Probability and Statistics

Solution sheet 6

Solution 6.1

(a) We have by Chebyshev inequality that

$$\begin{aligned}\mathbb{P}\left(\left|\frac{S_n}{n}-1\right|\leq 0.5\right) &= 1-\mathbb{P}\left(\left|\frac{S_n}{n}-1\right|\geq 0.5\right)\\ &\geq 1-\frac{\mathbb{E}\left(\left(\frac{S_n}{n}-1\right)^2\right)}{0.25}\\ &= 1-\frac{\mathrm{Var}\left(\frac{S_n}{n}\right)}{0.25}\\ &= 1-\frac{8}{n}.\end{aligned}$$

When $n = 40$, the bound is 0.8.

(b) Thanks to Central Limit Theorem, we have that:

$$\frac{\sqrt{n}}{\sqrt{\mathrm{Var}(X_i)}}\left(\frac{S_n}{n}-\mathbb{E}(X_i)\right)\overset{(d)}{\to} N(0,1).$$

Then, using this property we have

$$\begin{aligned}\mathbb{P}\left(\left|\frac{S_n}{n}-1\right|\leq 0.5\right) &= \mathbb{P}\left(\frac{\sqrt{n}}{\sqrt{2}}\left|\frac{S_n}{n}-1\right|\leq\frac{\sqrt{n}}{\sqrt{2}}0.5\right)\\ &\approx \mathbb{P}(-\sqrt{5}\leq N(0,1)\leq\sqrt{5})\\ &= \phi\left(\sqrt{5}\right)-\phi\left(-\sqrt{5}\right)\approx 0.97.\end{aligned}$$

Solution 6.2

(a) Take $\epsilon > 0$, we know by continuity of f that there exists $\delta > 0$ so that for all $x \in [c-\delta, c+\delta]$, $|f(x) - f(c)| \leq \epsilon$. Then

$$\begin{aligned}|\mathbb{E}\left(f(Z_n)-f(c)\right)| &\leq \mathbb{E}\left(|f(Z_n)-f(c)|\right)\\ &\leq \mathbb{E}\left(|f(Z_n)-f(c)|\mathbb{1}_{|Z_n-c|\leq\delta}\right)+\mathbb{E}\left(|f(Z_n)-f(c)|\mathbb{1}_{|Z_n-c|>\delta}\right)\\ &\leq \epsilon+2\|f\|_\infty\mathbb{P}(|Z_n-c|>\delta)\xrightarrow[n\to\infty]{}\epsilon.\end{aligned}$$

(b) Take $\epsilon > 0$ and define

$$f_\epsilon(x)\mapsto\min\left\{\frac{1}{\epsilon}d(x,[c-\epsilon,c+\epsilon]),1\right\}.$$

f_ϵ is clearly a continuous function. Note that $f_\epsilon(x) = 0$ if $x \in [c-\epsilon, c+\epsilon]$ and $f(x) = 1$ if $|x-c| \geq 2\epsilon$. Then, we have that:

$$\mathbb{P}(|X_n-c|\geq 2\epsilon)\leq\mathbb{E}\left[f_\epsilon(X_n)\right]\xrightarrow{n\to\infty}f_\epsilon(c)=0.$$

Solution 6.3

(a) We know that for all $\epsilon \leq \frac{1}{2}$

$$\mathbb{P}\left[\bigcap_{0<\epsilon\leqslant\frac{1}{2}} \{|X_n| \leq \epsilon\}\right] = \mathbb{P}\left[|X_n| = 0\right] = \mathbb{P}\left[A_n^c\right],$$

so $X_n \xrightarrow{\mathbb{P}} 0$ if and only if $\mathbb{P}(A_n^c) \to 1$.

(b) Given that X_n takes only values in $\{0,1\}$, the sequence converges if and only if, from a point onward, it only takes the value 0. Therefore

$$\{\omega : \lim X_n(\omega) = 0\} = \bigcup_{k\in\mathbb{N}} \bigcap_{n\geq k} A_n^c = \liminf A_n^c.$$

(c) For $n \in \mathbb{N}$, define $r_n = \lfloor \log_2(n) \rfloor$ and $k_n = n - 2^{r_n}$. Take

$$A_n = \left[\frac{k_n}{2^{r_n}}, \frac{k_n+1}{2^{r_n}}\right],$$

note that $\mathbb{P}(A_n) = 2^{-r_n} \to 0$, so $X_n \xrightarrow{\mathbb{P}} 0$. Moreover, there are 2^{r_n} integers n_i such that $r_n = \lfloor \log_2(n_i) \rfloor$, and we have

$$\mathbb{P}\left(\bigcup_{n:r_n=r} A_n\right) = 2^{r_n}\frac{1}{2^{r_n}} = 1,$$

so $\bigcup_{n:r_n=r} A_n = [0,1]$. Then we know that for each $r \in \mathbb{N}$ and for all $x \in [0,1]$ there exits $n \in \mathbb{N}$ so that $r_n = r$ and $x \in A_n$, so $X_n(x)$ is 1 infinitely many times. Thus, $\{\omega : X_n(\omega) \to 0\} = \emptyset$.

Solution 6.4

(a) With Chebyshev inequality

$$P\left[\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right] \leq \frac{1}{\varepsilon^2} Var\left(\frac{S_n}{n}\right)$$

it is enough to prove that $\mathrm{Var}(\frac{S_n}{n}) \to 0\,(n \to \infty)$.

Computing the variance we have:

$$\begin{aligned}
\mathrm{Var}\left(\frac{S_n}{n}\right) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^n \mathrm{Var}(X_i) + 2\sum_{i<j}\mathrm{Cov}(X_i, X_j)\right) \\
&= \frac{1}{n^2}\left(n\sigma^2 + 2\sum_{k=1}^{n-1}(n-k)\,R(k)\right) \\
&= \frac{1}{n}\left(\sigma^2 + 2\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)R(k)\right)
\end{aligned}$$

The result obtains if we can prove

$$\lim_{n\to\infty} \frac{2}{n}\sum_{k=1}^{n-1}\left(\frac{n-k}{n}\right)R(k) = 0\,.$$

This can be proved as Cesaro Lemma. Let $\epsilon > 0$. As $\lim_{k\to\infty} R(k) = 0$, there exists $N \in \mathbb{N}$ such that for $n \geq N$ it holds that $|R(n)| \leqslant \epsilon$. Then, for $n \geqslant N+1$,

$$\frac{2}{n}\sum_{k=1}^{n-1}\Big(\frac{n-k}{n}\Big)R(k) = \frac{2}{n}\sum_{k=1}^{N}\Big(\frac{n-k}{n}\Big)R(k) + \frac{2}{n}\sum_{k=N+1}^{n-1}\Big(\frac{n-k}{n}\Big)R(k)$$

There exists $N' \in \mathbb{N}$ such that for $n \geqslant N'$, $\left|\frac{2}{n}\sum_{k=1}^{N}\Big(\frac{n-k}{n}\Big)R(k)\right| \leqslant \epsilon$. Then for $n \geqslant \max\{N, N'\}$, we have

$$\left|\frac{2}{n}\sum_{k=1}^{n-1}\Big(\frac{n-k}{n}\Big)R(k)\right| \leqslant \epsilon + 2\epsilon,$$

and we have proved the statement.

(b) We compute

$$\begin{aligned}\lim_{n\to\infty} n\,\mathrm{Var}\left(\frac{S_n}{n}\right) &= \lim_{n\to\infty}\left(\sigma^2 + 2\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)R(k)\right) \\ &= \sigma^2 + 2\sum_{k=1}^{\infty}R(k) - 2\lim_{n\to\infty}\sum_{k=1}^{n-1}\frac{k}{n}R(k)\,.\end{aligned}$$

Define:

$$a_n(k) := \begin{cases} \frac{k}{n}R(k) & (k<n) \\ 0 & (k \geq n)\end{cases}$$

it holds that $a_n(k) \to 0 (n \to \infty)$ for all k. Then we just have to use the dominated convergence theorem, to prove that this part converges to 0. Note that $|a_n(k)| \leq |R(k)|$and $|R(k)|$ is absolutely convergent. Therefore

$$\lim_{n\to\infty}\sum_{k=1}^{n-1}\frac{k}{n}R(k) = \lim_{n\to\infty}\sum_{k=1}^{n-1}a_n(k) \leqslant \lim_{n\to\infty}\sum_{k=1}^{\infty}a_n(k) = \sum_{k=1}^{\infty}\lim_{n\to\infty}a_n(k) = 0$$

Then

$$\lim_{n\to\infty} n\,Var\Big(\frac{S_n}{n}\Big) = \sigma^2 + 2\sum_{k=1}^{\infty}R(k).$$

Solution 6.5

(a) Take $f : \mathbb{R} \to \mathbb{R}$ a continuous and bounded function

$$\begin{aligned}\left|\int f d((1-\epsilon_n)\mu_n + \epsilon_n\nu_n) - \int f d\mu\right| &\leq \left|\int f d\mu_n - \int f d\mu\right| + \epsilon_n\left|\int f d\nu_n - \int f d\mu_n\right| \\ &\leq \left|\int f d\mu_n - \int f d\mu\right| + 2\epsilon_n\|f\|_\infty \to 0.\end{aligned}$$

(b) Take $\mu_n = \delta_0$, i.e. $\mu(A) = \mathbb{1}_{0\in A}$ and $\nu_n = \delta_n$. It is clear that $\mu_n \to \delta_0$ (it is a constant sequence), so $\left(1-\frac{1}{n}\right)\mu_n + \frac{1}{n}\nu_n \to \delta_0$, but:

$$\int |x| d\left(\left(1-\frac{1}{n}\right)\mu_n + \frac{1}{n}\nu_n\right)(x) = \frac{1}{n}n = 1 \neq 0 = \int |x| d\delta_0(x).$$

(c) We prove first the two claims of the Hint. We know that $\min\{|\cdot|, M\}$ is a bounded continuous function, therefore

$$\int \min\{|x|, M\} d\mu_n(x) \to \int \min\{|x|, M\} d\mu(x),$$

and

$$\begin{aligned}
&\int |x| d\mu_n(x) - \int \min\{|x|, M\} d\mu_n(x) \\
&= \int (|x| - M) \mathbb{1}_{|x| \geq M} d\mu_n(x) \\
&\leq \int |x| \mathbb{1}_{|x| \geq M} d\mu_n(x) \\
&\leq \sqrt{\int x^2 d\mu_n(x) \int \mathbb{1}_{|x| \geq M} d\mu_n(x)} \\
&\leq \sqrt{K} \sqrt{\int \mathbb{1}_{|x|^2 \geq M^2} d\mu_n(x)} \\
&\leq \sqrt{K} \sqrt{K/M^2} \\
&= \frac{K}{M}
\end{aligned}$$

where we used Cauchy-Schwarz and Chebychev inequalities successively. The above difference is clearly non-negative. By the monotone convergence theorem

$$\int \min\{|x|, M\} d\mu(x) \overset{M\to\infty}{\nearrow} \int |x| d\mu(x)$$

To finish, take $\epsilon > 0$, and M so that $K/M \leq \epsilon$, and that

$$\left| \int \min\{|x|, M\} d\mu(x) - \int |x| d\mu(x) \right| \leq \epsilon.$$

Take n_0 such that for all $n \geq n_0$,

$$\left| \int \min\{|x|, M\} d\mu_n(x) - \int \min\{|x|, M\} d\mu(x) \right| \leq \epsilon.$$

Finally,

$$\begin{aligned}
&\left| \int |x| d\mu_n(x) - \int |x| d\mu(x) \right| \\
\leq &\left| \int |x| d\mu_n(x) - \int \min\{|x|, M\} d\mu_n(x) \right| + \left| \int \min\{|x|, M\} d\mu_n(x) - \int \min\{|x|, M\} d\mu(x) \right| \\
&+ \left| \int \min\{|x|, M\} d\mu(x) - \int |x| d\mu(x) \right| \\
\leq &K/M + \epsilon + \epsilon = 3\epsilon.
\end{aligned}$$

Since ϵ is arbitrary, the statement is proved.