2.10 Предельные теоремы теории вероятностей

2.10.1 Закон больших чисел

Нельзя заранее предвидеть, какое из возможных значений примет СВ в результате испытания.

Однако, при некоторых, сравнительно широких условиях, при очень большом числе испытаний средний их результат практически перестает быть случайным и может быть предсказан с большой степенью определенности.

Эта «устойчивость средних» – физическое содержание ***закона больших чисел*** (ЗБЧ), понимаемого в широком смысле.

В узком смысле под ЗБЧ понимается ряд теорем, в каждой из которых для тех или иных условий устанавливается факт приближения средних характеристик большого числа испытаний к определенным постоянным величинам.

Все эти теоремы опираются на *неравенство Чебышева*.

Неравенство Чебышева

Лемма (*неравенство Чебышева*).

Пусть имеется СВ X с математическим ожиданием m_x и дисперсией D_x.

Каково бы ни было $\varepsilon > 0$, вероятность того, что СВ X отклонится от своего математического ожидания не менее, чем на ε, ограничена сверху величиной $\frac{D_x}{\varepsilon^2}$:

$$P(|X - m_x| \geq \varepsilon) \leq \frac{D_x}{\varepsilon^2}.$$

Неравенство Чебышева

Замечание.

Неравенство Чебышева может быть записано и в другой форме:

$$P(|X - m_x| < \varepsilon) \geq 1 - \frac{D_x}{\varepsilon^2}.$$

Неравенство Чебышева справедливо и для дискретных, и для непрерывных СВ.

Неравенство Чебышева

Пример.

С помощью неравенства Чебышева можно показать, что для любой СВ вероятность невыполнения правила «трех сигма» не превышает 1/9:

$$P(|X - m_x| \geq 3\sigma) \leq \frac{D_x}{(3\sigma)^2} = \frac{1}{9}.$$

В действительности для большинства СВ, встречающихся на практике, ошибка правила «трех сигма» существенно меньше

Сходимость по вероятности

Пусть имеется последовательность СВ

$$X_1, X_2, \ \ldots, X_n, \ldots$$

Говорят, что эта последовательность ***сходится по вероятности к неслучайной величине а***, если для любого $\varepsilon > 0$ и для любого $\delta > 0$ найдется номер N, такой что для всех $n > N$

$$P(|X_n - a| < \varepsilon) > 1 - \delta.$$

При неограниченном увеличении n вероятность события $|X_n - a| < \varepsilon$, где $\varepsilon > 0$ – произвольно малое фиксированное число, стремится к 1

Обозначение: $X_n \xrightarrow[p]{n\to\infty} a$ или $X_n \xrightarrow[n\to\infty]{p} a$

Теорема Чебышева (ЗБЧ)

Теорема Чебышева (называемая ЗБЧ).

Пусть имеется последовательность $X_1, X_2, \ldots, X_n, \ldots$ попарно независимых СВ с одним и тем же математическим ожиданием m, дисперсии которых равномерно ограничены (т. е. существует постоянная C, такая что $D_i \le C$, $i = 1, 2, \ldots, n, \ldots$).

Рассмотрим последовательность $Y_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ – среднее арифметическое $X_1, X_2, \ldots, X_n$ при разных n).

Теорема Чебышева (ЗБЧ)

Теорема Чебышева (продолжение).

Тогда

$$Y_n \xrightarrow[p]{n\to\infty} m,$$

т. е. для любого $\varepsilon > 0$ и для любого $\delta > 0$ найдется номер N, такой что для всех $n > N$

$$P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - m\right| < \varepsilon\right) > 1 - \delta.$$

Доказательство проводится с применением неравенства Чебышева

Теорема Чебышева (ЗБЧ)

Практическое значение теоремы Чебышева можно пояснить следующим примером.

Пусть требуется измерить некоторую физическую величину, истинное значение которой равно m. Результат измерения будет СВ X, математическое ожидание которой равно m, а дисперсия определяется точностью измерительного прибора.

Предположим: произведено n измерений, в результате которых получены значения $X_1, X_2, \ldots, X_n$ (n «экземпляров» СВ X).

Теорема Чебышева (ЗБЧ)

Если измерения производились независимо (результат каждого измерения не зависит от результатов остальных), то СВ $X_1, X_2, \ldots, X_n$ будут независимы.

Если измерения производились без систематических (т. е. одного знака) ошибок, то СВ $X_1, X_2, \ldots, X_n$ имеют одно и то же математическое ожидание m.

Если прибор обеспечивает определенную точность измерений, то дисперсии СВ $X_1, X_2, \ldots, X_n$ равны, а значит, равномерно ограничены.

При выполнении трех указанных условий при достаточно большом n среднее арифметическое результатов измерений перестает быть случайным, приближаясь к m.

Теорема Чебышева (ЗБЧ)

Этим обосновывается рекомендуемый в практической деятельности способ получения более точных результатов измерений:

одна и та же величина измеряется многократно, и среднее арифметическое полученных результатов принимается в качестве искомого значения.

На теореме Чебышева основан широко применяемый в статистике выборочный метод, суть которого состоит в том, что по сравнительно небольшой случайной выборке судят о всей совокупности исследуемых объектов.

Обобщенная теорема Чебышева

Теорема (обобщенная теорема Чебышева).

Пусть имеется последовательность $X_1, X_2, \ldots, X_n, \ldots$ попарно независимых СВ с различными, в общем случае, математическими ожиданиями

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

и дисперсиями $D_1, D_2, \ldots, D_n, \ldots$

Если дисперсии равномерно ограничены (существует постоянная C, такая что $D_i \le C$, $i = 1, 2, \ldots, n, \ldots$),

а $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, то

$$Y_n - \frac{1}{n}\sum_{i=1}^{n} m_i \xrightarrow[p]{n\to\infty} 0,$$

Обобщенная теорема Чебышева

Теорема (обобщенная теорема Чебышева) – продолжение.

т. е. для любого $\varepsilon > 0$ и для любого $\delta > 0$ найдется номер N, такой что для всех $n > N$

$$P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \frac{m_1 + m_2 + \cdots + m_n}{n}\right| < \varepsilon\right) > 1 - \delta .$$

Таким образом:

хотя отдельные независимые СВ могут принимать значения, далекие от своих математических ожиданий, среднее арифметическое достаточно большого числа СВ с высокой вероятностью принимает значения, близкие к вполне определенному постоянному числу.

Обобщение ЗБЧ на случай зависимых СВ (теорема Маркова) принадлежит А.А. Маркову – за пределами настоящего курса.

Следствия из ЗБЧ

Теорема Бернулли.

Пусть производится n независимых опытов, в каждом из которых вероятность появления некоторого события A постоянна и равна p.

Тогда относительная частота появления события A при неограниченном увеличении n сходится по вероятности к p.

Это свойство уже упоминалось ранее (без обоснования) при определении статистической вероятности

Следствия из ЗБЧ

Теорема Пуассона.

Пусть производится n независимых опытов, причем вероятность появления события A в i-м опыте равна p_i, $i = 1, 2, \ldots, n$.

Тогда при неограниченном увеличении n разность между относительной частотой события A и величиной $\frac{1}{n}\sum_{i=1}^{n} p_i$ сходится по вероятности к нулю.

2.10.2 Центральная предельная теорема

Возможности предсказания результатов массовых случайных явлений, которые обеспечивает применение ЗБЧ, могут быть еще более расширены за счет другой группы предельных теорем.

Эта группа касается уже не предельных значений СВ, а *предельных законов распределения*.

Данная группа теорем известна под названием ***центральной предельной теоремы*** (ЦПТ).

Как и ЗБЧ, ЦПТ имеет ряд форм

ЦПТ в различных формах устанавливает условия, при которых возникает нормальное распределение.

Такие условия часто возникают на практике, что и объясняет широкую распространенность нормального закона в случайных явлениях.

Ранее уже говорилось: нормальное распределение возникает тогда, когда суммируется большое число независимых (или слабо зависимых) СВ $X_1, X_2, \ldots, X_n$, сравнимых по порядку своего влияния на рассеивание суммы.

Различные формы ЦПТ отличаются друг от друга условиями, накладываемыми на распределения случайных слагаемых $X_1, X_2, \ldots, X_n$.

Чем жестче эти условия, тем легче доказывается утверждение.

Одна из самых простых форм ЦПТ – ЦПТ для одинаково распределенных слагаемых.

ЦПТ для одинаково распределенных слагаемых

Теорема (ЦПТ для одинаково распределенных слагаемых).

Пусть $X_1, X_2, \ldots, X_n$ – независимые СВ, имеющие одно и то же распределение с математическим ожиданием m и дисперсией σ^2.

Тогда СВ

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

имеет асимптотически нормальное распределение с параметрами $m_Y = n \cdot m$ и $\sigma_Y = \sigma\sqrt{n}$.

ЦПТ для одинаково распределенных слагаемых

Замечание 1.

Закон распределения СВ $X_1, X_2, \ldots, X_n$ может быть любым (непрерывным или дискретным).

При этом СВ

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

при достаточно большом n будет иметь распределение, близкое к нормальному.

ЦПТ для одинаково распределенных слагаемых

Замечание 2.

Заключение теоремы может быть сформулировано иначе:

СВ

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$$

имеет асимптотически нормальное распределение с параметрами $m_{\bar{X}} = m$ и $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$.

ЦПТ для одинаково распределенных слагаемых

Пример 1.

СВ $X_1, X_2, \ldots, X_n$ имеют показательный закон распределения с параметром $\lambda = 2$.

Рассмотрим СВ

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$$

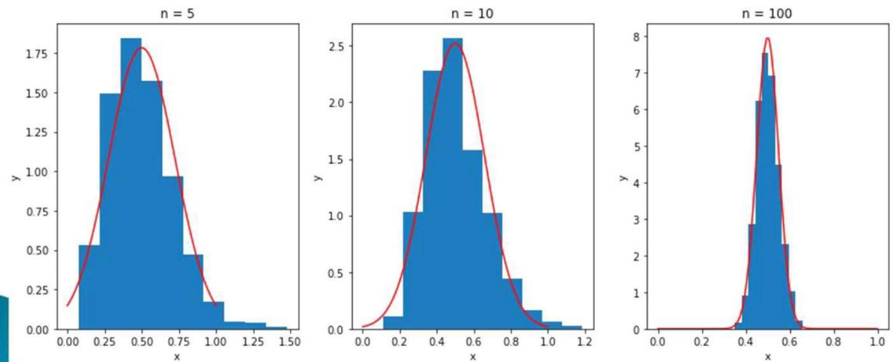
при $n = 5$, $n = 10$, и $n = 100$.

С помощью генератора псевдослучайных чисел сгенерировано по 1000 значений СВ $\bar{X}$ для каждого значения n.

ЦПТ для одинаково распределенных слагаемых

Пример 1 (продолжение).

Гистограммы полученных распределений и кривая плотности предельного нормального распределения:



ЦПТ для одинаково распределенных слагаемых

Пример 2 (машина Гальтона).

Машина Гальтона (*доска Гальтона*, *Galton board*) – первый экземпляр сконструирован в 1873 г. для демонстрации закономерности, описываемой ЦПТ.

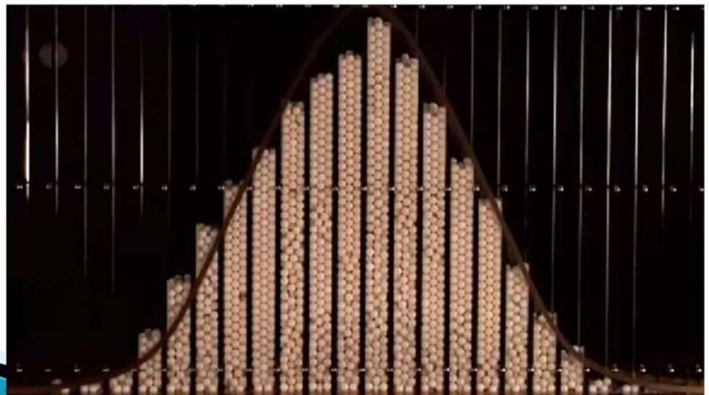
Представляет собой ящик, в заднюю стенку которого вбиты штыри в шахматном порядке. Сверху через воронку, расположенную ровно посередине между боковыми стенками, в ящик засыпается большое количество шариков.

При столкновении со штырьками шарики с одинаковой вероятностью могут падать справа или слева от каждого штырька.

ЦПТ для одинаково распределенных слагаемых

Пример 2 (машина Гальтона – продолжение).

Упражнение:
попытаться построить математическую модель работы машины Гальтона.



ЦПТ для одинаково распределенных слагаемых

Интегральная теорема Лапласа, рассмотренная выше, фактически, является частным случаем ЦПТ для одинаково распределенных слагаемых, где СВ X_i – число наступлений рассматриваемого события в i–м испытании, $i = 1, 2, \ldots, n$ – имеют распределение Бернулли с параметром p;

СВ

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

число наступлений события в n испытаниях, согласно ЦПТ, имеет асимптотически нормальное распределение с параметрами $m = np$ и $\sigma = \sqrt{npq}$.