Функциональная предельная теорема

На прошлой лекции мы с вами сформулировали теорему Донскера-Прохорова о сходимости нормированного случайного блуждания к броуновскому движению
$$Y_n \xrightarrow{d} W,$$
где
$$Y_n(t) = \frac{S_{[nt]} + (nt - [nt])X_{[nt]+1} - \mu nt}{\sigma\sqrt{n}}.$$

Как мы уже сказали, такого результаты достаточно ценны, ведь доказав одну такую теорему, мы получаем массу результатов о сходимости. Хорошим примером такого рода результатов является следующее известное утверждение:
Пример 20.1. Пусть X_i — н.о.р. величины с непрерывной ф.р. $F(x)$. Тогда при любом $y > 0$
$$P(\sup|\hat{F}_n(x) - F(x)| \le y) \to K(y),$$
где $\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{X_i \le x}$ — ЭФР, $K(x)$ — функция распределения Колмогорова. Это — хорошо известная вам из статистики теорема Колмогорова. Доказывать мы ее не будем, но расскажем, откуда возникает этот результат. Базовым фактом для теоремы Колмогорова является аналогичная теореме Донскера теорема о том, что для $X_i \sim R[0,1]$ выполнено соотношение $\sqrt{n}(\hat{F}_n(x) - x) \xrightarrow{d} W^0$, где $W_t^0 = W_t - tW_1$ — так называемый процесс броуновского моста. Правда, ни о какой сходимости в $C[0,1]$ здесь не может идти и речи, ведь $\hat{F}_n(x)$ — разрывная функция, взамен используется сходимость в $D[0,1]$ — пространстве непрерывных справа функций, а вместо равномерной нормы — более сложная топология Скорохода. Тогда поскольку sup — непрерывный функционал в $D[0,1]$ около этой топологии. Из этого следует, что
$$P(\sup_x(\hat{F}_n(x) - x) \le y) \to P(\sup_x W_x^0 \le y).$$

Явное выражение для ф.р. $\sup_x W_x^0$ получается напрямую, один из вариантов получения такого результата мы покажем чуть позже. При этом, доказав функциональную предельную теорему для $X_i \sim R[0,1]$, мы получаем ее для любого другого непрерывного распределения. Кроме теоремы Колмогорова, та же функциональная предельная теорема позволяет искать предельные распределения для других функционалов, например,
$$f_1(g) = \int_{-\infty}^{\infty} g^2(x)dF(x),\ f_2(g) = \int_{\infty}^{\infty} g^2(x)\frac{F(x)(1-F(x)}{d}F(x).$$

Тогда $f_i(\sqrt{n}(\hat{F}_n - F)) \xrightarrow{d} f_i(W^0)$, $i = 1, 2$. На основе первого утверждения строится так называемый критерий Крамера-Мизеса, а на основе второго — критерий Андерсона-Дарлинга для проверки простой гипотезы $F = F_0$, где F_0 непрерывна.

Броуновским мостом называют процесс W_t^0, имеющий то же распределение, что и $W_t - tW_1$ при $t \in [0,1]$. Нетрудно вывести, что конечномерные распределения этого процесса задаются соотношением:
$$P(W_{t_1}^0 \in [a_1, b_1], ..., W_{t_k}^0 \in [a_k, b_k]) = \int_{a_1}^{b_1} ... \int_{a_k}^{b_k} e^{-\frac{x_k^2}{2(1-t_k)}} e^{-\frac{(x_k - x_{k-1})^2}{2(t_k - t_{k-1})}} ... e^{-\frac{x_1^2}{2t_1}} dx_k ... dx_1.$$

Указание Найдите распределение вектора $(W_{t_1}, W_{t_2} - W_{t_1}, ..., W_{t_k} - W_{t_{k-1}}, W_1 - W_{t_k})$ и посмотрите как оно изменится при рассматриваемой линейной замене.

Рассмотрим последовательность
$$Y_n(t) = \frac{S_{[nt]} + X_{[nt]+1}(nt - [nt]) - \mu nt}{\sigma\sqrt{n}},$$
где $EX_i = \mu$, $DX_i = \sigma^2$.
Теорема 20.1. Тогда соотношение
$$\mathbf{P}(Y_n(t) \in A | S_n \in [x, x + \Delta_n]) \to \mathbf{P}(W_t^0 \in A),\ n \to \infty,\ \Delta_n \to 0,$$
где $A : \mathbf{P}(W_t^0 \in \partial A)$, выполнено равномерно по $x \in [\mu n - t\sqrt{n}, \mu n + t\sqrt{n}]$ при любом t.

Доказательство Теоремы 20.1. Без ограничения общности будем считать, что $\mu = 0$, $\sigma = 1$. Для доказательства мы покажем
1) что конечномерные распределения Y_n при условии $S_n \in [x, x + \Delta_n)$ сходятся к конечномерным распределениям W^0
2) что семейство мер $P_n(A)$ плотно.
1) Докажем сперва следующую лемму
Лемма 20.1. При любых $0 = t_0 < t_1 < ... < t_k \le 1$
$$P(S_{[nt_1]} \in [x_1, x_1 + \Delta_n), ..., S_{nt_k} \in [x_k, x_k + \Delta_n)) \sim \frac{\Delta_n^k}{(2\pi n)^{k/2}\prod_{i=1}^{k}\sqrt{t_i - t_{i-1}}} e^{-\frac{x_1^2}{2nt_1}} e^{-\frac{(x_2 - x_1)^2}{2n(t_2 - t_1)}} ... e^{-\frac{(x_k - x_{k-1})^2}{2n(t_k - t_{k-1})}}$$
равномерно по $|x_i| \le t\sqrt{n}$ при всех Δ_n, достаточно медленно стремящихся к 0.
Доказательство

Будем доказывать индукцией по k. Для $k=1$ это просто интегролокальная теорема. Пусть для $k=l$ доказано, докажем для $k=l+1$. Рассмотрим $\tilde{\Delta}_n = o(\Delta_n)$. Тогда

$$P(S_{[nt_1]} \in [x_1, x_1+\Delta_n), ..., S_{nt_{l+1}} \in [x_{l+1}, x_{l+1}+\Delta_n)) \le \sum_{i \le I} P(S_{[nt_1]} \in [x_1 + i\tilde{\Delta}_n, x_1 + (i+1)\tilde{\Delta}_n)) \times$$

$$P(S_{[nt_2]-[nt_1]} \in [x_2 - x_1 - (i+1)\tilde{\Delta}_n, x_2 - x_1 - i\tilde{\Delta} + \Delta_n], ..., S_{[nt_{l+1}]-[nt_1]} \in [x_{l+1} - x_1 - (i+1)\tilde{\Delta}_n, x_{l+1} - x_1 - i\tilde{\Delta} + \Delta_n]) \sim$$

$$\frac{\tilde{\Delta}_n \Delta_n^l}{(2\pi n)^{(l+1)/2} \prod_{j=1}^{l+1} \sqrt{t_j - t_{j-1}}} \sum_{i \le I} e^{-\frac{(x_1+i\tilde{\Delta}_n)^2}{2nt_1}} e^{-\frac{(x_2-x_1-i\tilde{\Delta}_n)^2}{2n(t_2-t_1)}} e^{-\frac{(x_3-x_2)^2}{2n(t_3-t_2)}} ... e^{-\frac{(x_{l+1}-x_l)^2}{2n(t_{l+1}-t_l)}},$$

где $I = \Delta_n/\tilde{\Delta}_n - 1$. При этом

$$e^{-\frac{(x_1+i\tilde{\Delta}_n)^2}{2nt_1}} \sim e^{-\frac{x_1^2}{2nt_1}}, e^{-\frac{(x_2-x_1-i\tilde{\Delta}_n)^2}{2n(t_2-t_1)}} \sim e^{-\frac{(x_2-x_1)^2}{2n(t_2-t_1)}},$$

откуда и получаем требуемую оценку сверху. Оценка снизу совершенно аналогична.

Для доказательства сходимости конечномерных распределений нам требуется доказать, что

$$P(Y_n(t_1) \in [a_1, b_1], ..., Y_n(t_k) \in [a_k, b_k] | S_n \in [x, x+\Delta_n)) \to P(W^0_{t_1} \in [a_1, b_1], ..., W^0_{t_k} \in [a_k, b_k])$$

при любых a_i, b_i, t_i, причем сходимость равномерна по x. Заметим, что

$$P(Y_n(t_1) \in [a_1, b_1], ..., Y_n(t_k) \in [a_k, b_k], S_n \in [x, x+\Delta_n)) =$$
$$P\left(\frac{S_{[nt_1]}}{\sqrt{n}} + \frac{(nt-[nt])X_{[nt_1]+1}}{\sqrt{n}} \in [a_1, b_1], ..., \frac{S_{[nt_k]}}{\sqrt{n}} + \frac{(nt-[nt])X_{[nt_1]+1}}{\sqrt{n}} \in [a_k, b_k], S_n \in [x, x+\Delta_n)\right).$$

При этом $P(X_{[nt_1]+1} > \varepsilon\sqrt{n}) \le \varepsilon^{-1} n^{-1} DX$, то есть $o(P(S_n \in [x, x+\Delta_n)))$ при всех достаточно медленно стремящихся к 0 Δ_n. Следовательно,

$$P(Y_n(t_1) \in [a_1, b_1], ..., Y_n(t_k) \in [a_k, b_k], S_n \in [x, x+\Delta_n)) \le o(P(S_n \in [x, x+\Delta_n)) +$$
$$P\left(\frac{S_{[nt_1]}}{\sqrt{n}} \in [a_1 - \varepsilon, b_1 + \varepsilon], ..., \frac{S_{[nt_k]}}{\sqrt{n}} \in [a_k, b_k], S_n \in [x, x+\Delta_n)\right).$$

Последняя вероятность в силу Леммы есть

$$\sum_{i_1 \in I_1} ... \sum_{i_k \in I_k} P\left(S_{[nt_1]} \in [i_1\Delta_n, (i_1+1)\Delta_n), ..., S_{[nt_k]} \in [i_k\Delta_n, (i_k+1)\Delta_n), S_n \in [x, x+\Delta_n)\right) \sim$$
$$\frac{\Delta_n^{k+1}}{(2\pi n)^{(k+1)/2}\sqrt{1-t_k} \prod_{i=1}^k \sqrt{t_i - t_{i-1}}} \sum_{i_1 \in I_1} ... \sum_{i_k \in I_k} e^{-\frac{i_1^2\Delta_n^2}{2nt_1}} e^{-\frac{(i_2-i_1)^2\Delta_n^2}{2n(t_2-t_1)}} ... e^{-\frac{(i_k-i_{k-1})^2\Delta_n^2}{2n(t_k-t_{k-1})}} e^{-\frac{(x-i_k\Delta_n)^2}{2n(1-t_k)}},$$

где $I_j = \left\{\frac{a_j\sqrt{n}}{\Delta_n}, \frac{a_j\sqrt{n}}{\Delta_n} + 1 ..., \frac{b_j\sqrt{n}}{\Delta_n}\right\}$ (для простоты записи мы опустили целые части, которые должны были здесь стоять). При этом

$$e^{-\frac{(x-i_k\Delta_n)^2}{2n(1-t_k)}} \sim e^{-\frac{i_k^2\Delta_n^2}{2n(1-t_k)}},$$

откуда, переходя от сумм по I_j к интегралам, для которых они являются интегральными суммами, имеем

$$P(Y_n(t_1) \in [a_1, b_1], ..., Y_n(t_k) \in [a_k, b_k], S_n \in [x, x+\Delta_n)) \sim \frac{\Delta_n}{(2\pi)^{k+1}\sqrt{n}\sqrt{1-t_k}\prod_{i=1}^k \sqrt{t_i - t_{i-1}}} \times$$
$$\int_{a_1}^{b_1} ... \int_{a_k}^{b_k} e^{-\frac{x_1^2}{2t_1}} ... e^{-\frac{(x_k-x_{k-1})^2}{2(t_k-t_{k-1})}} e^{-\frac{x_k^2}{2(1-t_k)}} dx_1 ... dx_k.$$

Деля полученное соотношение на $P(S_n \in [x, x+\Delta_n)) \sim \frac{\Delta_n}{\sqrt{2\pi n}}$, получаем, что рассматриваемые к.м. распределения сходятся к

$$\frac{1}{(2\pi)^k\sqrt{1-t_k}\prod_{i=1}^k \sqrt{t_i - t_{i-1}}} \int_{a_1}^{b_1} ... \int_{a_k}^{b_k} e^{-\frac{x_1^2}{2t_1}} ... e^{-\frac{(x_k-x_{k-1})^2}{2(t_k-t_{k-1})}} e^{-\frac{x_k^2}{2(1-t_k)}} dx_1 ... dx_k.$$

Это требуемые конечномерные распределения броуновского моста.

2) Для доказательства плотности воспользуемся следующим достаточным условием:
a) Для любого $\varepsilon > 0$ найдется a, т.ч. $P(|Y_n(0)| > a) < \varepsilon$ при всех n.
b) при всех $\varepsilon, \varepsilon_1 > 0$ и всех достаточно малых δ справедливо неравенство

$$P(\sup_{|t-s|<\delta} |Y_n(t) - Y_n(s)| > \varepsilon) < \varepsilon_1, \ \forall n.$$

Условие a) очевидно выполнено, проверим b). Будем использовать неравенство Этемади:

$$\mathbf{P}(\max_{i \le n} S_n \ge 3x) \le 3 \max_{i \le n} P(|S_i| \ge x).$$

Это неравенство довольно очевидно вытекает из соотношений

2) Для доказательства плотности воспользуемся следующим достаточным условием плотности в $C[0,1]$ (см. Биллингсли, Сходимость вероятностных мер, 1977, теорема 8.3).

Теорема A. Последовательность мер P_n плотна, если выполняются два соотношения:

1) Для любого $\varepsilon > 0$ найдется t, такое что

$$P_n(x : |x(0)| > a) \le \varepsilon.$$

2) Для любых $\varepsilon, \varepsilon_1 > 0$ найдутся $\delta \in (0,1)$, $n_0 \in \mathbb{N}$, т.ч. при всех t, $n > n_0$

$$\frac{1}{\delta} P_n(\sup_{t \le s \le t+\delta} |x(s) - x(t)| \ge \varepsilon) \le \varepsilon_1.$$

Кроме того, нам понадобится неравенство Этемади

Лемма 20.2. Для любых н.о.р. ξ_i, $x > 0$

$$P(\max_{i \le n} |S_i| \ge 3x) \le 3 \max P(|S_i| \ge x).$$

Для доказательства положим $B_i = P(|S_i| \ge 3x, |S_j| < 3x, j \le i)$ и заметим, что

$$\sum_{i=1}^{n} P(B_i) \le P(|S_n| \ge x) + \sum_{i=1}^{n-1} P(B_i, |S_n| < x) \le P(|S_n| \ge x) + \sum_{i=1}^{n-1} P(B_i) P(|S_n - S_i| \ge 2x) \le$$

$$P(|S_n| \ge x) + \sum_{i=1}^{n-1} P(B_i)(P(|S_n| \ge x) + P(|S_i| \ge x)) \le 3 \max P(|S_i| \ge x).$$

Итак, докажем, что при любых ε, ε_1 найдется δ, т.ч.

$$\frac{1}{\delta} P\left(\sup_{t \le s \le t+\delta} |Y_n(s) - Y_n(t)| \ge \varepsilon \middle| S_n \in [x, x+\Delta_n)) \right) < \varepsilon_1$$

при всех достаточно больших n. Заметим, что Заметим, что

$$|Y_n(s) - Y_n(t)| \le \frac{\max(|S_{[ns]} - S_{[nt]}|, |S_{[ns]+1} - S_{[nt]}|, |S_{[ns]} - S_{[nt]+1}|, |S_{[ns]+1} - S_{[nt]+1}|}{\sigma(h_\theta)\sqrt{n}},$$

поскольку линейно интерполированная функция принимает максимум в одной из исходных точек. Тем самым, нам достаточно оценить вероятность

$$P\left(\sup_{t \le s \le t+\delta} |S_{[ns]} - S_{[nt]}| \ge \varepsilon\sigma\sqrt{n} \middle| S_n \in [x, x+\Delta_n) \right) = \frac{P\left(\sup_{s \le \delta} |S_{[ns]}| \ge \varepsilon\sigma\sqrt{n}, S_n \in [x, x+\Delta_n) \right)}{P(S_n \in [x, x+\Delta_n))}.$$

Заметим, что при любом δ $P(|S_{[n\delta]}| > t\sigma\sqrt{n} | S_n \in [x, x+\Delta_n)) < \delta\varepsilon/2$ при достаточно большом t и n в силу части 1), поскольку

$$P\left(\left| \frac{S_{[n\delta]}}{\sigma\sqrt{n}} \right| > t | S_n \in [x, x+\Delta_n) \right) \to P(|W_\delta^0| > t),$$

где правая часть может быть сделана сколь угодно малой за счет выбора t.

Аналогичным образом доказывается следующая теорема:

Теорема 20.2. Пусть X_i — н.о.р., $\mathbf{E}X_i = \mu$, $\mathbf{D}X_i = \sigma^2$, X_i нерешетчаты, $R(h) = \mathbf{E}e^{hX} < \infty$, $h \in [0, h^+)$. Тогда

$$\mathbf{P}(Y_{n,x}(t) \in A | S_n \in [x, x+\Delta_n]) \to \mathbf{P}(W_t^0 \in A),\ n \to \infty,\ \Delta_n \to 0,$$

где $A : \mathbf{P}(W_t^0 \in \partial A)$, выполнено равномерно по $x/n \in [\mu, m_2]$ при любом $m_2 < m^+$. Здесь

$$Y_n(t) = \frac{S_{[nt]} + (nt - [nt]) X_{[nt]+1} - xt}{\sigma(h_{x/n})\sqrt{n}}.$$

Доказательство Теоремы 14.2 близко к доказательству Теоремы 14.1 и в курсе опущено.

Аналогичным образом можно получить теорему для решетчатых величин.

Теорема 20.3. Пусть X_i — н.о.р., $\mathbf{E}X_i = \mu$, $\mathbf{D}X_i = \sigma^2$, X_i решетчаты с шагом 1, $R(h) = \mathbf{E}e^{hX} < \infty$, $h \in [0, h^+)$. Тогда при целых x

$$\mathbf{P}(Y_{n,x}(t) \in A | S_n = x) \to \mathbf{P}(W_t^0 \in A),\ n \to \infty,\ \Delta_n \to 0,$$

где $A : \mathbf{P}(W_t^0 \in \partial A)$, выполнено равномерно по $x/n \in [\mu, m_2]$ при любом $m_2 < m^+$.