x^2}{2}} dx = 2\Phi(z) - 1 .$$

Тако се добија да је површина у интервалу $(-1, 1]$ једнака:

$$P = 0{,}6826 = 68{,}3\% ,$$

а у интервалу $(-2, 2]$:

$$P = 0{,}9544 = 95{,}4\% ,$$

док је за интервал $(-3, 3]$

$$P = 0{,}9974 = 99{,}7\% ,$$

што значи да се 99,7% од укупне површине између Гаусове криве $\phi(x)$ и осе x, налази у интервалу (-3; 3] док је изван тог интервала свега 0,03% те површине.

4.3. БЕРНУЛИЈЕВ И БОРЕЛ-КАНТЕЛИЈЕВ ЗАКОН ВЕЛИКИХ БРОЈЕВА

Посматраћемо Бернулијеву шему независних понављања експеримента са случајним догађајем A и вероватноћом p реализације догађаја A у сваком понављању.

Посматрајмо следећи проблем: ***Ако се n пута понавља експеримент, да ли је између релативне фреквенцијое остварења посматраног догађаја и вероватноће p може уочити нека зависност?***

Одговор на постављено питање дао је J. Bernoulli (1654-1705. године) у познатом делу „*Ars Conjectandi*” објављеном 1713. године. Бернулијеве резултате уопштио је Моавр (Abraham De Moivre) 1718. године, а затим је Лаплас (Pierre S. Laplace) 1812. године дао најопштију формулацију одговора на постављено питање.

Ради бољег разумевања правог значења ових резултата, размотримо прво Бернулијеву формулацију.

Означимо са X број понављања експеримента у којима се реализовао догађај A. Ако се експеримент понови n пута X ће узети једну вредност из низа могућих вредности

$$\{0,1,2,...,n\}$$

Догађај $D=\{X=k\}$ је случајни догађај који се у Бернулијевом експерименту реализује са вероватноћом

$$P_n(k)=P\{X=k\}=\frac{n!}{k!(n-k)!}p^k q^{n-k} \tag{4.41}$$

Количник

$$\frac{X}{n} \tag{4.42}$$

је **фреквенција остварења догађаја A** *у* ***n*** *пута поновљеном експерименту.*

Фреквенција остварења догађаја A може узети једну вредност из низа могућих вредности

$$\{0,\frac{1}{n},\frac{2}{n},...,1\}.$$

Догађај $D`=\left\{\frac{X}{n}=\frac{k}{n}\right\}$ је случајни догађај који се у Бернулијевом експерименту реализује са вероватноћом датом изразом (4.41).

Интересује нас разлика између фреквенције (4.42) и вероватноће p. Посматрајмо апсолутну вредност те разлике.

$$\left|\frac{X}{n}-p\right|. \tag{4.43}$$

Ако експеримент поновимо n пута, разлика (4.43) ће узети неку вредност из интервала [0; 1]. Шта се дешава са разликом када се повећава број понављања експеримента?

Изаберимо произвољно мали позитиван број ε и посматрајмо догађај да ће разлика (4.43) бити већа од тога броја, тј. посматрајмо догађај

$$D = \left\{ \left| \frac{X}{n} - p \right| > \varepsilon \right\}, \tag{4.44}$$

и одредимо вероватноћу остварења овог догађаја у n понављања експеримента.

Догађај D ће се остварити кад је

$$X > n(p+\varepsilon) \tag{4.45}$$

или кад је

$$X < n(p-\varepsilon) \tag{4.46}$$

(Види слику 4.1.).

Слика 4.1.

Вероватноћа догађаја датог изразом (4.45) је једнака збиру биномних вероватноћа

$$P\{X > n(p+\varepsilon)\} = \sum_{i=r}^{n} P_n(i) \tag{4.47}$$

при чему је r најмањи цео број који је већи од производа $n(p+\varepsilon)$.

Могуће је показати да је посматрана вероватноћа ограничена са горње стране тако да је

$$P\{X > n(p+\varepsilon)\} < P_n(r)\frac{n(p+\varepsilon)+q}{n\varepsilon + q}. \tag{4.48}$$

Када број понављања експеримента n неограничено расте, израз на десној страни (4.8) остаје ограничен. Из дефиниције биномних вероватноћа се закључује да је:

$$\lim_{n\to\infty} P_n(r) = 0,$$

Следи да је

$$\lim_{n\to\infty} P\{X > n(p+\varepsilon)\} = 0, \tag{4.49}$$

На сличан начин се може утврдити да је гранична вредност вероватноће догађаја датог изразом (4.46) такође једнака нули, тј. да је

$$\lim_{n\to\infty} P\{X < n(p-\varepsilon)\} = 0. \tag{4.50}$$

Из наведеног следи да је:

$$\lim_{n\to\infty} P\left\{\left|\frac{X}{n} - p\right| > \varepsilon\right\} = 0 ,$$

одакле следи:

Теорема 4.2. ***(Бернулијева теорема).*** *Вероватноћа да ће фреквенција остварења догађаја А одступити од његове вероватноће p за вредност мању од произвољно малог броја* ε *тежи јединици кад број понављања експеримента n неограничено расте.тј.*

$$\lim_{n\to\infty} P\{|\frac{X}{n} - p| \le \varepsilon\} = 1 \tag{4.51}$$

Теорема 4.2. представља **Бернулијев закон великих бројева**.

Ако је број понављања n довољно велики, теорема тврди да је скоро сигурно да ће разлика релативне фреквенције остварења догађаја А и његове вероватноће бити произвољно мала. То значи да се на основу релативне фреквенције може скоро са сигурношћу утврдити вероватноћа неког случајног догађаја.

Тек када је доказан овај закон, било је могуће прихватити дефиницију емпиријске вероватноће и на тај начин, експерименталним путем, одређивати вероватноће случајних догађаја.

Овај закон одиграо је пресудну улогу у развоју и примени теорије вероватноће, односно у развоју статистике, јер је тек његовим доказивањем постало јасно да постоје законитости код појава и процеса које нису детерминистичког карактера, тј. код појава и процеса стохастичког карактера.

Са развојем теорије вероватноће утврђене су и општије формулације закона великих бројева. Са становишта одређивања самог појма вероватноће и њене суштине, интересантно је уопштење које су дали **Борел и Кантели** и које је познато под називом *„**јаки закон великих бројева**”*, а гласи:

Ако је X број експеримента у којима се случајни догађај А остварио, вероватноћа да ће гранична вредност фреквенција остварења догађаја А, када се број понављања експеримената неограничено повећава, бити једнака вероватноћи p је једнака јединици, тј.

$$P\{\lim_{n\to\infty} \frac{X}{n} = p\} = 1 \tag{4.52}$$

Закони дати изразима (4.51) и (4.52) потпуно описују фундаментално својство случајности које је могуће проучавати преко вероватноће.

4.4. МОАВР-ЛАПЛАСОВО РЕШЕЊЕ БЕРНУЛИЈЕВОГ ПРОБЛЕМА

У разматрањима Бернулијевог експеримента и закона великих бројева остао је нерешен проблема израчунавања биномних вероватноћа за велике вредности n и k. Било би пожељно да се утврди начин приближног рачунања ових вероватноћа, али тако да буде обезбеђена жељена прецизност.

Решења за овај проблем први је дао Моавр за случај када је $p = 0{,}5$, а затим је Лаплас то решење уопштио и за произвољне вредности p. Та решења дата су у Теоремама 4.3 и 4.4.

Теорема 4.3. (Локална теорема Моавр-Лапласа): *Нека је вероватноћа остварења неког догађаја A у сваком од n независних понављања једног експеримента једнака p, (0 < p < 1), и нека је X број понављања експеримента у којима се реализовао догађај A. Означимо са*

$$P\{X = k\}$$

вероватноћу да ће се догађај A реализовати k пута.

Тада је низ

$$\sqrt{npq}P\{X = k\} = \sqrt{npq}P_n(k)$$

униформно конвергентан за свако k за које се

$$\frac{k - np}{\sqrt{npq}} \tag{4.38}$$

налази у неком коначном интервалу. При томе је

$$\lim_{n\to\infty}\sqrt{npq}P_n(k) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(k-np)^2}{npq}} = \phi\left(\frac{k-np}{\sqrt{npq}}\right) \tag{4.39}$$

где је функција $\phi(x)$ Гаусова крива.

Из теореме 4.3. непосредно следи:

Биномне вероватноће $P_n(k)$ за довољно велико n имају приближне вредности

$$P_n(k) \approx \frac{1}{\sqrt{npq}}\phi(x). \tag{4.55}$$

при чему је

$$x = \frac{k - np}{\sqrt{npq}} \tag{4.56}$$

А функција $\phi(x)$ ја Гаусова крива

(4.57) $$\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

односно закон вероватноћа Нормалне расподеле.

Приближне вредности биномних вероватноћа дате преко Гаусове криве изразом (4.55) су боље што је *n* веће. Међутим, чак и за релативно мале вредности n приближне вредности су довољно добра апроксимација.

Теорема 4.4. (Интегрална теорема Моавр-Лапласа). *Нека је X број понављања експеримента у којима се реализовао догађај A, а p вероватноћа његовог остварења у сваком понављању, (0 < p < 1). Тада ће вероватноћа да се вредност*

(4.58) $$\frac{X-np}{\sqrt{npq}}$$

нађе у интервалу (a,b] тежити интегралу

$$\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx=\int_a^b \phi(x)dx$$

тј. тада је

(4.59) $$\lim_{n\to\infty}P\left\{a<\frac{X-np}{\sqrt{npq}}\le b\right\}=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx=\int_a^b \phi(x)dx\,.$$

Интегрална теорема Моавр-Лапласа има веома широку примену у одређивању вероватноћа различитих догађаја везаних за Бернулијеву шему независних понављања једног експеримента.

На основу граничне вредности дате у теореми 4.4 можемо приближно одређивати збирове биномних вероватноћа. Наиме, из (4.59) следи да је за довољно велико *n*

(4.61) $$P\left\{a<\frac{X-np}{\sqrt{npq}}\le b\right\}\approx\int_a^b \phi(x)dx$$

Пошто је вероватноћа на левој страни (4.61) једнака збиру одговарајућих биномних вероватноћа, то значи да је

(4.62) $$\sum_k P_n(k)\approx\int_a^b \phi(x)dx$$

при чему збир иде по k од $(np-a\sqrt{npq})$ до $(np+b\sqrt{npq})$.

Интеграл на десној страни (4.61) и (4.62) може се одредити преко таблица за функцију Нормалне расподеле $\Phi(z)$. Вероватноћа у (4.61) је приближно једнака

(4.63) $$P\left\{a<\frac{X-np}{\sqrt{npq}}\le b\right\}\approx\Phi(b)-\Phi(a)$$

тј. $$\sum_k P_n(k)\approx\Phi(b)-\Phi(a)$$

4.5. ПУАСОНОВО РЕШЕЊЕ ЗА РЕТКЕ ДОГАЂАЈЕ

Одређивање биномних вероватноћа $P_n(k)$ у Бернулијевој шеми независних понављања експеримента није сасвим једноставно за велике бројеве n. Моавр-Лапласова теорема омогућава да се приближно одреде посматране вероватноће за случај када np и npq расту и кад $n \to \infty$.

Међутим, постоје случајеви када је Моавр-Лапласова апроксимација незадовољавајућа. За те случајеве користи се једина апроксимација која је математички много једноставнија од претходне.

Пуасон је дао формулу која представља веома добру апроксимацију биномних вероватноћа за ретке догађаје, тј. за догађаје чија вероватноћа реализације у једном извођењу је мала.

Означимо са λ производ

$$\lambda = np,$$

при чему је n број понављања експеримента, а p вероватноћа реализације догађаја A у сваком понављању експеримента. Тада се биномне вероватноће

$$P_n(k)$$

могу написати у облику:

$$. \lim_{n\to+\infty} P_n(k) = \frac{\lambda^k}{k!} e^{-\lambda} \tag{4.74}$$

Израз (4.74) омогућава да се изврши апроксимација биномних вероватноћа.

Ако је n довољно велико, а p мало, тако да је вредност производа np мала (np <10), тада се биномне вероватноће могу приближно одређивати по формули:

(4.76.) $$P_n(k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

при чему је $\lambda = np$.