§11. Сходимость случайных величин и функций распределений

Лекция 17

Для того, чтобы продвинуться дальше в построении новых вероятностных моделей и возможно в большей степени дать математическое основание для применения методов математической статистики к идентификации вероятностных моделей, мы должны изучить проблемы сходимости последовательностей случайных величин и их функций распределения.

Пусть на вероятностном пространстве $(\Omega, \mathcal{A}, P)$ заданы две случайные величины $X = X(\omega)$ и $Y = Y(\omega)$, имеющих одно и то же распределение вероятностей. В таком случае, с точки зрения приложений теории вероятностей такие случайные величины следует считать эквивалентными (неразличимыми).

Определение 11.1. Если $P(X(\omega) = Y(\omega)) = 1$, то случайные величины X и Y называются *равными почти наверное* или *эквивалентными* и пишется $X \underset{п.н.}{=} Y$.

При исследовании предела последовательности случайных величин $\{X_n, n \geq 1\}$, заданных на одном и том же вероятностном пространстве $(\Omega, \mathcal{A}, P)$, мы имеем дело, по существу, с проблемой сходимости последовательности функций $\{X_n(\omega), n \geq 1\}$, но при этом мы можем не обращать внимания на множество точек ω нулевой вероятности, в которых соответствующие числовые последовательности не имеют предела. Поэтому поточечная (в каждой точке $\omega \in \Omega$)) сходимость функций претерпевает существенное изменение и превращается в следующее

Определение 11.2. Если для последовательности случайных величин $\{X_n(\omega), n \geq 1\}$, заданных на вероятностном пространстве $(\Omega, \mathcal{A}, P)$, существует такая случайная величина $X = X(\omega)$, $\omega \in \Omega$ что

$$P(\{\lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1,$$

то X называется *пределом последовательности* $\{X_n, n \geq 1\}$ *почти наверное* и пишется $X_n \underset{п.н.}{\longrightarrow} X$.

В общей теории меры сходимость почти наверное называется сходимостью *почти всюду*, и это наиболее *сильная* из известных нам форма сходимости функций – случайных величин. Для нее, естественно, справедлив критерий сходимости Коши: если $|X_n - X_m| \underset{п.н.}{\longrightarrow} 0$ при $n, m \to 0$, то последовательность $\{X_n,\, n \geq 1\}$ почти наверное сходится к некоторому пределу X. Существуют довольно сложные в применениях достаточные признаки сходимости почти наверное, однако мы практически не будем в дальнейшем касаться этой формы сходимости, поскольку конкретные результаты (например, усиленный закон больших чисел) требуют для своего доказательства огромных временных затрат, – мы не можем позволить себе такой роскоши в рамках того скудного промежутка времени, который отведен нам учебным планом для изучения теории вероятностей и математической статистики. Мы будем, в основном, иметь дело с более „слабой" формой сходимости, которая в общей теории меры называется *сходимостью по мере*, а в теории вероятностей, как вы наверное уже догадываетесь, *сходимостью по вероятности.*

Определение 11.3. Последовательность $\{X_n,\, n \geq 1\}$ называется сходящейся к пределу X *по вероятности*, если для любого $\varepsilon > 0$

$$\lim_{n\to\infty} P(|\, X_n(\omega) - X)\omega)\,| > \varepsilon) = 0.$$

Сходимость по вероятности обозначается $X_n \underset{P}{\to} X$.

Напомним известный вам из общей теории меры факт: сходимость почти наверное влечет сходимость по вероятности, но обратное, вообще говоря, не верно, и существуют примеры последовательностей, сходящихся по вероятности, но не имеющих предела почти наверное. Однако из всякой сходящейся по вероятности последовательности случайных величин можно извлечь подпоследовательность, сходящуюся к тому же пределу почти наверное.

Мы уже имели дело со сходимостью по вероятности, когда доказывали закон больших чисел Бернулли. Следующий результат обобщает этот закон на суммы независимых случайных величин с достаточно произвольным общим распределением.

Теорема 11.1 (закон больших чисел Чебышева) *Пусть* $\{X_n,$ $n \geq 1\}$ *– последовательность независимых одинаково распределенных случайных величин с конечным вторым моментом. Тогда*

$$\overline{X}_n = \frac{1}{n}\sum_{1}^{n} X_k \underset{P}{\to} \mathbf{E}X_1.$$

Д о к а з а т е л ь с т в о проводится столь же просто, как и в случае закона больших чисел Бернулли, когда $X_k \sim \mathrm{B}(1,\, p)$, и также основано на использовании неравенства Чебышева $P(g(X) > c) \leq \mathbf{E}g(X)/c$. Положим $g(X) = (\overline{X}_n - \mathbf{E}X_1)^2$ и $c = \varepsilon^2$, где ε – произвольное положительное число. Тогда

$$P(|\,\overline{X}_n - \mathbf{E}X_1\,| > \varepsilon) = P\left((\overline{X}_n - \mathbf{E}X_1)^2 > \varepsilon^2\right) \leq \frac{\mathbf{E}(\overline{X}_n - \mathbf{E}X_1)^2}{\varepsilon^2}.$$

Поскольку $\overline{X}_n$ есть нормированная на n сумма независимых случайных величин с общим конечным средним $\mu = \mathbf{E}X_1$ и общей дисперсией $\sigma^2 = \mathbf{D}X_1$, то

$$\mathbf{E}(\overline{X}_n - \mathbf{E}X_1)^2 = \mathbf{E}\left[\frac{1}{n}\sum_1^n (X_k - \mu)\right]^2 = \frac{1}{n^2}\mathbf{D}\sum_1^n X_k = \frac{\sigma^2}{n}.$$

Следовательно,

$$P(|\,\overline{X}_n - \mathbf{E}X_1\,|) \leq \frac{\sigma^2}{n\varepsilon^2} \to 0,$$

каково бы ни было $\varepsilon > 0$.

Мы покажем в дальнейшем, что в законе больших чисел можно убрать требование о существовании второго момента, – достаточно только существование среднего, но для этого мы должны будем разработать более совершенный математический аппарат анализа сходимости по вероятности последовательностей случайных величин. Отметим пока несомненную практическую ценность установленного закона: при независимых равноточных наблюдениях некоторой постоянной μ, характеризующей состояние исследуемого объекта (например, общего содержания серы в партии дизельного топлива), арифметическое среднее $\overline{X}_n$ результатов параллельных наблюдений, отягщенных случайной ошибкой с нулевым средним, является асимптотически точной оценкой μ в том смысле, что $\overline{X}_n \underset{P}{\to} \mu$.

Докажем еще два утверждения, касающиеся сходимости по вероятности, которые обычно называются в монографиях по теории вероятностей *теоремами типа Слуцкого*.

Предложение 11.1. *Пусть последовательность случайных величин $\{X_n,\, n \geq 1\}$ сходится по вероятности к случайной величине X,* *а последовательность случайных величин* $\{\xi_n,\, n \ge 1\}$ *сходится по вероятности к нулю. Тогда* (1) $X_n + \xi_n \underset{P}{\to} X$, (2) $X_n \xi_n \underset{P}{\to} 0$.

Доказательство. (1) Требуется показать, что $P(|\,X_n + \xi_n - X\,| > \varepsilon) \to 0$, каково бы ни было $\varepsilon > 0$.

Используя неравенство $|\,a+b\,| \le |\,a\,| + |\,b\,|$, получаем

$$0 \le P\left(|\,X_n + \xi_n - X\,| > \varepsilon\right) \le P\left(|\,X_n - X\,| + |\,\xi_n\,| > \varepsilon\right) \le$$

$$P\left(\left\{|\,X_n - X\,| > \frac{\varepsilon}{2}\right\} \bigcup \left\{|\,\xi_n\,| > \frac{\varepsilon}{2}\right\}\right) \le P\left(|\,X_n - X\,| > \frac{\varepsilon}{2}\right) +$$

$$P\left(|\,\xi_n\,| > \frac{\varepsilon}{2}\right) \to 0,$$

так как по условию предложения $|\,X_n - X\,| \underset{P}{\to} 0$ и $|\,\xi_n\,| \underset{P}{\to} 0$.

(2) По аналогии с (1) сначала делаем оценку вероятности

$$P\left(|\,X_n \xi_n\,| > \varepsilon\right) = P\left(|\,((X_n - X) + X)\,\xi\,| > \varepsilon\right) \le$$

$$P\left(|\,X_n - X\,|\,|\,\xi_n\,| + |\,X\,|\,|\,\xi_n\,| > \varepsilon\right) \le$$

$$P\left(|\,X_n - X\,|\,|\,\xi_n\,| > \frac{\varepsilon}{2}\right) + P\left(|\,X\,|\,|\,\xi_n\,| > \frac{\varepsilon}{2}\right). \tag{1}$$

Покажем, что первое слагаемое в правой части (1) стремится к нулю, когда $n \to \infty$:

$$P\left(|\,X_n - X\,|\,|\,\xi_n\,| > \varepsilon/2\right) \le P\left(\left\{|\,X_n - X\,| > \sqrt{\varepsilon/2}\right\} \bigcup \left\{|\,\xi\,| > \sqrt{\varepsilon/2}\right\}\right) \le$$

$$P\left(|\,X_n - X\,| > \sqrt{\varepsilon/2}\right) + P\left(|\,\xi_n\,| > \sqrt{\varepsilon/2}\right) \to 0.$$

Теперь покажем, что второе слагаемое в (1) можно сделать меньше любого $\delta > 0$, для всех $n \ge N = N(\delta)$, иными словами, покажем, что второе слагаемое также стремится к нулю при $n \to \infty$. С этой целью введем два противоположных события $\{|\,X\,| > A\}$ и $\{|\,X\,| \le A\}$, в которых число A будет выбрано в дальнейшем по заданному δ, и представим второе слагаемое в виде

$$P\left(|\,X\,|\,|\,\xi_n\,| > \frac{\varepsilon}{2}\right) = P\left(\left\{|\,X\,|\,|\,\xi_n\,| > \frac{\varepsilon}{2}\right\} \bigcap \{|\,X\,| > A\}\right) +$$

$$P\left(\left\{|\,X\,|\,|\,\xi_n\,| > \frac{\varepsilon}{2}\right\} \bigcap \{|\,X\,| \le A\}\right). \tag{2}$$

Первое слагаемое в правой части этого представления не превосходит $P(|X| > A)$, и мы можем выбрать $A = A(\delta)$ по заданному δ настолько большим, что $P(|X| > A) < \delta/2$. Для выбранного таким образом A, которое не зависит от n, оценим второе слагаемое в представлении (2):

$$P\left(\left\{|X| \cdot |\xi_n| > \frac{\varepsilon}{2}\right\} \bigcap \{|X| \leq A\}\right) \leq$$

$$P\left(A \cdot |\xi_n| > \frac{\varepsilon}{2}\right) = P\left(|\xi_n| > \frac{\varepsilon}{2A}\right).$$

Поскольку $\xi_n \underset{P}{\to} 0$, то существует такое $N = N(\delta)$, что для всех $n > N$ вероятность

$$P\left(|\xi_n| > \frac{\varepsilon}{2A}\right) \leq \frac{\delta}{2}.$$

Следует обратить особое внимание на то, как в изучаемых нами видах сходимости почти наверное и по вероятности играет существенную роль задание последовательностей случайных величин на едином вероятностном пространстве $(\Omega, \mathcal{A}, P)$. По существу, близость членов X_n с большими значениями n к их пределу X зависит не столько от совпадения распределений X_n и X, сколько от близости функций $X_n(\omega)$ и $X(\omega)$. Сейчас мы введем еще один вид сходимости случайных величин, более слабый, чем сходимость по вероятности, и для этого вида близость распределений случайных величин становится доминирующей, по сравнению с близостью их как функций на едином пространстве Ω; в этом виде сходимости случайные величины, как компоненты некоторой последовательности, могут быть определены даже на разных пространствах элементарных исходов.

Определение 11.4. Последовательность случайных величин $\{X_n, n \geq 1\}$ сходится к случайной величине X *слабо* или *по распределению*, если соответствующая последовательность их функций распределения $\{F_n(x), n \geq 1\}$ сходится к функции распределения $F(x)$ случайной величины X в каждой точке непрерывности функции $F(x)$.

Слабая сходимость обозначается двойной стрелкой: $X_n \Rightarrow X$, и поскольку речь идет не столько о сходимости случайных величин, сколько об их распределениях, то аналогичное обозначение сохраняется и для последовательности функций распределения: $F_n \Rightarrow F$, и при этом говорят, что последовательность функций распределения $\{F_n,\, n \geq 1\}$ *слабо сходится* к функции $F(x)$.

Возникает естественный вопрос: почему в определении слабой сходимости ограничиваются только точками непрерывности предельной функции $F(x)$? Дело в том, что без этого условия последовательность $\{X_n,\, n \geq 1\}$ и, например, , ,сдвинутая" на бесконечно малую величину последовательность $\{X_n + 1/n,\, n \geq 1\}$ могут иметь различные , ,слабые" пределы, а это – нехорошо.

Приведем пример последовательности случайных величин, которая сходится по распределению, но не имеет предела по вероятности.

Пример 11.1. Пусть $\Omega = [0; 1]$, $\mathcal{A}$ – борелевская σ-алгебра подмножеств этого отрезка и P – раномерное распределение $\mathrm{U}(0, 1)$. На вероятностном пространстве $(\Omega,\, \mathcal{A},\, P)$ введем последовательность случайных величин $\{X_n,\, n \geq 1\}$, полагая $X_n(\omega) = (-1)^n$, если $0 \leq \omega \leq 1/2$, и $X_n(\omega) = (-1)^{n-1}$, если $1/2 < \omega \leq 1$. Все случайные величины последовательности имеют одну и ту же функцию распределения $F(x)$, определяемую вероятностями $P(X_n = -1) = P(X_n = +1) = 1/2$, так что $\{X_n,\, n \geq 1\}$ сходится по распределению. Однако эта последовательность состоит из чередующейся пары различных случайных величин: все члены последовательности с четными номерами X_{2k} принимают значения $+1$ при $0 \leq \omega \leq 1/2$, в то время как соседний член X_{2k+1} с нечетным номером на этом отрезке равен -1. Следовательно, не существует случайной величины $X = X(\omega)$, $\omega \in [0; 1]$ одинаково близкой при больших n ко всем X_n в смысле малости вероятности $P(|\, X_n - X\, | > \varepsilon)$.

Естественно, представляют несомненный интерес достаточные условия, при выполнении которых слабая сходимость влечет сходимость по вероятности. Одно из таких условий дает

Предложение 11.2. *Если последовательность* $X_n \Rightarrow C(\mathrm{const})$, *то* $X_n \underset{P}{\to} C$.

Доказательство. Будем трактовать постоянную C как вырожденную случайную величину с функцией распределения $F(x) = 0$ при $x \leq C$ и $F(x) = 1$ при $x > C$. По условию предложения соответствующая последовательность функций распределения $F_n(x) \to F(x)$ при любом $x \neq C$, ибо C – единственная точка разрыва предельной функции распределения $F(x)$. Требуется показать, что $P(|\, X_n - C\, | > \varepsilon) \to$

0, когда $n \to \infty$, каково бы ни было $\varepsilon > 0$.

Выразим вероятность $P(|\,X_n - C\,| > \varepsilon)$ через функцию распределения $F_n(x)$:

$$P(|\,X_n - C\,| > \varepsilon) = P(X_n - C > \varepsilon) + P(X_n - C < -\varepsilon) =$$

$$P(X_n < C - \varepsilon) + 1 - P(X_n \le C + \varepsilon) =$$

$$F_n(C - \varepsilon) + 1 - F_n(C + \varepsilon) - P(X_n = C + \varepsilon).$$

Так как $C - \varepsilon$ и $C + \varepsilon$ – точки непрерывности функции $F(x)$, то $F_n(C - \varepsilon) \to 0$, а $F_n(C + \varepsilon) \to 1$. Остается показать, что $P(X_n = C + \varepsilon) \to 0$. Имеем

$$0 \le P(X_n = C + \varepsilon) \le P(C + \frac{\varepsilon}{2} \le X_n \le C + \frac{3\varepsilon}{2}) =$$

$$F_n(C + \frac{3\varepsilon}{2}) - F_n(C + \frac{\varepsilon}{2}) \to 0,$$

поскольку $C + 3\varepsilon/2$ и $C + \varepsilon/2$ есть точки непрерывности предельной функции $F(x)$, в которых она принимает одно и то же значение 1.

Теперь мы приступим к построению критерия слабой сходимости, развивая попутно новую и очень сильную технику построения вероятностных моделей.