

Урок 14. Закон больших чисел (ЗБЧ), центральная предельная теорема (ЦПТ) и p -значение

Хакимов Р.И. + ChatGPT

Закон больших чисел (ЗБЧ)

Закон больших чисел – это фундаментальная теорема в теории вероятностей и статистике, которая описывает поведение выборочных средних при большом числе наблюдений. В общих чертах он утверждает, что по мере увеличения числа независимых и одинаково распределённых случайных величин их среднее значение стремится к математическому ожиданию этой величины.

Есть две основные формы закона больших чисел: **слабый** и **сильный**.

Закон больших чисел (ЗБЧ)

Слабый закон больших чисел утверждает, что выборочное среднее стремится к математическому ожиданию случайной величины с вероятностью, стремящейся к единице при увеличении числа наблюдений. Это можно записать так:

$$P\left(\left|\frac{1}{n}\sum_{i=1}^n X_i - \mathbb{E}(X)\right| \geq \varepsilon\right) \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty$$

Здесь X_1, X_2, \dots, X_n — независимые одинаково распределённые случайные величины с конечным математическим ожиданием $\mathbb{E}(X)$, а ε — произвольно малое положительное число.

Закон больших чисел (ЗБЧ)

Сильный закон больших чисел утверждает, что при достаточно большом числе повторений случайного эксперимента среднее арифметическое его результатов с вероятностью, стремящейся к единице, будет приближаться к математическому ожиданию (ожидаемому среднему) случайной величины. Это более строгая форма, чем слабый закон, так как утверждает сходимость для почти всех реализаций случайной величины:

$$P\left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = \mathbb{E}(X)\right) = 1$$

Закон больших чисел (ЗБЧ)

Пример 1: Подбрасывание монеты

Предположим, что у нас есть честная монета, и вероятность выпадения орла равна $p = 0.5$. При большом количестве подбрасываний выборочная доля орлов должна стремиться к 0.5.

Процесс:

1. Вы подбрасываете монету 10 раз и получаете 6 орлов и 4 решки. Отношение орлов: $\frac{6}{10} = 0.6$.
2. Вы подбрасываете монету 100 раз и получаете 52 орла и 48 решек. Отношение орлов: $\frac{52}{100} = 0.52$.
3. При 1000 подбрасываниях вы получаете 503 орла, и отношение орлов $\frac{503}{1000} = 0.503$.

Закон больших чисел (ЗБЧ)

Пример 2: Средний рост в классе

Предположим, вы хотите узнать средний рост учеников в школе. Закон больших чисел утверждает, что если вы выберете достаточно большое количество учеников (независимых случайных наблюдений), их средний рост будет приближаться к истинному среднему росту всех учеников в школе.

Процесс:

1. Вы выбираете случайные 5 человек и получаете средний рост 165 см.
2. Выбираете 50 человек, и средний рост уже ближе к 170 см.
3. Наконец, при выборке из 500 человек средний рост приближается к 171 см, что является реальным средним для всей школы.

Закон больших чисел (ЗБЧ)

Пример 3: Казино

Казино основаны на идее, что в долгосрочной перспективе выигрышные и проигрышные ставки игроков подчиняются закону больших чисел.

Например, если вероятность выигрыша в рулетке — 47.37%, то в краткосрочной перспективе игрок может выиграть или проиграть значительно больше или меньше этой вероятности. Но по мере увеличения числа игр средний процент выигрышей будет стремиться к 47.37%, что обеспечивает казино стабильный доход в долгосрочной перспективе.

Процесс:

1. Игрок играет 10 раз и выигрывает 8 раз. Выигрыши составляют 80%.
2. Игрок играет 100 раз и выигрывает 52 раза. Выигрыши уже составляют 52%.
3. При 1000 играх выигрыши составляют 473 раза — это уже почти 47.37%.

Закон больших чисел (ЗБЧ)

Пример 4: Производственный процесс

Представим, что завод производит детали, и на каждую 1000 деталей в среднем приходится 10 дефектных. Однако при производстве малой партии (например, 100 деталей) вы можете получить либо ни одной, либо несколько дефектных деталей. Если вы производите 100 000 деталей, число дефектных будет близким к среднему (около 1%).

Процесс:

1. В небольшой партии из 100 деталей 0 дефектных.
2. В партии из 1000 деталей 9 дефектных (0.9%).
3. В партии из 100 000 деталей число дефектных 1002 (около 1%).

Закон больших чисел (ЗБЧ)

Пример 5: Оценка среднего веса фруктов

В магазине продаются яблоки, и известно, что их средний вес — 150 граммов.

Процесс:

1. 10 яблок: Средний вес может составить 148 граммов.
2. 100 яблок: Средний вес может составить 149.5 граммов.
3. 1000 яблок: Средний вес приблизится к 150 граммам.

Чем большее количество яблок взято для измерения, тем ближе средний вес к истинному значению 150 граммов.

Центральная предельная теорема (ЦПТ)

Центральная предельная теорема (ЦПТ) — одно из важнейших утверждений в теории вероятностей и математической статистике. Она объясняет, почему многие распределения данных в природе и на практике имеют форму, близкую к нормальному распределению (колоколообразной кривой), даже если исходные данные не являются нормально распределёнными.

Центральная предельная теорема (ЦПТ)

Суть теоремы:

Центральная предельная теорема утверждает, что если у нас есть достаточно большое количество независимых случайных величин, имеющих одинаковое распределение с конечной дисперсией, то сумма (или среднее) этих величин будет приближенно иметь нормальное распределение. Это справедливо независимо от того, как распределены исходные величины.

Центральная предельная теорема (ЦПТ)

Формальное утверждение (вариант 1):

Пусть X_1, X_2, \dots, X_n — независимые одинаково распределённые случайные величины с математическим ожиданием μ и дисперсией σ^2 . Тогда по мере роста числа n , стандартная сумма этих случайных величин:

$$S_n = \frac{\sum_{i=1}^n X_i - n\mu}{\sigma\sqrt{n}}$$

будет стремиться к стандартному нормальному распределению $N(0, 1)$ при $n \rightarrow \infty$.

Центральная предельная теорема (ЦПТ)

Формальное утверждение (вариант 2):

Пусть X_1, X_2, \dots, X_n — это независимые одинаково распределённые случайные величины с математическим ожиданием $E[X_i] = \mu$ и дисперсией $Var[X_i] = \sigma^2$. Тогда для суммы этих случайных величин:

$$S_n = X_1 + X_2 + \dots + X_n$$

стандартная нормировка этой суммы:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}}$$

будет приближаться к стандартному нормальному распределению, когда $n \rightarrow \infty$. Это можно записать так:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1), \quad \text{при } n \rightarrow \infty,$$

где \xrightarrow{d} обозначает сходимость по распределению, а $N(0, 1)$ — это стандартное нормальное распределение.

Центральная предельная теорема (ЦПТ)

Пример с подбрасыванием монеты

Пусть мы подбрасываем монету n раз. Подброс монеты можно моделировать как случайную величину с двумя исходами: 'орёл' (1) и 'решка' (0). Математическое ожидание для каждой подбрасываемой монеты равно $E[X] = 0.5$, а дисперсия $Var[X] = 0.25$.

Если мы будем подбрасывать монету много раз (большое n), то распределение среднего количества 'орлов' будет приближаться к нормальному распределению, даже несмотря на то, что исходное распределение для каждого отдельного подбрасывания — это не нормальное, а биномиальное.

Центральная предельная теорема (ЦПТ)

Пример с измерениями роста

Допустим, мы измеряем рост группы людей. Каждый рост можно рассматривать как случайную величину. Рост людей в реальной жизни может не подчиняться строго нормальному распределению. Однако если мы возьмем достаточно большую выборку людей, средний рост людей в этой выборке будет приближаться к нормальному распределению в силу центральной предельной теоремы.

Центральная предельная теорема (ЦПТ)

Пример с играми в казино

В казино при игре в рулетку каждый выигрыш или проигрыш может рассматриваться как случайная величина, которая имеет своё специфическое распределение (например, распределение Бернулли или биномиальное). Если игрок играет много раз, то итоговая сумма выигрышей и проигрышей, нормированная на число игр, будет распределена нормально, даже если отдельные исходы игр имеют совершенно другие распределения.

Центральная предельная теорема (ЦПТ)

Пример с бросками кубика

Если мы бросаем шестигранный кубик, результаты каждого броска равномерно распределены на числах от 1 до 6. Однако, если мы будем бросать кубик очень много раз и считать среднее значение всех бросков, это среднее будет приближаться к нормальному распределению по ЦПТ, даже несмотря на то, что распределение одного броска кубика явно не является нормальным.

Важность ЦПТ

Статистические выводы: ЦПТ объясняет, почему многие статистические процедуры (например, построение доверительных интервалов или проверка гипотез) основаны на предположении нормальности выборки. Даже если данные не нормальны, их средние будут приближаться к нормальному распределению при увеличении размера выборки.

Реальные приложения: Это объясняет многие явления в реальной жизни, когда суммирование различных случайных факторов приводит к результатам, которые распределены близко к нормальному распределению.

Ограничения ЦПТ

ЦПТ работает хорошо для больших выборок (большое n), но для малых выборок результат может существенно отличаться от нормального распределения.

Если исходные данные сильно асимметричны или имеют 'тяжёлые хвосты' (например, распределение Парето), требуется намного больше наблюдений, чтобы сумма начала приближаться к нормальному распределению.

P-value (уровень значимости)

P-value (уровень значимости) — это статистическая величина, которая показывает вероятность того, что наблюдаемый результат или более экстремальный результат мог бы быть получен случайно, при условии, что нулевая гипотеза верна.

Основные моменты:

Нулевая гипотеза обычно предполагает, что никакого эффекта или связи не существует (например, что две переменные не связаны или что лекарство не оказывает влияния).

P-value помогает определить, есть ли статистически значимые доказательства против нулевой гипотезы.

P-value (уровень значимости)

Интерпретация:

Малое p -value (например, меньше 0.05) говорит о том, что маловероятно получить такой результат случайно, и у исследователя есть основания отвергнуть нулевую гипотезу в пользу альтернативной.

Высокое p -value (например, больше 0.05) означает, что данные не дают достаточных доказательств для отклонения нулевой гипотезы, и она остается допустимой.

Важно помнить, что p -value не показывает силу эффекта или важность результата, оно лишь говорит о том, насколько результат является случайным.

P-value. Пример 1

Предположим, что мы проводим исследование, чтобы проверить гипотезу о среднем росте мужчин в каком-то городе. Мы знаем, что средний рост мужчин в общем населении равен $\mu_0 = 175$ см (это нулевая гипотеза).

Мы хотим проверить, отличается ли средний рост мужчин в этом городе.

Шаги:

1. Нулевая гипотеза (H_0):

Средний рост мужчин в городе равен 175 см ($\mu = 175$).

2. Альтернативная гипотеза (H_1):

Средний рост мужчин в городе не равен 175 см ($\mu \neq 175$).

3. Сбор данных:

Пусть мы взяли случайную выборку из 50 мужчин и измерили их средний рост. Получили среднее значение $\bar{x} = 177$ см и стандартное отклонение $s = 5$ см.

P-value. Пример 1

4. Рассчитаем тестовую статистику (z-статистику или t-статистику):

Так как выборка достаточно большая (50 человек), можно использовать z-статистику.

Формула для неё:

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

где: \bar{x} — среднее выборки,

μ_0 — среднее по нулевой гипотезе,

s — стандартное отклонение выборки,

n — размер выборки.

Подставляем данные:

$$z = \frac{177 - 175}{\frac{5}{\sqrt{50}}} = \frac{2}{\frac{5}{7.07}} = \frac{2}{0.707} \approx 2.83$$

P-value. Пример 1

5. Рассчитываем P-value:

Используя таблицу стандартного нормального распределения для z-значения $z = 2.83$, находим вероятность P-value для двухстороннего теста.

Для $z = 2.83$ вероятность попасть в такой или более экстремальный результат с обеих сторон примерно равна 0.0046.

6. Интерпретация:

Если уровень значимости $\alpha = 0.05$, то $P\text{-value} = 0.0046 < 0.05$. Это значит, что мы отвергаем нулевую гипотезу и считаем, что средний рост мужчин в этом городе статистически отличается от 175 см.

Вывод:

При $P\text{-value} = 0.0046$ мы можем сделать вывод, что средний рост мужчин в городе, скорее всего, не равен 175 см.

P-value. Пример 2

Пусть исследуется, влияет ли новый препарат на снижение уровня сахара в крови.

Нулевая гипотеза (H_0): препарат не влияет на уровень сахара (среднее изменение $= 0$).

Альтернативная гипотеза (H_1): препарат снижает уровень сахара (среднее изменение ≤ 0).

Предположим, исследование показало, что среднее снижение уровня сахара в контрольной группе равно 5 мг/дл, а стандартное отклонение равно 2 мг/дл. В выборке из 30 человек среднее снижение уровня сахара оказалось 6 мг/дл. Теперь проведем t-тест для проверки гипотезы.

P-value. Пример 2

Шаг 1: Построение статистики t:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

где:

$\bar{x} = 6$ — среднее снижение уровня сахара в выборке,

$\mu = 5$ — среднее снижение в контрольной группе (нулевая гипотеза),

$s = 2$ — стандартное отклонение,

$n = 30$ — размер выборки.

$$t = \frac{6 - 5}{2/\sqrt{30}} = \frac{1}{2/\sqrt{30}} \approx 2.74$$

Шаг 2: Определение p-value:

Используя таблицу распределения Стьюдента для t-распределения с 29 степенями свободы, находим p-value. По таблице для $t = 2.74$ значение p примерно равно 0.005.

P-value. Пример 2

Шаг 3: Интерпретация:

Если уровень значимости $\alpha = 0.05$, то p-value (0.005) меньше α . Это означает, что мы отвергаем нулевую гипотезу и принимаем альтернативную, что препарат действительно влияет на снижение уровня сахара.

Заключение: Новый препарат статистически значимо снижает уровень сахара в крови.

P-value. Пример 3

Предположим, мы хотим проверить гипотезу о среднем весе фруктов в корзине. Мы подозреваем, что средний вес фруктов отличается от 100 граммов, и собираем выборку для проверки.

Шаг 1: Формулировка гипотез

Нулевая гипотеза (H_0): средний вес фруктов = 100 граммов.

Альтернативная гипотеза (H_1): средний вес фруктов \neq 100 граммов.

Шаг 2: Сбор данных

Выбираем 30 фруктов, и их средний вес составляет 102 грамма с стандартным отклонением 5 граммов.

P-value. Пример 3

Шаг 3: Вычисление тестовой статистики

Используем t-тест для выборки:

- Размер выборки $n = 30$,
- Среднее выборки $\bar{x} = 102$,
- Предполагаемое среднее $\mu_0 = 100$,
- Стандартное отклонение $s = 5$.

Тестовая статистика t :

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{102 - 100}{\frac{5}{\sqrt{30}}} = \frac{2}{\frac{5}{5.477}} \approx 2.19$$

P-value. Пример 3

Шаг 4: Поиск p-value

Для тестовой статистики $t \approx 2.19$ и $n - 1 = 29$ степеней свободы, мы находим p-value, используя таблицы t-распределения или статистический софт. Допустим, $p\text{-value} \approx 0.037$.

Шаг 5: Интерпретация

Предположим, что уровень значимости $\alpha = 0.05$. Так как $p\text{-value} = 0.037 < 0.05$, мы отвергаем нулевую гипотезу. Это означает, что у нас есть статистически значимые основания считать, что средний вес фруктов отличается от 100 граммов.