KTH, School of Electrical Engineering and Computer Science

Probability and random variables and the law of large numbers

Yuchao Li

December 17, 2018

Question 1

Q: Define/explain the concepts: probability space, event, probability, and independence (pairwise and mutual).
A: Given a measurable space $(\Omega, \mathscr{A})$, a probability space is a measure space $(\Omega, \mathscr{A}, \mathbf{P})$ where $\mathbf{P}$ is a probability measure defined on the aforementioned measurable space such that $\mathbf{P}(\Omega) = 1$. A event is any $A \in \mathscr{A}$. Probability of event A is given as $\mathbf{P}(A)$. For two events $E, F \in \mathscr{A}$, we say they are independent if $\mathbf{P}(E \cap F) = \mathbf{P}(E)\mathbf{P}(F)$. For a collection of events $A_1, ..., A_n$, they are pairwise independent if $\mathbf{P}(A_i \cap A_j) = \mathbf{P}(A_i)\mathbf{P}(A_j)$ for $i \neq j$; they are mutually independent if for any $\{i_1, ..., i_k\} \subset \{1, ..., n\}$, it holds that $\mathbf{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \mathbf{P}(A_{i_1}) \cdots \mathbf{P}(A_{i_k})$.

Question 2

Q: Define/explain the concepts: random variable, distribution, probability distribution function, expectation.
A: Given probability space $(\Omega, \mathscr{A}, \mathbf{P})$, a random variable is a $\mathscr{A}$-function $X : \Omega \to \mathbf{R}$. A distribution μ_X is an induced probability measure on $(\mathbf{R}, \mathscr{B})$ by random variable X. A probability distribution function is $F_X : \mathbf{R} \to [0, 1]$ that $F_X(x) = \mu_X((-\infty, x])$. The expectation of X defined on the aforementioned probability space is given by $E[X] = \int_\Omega X d\mathbf{P}$.

Question 3

Q: Prove the Borel–Cantelli lemma.

A: Given a probability space $(\Omega, \mathscr{A}, \mathbf{P})$ and a sequence of events $\{A_n\}_{n=1}^{\infty}$, we denote $F_n = \cup_{k=n}^{\infty} A_k$ where $F_n \downarrow$ and $E = \{A_n \,\text{i.o.}\} = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k = \cap_{n=1}^{\infty} F_n$.

1. Since $F_n \downarrow$, we have

$$\mathbf{P}(E) = \mathbf{P}(\cap_{n=1}^{\infty} F_n) = \lim_{n\to\infty} \mathbf{P}(F_n) = \lim_{n\to\infty} \mathbf{P}(\cup_{k=n}^{\infty} A_k). \tag{0.1}$$

   Due to subadditivity of probability measure, we have

$$\mathbf{P}(\cup_{k=n}^{\infty} A_k) \le \sum_{k=n}^{\infty} \mathbf{P}(A_k). \tag{0.2}$$

   Since $\sum_{k=1}^{\infty} \mathbf{P}(A_k) < \infty$, then if we denote $S_n = \sum_{k=1}^{n} \mathbf{P}(A_k)$, which is the partial sum, then we know that $\lim_{n\to\infty} S_n = \sum_{k=1}^{\infty} \mathbf{P}(A_k) < \infty$, namely $\{S_n\}_{n=1}^{\infty}$ is convergent and consequently, Cauchy. As a result, we have

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbf{P}(A_k) = 0, \tag{0.3}$$

   which concludes the proof of this direction.

2. If $\{A_n\}_{n=1}^{\infty}$ are mutually independent, then we have

$$\mathbf{P}(\cap_{k=n}^{\infty} A_k^c) = \prod_{k=n}^{\infty} \mathbf{P}(A_k^c) = \prod_{k=n}^{\infty} \big(1 - \mathbf{P}(A_k)\big). \tag{0.4}$$

   Since $1 - x \le e^{-x}$, then we have

$$\prod_{k=n}^{\infty} \big(1 - \mathbf{P}(A_k)\big) \le \exp\big(-\sum_{k=n}^{\infty} \mathbf{P}(A_k)\big). \tag{0.5}$$

   Since $\sum_{k=1}^{\infty} \mathbf{P}(A_k) = \infty$, then we have $\sum_{k=n}^{\infty} \mathbf{P}(A_k) = \infty \ \forall n \in \mathbf{R}$. Therefore, by Eq. (0.4), we have

$$\mathbf{P}(\cap_{k=n}^{\infty} A_k^c) = 0 \,\forall n \in \mathbf{R} \implies \mathbf{P}(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k^c) = 0.$$

   By De Morgan's rules, we have

$$\mathbf{P}(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) = 1 - \mathbf{P}(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k^c) = 1. \tag{0.6}$$

QUESTION 4

Q: Given $(\Omega, \mathscr{A}, P)$ and a sequence of random variables $\{X_n\}$, show that

$$\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} = \cap_{m=1}^{\infty} \cup_{n=1}^{\infty} \cap_{k=1}^{\infty} \{\omega : |X_n(\omega) - X_{n+k}(\omega)| < \frac{1}{m}\}.$$

A: Denote a real sequence $\{a_n\}_{n=1}^{\infty}$ where $a_n \in \mathbf{R}$. Since the real line $\mathbf{R}$ is complete, then we have that $\{a_n\}_{n=1}^{\infty}$ is convergent iff $\forall \varepsilon > 0$, $\exists N_\varepsilon$ such that $|a_i - a_j| < \varepsilon$ holds $\forall i, j \ge N_\varepsilon$ (statement 1). Besides $\forall \varepsilon > 0$, $\exists m_\varepsilon \in \mathbf{N}_+$ such that $\varepsilon > \frac{1}{m_\varepsilon}$ and $\forall m \in \mathbf{N}_+$, $\exists \varepsilon_m > 0$ such that

$\frac{1}{m} > \varepsilon_m$. Then we can see that for a real sequence $\{a_n\}_{n=1}^{\infty}$, statement 1 holds iff $\forall m \in \mathbf{N}_+$, $\exists N_m$ such that $|a_i - a_j| < \frac{1}{m}$ holds $\forall i, j \geq N_m$ (statement 2). Due to the triangle inequality, that is $|a_i - a_j| \leq |a_i - a_l| + |a_l - a_j|$, then we have statement 2 holds iff $\forall m \in \mathbf{N}_+$, $\exists n_m$ such that $|a_{n_m} - a_{n_m+k}| < \frac{1}{m}$ holds $\forall k \in \mathbf{N}_+$ (statement 3). This is due to that statement 2 obviously implies statement 3; and if statement 3 holds, then $\forall m \in \mathbf{N}_+$, $\exists n_{2m}$ such that $|a_{n_{2m}} - a_{n_{2m}+k}| < \frac{1}{2m}$, which means $\forall i, j \geq n_{2m}+1$, $|a_i - a_j| \leq |a_{n_{2m}} - a_i| + |a_{n_{2m}} - a_j| < \frac{1}{m}$ Then we proceed to prove the set equality, which is given as

$$\begin{aligned}
\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} &= \{\omega : \forall m \in \mathbf{N}_+, \exists n_m \text{ such that } |X_{n_m}(\omega) - X_{n_m+k}(\omega)| < \frac{1}{m}\ \forall k \in \mathbf{N}_+\} \\
&= \cap_{m=1}^{\infty}\{\omega : \exists n_m \text{ such that } |X_{n_m}(\omega) - X_{n_m+k}(\omega)| < \frac{1}{m}\ \forall k \in \mathbf{N}_+\} \\
&= \cap_{m=1}^{\infty} \cup_{n=1}^{\infty} \{\omega : |X_n(\omega) - X_{n+k}(\omega)| < \frac{1}{m}\ \forall k \in \mathbf{N}_+\} \\
&= \cap_{m=1}^{\infty} \cup_{n=1}^{\infty} \cap_{k=1}^{\infty}\{\omega : |X_n(\omega) - X_{n+k}(\omega)| < \frac{1}{m}\}.
\end{aligned}$$

For a similar result, please refer to P. 180, sidenote regarding Eq. (6.27) in [1].

QUESTION 5

Q: For mutually independent random variables $\{X_n\}$ with $E[X_n] = 0$ and $\sum_n \text{Var}[X_n] < \infty$ use the result in preceeding result and Kolmogorov's inequality (without proof) to show that $\sum_n X_n$ converges with probability one.

A: Given mutually independent random variables $\{X_n\}_{n=1}^{\infty}$ with $E[X_i] = 0$, $\sum_n \text{Var}[X_n] < \infty$, we denote $S_n = \sum_{i=1}^{n} X_i$. Then we have $\forall n, m \in \mathbf{N}_+$, it holds that

$$\begin{aligned}
\mathbf{P}(\cup_{k=1}^{\infty}\{|S_{n+k} - S_n| \geq \frac{1}{m}\}) &= \lim_{l\to\infty} \mathbf{P}(\cup_{k=1}^{l}\{|S_{n+k} - S_n| \geq \frac{1}{m}\}) \\
&= \lim_{l\to\infty} \mathbf{P}(\{\max_{1\leq k\leq l} |S_{n+k} - S_n| \geq \frac{1}{m}\}).
\end{aligned}$$

Denote $T_{n,k} = \sum_{j=1}^{k} X_{n+j} = S_{n+k} - S_n$, then we have $E[T_{n,k}] = 0$. By Kolmogorov inequality and $X_n{}_{n=1}^{\infty}$ iid, we have

$$\begin{aligned}
\mathbf{P}(\{\max_{1\leq k\leq l} |S_{n+k} - S_n| \geq \frac{1}{m}\}) &= \mathbf{P}(\{\max_{1\leq k\leq l} |T_{n,k} - E[T_{n,k}]| \geq \frac{1}{m}\}) \\
&\leq m^2 \text{Var}[T_{n,l}] \\
&= m^2 \sum_{k=1}^{l} \text{Var}[X_{n+k}].
\end{aligned}$$

Then we have

$$\begin{aligned}
\lim_{l\to\infty} \mathbf{P}(\{\max_{1\leq k\leq l} |S_{n+k} - S_n| \geq \frac{1}{m}\}) \leq \lim_{l\to\infty} m^2 \sum_{k=1}^{l} \text{Var}[X_{n+k}] \implies \\
\mathbf{P}(\cup_{k=1}^{\infty}\{|S_{n+k} - S_n| \geq \frac{1}{m}\}) \leq m^2 \sum_{k=1}^{\infty} \text{Var}[X_{n+k}].
\end{aligned}$$

Since $\sum_n \mathrm{Var}[X_n] < \infty$, then we have

$$\lim_{n\to\infty} \mathbf{P}(\cup_{k=1}^{\infty}\{|S_{n+k} - S_n| \geq \frac{1}{m}\}) \leq \lim_{n\to\infty} m^2 \sum_{k=1}^{\infty} \mathrm{Var}[X_{n+k}] = 0 \; \forall m \in \mathbf{N}_+.$$

Denote $F = \cup_{m=1}^{\infty} \cap_{n=1}^{\infty} \cup_{k=1}^{\infty} \{\omega : |S_n(\omega) - S_{n+k}(\omega)| \geq \frac{1}{m}\}$, then we have

$$\begin{aligned}
\mathbf{P}(\cap_{n=1}^{\infty} \cup_{k=1}^{\infty} \{\omega : |S_n(\omega) - S_{n+k}(\omega)| \geq \frac{1}{m}\}) &= \mathbf{P}(\cap_{n=1}^{\infty} \cap_{l=1}^{n} \cup_{k=1}^{\infty} \{\omega : |S_l(\omega) - S_{l+k}(\omega)| \geq \frac{1}{m}\}) \\
&= \lim_{n\to\infty} \mathbf{P}(\cap_{l=1}^{n} \cup_{k=1}^{\infty} \{\omega : |S_l(\omega) - S_{l+k}(\omega)| \geq \frac{1}{m}\}) \\
&\leq \lim_{n\to\infty} \mathbf{P}(\cup_{k=1}^{\infty} \{\omega : |S_n(\omega) - S_{n+k}(\omega)| \geq \frac{1}{m}\}) \\
&= 0 \; \forall m \in \mathbf{N}_+.
\end{aligned}$$

Due to subadditivity of probability measure, we have

$$\mathbf{P}(F) \leq \sum_{m=1}^{\infty} \mathbf{P}(\cap_{n=1}^{\infty} \cup_{k=1}^{\infty} \{\omega : |S_n(\omega) - S_{n+k}(\omega)| \geq \frac{1}{m}\}) = 0.$$

Therefore, we have $\mathbf{P}(F^c) = 1$, which concludes the proof.

QUESTION 6

Q: Given an iid sequence of zero-mean random variables $\{X_n\}$, let $Y_n = X_n \chi_{\{|X_n|<n\}}$, show that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \mathrm{Var}[Y_n] < \infty.$$

A: The proof here follows the approach given in P. 264 [2]. Since X_i is iid zero mean random variables, and $Y_n = X_n \chi_{\{|X_n|<n\}} \; \forall n$, then we have $E[Y_n] = E[X_n \chi_{\{|X_n|<n\}}] = E[X_1 \chi_{\{|X_1|<n\}}]$. Then we have $\lim_{n\to\infty} E[Y_n] = E[X_1] = 0$ due to DCT. Refer to the aforementioned source for

details. Then we have

$$\sum_{n=1}^{\infty}\frac{1}{n^2}\mathrm{Var}[Y_n] = \sum_{n=1}^{\infty}\frac{1}{n^2}\int_{\Omega}(X_n\chi_{\{|X_n|<n\}} - E[X_n\chi_{\{|X_n|<n\}}])^2 d\mathbf{P} \tag{0.7}$$

$$\leq \sum_{n=1}^{\infty}\frac{1}{n^2}\int_{\Omega}(X_n\chi_{\{|X_n|<n\}})^2 d\mathbf{P} \tag{0.8}$$

$$= \sum_{n=1}^{\infty}\frac{1}{n^2}\int_{\Omega}(X_1\chi_{\{|X_1|<n\}})^2 d\mathbf{P} \tag{0.9}$$

$$= \sum_{n=1}^{\infty}\frac{1}{n^2}\int_{\{|X_1|<n\}} X_1^2 d\mathbf{P} \tag{0.10}$$

$$= \sum_{n=1}^{\infty}\frac{1}{n^2}\sum_{m=1}^{n}\int_{\{m-1\leq|X_1|<m\}} X_1^2 d\mathbf{P} \tag{0.11}$$

$$= \sum_{m=1}^{\infty}\int_{\{m-1\leq|X_1|<m\}} X_1^2 d\mathbf{P}\sum_{n=m}^{\infty}\frac{1}{n^2} \tag{0.12}$$

$$\leq \sum_{m=1}^{\infty} m\int_{\{m-1\leq|X_1|<m\}} |X_1| d\mathbf{P}\sum_{n=m}^{\infty}\frac{1}{n^2} \tag{0.13}$$

$$\leq \sum_{m=1}^{\infty} m\int_{\{m-1\leq|X_1|<m\}} |X_1| d\mathbf{P}\frac{2}{m} \tag{0.14}$$

$$= 2\sum_{m=1}^{\infty}\int_{\{m-1\leq|X_1|<m\}} |X_1| d\mathbf{P} \tag{0.15}$$

$$= 2E[|X_1|] < \infty, \tag{0.16}$$

where in step (0.12), the summation is interchanged. The reason that the interchange is valid is due to the Tonelli's theorem. More precisely, if we denote $\{c_n\}_{n=1}^{\infty}$ where $c_n = \frac{1}{n^2}$ and $\{d_n\}_{n=1}^{\infty}$ where $d_n = \int_{\{n-1\leq|X_1|<n\}} X_1^2 d\mathbf{P}$. Then we have $\forall n$, $0 \leq c_n < \infty$ and $0 \leq d_n = \int_{\{n-1\leq|X_1|<n\}} X_1^2 d\mathbf{P} \leq n\int_{\{n-1\leq|X_1|<n\}} |X_1| d\mathbf{P} < \infty$ since $E[|X_1|] < \infty$. Then we define function $\mu : \mathbf{N}_+ \times \mathbf{N}_+ \to \{0,1\}$ as

$$\mu(n,m) = \begin{cases} 1, & m \leq n, \\ 0, & \text{o.w..} \end{cases}$$

Then we define $a_{mn} = c_n d_m \mu(n,m)$ which is nonnegative $\forall m,n$. Then we apply the technique used in P. 215, Example 6.12 b), Eq. (6.29), [2], by doing so we prove that the interchange of summation is valid. Regarding step (0.14), we apply that $\sum_{n=m}^{\infty}\frac{1}{n^2} \leq \frac{2}{m}$. This is because for $m \geq 2$, we have

$$\sum_{n=m}^{\infty}\frac{1}{n^2} < \sum_{n=m}^{\infty}\frac{1}{(n-1)n} = \frac{1}{m-1} \leq \frac{2}{m}\ \forall m \in \mathbf{N}_+. \tag{0.17}$$

In fact, the above inequality also holds when $m = 1$.

Question 7

Q: Using above results and prove the (strong) law of large numbers.

A: First we prove a Lemma, that is given sequence $\{a_n\}_{n=1}^{\infty}$, $\{b_n\}_{n=1}^{\infty}$ and $\exists N < \infty$ such that $a_n = b_n \ \forall n > N$, and we assume that $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_n$ exists, then we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_k = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} b_k.$$

To prove it, we need another theorem, which states for any convergent sequence $\{c_n\}$, we introduce a new sequence $\{d_n\}$, which is given simply by removing finite number of elements of $\{c_n\}$, then $\{d_n\}$ is also convergent and converges to the same limit. With this theorem, we denote $c_n = \frac{1}{n+N}\sum_{k=1}^{n+N} a_k$. Then we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_k = \lim_{n\to\infty} c_n = \lim_{n\to\infty} \frac{1}{n+N}\big(\sum_{k=1}^{N} a_k + \sum_{k=N+1}^{n+N} a_k\big) = 0 + \lim_{n\to\infty} \frac{1}{n+N}\sum_{k=N+1}^{n+N} a_k.$$

Replace the 0 with $\lim_{n\to\infty} \frac{1}{n+N}\sum_{k=1}^{N} b_k$, we obtain a truncated sequence of b_n, which has the same limit as $\{b_n\}$. Therefore, the proof is done.
Due to $\{X_i\}_{i=1}^{\infty}$ as iid, we first apply Lemma 7.4, P. 263, [2], which gives that

$$\sum_{n=1}^{\infty} \mathbf{P}(|X_n| \geq n) = \sum_{n=1}^{\infty} \mathbf{P}(|X_1| \geq n) < \infty.$$

Then denote $E = \{|X_k| \geq k \text{ i.o.}\}$. Due to Borel-Cantelli, we have $\mathbf{P}(E) = 0$. Therefore, we can foucs on E^c. Denote $Y_n = X_n \chi_{\{|X_n|<n\}}$, then it can be shown that $\lim_{n\to\infty} E[Y_n] = E[X_1] = 0$ following the arguments given on top of P. 264, [2]. Besdies, it also can be shown that $\forall \omega \in E^c$, $\exists N$ such that $Y_n = X_n \ \forall n > N$. Then due to the lemma above, we know $\forall \omega \in E^c$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = 0 \iff \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} Y_k = 0.$$

Denote $Z_n = n^{-1} Y_n$, then by Question 6, we know $\sum_{n=1}^{\infty} \mathrm{Var}[Z_n] < \infty$, then due to Question 5, we know that $\{\sum_{k=1}^{n}(Z_k - E[Z_k])\}_{n=1}^{\infty}$ converges a.s.. Denote the domain that the aforementioned sequence converges as F^c. Then we have $\mathbf{P}(F) = 0$. Recall that $\mathbf{P}(E) = 0$, due that subadditivity of probability measure, we have

$$\mathbf{P}(E^c \cap F^c) = 1 - \mathbf{P}\big((E \cup F)^c\big) \geq 1 - \mathbf{P}(E) - \mathbf{P}(F) = 1.$$

Therefore, we have $\mathbf{P}(E^c \cap F^c) = 1$. By Kronecker's lemma, Lemma 7.3, P. 259, [2], we have that $\forall \omega \in E^c \cap F^c$, it holds that

$$\lim_{n\to\infty} \sum_{k=1}^{n} (Y_k - E[Y_k]) = \lim_{n\to\infty} \sum_{k=1}^{n} k(Z_k - E[Z_k]) = 0.$$

Due to the Question 1.4, Homework 1, since $\lim_{n\to\infty} E[Y_n] = E[X_1] = 0$, we have

$$\lim_{n\to\infty} \sum_{k=1}^{n} E[Y_k] = \lim_{n\to\infty} E[Y_n] = 0.$$

As a result, we have that $\lim_{n\to\infty} \sum_{k=1}^{n} Y_k = 0$ holds $\forall \omega \in E^c \cap F^c$, where $\mathbf{P}(E^c \cap F^c) = 1$, which concludes the proof.

References

[1] Timo Koski, *Lecture notes: Probability and random processes at KTH*, 2017.

[2] John McDonald and Neil A. Weiss, *A course in real analysis*, 2nd Edition, 2012.