# *Тема № 11. Неравенство Чебышева. Теорема Чебышева. Теорема Бернулли. Центральная предельная теорема. Интегральная теорема Муавра-Лапласа*

## Содержание

# Предельные теоремы теории вероятности

Рассмотрим ряд утверждений и теорем из большой группы так называемых предельных теорем теории вероятностей, устанавливающие связь между теоретическими и экспериментальными характеристиками случайных величин при большом числе испытаний над ними. Они составляют основу математической статистики. Предельные теоремы условно делят на две группы. Первая группа теорем, называемая ***законом больших чисел*** (коротко:ЗБЧ), устанавливает устойчивость средних значений: при большом числе испытаний их средний результат перестает быть случайным и может быть предсказан с достаточной точностью. Вторая группа теорем, называемая ***центральной предельной теоремой*** (коротко:ЦПТ), устанавливает условия, при которых закон распределения суммы большого числа случайных величин неограниченно приближается к нормальному.

В начале рассмотрим неравенство Чебышева, которое можно использовать:

а) для грубой оценки вероятностей событий, связанных с с. в., распределение которых неизвестно;

б) доказательства ряда теорем ЗВЧ.

# Неравенство Чебышева

**Теорема 1** *Если с. в. $X$ имеет м. о. $MX = a$ и дисперсию $DX$, то для любого $\varepsilon > 0$ справедливо неравенство Чебышева*

$$P\{|X - MX| \geqslant \varepsilon\} \leqslant \frac{DX}{\varepsilon^2} \tag{11.1}$$

Докажем неравенство (11.1) для непрерывной с. в. $X$ с плотностью $f(x)$. Вероятность $P\{|X - a| \geqslant \varepsilon\}$ есть вероятность попадания с. в. $X$ в область, лежащую вне промежутка $[a - \varepsilon, a + \varepsilon]$. Можно записать

$$P\{|X-a| \geqslant \varepsilon\} = \int_{-\infty}^{a-\varepsilon} f(x)dx + \int_{a+\varepsilon}^{+\infty} f(x)dx = \int_{|x-a|\geqslant\varepsilon} f(x)dx =$$

$$= \int_{|x-a|\geqslant\varepsilon} 1 \cdot f(x)dx \leqslant \int_{|x-a|\geqslant\varepsilon} \frac{(x-a)^2}{\varepsilon^2} f(x)dx,$$

так как область интегрирования $|x-a| \geqslant \varepsilon$ можно записать в виде $(x-a)^2 \geqslant \varepsilon^2$, откуда следует $1 \leqslant \dfrac{(x-a)^2}{\varepsilon^2}$. Имеем

$$P\{|X-a| \geqslant \varepsilon\} = \frac{1}{\varepsilon^2} \int_{|x-a|\geqslant\varepsilon} (x-a)^2 f(x)dx \geqslant \frac{1}{\varepsilon^2} \int_{-\infty}^{\infty} (x-a)^2 f(x)dx,$$

так как интеграл от неотрицательной функции при расширении области интегрирования может только возрасти. Таким образом,

$$P\{|X-a| \geqslant \varepsilon\} = \frac{1}{\varepsilon^2} \int_{-\infty}^{\infty} (x-a)^2 f(x)dx = \frac{1}{\varepsilon^2} DX.$$

Аналогично доказывается неравенство Чебышева и для дискретной с. в. X, принимающей значения $x_1, x_2, x_3, \ldots$ с вероятностями $p_1, p_2, p_3, \ldots$, только интегралы (вида $\int\limits_{|x-a|\geqslant\varepsilon}$) заменяются соответствующими суммами (вида $\sum\limits_{|x_i-a|}$).

Отметим, что неравенство Чебышева можно записать в другой форме:

$$P\{|X - MX| < \varepsilon\} = 1 - \frac{DX}{\varepsilon^2} \quad , \tag{11.2}$$

В форме (11.2) оно устанавливает нижнюю границу вероятности события, а в форме (11.1) — верхнюю.

Неравенство Чебышева справедливо *для любых с. в.* В частности: для с. в. $X = m$, имеющей ***биномиальное распределение*** с матожиданием $MX = a = np$ дисперсией $DX = npq$ (п. 2.7), оно принимает вид

$$P\{|m - np| < \varepsilon\} = \frac{npq}{\varepsilon^2}; \tag{11.3}$$

для частости $m/n$ (или $\dfrac{n_A}{n}$, п. 1.5) события в $n$ независимых испытаниях, в каждом из которых оно может произойти с вероятностью $p = M(\dfrac{m}{n}) = a$, дисперсия которых $D(\dfrac{m}{n}) = \dfrac{pq}{n}$, неравенство Чебышева имеет вид

$$P\{|\frac{m}{n} - p| < \varepsilon\} = 1 - \frac{pq}{n\varepsilon^2}. \tag{11.4}$$

Оценку вероятности попадания с. в. $X$ в промежуток $(\varepsilon, \infty)$ дает неравенство Маркова.

**Теорема 2** *(Неравенство Маркова). Для любой неотрицательной с.в. $X$, имеющей м.о. $MX$ и $\varepsilon > 0$, справедливо неравенство:*

$$P\{X \geqslant \varepsilon\} \leqslant \frac{MX}{\varepsilon}. \tag{11.5}$$

$$P\{X \geqslant \varepsilon\} = \int\limits_{\varepsilon}^{\infty} f(x)dx \leqslant \int\limits_{\varepsilon}^{\infty} \frac{x}{\varepsilon} f(x)dx = \frac{1}{3} \int\limits_{\varepsilon}^{\infty} xf(x)dx \leqslant \frac{1}{3} \int\limits_{0}^{\infty} xf(x)dx = \frac{MX}{\varepsilon}$$

Неравенство (11.5) можно записать в форме

$$P\{X < \varepsilon\} > 1 - \frac{MX}{\varepsilon}$$

**Пример 5.1** Оценить с помощью неравенства Чебышева вероятность того, что отклонение с. в. $X$ от своего м. о. будет меньше трех с. к. от е. меньше $3\sigma_x$.

Полагая $\varepsilon = 3\sigma_x$ в формуле (11.2), получаем

$$P\{|X - MX| < 3\sigma_x\} \geqslant 1 - \frac{\sigma_x^2}{(3\sigma_x)} = 1 - \frac{1}{9} = \frac{8}{9} \approx 0,8889.$$

Эта оценка, как известно (п. 2.7), называется ***правилом трех сигм***; для с. в. $X \sim N(a, \sigma)$ эта вероятность равна 0,9973.

## Теорема Чебышева

Основное утверждение ЗБЧ содержится в теореме Чебышева. В ней и других теоремах ЗБЧ используется понятие «сходимости случайных величин по вероятности».

Случайные величины $X_1, X_2, \ldots, X_n, \ldots$ *сходятся по вероятности* к величине $A$ (случайной или неслучайной), если для любого $\varepsilon > 0$ вероятность события $\{|X_n - A| < \varepsilon\}$ при $n \to \infty$ стремится к единице, т.е.

$$\lim_{n\to\infty} P\{|X_n| - A < \varepsilon\} = 1$$

(или $P\{|X_n - A| < \varepsilon\} \to 1$). Сходимость по вероятности символически записывают так:

$$X_n \xrightarrow{P} A \quad \text{при} \quad n \to \infty.$$

Следует отметить, что ***сходимость по вероятности*** требует, чтобы неравенство $|X_n - A| < \varepsilon$ выполнялось *для подавляющего числа членов* последовательности (в математическом анализе — для всех $n > N$, где $N$ — некоторое число), а при $n \to \infty$ практически все члены последовательности должны попасть в $\varepsilon$-окрестность $A$.

**Теорема 3** *(ЗБЧ в формуле П.Л.Чебышева,1886г.). Если случайные величины $X_1, X_2, \ldots, X_n, \ldots$ независимы и существует такое число $C > 0$, что $DX_i \leqslant C$ $i = 1, 2, \ldots$, то для любого $\varepsilon > 0$*

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} MX_i\right| < \varepsilon\right\} = 1, \tag{11.6}$$

*т.е. среднее арифметическое этих случайных величин сходится по вероятности к среднему арифметическому их м.о.:*

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{P} \frac{1}{n}\sum_{i=1}^{n} MX_i.$$

Так как $DX_i \leqslant C$, $I = 1, 2, \ldots$, то

$$D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} D\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} DX_i =$$

$$= \frac{1}{n^2}(DX_1 + DX_2 + \ldots + DX_n) \leqslant \frac{1}{n^2}(C + C + \ldots + C) = \frac{1}{n^2}Cn = \frac{C}{n}.$$

Тогда, применяя к с. в.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

неравенство Чебышева (11.2), имеем

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} MX_i\right| < \varepsilon\right\} \geqslant 1 - \frac{D\left(\frac{1}{n}\sum\limits_{i=1}^{n} X_i\right)}{\varepsilon^2} \geqslant 1 - \frac{C}{n\varepsilon^2}. \tag{11.7}$$

Переходя к пределу при $n \to \infty$ и учитывая, что вероятность любого события не превышает 1, получаем

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} MX_i\right| < \varepsilon\right\} = 1.$$

**Следствие.** Если с. в. $X_1, X_2, \ldots, X_n, \ldots$ независимы и одинаково распределены, $MX_i = a$, $DX_i = \sigma^2$, то для любого $\varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - a\right| < \varepsilon\right\} = 1, \tag{11.8}$$

т. е. среднее арифметическое с. в. сходится по вероятности к математическому ожиданию $a$:

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{P} a.$$

Так как

$$\frac{1}{n}\sum_{i=1}^{n} MX_i = \frac{1}{n}(MX_1 + MX_2 + \ldots + MX_n) = \frac{1}{n^2}(a + a + \ldots + a) = \frac{1}{n}na = a,$$

а дисперсии с. в. $X$ равны числу $\sigma^2$, т. е. ограничены, то, применив ЗБЧ в форме Чебышева (11.6), получим утверждение (11.8).

Следствие (11.8) теоремы Чебышева обосновывает «принцип среднего арифметического с.в. $X_i$», постоянно используемый на практике. Так, пусть произведено $n$ независимых измерений некоторой величины, истинное значение которой $a$ (оно неизвестно). Результат каждого с.в. измерения есть с.в. $X_i$. Согласно следствию (11.8) в качестве приближенного значения величины а можно взять среднее арифметическое результатов измерений:

$$a = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}.$$

Равенство тем точнее, чем больше $n$.

На теореме Чебышева основан также широко применяемый в статистике ***выборочный метод***, суть которого в том, что о качестве большого количества однородного материала можно судить по небольшой его пробе.

Теорема Чебышева подтверждает связь между случайностью и необходимостью: среднее значение *случайной величины*

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

практически не отличается от неслучайной величины

$$\frac{1}{n}\sum_{i=1}^{n} MX_i.$$

**Пример 5.2** Глубина моря измеряется прибором, не имеющим систематической ошибки. Среднее квадратическое отклонение измерений не превосходит 15 м. Сколько нужно сделать независимых измерений, чтобы с вероятностью, не меньшей 0,9, можно было утверждать, что среднее арифметическое этих измерений отличается от $a$ (глубины моря) по модулю меньше, чем на 5 м?

Обозначим через $X_i$ результаты $n$ независимых измерений глубины моря. Нужно найти число $n$, которое удовлетворяет неравенству (11.5):

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} MX_i\right| < \varepsilon\right\} \geqslant 1 - \frac{C}{n\varepsilon^2} \quad ,$$

где $MX_i = a$, что означает отсутствие при измерениях систематической ошибки (т. е. измерения производятся с одинаковой точностью). По условию $\varepsilon = 5$, $C = 225$, так как $\sigma = \sqrt{DX} = 15$ м. Отсюда

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - a\right| < 5\right\} \geqslant 1 - \frac{225}{25n} \geqslant 0,9,$$

т. е. $0,1 \geqslant 9/n$, $n \geqslant 90$. Измерение нужно проводить не менее 90 раз.

# Теорема Бернулли

Теорема Бернулли исторически является первой и наиболее простой формой закона больших чисел. Она теоретически обосновывает свойство устойчивости относительной частоты (см. п. 1.5).

**Теорема 4** *(ЗБЧ в форме Я. Бернулли, 1713 г.). Если вероятность появления события $A$ в одном испытании равна $p$, число наступления этого события при $n$ независимых испытаниях равно $n_A$, то для любого числа $\varepsilon > 0$ имеет место равенство*

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1, \tag{11.9}$$

*т.е. относительная частота $P^*(A)$ события $A$ сходится по вероятности $p$ события $A$:* $P^*(A)\xrightarrow[n\to\infty]{P} P(A)$ .

Введем с. в. $X_1, X_2, \ldots, X_n$ следующим образом: $X_i = 1$, если в $i$-м испытании появилось событие $A$, а если не появилось, то $X_i = 0$. Тогда число $n_A$ (число успехов) можно представить в виде

$$n_A = \sum_{i=1}^{n} X_i.$$

М.о. и дисперсия с.в. $X_i$ равны: $MX_i = 1\cdot p + 0\cdot(1-p) = pDXi = (0-p)^2(1-p) + (1-p)^2 p = p(1-p) = pq$. Закон распределения с. в. $X_i$ имеет вид

| $X_i$ | $0$ | $1$ |
|---|---|---|
| $P$ | $1-p$ | $p$ |

при любом $i$. Таким образом, с. в. $X_i$ независимы, их дисперсии ограничены одним и тем же числом $1/4$, так как

$$p(1-p) = p - p^2 = 1/4 - (p - 1/2)^2 \leqslant 1/4.$$

Поэтому к этим с.в. можно применить теорему Чебышева (11.6):

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} MX_i\right| < \varepsilon\right\} = 1.$$

Но

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{n_A}{n}, \quad \frac{1}{n}\sum_{i=1}^{n} MX_i = \frac{1}{n}np = p.$$

Следовательно,

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1.$$

Теорема Бернулли теоретически обосновывает возможность приближенного вычисления вероятности события с помощью его относительной частоты. Так, например, за вероятность рождения девочки можно взять относительную частоту этого события, которая, согласно статистическим данным, приближенно равна 0,485.

Неравенство Чебышева (11.2) для случайных величин

$$\frac{1}{n}\sum_{i=1}^{n} X_i = n_A$$

принимает вид

$$P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} \geqslant 1 = \frac{pq}{n\varepsilon^2}. \tag{11.10}$$

Обобщением теоремы Бернулли на случай, когда вероятности $p_i$, появления события $A$ в каждом из $n$ испытаний различны, является *теорема Пуассона*:

$$P^*(A)\frac{P}{n \to \infty}\frac{1}{n}\sum_{i=1}^{n} p_i, \tag{11.11}$$

где $p_i$ — вероятность события $A$ в $i$-м испытании.

**Пример 5.3**. Вероятность наличия опечатки на одной странице рукописи равна 0,2. Оценить вероятность того, что в рукописи, содержащей 400 страниц, частость появления опечатки отличается от соответствующей вероятности по модулю меньше чем 0,05.

Воспользуемся формулой (11.10). В данном случае $p = 0,2$, $q = 0,8$, $n = 400$, $\varepsilon = 0,05$. Имеем

$$P\left\{\left|\frac{n_A}{n} - 0,2\right| < 0,05\right\} \geqslant \left[1 - \frac{pq}{n\varepsilon^2}\right] = 1 - \frac{0,2 \cdot 0,8}{400 \cdot 0,05^2} = 0,84,$$

т. е. $p \geqslant 0,84$.

## Центральная предельная теорема

Центральная предельная теорема (ЦПТ) представляет собой вторую группу предельных теорем, которые устанавливают связь между законом распределения суммы с. в. и его предельной формой — нормальным законом распределения.

Сформулируем ЦПТ для случая, когда члены суммы имеют одинаковое распределение (именно эта теорема чаще других используется на практике, так, в математической статистике выборочные случайные величины имеют одинаковые распределения, так как получены из одной и той же генеральной совокупности).

**Теорема 5** *Пусть св.* $X_1, X_2, \dots, X_n$ *независимы, одинаково распределены, имеют конечные математическое ожидание* $MX_i = a$ *и дисперсию* $DX_i = \alpha^2$, $i = \overline{1, n}$. *Тогда функция распределения центрированной и нормированной суммы этих случайных величин стремится при* $n \to \infty$ *к функции распределения стандартной нормальной случайной величины:*

$$Z_n = \frac{\sum\limits_{i=1}^{n} X_i - M\left(\sum\limits_{i=1}^{n} X_i\right)}{\sqrt{D\left(\sum\limits_{i=1}^{n} X_i\right)}} = \frac{\sum\limits_{i=1}^{n} X_i - na}{\sigma\sqrt{n}}, \tag{11.12}$$

$$F_{Z_n}(x) = P\{Z_n < x\} \xrightarrow[n\to\infty]{} \Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt.$$

Из соотношения (11.12) следует, что при достаточно большом $n$ сумма $Z_n$ приближенно распределена по нормальному закону: $Z_n \sim N(0, 1)$. Это означает, что сумма $S_n = X_1, X_2, \dots, X_n$ приближенно распределена по нормальному закону: $S_n \sim N(na, \sqrt{n}\sigma)$. Говорят, что при $n \to \infty$ с. в. $\sum\limits_{i=1}^{n} X_i$ *асимптотически нормальна.*

Напомним, что:

1. С.в. $X$ называется центрированной и нормированной (т.е. стандартной), если $MX = 0$, а $DX = 1$.

2. Если с. в. $X_i, i = 1, n$ независимы, $MX_i = a$, $DX_i = \sigma^2$ , то

$$M\left(\sum_{i=1}^{n} X_i\right) = a + a + \ldots + a = na,$$

$$D\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} DX_i = \sigma^2 + \sigma^2 + \ldots + \sigma^2 = n\sigma^2.$$

3. $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\dfrac{t^2}{2}} dt$ — функция Лапласа; $\Phi(x) = 1/2 + \Phi_0(x)$, где

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}} dt.$$

Формула (11.12) позволяет при больших $n$ вычислять вероятности различных событий, связанных с суммами случайных величин. Так, перейдя от с.в.

$$S_n = \left(\sum_{i=1}^{n} \chi_i\right) = \left(\sum_{i=1}^{n} X_i\right)$$

к стандартной с.в., получим:

$$P\left\{\alpha \leqslant \sum_{i=1}^{n} X_i \leqslant \beta\right\} = P\left\{\frac{\alpha - na}{\sigma\sqrt{n}} \leqslant \frac{\sum\limits_{i=1}^{n} X_i - na}{\sigma\sqrt{n}} \leqslant \frac{\beta - na}{\sigma\sqrt{n}}\right\} \approx \Phi\left(\frac{\beta - na}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{\alpha - na}{\sigma\sqrt{n}}\right),$$

$$P\{\alpha \leqslant S_n \leqslant \beta\} \approx \Phi\left(\frac{\beta - MS_n}{\sqrt{DS_n}}\right) - \Phi\left(\frac{\alpha - MS_n}{\sqrt{DS_n}}\right), \tag{11.13}$$

формула для определения вероятности того, что сумма нескольких с. в. окажется в заданных пределах (см. (2.43) и (2.46)). Часто ЦПТ используют, если $n > 10$.

**Пример 5.4** Независимые с. в. $X_i$ распределены равномерно на отрезке $[0, 1]$. Найти закон распределения с. в.

$$Y = \sum_{i=1}^{100} X_i,$$

а также вероятность того, что $55 < Y < 70$.

Условия ЦПТ соблюдаются, поэтому с. в. $Y$ имеет приближенно плотность распределения

$$f_Y(y) \approx \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(y-m_y)^2}{2\sigma_y^2}}.$$

По известным формулам для м. о. и дисперсии в случае равномерного распределения находим: $MX_i = (0+1)/2 = 1/2$, $DX_i = (1-0)^2/12 = 12$, $\sigma X_i = \dfrac{1}{\sqrt{12}} = \dfrac{1}{2\sqrt{3}}$. Тогда:

$$m_y = M\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} MX_i = 100 \cdot \frac{1}{2} = 50,$$

$$\sigma_y^2 = D\left(\sum_{i=1}^{100} X_i\right) = \sum_{i=1}^{100} DX_i = 100 \cdot \frac{1}{12} = \frac{25}{3}, \quad \sigma_y = \frac{5\sqrt{3}}{3}.$$

Поэтому

$$f_Y(y) \approx \frac{3}{5\sqrt{6\pi}} e^{-\frac{3(y-50)^2}{50}}.$$

Используя формулу (11.13), находим

$$P\{55 < Y < 70\} \approx \Phi\left(\frac{70-50}{\frac{5\sqrt{3}}{3}}\right) - \Phi\left(\frac{55-50}{\frac{5\sqrt{3}}{3}}\right) = \Phi(4\sqrt{3}) - \Phi(\sqrt{3}) = \Phi(6,9282) - \Phi(1,73) \approx 0,04,$$

т.е. $P\{55 < Y < 70\} \approx 0.04$

# Интегральная теорема Муавра-Лапласа

Следствиями ЦПТ являются рассмотренные ранее (п. 1.21) локальная и интегральная теоремы Муавра-Лапласа. Приведем вывод интегральной теоремы Муавра-Лапласа.

Рассмотрим схему Бернулли. Пусть $n_A$ — число появления события $A$ (число успехов) в $n$ независимых испытаниях, в каждом из которых оно может появиться с вероятностью $p$ (не появиться — с вероятностью $q = 1 - p$). Случайную величину $n_A$ можно представить в виде суммы $n$ независимых с.в.

$X_1, X_2, \ldots, X_n$ , таких, что $X_i = 1$, если в $i$-м испытании событие $A$ наступило, и $X_i = 0$ в противном случае, т.е.

$$n_A = \sum_{i=1}^{n} X_i.$$

Так как $MX_i = p$, $DX_i = pq$ (п. 5.3), то

$$Mn_A = M\sum_{i=1}^{n} X_i = np, \quad Dn_A = D\sum_{i=1}^{n} X_i = npq,$$

(да это уже давно известно, см. (2.24), ведь с. в. $n_A$ имеет биномиальный закон распределения). Тогда с.в.

$$Z_n = \frac{\sum\limits_{i=1}^{n} X_i - M\left(\sum\limits_{i=1}^{n} X_i\right)}{\sqrt{D\left(\sum\limits_{i=1}^{n} X_i\right)}} = \frac{n_A - np}{\sqrt{npq}}$$

представляет также сумму $n$ независимых, одинаково распределенных случайных величин. При этом $Zn \sim N(0,1)$, действительно:

$$MZ_n = M\left(\frac{n_A - np}{\sqrt{npq}}\right) = \frac{np - np}{\sqrt{npq}} = 0$$

и

$$DZ_n = D\left(\frac{n_A - np}{\sqrt{npq}}\right) = \frac{1}{npq} Dn_A = \frac{npq}{npq} = 1 \quad .$$

Следовательно, на основании ЦПТ (11.12) с.в. $Z_n$ при большом числе $n$ имеет приближенно нормальное распределение. Согласно свойству (2.46) нормального закона, записываем

$$P\{z_1 \leqslant Z_n \leqslant z_2\} \approx \Phi(z_2) - \Phi(z_1)$$

($\Phi(z)$ — функция Лапласа). Полагая

$$z_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad z_2 = \frac{k_2 - np}{\sqrt{npq}}$$

двойное неравенство

$$\frac{k_1 - np}{\sqrt{npq}} \leqslant \frac{n_A - np}{\sqrt{npq}} \leqslant \frac{k_2 - np}{\sqrt{npq}}$$

можно переписать в эквивалентном виде $k_1 \leqslant n_A \leqslant k_2$ Таким образом, получаем

$$P\{k_1 \leqslant n_A \leqslant k_2\} = \Phi(z_2) - \Phi(z_1),$$

т. е. *интегральную формулу Муавра-Лапласа.*

Заметим, что интегральная предельная теорема Муавра-Лапласа

$$P\left\{\frac{n_A - np}{\sqrt{npq}} < x\right\} \to \Phi(x), \tag{11.14}$$

для схемы Бернулли непосредственно вытекает из ЦПТ (11.12) с учетом результатов, полученных в п. 5.3 ($n_A = \sum\limits_{i=1}^{m} X_i$, $MX_i = p$, $DX_i = pq$, тогда $na = np$, $\sigma\sqrt{n} = \sqrt{pq}\sqrt{n} = \sqrt{npq}$).

Для подсчета сумм биномиальных вероятностей можно воспользоваться приближенной формулой

$$\sum_{k=0}^{m} P_{n,k} \approx \Phi\left(\frac{m-np}{\sqrt{npq}}\right). \tag{11.15}$$

Действительно,

$$\sum_{k=0}^{m} P_{n,k} = P\{n_A \leqslant m\} = P\{-\infty < n_A \leqslant m\} =$$

$$= P\left\{-\infty < \frac{n_A-np}{\sqrt{npq}} < \frac{m-np}{\sqrt{npq}}\right\} = \Phi\left(\frac{m-np}{\sqrt{npq}}\right) - \Phi(-\infty) = \Phi\left(\frac{m-np}{\sqrt{npq}}\right),$$

где $\Phi(x)$ — функция Лапласа.

**Пример 5.5** Машинистке требуется напечатать текст, содержащий 8000 слов, состоящих из четырех и более букв. Вероятность сделать ошибку в любом из этих слов равна 0,01. Какова вероятность того, что при печатании будет сделано не более 90 ошибок?

Применим формулу (11.15). Так как $n = 8000$, $p = 0,01$, $q = 0,99$, $m = 90$, то

$$\frac{1}{\sqrt{npq}} = \frac{1}{\sqrt{8000 \cdot 0.01 \cdot 0.99}} = \frac{1}{\sqrt{79.2}} \propto 0.112,$$

$$\frac{m-np}{\sqrt{npq}} \approx \frac{10}{8.9} \approx 1.12,$$

$$\Phi(1,12) = 0,8686.$$

Следовательно, $P\{n_A \leqslant 90\} = \sum\limits_{k=0}^{90} P_{n,k} \approx 0.869$ .

# Предметный указатель