21.04.2021

W poprzednim odcinku:

$$\mathbb{E}X = \int_{\mathbb{R}} x \, d\mu(x), \quad X \sim \mu$$

$$\operatorname{Var} X = \mathbb{E}[X - \mathbb{E}X]^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2$$

Jeżeli X, Y nzl., to $\mathbb{E}X \cdot Y = \mathbb{E}X \cdot \mathbb{E}Y$,

$$\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y)$$

$$\operatorname{Cov}(X,Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}X)]$$
$$= \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y$$

$$\operatorname{Var} X = \operatorname{Cov}(X, X)$$

Definicja 5.15 Niech $X = (X_1, \dots, X_n)$ będzie n-wymiarową zmienną losową taką, że $\mathbb{E}X_i^2 < \infty$. Macierz

$$Q = \begin{bmatrix} \operatorname{Cov}(X_1,X_1) & \operatorname{Cov}(X_1,X_2) & \dots & \operatorname{Cov}(X_1,X_n) \\ \operatorname{Cov}(X_2,X_1) & \dots & & \vdots \\ \vdots & & \ddots & \\ \operatorname{Cov}(X_n,X_1) & \dots & & \operatorname{Cov}(X_n,X_n) \end{bmatrix}$$

nazywamy **macierzą kowariancji** zmiennej losowej X. Jest to wielowymiarowe uogólnienie wariancji.

<u>Uwaga</u> Jeżeli X_i są parami nieskorelowane, to Q diagonalne.

<u>TW. 5.16</u> Macierz kowariancji Q zm. los. X jest symetryczna oraz dodatnio określona (tzn. dla każdych $t_1, \dots, t_n$, $\sum t_i t_j q_{ij} \geq 0$).

<u>D-d</u> Ustalmy $t_1, \dots, t_n \in \mathbb{R}$. Definiujemy
$$Y = \sum_{i=1}^{n} t_i X_i .$$

$$0 \leq \operatorname{Var} Y = \mathbb{E}\left[(Y - \mathbb{E}Y)^2\right] =$$

$$= \mathbb{E}\left[\left(\sum t_j X_j - \sum t_j \mathbb{E}X_j\right)^2\right] =$$

$$= \mathbb{E}\left[\left(\sum t_j (X_j - \mathbb{E}X_j)\right)^2\right] =$$

$$\mathbb{E}\Big[\sum_{i,j} t_i (X_i - \mathbb{E}X_i) t_j (X_j - \mathbb{E}X_j)\Big] =$$

$$= \sum_{i,j} t_i t_j \, \mathbb{E}\big[(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)\big] =$$

$$= \sum_{i,j} t_i t_j \, \mathrm{Cov}(X_i, X_j) = \sum_{i,j} t_i t_j q_{ij}$$

$\blacksquare$

<u>Zadanie</u> X – wielowym. zm. los. $\sim \mathcal{N}(m, A^{-1})$, gdzie $m \in \mathbb{R}^d$, A jest macierzą sym. i nieujemnie określoną, tzn. gęstość X jest dana wzorem

$$f(x) = \frac{\sqrt{\det A}}{(2\pi)^{\frac{d}{2}}} e^{-\frac{1}{2}\langle x-m, x-m\rangle}$$

- Pokaż, że $\mathbb{E}X = m$, a macierz kow. jest równa A^{-1}
- Pokaż, że jeżeli X ma rozkład normalny, to $X_1, \dots, X_d$ są niezależne iff są nieskorelowane.

Przykład Regresja liniowa. Dane są zm. los. X i Y, są skorelowane.

$$f(x) = y = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}\, X}(x - \mathbb{E}X) + \mathbb{E}Y$$

Zał., że łatwo jest obserwować jedną z tych zm. los., np. X (łatwo jest mierzyć przebieg samochodu, trudno liczyć koszty eksploatacji). Szukamy funkcji f, która dobrze aproksymuje Y, to możemy przyjąć $Y = f(x)$.

Jak wyznaczyć ten wzór na f?

Chcemy zminimalizować $\underbrace{\mathbb{E}\left[(Y-(aX+b))^2\right]}_{=g(a,b)}$.

Szukamy a i b.

$$\begin{cases} \dfrac{\partial g}{\partial a} = 2a\mathbb{E}X^2 - 2\mathbb{E}XY + 2b\mathbb{E}X = 0 \\[2ex] \dfrac{\partial g}{\partial b} = 2b - 2\mathbb{E}Y + 2a\mathbb{E}X = 0 \end{cases}$$

Stąd $\begin{cases} a = \dfrac{\mathbb{E}XY - b\mathbb{E}X}{\mathbb{E}X^2} \\[3ex] b = \mathbb{E}Y - a\mathbb{E}X \end{cases}$

$$a\mathbb{E}X^2 - \mathbb{E}XY + \mathbb{E}Y\mathbb{E}X - a(\mathbb{E}X)^2 = 0$$

$$a \operatorname{Var} X - \operatorname{Cov}(X,Y) = 0$$

$$a = \frac{\operatorname{Cov}(X,Y)}{\operatorname{Var} X}$$

TW. 6.1 Jeżeli $\mathbb{E}X^2 < \infty$, $\mathbb{E}Y^2 < \infty$, to

$$\mathbb{E}|XY| \leq \left(\mathbb{E}X^2\right)^{\frac{1}{2}}\left(\mathbb{E}Y^2\right)^{\frac{1}{2}}$$

TW. 6.2 Nierówność Höldera: Jeżeli $\mathbb{E}|X|^p < \infty$ i $\mathbb{E}|Y|^q < \infty$, dla $p, q > 1$ t. że $\frac{1}{p} + \frac{1}{q} = 1$, to

$$\mathbb{E}|XY| \leq \left(\mathbb{E}|X|^p\right)^{\frac{1}{p}}\left(\mathbb{E}|Y|^q\right)^{\frac{1}{q}}$$

TW. 6.3 Nierówność Czebyszewa: Niech X będzie zm. los. i niech $f: [0, \infty) \to [0, \infty)$ będzie niemalejącą funkcją taką, że $f(x) > 0$ dla każdego $x > 0$. Wtedy dla każdego $\lambda > 0$ zachodzi

$$\mathbb{P}[|X| \geq \lambda] \leq \frac{\mathbb{E}f(|X|)}{f(\lambda)}$$

D-d. $\mathbb{P}[|X| \geq \lambda] = \mathbb{E}\, \mathbb{1}_{\{|X| \geq \lambda\}} \leq$

$$\leq \mathbb{E}\left[\frac{f(|X|)}{f(\lambda)} \cdot \mathbb{1}_{\{|X| \geq \lambda\}}\right] \leq \frac{1}{f(\lambda)} \mathbb{E}\left[f(|X|)\right]$$

Wniosek 8.4

- Nierówność Markowa $\left(f(x) = x^p \text{ dla } p > 0\right)$:

$$\mathbb{P}[|X| > \lambda] \leq \frac{\mathbb{E}|X|^p}{\lambda^p}, \quad \forall \lambda > 0$$

- Nierówność Czebyszewa $\left(f(x) = x^2,\right.$ zamiast X bierzemy $\left.X - \mathbb{E}X\right)$:

$$\mathbb{P}[|X - \mathbb{E}X| \geq \lambda] \leq \frac{\operatorname{Var} X}{\lambda^2}, \quad \forall \lambda > 0$$

- Wykładnicza nierówność Czebyszewa $\left(f(x) = e^{px}\right)$: Jeżeli $\mathbb{E}e^{pX} < \infty$, $p > 0$, to

$$\mathbb{P}[X \geq \lambda] \leq \frac{\mathbb{E}e^{pX}}{e^{p\lambda}}, \quad \forall \lambda > 0$$

Rodzaje zbieżności zmiennych losowych.

<u>Definicja 8.5</u> Załóżmy, że X_n jest ciągiem zm. losowych. Mówimy, że

1. X_n zbiega do X **prawie na pewno**, jeżeli $\mathbb{P}(\{\omega : X_n(\omega) \to X(\omega)\}) = 1$.

   Piszemy wówczas $X_n \to X$ (lub $X_n \xrightarrow{p.n.} X$)

2. X_n zbiega do X **według prawdopodobieństwa**, jeżeli dla każdego $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}[\{\omega : |X_n(\omega) - X(\omega)| \geq \varepsilon\}] = \lim_{n\to\infty} \mathbb{P}[|X_n - X| \geq \varepsilon] = 0$$

   Piszemy wówczas $X_n \xrightarrow{\mathbb{P}} X$

3. X_n zbiega do X w $\mathcal{L}^p$, jeżeli $X_n \in \mathcal{L}^p$ (tzn. $\|X_n\|_p = (\mathbb{E}|X_n|^p)^{\frac{1}{p}} < \infty$) oraz

$$\lim_{n\to\infty} \|X_n - X\|_p = \lim_{n\to\infty} (\mathbb{E}|X_n - X|^p)^{\frac{1}{p}} = 0$$

   Piszemy wówczas $X_n \xrightarrow{\mathcal{L}^p} X$.

TW. 6.6 Jeżeli $X_n \xrightarrow{p.n.} X$, $\varphi \xrightarrow{p.n.} \varphi$, to

1. $aX_n + b\varphi_n \xrightarrow{p.n.} aX + b\varphi$
2. $X_n \varphi_n \xrightarrow{p.n.} X\varphi$

TW. 6.7 Jeżeli $X_n \longrightarrow X$, to $X_n \xrightarrow{\mathbb{P}} X$. Implikacja odwrotna nie jest prawdziwa.

$\Rightarrow$ było na teorii miary

$\Leftarrow$ $(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1], Bor([0,1]), \lambda)$.

$X_1 = \mathbb{1}_{[0,1]}$, $X_2 = \mathbb{1}_{[0,\frac{1}{2}]}$, $X_3 = \mathbb{1}_{[\frac{1}{2},1]}$,

$X_4 = \mathbb{1}_{[0,\frac{1}{4}]}$, ..., $X_7 = \mathbb{1}_{[\frac{3}{4},1]}$, ...

Wtedy $X_n \xrightarrow{\mathbb{P}} 0$, ale nie $X \to 0$.

TW. 6.8 Jeżeli $X_n \xrightarrow{L^p} X$, to $X_n \xrightarrow{\mathbb{P}} X$. Implikacja odwrotna nie jest prawdziwa.

D-d. z nierówności Czebyszewa

$\Rightarrow$

$$\mathbb{P}[|X_n - X| < \varepsilon] \leq \frac{\mathbb{E}|X_n - X|^p}{\varepsilon^p} \longrightarrow 0$$

$\Leftarrow$ $X_n = n^{\frac{1}{p}} \cdot \mathbb{1}_{[0,\frac{1}{n}]}$. $X_n \xrightarrow{\mathbb{P}} 0$, $X_n \not\xrightarrow{L^p} 0$

TW. 6.9 Jeżeli $X_n \xrightarrow{\mathbb{P}} X$, to istnieje podciąg $\{X_{n_k}\}_{k \in \mathbb{N}}$ taki, że $X_{n_k} \xrightarrow{p.n.} X$.

D-d. Ustalmy $\varepsilon = 2^{-k}$.

$$\mathbb{P}[|X_n - X| \geq 2^{-k}] \leq 2^{-k} \quad (\text{dla } \underbrace{\text{dużych } n}_{n \geq n_k}).$$

Możemy założyć, że $\{n_k\}$ jest rosnący.

Mamy $\sum_{k=1}^{\infty} \mathbb{P}[|X_{n_k} - X| > 2^{-k}] \leq \sum_{k=1}^{\infty} 2^{-k} < \infty$

Z lematu Borela-Cantelliego wynika, że $\mathbb{P}[|X_{n_k} - X| \geq 2^{-k} \text{ i.o.}] = 0$.

Stąd wynika, że $|X_{n_k} - X| \geq 2^{-k}$ tylko dla skończenie wielu indeksów. Tzn. że od pewnego miejsca $|X_{n_k} - X| < 2^{-k}$

$\Rightarrow X_{n_k} \to X$.

Słabe Prawo Wielkich Liczb (SPWL)

TW. Dany jest ciąg niezależnych zmiennych losowych $\{X_n\}$ o takiej samej wartości oczekiwanej i wariancji. Wówczas $\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}X_1$.

Naszym celem będzie pokazanie Mocnego PWL, które mówi o zbieżności p.w.

TW. 7.1 Jeżeli ciąg $X_1, X_2, \ldots$ spełnia $\mathbb{E}X_i^2 < \infty$, zmienne losowe są nieskorelowane oraz mają wspólnie ograniczoną wariancję, to

$$\frac{X_1 + \ldots + X_n - \mathbb{E}[X_1 + \ldots + X_n]}{n} \xrightarrow{\mathbb{P}} 0$$

W szczególności, jeżeli wszystkie zmienne losowe X_i mają tę samą wartość oczekiwaną, to

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}X_1 .$$

D-d. Z nierówności Czebyszewa, dla $\varepsilon > 0$

$$\mathbb{P}\left[\frac{X_1 + \ldots + X_n - \mathbb{E}[X_1 + \ldots + X_n]}{n} \geqslant \varepsilon \right] \leqslant$$

$$1/\varepsilon^2 \cdot \operatorname{Var}\left(\frac{X_1 + \ldots + X_n}{n} \right) =$$

$$\frac{1}{(n\varepsilon)^2} \operatorname{Var}(X_1 + \ldots + X_n) \overset{\text{nieskorelowane}}{=}$$

$$\frac{1}{(n\varepsilon)^2} (\operatorname{Var} X_1 + \ldots + \operatorname{Var} X_n) \leqslant$$

$$\frac{M \cdot n}{(n\varepsilon)^2} \xrightarrow[n \to \infty]{} 0$$

Przykład. Weźmy $\{X_k\}$ ciąg nzl. zm. los. oraz że $X_k \sim \mathcal{U}([-1, 1])$. Wtedy $X_1^2, X_2^2, \ldots$ są niezależne. Ponadto

$$\mathbb{E}X_1^2 = \int_{-1}^{1} x^2 \frac{dx}{2} = \frac{1}{2} \cdot \frac{1}{3} x^3 \Big|_{-1}^{1} = 1/3 .$$

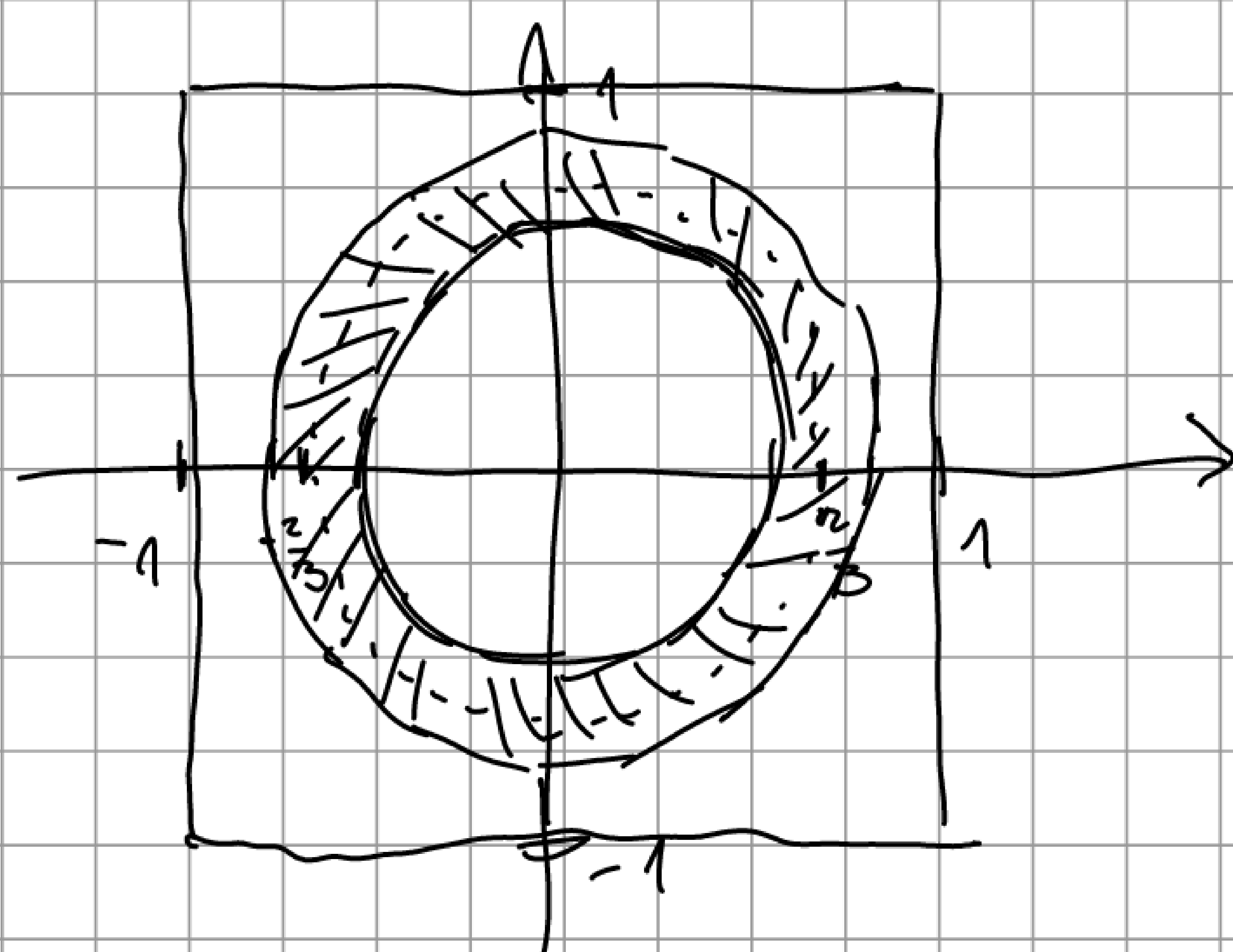
$$\operatorname{Var} X_1^2 \leqslant \mathbb{E} X_1^4 \leqslant 1 .$$

Ze SPWL mamy

$$\frac{X_1^2 + \ldots + X_n^2}{n} \xrightarrow{\mathbb{P}} 1/3 .$$

Ustalmy $\varepsilon \in (0, 1/2)$. Zdefiniujmy

$$A_{n,\varepsilon} = \{x \in \mathbb{R}^n : (1-\varepsilon)\sqrt{\tfrac{n}{3}} \leq \|x\| \leq (1+\varepsilon)\sqrt{\tfrac{n}{3}}\}$$



Dla $n \geq 4$ będzie tak, że ten pierścień będzie trochę wychodził poza kostkę, a trochę jeszcze w nim będzie.

Teror

$$\frac{1}{2^n}\cdot\lambda\left(A_{n,\varepsilon}\cap[-1,1]^n\right) =$$

$$= \mathbb{P}\left[(X_1,\dots,X_n)\in A_{n,\varepsilon}\right] =$$

$$= \mathbb{P}\left[(1-\varepsilon)\sqrt{\frac{n}{3}} \le \sqrt{X_1^2+\dots+X_n^2} \le (1+\varepsilon)\sqrt{\frac{n}{3}}\right] =$$

$$= \mathbb{P}\left[\frac{1}{3}(1-\varepsilon)^2 \le \frac{X_1^2+\dots+X_n^2}{n} \le \frac{1}{3}(1+\varepsilon)^2\right] \geq$$

$$\geq \mathbb{P}\left[\left|\frac{1}{n}\sum_{i=1}^{n}X_i^2 - \frac{1}{3}\right| \le \frac{1}{3}(2\varepsilon-\varepsilon^2)\right] \xrightarrow{\text{SPWL}} 1$$