DPT Exercise 8

github.com/aiulus/dpt-ss25

START: 12:15

Contents:
- Estimating prob.'s:
  - > Markov's inequality
  - > Chebyshev's inequality
  - CLT, WLLN
- Convergence of RVRV's
- Characteristic functions

DPT Exercises

Exercise sheet 8 Niki Kilbertus

Exercise 1 (heavy planes)**.** A plane that can transport at most 100 passengers is filled with just enough jet fuel to carry up to 8 tons. For a fully booked flight, let X_i be a random variable modeling the i-th passengers weight in kg.

(i) Write down one (convincing) argument for why the $(X_i)_{i=1}^{100}$ should be modeled as independent or not as independent. From now on, assume they are iid with $\mathbb{E}[X_i] = 75$, and $\sigma[X_i] = 20$ for all $i \in [100]$.

(ii) Can you compute the exact probability that the maximum weight capacity of 8 tons is exceeded? Either compute it or explain why you can't?

7

(iii) Use the clt to compute an approximate probability that the maximum weight capacity will be exceeded.
[Hint: You can use wolframalpha or your favorite programming language/calculator to evaluate the cdf of standard Gaussians $\Phi(\cdot)$. Instead, you can also keep the results in terms of $\Phi(\cdot)$.]

(iv) If the profit margin per passenger is 3 Euros and the cost of going over capacity is 3000 Euros per flight (for topping up jet fuel), is the plane on average profitable for the airline (assuming it's always fully booked)?

(v) **[Optional:]** If you were working for the airline trying to maximize profits. How would you improve the model (e.g., model how fully planes are booked, potential dependencies, varying ticket costs, etc) and what would you optimize? (This is a purely imaginative exercise, no need to do any computations. Just train your "modeling and assumptions muscles.")

Exercise 2 (insurance claims and risk)**.** An insurance company has $n = 10^4$ clients. For each client i, let X_i be the amount claimed in a year. Assume the claims $X_1, \dots, X_n$ are iid RVRVs with mean $\mathbb{E}[X_i] = \mu = 500$ and variance $\mathrm{var}[X_i] = \sigma^2 = 800^2 = 640,000$. The company charges each client a premium of Z Euros.

(i) Let $S_n = \sum_{i=1}^n X_i$ be the total amount claimed. Calculate the minimum premium Z the company needs to charge to make an expected profit of half a million per year.

(ii) Use the weak law of large numbers to describe the behavior of the average claim per client $\bar{X}_n = S_n/n$ as n becomes very large.

(iii) The company wants to be at least 95% sure that the total premiums collected will cover the total claims S_n. Use Chebyshev's inequality to assess if the current number of clients $n = 10^4$ is sufficient with a premium of $Z = 550$.
[Hint: The company makes a loss if $S_n > n \cdot Z$. You want to bound $P(S_n > n \cdot Z)$.]

(iv) Now, use the clt to approximate the probability that the total claims S_n exceed the total premiums $n \cdot Z$. Compare this approximation with the bound obtained from Chebyshev's inequality. Why might the clt provide a different (and potentially more accurate) value for this probability? Which would you trust more?
[Hint: We need $P(S_n > n \cdot Z)$, which is easiest when S_n is standardized. You can use wolframalpha or your favorite programming language/calculator to evaluate the cdf of standard Gaussians.]

Exercise 3. Let X, Y RVRVs and $(X_n)_{n\in\mathbb{N}}, (Y_n)_{n\in\mathbb{N}}$ sequences of RVRVs. Verify the following statements.

(i) If $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$ then $X_n + Y_n \xrightarrow{P} X + Y$.
[Hint: Use the triangle inequality $|(X_n - X) + (Y_n - Y)| \leq |X_n - X| + |Y_n - Y|$. Decompose the probability to be bounded by $P(A \cup B) \leq P(A) + P(B)$ (this is called a *union bound* where A is an event only about X_n, X and B is only about Y_n, Y.]

(ii) If $\mathbb{E}[|X_n - X|] \to 0$, then $X_n \xrightarrow{P} X$.
[Hint: Use Markov's inequality for $|X_n - X|$.]

(iii) If $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{P} Y$ then $P(X = Y) = 1$, i.e., X and Y are equal "almost surely."
[Hint: Again use the triangle inequality $|X - Y| \leq |X - X_n| + |X_n - Y|$, decompose $P(|X - Y| \geq \epsilon)$ via a union bound (for any n), to get to $P(|X - Y| \geq \epsilon) = 0$ for any $\epsilon > 0$ and use a limiting argument to get the statement also for $\epsilon = 0$ (for example via σ-subadditivity.]

(iv) If $\mathbb{E}[|X_n - X|] \to 0$ and $\mathbb{E}[|X_n - Y|] \to 0$ then $P(X = Y) = 1$.

Exercise 4 (convergence of normal random variables)**.** Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of RVRVs such that $X_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$ for each $n \in \mathbb{N}$.

(i) Show that if $\mu_n \to \mu \in \mathbb{R}$ and $\sigma_n^2 \to \sigma^2 > 0$ as $n \to \infty$, then $X_n \xrightarrow{d} X$, where $X \sim \mathcal{N}(\mu, \sigma^2)$.
[Hint: Use characteristic functions with the uniqueness and continuity properties. The cf of $\mathcal{N}(m, s^2)$ is $\phi(t) = \exp(imt - \frac{1}{2}s^2t^2)$.]

(ii) Show that if $\mu_n \to \mu \in \mathbb{R}$ and $\sigma_n^2 \to 0$ as $n \to \infty$, then $X_n \xrightarrow{P} \mu$.
[Hint: You can use either Chebyshev's inequality or the fact that convergence in distribution to a constant implies convergence in probability to that constant.]

Exercise 5 (**bonus:** basic properties of ϕ_X)**.** Let X and $(X_i)_{i=1}^n$ be RVRVs. Verify the following properties of the characteristic function.

(i) $|\phi_X(t)| \leq 1$ for all $t \in \mathbb{R}$ and $\phi_X(0) = 1$.

(ii) $\phi_{aX+b}(t) = e^{ibt}\phi_X(at)$ for all $a \in \mathbb{R} \setminus \{0\}, b \in \mathbb{R}$.

(iii) We have that $P_X = P_{-X}$ if and only if ϕ_X is real-valued.

(iv) If $X_1, \ldots, X_n$ are independent, then

$$\phi_{X_1 + \cdots + X_n}(t) = \phi_{X_1}(t) \cdots \phi_{X_n}(t) \,.$$

Exercise 1 (heavy planes)**.** A plane that can transport at most 100 passengers is filled with just enough jet fuel to carry up to 8 tons. For a fully booked flight, let X_i be a random variable modeling the i-th passengers weight in kg.

(i) Write down one (convincing) argument for why the $(X_i)_{i=1}^{100}$ should be modeled as independent or not as independent. From now on, assume they are iid with $\mathbb{E}[X_i] = 75$, and $\sigma[X_i] = 20$ for all $i \in [100]$.

(ii) Can you compute the exact probability that the maximum weight capacity of 8 tons is exceeded? Either compute it or explain why you can't?

$P(\sum_{j=1}^{100} X_j \geq 8000)$ not exactly computable w/o P_{X_i}

(iii) Use the clt to compute an approximate probability that the maximum weight capacity will be exceeded.

Theorem 2.64 (central limit theorem (clt)). *Let $X_1, X_2, \ldots$ be a sequence of iid RVRVs with finite mean $\mathbb{E}[X_i] = \mu$ and finite, non-zero variance $\mathrm{var}[X_i] = \sigma^2 \in \mathbb{R}_{>0}$. Let $S_n := \sum_{i=1}^n X_i$ be the sum, and $\bar{X}_n = S_n/n$ be the sample mean. Consider the* standardized sum *(or* standardized mean*):*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathrm{var}[S_n]}} = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} .$$

Then

$$Z_n \xrightarrow{d} \mathcal{N}(0,1) .$$

$$S_{100} := \sum_{j=1}^{100} X_j$$

$$\mathbb{E}[S_{100}] = \mathbb{E}\left[\sum_{j=1}^{100} X_j\right] = \sum_{j=1}^{100} \underbrace{\mathbb{E}[X_j]}_{=\mu \ \forall j \in [100]} = 100 \cdot 75$$

$$\mathrm{Var}[S_{100}] = \mathrm{Var}\left[\sum_{j=1}^{100} X_j\right] = \sum_{j=1}^{100} \underbrace{\mathrm{Var}[X_j]}_{=\sigma^2 = 20^2 \ \forall j}$$

$$= 100 \cdot 20^2$$

$$\sqrt{\mathrm{Var}[S_{100}]} = 200$$

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathrm{Var}[S_n]}} = \frac{S_n - 7500}{200} \xrightarrow[\text{CLT}]{d} \mathcal{N}(0,1)$$

$$P\left(\underbrace{\sum_{j=1}^{100} X_j}\geq 8000\right) = P(S_n \geq 8000)$$

$$= P\left(\frac{S_n - 7500}{200} := Z_n \geq \frac{8000 - 7500}{200}\right)$$

$$= P(Z_n \geq 2.5)$$

continuous RV!

$$= 1 - P(Z_n \leq 2.5)$$

$$= 1 - \Phi(2.5) \approx 0.00621$$

(iv) If the profit margin per passenger is 3 Euros and the cost of going over capacity is 3000 Euros per flight (for topping up jet fuel), is the plane on average profitable for the airline (assuming it's always fully booked)?

(v) **[Optional:]** If you were working for the airline trying to maximize profits. How would you improve the model (e.g., model how fully planes are booked, potential dependencies, varying ticket costs, etc) and what would you optimize? (This is a purely imaginative exercise, no need to do any computations. Just train your "modeling and assumptions muscles.")

(iv) $3 \cdot 100 - 3000 \cdot 0{,}00121 = 300 - 18.13 > 0$

(v)

Exercise 2 (insurance claims and risk)**.** An insurance company has $n = 10^4$ clients. For each client i, let X_i be the amount claimed in a year. Assume the claims $X_1, \ldots, X_n$ are iid RVRVs with mean $\mathbb{E}[X_i] = \mu = 500$ and variance $\mathrm{var}[X_i] = \sigma^2 = 800^2 = 640,000$. The company charges each client a premium of Z Euros.

(i) Let $S_n = \sum_{i=1}^n X_i$ be the total amount claimed. Calculate the minimum premium Z the company needs to charge to make an expected profit of half a million per year.

$$\text{Profit} = \text{Income} - \text{Costs} = n \cdot Z - S_n$$

$$\mathbb{E}[\text{"Profit"}] = \mathbb{E}[n \cdot Z - S_n]$$

$$= n \cdot Z - n \cdot \mu \overset{!}{=} 500\,000$$

$$\Leftrightarrow Z = \frac{5 \cdot 10^5 + 5 \cdot 10^6}{10^4} = 550$$

$$\mathbb{E}[S_n] = \mathbb{E}\left[\sum_{j=1}^n X_j\right] = n \cdot \mu = 5 \cdot 10^6$$

$$\mathrm{Var}[S_n] = \mathrm{Var}\left[\sum_{j=1}^n X_j\right] \overset{iid}{=} n \cdot \sigma^2 = 10^4 \cdot 800^2$$

$$= 64 \cdot 10^8$$

Exercise 2 (insurance claims and risk). An insurance company has $n = 10^4$ clients. For each client i, let X_i be the amount claimed in a year. Assume the claims $X_1, \ldots, X_n$ are iid RVRVs with mean $\mathbb{E}[X_i] = \mu = 500$ and variance $\mathrm{var}[X_i] = \sigma^2 = 800^2 = 640,000$. The company charges each client a premium of Z Euros.

(ii) Use the weak law of large numbers to describe the behavior of the average claim per client $\bar{X}_n = S_n/n$ as n becomes very large.

Theorem 2.61 (weak law of large numbers (wlln)). *Let $X_1, X_2, \ldots$ be a sequence of identically distributed RVRVs with finite mean $\mathbb{E}[X_i] = \mu < \infty$ that are pairwise uncorrelated. Let $\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ be the sample mean. Then,*

$$\bar{X}_n \xrightarrow{P} \mu\,.$$

$$\bar{X}_n \xrightarrow{P} 500 \quad \text{as } n \to \infty$$

$$\left(:\Leftrightarrow \quad P(|\bar{X}_n - 500| \geq \varepsilon) = 0 \quad \text{as } n \to \infty \quad \text{for arbitrary } \varepsilon \in \mathbb{R}_{>0} \right)$$

(iii) The company wants to be at least 95% sure that the total premiums collected will cover the total claims S_n. Use Chebyshev's inequality to assess if the current number of clients $n = 10^4$ is sufficient with a premium of $Z = 550$.
[Hint: The company makes a loss if $S_n > n \cdot Z$. You want to bound $P(S_n > n \cdot Z)$.]

$$\mathbb{E}[S_n] = \mathbb{E}\left[\sum_{j=1}^{n} X_j\right] = n \cdot \mu = 5 \cdot 10^6$$

$$\mathrm{Var}[S_n] = \mathrm{Var}\left[\sum_{j=1}^{n} X_j\right] \overset{iid}{=} n \cdot \sigma^2 = 10^4 \cdot 800^2 = 64 \cdot 10^8$$

Theorem 2.40 (Chebyshev's inequality). *Let X be an RVRV with finite mean $\mathbb{E}[X]$ and finite variance $\mathrm{var}[X] > 0$. Then for any $\epsilon > 0$,*

$$P(|X - \mathbb{E}[X]| \geq \epsilon) \leq \frac{\mathrm{var}[X]}{\epsilon^2}.$$

$$P(n \cdot Z \geq S_n) \geq 0.95 \iff P(S_n \leq n \cdot Z) \geq 0.95$$

$$\iff 1 - P(S_n \geq n \cdot Z) \geq 0.95$$

$$\iff P(S_n \geq n \cdot Z) \leq 0.05$$

$$\iff P(S_n - \mathbb{E}[S_n] \geq \underbrace{10^4 \cdot 550 - 5 \cdot 10^6}_{= 5{,}5 \cdot 10^6 - 5 \cdot 10^6 = 5 \cdot 10^5})$$

$$\leq P(|S_n - \mathbb{E}[S_n]| \geq 5 \cdot 10^5) \underset{\text{Chebyshev}}{\leq} \frac{\mathrm{Var}[S_n]}{\epsilon^2} = \frac{64 \cdot 10^8}{(5 \cdot 10^5)^2} = \left(\frac{8 \cdot 10^4}{5 \cdot 10^5}\right)^2 = 0.16^2 = 0.0256$$

(iv) Now, use the clt to approximate the probability that the total claims S_n exceed the total premiums $n \cdot Z$. Compare this approximation with the bound obtained from Chebyshev's inequality. Why might the clt provide a different (and potentially more accurate) value for this probability? Which would you trust more?

Theorem 2.64 (central limit theorem (clt)). *Let $X_1, X_2, \ldots$ be a sequence of iid RVRVs with finite mean $\mathbb{E}[X_i] = \mu$ and finite, non-zero variance $\mathrm{var}[X_i] = \sigma^2 \in \mathbb{R}_{>0}$. Let $S_n := \sum_{i=1}^n X_i$ be the sum, and $\bar{X}_n = S_n/n$ be the sample mean. Consider the standardized sum (or standardized mean):*

$$Z_n = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathrm{var}[S_n]}} = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} .$$

Then

$$Z_n \xrightarrow{d} \mathcal{N}(0,1) .$$

$$\mathbb{E}[S_n] = \mathbb{E}\left[\sum_{j=1}^n X_j\right] = n \cdot \mu = 5 \cdot 10^6 := \mu_n$$

$$\mathrm{Var}[S_n] = \mathrm{Var}\left[\sum_{j=1}^n X_j\right] \overset{iid}{=} n \cdot \sigma^2 = 64 \cdot 10^8$$

$$\sigma_n = \sqrt{\mathrm{Var}[S_n]} = 8 \cdot 10^4$$

$$Z_n = \frac{S_n - 5 \cdot 10^6}{8 \cdot 10^4}$$

$$P(S_n \geq n \cdot Z) = P\left(Z_n \geq \underbrace{\frac{10^4 \cdot 550 - 5 \cdot 10^6}{8 \cdot 10^4}}_{\frac{5.5 \cdot 10^4 - 5 \cdot 10^6}{8 \cdot 10^4} = \frac{50}{8} = 6.25}\right) = 1 - \Phi(6.25) \approx \underline{0.00621}$$

Exercise 3. Let X, Y RVRVs and $(X_n)_{n\in\mathbb{N}}, (Y_n)_{n\in\mathbb{N}}$ sequences of RVRVs. Verify the following statements.

(i) If $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$ then $X_n + Y_n \xrightarrow{P} X + Y$.
[Hint: Use the triangle inequality $|(X_n - X) + (Y_n - Y)| \leq |X_n - X| + |Y_n - Y|$. Decompose the probability to be bounded by $P(A \cup B) \leq P(A) + P(B)$ (this is called a *union bound* where A is an event only about X_n, X and B is only about Y_n, Y.]

(i) *converges in probability* to X, denoted $X_n \xrightarrow{P} X$, if for every $\epsilon > 0$,

$$\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0\,.$$

Intuitively, this means that for large n, it is very unlikely that X_n is far from X.

t.s.: $P(|X_n + Y_n - (X+Y)| \geq \varepsilon) = 0$ as $n \to \infty$

1. triangle ineq: $|X_n + Y_n - (X+Y)| = |(X_n - X) + (Y_n - Y)| \leq |X_n - X| + |Y_n - Y|$

2. $|X_n - X| + |Y_n - Y| \geq |X_n + Y_n - (X+Y)| \geq \varepsilon$ $\quad Y \quad X$

event 2 $\subseteq$ event 1

$Y \subseteq X$

$|X_n - X| \geq \varepsilon/2$ OR $|Y_n - Y| \geq \varepsilon/2$

3. $\{|(X_n + Y_n) - (X+Y)| \geq \varepsilon\} \subseteq \{|X_n - X| + |Y_n - Y| \geq \varepsilon\} = \{|X_n - X| \geq \varepsilon/2\} \cup \{|Y_n - Y| \geq \varepsilon/2\}$

4. Union bound: $P(|(X_n + Y_n) - (X+Y)| \geq \varepsilon) \leq \underbrace{P(|X_n - X| \geq \varepsilon/2)}_{\to 0} + \underbrace{P(|Y_n - Y| \geq \varepsilon/2)}_{\to 0} = 0$ as $n \to \infty$

5. $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$, so:

Exercise 3. Let X, Y RVRVs and $(X_n)_{n\in\mathbb{N}}, (Y_n)_{n\in\mathbb{N}}$ sequences of RVRVs. Verify the following statements.

(ii) If $\mathbb{E}[|X_n - X|] \to 0$, then $X_n \xrightarrow{P} X$.
[Hint: Use Markov's inequality for $|X_n - X|$.]

Theorem 2.38 (Markov's inequality). *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $X : \Omega \to \mathbb{R}$ be an RVRV (whose density exists and is integrable) and let $f : S_X \to \mathbb{R}_{\geq 0}$ be monotonically increasing, where $S_X = \mathrm{im}(X) := X(\Omega) \subseteq \mathbb{R}$. Then for each $\epsilon > 0$ with $f(\epsilon) > 0$ we have*

$$P(X \geq \epsilon) \leq \frac{1}{f(\epsilon)} \mathbb{E}[f(X)] .$$

In the literature, the Markov inequality is often shown only as a special case: Let X be a non-negative RVRV (i.e., $P(X \geq 0) = 1$). Then for any constant $\epsilon > 0$,

$$P(X \geq \epsilon) \leq \frac{\mathbb{E}[X]}{\epsilon} .$$

$\forall\, \varepsilon > 0$

t.s.: $\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$ for any $\varepsilon \in \mathbb{R}_{>0}$.

t.s.: $\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$ for any $\varepsilon \in \mathbb{R}_{>0}$.

$$\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) \overset{\text{M.I.}}{\leq} \lim_{n\to\infty} \frac{\overbrace{\mathbb{E}[|X_n - X|]}^{\to 0 \text{ by assumption}}}{\varepsilon > 0} = 0$$

Exercise 3. Let X, Y RVRVs and $(X_n)_{n\in\mathbb{N}}, (Y_n)_{n\in\mathbb{N}}$ sequences of RVRVs. Verify the following statements.

(iii) If $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{P} Y$ then $P(X = Y) = 1$, i.e., X and Y are equal "almost surely."
[Hint: Again use the triangle inequality $|X-Y| \leq |X-X_n|+|X_n-Y|$, decompose $P(|X-Y| \geq \epsilon)$ via a union bound (for any n), to get to $P(|X-Y| \geq \epsilon) = 0$ for any $\epsilon > 0$ and use a limiting argument to get the statement also for $\epsilon = 0$ (for example via σ-subadditivity.]

1. Same analysis as in (i) – consider the event $\{|X-Y| \geq \epsilon\}$
   - triangle ineq. ; $\epsilon/2$; union bound

2. Consider the event $\{X \neq Y\} = \bigcup_{k=1}^{\infty} \{|X-Y| \geq 1/k\}$

3. ... $P(X \neq Y) = 0$

Exercise 3. Let X, Y RVRVs and $(X_n)_{n\in\mathbb{N}}, (Y_n)_{n\in\mathbb{N}}$ sequences of RVRVs. Verify the following statements.

(iv) If $\mathbb{E}[|X_n - X|] \to 0$ and $\mathbb{E}[|X_n - Y|] \to 0$ then $P(X = Y) = 1$.

follows directly from (ii)

Exercise 4 (convergence of normal random variables). Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of RVRVs such that $X_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$ for each $n \in \mathbb{N}$.

(i) Show that if $\mu_n \to \mu \in \mathbb{R}$ and $\sigma_n^2 \to \sigma^2 > 0$ as $n \to \infty$, then $X_n \xrightarrow{d} X$, where $X \sim \mathcal{N}(\mu, \sigma^2)$.
[Hint: Use characteristic functions with the uniqueness and continuity properties. The cf of $\mathcal{N}(m, s^2)$ is $\phi(t) = \exp(imt - \frac{1}{2}s^2t^2)$.]

$t \cdot \frac{1}{7}$:

$$\lim_{n\to\infty} \phi_{X_n}(t) = \phi_X(t) \; ; \; X \sim N(\mu, \sigma^2)$$

Example 2.48 (cf of normal distribution). The characteristic function of $X \sim \mathcal{N}(\mu, \sigma^2)$ is

$$\phi_X(t) = \mathbb{E}[e^{itX}] = \exp\left(i\mu t - \frac{1}{2}\sigma^2 t^2\right).$$

$$\phi_{X_n}(t) = \exp\left(i\mu_n t - \frac{1}{2}\sigma_n^2 t^2\right) \; ; \; \lim_{n\to\infty} \phi_{X_n}(t) = \lim_{n\to\infty} \exp\left(i\mu_n t - \frac{1}{2}\sigma_n^2 t^2\right) = \exp\left(i\mu t - \frac{1}{2}\sigma^2 t\right) = \phi_X(t)$$

$\mu_n \to \mu$, $\sigma_n^2 \to \sigma$ as $n \to \infty$

1:1 distr $\Leftrightarrow$ pdf. / cdf

$\updownarrow$

cf.

Exercise 4 (convergence of normal random variables). Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of RVRVs such that $X_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$ for each $n \in \mathbb{N}$.

(ii) Show that if $\mu_n \to \mu \in \mathbb{R}$ and $\sigma_n^2 \to 0$ as $n \to \infty$, then $X_n \xrightarrow{P} \mu$.

[Hint: You can use either Chebyshev's inequality or the fact that convergence in distribution to a constant implies convergence in probability to that constant.]

t.s.: $\lim_{n\to\infty} P(|X_n - \mu| \geq \varepsilon) = 0$ for arbitrary $\varepsilon > 0$.

aybuke.uluusarslan@tum.de

1. Consider $Y_n := X_n - \mu_n$; $Y_n \sim \mathcal{N}(0, \sigma_2^2)$

2. Use $X_n - \mu = (X_n - \mu_n) + (\mu_n - \mu)$

3. Consider $P(|X_n - \mu| \geq \varepsilon)$ + Chebyshev

Definition 2.56 (modes of convergence). Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of RVRV and let X be another RVRV (which can also be a constant), all defined on the same probability space $(\Omega, \mathcal{A}, P)$. The sequence (X_n)

(i) *converges in probability* to X, denoted $X_n \xrightarrow{P} X$, if for every $\epsilon > 0$,

$$\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0\,.$$

Intuitively, this means that for large n, it is very unlikely that X_n is far from X.

(ii) *converges almost surely* to X, denoted $X_n \xrightarrow{a.s.} X$ or $X_n \to X$ a.s., if

$$P\left(\left\{\omega \in \Omega \,\middle|\, \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1\,.$$

Intuitively, this means that for almost all outcomes ω, the sequence of numbers $X_n(\omega)$ converges to $X(\omega)$.

(iii) *converges in distribution* (or *converges weakly*) to X, denoted $X_n \xrightarrow{d} X$ or $X_n \Rightarrow X$, if

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

for every $x \in \mathbb{R}$ at which F_X is continuous. This means that the overall shape of the distribution of X_n approaches the shape of the distribution of X. Equivalently, $\lim_{n\to\infty} \phi_{X_n}(t) = \phi_X(t)$ for all $t \in \mathbb{R}$.

Definition 2.46 (characteristic function (cf)). Let X be an RVRV. The *characteristic function (cf)* of X, denoted $\phi_X(t)$ or $\varphi_X(t)$, is defined as

$$\phi_X(t) = \mathbb{E}[e^{itX}] = \mathbb{E}[\cos(tX) + i\sin(tX)]\,,$$

where $i = \sqrt{-1}$ is the imaginary unit.

(i) If X is discrete with pmf $p_X(x)$, $\phi_X(t) = \sum_x e^{itx} p_X(x)$.

(ii) If X is continuous with pdf $p_X(x)$, $\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} p_X(x)\,dx$.

The characteristic function always exists for any random variable X and any $t \in \mathbb{R}$, because $|e^{itX}| = |\cos(tX) + i\sin(tX)| = \sqrt{\cos^2(tX) + \sin^2(tX)} = 1$, so the expectation is of a bounded function.

1:1

distr. $\Leftrightarrow$ pdf / cdf

distr. $\Leftrightarrow$ c.f.

Remark 2.47 (properties of cfs). Let X, Y be RVRVs.

(i) $\phi_X(0) = 1$ and $|\phi_X(t)| \leq 1$.

(ii) $\phi_{aX+b}(t) = e^{itb}\phi_X(at)$.

(iii) We can compute *moments* from the derivatives of $\phi(t)$ via $\mathbb{E}[X^n] = i^{-n}\phi_X^{(n)}(0)$ (if the moments exist).

(iv) If $X \perp\!\!\!\perp Y$, then $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$.

(v) **Uniqueness:** The cf uniquely determines the distribution. There are explicit inversion formulas to recover the cdf or pdf from the cf (e.g., Lévy's inversion theorem).

(vi) **(Lévy's) continuity theorem:** A sequence of RVRVs X_n converges in distribution to X if and only if their characteristic functions $\phi_{X_n}(t)$ converge pointwise to $\phi_X(t)$ for all t, and $\phi_X(t)$ is continuous at $t = 0$. This is crucial for proving the central limit theorem (clt). For more concrete statements see the wikipedia article.

Example 2.48 (cf of normal distribution). The characteristic function of $X \sim \mathcal{N}(\mu, \sigma^2)$ is

$$\phi_X(t) = \mathbb{E}[e^{itX}] = \exp\Big(i\mu t - \frac{1}{2}\sigma^2 t^2\Big)\,.$$