# Solution to exercise 3

Let $(X_i)_{i\in\{1,2,\dots,n\}}$ be $n$ independent and equally distributed random variables in $L^2(\Omega,\mathbb{R})$, with expectation $\mu$ and standard deviation $\sigma$. Call

$$X = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Note that $X$ can be seen as the average of a sample of length $n$ of drawings of a random variable with expectation $\mu$ and standard deviation $\sigma$.
It holds

$$\mathbb{E}[X] = \mu, \quad Var(X) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}. \tag{1}$$

We want to use first Chebyshev's inequality and then the Central limit theorem, together with (1), to determine the confidence interval at a given confidence level $\ell \in (0,1)$ for $X$, that is, we want to find $a, b$ with $a < \mu$, $b > \mu$, such that

$$P\left(X \in [a,b]\right) \geq \ell. \tag{2}$$

We start with Chebyshev's inequality, that tells us that for a given $\epsilon > 0$ it holds

$$P\left(|X - \mathbb{E}[X]| > \epsilon\right) \leq \frac{Var(X)}{\epsilon^2},$$

possibly with an approximation. Let $\epsilon = \frac{\sqrt{Var(X)}}{\sqrt{1-\ell}}$. Then we have

$$\begin{aligned}
1-\ell &\geq P\left(|X - \mathbb{E}[X]| > \frac{\sqrt{Var(X)}}{\sqrt{1-\ell}}\right) \\
&= P\left(X > \mathbb{E}[X] + \frac{\sqrt{Var(X)}}{\sqrt{1-\ell}} \text{ or } X < \mathbb{E}[X] - \frac{\sqrt{Var(X)}}{\sqrt{1-\ell}}\right) \\
&= P\left(X > \mu + \frac{\sigma}{\sqrt{n(1-\ell)}} \text{ or } X < \mu - \frac{\sigma}{\sqrt{n(1-\ell)}}\right) \\
&= 1 - P\left(\mu - \frac{\sigma}{\sqrt{n(1-\ell)}} \leq X \leq \mu + \frac{\sigma}{\sqrt{n(1-\ell)}}\right).
\end{aligned}$$

So we have

$$P\left(\mu - \frac{\sigma}{\sqrt{n(1-\ell)}} \leq X \leq \mu + \frac{\sigma}{\sqrt{n(1-\ell)}}\right) \geq \ell,$$

and we can write

$$a = \mu - \frac{\sigma}{\sqrt{n(1-\ell)}}, \quad b = \mu + \frac{\sigma}{\sqrt{n(1-\ell)}}$$

such that (2) holds true.
We now consider the Central limit theorem, which tells us that the distribution of $X = \frac{1}{n}\sum_{i=1}^{n} X_i$ can be approximated by $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ for large $n$. We want to find $\delta > 0$ such that

$$\ell \approx P(\mu - \delta \leq X \leq \mu + \delta).$$

For large $n$ it holds

$$
\begin{aligned}
P(|X-\mu|>\delta) &\approx P\left(\left|\left(\frac{\sigma}{\sqrt{n}}Z+\mu\right)-\mu\right|>\delta\right)=P\left(\left|\frac{\sigma}{\sqrt{n}}Z\right|>\delta\right)=2\cdot P\left(Z>\frac{\sqrt{n}}{\sigma}\delta\right)\\
&=2\cdot\left(1-\Phi\left(\frac{\sqrt{n}}{\sigma}\delta\right)\right),
\end{aligned}
$$

where $Z\sim\mathcal{N}(0,1)$ and $\Phi$ is the standard normal cumulative distribution function. Note that the second last inequality comes from the fact that the distribution of $Z$ is symmetric. We have

$$\ell=P(\mu-\delta\le X\le\mu+\delta)=1-P(|X-\mu|>\delta)\approx-1+2\Phi\left(\frac{\sqrt{n}}{\sigma}\delta\right),$$

which implies

$$\Phi\left(\frac{\sqrt{n}}{\sigma}\delta\right)\approx\frac{\ell+1}{2}.$$

Then we find

$$\delta\approx\frac{\sigma}{\sqrt{n}}\Phi^{-1}\left(\frac{\ell+1}{2}\right).$$