# AMS 261: Probability Theory (Fall 2017)

Homework 1 solutions

1. Consider a sample space $\Omega$.

(a) Prove that any intersection of $\sigma$-fields (of subsets of $\Omega$) is a $\sigma$-field. That is, if $\mathcal{F}_j$, $j \in J$, are $\sigma$-fields on $\Omega$ (with $J$ an arbitrary index set, countable or uncountable), then show that $\mathcal{F} = \bigcap_{j\in J} \mathcal{F}_j$ is a $\sigma$-field.

**Solution:** First, because each $\mathcal{F}_j$ is a $\sigma$-field on $\Omega$, we have that $\Omega \in \mathcal{F}_j$, for all $j \in J$, and therefore $\Omega \in \mathcal{F}$. For the second condition of the definition, consider an $A \in \mathcal{F}$. Then, $A \in \mathcal{F}_j$, for all $j \in J$, and therefore $A^c \in \mathcal{F}_j$, for all $j \in J$ (since the $\mathcal{F}_j$, $j \in J$, are $\sigma$-fields), which yields that $A^c \in \mathcal{F}$. Finally, consider a countable collection $\{A_i : i = 1, 2, ...\}$ of members of $\mathcal{F}$. Then, for each $j \in J$, $A_i \in \mathcal{F}_j$, for all $i$, which implies that for each $j \in J$, $\cup_{i=1}^{\infty} A_i \in \mathcal{F}_j$ (since the $\mathcal{F}_j$, $j \in J$, are $\sigma$-fields), and therefore $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

(b) Show by counterexample that a union of $\sigma$-fields may not be a $\sigma$-field.

**Solution:** Consider a finite sample space with three sample points, say, $\Omega = \{a, b, c\}$, and two $\sigma$-fields on $\Omega$ given by $\mathcal{F}_1 = \{\emptyset, \Omega, \{a\}, \{b, c\}\}$ and $\mathcal{F}_2 = \{\emptyset, \Omega, \{b\}, \{a, c\}\}$ (that is, the $\sigma$-fields generated by $\{a\}$ and $\{b\}$, respectively). Then, the union of the two $\sigma$-fields, $\mathcal{F}_1 \cup \mathcal{F}_2 = \{\emptyset, \Omega, \{a\}, \{b\}, \{b, c\}, \{a, c\}\}$, is not closed under (finite) unions (it contains $\{a\}$ and $\{b\}$ but not $\{a\} \cup \{b\} = \{a, b\}$), and hence it is not a $\sigma$-field.

2. Given a sample space $\Omega$ and a collection $\mathcal{E}$ of subsets of $\Omega$, the $\sigma$-field generated by $\mathcal{E}$, $\sigma(\mathcal{E})$, is defined as the intersection of all $\sigma$-fields on $\Omega$ that contain $\mathcal{E}$. (As discussed in class, $\sigma(\mathcal{E})$ is the smallest $\sigma$-field that contains $\mathcal{E}$.)

(a) Consider two collections $\mathcal{E}_1$ and $\mathcal{E}_2$ of subsets of $\Omega$. Show that if $\mathcal{E}_1 \subseteq \mathcal{E}_2$, then $\sigma(\mathcal{E}_1) \subseteq \sigma(\mathcal{E}_2)$.

**Solution:** Because $\mathcal{E}_1 \subseteq \mathcal{E}_2$, any $\sigma$-field on $\Omega$ that contains $\mathcal{E}_2$ will also contain $\mathcal{E}_1$. Therefore, to define $\sigma(\mathcal{E}_2)$ we are intersecting over a smaller number of sets relative to $\sigma(\mathcal{E}_1)$, and thus by their definition, we have $\sigma(\mathcal{E}_1) \subseteq \sigma(\mathcal{E}_2)$.

(b) As in part (a), let $\mathcal{E}_1$ and $\mathcal{E}_2$ be collections of subsets of the sample space $\Omega$. Prove that if $\mathcal{E}_1 \subseteq \sigma(\mathcal{E}_2)$ and $\mathcal{E}_2 \subseteq \sigma(\mathcal{E}_1)$, then $\sigma(\mathcal{E}_1) = \sigma(\mathcal{E}_2)$.

**Solution:** Using the result from part (a), the assumptions $\mathcal{E}_1 \subseteq \sigma(\mathcal{E}_2)$ and $\mathcal{E}_2 \subseteq \sigma(\mathcal{E}_1)$ yield $\sigma(\mathcal{E}_1) \subseteq \sigma(\sigma(\mathcal{E}_2))$ and $\sigma(\mathcal{E}_2) \subseteq \sigma(\sigma(\mathcal{E}_1))$, respectively. Because $\sigma(\mathcal{E}_1)$ and $\sigma(\mathcal{E}_2)$ are $\sigma$-fields, we have $\sigma(\sigma(\mathcal{E}_1)) = \sigma(\mathcal{E}_1)$ and $\sigma(\sigma(\mathcal{E}_2)) = \sigma(\mathcal{E}_2)$. Therefore, $\sigma(\mathcal{E}_1) \subseteq \sigma(\mathcal{E}_2)$ and $\sigma(\mathcal{E}_2) \subseteq \sigma(\mathcal{E}_1)$, which provides the result.

3. Let $\mathcal{F}$ be a collection of subsets of a sample space $\Omega$.

(a) Suppose that $\Omega \in \mathcal{F}$, and that when $A, B \in \mathcal{F}$ then $A \cap B^c \in \mathcal{F}$. Show that $\mathcal{F}$ is a field.

**Solution:** First, $\Omega \in \mathcal{F}$, by assumption. For the second condition, let $A \in \mathcal{F}$. Because $A \in \mathcal{F}$ and $\Omega \in \mathcal{F}$, we have by assumption that $\Omega \cap A^c = A^c \in \mathcal{F}$. Finally, for the third condition, consider $A \in \mathcal{F}$ and $B \in \mathcal{F}$. We have shown that $A^c \in \mathcal{F}$, hence by assumption, $A^c \cap B^c = (A \cup B)^c \in \mathcal{F}$. Now using the second condition, we obtain that $A \cup B = ((A \cup B)^c)^c \in \mathcal{F}$.

(b) Suppose that $\Omega \in \mathcal{F}$, and that $\mathcal{F}$ is closed under the formation of complements and finite pairwise disjoint unions. Show by counterexample that $\mathcal{F}$ need not be a field.

**Solution:** For a counterexample, consider again a finite sample space, now, with four sample points, say, $\Omega = \{1, 2, 3, 4\}$. Let $\mathcal{F} = \{\emptyset, \Omega, \{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\}$. It is straightforward to check that $\mathcal{F}$ is closed under taking complements and is closed under finite disjoint unions (it is also defined to contain $\Omega$). However, although, say, $\{1, 2\} \in \mathcal{F}$ and $\{1, 3\} \in \mathcal{F}$, their union, $\{1, 2, 3\}$, does not belong to $\mathcal{F}$, and thus $\mathcal{F}$ is not a field.

4. Consider the sample space $\Omega = (0,1]$ and the collection $\mathcal{B}_0$ of all finite pairwise disjoint unions of subintervals of $(0,1]$. That is, any member $B$ of $\mathcal{B}_0$ is of the form $B = \bigcup_{i=1}^{n}(a_i, b_i]$, where $n$ is finite, and for each $i = 1, ..., n$, $0 \leq a_i < b_i \leq 1$, with $(a_i, b_i] \cap (a_j, b_j] = \emptyset$ for any $i \neq j$. Show that $\mathcal{B}_0$ augmented by the empty set is a field, but not a $\sigma$-field.
**Solution:** First, taking $n = 1$ with $a_1 = 0$ and $b_1 = 1$, we obtain that $\Omega \in \mathcal{B}_0$. Next, consider $B \in \mathcal{B}_0$, which therefore is of the form $B = \bigcup_{i=1}^{n}(a_i, b_i]$ for disjoint $(a_i, b_i]$. Without loss of generality, assume that $a_1 \leq a_2 \leq ... \leq a_n$. Then, $B^c = (0, a_1] \cup (b_1, a_2] \cup ... (b_{n-1}, a_n] \cup (b_n, 1]$ which belongs to $\mathcal{B}_0$ (note that some of the intervals in the expression for $B^c$ may be given by the empty set, since we can have $b_{i-1} = a_i$ for some $i$). Finally, consider $B \in \mathcal{B}_0$ and $C \in \mathcal{B}_0$, say, $B = \bigcup_{i=1}^{n}(a_i, b_i]$ for disjoint $(a_i, b_i]$, and $C = \bigcup_{j=1}^{m}(c_j, d_j]$ for disjoint $(c_j, d_j]$. Note that for the third condition of the definition for a field, rather than proving that $B \cup C \in \mathcal{B}_0$, it is equivalent to show that $B \cap C \in \mathcal{B}_0$. Now, $B \cap C = \bigcup_{i=1}^{n}\bigcup_{j=1}^{m}\{(a_i, b_i] \cap (c_j, d_j]\}$, where the $(a_i, b_i] \cap (c_j, d_j]$ are disjoint, and each set $(a_i, b_i] \cap (c_j, d_j]$ is either the empty set or an interval of the form $(e, f]$. Hence, $B \cap C \in \mathcal{B}_0$, completing the proof that $\mathcal{B}_0$, along with the empty set, is a field. To show that $\mathcal{B}_0$ is not a $\sigma$-field, note that for any $x \in (0,1)$, we can write $\{x\} = \bigcap_{n=1}^{\infty}(x - n^{-1}, x]$, i.e., $\{x\}$ can be expressed as a countable intersection of members of $\mathcal{B}_0$, but $\{x\}$ does not belong to $\mathcal{B}_0$.

5. Let $\Omega = \{\omega_1, \omega_2, ...\}$ be a countable set, $\{p_n : n = 1, 2, ...\}$ be a sequence of non-negative real numbers such that $\sum_{n=1}^{\infty} p_n = 1$, and $\mathcal{F}$ be the collection of all subsets of $\Omega$. For each $A \in \mathcal{F}$, define the set function

$$P(A) = \sum_{\{n:\, \omega_n \in A\}} p_n.$$

Show that $(\Omega, \mathcal{F}, P)$ is a probability space.
**Solution:** Using directly the definition, $\mathcal{F}$ is a $\sigma$-field (the default $\sigma$-field for countable sample spaces). Therefore, to establish that $(\Omega, \mathcal{F}, P)$ is a probability space, we need to show that $P$ is a probability measure on $(\Omega, \mathcal{F})$. By definition, $P(\Omega) = \sum_{\{n:\omega_n \in \Omega\}} p_n = \sum_{n=1}^{\infty} p_n$, which is equal to 1 by assumption. Moreover, since the $p_n$ are non-negative, we have $P(A) \geq 0$ for any $A \in \mathcal{F}$; also, for any $A \in \mathcal{F}$, i.e., any $A \subseteq \Omega$, we have $P(A) = \sum_{\{n:\omega_n \in A\}} p_n \leq \sum_{\{n:\omega_n \in \Omega\}} p_n = 1$. Finally, to show countable additivity, consider a countable sequence, $\{A_m : m = 1, 2, ...\}$, of pairwise disjoint subsets of $\Omega$. For any $\omega_n \in \bigcup_{m=1}^{\infty} A_m$, we have that $\omega_n \in A_k$ for some $k$ and $\omega_n \notin A_\ell$ for all $\ell \neq k$, since the $A_m$ are pairwise disjoint. Therefore,

$$P\left(\bigcup_{m=1}^{\infty} A_m\right) = \sum_{\{n:\, \omega_n \in \bigcup_{m=1}^{\infty} A_m\}} p_n = \sum_{m=1}^{\infty} \sum_{\{n:\, \omega_n \in A_m\}} p_n = \sum_{m=1}^{\infty} P(A_m).$$

# AMS 261: Probability Theory (Fall 2017)

Homework 2 solutions

1. Let $\{A_n : n = 1, 2, ...\}$ be a countable sequence of subsets of a sample space $\Omega$.
(a) Assume that $\{A_n : n = 1, 2, ...\}$ is an increasing sequence, that is, $A_n \subseteq A_{n+1}$, for all $n \geq 1$. Show that $\lim_{n\to\infty} A_n$ exists, and $\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$.
(b) Assume that $\{A_n : n = 1, 2, ...\}$ is a decreasing sequence, that is, $A_{n+1} \subseteq A_n$, for all $n \geq 1$. Show that $\lim_{n\to\infty} A_n$ exists, and $\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$.
**Solution:** (a) Let $B = \bigcup_{n=1}^{\infty} A_n$. Based on the assumption, we have $1_{A_n}(\omega) \leq 1_{A_{n+1}}(\omega)$, for all $\omega \in \Omega$, and for all $n$. Therefore, for each $\omega \in \Omega$, $\{1_{A_n}(\omega) : n = 1, 2, ...\}$ is an increasing sequence of reals, which is bounded from above by 1. Hence, for each $\omega \in \Omega$, $\lim_{n\to\infty} 1_{A_n}(\omega)$ exists, and thus, by its definition, $\lim_{n\to\infty} A_n$ exists. Let $A = \lim_{n\to\infty} A_n$.
Consider a fixed $\omega \in \Omega$. If $\omega$ does not belong to any of the $A_n$, we have $1_B(\omega) = 0$ as well as $1_{A_n}(\omega) = 0$ for all $n$, which yields $\lim_{n\to\infty} 1_{A_n}(\omega) = 0$. Next, assume that $\omega \in A_n$ for at least one $n$. Because $1_{A_n}(\omega) \leq 1_{A_{n+1}}(\omega)$, for all $n$, there must exist some $k$ (that depends on $\omega$) such that $1_{A_n}(\omega) = 1$ for all $n \geq k$. Hence, $\lim_{n\to\infty} 1_{A_n}(\omega) = 1$, and $1_B(\omega) = \max_n\{1_{A_n}(\omega)\} = 1$.
Therefore, we have shown that for all $\omega \in \Omega$, $\lim_{n\to\infty} 1_{A_n}(\omega) = 1_B(\omega)$, i.e., $1_A(\omega) = 1_B(\omega)$, and thus $\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$. The approach is similar for part (b).

2. Consider countable sequences, $\{A_n : n = 1, 2, ...\}$, $\{B_n : n = 1, 2, ...\}$ and $\{C_n : n = 1, 2, ...\}$, of subsets of the same sample space $\Omega$. Assume that $A_n \subseteq B_n \subseteq C_n$, for all $n \geq K$ for some sufficiently large positive integer $K$. Moreover, suppose that $\limsup_{n\to\infty} C_n \subseteq \liminf_{n\to\infty} A_n$. Prove that each of $\lim_{n\to\infty} A_n$, $\lim_{n\to\infty} B_n$ and $\lim_{n\to\infty} C_n$ exists, and that all three limits are the same.
**Solution:** From the assumption $A_n \subseteq B_n$, for all $n \geq K$, we obtain $\cup_{n=K}^{\infty} A_n \subseteq \cup_{n=K}^{\infty} B_n$, and taking intersection over $K = 1, 2, ...$ on both sides, we have

$$\limsup_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} B_n. \tag{2.1}$$

Using similar arguments, it can be shown that:

$$\limsup_{n\to\infty} B_n \subseteq \limsup_{n\to\infty} C_n; \qquad \liminf_{n\to\infty} A_n \subseteq \liminf_{n\to\infty} B_n; \qquad \liminf_{n\to\infty} B_n \subseteq \liminf_{n\to\infty} C_n. \tag{2.2}$$

Now, combining the assumption $\limsup_{n\to\infty} C_n \subseteq \liminf_{n\to\infty} A_n$ with (2.1) and the first result in (2.2), we obtain that $\limsup_{n\to\infty} A_n \subseteq \liminf_{n\to\infty} A_n$. Hence, $\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = A$, and therefore $A = \lim_{n\to\infty} A_n$ exists. From the assumption $\limsup_{n\to\infty} C_n \subseteq \liminf_{n\to\infty} A_n$ and the second and third results in (2.2), $\limsup_{n\to\infty} C_n \subseteq \liminf_{n\to\infty} C_n$, and therefore $\liminf_{n\to\infty} C_n = \limsup_{n\to\infty} C_n = C$ and $C = \lim_{n\to\infty} C_n$ exists. In addition, from (2.1) and the first result in (2.2), $\limsup_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} C_n$, and therefore $\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} C_n$, which, along with the assumption $\limsup_{n\to\infty} C_n \subseteq \liminf_{n\to\infty} A_n$, implies that $\limsup_{n\to\infty} C_n = \liminf_{n\to\infty} A_n$, that is, $A = C$. Using similar arguments, it can be shown that $B = \lim_{n\to\infty} B_n$ exists and that $A = B = C$.

3. Consider a measurable space $(\Omega, \mathcal{F})$ and a set function $P$: $\mathcal{F} \longrightarrow [0,1]$, which satisfies $P(\Omega) = 1$, and $P(A \cup B) = P(A) + P(B)$ for any $A$ and $B$ in $\mathcal{F}$ with $A \cap B = \emptyset$. Moreover, assume that $P$ is continuous, that is, $P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$, for any sequence $\{A_n : n = 1, 2, ...\}$ of sets in $\mathcal{F}$ for which $\lim_{n\to\infty} A_n$ exists. Prove that $P$ is a probability measure on $(\Omega, \mathcal{F})$.
**Solution:** We basically need to prove that $P$ is countably additive if it is continuous and finitely additive. Let $\{A_n : n = 1, 2, ...\}$ be a countable pairwise disjoint sequence of events in $\mathcal{F}$. Define $B_n = \bigcup_{k=1}^{n} A_k$, for $n \geq 1$. This is an increasing sequence of events with $\lim_{n\to\infty} B_n = \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. Hence,

$$P(\bigcup_{n=1}^{\infty} A_n) = P(\bigcup_{n=1}^{\infty} B_n) = P(\lim_{n\to\infty} B_n) = \lim_{n\to\infty} P(B_n),$$

using the assumption of continuity. Now, $\lim_{n\to\infty} P(B_n) = \lim_{n\to\infty} P(\bigcup_{k=1}^{n} A_k) = \lim_{n\to\infty} \sum_{k=1}^{n} P(A_k)$, using the assumption of finite additivity. (Finite additivity for general finite $n$ results by induction from the assumption, which involves the case with $n = 2$.) Hence the result is established noting that, by definition, $\lim_{n\to\infty} \sum_{k=1}^{n} P(A_k) = \sum_{n=1}^{\infty} P(A_n)$.

4. Prove that any non-decreasing function from $\mathbb{R}$ to $\mathbb{R}$ is measurable. (Assume the usual Borel $\sigma$-field on $\mathbb{R}$.)
**Solution:** Let $f$ denote the non-decreasing function. First, since the collection of intervals $\{(-\infty, b] : b \in \mathbb{R}\}$ generates the Borel $\sigma$-field on the real line, it suffices to show that $f^{-1}((-\infty, b]) = \{\omega \in \mathbb{R} : f(\omega) \leq b\}$ is a Borel subset on the real line. This is fairly straightforward to check *graphically* by considering the different possible shapes that $f$ could have, e.g., strictly increasing and continuous; non-decreasing and continuous; non-decreasing with discontinuities.
Alternatively, let $\alpha$ be the least upper bound of $f^{-1}((-\infty, b])$. Therefore, $\omega \leq \alpha$ for all $\omega \in f^{-1}((-\infty, b])$, and hence $f^{-1}((-\infty, b]) \subseteq (-\infty, \alpha]$. Moreover, using the definition for least upper bounds, we have that for each $\epsilon > 0$, there exists some $\omega \in f^{-1}((-\infty, b])$ such that $\omega > \alpha - \epsilon$, and thus $f(\alpha - \epsilon) \leq f(\omega) \leq b$. That is, for each $\epsilon > 0$, $\alpha - \epsilon \in f^{-1}((-\infty, b])$, which yields that $(-\infty, \alpha) \subseteq f^{-1}((-\infty, b])$. Hence, $(-\infty, \alpha) \subseteq f^{-1}((-\infty, b]) \subseteq (-\infty, \alpha]$, which implies that $f^{-1}((-\infty, b])$ must be either $(-\infty, \alpha)$ or $(-\infty, \alpha]$ both of which are Borel sets on the real line.

5. Let $(\Omega_j, \mathcal{F}_j)$, $j = 1, 2, 3$, be measurable spaces. Consider measurable functions $X : \Omega_1 \to \Omega_2$ and $Y : \Omega_2 \to \Omega_3$, and define the composition function $Y \circ X : \Omega_1 \to \Omega_3$ by $Y \circ X(\omega_1) = Y(X(\omega_1))$, for any $\omega_1 \in \Omega_1$. Show that $Y \circ X$ is a measurable function.
**Solution:** Let $B \in \mathcal{F}_3$. We need to show that $Y \circ X^{-1}(B) = \{\omega_1 \in \Omega_1 : Y \circ X(\omega_1) \in B\} \in \mathcal{F}_1$. Because $Y$ is measurable, $Y^{-1}(B) = \{\omega_2 \in \Omega_2 : Y(\omega_2) \in B\} \in \mathcal{F}_2$. Now, because $X$ is measurable, $X^{-1}(Y^{-1}(B))$ $\in \mathcal{F}_1$, and this establishes the result, since $X^{-1}(Y^{-1}(B)) = \{\omega_1 \in \Omega_1 : X(\omega_1) \in Y^{-1}(B)\} = \{\omega_1 \in \Omega_1 : Y(X(\omega_1)) \in B\} = Y \circ X^{-1}(B)$.

6. Consider a sequence $\{X_n : n = 1, 2, ...\}$ of $\mathbb{R}$-valued random variables defined on the same probability space $(\Omega, \mathcal{F}, P)$. Let $C$ be the set of $\omega \in \Omega$ such that $\{X_n(\omega) : n = 1, 2, ...\}$ is a convergent numerical sequence. Prove that $C \in \mathcal{F}$.
**Solution:** Recall that a characterization of convergence for a numerical sequence is through the Cauchy criterion, specifically, sequence $\{a_n : n = 1, 2, ...\}$ converges to some limit if and only if it is a Cauchy sequence, that is, for any $\varepsilon > 0$, there exists $N$ such that for all $n, m > N$, $|a_n - a_m| < \varepsilon$. Therefore,

$$C = \{\omega \in \Omega : \{X_n(\omega) : n = 1, 2, ...\} \text{ Cauchy}\} = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n,m>N} B_{nm,k},$$

where $B_{nm,k} = \{\omega \in \Omega : |X_n(\omega) - X_m(\omega)| < k^{-1}\}$. Since $|X_n - X_m| = \max\{X_n - X_m, X_m - X_n\}$ is a random variable, $B_{nm,k} = |X_n - X_m|^{-1}((-\infty, k^{-1})) \in \mathcal{F}$, for all $n, m, k$, and thus, $C \in \mathcal{F}$.
Note that working with the Cauchy criterion avoids the need to refer to the limit of the sequence, which does not necessarily correspond to a well-defined function on $\Omega$ (there may be many $\omega$ for which the limit $\lim_{n\to\infty} X_n(\omega)$ does not exist).

7. Let $X$ and $Y$ be $\mathbb{R}$-valued random variables defined on the same probability space $(\Omega, \mathcal{F}, P)$, and consider the subset of $\Omega$ defined by $A = \{\omega \in \Omega : X(\omega) \neq Y(\omega)\}$.
(a) Prove that $A$ is an event in $\mathcal{F}$.
(b) Assume that $P(A) = 0$. Prove that $P(X^{-1}(B)) = P(Y^{-1}(B))$ for any Borel subset $B$ of $\mathbb{R}$ (in which case, we say that the distributions of $X$ and $Y$ are equal).
**Solution:** (a) Let $Q$ be the (countable) set of rational numbers. We can write $A = A_1 \cup A_2$, where $A_1 = \bigcup_{q\in Q} (\{\omega \in \Omega : X(\omega) < q\} \cap \{\omega \in \Omega : Y(\omega) > q\})$, and $A_2 = \bigcup_{q\in Q} (\{\omega \in \Omega : Y(\omega) < q\} \cap \{\omega \in \Omega : X(\omega) > q\})$. (This is based on the *Archimedean Property* of the real numbers: for any real numbers $a$ and $b$ with $a < b$, there exists a rational number $q$ such that $a < q < b$.) Because $A_1$ and $A_2$ are expressed through countable operations on events, we have that $A_1 \in \mathcal{F}$ and $A_2 \in \mathcal{F}$, and thus $A \in \mathcal{F}$.
(b) Consider a Borel subset $B$ of $\mathbb{R}$ and let $D_1 = X^{-1}(B)$ and $D_2 = Y^{-1}(B)$, both of which are events in $\mathcal{F}$. We have $D_1 = (D_1 \cap A) \cup (D_1 \cap A^c)$, and hence $P(D_1) = P(D_1 \cap A^c)$ (note that $P(D_1 \cap A) = 0$, since $P(D_1 \cap A) \leq P(A) = 0$). Similarly, we can show that $P(D_2) = P(D_2 \cap A^c)$. Now $P(D_1 \cap A^c) = P(\{\omega \in \Omega : X(\omega) \in B \text{ and } X(\omega) = Y(\omega)\})$, $P(D_2 \cap A^c) = P(\{\omega \in \Omega : Y(\omega) \in B \text{ and } X(\omega) = Y(\omega)\})$, and thus $P(D_1 \cap A^c) = P(D_2 \cap A^c) = P(\{\omega \in \Omega : X(\omega) \in B \text{ and } Y(\omega) \in B \text{ and } X(\omega) = Y(\omega)\})$. Therefore $P(D_1) = P(D_2)$, i.e., $P(X^{-1}(B)) = P(Y^{-1}(B))$.

# AMS 261: Probability Theory (Fall 2017)

Homework 3 solutions

1. Consider a countable sequence $\{X_n : n = 1, 2, ...\}$ of $\mathbb{R}^+$-valued random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. Assume that all random variables $X_n$ have the same distribution, with distribution function given by $F(x) = 1 - \exp(-x)$, $x \in \mathbb{R}^+$.
   - Show that
$$P\left(\liminf_{n\to\infty} \{\omega \in \Omega : X_n(\omega) \le (1+\delta)\log(n)\}\right) = 1,$$
   for any fixed $\delta > 0$.
   **Solution:** Fix $\delta > 0$, and for each $n$, let $A_n = \{\omega \in \Omega : X_n(\omega) > (1+\delta)\log(n)\}$. (Note that each $A_n \in \mathcal{F}$, since each $X_n$ is a random variable.) We need to show that $P(\liminf_{n\to\infty} A_n^c) = 1$. We have
$$P(A_n) = 1 - P(X_n \le (1+\delta)\log(n)) = \exp\{-(1+\delta)\log(n)\} = \frac{1}{n^{(1+\delta)}}$$
   using the form of the distribution function for each $X_n$, $P(X_n \le x) = 1 - \exp(-x)$, $x \in \mathbb{R}^+$. Hence,
$$\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} \frac{1}{n^{(1+\delta)}} < \infty,$$
   since $\delta > 0$, and the Borel lemma yields $P(\limsup_{n\to\infty} A_n) = 0$, i.e., $P(\liminf_{n\to\infty} A_n^c) = 1$.

2. Let $F$ and $G$ be distribution functions on $\mathbb{R}$ such that $G(t) \le F(t)$, for all $t \in \mathbb{R}$ (in which case, $G$ is said to be *stochastically larger* than $F$).
   - Construct two $\mathbb{R}$-valued random variables $X$ and $Y$, defined on the same probability space $(\Omega, \mathcal{F}, P)$, such that the distribution function of $X$ is $G$, the distribution function of $Y$ is $F$, and $P(X \ge Y) = 1$.
   **Solution:** Consider $\Omega = (0,1)$, $\mathcal{F}$ the Borel $\sigma$-field on $(0,1)$, and $P$ the uniform distribution on $(0,1)$. Next, define $\mathbb{R}$-valued functions $X$ and $Y$ on $(\Omega, \mathcal{F}, P)$ by $X(\omega) = \inf\{t \in \mathbb{R} : G(t) \ge \omega\}$ and $Y(\omega) = \inf\{t \in \mathbb{R} : F(t) \ge \omega\}$, for each $\omega \in (0,1)$. Then, $X$ and $Y$ are random variables with distribution functions $G$ and $F$, respectively. Moreover, because $G(t) \le F(t)$ for any $t \in \mathbb{R}$, we have that $\{t \in \mathbb{R} : G(t) \ge \omega\} \subseteq \{t \in \mathbb{R} : F(t) \ge \omega\}$ for any $\omega \in (0,1)$, and thus, $\inf\{t \in \mathbb{R} : G(t) \ge \omega\} \ge \inf\{t \in \mathbb{R} : F(t) \ge \omega\}$ for any $\omega \in (0,1)$. Therefore, $X(\omega) \ge Y(\omega)$ for any $\omega \in (0,1)$, which yields the result.

3. Consider a simple random variable $X$ defined on some probability space $(\Omega, \mathcal{F}, P)$, and let $F$ be its distribution function. Denote by $F(x^-) = \lim_{y \nearrow x} F(y)$ (or equivalently, $F(x^-) = \lim_{n\to\infty} F(x_n)$ for an arbitrary increasing sequence $\{x_n : n = 1, 2, ...\}$ converging to $x$).
   - Show that the expectation of $X$ can be written in the form
$$\mathrm{E}(X) = \sum_{x \in \mathbb{R}} x\{F(x) - F(x^-)\}.$$

   **Solution:** Because $X$ is a simple random variable, it can be expressed in the form $\sum_{j=1}^{k} c_j 1_{C_j}$, where the $c_j$ are the distinct values of $X$ and the $C_j$ partition $\Omega$, in particular, $C_j = \{\omega \in \Omega : X(\omega) = c_j\}$. Denote by $Q_X$ the probability measure induced on $\mathbb{R}$ by $X$. Consider an arbitrary increasing sequence $\{x_n : n = 1, 2, ...\}$ converging to $c_j$. If we define $A_n = (-\infty, x_n]$, $n = 1, 2, ...$, we have $\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n = (-\infty, c_j)$. Hence, using continuity of probability measure, we obtain $\lim_{n\to\infty} Q_X(A_n) = Q_X((-\infty, c_j))$, i.e., $\lim_{n\to\infty} F(x_n) = P(\{\omega \in \Omega : X(\omega) < c_j\})$, and thus $F(c_j^-) = P(\{\omega \in \Omega : X(\omega) < c_j\})$ (by the definition of $F(c_j^-)$). Now, $F(c_j) = P(\{\omega \in \Omega : X(\omega) \le c_j\}) = P(\{\omega \in \Omega : X(\omega) < c_j\}) + P(C_j)$, and hence $P(C_j) = F(c_j) - F(c_j^-)$. Therefore, $\mathrm{E}(X) = \sum_{j=1}^{k} c_j P(C_j) = \sum_{j=1}^{k} c_j\{F(c_j) - F(c_j^-)\}$, which yields the result, since for any $x \in \mathbb{R}$ that is not a value of $X$ (i.e., not one of the $c_j$) we have $F(x) - F(x^-) = P(\{\omega \in \Omega : X(\omega) = x\}) = 0$.

4. Let $X$ be a simple random variable (taking both negative and positive values) defined on some probability space $(\Omega, \mathcal{F}, P)$.
   - Show that expectation definitions 1 (for simple random variables) and 3 (for general random variables taking values on the extended real line) are equivalent.

   **Solution:** We can write $X$ in the form $\sum_{j=1}^{k} c_j 1_{C_j}$, where the $c_j$ are the distinct values of $X$ ($c_j \in \mathbb{R}$), and $\{C_j : j = 1, ..., k\}$ is a finite measurable partition of $\Omega$ (in particular, $C_j = \{\omega \in \Omega : X(\omega) = c_j\}$). Hence, $X^+ = \sum_{j\in J_1} c_j 1_{C_j}$, and $X^- = -\sum_{j\in J_2} c_j 1_{C_j}$, where $J_1 = \{j : c_j \geq 0\}$, and $J_2 = \{j : c_j < 0\}$. Because $X^+$ and $X^-$ are both simple random variables, based on definition 1, we have $\mathrm{E}_1(X^+) = \sum_{j\in J_1} c_j P(C_j)$, and $\mathrm{E}_1(X^-) = -\sum_{j\in J_2} c_j P(C_j)$, both finite. Moreover, as shown in class, expectation definitions 1 and 2 are equivalent, and thus $\mathrm{E}_2(X^+) = \mathrm{E}_1(X^+)$ and $\mathrm{E}_2(X^-) = \mathrm{E}_1(X^-)$. Therefore, the expectation of $X$ according to definition 3 exists and is given by $\mathrm{E}_3(X) = \mathrm{E}_2(X^+) - \mathrm{E}_2(X^-) = \sum_{j\in J_1} c_j P(C_j) + \sum_{j\in J_2} c_j P(C_j) = \sum_{j=1}^{k} c_j P(C_j) = \mathrm{E}_1(X)$.

5. Consider an $\overline{\mathbb{R}}^+$-valued random variable $X$ defined on some probability space $(\Omega, \mathcal{F}, P)$, and assume that $\mathrm{E}(X) < \infty$. Let $A = \{\omega \in \Omega : X(\omega) = +\infty\}$, and note that, based on the general definition for $\overline{\mathbb{R}}^+$-valued measurable functions, we have $A \in \mathcal{F}$.
   - Show that $X$ is almost surely finite, that is, $P(A) = 0$.

   **Solution:** Consider the countable sequence of random variables $\{X_n : n = 1, 2, ...\}$, with $X_n$ defined by $X_n(\omega) = n 1_A(\omega)$, $\omega \in \Omega$. By construction, $X_n(\omega) \leq X(\omega)$, for all $n$ and for all $\omega \in \Omega$. Note also that each $X_n$ is a simple random variable with $\mathrm{E}(X_n) = nP(A)$. Therefore, we have that, for all $n$, $nP(A) = \mathrm{E}(X_n) \leq \mathrm{E}(X) < \infty$, which can only occur if $P(A) = 0$.

6. Consider a sequence $\{X_n : n = 1, 2, ...\}$ of $\overline{\mathbb{R}}^+$-valued random variables defined on the same probability space $(\Omega, \mathcal{F}, P)$. Assume that the sequence is (pointwise) increasing, that is, for all $n$ and for each $\omega \in \Omega$, $X_n(\omega) \leq X_{n+1}(\omega)$. Denote by $X$ the pointwise limit of $\{X_n : n = 1, 2, ...\}$, that is, for each $\omega \in \Omega$, $X(\omega) = \lim_{n\to\infty} X_n(\omega)$, and assume that $\mathrm{E}(X) < \infty$. Define the *variance* for $X$ by $\mathrm{Var}(X) = \mathrm{E}\{(X - \mathrm{E}(X))^2\}$, and similarly, for each $n$, $\mathrm{Var}(X_n) = \mathrm{E}\{(X_n - \mathrm{E}(X_n))^2\}$. (In general, the variance for a random variable $Y$ with finite expectation $\mathrm{E}(Y)$ is given by $\mathrm{Var}(Y) = \mathrm{E}\{(Y - \mathrm{E}(Y))^2\}$, whether finite or infinite.)
   - Prove that $\mathrm{Var}(X) = \lim_{n\to\infty} \mathrm{Var}(X_n)$.

   **Solution:** Let $\mu = \mathrm{E}(X)$ and, for $n = 1, 2, ...$, $\mu_n = \mathrm{E}(X_n)$. The increasing structure of the $X_n$ and the assumption $\mu = \mathrm{E}(X) < \infty$ imply that $\mu_n < \infty$ for each $n$. Hence, $\mathrm{Var}(X)$ and $\mathrm{Var}(X_n)$, for each $n$, are well defined, and, based on the variance definition, we can write $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mu^2$, and $\mathrm{Var}(X_n) = \mathrm{E}(X_n^2) - \mu_n^2$, for $n = 1, 2, ...$ (by expanding the square and using additivity of expectation).
   Next, applying the MCT to the sequence $\{X_n : n = 1, 2, ...\}$, we have $\lim_{n\to\infty} \mu_n = \mu$, and thus $\lim_{n\to\infty} \mu_n^2 = \mu^2$. Note that the sequence $\{X_n^2 : n = 1, 2, ...\}$ also satisfies the MCT assumptions; it is a pointwise increasing sequence with pointwise limit $X^2$. Therefore, $\lim_{n\to\infty} \mathrm{E}(X_n^2) = \mathrm{E}(X^2)$. Combining the above results/expressions, we obtain

$$\lim_{n\to\infty} \mathrm{Var}(X_n) = \lim_{n\to\infty} \{\mathrm{E}(X_n^2) - \mu_n^2\} = \lim_{n\to\infty} \mathrm{E}(X_n^2) - \lim_{n\to\infty} \mu_n^2 = \mathrm{E}(X^2) - \mu^2 = \mathrm{Var}(X).$$

# AMS 261: Probability Theory (Fall 2017)

Homework 4 solutions

1. Consider a sequence $\{X_n : n = 1, 2, ...\}$ of $\overline{\mathbb{R}}$-valued random variables defined on the same probability space $(\Omega, \mathcal{F}, P)$. Assume that the sequence is (pointwise) increasing, that is, for all $n$ and for each $\omega \in \Omega$, $X_n(\omega) \leq X_{n+1}(\omega)$. Moreover, assume that $\mathrm{E}(X_1) > -\infty$. Denote by $X$ the pointwise limit of $\{X_n : n = 1, 2, ...\}$, that is, for each $\omega \in \Omega$, $X(\omega) = \lim_{n\to\infty} X_n(\omega)$. Prove that $\mathrm{E}(X) = \lim_{n\to\infty} \mathrm{E}(X_n)$.
**Solution:** Since $\{X_n : n = 1, 2, ...\}$ is a (pointwise) increasing sequence of random variables, it is easy to show that the sequence of the corresponding positive parts, $\{X_n^+ : n = 1, 2, ...\}$, is increasing with limit given by $X^+$. Therefore, applying the MCT to the $\overline{\mathbb{R}}^+$-valued random variables $X_n^+$, we obtain

$$\lim_{n\to\infty} \mathrm{E}(X_n^+) = \mathrm{E}(X^+). \tag{1.1}$$

Similarly, note that $\{-X_n^- : n = 1, 2, ...\}$ is an increasing sequence of $\overline{\mathbb{R}}^-$-valued random variables. Since $X_1 \leq X_2$ and $\mathrm{E}(X_1) > -\infty$, we have that $\mathrm{E}(X_2)$ exists and $-\infty < \mathrm{E}(X_1) \leq \mathrm{E}(X_2)$ (Fristedt & Gray, 1997, Chapter 4, Theorem 9(iv)). Applying the same argument, we get that $\mathrm{E}(X_n) > -\infty$, for each $n$, as well as that $\mathrm{E}(X) > -\infty$, which implies that $\mathrm{E}(X_n^-) < \infty$, for all $n$, as well as $\mathrm{E}(X^-) < \infty$. Next, since $\mathrm{E}(X_1^-) < \infty$, we conclude that $X_1^-$ is almost surely finite, that is, $-X_1^- > -\infty$, almost surely, and thus $c = \inf\{-X_1^-(\omega) : \omega \in \Omega\} > -\infty$. Now, $\{-X_n^- - c : n = 1, 2, ...\}$ is an increasing sequence of $\overline{\mathbb{R}}^+$-valued random variables, and the MCT yields

$$\lim_{n\to\infty} \mathrm{E}(X_n^-) = \mathrm{E}(X^-). \tag{1.2}$$

The result can now be obtained by combining (1.1) and (1.2), noting that $\lim_{n\to\infty}(\mathrm{E}(X_n^+) - \mathrm{E}(X_n^-))$ is well defined because $\mathrm{E}(X_n^-) < \infty$, for all $n$.

2. Let $\{X_n : n = 1, 2, ...\}$ be a countable sequence of $\overline{\mathbb{R}}^+$-valued random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$, and assume that $\mathrm{E}(\sum_{n=1}^{\infty} X_n) < \infty$. Show that $\mathrm{E}\left(\sum\limits_{n=1}^{\infty} X_n\right) = \sum\limits_{n=1}^{\infty} \mathrm{E}(X_n)$.
**Solution:** For $n = 1, 2, ...,$ define $Y_n = \sum_{j=1}^{n} X_j$. Then the sequence of $\overline{\mathbb{R}}^+$-valued random variables $\{Y_n : n = 1, 2, ...\}$, defined on $(\Omega, \mathcal{F}, P)$, is increasing, since each of the $X_j$ is $\overline{\mathbb{R}}^+$-valued. Denote by $Y$ the pointwise limit of the $Y_n$, i.e., for each $\omega \in \Omega$, $Y(\omega) = \lim_{n\to\infty} \sum_{j=1}^{n} X_j(\omega) = \sum_{n=1}^{\infty} X_n(\omega)$. Then, using the MCT and additivity of expectation,

$$\mathrm{E}\left(\sum_{n=1}^{\infty} X_n\right) = \mathrm{E}(Y) = \lim_{n\to\infty} \mathrm{E}(Y_n) = \lim_{n\to\infty} \sum_{j=1}^{n} \mathrm{E}(X_j) = \sum_{n=1}^{\infty} \mathrm{E}(X_n).$$

(Note that the assumption $\mathrm{E}(\sum_{n=1}^{\infty} X_n) < \infty$ implies that $\sum_{n=1}^{\infty} X_n$ is an almost surely finite random variable, but is not strictly needed.)

3. Let $\{X_n : n = 1, 2, ...\}$, $\{Y_n : n = 1, 2, ...\}$, and $\{Z_n : n = 1, 2, ...\}$ be sequences of $\mathbb{R}$-valued random variables (all the random variables are defined on the same probability space). Assume that: (a) $\mathrm{E}(X_n)$ and $\mathrm{E}(Z_n)$ exist for all $n$ and are finite; (b) each of the three sequences converges almost surely (denote by $X$, $Y$, and $Z$ the respective almost sure limits); (c) $\mathrm{E}(X)$, $\mathrm{E}(Y)$, and $\mathrm{E}(Z)$ exist and are finite; (d) $X_n \leq Y_n \leq Z_n$ almost surely; (e) $\lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}(X)$, and $\lim_{n\to\infty} \mathrm{E}(Z_n) = \mathrm{E}(Z)$. Show that $\lim_{n\to\infty} \mathrm{E}(Y_n) = \mathrm{E}(Y)$.
**Solution:** Consider the sequence of random variables $\{Z_n - Y_n : n = 1, 2, ...\}$. Based on assumption (d), $Z_n - Y_n \geq 0$, almost surely, and, therefore, using the Fatou lemma,

$$\mathrm{E}(\liminf_{n\to\infty}(Z_n - Y_n)) \leq \liminf_{n\to\infty} \mathrm{E}(Z_n - Y_n). \tag{3.1}$$

Using assumption (b), we obtain that the almost sure limit of the sequence $\{Z_n - Y_n : n = 1, 2, ...\}$ is given by $Z - Y$, and so $\liminf_{n\to\infty}(Z_n - Y_n) = \lim_{n\to\infty}(Z_n - Y_n) = Z - Y$, almost surely. Therefore, using properties of the lim inf for numerical sequences, (3.1) yields

$$\begin{aligned}\mathrm{E}(Z - Y) &\leq \liminf_{n\to\infty} \mathrm{E}(Z_n - Y_n) = \liminf_{n\to\infty}\{\mathrm{E}(Z_n) - \mathrm{E}(Y_n)\} \\ &= \liminf_{n\to\infty} \mathrm{E}(Z_n) + \liminf_{n\to\infty}\{-\mathrm{E}(Y_n)\} = \mathrm{E}(Z) - \limsup_{n\to\infty} \mathrm{E}(Y_n),\end{aligned}$$

since $\mathrm{E}(Z) = \lim_{n\to\infty} \mathrm{E}(Z_n) = \liminf_{n\to\infty} \mathrm{E}(Z_n)$ (assumption (e)). Rearranging terms in the above inequality, we have $\mathrm{E}(Y) \geq \limsup_{n\to\infty} \mathrm{E}(Y_n)$.
Analogously, consider the sequence $\{Y_n - X_n : n = 1, 2, ...\}$, which is, almost surely, non-negative, and converges, almost surely, to $Y - X$, based on assumptions (d) and (b), respectively. Hence, $\mathrm{E}(Y) - \mathrm{E}(X) = \mathrm{E}(Y - X) = \mathrm{E}(\lim_{n\to\infty}(Y_n - X_n)) = \mathrm{E}(\liminf_{n\to\infty}(Y_n - X_n))$, and, thus, using, again, the Fatou lemma and properties of the lim inf, $\mathrm{E}(Y) - \mathrm{E}(X) \leq \liminf_{n\to\infty} \mathrm{E}(Y_n - X_n) = \liminf_{n\to\infty}\{\mathrm{E}(Y_n) - \mathrm{E}(X_n)\} = \liminf_{n\to\infty} \mathrm{E}(Y_n) - \mathrm{E}(X)$, since $\liminf_{n\to\infty} \mathrm{E}(X_n) = \lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}(X)$ from assumption (e). Hence, $\mathrm{E}(Y) \leq \liminf_{n\to\infty} \mathrm{E}(Y_n)$, which, combined with $\mathrm{E}(Y) \geq \limsup_{n\to\infty} \mathrm{E}(Y_n)$, proves the result.

4. Let $\{X_n : n = 1, 2, ...\}$ be a countable sequence of $\mathbb{R}$-valued random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. Assume that there exist finite real constants $p > 1$ and $K > 0$ such that $\sup_n \mathrm{E}(|X_n|^p) \leq K$. Show that $\{X_n : n = 1, 2, ...\}$ is uniformly integrable.
**Solution:** For any $c > 0$, we can write $\mathrm{E}(|X_n| 1_{(|X_n|\geq c)}) = \mathrm{E}(|X_n|^p |X_n|^{1-p} 1_{(|X_n|\geq c)}) \leq c^{1-p}\mathrm{E}(|X_n|^p)$, since $p > 1$. Therefore, $\sup_n \mathrm{E}(|X_n| 1_{(|X_n|\geq c)}) \leq c^{1-p} \sup_n \mathrm{E}(|X_n|^p) \leq Kc^{1-p}$, using the assumption. Hence, finally, $\lim_{c\to\infty} \sup_n \mathrm{E}(|X_n| 1_{(|X_n|\geq c)}) \leq \lim_{c\to\infty}(Kc^{1-p}) = 0$, proving the result.

5. Let $X$ be an $\mathbb{R}$-valued random variable, defined on probability space $(\Omega, \mathcal{F}, P)$, with finite expectation $\mu = \mathrm{E}(X)$ and finite standard deviation $\sigma = \{\mathrm{Var}(X)\}^{1/2}$. Prove that for any $0 \leq z \leq \sigma$,

$$P(\{\omega \in \Omega : |X(\omega) - \mu| \geq z\}) \geq \frac{\sigma^4\{1 - (z/\sigma)^2\}^2}{\mathrm{E}(|X - \mu|^4)}.$$

**Solution:** Let $Y = |X - \mu|^2$. We have $\mathrm{E}(Y) = \mathrm{E}(|X - \mu|^2) = \mathrm{Var}(X) < \infty$, by assumption. If $\mathrm{E}(Y^2) = \mathrm{E}(|X - \mu|^4) = \infty$, the inequality holds true (the right hand side is 0 in this case). The case $\mathrm{E}(Y^2) = 0$ is not of interest for the inequality (the right hand side is not well defined in this case); note that if $\mathrm{E}(Y^2) = 0$ (and since $\mathrm{E}(Y) < \infty$), $Y$ is almost surely equal to a finite constant. Therefore, consider the case $0 < \mathrm{E}(Y^2) < \infty$. The result is obtained by applying to random variable $Y$ the inequality that can be viewed as a complement to Chebyshev inequality (Fristedt & Gray, 1997, Corollary 5.5; proved in class). In particular, setting $\lambda = z^2/\sigma^2$, for any $0 \leq z \leq \sigma$, we have (note that $\lambda \in [0, 1]$)

$$P(\{\omega \in \Omega : |X(\omega) - \mu|^2 \geq z^2\sigma^{-2}\mathrm{E}(|X - \mu|^2)\}) \geq \left(1 - \frac{z^2}{\sigma^2}\right)^2 \frac{\{\mathrm{E}(|X - \mu|^2)\}^2}{\mathrm{E}(|X - \mu|^4)},$$

which yields the result noting that $\sigma^2 = \mathrm{E}(|X - \mu|^2) < \infty$.

6. Let $\{X_n : n = 1, 2, ...\}$ be a sequence of $\mathbb{R}$-valued random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. Suppose there exists an $\mathbb{R}^+$-valued random variable $Y$, defined on $(\Omega, \mathcal{F}, P)$, such that $\mathrm{E}(Y) < \infty$ and $|X_n| \leq Y$, almost surely, for all $n$. Show that $\{X_n : n = 1, 2, ...\}$ is uniformly integrable.
**Solution:** Fix $c > 0$. Because $|X_n| \leq Y$, almost surely, for all $n$, we have $1_{(|X_n|\geq c)} \leq 1_{(Y\geq c)}$, almost surely, for all $n$. By combining the above inequalities, $|X_n| 1_{(|X_n|\geq c)} \leq Y 1_{(Y\geq c)}$, almost surely, for all $n$. Therefore, $\mathrm{E}(|X_n| 1_{(|X_n|\geq c)}) \leq \mathrm{E}(Y 1_{(Y\geq c)})$, for all $n$, and so $\sup_n \mathrm{E}(|X_n| 1_{(|X_n|\geq c)}) \leq \mathrm{E}(Y 1_{(Y\geq c)})$. Next, $\lim_{c\to\infty} \sup_n \mathrm{E}(|X_n| 1_{(|X_n|\geq c)}) \leq \lim_{c\to\infty} \mathrm{E}(Y 1_{(Y\geq c)}) = 0$, and thus $\lim_{c\to\infty} \sup_n \mathrm{E}(|X_n| 1_{(|X_n|\geq c)}) = 0$. (Note that the result $\lim_{c\to\infty} \mathrm{E}(Y 1_{(Y\geq c)}) = 0$ was proved in class, using the assumptions that $Y \geq 0$ and $\mathrm{E}(Y) < \infty$, and applying the DCT to the sequence $Z_k = Y 1_{(Y\geq k)} \leq Y$.)

7. Consider a countable sequence $\{X_n : n = 1, 2, ...\}$ of $\overline{\mathbb{R}}$-valued random variables, defined on a common probability space $(\Omega, \mathcal{F}, P)$, and an increasing function $G : [0, \infty) \to [0, \infty)$, which satisfies $\lim_{t\to\infty}\{t^{-1}G(t)\} = \infty$ and $0 < \sup_n \mathrm{E}\{G(|X_n|)\} < \infty$. Prove that $\{X_n : n = 1, 2, ...\}$ is uniformly integrable.
**Solution:** Fix $\varepsilon > 0$ and let $A = \varepsilon^{-1} \sup_n \mathrm{E}\{G(|X_n|)\}$ (we have $0 < A < \infty$, by assumption). Because $\lim_{t\to\infty}\{t^{-1}G(t)\} = \infty$, we can find large $c$ (which depends on $\varepsilon$) such that

$$t^{-1}G(t) \geq A, \quad \forall t \geq c. \tag{7.1}$$

For $n = 1, 2, ...,$ let $Y_n = |X_n| 1_{(|X_n|\geq c)}$. For any $\omega \in \Omega$ with $|X_n(\omega)| \geq c$, we have $Y_n(\omega) \geq c$, and using (7.1), $G(Y_n(\omega)) \geq AY_n(\omega)$. Moreover, for any $\omega \in \Omega$ with $|X_n(\omega)| < c$, we have $Y_n(\omega) = 0$, and since $G(0) \geq 0$, the inequality $G(Y_n(\omega)) \geq AY_n(\omega)$ is still valid. Therefore, for any $n = 1, 2, ...,$ $A|X_n| 1_{(|X_n|\geq c)} \leq G(|X_n| 1_{(|X_n|\geq c)}) \leq G(|X_n|)$, using the assumption that $G$ is increasing. Taking expectations, $\mathrm{E}(|X_n| 1_{(|X_n|\geq c)}) \leq A^{-1}\mathrm{E}\{G(|X_n|)\}$, and therefore, $\sup_n \mathrm{E}(|X_n| 1_{(|X_n|\geq c)}) \leq A^{-1} \sup_n \mathrm{E}\{G(|X_n|)\} = \varepsilon$, which provides the result, since the inequality above holds true for any $\varepsilon > 0$ and any $c' > c$.

# AMS 261: Probability Theory (Fall 2017)

Homework 5 solutions

1. Let $X$ be a random variable defined on a probability space $(\Omega, \mathcal{F}, P)$ and taking values in a measurable space $(\Psi, \mathcal{G})$, where $\mathcal{G}$ is the $\sigma$-field on space $\Psi$. Consider the collection $\mathcal{A}$ of subsets of $\Omega$ consisting of $X^{-1}(B)$ for all $B \in \mathcal{G}$. Show that $\mathcal{A}$ is a $\sigma$-field on $\Omega$.
**Solution:** First, note that because $X$ is a random variable, we know that $\mathcal{A} \subseteq \mathcal{F}$. To show that $\mathcal{A}$ is a $\sigma$-field, we need to verify the three conditions of the definition of a $\sigma$-field. First, because $\Psi \in \mathcal{G}$, we have $X^{-1}(\Psi) = \Omega \in \mathcal{A}$. Next, consider $A \in \mathcal{A}$. We have $A = X^{-1}(B)$ for some $B \in \mathcal{G}$. Using properties of inverse images, $X^{-1}(B^c) = (X^{-1}(B))^c = A^c$. Because $\mathcal{G}$ is a $\sigma$-field, we have $B^c \in \mathcal{G}$, which implies that $X^{-1}(B^c) \in \mathcal{A}$, and therefore $A^c \in \mathcal{A}$. Finally, let $\{A_n : n = 1, 2, ...\}$ be a countable collection of members of $\mathcal{A}$. For each $n$, $A_n = X^{-1}(B_n)$ for $B_n \in \mathcal{G}$. Now, $\bigcup_{n=1}^{\infty} B_n \in \mathcal{G}$, since $\mathcal{G}$ is a $\sigma$-field. Hence, $X^{-1}(\bigcup_{n=1}^{\infty} B_n) \in \mathcal{A}$, which yields the third condition, since $X^{-1}(\bigcup_{n=1}^{\infty} B_n) = \bigcup_{n=1}^{\infty} X^{-1}(B_n) = \bigcup_{n=1}^{\infty} A_n$, using again properties of inverse images.

2. For $k = 1, 2, ...$, consider random variables $X_k : (\Omega, \mathcal{F}, P) \to (\Psi_k, \mathcal{G}_k)$ and measurable functions $\varphi_k : (\Psi_k, \mathcal{G}_k) \to (\Theta_k, \mathcal{H}_k)$. Assume that the countable sequence of random variables $\{X_k : k = 1, 2, ...\}$ is independent. Prove that the sequence $\{\varphi_k \circ X_k : k = 1, 2, ...\}$ is independent.
**Solution:** We are given that $\{X_k : k = 1, 2, ...\}$ is independent, i.e., $\{\sigma(X_k) : k = 1, 2, ...\}$ is independent, i.e., for any finite index set $J$ (with $J \subset \{1, 2, ...\}$), $\{\sigma(X_j) : j \in J\}$ is independent, which implies that for any $B_j \in \mathcal{G}_j$,
$$P(\bigcap_{j \in J} X_j^{-1}(B_j)) = \prod_{j \in J} P(X_j^{-1}(B_j)). \tag{2.1}$$
Consider an arbitrary finite index set $J$ and $C_j \in \mathcal{H}_j$. We have
$$P(\bigcap_{j \in J} (\varphi_j \circ X_j)^{-1}(C_j)) = P(\bigcap_{j \in J} X_j^{-1}(\varphi_j^{-1}(C_j))) = \prod_{j \in J} P(X_j^{-1}(\varphi_j^{-1}(C_j))) = \prod_{j \in J} P((\varphi_j \circ X_j)^{-1}(C_j))$$
using (2.1) (with $B_j = \varphi_j^{-1}(C_j)$). Hence, $\{\sigma(\varphi_j \circ X_j) : j \in J\}$ is independent for any finite index set $J$, and therefore $\{\sigma(\varphi_k \circ X_k) : k = 1, 2, ...\}$ is independent.

3. Let $\{A_n : n = 1, 2, ...\}$ be a countable independent sequence of events on a probability space $(\Omega, \mathcal{F}, P)$. Prove that $P(\bigcap_{n=1}^{\infty} A_n) = \prod_{n=1}^{\infty} P(A_n)$. (**Note:** For a countable sequence of reals, $\{b_n : n = 1, 2, ...\}$, the infinite product $\prod_{n=1}^{\infty} b_n$ is defined by $\lim_{n\to\infty} \prod_{k=1}^{n} b_k$, provided this limit exists.)
**Solution:** Consider the new sequence of events $\{B_n : n = 1, 2, ...\}$, where $B_n = \bigcap_{k=1}^{n} A_k$. This is a decreasing sequence of events with $\lim_{n\to\infty} B_n = \bigcap_{n=1}^{\infty} B_n = \bigcap_{n=1}^{\infty} \bigcap_{k=1}^{n} A_k = \bigcap_{n=1}^{\infty} A_n$. Therefore, using continuity of probability measure and the assumption of independence for $\{A_n : n = 1, 2, ...\}$, we have
$$P(\bigcap_{n=1}^{\infty} A_n) = P(\lim_{n\to\infty} B_n) = \lim_{n\to\infty} P(B_n) = \lim_{n\to\infty} P(\bigcap_{k=1}^{n} A_k) = \lim_{n\to\infty} \prod_{k=1}^{n} P(A_k) = \prod_{n=1}^{\infty} P(A_n).$$
Note that for the sequence $q_n = \prod_{k=1}^{n} P(A_k)$ we have $1 \geq q_n \geq q_{n+1} \geq ... \geq 0$, and therefore the infinite product $\prod_{n=1}^{\infty} P(A_n) = \lim_{n\to\infty} q_n$ exists as either a strictly positive constant or 0.

4. Consider two countable sequences of events, $\{A_n : n = 1, 2, ...\}$ and $\{B_n : n = 1, 2, ...\}$, on the same probability space $(\Omega, \mathcal{F}, P)$. Assume that, for each $n$, $A_n$ and $B_n$ are independent. Moreover, assume that $A = \lim_{n\to\infty} A_n$ and $B = \lim_{n\to\infty} B_n$ exist. Show that $A$ and $B$ are independent.
**Solution:** We have $\lim_{n\to\infty} 1_{A_n}(\omega) = 1_A(\omega)$ and $\lim_{n\to\infty} 1_{B_n}(\omega) = 1_B(\omega)$, for each $\omega \in \Omega$. Therefore $1_{A\cap B}(\omega) = 1_A(\omega)1_B(\omega) = \lim_{n\to\infty}(1_{A_n}(\omega)1_{B_n}(\omega)) = \lim_{n\to\infty} 1_{A_n \cap B_n}(\omega)$, for each $\omega \in \Omega$, and thus $\lim_{n\to\infty}(A_n \cap B_n) = A \cap B$. Hence,
$$\begin{aligned} P(A \cap B) &= P(\lim_{n\to\infty}(A_n \cap B_n)) = \lim_{n\to\infty} P(A_n \cap B_n) = \lim_{n\to\infty}(P(A_n)P(B_n)) \\ &= (\lim_{n\to\infty} P(A_n))(\lim_{n\to\infty} P(B_n)) = P(\lim_{n\to\infty} A_n)P(\lim_{n\to\infty} B_n) = P(A)P(B) \end{aligned}$$
using continuity of probability measure (twice) and the independence of $A_n$ and $B_n$, for each $n$.

5. A sequence $\{X_n : n = 1, 2, ...\}$ of $\mathbb{R}$-valued random variables, defined on a common probability space $(\Omega, \mathcal{F}, P)$, is said to converge completely if for any $k = 1, 2, ...$, $\sum_{n=1}^{\infty} P(|X_n| > k^{-1}) < \infty$. Show that if $\{X_n : n = 1, 2, ...\}$ converges completely, then $\lim_{n\to\infty} X_n = 0$ almost surely.
**Solution:** The assumption of complete convergence yields that

$$P(\limsup_{n\to\infty}\{\omega \in \Omega : |X_n(\omega)| > k^{-1}\}) = 0, \quad \text{for} \quad k = 1, 2, ...$$

using the Borel lemma. Now the result follows using one of the equivalent definitions of almost sure convergence proved in class.

6. Construct a sequence $\{X_n : n = 1, 2, ...\}$ of $\mathbb{R}^+$-valued random variables (i.e., $X_n \geq 0$, for all $n$) that satisfies $\sum_{n=1}^{\infty} P(X_n > k^{-1}) < \infty$, for any $k = 1, 2, ...$, but for which $\lim_{n\to\infty} \mathrm{E}(X_n) \neq 0$.
**Solution:** For each $n = 1, 2, ...$, define $X_n$ so that it takes value $3^n$ with probability $2^{-n}$, and value 0 with probability $1 - 2^{-n}$. (For example, $X_n$ can be defined on $\Omega = (0, 1]$, with $\mathcal{F}$ the Borel $\sigma$-field on $(0, 1]$ and $P$ the uniform distribution, such that $X_n(\omega) = 3^n$, if $\omega \in (0, 2^{-n}]$, and $X_n(\omega) = 0$, otherwise.) Then, for any $k = 1, 2, ...$, $\sum_{n=1}^{\infty} P(X_n > k^{-1}) = \sum_{n=1}^{\infty} 2^{-n} = 1 < \infty$, but $\lim_{n\to\infty} \mathrm{E}(X_n) = \lim_{n\to\infty}(3/2)^n = \infty$.

7. Consider a countable sequence $\{X_n : n = 1, 2, ...\}$ of random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. Assume that each random variable $X_n$ is uniformly distributed on $(0, 1)$, hence, $P(c < X_n < d) \equiv P(\{\omega \in \Omega : X_n(\omega) \in (c, d)\}) = d - c$, for any $0 \leq c < d \leq 1$. Show that the sequence $\{1/(n^2 X_n) : n = 1, 2, ...\}$ converges almost surely to 0 as $n \to \infty$.
**Solution:** We need to show that $P\left(\left\{\omega \in \Omega : \forall k, \exists j, \forall n \geq j, \frac{1}{n^2 X_n(\omega)} < \frac{1}{k}\right\}\right) = 1$, or, equivalently, that

$$P\left(\bigcap_{k=1}^{\infty}\bigcup_{j=1}^{\infty}\bigcap_{n=j}^{\infty}\left\{\omega \in \Omega : \frac{1}{n^2 X_n(\omega)} < \frac{1}{k}\right\}\right) = P\left(\bigcap_{k=1}^{\infty} \liminf_{j\to\infty} A_{j,k}\right) = 1,$$

or, equivalently, that

$$0 = P\left((\bigcap_{k=1}^{\infty} \liminf_{j\to\infty} A_{j,k})^c\right) = P\left(\bigcup_{k=1}^{\infty} (\liminf_{j\to\infty} A_{j,k})^c\right) = P\left(\bigcup_{k=1}^{\infty} \limsup_{j\to\infty} A_{j,k}^c\right) \tag{7.1}$$

Here, for each $j = 1, 2, ...$, $k = 1, 2, ...$, $A_{j,k}$ is the event $\{\omega \in \Omega : \frac{1}{j^2 X_j(\omega)} < \frac{1}{k}\}$.

Now, if we fix $k$, there exists some $M = M(k)$ such that $k/j^2 < 1$, for any $j \geq M$. Then, for any such $j \geq M$,

$$P\left(A_{j,k}^c\right) = P\left(\left\{\omega \in \Omega : \frac{1}{j^2 X_j(\omega)} \geq \frac{1}{k}\right\}\right) = P\left(\left\{\omega \in \Omega : X_j(\omega) \leq \frac{k}{j^2}\right\}\right) = \frac{k}{j^2},$$

since each $X_j$ is uniformly distributed on $(0, 1)$. Hence, the series $\sum_{j=1}^{\infty} P(A_{j,k}^c)$ converges, since $\sum_{j=M}^{\infty} P(A_{j,k}^c) = k\sum_{j=M}^{\infty} j^{-2} < \infty$. Therefore, the Borel lemma yields that $P(\limsup_{j\to\infty} A_{j,k}^c) = 0$, for any $k$. Finally, (7.1) is established if we note that $P\left(\bigcup_{k=1}^{\infty} \limsup_{j\to\infty} A_{j,k}^c\right) \leq \sum_{k=1}^{\infty} P(\limsup_{j\to\infty} A_{j,k}^c) = 0$.

# AMS 261: Probability Theory (Fall 2017)

## Modes of convergence for sequences of random variables

**Definitions.** Consider $\mathbb{R}$-valued random variables $X$ and $\{X_n : n = 1, 2, ...\}$ defined on a common probability space $(\Omega, \mathcal{F}, P)$. The following four definitions are commonly used to study convergence for the sequence of random variables, "$X_n \to X$ as $n \to \infty$", and to obtain various limiting results for random variables and stochastic processes.

**Almost sure convergence** ($X_n \to^{\text{a.s.}} X$). $\{X_n : n = 1, 2, ...\}$ converges almost surely to $X$ if

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

**Convergence in probability** ($X_n \to^{\text{P}} X$). $\{X_n : n = 1, 2, ...\}$ converges in probability to $X$ if, for any $\epsilon > 0$,

$$\lim_{n\to\infty} P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \epsilon\}) = 0.$$

**Convergence in $r$th mean** ($X_n \to^{r-\text{mean}} X$). $\{X_n : n = 1, 2, ...\}$ converges in mean of order $r \geq 1$ (or in $r$th mean) to $X$ if

$$\lim_{n\to\infty} \mathrm{E}(|X_n - X|^r) = 0,$$

provided $\mathrm{E}(|X_n - X|^r) < \infty$, for each $n$.

**Convergence in distribution:** ($X_n \to^{\text{d}} X$). Denote by $F_{X_n}$ and $F_X$ the distribution function of $X_n$ and $X$, respectively. $\{X_n : n = 1, 2, ...\}$ converges in distribution to $X$ if

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

for all points $x$ at which $F_X$ is continuous.

**Equivalent definitions for almost sure convergence.** We have proved that each of the following are necessary and sufficient conditions for $\{X_n : n = 1, 2, ...\}$ to converge almost surely to $X$.
(1) For any $\epsilon > 0$, $P(\limsup_{n\to\infty}\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \epsilon\}) = 0$.
(2) For any $\epsilon > 0$, $\lim_{n\to\infty} P(\cup_{j=n}^{\infty}\{\omega \in \Omega : |X_j(\omega) - X(\omega)| > \epsilon\}) = 0$.
(3) For any $\epsilon > 0$, $\lim_{n\to\infty} P(\{\omega \in \Omega : \sup_{j\geq n} |X_j(\omega) - X(\omega)| > \epsilon\}) = 0$
(that is, $\sup_{j\geq n} |X_j - X| \to^{\text{P}} 0$, as $n \to \infty$).

**Comparisons between the different types of convergence.** We have shown that:

- Almost sure convergence implies convergence in probability.
- Convergence in $r$th mean implies convergence in probability, for any $r \geq 1$.
- Convergence in probability implies convergence in distribution.

It is also immediate from the definition that convergence in $r$th mean implies convergence in $s$th mean, for $r > s \geq 1$. No other implications hold without further assumptions on $\{X_n : n = 1, 2, ...\}$ and/or on $X$, as can be demonstrated with counterexamples.

*Example 1* ($X_n \to^{\text{p}} X$ does not imply $X_n \to^{\text{a.s.}} X$).
Let $\{X_n : n = 1, 2, ...\}$ be a sequence of independent random variables on $(\Omega, \mathcal{F}, P)$ such that, for each $n$, $X_n$ takes the value 0 with probability $1 - n^{-1}$ and the value $n$ with probability $n^{-1}$. (Note that, to define such $X_n$, we can take $\Omega = (0, 1]$ with the Borel $\sigma$-field, the uniform distribution for $P$, and set $X_n(\omega) = n$ if $0 < \omega \leq n^{-1}$, and $X_n(\omega) = 0$, otherwise). Then, from the definition, we have that $\{X_n : n = 1, 2, ...\}$ converges in probability to 0. However, using the first equivalent definition of almost sure convergence, we obtain that the sequence does not converge to 0 almost surely.

$P(|X_n| > \epsilon) = \frac{1}{n} \to 0$

$X_n \xrightarrow{p} 0$

*Example 2* ($X_n \to^{\text{d}} X$ does not imply $X_n \to^{\text{p}} X$).
Consider two independent random variables $X$ and $Y$ on $(\Omega, \mathcal{F}, P)$ both taking values 0 and 1 with probability 0.5 each. Set $X_n = Y$, for $n = 1, 2, ...$, which trivially implies that $\{X_n : n = 1, 2, ...\}$ converges in distribution to $X$. However, $|X_n - X| = |Y - X|$ takes values 0 and 1 with probability 0.5 each, therefore $P(|X_n - X| > \epsilon) = 0.5$ for any small $\epsilon$, and thus $\{X_n : n = 1, 2, ...\}$ does not converge in probability to $X$.

*Example 3* ($X_n \to^{\text{a.s.}} X$ does not imply $X_n \to^{r-\text{mean}} X$).
Let $\{X_n : n = 1, 2, ...\}$ be a sequence of random variables on $(\Omega, \mathcal{F}, P)$ such that, for each $n$, $X_n$ takes the value 0 with probability $1 - n^{-2}$ and the value $n$ with probability $n^{-2}$. Then, using the second equivalent definition of almost sure convergence, we obtain that $\{X_n : n = 1, 2, ...\}$ converges almost surely to 0. However, based on the definition, the sequence does not converge in mean of order 2 (and therefore it also does not converge in mean of any order greater than 2).

*Example 4* ($X_n \to^{r-\text{mean}} X$ does not imply $X_n \to^{\text{a.s.}} X$).
Let $\{X_n : n = 1, 2, ...\}$ be a sequence of independent random variables on $(\Omega, \mathcal{F}, P)$ such that, for each $n$, $X_n$ takes value 0 and 1 with probability $1 - n^{-1}$ and $n^{-1}$, respectively. Then, from the definition, $\{X_n : n = 1, 2, ...\}$ converges to 0 in mean of order $r$, for any $r \geq 1$. However, using the Borel-Cantelli lemma, $P(\limsup_{n\to\infty}\{\omega \in \Omega : |X_n(\omega)| > \epsilon\}) = 1$, for any $\epsilon > 0$, and therefore the sequence does not converge almost surely.

# AMS 261: Probability Theory (Fall 2017)

## Random infinite series

Consider a countable sequence $\{X_n : n = 1, 2, ...\}$ of **independent** $\mathbb{R}$–valued random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. A key theoretical question for the associated infinite series, $\sum_{n=1}^{\infty} X_n$, involves study of conditions for its almost sure (a.s.) convergence.

By definition, $\sum_{n=1}^{\infty} X_n$ converges a.s. if the sequence of random variables $\{S_n : n = 1, 2, ...\}$, where $S_n = \sum_{i=1}^{n} X_i$, converges a.s. Hence, the definition exploits the relation between series of random variables and series of reals (note that for each $\omega \in \Omega$, $\sum_{n=1}^{\infty} X_n(\omega)$ is a series of reals). In fact, it can be shown that for independent $\{X_n : n = 1, 2, ...\}$, $\sum_{n=1}^{\infty} X_n$ either converges a.s. or diverges a.s., with some key convergence results including:

**Theorem 1:** For a sequence $\{X_n : n = 1, 2, ...\}$ of independent $\mathbb{R}$–valued random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$, $\sum_{n=1}^{\infty} X_n$ converges a.s. to an $\mathbb{R}$–valued random variable $Z$ if and only if $\sum_{n=1}^{\infty} X_n$ converges in probability to $Z$.

**Theorem 2:** Consider a sequence $\{X_n : n = 1, 2, ...\}$ of independent $\mathbb{R}$–valued random variables, defined on a common probability space $(\Omega, \mathcal{F}, P)$. Assume that each $X_n$ has finite variance and that $\sum_{n=1}^{\infty} \mathrm{Var}(X_n) < \infty$. Then, $\sum_{n=1}^{\infty}\{X_n - \mathrm{E}(X_n)\}$ converges a.s.

**Kolmogorov three-series theorem:** Let $\{X_n : n = 1, 2, ...\}$ be an independent sequence of $\mathbb{R}$–valued random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. For each $n$, consider the truncated version of $X_n$ defined by $Y_n = X_n 1_{(|X_n| \leq b)}$, where $b$ is a positive real constant. Then $\sum_{n=1}^{\infty} X_n$ converges a.s. if and only if each of the following three series converges: $\sum_{n=1}^{\infty} \mathrm{E}(Y_n)$; $\sum_{n=1}^{\infty} \mathrm{Var}(Y_n)$; and $\sum_{n=1}^{\infty} P(X_n \neq Y_n) = \sum_{n=1}^{\infty} P(|X_n| > b)$.

## Weak and strong laws of large numbers

Laws of large numbers involve convergence results for functionals of $S_n = \sum_{i=1}^{n} X_i$ (the average, $n^{-1}\sum_{i=1}^{n} X_i$, being a standard example), where the sequence of random variables $\{X_n : n = 1, 2, ...\}$ is independent. Various versions of laws of large numbers exist, but the key results include the "Weak law of large numbers" (WLLN) (yielding convergence in probability) and the "Strong law of large numbers" (SLLN) (resulting in almost sure convergence). In particular, the two standard versions for the SLLN correspond to different assumptions for the sequence $\{X_n : n = 1, 2, ...\}$.

## Weak law of large numbers

Consider an independent sequence of $\mathbb{R}$–valued random variables $\{X_n : n = 1, 2, ...\}$, defined on a common probability space $(\Omega, \mathcal{F}, P)$, and let $S_n = \sum_{i=1}^{n} X_i$. Assume that for each $n$, $\mathrm{E}(X_n^2) < \infty$, and that $\lim_{n\to\infty} b_n^{-2} \sum_{i=1}^{n} \mathrm{Var}(X_i) = 0$, where $\{b_n : n = 1, 2, ...\}$ is a sequence of reals. Then, $b_n^{-1}(S_n - \mathrm{E}(S_n))$ converges to 0 in probability.

*Proof.* Application of Chebyshev's inequality to random variable $b_n^{-1} S_n$. ($b_n$ is usually $n$)

## Strong law of large numbers

Consider an independent sequence of $\mathbb{R}$–valued random variables $\{X_n : n = 1, 2, ...\}$, defined on a common probability space $(\Omega, \mathcal{F}, P)$, and let $S_n = \sum_{i=1}^{n} X_i$. Moreover, let $\{b_n : n = 1, 2, ...\}$ be an increasing sequence of positive reals such that $\lim_{n\to\infty} b_n = \infty$. Assume that for each $n$, $\mathrm{E}(X_n) = 0$ and $\mathrm{E}(X_n^2) < \infty$, and that $\sum_{i=1}^{\infty} b_i^{-2} \mathrm{E}(X_i^2) < \infty$. Then, $b_n^{-1} S_n$ converges to 0 almost surely.

*Proof.* Based on Theorem 2 and a result from series of real numbers, the Kronecker lemma. **Kronecker lemma:** Consider a sequence $\{x_n : n = 1, 2, ...\}$ of reals such that $\sum_{i=1}^{\infty} x_i < \infty$, and another sequence $\{b_n : n = 1, 2, ...\}$ of positive reals which is increasing, with $\lim_{n\to\infty} b_n = \infty$. Then, $\lim_{n\to\infty} b_n^{-1} \sum_{i=1}^{n} b_i x_i = 0$.

## Kolmogorov strong law of large numbers

Let $\{X_n : n = 1, 2, ...\}$ be independent and identically distributed (i.i.d.) $\mathbb{R}$–valued random variables, defined on a common probability space $(\Omega, \mathcal{F}, P)$. If $\mathrm{E}(|X_1|) < \infty$, then $n^{-1} \sum_{i=1}^{n} X_i$ converges to $\mathrm{E}(X_1)$ almost surely. Moreover, if $\mathrm{E}(|X_1|) = \infty$, then $n^{-1} \sum_{i=1}^{n} X_i$ diverges a.s.

*Proof.* For the case where $\mathrm{E}(|X_1|) < \infty$, assume without loss of generality that $\mathrm{E}(X_1) = 0$. The key idea is that to prove $n^{-1} \sum_{i=1}^{n} X_i \to^{\text{a.s.}} 0$ it suffices to prove $n^{-1} \sum_{i=1}^{n} Y_i \to^{\text{a.s.}} 0$, where $Y_n = X_n 1_{\{|X_n| < n\}}$ is a truncated version of $X_n$. This is based on the following lemmas:

**Lemma 1:** Let $\{X_n : n = 1, 2, ...\}$ be i.i.d. $\mathbb{R}$–valued random variables on probability space $(\Omega, \mathcal{F}, P)$. Then, $\mathrm{E}(|X_1|) < \infty$ if and only if $P(\limsup_{n\to\infty} \{|X_n| \geq n\}) = 0$.

**Lemma 2:** Let $Y$ be an $\mathbb{R}^+$–valued random variable on probability space $(\Omega, \mathcal{F}, P)$. Then, $\sum_{n=1}^{\infty} P(Y \geq n) \leq \mathrm{E}(Y) \leq 1 + \sum_{n=1}^{\infty} P(Y \geq n)$.

Almost sure convergence for $n^{-1} \sum_{i=1}^{n} Y_i$ can be established using Theorem 2 (truncation ensures finiteness of $\mathrm{Var}(Y_n)$, even though no assumption is made on finiteness of $\mathrm{Var}(X_n)$).