Лекции 11-12

Фундаментальная теорема математической статистики

Пусть $\eta_1, \eta_2, \dots$ – независимые одинаково распределенные случайные величины с функцией распределения $F(t)$. Введем для каждого $n \in \mathbf{N}$ *эмпирическую функцию распределения*

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} I_{\{\eta_i \le t\}}, \qquad t \in \mathbb{R},$$

где $I_A(\cdot)$ означает индикатор случайного события A. Ясно, что $F_n(t)$ – случайная величина, и по усиленному закону больших чисел почти наверное

$$\lim_{n\to\infty} F_n(t) = F(t).$$

По теореме Гливенко-Кантелли почти наверное

$$\lim_{n\to\infty}\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| = 0.$$

Важно понять, с какой скоростью величина $\sup_{t\in\mathbb{R}} |F_n(t) - F(t)|$, получившая название *статистики Колмогорова*, стремится к нулю. Следующий результат, установленный А.Н. Колмогоровым, является одной из *фундаментальных теорем математической статистики.*

Теорема 1. *Пусть $\eta_1, \eta_2, \dots$ – независимые одинаково распределенные случайные величины с функцией распределения $F(t)$. Тогда для произвольного $x \ge 0$*

$$\lim_{n\to\infty} \mathbf{P}\left(\sqrt{n}\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| \le x\right) = K(x),$$

где $K(\cdot)$ – функция распределения Колмогорова, т.е.

$$K(0) = 0, \qquad K(x) = 1 + 2\sum_{k=1}^{\infty} (-1)^k \exp\left(-2k^2x^2\right), \quad x > 0.$$

Приведем доказательство, основываясь на принципе инвариантности Лиггетта. Сначала установим лемму, сводящую все к частному случаю, когда случайные величины $\eta_1, \eta_2, \dots$ имеют равномерное распределение на интервале $(0, 1)$.

Лемма 1. *Случайная величина $\sup_{t\in\mathbb{R}} |F_n(t) - F(t)|$ имеет такое же распределение, как $\sup_{t\in[0,1]} \left|\widetilde{F}_n(t) - t\right|$, где $\widetilde{F}_n(t)$ – эмпирическая функция распределения, построенная по $\widetilde{\eta}_1, \widetilde{\eta}_2, \dots$ – независимым одинаково распределенным случайным величинам, имеющим равномерное распределение на интервале $(0, 1)$.*

Доказательство. Для простоты изложения будем считать, что функция распределения $F(t)$, $t \in \mathbb{R}$, строго возрастает, тогда существует обратная функция $F^{-1}(y)$, $y \in (0,1)$. Случайные величины $\widetilde{\eta}_i = F(\eta_i)$, $i \in \mathbb{N}$, имеют равномерное распределение на интервале $(0,1)$. Действительно, при $y \in (0,1)$

$$\mathbf{P}(\widetilde{\eta}_i \le y) = \mathbf{P}(F(\eta_i) \le y) = \mathbf{P}\left(\eta_i \le F^{-1}(y)\right) = F\left(F^{-1}(y)\right) = y.$$

Далее,

$$\sup_{t\in\mathbb{R}} |F_n(t) - F(t)| = \sup_{t\in\mathbb{R}} \left| \frac{1}{n} \sum_{i=1}^{n} I_{\{\eta_i \le t\}} - F(t) \right| =$$

$$= \sup_{t\in\mathbb{R}} \left| \frac{1}{n} \sum_{i=1}^{n} I_{\{F(\eta_i) \le F(t)\}} - F(t) \right| = \sup_{u\in[0,1]} \left| \frac{1}{n} \sum_{i=1}^{n} I_{\{\widetilde{\eta}_i \le u\}} - u \right|.$$

Лемма доказана.

Теорема 1 непосредственно вытекает из следующего принципа инвариантности.

Теорема 2. *Если $\eta_1, \eta_2, \ldots$ – независимые случайные величины, имеющие равномерное распределение на интервале $(0,1)$, то при $n \to \infty$*

$$\left\{\sqrt{n}\,(F_n(t) - t)\,,\ t \in [0,1]\right\} \xrightarrow{D} W_0, \tag{1}$$

где W_0 - броуновский мост, знак $\xrightarrow{D}$ означает сходимость по распределению в пространстве $D[0,1]$.

Следствие 1. *В условиях теоремы 1 при $n \to \infty$*

$$\sqrt{n} \sup_{t\in\mathbb{R}} |F_n(t) - F(t)| \xrightarrow{D} \sup_{t\in[0,1]} |W_0(t)|\,. \tag{2}$$

Доказательство. Применим к обеим частям соотношения (1) отображение $f : x \to \sup_{t\in[0,1]} |x(t)|$, где $x \in D[0,1]$. Ранее было показано (см. доказательство теоремы 2 лекций 7-8), что это отображение непрерывно. Поэтому из (1) следует, что при $n \to \infty$

$$\sqrt{n} \sup_{t\in[0,1]} |F_n(t) - t| \xrightarrow{D} \sup_{t\in[0,1]} |W_0(t)| \tag{3}$$

для случая, когда $\eta_1, \eta_2, \ldots$ – независимые случайные величины, имеющие равномерное распределение на интервале $(0,1)$. По лемме 1 левая часть соотношения (3) имеет такое же распределение, как и левая часть соотношения (2), поэтому из (3) следует (2). Следствие доказано.

Следствие 1 означает, что для завершения доказательства теоремы 1 требуется показать, что функцией распределения правой части соотношения (3) является функция $K(\cdot)$. Это мы сделаем позже, а сначала докажем теорему 2. С этой целью установим вспомогательные утверждения.

Положим

$$Y_n(t) = \sum_{i=1}^{n} I_{\{\eta_i \le t\}}.$$

Предположим, что $m \in \mathbb{N}$ и $p_1, \ldots, p_m$ – неотрицательные числа, такие что $p_1 + \ldots + p_m = 1$. Напомним, что случайный вектор $\{\xi_1, \ldots, \xi_m\}$, $m \in \mathbb{N}$, имеет полиномиальное распределение с параметрами n и $p_1, \ldots, p_m$, если для произвольных целых неотрицательных чисел $k_1, \ldots, k_m$ таких, что $\sum_{i=1}^m k_i = n$,

$$\mathbf{P}\left(\xi_1 = k_1, \ldots, \xi_m = k_m\right) = \frac{n!}{k_1! \ldots k_m!} p_1^{k_1} \ldots p_m^{k_m}.$$

Лемма 2. *Пусть* $m \in \mathbb{N}$ *и* $0 = t_0 < t_1 < \ldots < t_m < t_{m+1} = 1$. *Тогда случайный вектор* $\{Y_n(t_i) - Y_n(t_{i-1}),\ i \in\ I_{m+1}\}$ *имеет полиномиальное распределение с параметрами* n *и* $p_1, \ldots, p_{m+1}$, где $p_i = t_i - t_{i-1}$, $i \in\ I_{m+1}$.

Доказательство. Рассмотрим случайную величину η, имеющую равномерное распределение на интервале $(0, 1)$. Будем говорить, что достигается успех i-го типа, если $t_{i-1} < \eta \leq t_i$, $i \in\ I_{m+1}$. Тогда $Y_n(t_i) - Y_n(t_{i-1})$ – число успехов i-го типа в n независимых испытаниях, $i \in\ I_{m+1}$. Хорошо известно, что случайный вектор, состоящий из таких компонент имеет полиномиальное распределение. Лемма доказана.

Положим

$$\widetilde{Y}_n(t) = Y_n\left(\frac{\lfloor nt \rfloor}{n}\right).$$

Следствие 2. *Пусть* $m \in \mathbb{N}$ *и* $0 = t_0 < t_1 < \ldots < t_m < t_{m+1} = 1$. *Тогда случайный вектор* $\left\{\widetilde{Y}_n(t_i) - \widetilde{Y}_n(t_{i-1}),\ i \in\ I_{m+1}\right\}$ *имеет полиномиальное распределение с параметрами* n *и* $\widetilde{p}_1, \ldots, \widetilde{p}_{m+1}$, *где* $\widetilde{p}_i = \left(\lfloor nt_i \rfloor - \lfloor nt_{i-1} \rfloor\right)/n$, $i \in\ I_{m+1}$.

Рассмотрим последовательность независимых случайных величин X_i, $i \in \mathbb{N}$, имеющих распределение Пуассона с параметром $a > 0$, и положим $S_n = X_1 + \ldots + X_n$, $n \in \mathbb{N}$.

Лемма 3. *Пусть* $m, n \in \mathbb{N}$ *и натуральные числа* $n_1, n_2, \ldots, n_m$ *возрастают. Тогда случайный вектор* $\left\{ S_{n_1}, S_{n_2} - S_{n_1}, \ldots, S_{n_m} - S_{n_{m-1}} \;\middle|\; S_{n_m} = n\right\}$ *имеет полиномиальное распределение с параметрами* n *и* $p_1, \ldots, p_m$, *где* $p_i = \left(n_i - n_{i-1}\right)/n_m$, $i \in\ I_m$, *причем* $n_0 = 0$.

Доказательство. При целых неотрицательных числах $k_1, \ldots, k_m$ таких, что $\sum_{i=1}^m k_i = n$, находим, что

$$\mathbf{P}\left(S_{n_i} - S_{n_{i-1}} = k_i,\ i \in I_m \;\middle|\; S_{n_m} = n\right) =$$

$$= \frac{1}{\mathbf{P}\left(S_{n_m} = n\right)} \prod_{i=1}^m \mathbf{P}\left(S_{n_i} - S_{n_{i-1}} = k_i\right) =$$

$$= \frac{n!}{\left(an_m\right)^n e^{-an_m}} \prod_{i=1}^m \frac{\left(a\left(n_i - n_{i-1}\right)\right)^{k_i} e^{-a\left(n_i - n_{i-1}\right)}}{k_i!} =$$

$$= \frac{n!}{k_1! \ldots k_m!} \prod_{i=1}^m \left(\frac{n_i - n_{i-1}}{n_m}\right)^{k_i} = \frac{n!}{k_1! \ldots k_{m+1}!} \prod_{i=1}^m p_i^{k_i}.$$

Лемма доказана.

Следствие 3. *Пусть* $m \in \mathbb{N}$ *и* $0 = t_0 < t_1 < \ldots < t_m < t_{m+1} = 1$. *Тогда случайный вектор* $\left\{ S_{\lfloor nt_1 \rfloor}, S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor}, \ldots, S_{\lfloor nt_{m+1} \rfloor} - S_{\lfloor nt_m \rfloor} \;\middle|\; S_n = n \right\}$ *имеет полиномиальное распределение с параметрами* n *и* $\widetilde{p}_1, \ldots, \widetilde{p}_{m+1}$, *где* $\widetilde{p}_i = (\lfloor nt_i \rfloor - \lfloor nt_{i-1} \rfloor)/n$, $i \in I_{m+1}$.

Приступим к непосредственному доказательству теоремы 2. Положим

$$Z_n(t) = \sqrt{n}\,(F_n(t) - t).$$

Наша задача показать, что при $n \to \infty$

$$Z_n \xrightarrow{D} W_0. \tag{4}$$

Очевидно, что $F_n(t) = Y_n(t)/n$. Откуда следует, что

$$Z_n(t) = \frac{Y_n(t) - nt}{\sqrt{n}}.$$

Положим

$$\widetilde{Z}_n(t) = Z_n\left(\frac{\lfloor nt \rfloor}{n}\right).$$

Тогда

$$\widetilde{Z}_n(t) = \frac{\widetilde{Y}_n(t) - \lfloor nt \rfloor}{\sqrt{n}}.$$

Докажем сначала, что при $n \to \infty$

$$\widetilde{Z}_n \xrightarrow{D} W_0. \tag{5}$$

Ввиду следствий 2 и 3 совпадают конечномерные распределения процессов

$$\left\{\widetilde{Y}_n(t),\ t \in [0,1]\right\} \text{ и } \left\{ S_{\lfloor nt \rfloor},\ t \in [0,1] \;\middle|\; S_n = n \right\}.$$

В дальнейшем будем рассматривать частный случай второго из этих процессов, когда $a = 1$. Положим $X_i^* = X_i - 1$ при $i \in \mathbb{N}$ и

$$S_n^* = X_1^* + \ldots + X_n^*, \quad n \in \mathbb{N}.$$

Поскольку $S_n^* = S_n - n$ при $n \in \mathbb{N}$, то

$$\left\{ S_{\lfloor nt \rfloor},\ t \in [0,1] \;\middle|\; S_n = n \right\} = \left\{ S^*_{\lfloor nt \rfloor} + \lfloor nt \rfloor,\ t \in [0,1] \;\middle|\; S_n^* = 0 \right\}.$$

Таким образом, совпадают конечномерные распределения процессов

$$\left\{\widetilde{Y}_n(t),\ t \in [0,1]\right\} \text{ и } \left\{ S^*_{\lfloor nt \rfloor} + \lfloor nt \rfloor,\ t \in [0,1] \;\middle|\; S_n^* = 0 \right\}$$

и, значит, совпадают конечномерные распределения процессов

$$\left\{\widetilde{Z}_n(t),\ t \in [0,1]\right\} \text{ и } \left\{ S^*_{\lfloor nt \rfloor}/\sqrt{n},\ t \in [0,1] \;\middle|\; S_n^* = 0 \right\}.$$

Заметим, что $\mathbf{E}X_1^* = 0$, $\mathbf{E}\left(X_1^*\right)^2 = 1$ и, следовательно, по принципу инвариантности Лиггетта при $n \to \infty$

$$\left\{ S^*_{\lfloor nt \rfloor}/\sqrt{n},\ t \in [0,1] \ \middle|\ S^*_n = 0 \right\} \xrightarrow{D} W_0.$$

Таким образом, справедливо соотношение (5).

Отображение $x \to w_x(\delta)$, где $x \in D[0,1]$, непрерывно при $x \in C[0,1]$. Поскольку траектории броуновского моста почти наверное непрерывны, то $w_{\widetilde{Z}_n}(\delta) \xrightarrow{D} w_{W_0}(\delta)$ при $n \to \infty$ и, следовательно, для всех положительных ε (за исключением не более чем счетного множества)

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mathbf{P}\left(w_{\widetilde{Z}_n}(\delta) \geq \varepsilon\right) = \lim_{\delta\to 0}\mathbf{P}\left(w_{W_0}(\delta) \geq \varepsilon\right) = 0. \tag{6}$$

Поскольку функция $Y_n(t)$, $t \geq 0$, не убывает, то при $t \in [0,1]$

$$0 \leq Y_n(t) - \widetilde{Y}_n(t) \leq \widetilde{Y}_n(t+1/n) - \widetilde{Y}_n(t) \leq w_{\widetilde{Y}_n}(1/n)$$

и, следовательно,

$$\rho_{\text{равн}}\left(Z_n, \widetilde{Z}_n\right) \leq \frac{w_{\widetilde{Y}_n}(1/n)+1}{\sqrt{n}} \leq \frac{\sqrt{n}w_{\widetilde{Z}_n}(1/n)+2}{\sqrt{n}} = w_{\widetilde{Z}_n}\left(\frac{1}{n}\right) + \frac{2}{\sqrt{n}}. \tag{7}$$

Из соотношений (6), (7) следует, что для любого $\varepsilon > 0$

$$\lim_{n\to\infty}\mathbf{P}\left(\rho_{\text{равн}}\left(Z_n, \widetilde{Z}_n\right) \geq \varepsilon\right) = 0.$$

Но это вместе с соотношением (5) означает, как нетрудно понять, справедливость (4). Теорема 2 доказана.

Для завершения доказательства теоремы 1 требуется показать, что

$$K(x) = \mathbf{P}\left(\sup_{t\in[0,1]}|W_0| \leq x\right),\ x \geq 0. \tag{8}$$

По теореме 1 лекций 9-10 при $x \geq 0$

$$\mathbf{P}\left(\sup_{t\in[0,1]}|W_0(t)| \leq x\right) = \lim_{\varepsilon\to 0}\mathbf{P}\left(\sup_{t\in[0,1]}|W(t)| \leq x \ \middle|\ |W(1)| \leq \varepsilon\right) =$$

$$= \lim_{\varepsilon\to 0}\mathbf{P}\left(\sup_{t\in[0,1]}|W(t)| \leq x,\ \ |W(1)| \leq \varepsilon\right)/\mathbf{P}\left(|W(1)| \leq \varepsilon\right). \tag{9}$$

Пусть верхний индекс у $\mathbf{P}^{(x)}$ означает, что броуновское движение W стартует из точки x, а не из точки 0.

Лемма 4. *Для $x > 0$ и борелевского множества $D \subset (-x, x)$ справедливо равенство*

$$\mathbf{P}\left(\sup_{t\in[0,1]}|W(t)| \geq x,\ W(1) \in D\right) =$$

$$= \sum_{k=1}^{\infty} \Big(\mathbf{P}^{((4k-2)x)} \left(W(1) \in D \right) - \mathbf{P}^{(4kx)} \left(W(1) \in D \right) \Big) +$$

$$+ \sum_{k=1}^{\infty} \Big(\mathbf{P}^{(-(4k-2)x)} \left(W(1) \in D \right) - \mathbf{P}^{(-4kx)} \left(W(1) \in D \right) \Big) .$$

Доказательство. Для броуновского движения справедлив принцип отражения: для произвольного борелевского множества $C \subset (-\infty, x)$ при $x > 0$

$$\mathbf{P} \left(\, M \geq x, \; W\left(1\right) \in C \right) = \mathbf{P}^{(2x)} \left(W\left(1\right) \in C \right), \tag{10}$$

где $M = \sup_{t \in [0,1]} W\left(t\right)$.

Действительно, если броуновская траектория, рассматриваемая при $t \in [0,1]$, стартует из точки 0 и $M \geq x$, то она достигает состояния x (пусть впервые это происходит в момент τ_x). Если отразить часть броуновской траектории до момента τ_x относительно прямой l_x, ординаты точек которой равны x, то получится броуновская траектория, стартующая из точки $2x$. Наоборот, если броуновская траектория стартует из точки $2x$ и завершается в точке из множества $C \subset (-\infty, x)$, то она достигает состояния x (пусть впервые это происходит в момент τ'_x). Если отразить часть броуновской траектории до момента τ'_x относительно прямой l_x, то получится броуновская траектория, стартующая из точки 0 и у которой максимум больше или равен x.

Справедливо обобщение соотношения (10). Пусть A_0 – случайное событие, состоящее в том, что броуновская траектория, стартующая из точки 0, достигает состояния x раньше, чем состояния $(-x)$, и завершается в точке из множества $D \subset (-x, x)$.

Применяя отражение броуновской траектории относительно прямой l_x, получаем, что

$$\mathbf{P}\left(A_0\right) = \mathbf{P}\left(A_2\right),$$

где A_2 – случайное событие, состоящее в том, что броуновская траектория, стартующая из точки $2x$, достигает состояния x раньше, чем состояния $3x$, и завершается в точке из множества D. Ясно, что

$$\mathbf{P}\left(A_2\right) = \mathbf{P}\left(A'_2\right) - \mathbf{P}\left(A''_2\right),$$

где A'_2 – случайное событие, состоящее в том, что броуновская траектория, стартующая из точки $2x$, завершается в точке из множества D; A''_2 – случайное событие, состоящее в том, что броуновская траектория, стартующая из точки $2x$, достигает состояния $3x$ раньше, чем состояния x, и завершается в точке из множества D.

Далее, применяя отражение броуновской траектории относительно прямой l_{3x}, получаем, что

$$\mathbf{P}\left(A''_2\right) = \mathbf{P}\left(A_4\right),$$

где A_4 – случайное событие, состоящее в том, что броуновская траектория, стартующая из точки $4x$, достигает состояния $3x$ раньше, чем состояния $5x$, и завершается в точке из множества D. Ясно, что

$$\mathbf{P}\left(A_4\right) = \mathbf{P}\left(A'_4\right) - \mathbf{P}\left(A''_4\right),$$

где A_4' – случайное событие, состоящее в том, что броуновская траектория, стартующая из точки $4x$, завершается в точке из множества D; A_4'' – случайное событие, состоящее в том, что броуновская траектория, стартующая из точки $4x$, достигает состояния $5x$ раньше, чем состояния $3x$, и завершается в точке из множества D.

Продолжая эти рассуждения, в итоге получаем, что

$$\begin{aligned}\mathbf{P}(A_0) = \mathbf{P}(A_2') - \mathbf{P}(A_2'') = \mathbf{P}(A_2') - (\mathbf{P}(A_4') - \mathbf{P}(A_4'')) = \\ = \mathbf{P}(A_2') - \mathbf{P}(A_4') + (\mathbf{P}(A_6') - \mathbf{P}(A_6'')) = \\ = \mathbf{P}(A_2') - \mathbf{P}(A_4') + \mathbf{P}(A_6') - (\mathbf{P}(A_8') - \mathbf{P}(A_8'')) = \ldots = \\ = \sum_{k=1}^{\infty} \left(\mathbf{P}^{((4k-2)x)}(W(1) \in D) - \mathbf{P}^{(4kx)}(W(1) \in D) \right).\end{aligned} \tag{11}$$

Пусть B_0 – случайное событие, состоящее в том, что броуновская траектория, стартующая из точки 0, достигает состояния $(-x)$ раньше, чем состояния x, и завершается в точке из множества $D \subset (-x, x)$. Аналогично показывается, что

$$\mathbf{P}(B_0) = \sum_{k=1}^{\infty} \left(\mathbf{P}^{(-(4k-2)x)}(W(1) \in D) - \mathbf{P}^{(-4kx)}(W(1) \in D) \right). \tag{12}$$

Поскольку

$$\left\{ \sup_{t \in [0,1]} |W(t)| \geq x,\ W(1) \in D \right\} = A_0 + B_0 \text{ и } A_0 \cap B_0 = \varnothing,$$

то из формул (11) и (12) следует утверждение леммы.

Из леммы 4 следует, что

$$\begin{aligned}\mathbf{P}(\sup_{t \in [0,1]} |W(t)| \leq x,\ |W(1)| \leq \varepsilon) = \\ = \mathbf{P}(\ |W(1)| \leq \varepsilon) - \mathbf{P}(\sup_{t \in [0,1]} |W(t)| \geq x,\ |W(1)| \leq \varepsilon) = \\ = \mathbf{P}(|W(1)| \leq \varepsilon) - 2\sum_{k=1}^{\infty} \left(\mathbf{P}^{((4k-2)x)}(|W(1)| \leq \varepsilon) - \mathbf{P}^{(4kx)}(|W(1)| \leq \varepsilon) \right).\end{aligned} \tag{13}$$

Заметим, что для каждого $a \in \mathbb{R}$

$$\lim_{\varepsilon \to 0} \frac{\mathbf{P}^{(a)}(|W(1)| \leq \varepsilon)}{\mathbf{P}(|W(1)| \leq \varepsilon)} = e^{-\frac{a^2}{2}}. \tag{14}$$

Из соотношений (13) и (14) находим, что

$$\begin{aligned}\lim_{\varepsilon \to 0} \mathbf{P}\left(\sup_{t \in [0,1]} |W(t)| \leq x,\ |W(1)| \leq \varepsilon \right) / \mathbf{P}(|W(1)| \leq \varepsilon) = \\ = 1 - 2\sum_{k=1}^{\infty} \left(\exp\left(-2(2k-1)^2 x^2\right) - \exp\left(-2\,(2k)^2\, x^2\right) \right) = \\ = 1 + 2\sum_{k=1}^{+\infty} (-1)^k\, e^{-2k^2x^2} = K(x).\end{aligned} \tag{15}$$

Из соотношений (9) и (15) следует (8). Теорема 1 доказана.