**Московский государственный университет
экономики, статистики и информатики
Московский международный институт эконометрики,
информатики, финансов и права**

Иванов В.С.

Теория вероятностей

(ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ)

Москва, 2002

УДК
ББК
И 907

Иванов В.С., Теория вероятностей. /М. Московский международный институт эконометрики, информатики, финансов и права. 2002.

Содержание

1. Закон больших чисел и центральная предельная теорема.

Математические законы теории вероятностей получены абстрагированием реальных статистических закономерностей, свойственных массовым случайным явлениям. Наличие этих закономерностей связано именно с массовостью явлений, то есть с большим числом выполняемых однородных опытов или с большим числом складывающихся случайных воздействий, порождающих в своей совокупности случайную величину, подчиненную вполне определенному закону. Свойство устойчивости массовых случайных явлений известно человечеству еще с глубокой древности. В какой бы области оно ни проявлялось, суть его сводится к следующему: конкретные особенности каждого отдельного случайного явления почти не сказываются на среднем результате массы таких явлений; случайные отклонения от среднего, неизбежные в каждом отдельном явлении, в массе взаимно погашаются, нивелируются, выравниваются. Именно эта *устойчивость средних* и представляет собой физическое содержание “закона больших чисел”, понимаемого в широком смысле слова: при очень большом числе случайных явлений средний их результат практически перестает быть случайным и может быть предсказан с большой степенью определенности.

В узком смысле слова под “законом больших чисел” в теории вероятностей понимается ряд математических теорем, в каждой из которых для тех или иных условий устанавливается факт приближения средних характеристик большого числа опытов к некоторым определенным постоянным.

Возможности предсказаний в области массовых случайных явлений еще больше расширяются наличием другой группы предельных теорем, касающихся уже не предельных значений случайных величин, а *предельных законов распределения.* Речь идет о группе теорем, известных под названием “центральной предельной теоремы.” Мы уже говорили о том, что при суммировании достаточно большого числа случайных величин закон распределения суммы неограниченно приближается к нормальному при соблюдении некоторых условий. Эти условия, которые математически можно формулировать различным образом - в более или менее общем виде, - по существу сводятся к требованию, чтобы влияние на сумму отдельных слагаемых было *равномерно малым*, т. е. чтобы в состав суммы не входили члены, явно преобладающие над совокупностью остальных по своему влиянию на рассеивание суммы. Различные формы центральной предельной теоремы различаются между собой теми условиями, для которых устанавливается это предельное свойство суммы случайных величин.

Различные формы закона больших чисел вместе с различными формами центральной предельной теоремы образуют совокупность так называемых *предельных теорем* теории вероятностей. Предельные теоремы дают возможность не только осуществлять научные прогнозы в области случайных явлений, но и оценивать точность этих прогнозов.

2. Неравенство Чебышева.

В качестве леммы, необходимой для доказательства теорем, относящихся к группе "закона больших чисел" , мы докажем одно весьма общее неравенство, известное
под названием *неравенства Чебышева.*

Пусть имеется случайная величина X с математическим ожиданием m_x и дисперсией D_x . Неравенство Чебышева утверждает, что, каково бы ни было положительное число α, вероятность того, что величина X отклонится от своего математического ожидания не меньше чем на α, ограничена сверху величиной $\frac{Dx}{\alpha^2}$

$$P(|X - m_x| \geq \alpha) \leq \frac{Dx}{\alpha^2}. \tag{2.1}$$

Доказательство. 1. Пусть величина X прерывная, с рядом распределения

x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

Изобразим возможные значения величины X и ее математическое ожидание m_x в виде точек на числовой оси Ох (рис.2.1).

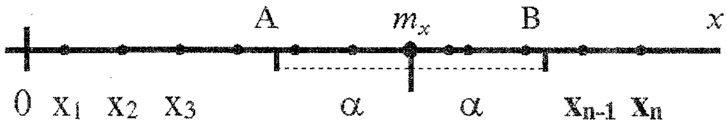
Зададимся некоторым значением α>0 и вычислим вероятность того, что величина X отклонится от своего математического ожидания не меньше чем на α:

$$P(|X - m_x| \geq \alpha). \tag{2.2}$$

Для этого отложим от точки m_x вправо и влево по отрезку длиной α; получим отрезок АВ. Вероятность (2.2) есть не что иное, как вероятность того, что случайная точка X попадет не внутрь отрезка АВ, а вовне его:

$$P(|X - m_x| \geq \alpha) = P(X \not\subset AB)^{1}).$$

[1] Концы отрезка АВ мы в него не включаем.



(рис. 2.1)

Для того чтобы найти эту вероятность, нужно просуммировать вероятности всех тех значений $\mathbf{x}_i$, которые лежат *вне* отрезка AB. Это мы запишем следующим образом:

$$P(|X - m_x| \ge \alpha) = \sum_{|\mathbf{x}_i - m_x| \ge \alpha} P_i , \qquad (2.3)$$

где запись $|\mathbf{x}_i - m_x| \ge \alpha$ под знаком суммы означает, что суммирование распространяется на все те значения $\boldsymbol{i}$, для которых точки $\mathbf{x}_i$ лежат вне отрезка AB.

С другой стороны, напишем выражение дисперсии величины X. По определению:

$$D_x = M[(X - m_x)^2] = \sum_{i=1}^{n} (\mathbf{x}_i - m_x)^2 P_i = \sum_{i=1}^{n} |\mathbf{x}_i - m_x|^2 P_i \qquad (2.4)$$

Так как все члены суммы (2.4) неотрицательны, она может только уменьшиться, если мы распространим её не на все значения $\mathbf{x}_i$, а только на некоторые, в частности на те, которые лежат вне отрезка AB:

$$D_x \ge \sum_{|\mathbf{x}_i - m_x| \ge \alpha} |\mathbf{x}_i - m_x|^2 P_i \qquad (2.5)$$

Заменим под знаком суммы выражение $|\mathbf{x}_i - m_x|$ через α. Так как для всех членов суммы $|\mathbf{x}_i - m_x| \ge \alpha$, то от такой замены сумма тоже может только уменьшиться; значит,

$$D_x \ge \sum_{|\mathbf{x}_i - m_x| \ge \alpha} \alpha^2 P_i = \alpha^2 \sum_{|\mathbf{x}_i - m_x| \ge \alpha} P_i \qquad (2.6)$$

Но согласно формуле (2.3) сумма, стоящая в правой части (2.6), есть не что иное, как вероятность попадания случайной точки вовне отрезка AB; следовательно,

$$D_x \ge \alpha^2 P(|X - m_x| \ge \alpha),$$

откуда непостедственно вытекает доказываемое неравенство.

2. В случае, когда величина X непрерывна, доказательство проводится аналогичным образом с заменой вероятностей P_i елементом вероятности, а конечных сумм - интегралами. Действительно,

$$P(|X - m_x| \geq \alpha) = \int\limits_{|x - mx| > \alpha} f(x)dx^{2}), \qquad (2.7)$$

где $f(x)$ - плотность распределения величины X. Далее, имеем:

$$D_x = \int\limits_{-\infty}^{\infty} (x - m_x)^2 f(x)dx = \int\limits_{-\infty}^{\infty} |x - m_x|^2 f(x)dx \geq \int\limits_{|x - mx| > \alpha} |x - m_x|^2 f(x),$$

где знак $|x - m_x| > \alpha$ под интегралом означает, что интегрирование распространяется на внешнюю часть отрезка AB.

Заменяя $|x - m_x|$ под знаком интеграла через α, получим:

$$D_x \geq \alpha^2 \int\limits_{|x - mx| > \alpha} f(x)dx = \alpha^2 P(|X - m_x| > \alpha),$$

откуда и вытекает неравенство Чебышева для непрерывных величин.

Примечание. Неравенство Чебышева даёт только верхнюю границу вероятности данного отклонения. Выше этой границы вероятность не может быть ни при каком законе распределения. На практике в большинстве случаев вероятность того, что величина X выйдет за пределы участка $m_x \pm 3\sigma_x$, значительно меньше 1/9. Например, для нормального закона эта вероятность приблизительно равна 0,003. На практике чаще всего мы имеем дело со случайными величинами, значения которых только крайме редко выходят за пределы $m_x \pm 3\sigma_x$. Если закон распределения случайной величины неизвестен, а известны только m_x и σ_x , на практике обычно считают отрезок $m_x \pm 3\sigma_x$ участком практически возможных значений случайной величины (так называемое “правило трёх сигма”).

3. Закон больших чисел (теорема Чебышева)

В данном n^0 мы докажем одну из простейших, но вместе с тем наиболее важных форм закона больших чисел - теорему Чебышева. эта теорема устанавливает связь между средним арифметическим наблюденных значений случайной величины и её математическим ожиданием.

Предварительно решим следующую вспомогательную задачу.

Имеется случайная величина X с математическим ожиданием m_x и дисперсией D_x . Над этой величиной производится n независимых

---

[2] Знак $\geq$ заменён знаком >, так как для непрерывной величины вероятность точного равенства равна нулю.

опытов и вычисляется среднее арифметическое всех наблюденных значений величины X. Требуется найти числовые характеристики этого среднего арифметического - математическое ожидание и дисперсию - и выяснить, как они изменяются с увеличением n.

Обозначим:

X_1 - значение величины X в первом опыте;

X_2 - значение величины X во втором опыте, и т. д.

Очевидно, совокупность величин $X_1, X_2, \ldots, X_n$ представляет собой n независимых случайных величин, каждая из которых распределена по тому е закону, что и сама величина X. Рассмотрим среднее арифметическое этих величин:

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

Случайная величина Y есть линейная функция независимых случайных величин $X_1, X_2, \ldots, X_n$. Найдём математическое ожидание и дисперсию этой величины. Согласно правилам n^0 10 для определения числовых характеристик линейных функций получим:

$$m_y = M[Y] = \frac{1}{n} \sum_{i=1}^{n} M[X_i] = \frac{1}{n}\, n\, m_x = m_x \;;$$

$$D_y = \frac{1}{n^2} \sum_{i=1}^{n} D[X_i] = \frac{D_x}{n}.$$

Итак, математическое ожидание величины У не зависит от числа опытов n и равно математическому ожиданию наблюдаемой величины X; что касается дисперсии величины Y, то она неограниченно убывает с увеличением числа опытов и при достаточно большом n может быть сделана сколь угодно малой. Мы убеждаемся, что среднее арифметическое есть случайная величина со сколько угодно малой дисперсией и при большом числе опытов ведёт себя почти как не случайная.

Теорема Чебышева и устанавливает в точной количественной форме это свойство устойчивости среднего арифметического. Она формулируется следующим образом:

При достаточно большом числе независимых опытов среднее арифметическое наблюденных значений случайной величины сходится по вероятности к её математическому ожиданию.

Запишем теорему Чебушева в виде формулы. Для этого напомним смысл термина “сходится по вероятности”. Говорят, что случайная величина X_n сходится по вероятности к величине α, если при увеличении n вероятность того, что X_n и α будут сколь угодно близкими, неограниченно приближается к единице, а это значит, что при достаточно большом n

$$P(|X_n - \alpha| < \varepsilon) > 1 - \delta,$$

где ε и δ - произвольно малые положительные числа.

Запишем в аналогичной форме теорему Чебышева. Она утверждает, что при увеличении n среднее арифметическое

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

сходится по вероятности к m_x, т. е.

$$P\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - m_x\right| < \varepsilon\right] > 1 - \delta \qquad (3.1)$$

Докажем это неравенство.

Доказательство. Выше было показано, что величина

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

имеет числовые характеристики

$$D_y = \frac{D_x}{n} \; ; \quad m_y = m_x .$$

Применим к случайной величине У неравенство Чебышева, полагая $\alpha = \varepsilon$:

$$P(|Y - m_y| \geq \varepsilon) \leq \frac{D_y}{\varepsilon^2} = \frac{D_x}{n\varepsilon^2} .$$

Как бы мало ни было число ε, можно взять n таким большим, чтобы выполнялось неравенство

$$\frac{D_x}{n\varepsilon^2} < \delta ,$$

где δ - сколь угодно малое число.

Тогда

$$P\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - m_x\right| \geq \varepsilon\right] < \delta,$$

откуда, переходя к противоположному событию, имеем:

$$P\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - m_x\right| < \varepsilon\right] > 1 - \delta,$$

что и требовалось доказать.

4. Обобщенная теорема Чебышева. Теорема Маркова.

Теорема Чебышева легко может быть обобщена на более сложный случай, а именно когда закон распределения случайной величины X от опыта к опыту не остается одним и тем же, а изменяется. Тогда вместо среднего арифметического наблюденных значений одной и той же величины X с постоянными математическим ожиданием и дисперсией мы имеем дело со средним арифметическим n различных случайных величин, с различными математическими ожиданиями и дисперсиями. Оказывается, что и в этом случае при соблюдении некоторых условий среднее арифметическое является устойчивым и сходится по вероятности к определенной неслучайной величине.

Обобщенная теорема Чебышева формулируется следующим образом. *Если*

$$X_1, X_2, \ldots, X_n \text{ -}$$

независимые случайные величины с математическими ожиданиями

$$m_{x1}, m_{x2}, \ldots, m_{xn}$$

и дисперсиям

$$D_{x1}, D_{x2}, \ldots, D_{xn}$$

и если все дисперсии ограничены сверху одним и тем же числом L:

$$D_{xi} < L \qquad (i=1,2,\ldots,n),$$

то при возрастании n среднее арифметическое наблюденных значений величин $X_1, X_2, \ldots, X_n$ сходится по вероятности к среднему арифметическому их математических ожиданий[3]).

Запишем эту теорему в виде формулы. Пусть ε, δ — сколь угодно малые положительные числа. Тогда при достаточно большом n

[3] То есть разность между тем и другим средним арифметическим сходится по вероятности к 0.

$$P\left(\left|\frac{\sum_{i=1}^{n} x_i}{n} - \frac{\sum_{i=1}^{n} m_{x_i}}{n}\right| < \varepsilon\right) > 1-\delta$$

(4.1)

Доказательство. Рассмотрим величину

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

Ее математическое ожидание равно:

$$m_y = \frac{\sum_{i=1}^{n} m_{x_i}}{n},$$

А дисперсия

$$D_y = \frac{1}{n^2}\sum_{i=1}^{n} D_{x_i}.$$

Применим к величине Y неравенство Чебышева:

$$P(|Y - m_y| \geq \varepsilon) \leq \frac{D_y}{\varepsilon^2}.$$

Или

$$P\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} m_{xi}}{n}\right| \geq \varepsilon\right] \leq \frac{\sum_{i=1}^{n} D_{xi}}{n^2\varepsilon^2}. \qquad (4.2)$$

Заменим в правой части неравенства (4.2) каждую из величин D_{x_i} большей величиной L. 'Тогда неравенство только усилится:

$$P\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} m_{xi}}{n}\right| \geq \varepsilon\right] < \frac{L}{n\varepsilon^2}.$$

Как бы мало ни было ε можно выбрать n настолько большим, чтобы выполнялось неравенство

$$\frac{L}{n\varepsilon^2} < \delta$$

тогда

$$P\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} m_{xi}}{n}\right| \geq \varepsilon\right] < \delta.$$

откуда, переходя к противоположному событию, получим доказываемое неравенство (4.1).

Закон больших чисел может быть распространен и на зависимые случайные величины. Обобщение закона больших чисел на случай зависимых случайных величин принадлежит А. А.Маркову.

Теорема Маркова. *Если имеются зависимые случайные величины X_1, X_2, ...,X_N и если при* $n \to \infty$ $\frac{D\left[\sum_{i=1}^{n} X_i\right]}{n^2} \to 0$

то среднее арифметическое наблюденных значений случайных величин X_1, X_2, ...,X_N сходятся по вероятности к среднему арифметическому их математических ожиданий

Доказательство. Рассмотрим величину

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

Очевидно,

$$D_y = \frac{D\left[\sum_{i=1}^{n} X_i\right]}{n^2}.$$

Применим к величине Y неравенство Чебышева:

$$P(|Y - m_y| \geq \varepsilon) \leq \frac{D_y}{\varepsilon^2}.$$

Так как по условию теоремы при $n\to\infty$ $D_y\to0$, то при достаточно большом n

$$P(|Y - m_y| \geq \varepsilon) < \delta,$$

или, переходя к противоположному событию,

$$P(|Y - m_y|) < \varepsilon = P\left[\left|\frac{\sum_{i=1}^{n} X_i}{n} - \frac{\sum_{i=1}^{n} m_{xi}}{n}\right| < \varepsilon\right] > 1-\delta.$$

что и требовалось доказать.

5. Следствия закона больших чисел: теоремы Бернулли и Пуассона

Известная теорема Я. Бернулли, устанавливающая связь между частотой события и его вероятностью, может быть доказана как прямое следствие закона больших чисел.

Пусть производится n независимых опытов, в каждом из которых может появиться или не появиться некоторое событие А, вероятность которого в каждом опыте равна *p.* Теорема Я. Бернулли утверждает, что *при неограниченном увеличении числа опытов n частота события A сходится по вероятности к его вероятности p.*

Обозначим частоту события A в n опытах через P^* и запишем теорему Я. Бернулли в виде формулы

$$P(\,|\,P^{*}\ \text{-}p\,|<\varepsilon\,)>1\text{-}\delta, \qquad (5.1)$$

где ε,δ—сколь угодно малые положительные числа.

Требуется доказать справедливость этой формулы при достаточно большом n.

Доказательство. Рассмотрим независимые случайные величины:

X_1 — число появлений события А в первом опыте;

X_2 — число появлений события А во втором опыте, и т. д.

Все эти величины прерывны и имеют один и тот же закон распределения, выражаемый рядом вида

0	1
q	p

где q==1—p. Математическое ожидание каждой из величин X_i равно p, а ее дисперсия pq .

Частота P* представляет собой не что иное. как среднее арифметическое величин $X_1, X_2, \ldots, X_N$:

$$P^* = \frac{\sum\limits_{i=1}^{n} X_i}{n},$$

и, согласно закону больших чисел, сходится по вероятности к общему математическому ожиданию этих случайных величин. Отсюда и следует справедливость неравенства (5.1)

Теорема Я. Бернулли утверждает устойчивость частоты при постоянных условиях опыта. Но при изменяющихся условиях опыта аналогичная устойчивость также существует. Теорема, устанавливающая свойство устойчивости частот при переменных условиях опыта, называется *теоремой Пуассона* и формулируется следующим образом:

Если производится n независимых опытов и вероятность появления события A в i-м опыте равна P_I , *то при увеличении n частота события A сходится по вероятности к среднему арифметическому вероятностей* p_i..

Теорема Пуассона выводится из обобщенной теоремы Чебышева точно так же, как теорема Бернулли была выведена из закона больших чисел.

Теорема Пуассона имеет большое принципиальное значение для практического применения теории вероятностей. Дело в том, что зачастую вероятностные методы применяются для исследования явлений, которые в одних и тех же условиях не имеют шансов повториться достаточно много раз, но повторяются многократно при весьма разнообразных условиях, причем вероятности интересующих нас событий сильно зависят от этих условий. Например, вероятность поражения цели в воздушном бою существенно зависит от дальности стрельбы, ракурса цели, высоты полета, скорости стреляющего самолета и цели и т. д. Комплекс этих условий слишком многочислен для того, чтобы можно было рассчитывать на многократное осуществление воздушного боя именно в данных фиксированных условиях. И все же, несмотря на это, в данном явлении налицо определенная устойчивость частот, а именно частота поражения цели в реальных воздушных боях, осуществляемых в самых разных условиях, будетприближаться к средней вероятности поражения цели, характерной для данной группы условий. Поэтому те методы организации стрельбы, которые основаны на максимальной вероятности поражения цели, будут оправданы и в данном случае, несмотря на то, что нельзя ожидать подлинной массовости опытов в каждом определенном комплексе условий.

Аналогичным образом обстоит дело в области опытной проверки вероятностных расчетов. На практике очень часто встречается случай, когда требуется проверить на опыте соответствие вычисленной вероятности какого-либо события A его фактической частоте. Чаще всего это делается для того, чтобы проверить правильность той или иной теоретической схемы, положенной в основу метода вычисления вероятности события. Зачастую при такой экспериментальной проверке не удается воспроизвести достаточно много раз одни и те же условия опыта. И все же эта проверка может быть осуществлена, если сравнить наблюденную в опыте частоту события не с его вероятностью для фиксированных условий, а со средним арифметическим вероятностей, вычисленных для различных условий.

6. Массовые случайные явления и центральная предельная теорема

В предыдущих n°n° мы рассмотрели различные формы закона больших чисел. Все эти формы, как бы они ни были различны, утверждают одно: факт сходимости по вероятности тех или иных случайных величин к определенным постоянным. Ни в одной из форм закона больших чисел мы не имеем дела с законами распределения случайных величин. Предельные законы распределения составляют предмет другой группы теорем — центральной предельной теоремы, которую иногда называют “количественной формой закона больших чисел”.

Все формы центральной предельной теоремы посвящены установлению условий, при которых возникает нормальный закон распределения. Так как эти условия на практике весьма часто выполняются, нормальный закон является самым распространенным из законов распределения, наиболее часто встречающимся в' случайных явлениях природы. Он возникает во всех случаях, когда исследуемая случайная величина может быть представлена в виде суммы достаточно большого числа независимых (или слабо зависимых) элементарных слагаемых, каждое из которых в отдельности сравнительно мало влияет на сумму.

В теории стрельбы нормальный закон распределения играет особо важную роль, так как в большинстве случаев практики координаты точек попадания и точек разрыва снарядов распределяются по нормальному закону. Объяснить это можно на следующем примере.

Пусть производится стрельба по некоторой плоской мишени, с центром которой (точкой прицеливания) связано начало координат. Точка попадания характеризуется двумя случайными величинами: X и Y. Рассмотрим одну из них, например отклонение X точки попадания от цели в направлении оси O_x. Это отклонение вызвано совокупным действием очень большого количества сравнительно малых факторов, как-то: ошибка наводки, ошибка в определении дальности до цели, вибрации орудия и установки при стрельбе, ошибки изготовления снаряда, атмосферные условия и т. д. Каждая из этих причин создает элементарную ошибку—отклонение снаряда îò цели, и координата снаряда X может быть представлена как сумма таких элементарных отклонений:

$$X=X_1+X_2,+\ldots+X_N+\ldots \qquad (6.1)$$

где $X_1,$, X_2, ...—отклонения, вызванные отдельными факторами. Так как этих факторов очень много, между собой они являются в основном независимыми и по влиянию на сумму отдельные слагаемые можно считать приблизительно равномерно малыми, то налицо условия применимости центральной предельной теоремы, и величина (6.1) должна, подчиняться закону распределения, близкому к нормальному.

Остановимся несколько подробнее на нашем утверждении о приблизительно равномерно малом влиянии каждого из слагаемых на сумму. Смысл его в том. что среди элементарных ошибок стрельбы нет ни одной резко превалирующей над суммой всех остальных. Действительно, если бы такая ошибка была, нужно думать, что, составляя правила стрельбы или конструируя прицельный прибор, мы постарались бы ликвидировать эту ошибку и учесть заранее самую значительную причину, отклоняющую снаряд от цели. Неучтенные случайные факторы, создающие рассеивание, обычно характерны своей равномерной малостью и отсутствием среди них резко преобладающих. Именно поэтому закон распределения точек попадания снарядов (или закон распределения точек разрыва снарядов при дистанционной стрельбе) обычно принимается нормальным[4]).

[4] В некоторых случаях стрельбы фактическое распределение точек попадания на плоскости может сильно отличаться от нормального, например, при стрельбе в резко

Нормальный закон распределения является доминирующим не только в теории стрельбы, но и во многих других областях, например в теории ошибок измерения. Именно исходя из теории ошибок измерения нормальный закон и был впервые обоснован Лапласом и Гауссом. Действительно, в большинстве случаев ошибки, возникающие при измерении тех или иных физических величин, распределяются именно по нормальному закону; причина этого в том, что такие ошибки, как правило, складываются из многочисленных независимых элементарных ошибок, порождаемых различными причинами. Долгое время нормальный закон считался единственным и универсальным законом ошибок. В настоящее время взгляд на нормальный закон как на единственный и универсальный должен быть пересмотрен (опыт показывает, что в ряде процессов измерения и производства наблюдаются законы распределения, отличные от нормального), но все же нормальный закон остается самым распространенным и самым важным для практики законом ошибок.

7. Характеристические функции.

Одна из наиболее общих форм центральной предельной теоремы была доказана А. М. Ляпуновым в 1900 г. Для доказательства этой теоремы А. М. Ляпунов создал специальный метод характеристических функций. В дальнейшем этот метод приобрел самостоятельное значение и оказался весьма мощным и гибким методом, пригодным для решения самых различных вероятностных задач.

Характеристической функцией случайной величины X называется функция

$$g(t)=M[e^{itx}], \qquad (7.1)$$

где i—мнимая единица. Функция g(t) представляет собой математическое ожидание некоторой комплексной случайной величины

$$U=e^{itx}.$$

функционально связанной с величиной X. При решении многих задач теории вероятностей оказывается удобнее пользоваться характеристическими функциями, чем законами распределения.

Зная закон распределения случайной величины X, легко найти её характеристическую фуикцию.

Если X—непрерывная случайная величина с рядом распределения

x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

то ее характеристическая функция

переменных условиях, когда центр рассеивания и вероятностное отклонение в процессе стрельбы заметно меняются. Однако в таких случаях мы фактически имеем дело не с законом распределения координат точки попадания при одном выстреле, а со средним из таких законов для различных выстрелов.

$$g(t)=\sum_{k=1}^{n} e^{itx_k} p_k \quad (7.2)$$

Если X—непрерывная случайная величина с плотностью распределения f(X), то ее характеристическая функция

$$g(t)=\int_{-\infty}^{\infty} e^{itx} f(x)\, dx \quad (7.3)$$

Пример. Случайная величина X имеет нормальное распределение:

$$f(x)=\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad (7.4)$$

Определить ее характеристическую функцию.

Решение. По формуле (7.3) имеем:

$$g(t)=\int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx-x^2/2}\, dx \quad . \quad (7.5)$$

Пользуясь известной формулой

$$\int_{-\infty}^{\infty} e^{-Ax^2 \pm 2Bx - C}\, dx = \sqrt{(\pi / A)}\, e^{-(AC-B)^2/A}$$

и имея в виду, что $i^2 = -1$, получим:

$$g(t) = e^{-t^2/2} \quad . \quad (7.6)$$

Формула (7.3) выражает характеристическую функцию g(t) непрерывной случайной величины X через ее плотность распределения f(x). Преобразование (7.3), которому нужно подвергнуть f(x), чтобы получить g(t), называется *преобразованием Фурье*. В курсах математического анализа доказывается, что если функция g(t) выражается через f (x) с помощью преобразования Фурье, то, в свою очередь, функция f (x) выражается через g(t) с помощью так называемого *обратного преобразования Фурье*:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} g(t)dt. \quad (7.7)$$

Сформулируем и докажем основные свойства характеристических функций.

1. *Если случайные величины X и Y связаны соотношением*

Y=aX,

где *a—неслучайный множитель, то их характеристические функции связаны соотношением*:

$$g_y(t) = g_x(at). \quad (7.8)$$

Доказательство:

$$g_y(t) = M[e^{ity}] = M[e^{itax}] = M[e^{i(at)x}] = g_x(at).$$

2. *Характеристическая функция суммы независимых случайных величин равна произведению характеристических функций слагаемых.*

Доказательство. Даны $X_1, X_2, \ldots, X_n$—независимые случайные величины с характеристическими функциями

$$g_{x1}(t), g_{x2}(t), \ldots, g_{xn}(t)$$

и их сумма

$$Y = \sum_{k=1}^{n} X_k.$$

Требуется доказать, что

$$g_y(t) = \prod_{k=1}^{n} g_{xk}(t). \qquad (7.8)$$

Имеем

$$g_y(t) = M[e^{ity}] = M[e^{it\sum_{k=1}^{n} X_k}] = M[\prod_{k=1}^{n} e^{itX_k}]$$

Так как величины X_k независимы, то независимы и их функции e^{itXk}. По теореме умножения математических ожиданий получим:

$$g_y(t) = \prod_{k=1}^{n} M[e^{itXk}] = \prod_{k=1}^{n} g_{xk}(t).$$

что и требовалось доказать.

Аппарат характеристических функций часто применяется для композиции законов распределения. Пусть, например, имеются две независимые случайные величины X и Y с плотностями распределения $f_1(x)$ и $f_2(y)$. Требуется найти плотность распределения величины

$$Z=X+Y.$$

Это можно выполнить следующим образом: найти характеристические функции $g_x(t)$ и $g_y(t)$ случайных величин X и Y и, перемножив их, получить характеристическую функцию величины Z:

$$g_z(t) = g_x(t)\, g_y(t)$$

после чего, подвергнув $g_z(f)$ обратному преобразованию Фурье, найти плотность распределения величины Z:

1 ∞

$$f_3(z) = \frac{}{2\pi}\int_{-\infty} e^{-i\,t\,xz} g_z(t)dt.$$

Пример. Найти с помощью характеристических функций композицию двух нормальных законов:
$f_1(x)$ с характеристиками $m_x = 0$; σ_x;
$f_2(y)$ с характеристиками $m_y = 0$; σ_y;

Решение. Находим характеристическую функцию величины X. Для этого представим ее в виде

$X = \sigma_x U$,

где $m_u = 0$; $\sigma_u = 1$.

Пользуясь результатом предыдущего примера , найдем

$$g_u(t) = e^{-t^2/2}.$$

Согласно СВОЙСТВУ 1 характеристических функций,

$$g_x(t) = g_u(\sigma_x t) = e^{-(\sigma x\, t)^2/2}.$$

Аналогично

$$g_y(t) = e^{-(\sigma y\, t)^2/2}.$$

Перемножая $g_x(t)$ и $g_y(t)$ имеем:

$$g_z(t) = e^{-t^2(\sigma y^2 + \sigma x^2)/2},$$

а это есть характеристическая функция нормального закона с параметрами $m_z = 0$;
$\sigma_z = \sqrt{(\sigma_x^2 + \sigma_y^2)}$. Таким образом, получена композиция нормальных законов гораздо более простыми средствами.

8. Центральная предельная теорема для одинаково распределенных слагаемых.

Различные формы центральной предельной теоремы отличаются между собой условиями, накладываемыми на распределения образующих сумму случайных слагаемых. Здесь мы сформулируем и докажем одну из самых простых форм центральной предельной теоремы, относящуюся к случаю одинаково распределенных слагаемых.

Теорема. *Если $X_1, X_2, \dots\dots X_n$ —независимые случайнее величины, имеющие один и тот же закон распределения с математическим ожиданием m и дисперсией σ^2, то при неограниченном увеличении n закон, распределения суммы*

n

$$Y_n = \sum_{k=1} X_k. \qquad (8.1)$$

неограниченно приближается к нормальному.

Доказательство.

Проведем доказательство для случая непрерывных случайных величин $X_1, X_2, \dots\dots X_n$ (для прерывных оно будет аналогичным).

Согласно второму свойству характеристических функций, доказанному в предыдущем n°, характеристическая функция величины Y_n, представляет собой произведение характеристических функций слагаемых. Случайные величины $X_1, X_2, \dots\dots X_n$ имеют один и тот же закон распределения с плотностью $f(x)$ и, следовательно, одну и ту же характеристическую функцию

$$g(x)=\int_{-\infty}^{\infty} e^{itx} f(x)dx. \qquad (8.2)$$

Следовательно, характеристическая функция случайной величины Y_n будет $g_{yn}(t)=[g_x(t)]^n$ (8.3)

Исследуем более подробно функцию $g_x\{t)$. Представим ее в окрестности точки t=0 по формуле Маклорена с тремя членами:

$$g_x(t)=g_x(0)+ g'_x(0)t + [g''_x(0)/2 + \alpha(t)]t^2, \qquad (8.4)$$

где $\alpha \to 0$ при $t \to 0$.

Найдем величины $g_x(0)$, $g'_x(0)$, $g''_x(0)$. Полагая в формуле (8.2) t=0, имеем:

$$g_x(0) =\int_{-\infty}^{\infty} f (x) dx = 1. \qquad (8.5)$$

Продифференцируем (8.2) по t:

$$g_x(t)=\int_{-\infty}^{\infty} ixe^{itx} f(x)dx= i \int_{-\infty}^{\infty} x e^{i tx} f (x) dx. \qquad (8.6)$$

Полагая в (8.6) t=0, получим:

$$g'x(0)=i\int_{-\infty}^{\infty} x f(x)dx=iM[X] = im. \qquad (8.7)$$

Очевидно, не ограничивая общности, можно положить m=0 (для этого достаточно перенести начало отсчета в точку m). Тогда

$$g'x(0) = 0.$$

Продифференцируем (8.6) еще раз:

$$g''_x(t) = -\int_{-\infty}^{\infty} x^2 e^{i tx} f (x) dx,$$

отсюда

∞

$g''_x(0) = -\int x^2 f(x)\, dx.$ (8.8)

$-\infty$

При m=0 интеграл в выражении (8.8) есть не что иное. как дисперсия величины X с плотностью *f(x),* следовательно

$g''_x(0) = -\sigma^2$. (8.9)

Подставляя в (8.4) $g_x(0)=1$, $g'_x(0)=0$ и $g''_x(0)= -\sigma^2$, получим:

$g_x(t)=1-[\sigma^2/2 - \alpha(t)]t^2.$ (8.10)

Обратимся к случайной величине Y_n. Мы хотим доказать, что ее закон распределения при увеличении п приближается к нормальному. Для этого перейдем от величины Yn к другой (“нормированной”) случайной величине

$$Z_n= Y_n / (\sigma \sqrt{n}). \qquad (8.11)$$

Эта величина удобна тем, что ее дисперсия не зависит от n и равна единице при любом n. В этом нетрудно убедиться, рассматривая величину Zn как линейную функцию независимых случайных величин $X_1, X_2, \ldots\ldots X_n$, каждая из которых имеет дисперсию σ^2. Если мы докажем, что закон распределения величины Zn, приближается к нормальному, то, очевидно, это будет справедливо и для величины Y_n, связанной с Zn линейной зависимостью (8.11).

Вместо того чтобы доказывать, что закон распределения величины Zn при увеличении n приближается к нормальному, покажем, что ее характеристическая функция приближается к характеристической функции нормального закона[5].

Найдем характеристическую функцию величины Zn. Из соотношения (8.11), согласно первому свойству характеристических функций (7.8) получим

$$g_{zn}(t)=g_{yn} (t / (\sigma \sqrt{n})) \qquad (8.12)$$

где $g_{yn}(t)$ —характеристическая функция случайной величины Yn. Из формул (8.12) и (8.3) получим

$$g_{zn}(t) = [g_x (t / (\sigma \sqrt{n}))]^n \qquad (8.13)$$

или, пользуясь формулой (8.10),

$$g_{zn}(t)=\{ 1- [\alpha^2/2 - \alpha(t/(\sigma \sqrt{n}))]\, t^2 / n\sigma^2 \}^n. \qquad (8.14)$$

Прологарифмируем выражение (8.14):

$$\ln g_{zn} (t)= n \ln \{ 1 - [\alpha^2/2 - \alpha(t/(\sigma \sqrt{n}))]\, t^2 / n\sigma^2 \}.$$

[5] Здесь мы принимаем без доказательства, что из сходимости характеристических функций следует сходимость законов распределения. Доказательство см., например, Б.В.Гнеденко, Курс теории вероятностей, 1961.

Введем обозначение

$$[\alpha^2/2 - \alpha(t/(\sigma \sqrt{n}))]\, t^2 / n\sigma^2 = x. \qquad (8.15)$$

Тогда

$$\ln g_{zn}(t) = n \ln\{1-x\}. \qquad (8.16)$$

Будем неограниченно увеличивать n. При этом величина x, согласно формуле (8.15), стремится к нулю. При значительном n ее можно считать весьма малой. Разложим 1n{1-x) в ряд и ограничимся одним членом разложения (остальные при $n \to \infty$ станут пренебрежимо малыми):

$$1n\{1-x) \approx -x.$$

Тогда получим

$$\lim_{n\to\infty} \ln g_{zn}(t) = \lim_{n\to\infty} n(-x) = \lim_{n\to\infty} [-t^2/2 - \alpha(t/(\sigma \sqrt{n}))\, t^2 / n\sigma^2] =$$

$$= -t^2/2 + \lim_{n\to\infty} t^2/2\, \alpha(t/(\sigma \sqrt{n})).$$

По определению функция $\alpha(t)$ стремится к нулю при $t \to 0$; следовательно,

$$\mathrm{Lim}_{n\to\infty}\, \alpha(t/(\sigma \sqrt{n})) = 0$$

и

$$\lim \ln g_{zn}(t) = -t^2/2,$$

откуда

$$\lim g_{zn}(t) = e^{-t^2/2}. \qquad (8.17)$$

Это есть не что иное, как характеристическая функция нормального закона с параметрами m=0, σ =1(см. пример1, n^0 7).

Таким образом, доказано, что при увеличении n характеристическая функция' случайной величины Zn неограниченно приближается к характеристической функции нормального закона; отсюда заключаем, что и закон распределения величины Zn (а значит и величины Yn) неограниченно приближается к нормальному закону. Теорема доказана.

Мы доказали центральную предельную теорему для частного, но важного случая одинаково распределенных слагаемых. Однако в достаточно широком классе условий она справедлива и для неодинаково распределенных слагаемых. Например, А. М. Ляпунов доказал центральную предельную теорему для следующих условий:

$$\mathrm{Lim} \sum^{n} b_k \,/ \left(\sum^{n} D_k\right)^{3/2} = 0, \qquad (8.18)$$

$n \to \infty \quad k=1 \qquad k=1$

где b_k—третий абсолютный центральный момент величины Xk:

$b_k = \nu_3[X_k] = M[\,|X_k|^3]$ $\qquad (k = 1,\ldots,n),$

Dk - дисперсия величины X_k.

Наиболее общим (необходимым и достаточным) условием справедливости центральной предельной теоремы является условие Линдеберга: *при любом* $\tau > 0$

$$\lim_{n\to\infty} 1/B_n^2 \sum_{k=1}^{n} \int_{|x-m_k| > \tau B_n} (x-m_k)^2 f_k(x)dx = 0,$$

где m_k *— математическое ожидание,* $f_k(x)$ *— плотность распределения случайной величины* X_k*,* $B_n = \sqrt{\sum D_k}$*,* $\qquad (k = 1,\ldots,n)$.

9. Формулы, выражающие центральную предельную теорему и встречающиеся при ее практическом применении.

Согласно центральной предельной теореме, закон распределения суммы достаточно большого числа независимых случайных величин (при соблюдении некоторых нежестких ограничений) сколь угодно близок к нормальному.

Практически центральной предельной теоремой можно пользоваться и тогда, когда речь идет отлично сумме сравнительно небольшого числа случайных величин. При суммировании независимых случайных величин, сравнимых по своему рассеиванию, с увеличением числа слагаемых закон распределения суммы очень скоро становится приблизительно нормальным. На практике вообще широко применяется приближенная замена одних законов распределения другими; при той сравнительно малой точности, которая требуется от вероятностных расчетов, такая замена тоже может быть сделана крайне приближенно. Опыт показывает, что когда число слагаемых порядка десяти (а часто и меньше), закон распределения суммы обычно может быть заменен нормальным.

В практических задачах часто применяют центральную предельную теорему для вычисления вероятности того, что сумма нескольких случайных величин окажется в заданных пределах.

Пусть $X_1, X_2,\ldots\ldots X_n$ - независимые случайные величины с математическими ожиданиями

$$m_1, m_2,\ldots\ldots m_n$$

и дисперсиями

$$D_1, D_2,\ldots\ldots D_n \; .$$

Предположим, что условия предельной центральной теоремы выполнены (величины $X_1, X_2,\ldots\ldots X_n$ сравнимы по порядку своего влияния на рассеивание суммы) и число слагаемых n достаточно для того, чтобы закон распределения величины $\qquad n$

$$Y = \sum_{i=1} X_i. \qquad (9.1)$$

можно было считать приближенно нормальным.

Тогда вероятность того, что случайная величина Y попадет в пределы участка (α,β), выфражаетя формулой

$$P(\alpha<Y<\beta)= \Phi^*\left(\frac{\beta-m_y}{\sigma_y}\right) - \Phi^*\left(\frac{\alpha-m_y}{\sigma_y}\right), \qquad (9.2)$$

где m_y, σ_y - математическое ождание и среднее квадратическое отклонение величины Y, Ф* - нормальная функция распределения.

Согласно теоремам сложения математических ожиданий и дисперсий

$$m_Y = \sum_{i=1}^{n} m_i. \qquad (9.3)$$

$$\sigma_y = \sqrt{D_y} = \sqrt{\sum_{i=1}^{n} D_i}$$

Таким образом, для того, чтобы приближенно найти вероятность попадания суммы большого числа случайных величин на заданный участок,не требуется знать законы распределения этих величин; достаточно знать лишь их характеристики. Разумеется, это относится только к случаю, когда ыполнено основное условие центральной предельной теоремы - равномерно малое влияние слагаемых на рассеивание суммы.

Кроме формул типа (9.2), на практике часто применяются формулы, в которых вместо суммы случайных величин X_i фигурирует их нормированная сумма

$$Z = \frac{Y^0}{\sigma_y} = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} m_i}{\sum_{i=1}^{n} D_i} \qquad (9.4)$$

Очевидно,

$$M[Z]=0;\ D[Z]=\sigma_z = 1.$$

Если закон распределения величины Y близок к нормальному с

параметрами (9.3), то закон распределения величины Z близок к нормальному с параметрами $m_z=0$,
$\sigma_z = 1$. Отсюда

$$P(\alpha<Z<\beta)= \Phi^*(\beta) - \Phi^*(\alpha) . \qquad (9.5)$$

Заметим, что центральная предельная теорема может применяться не только к непрерывным, но и к дискретным случайным величинам при условии, что мы будем оперировать не плотностями, а функциями распределения. Действительно, если величины X_1, X_2,…… X_n дискретны, то их сумма X -также дискретная случайная величина и поэтому, строго говоря, не может подчиняться нормальному закону. Однако все формулы типа (9.2), (9.5) остаются в силе, т. к. в них фигурируют не плотности, а функции распределения. Можно доказать, что если дискретные случайные величины удовлетворяют условиям центральной предельной теоремы, то функция распределения их нормированной суммы Z (см. формулу (9.4)) при увеличении n неограниченно приближается к нормальной функции распределения с параметрами $m_z=0$,
$\sigma_z = 1$.

Частным случаем центральной предельной теоремы для дискретных случайных величин является теорема Лапласа.

Если производится n независимых опытов, в каждом из которых, событие A появляется с вероятностью p, то справедливо соотношение

$$P\left(\alpha<\frac{Y- np}{\sqrt{npq}}<\beta\right)= \Phi^*(\beta) - \Phi^*(\alpha) , \qquad (9.6)$$

где Y - число появлений события A в n опытах, q= 1-p.

Доказательство. Пусть производится n независимых опытов, в каждом из которых с вероятностью p может появиться событие A. Представим случайную величинуY - общее число появлений события в n опытах - в виде суммы

$$Y = \sum_{i=1}^{n} X_i. \qquad (9.7)$$

где $\mathbf{X_i}$ - число появлений события A в i-м опыте. Согласно доказанной в n^0 8 теореме, закон распределения суммы одинаково распределенных слагаемых при увеличении их числа приближается к нормальному закону. Следовательно при достаточно большом n справедлива формула (9.5), где

$$\frac{Y-m_y}{\sigma_y}$$

$$Z = \overline{\qquad} \tag{9.8}$$

Было доказано, что математическое ожидание и дисперсия числа появлений события в n независимых опытах равны: m_y=np; D_y = npq (q= 1-p).
Подставляя эти выражения в (9.8), получим

$$Z = \frac{Y - np}{\sqrt{npq}} \quad ,$$

и формула (9.5) примет вид:

$$P(\alpha < \frac{Y - np}{\sqrt{npq}} < \beta) = \Phi^*(\beta) - \Phi^*(\alpha) .$$

Теорема доказана.