Seminars 12 and 13 - 2025

Theoretical aspects:

Definition:
A sequence $(X_n)_n$ of random variables with $E|X_n| < \infty \; \forall\, n \in \mathbb{N}$, obeys the weak law of large numbers (WLLN), if

$$\frac{1}{n}\sum_{k=1}^{n}\Big(X_k - E(X_k)\Big) \xrightarrow{P} 0.$$

Definition:
A sequence $(X_n)_n$ of random variables with $E|X_n| < \infty, \forall\, n \geq 1$, obeys the strong law of large numbers (SLLN) if

$$\frac{1}{n}\sum_{k=1}^{n}\Big(X_k - E(X_k)\Big) \xrightarrow{a.s.} 0.$$

Theorem 1. *Let $(X_n)_n$ be a sequence of pairwise independent random variables satisfying the condition*

$$V(X_n) \leq L, \;\; \textit{for all } n \in \mathbb{N}^*,$$

where $L > 0$ is a constant. Then $(X_n)_n$ obeys the WLLN.

Theorem 2. *If $(X_n)_n$ is a sequence of independent random variables such that $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^2}V(X_n) < \infty$, then*

$$\frac{1}{n}\sum_{k=1}^{n}\Big(X_k - E(X_k)\Big) \xrightarrow{a.s.} 0,$$

i.e. $(X_n)_n$ obeys the SLLN.

Theorem 3. *Let $(X_n)_{n\geq 1}$ be a sequence of independent identically distributed random variables such that $E(X_n) = m$ for all $n \in \mathbb{N}$. Then*

$$\frac{1}{n}\sum_{k=1}^{n}X_k \xrightarrow{a.s.} m,$$

i.e. $(X_n)_{n\in\mathbb{N}}$ obeys the SLLN.

. . . .

1. Consider the sequence of independent identically distributed random variables $(X_n)_{n\geq 1}$ such that $X_n \sim Unif[1,3]$ for each $n \geq 1$. Compute the a.s. limit of the sequence which is
i) the arithmetic mean of $X_1, ..., X_n$, as $n \to \infty$;
ii) the geometric mean of $X_1, ..., X_n$, as $n \to \infty$;
iii) the harmonic mean of $X_1, ..., X_n$, as $n \to \infty$.

A: i) the SLLN (Th.3) $\Longrightarrow \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i \xrightarrow{a.s.} E(X_1) = \int_1^3 \frac{x}{2}\,dx = 2.$

ii) the SLLN (Th.3) $\Longrightarrow \sqrt[n]{\displaystyle\prod_{i=1}^{n}X_i} = e^{\frac{1}{n}\sum_{i=1}^{n}\ln X_i} \xrightarrow{a.s.} e^{E(\ln X_1)} = e^{\int_1^3 \frac{\ln x}{2}dx} = \dfrac{3\sqrt{3}}{e} \approx 1,91.$

iii) the SLLN (Th.3) $\Longrightarrow \dfrac{n}{\sum_{i=1}^{n}\frac{1}{X_i}} = \dfrac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i}} \xrightarrow{a.s.} \dfrac{1}{E(\frac{1}{X_1})} = \dfrac{1}{\int_1^3 \frac{1}{2x}dx} = \dfrac{2}{\ln 3} \approx 1{,}82.$

2. Let $(X_n)_{n\geq 1}$ be a sequence of random variables such that $P(X_n = n^2) = \frac{1}{n}$ and $P(X_n = 0) = 1 - \frac{1}{n}$, for all $n \geq 1$. Prove that:
a) $X_n \xrightarrow{P} 0$.
b) $(X_n)_{n\geq 1}$ does not converge in mean square.
A: a) For every $\varepsilon > 0$,

$$\lim_{n\to\infty} P(|X_n| \geq \varepsilon) = \lim_{n\to\infty} P(X_n = n^2) = \lim_{n\to\infty} \frac{1}{n} = 0.$$

b) We use the proof by contradiction: Assume that $(X_n)_{n\geq 1}$ converges in mean square. Then, by a theorem from the course and a), $X_n \xrightarrow{L^2} 0$. Since $E|X_n|^2 = \frac{n^4}{n} = n^3 \to \infty$, as $n \to \infty$, we get a contradiction. Hence, $(X_n)_{n\geq 1}$ does not converge in mean square.

3. A bank cashier serves customers in the queue one by one. It is known that the expected service time for each customer is 3 minutes, with a standard deviation of 2 minutes. We assume that the service times for the bank customers are independent. Let T be the total time the bank cashier spends serving 100 customers. Estimate the probability $P(240 < T < 320)$ by using values from the table below.

Hint: Let F denote the cdf of the $N(0, 1)$ distribution. In the table below there are computed the values $F(x)$ for $x \in \{-3, -2, -1, 0, 1, 2, 3\}$ in Python with `scipy.stats.norm.cdf`$(x, 0, 1)$

x	-3	-2	-1	0	1	2	3
$F(x)$	0.00135	0.02275	0.15866	0.5	0.84134	0.97725	0.99865

A: Denote X_i the service time for the i^{th} client, $i = \overline{1, 100}$. We have $\mu = E(X_i) = 3$, $\sigma = \sqrt{V(X_i)} = 2$, $i = \overline{1, 100}$, and $T = X_1 + ... + X_{100}$.

$$P(240 < T < 320) = P\left(\frac{240 - 100 \cdot 3}{2 \cdot \sqrt{100}} < \frac{(X_1 + ... + X_{100}) - 100 \cdot \mu}{2 \cdot \sqrt{100}} < \frac{320 - 100 \cdot 3}{2 \cdot \sqrt{100}}\right)$$
$$P\left(-3 < \frac{(X_1 + ... + X_{100}) - 100 \cdot \mu}{2 \cdot \sqrt{100}} < 1\right) \overset{CLT}{\approx} F(1) - F(-3) = 0.84134 - 0.00135 = 0.0.83999\,.$$

$$\Longrightarrow P(35 < X_1 + ... + X_{100} < 65) \approx 0.9973\,.$$

4. If $(X_n)_n$ is a sequence of independent normally distributed random variables such that $X_n \sim N(0, \frac{1}{n})$, for each $n \geq 1$. Prove that $(X_n)_n$ obeys the SLLN.

A: $X_n \sim N(0, \frac{1}{n}) \Rightarrow E(X_n) = 0$ and $V(X_n) = \frac{1}{n}$. We use Theorem 2 (the random variables from the sequence $(X_n)_n$ are *not* identically distributed).

We have $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} V(X_n) = \sum_{n=1}^{\infty} \frac{1}{n^3} < \infty$ (result from Analysis). Then by Theorem 1 it follows

$$\frac{1}{n} \sum_{k=1}^{n} \Big(X_k - 0\Big) \overset{a.s.}{\to} 0,$$

i.e. $(X_n)_n$ obeys the SLLN.

5. The measurement error (in millimeters) of a certain object produced in a factory is a continuous random variable X with the cumulative distribution function $F : \mathbb{R} \to [0,1]$,

$$F(x) = \begin{cases} 0, & x < -1 \\ \frac{1}{4}\left(2 + 3x - x^3\right), & x \in [-1,1] \\ 1, & x > 1. \end{cases}$$

Find: $P(-\frac{1}{2} < X < \frac{1}{2})$, $P(X < \frac{1}{2} | X > -\frac{1}{2})$, $E(X)$.

A: $P(-\frac{1}{2} < X < \frac{1}{2}) = F(\frac{1}{2}) - F(-\frac{1}{2}) = \frac{27}{32} - \frac{5}{32} = \frac{22}{32} = \frac{11}{16}$. $P(X < \frac{1}{2} | X > -\frac{1}{2}) = \frac{P(-\frac{1}{2} < X < \frac{1}{2})}{P(X > -\frac{1}{2})} = \frac{\frac{22}{32}}{1 - \frac{5}{32}} = \frac{22}{27}$.

$f(x) = \begin{cases} \frac{3}{4}(1 - x^2), & x \in [-1,1] \\ 0, & x \notin [-1,1] \end{cases} \implies E(X) = \int_{-1}^{1} \frac{3}{4}(x - x^3)dx = 0.$

6. A random value X is generated according to the density function $f_X : \mathbb{R} \to \mathbb{R}$, $f(x) = \dfrac{1}{2}e^{-|x|}$, for all $x \in \mathbb{R}$.

Compute:

a) the cumulative distribution function of X;

b) the cumulative distribution function of the random value X^2;

c) $P(X^2 \geq 1)$;

d) the mean value and the variance of X.

A: a) $F_X(x) = \begin{cases} \frac{1}{2}\int_{-\infty}^{x} e^t dt, & x < 0 \\ \frac{1}{2}\int_{-\infty}^{0} e^{-t}dt + \frac{1}{2}\int_{0}^{x} e^{-t}dt, & x \geq 0 \end{cases} = \begin{cases} \frac{e^x}{2}, & x < 0 \\ \frac{1}{2} + \frac{1 - e^{-x}}{2}, & x \geq 0 \end{cases} = \begin{cases} \frac{e^x}{2}, & x < 0 \\ 1 - \frac{e^{-x}}{2}, & x \geq 0 \end{cases}.$

b) $F_{X^2}(x) = P(X^2 \leq x) = \begin{cases} 0, & x < 0 \\ F(\sqrt{x}) - F(-\sqrt{x}), & x \geq 0 \end{cases} = \begin{cases} 0, & x < 0 \\ 1 - \frac{e^{-\sqrt{x}}}{2} - \frac{e^{-\sqrt{x}}}{2}, & x \geq 0 \end{cases} = \begin{cases} 0, & x < 0 \\ 1 - e^{-\sqrt{x}}, & x \geq 0 \end{cases}.$

c) $P(X^2 \geq 1) = 1 - F_{X^2}(1) = \frac{1}{e}$. d) $E(X) = \frac{1}{2}\int_{-\infty}^{\infty} xe^{-|x|}dx = 0$ (we integrate an odd function on a symmetric interval), $V(X) = E(X^2) - E^2(X) = \frac{1}{2}\int_{-\infty}^{\infty} x^2 e^{-|x|}dx = \int_{0}^{\infty} x^2 e^{-x}dx = -x^2e^{-x}\Big|_0^\infty - 2xe^{-x}\Big|_0^\infty - 2e^{-x}\Big|_0^\infty = 2$.

Another solution is to find first a density function for X^2: $f_{X^2}(y) = \begin{cases} 0, & y < 0 \\ (1 - e^{-\sqrt{y}})', & y \geq 0 \end{cases} = \begin{cases} 0, & y < 0 \\ \frac{e^{-\sqrt{y}}}{2\sqrt{y}}, & y \geq 0 \end{cases}.$

Then, if $Y = X^2$, $E(Y) = \int_0^\infty y \frac{e^{-\sqrt{y}}}{2\sqrt{y}}dy \overset{y=x^2}{=} \int_0^\infty x^2 e^{-x}dx = 2$.

7. For each $n \in \mathbb{N}, n \geq 2$, consider

$$X_n \sim \begin{pmatrix} -1 & 1 \\ \frac{1}{n} & 1 - \frac{1}{n} \end{pmatrix}$$

such that $(X_n)_{n \geq 2}$ is a sequence of pairwise independent random variables.

(a) Does $(X_n)_{n \geq 2}$ obey the weak law of large numbers?

(b) Compute $\lim\limits_{n \to \infty} V\left(\frac{1}{2}(X_{n-1} + X_n)\right)$.

A: (a) $V(X_n) = E(X_n^2) - (E(X_n))^2 = 1 - \left(1 - \frac{2}{n}\right)^2 = \frac{4}{n} - \frac{4}{n^2} \leq 4$.

$(X_n)_{n \geq 2}$ is a sequence of pairwise independent random variables, we use Theorem 1 to deduce that $(X_n)_{n \geq 2}$ obeys the weak law of large numbers.

(b) By the independence property $\Rightarrow V\left(\frac{1}{2}(X_{n-1} + X_n)\right) = \frac{1}{4}(V(X_{n-1}) + V(X_n)) = \frac{1}{n-1} - \frac{1}{(n-1)^2} + \frac{1}{n} - \frac{1}{n^2}$.

Therefore, $\lim\limits_{n \to \infty} V\left(\frac{1}{2}(X_{n-1} + X_n)\right) = 0$.

8. Consider a binary communication channel transmitting codes of n bits each. Assume that the probability of successful transmission of a single bit is $p \in (0, 1)$ and that the probability of an error is $1 - p$. Assume also that the channel is capable of correcting up to m errors, where $0 < m < n$. If we assume that the transmission of successive bits is independent, compute the probability of successful code transmission.

A: Let X be the number of number of errors in the code. The event A:"the code is transmitted with all errors corrected" is equivalent with $\{X \leq m\}$. Since $X \sim Bino(n, 1 - p)$,

$$P(A) = P(X \leq m) = \sum_{k=0}^{m} C_n^k p^{n-k}(1 - p)^k.$$